

# Copulas approximation and new families

V. Durrleman, A. Nikeghbali & T. Roncalli\*

Groupe de Recherche Opérationnelle  
Crédit Lyonnais  
France

August 21, 2000

## Abstract

In this paper, we study the approximation procedures introduced by LI, MIKUSINSKI, SHERWOOD and TAYLOR [1997]. We show that there exists a bijection between the set of the *discretized* copulas and the set of the *doubly stochastic* matrices. For the Bernstein and checkerboard approximations, we then provide analytical formulas for the Kendall's tau and Spearman's rho concordance measures. Moreover, we demonstrate that these approximations does not exhibit tail dependences. Finally, we consider the general case of approximations induced by *partitions of unity*.

## 1 Introduction

The inspiration of this article lies in the work of LI, MIKUSINSKI, SHERWOOD and TAYLOR [1997] on copulas approximations. They showed that approximations based on Bernstein polynomials and checkerboard method converge to the original copula in a *stronger* way.

We are here interested in studying these families and in obtaining new estimators for the Kendall's tau and Spearman's rho concordance measures. We show that these two dependence measures converge to the original ones too. Nevertheless, the tail dependences of these two approximations is **always** zero.

In a last part of the article, we consider new families induced by *partitions of unity*. Moreover, we have shown that the set of copulas induced by *partition of unity* is a Markov sub-algebra with respect to the  $*$ -product of DARSOW, NGUYEN and OLSEN [1992].

## 2 Doubly stochastic matrices and copulas

Suppose we are given a 2-dimensional copula  $\mathbf{C}$  (a positive and continuous function  $\mathbf{C}$  defined on  $[0, 1]^2$  which is 2-increasing). The issue we discuss here is to construct a copula from a discretization of  $\mathbf{C}$ . We suppose that we know the value of  $\mathbf{C}$  on each point of the grid  $\mathfrak{L} = \left\{ \left( \frac{i}{n}, \frac{j}{n} \right) : 0 \leq i, j \leq n \right\}$  for a given  $n$ .

**Definition 1** *We say that the function  $f$  is an approximation if we verify the following properties:*

- *the function  $f$  is defined in terms of the values of the discretized copula*

$$f(\mathbf{C}) = f(\{\mathbf{C}(u, v) : (u, v) \in \mathfrak{L}\}) \tag{1}$$

---

\* *Corresponding author:* Groupe de Recherche Opérationnelle, Bercy-Expo — Immeuble Bercy Sud — 4<sup>b</sup> étage, 90 quai de Bercy — 75613 Paris Cedex 12 — France; *E-mail adress:* thierry.roncalli@creditlyonnais.fr

- $f(\mathbf{C})$  is a copula;
- $f(\mathbf{C})$  converges to the original copula  $C$  when  $n$  increases.

We can see approximation of copulas from two points of view:

1. From an econometric point of view,  $\mathbf{C}$  is a copula which arises in a practical issue and the  $\mathbf{C}(\frac{i}{n}, \frac{j}{n})$  come from the estimation step. In this framework, approximations are an appropriate tool to generate regular copulas which ‘approximate’  $\mathbf{C}$  in a good way. Some applications could be found in DURRLEMAN, NIKEGHBALI and RONCALLI [2000].
2. From another point of view, we can study approximation of copulas for their own interest. In that case and in order to build a copula, we need a way to construct a family  $(C_{i,j})_{0 \leq i, j \leq n}$  which is a discretization of a copula.

In the second point of view, the construction is done by the following result which states a bijection between the set  $\mathcal{D}_n$  of *doubly stochastic*  $n \times n$  matrices and the set  $\mathcal{C}_n$  of *discretized* copulas at each point  $(\frac{i}{n}, \frac{j}{n})$  of the grid  $\mathcal{L}$ .

**Theorem 2** Let  $\mathcal{M}_n$  denotes the set of the  $n \times n$  matrices of real numbers and let  $n$  be a positive integer. We define the function  $\varpi$  as follows

$$\begin{aligned} \varpi : \mathcal{D}_n &\longrightarrow \mathcal{M}_{n+1} \\ A &\longmapsto B = \varpi(A) \end{aligned} \quad (2)$$

with

$$b_{0,i} = b_{i,0} = 0 \quad 0 \leq i \leq n \quad (3)$$

and

$$b_{i,j} = \frac{1}{n} \sum_{p=1}^i \sum_{q=1}^j a_{p,q} \quad 1 \leq i, j \leq n \quad (4)$$

then  $\varpi(\mathcal{D}_n) = \mathcal{C}_n$ .

**Proof.** We must first check that  $\varpi(\mathcal{D}_n) \subset \mathcal{C}_n$ . Let  $A = (a_{i,j}) \in \mathcal{D}_n$  and  $B = (b_{i,j})$ .  $B = \varpi(A)$  is the *discretization* of a copula if we check that

- $b_{0,i} = b_{i,0} = 0$  for all  $0 \leq i \leq n$  which is of course satisfied;
- $b_{n,i} = b_{i,n} = i/n$  for all  $0 \leq i \leq n$  which is also true:

$$b_{i,n} = \frac{1}{n} \sum_{p=1}^i \sum_{q=1}^n a_{p,q} = \frac{1}{n} \sum_{p=1}^i 1 = \frac{i}{n} \quad (5)$$

- $b_{i+1,j+1} - b_{i+1,j} - b_{i,j+1} + b_{i,j} \geq 0$  for all  $0 \leq i, j \leq n-1$  which is verified because

$$\begin{aligned} b_{i+1,j+1} - b_{i+1,j} - b_{i,j+1} + b_{i,j} &= \frac{1}{n} a_{i+1,j+1} \\ &\geq 0 \end{aligned} \quad (6)$$

Then, we have to verify that  $\varpi$  is a bijection. Given  $B = (b_{i,j})_{0 \leq i,j \leq n}$ , we calculate  $A = (a_{i,j})_{1 \leq i,j \leq n}$  as follows

$$a_{i,j} = n(b_{i,j} - b_{i,j-1} - b_{i-1,j} + b_{i-1,j-1}) \quad (7)$$

Because  $B \in \mathcal{C}_n$ , it comes that  $a_{i,j} \geq 0$ . On the other hand, we have

$$\begin{aligned} \sum_{i=1}^n a_{i,j} &= n \sum_{i=1}^n (b_{i,j} - b_{i-1,j}) + (b_{i-1,j-1} - b_{i,j-1}) \\ &= n(b_{n,j} - b_{n,j-1}) \\ &= 1 \end{aligned} \quad (8)$$

By symmetry, the same result holds for  $\sum_{j=1}^n a_{i,j}$ , and this completes the proof. ■

**Remark 3** In the rest of the paper, we will use the notations  $\mathcal{D}(\mathbf{C})$  — or  $\mathcal{D}_n(\mathbf{C})$  — for  $\varpi^{-1}(\mathbf{C})$  and  $\mathcal{C}(A)$  — or  $\mathcal{C}_n(\mathbf{C})$  — for  $\varpi(A)$ .

**Example 4** for  $\alpha \in [0, 1]$ , we have

$$\mathcal{C} \left( \begin{bmatrix} \alpha & 1-\alpha \\ 1-\alpha & \alpha \end{bmatrix} \right) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{\alpha}{2} & \frac{1}{2} \\ 0 & \frac{\alpha}{2} & 0 \end{bmatrix}$$

**Example 5** The doubly stochastic matrices of the lower and upper Fréchet bounds  $\mathbf{C}^-$  and  $\mathbf{C}^+$ , and the product copula  $\mathbf{C}^\perp$  are respectively

$$\mathcal{D}(\mathbf{C}^-) = [\mathbf{e}_n \quad \mathbf{e}_{n-1} \quad \cdots \quad \mathbf{e}_1] \quad (9)$$

$$\mathcal{D}(\mathbf{C}^\perp) = n^{-1} \mathbf{1}_{n \times n}$$

$$\mathcal{D}(\mathbf{C}^+) = \mathbf{I}_n \quad (10)$$

We remark that  $\mathcal{D}(\mathbf{C}^\perp)$  and  $\mathcal{D}(\mathbf{C}^+)$  correspond respectively to the null and unit elements of the algebra  $\mathcal{D}_n$  (DARSOW, NGUYEN and OLSEN [1992]). Moreover,  $\mathcal{D}(\mathbf{C}^-)$  and  $\mathcal{D}(\mathbf{C}^+)$  are two extreme points of  $\mathcal{D}_n$ , and we know that  $\mathbf{C}^-$  and  $\mathbf{C}^+$  are two extreme points of the set of copulas.

**Example 6** The doubly stochastic matrices associated to the Gaussian copula with parameter  $\rho = 0.5$  are

$$\begin{aligned} \mathcal{D}_2(\mathbf{C}) &= \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{bmatrix} \\ \mathcal{D}_3(\mathbf{C}) &= \begin{bmatrix} 0.5486 & 0.3112 & 0.1402 \\ 0.3112 & 0.3776 & 0.3112 \\ 0.1402 & 0.3112 & 0.5486 \end{bmatrix} \\ \mathcal{D}_4(\mathbf{C}) &= \begin{bmatrix} 0.4811 & 0.2783 & 0.1684 & 0.07210 \\ 0.2783 & 0.2955 & 0.2577 & 0.1684 \\ 0.1684 & 0.2577 & 0.2955 & 0.2783 \\ 0.07210 & 0.1684 & 0.2783 & 0.4811 \end{bmatrix} \end{aligned}$$

Note that we do not explore here the bijection between  $\mathcal{D}_n$  and  $\mathcal{C}_n$ . Nevertheless, it seems that there a lot of things to say about it. For example, we constate that the second largest eigenvalue of the  $\mathcal{D}(\mathbf{C})$  where  $\mathbf{C}$  is the Gaussian copula converges to its parameter  $\rho$ !

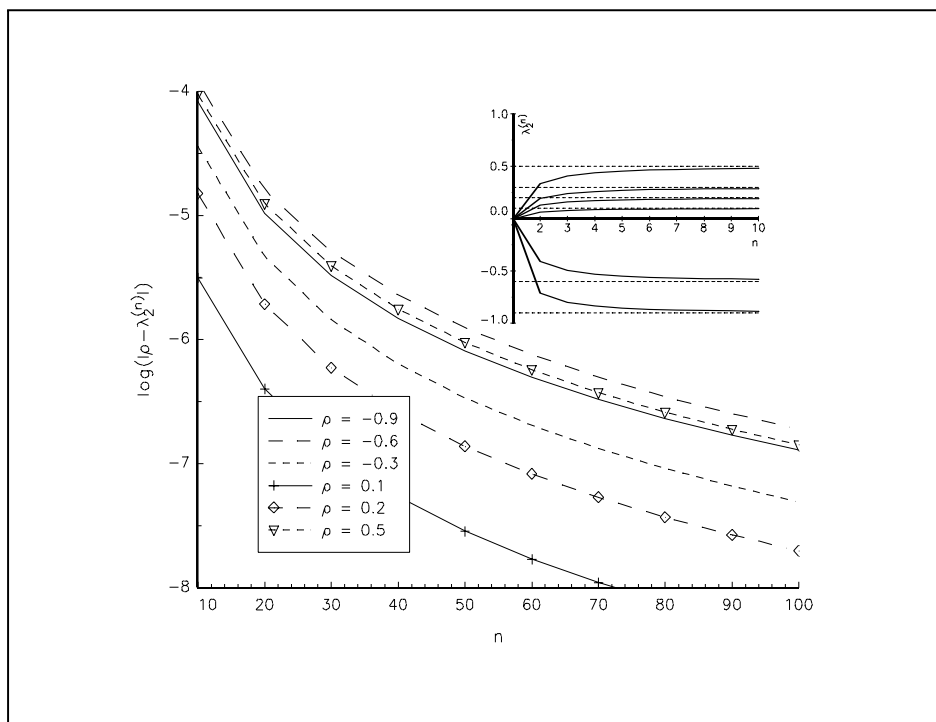


Figure 1: Convergence of the second largest eigenvalue to the parameter  $\rho$  of the Gaussian copula

### 3 The Bernstein polynomials approximation

Let  $B_{i,n}(x)$  denotes the Bernstein polynomial

$$B_{i,n}(x) = \binom{n}{i} x^i (1-x)^{n-i} \quad (11)$$

and  $\mathfrak{B}_n(\mathbf{C})$  the two place function defined by

$$\mathfrak{B}_n(\mathbf{C})(u, v) = \sum_{i=1}^n \sum_{j=1}^n B_{i,n}(u) B_{j,n}(v) \mathbf{C}\left(\frac{i}{n}, \frac{j}{n}\right) \quad (12)$$

The main interest of this construction is the following theorem:

**Theorem 7**  $\mathfrak{B}_n(\mathbf{C})$  is a copula for any copula  $\mathbf{C}$ .

**Proof.** First we shall see that the boundary conditions are satisfied.  $\mathfrak{B}_n(\mathbf{C})(0, v) = \sum_{i=1}^n \sum_{j=1}^n B_{i,n}(0) B_{j,n}(v) \mathbf{C}\left(\frac{i}{n}, \frac{j}{n}\right)$  and  $B_{i,n}(0) = 0$  for all  $i = 0, \dots, n$ . Thus,  $\mathfrak{B}_n(\mathbf{C})(0, v) = 0$ . By symmetry,  $\mathfrak{B}_n(\mathbf{C})(u, 0) = 0$ .  $\mathfrak{B}_n(\mathbf{C})(1, v) = \sum_{i=1}^n \sum_{j=1}^n B_{i,n}(1) B_{j,n}(v) \mathbf{C}\left(\frac{i}{n}, \frac{j}{n}\right)$  and  $B_{i,n}(1) = 0$  for all  $i$  except for  $i = n$  —  $B_{n,n}(1) = 1$ . So,  $\mathfrak{B}_n(\mathbf{C})(1, v) = \sum_{j=1}^n B_{j,n}(v) \mathbf{C}\left(1, \frac{j}{n}\right) = \sum_{j=1}^n \binom{n}{j} v^j (1-v)^{n-j}$ . Now we are interested in showing that this expression equals to  $v$ . For this, let us consider the expression  $(a+b)^n = \sum_{j=0}^n \binom{n}{j} a^j (1-b)^{n-j}$  and let's differentiate it with respect

to the variable  $a$ . We have  $\frac{\partial}{\partial a} (a+b)^n = n(a+b)^{n-1} = \sum_{j=0}^n j \binom{n}{j} a^{j-1} (1-b)^{n-j} = a^{-1} \sum_{j=1}^n j \binom{n}{j} a^j (1-b)^{n-j}$ .

Thus,  $\sum_{j=1}^n \frac{j}{n} \binom{n}{j} v^j (1-v)^{n-j} = \frac{v}{n} \times n \times (v+1-v)^{n-1} = v$ . So, we have  $\mathfrak{B}_n(\mathbf{C})(1, v) = v$ . Now, to complete the proof, as  $\mathfrak{B}_n(\mathbf{C})$  is infinitely differentiable, it suffices to check that for all  $(u, v) \in [0, 1]^2$ ,  $\frac{\partial^2}{\partial u \partial v} \mathfrak{B}_n(\mathbf{C})(u, v) \geq 0$ . We have

$$\begin{aligned} \frac{\partial^2}{\partial u \partial v} \mathfrak{B}_n(\mathbf{C})(u, v) &= \sum_{i=1}^n \sum_{j=1}^n \mathbf{C} \left( \frac{i}{n}, \frac{j}{n} \right) \left[ i \binom{n}{i} u^{i-1} (1-u)^{n-i} - (n-i) \binom{n}{i} u^i (1-u)^{n-i-1} \right] \times \\ &\quad \left[ j \binom{n}{j} v^{j-1} (1-v)^{n-j} - (n-j) \binom{n}{j} v^j (1-v)^{n-j-1} \right] \\ &= n^2 \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \Delta_{i+1, j+1} B_{i, n-1}(u) B_{j, n-1}(v) \\ &\geq 0 \end{aligned} \tag{13}$$

where  $\Delta_{i,j}$  is the mass of the square  $[\frac{i-1}{n}, \frac{i}{n}] \times [\frac{j-1}{n}, \frac{j}{n}]$  with the measure induced by the copula  $\mathbf{C}$  and which is thus nonnegative. ■

In order to illustrate this theorem, we have represented in the figure 2 the values taken by  $|\mathfrak{B}_n(\mathbf{C})(u, v) - \mathbf{C}(u, v)|$  when  $\mathbf{C}$  is the Gaussian copula with parameter 0.5, and we have plotted  $\ln(\sup |\mathfrak{B}_n(\mathbf{C})(u, v) - \mathbf{C}(u, v)|)$  for different copula functions in the figure 3.

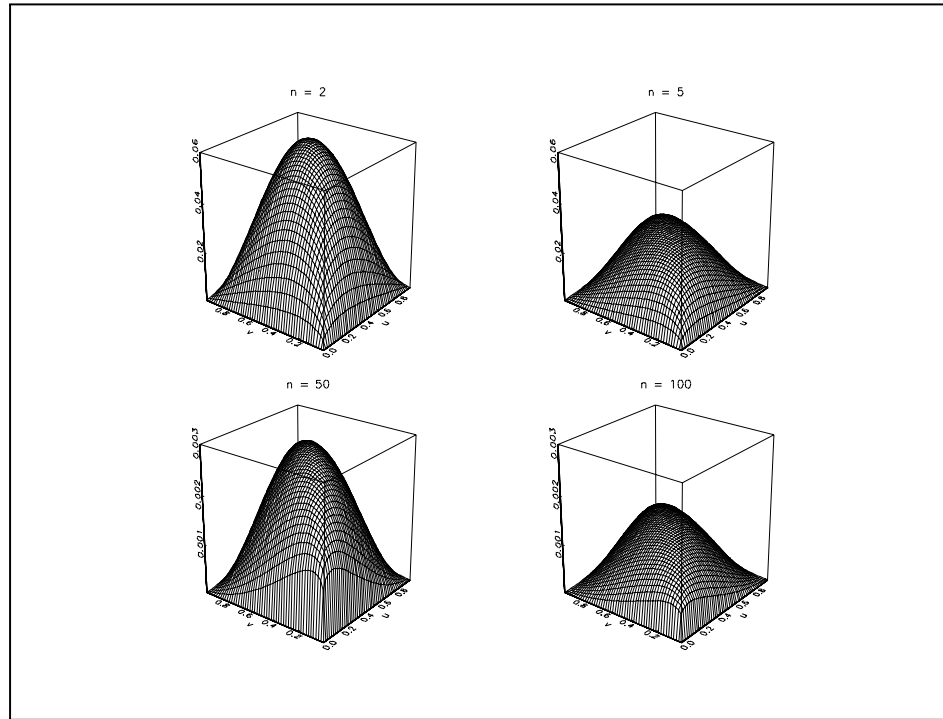


Figure 2: Surface plot of  $|\mathfrak{B}_n(\mathbf{C})(u, v) - \mathbf{C}(u, v)|$

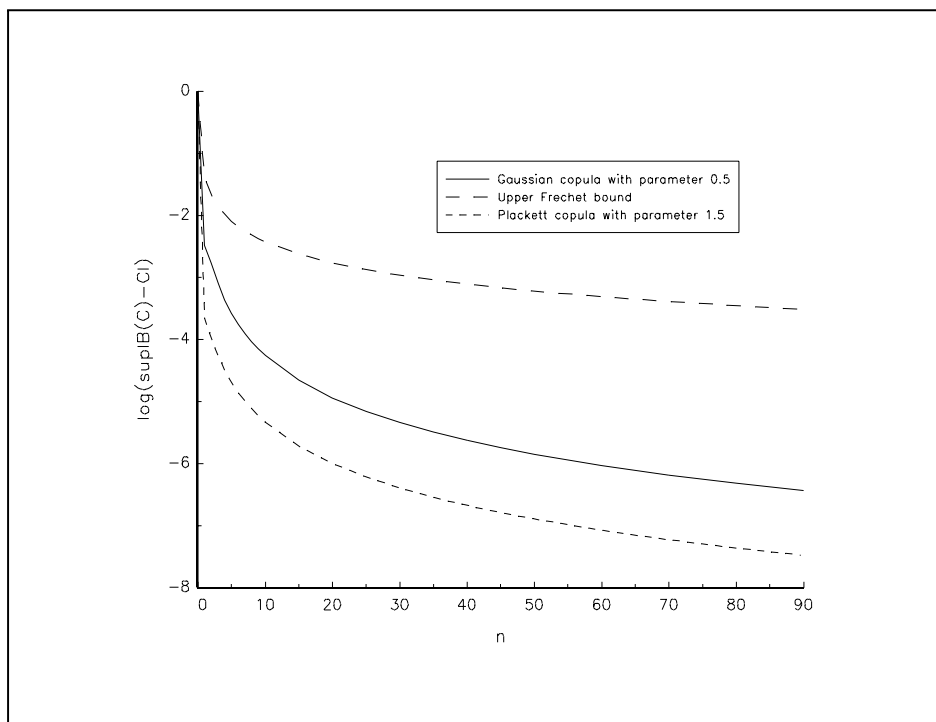


Figure 3: Illustration of the convergence for different copulas

### 3.1 How to generate Bernstein copulas?

**Definition 8** Let  $n$  be positive integer and  $D = (d_{i,j}) \in \mathcal{C}_n$ . A Bernstein copula associated with the matrix  $D$  is defined on  $[0, 1]^2$  by

$$\mathbf{C}(u, v) = \sum_{i=1}^n \sum_{j=1}^n B_{i,n}(u) B_{j,n}(v) d_{i,j} \quad (14)$$

By virtue of the above theorem,  $\mathbf{C}$  is indeed a copula.

### 3.2 The Kendall's tau concordance measure

Using the previous notations and results, we could show the following theorem:

**Theorem 9** Let  $\Theta \in \mathcal{M}_{n+1}$  where

$$\theta_{i,j} = \frac{(i-j) \binom{n}{i} \binom{n}{j}}{(2n-i-j) \binom{2n-1}{i+j-1}} \quad (15)$$

for all  $0 \leq i, j \leq n$  (with the convention  $0/0 = 1$ ), then the Kendall's tau of the Bernstein copula is

$$\tau = 1 - \text{tr}(\Theta D \Theta^\top) \quad (16)$$

**Proof.** We use the following definition of the Kendall's tau (NELSEN [1998]):

$$\tau = 4 \iint_{[0,1]^2} \mathbf{C}(u,v) \frac{\partial^2 \mathbf{C}}{\partial x \partial y}(u,v) \, du \, dv - 1 \quad (17)$$

Thanks to the Green's formula to obtain another tractable expression for  $\tau$

$$\tau = 4 \left( \int_{\partial[0,1]^2} \mathbf{C}(u,v) \frac{\partial \mathbf{C}}{\partial x}(u,v) \, d\sigma - \iint_{[0,1]^2} \frac{\partial \mathbf{C}}{\partial x}(u,v) \frac{\partial \mathbf{C}}{\partial y}(u,v) \, du \, dv \right) - 1 \quad (18)$$

and the first term can be evaluated (see NELSEN [1998])

$$\tau = 1 - 4 \iint_{[0,1]^2} \frac{\partial \mathbf{C}}{\partial x}(u,v) \frac{\partial \mathbf{C}}{\partial y}(u,v) \, du \, dv \quad (19)$$

Here we have

$$\frac{\partial \mathbf{C}}{\partial x}(u,v) = \sum_{i=1}^n \sum_{j=1}^n d_{i,j} \binom{n}{i} \binom{n}{j} u^{i-1} (1-u)^{n-i-1} v^j (1-v)^{n-j} (i-nu) \quad (20)$$

such that double integral in (19) equals now to

$$\begin{aligned} & \sum_{i=1}^n \sum_{j=1}^n \sum_{p=1}^n \sum_{q=1}^n d_{i,j} d_{p,q} \binom{n}{i} \binom{n}{j} \binom{n}{p} \binom{n}{q} \times \\ & \iint_{[0,1]^2} u^{i+p-1} (1-u)^{2n-i-p-1} v^{j+q-1} (1-v)^{2n-j-q-1} (i-nu)(q-nv) \, du \, dv \end{aligned} \quad (21)$$

Routine calculations of this last double integral give us

$$\iint_{[0,1]^2} \dots \, du \, dv = \frac{1}{4} \frac{(i-p)(q-j)}{(2n-i-p)(2n-j-q) \binom{2n-1}{i+p-1} \binom{2n-1}{q+j-1}} \quad (22)$$

so that Kendall's tau equals to

$$\tau = 1 - \sum_{i=1}^n \sum_{j=1}^n \sum_{p=1}^n \sum_{q=1}^n d_{i,j} d_{p,q} \theta_{i,p} \theta_{q,j} \quad (23)$$

which can be rewritten as

$$\tau = 1 - \text{tr}(\Theta D \Theta D^\top) \quad (24)$$

■

### 3.3 The Spearman's rho concordance measure

We could obtain a similar result for the Spearman's rho:

**Theorem 10** Let  $\Gamma \in \mathcal{M}_{n+1}$  where

$$\gamma_{i,j} = \frac{1}{(n+1)^2} \quad (25)$$

for all  $0 \leq i, j \leq n$ , then the Spearman's rho of the Bernstein copula is

$$\varrho = 12 \text{tr}(\Gamma D) - 3 \quad (26)$$

**Proof.** We use the following definition of the Spearman's rho (NELSEN [1998]):

$$\varrho = 12 \iint_{[0,1]^2} \mathbf{C}(u, v) \, du \, dv - 3 \quad (27)$$

Routine calculations give us

$$\iint_{[0,1]^2} \binom{n}{i} \binom{n}{j} u^i (1-u)^{n-i} v^j (1-v)^{n-j} \, du \, dv = \frac{1}{(n+1)^2} \quad (28)$$

Then, we have

$$\varrho = \frac{12}{(n+1)^2} \sum_{i=1}^n \sum_{j=1}^n d_{i,j} - 3 \quad (29)$$

■

### 3.4 The upper tail dependence measure

**Theorem 11** For any Bernstein copula  $\mathbf{C}$ , the upper tail dependence measure  $\lambda$  is

$$\lambda(\mathbf{C}) = 0 \quad (30)$$

**Proof.** We have

$$\lambda(\mathbf{C}) = 2 - \lim_{u \rightarrow 1^-} \frac{d}{du} \mathbf{C}(u, u) \quad (31)$$

Remark that

$$\frac{d}{du} \mathbf{C}(u, u) = \frac{\partial}{\partial u} \mathbf{C}(u, u) + \frac{\partial}{\partial v} \mathbf{C}(u, u) \quad (32)$$

Using the expression of the derivatives calculated for Kendall's tau, we have

$$\frac{\partial \mathbf{C}}{\partial u}(u, u) = \sum_{i=1}^n \sum_{j=1}^n d_{i,j} \binom{n}{i} \binom{n}{j} u^{i+j-1} (1-u)^{2n-i-j-1} (i-nu) \quad (33)$$

and

$$\frac{\partial \mathbf{C}}{\partial v}(u, u) = \sum_{i=1}^n \sum_{j=1}^n d_{i,j} \binom{n}{i} \binom{n}{j} u^{i+j-1} (1-u)^{2n-i-j-1} (j-nu) \quad (34)$$

so that

$$\lim_{u \rightarrow 1^-} \frac{\partial \mathbf{C}}{\partial u}(u, u) = \lim_{u \rightarrow 1^-} \frac{\partial \mathbf{C}}{\partial v}(u, u) = 1 \quad (35)$$

Then, we obtain the following result

$$\lambda = 0 \quad (36)$$

■



## 4 The checkerboard copula

As for Bernstein copulas, the construction of a checkerboard copula arises from empirical data. Let  $\Delta_{i,j}$  denotes the mass of  $[\frac{i-1}{n}, \frac{i}{n}] \times [\frac{j-1}{n}, \frac{j}{n}]$  with the measure induced by the copula  $\mathbf{C}$

$$\Delta_{i,j}(\mathbf{C}) = \mathbf{C}\left(\frac{i}{n}, \frac{j}{n}\right) - \mathbf{C}\left(\frac{i-1}{n}, \frac{j}{n}\right) - \mathbf{C}\left(\frac{i}{n}, \frac{j-1}{n}\right) + \mathbf{C}\left(\frac{i-1}{n}, \frac{j-1}{n}\right) \quad (37)$$

and  $\mathfrak{C}_n(\mathbf{C})$  the two place function defined by

$$\mathfrak{C}_n(\mathbf{C})(u, v) = n^2 \sum_{i=1}^n \sum_{j=1}^n \Delta_{i,j}(\mathbf{C}) \int_0^u \chi_{i,n}(x) dx \int_0^v \chi_{j,n}(y) dy \quad (38)$$

where  $\chi_{i,n}$  is the characteristic function of the interval  $[\frac{i-1}{n}, \frac{i}{n}]$ .

**Theorem 12**  $\mathfrak{C}_n(\mathbf{C})$  is a copula for any copula  $\mathbf{C}$ .

**Proof.** see [6], [7] or [8]. ■

As in Bernstein's copulas, we have represented in the figures 4 and 5 the values taken by  $|\mathfrak{B}_n(\mathbf{C})(u, v) - \mathbf{C}(u, v)|$  when  $\mathbf{C}$  is the Gaussian copula with parameter 0.5 and the values of  $\ln(\sup |\mathfrak{B}_n(\mathbf{C})(u, v) - \mathbf{C}(u, v)|)$  for different copula functions.

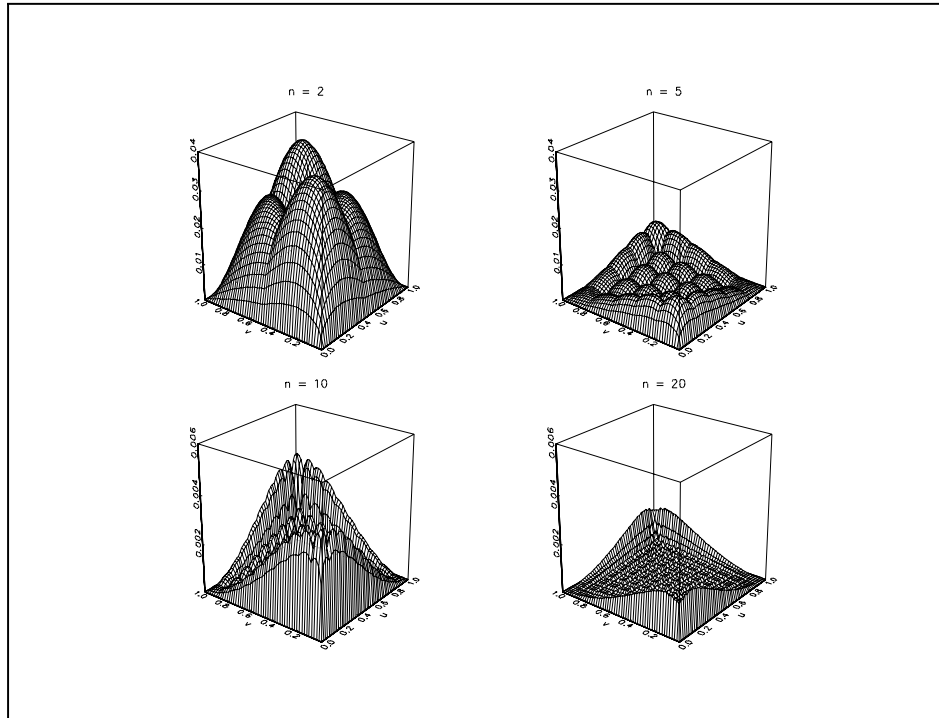


Figure 4: Surface plot of  $|\mathfrak{B}_n(\mathbf{C})(u, v) - \mathbf{C}(u, v)|$

### 4.1 How to generate checkerboard copulas?

We would like now to characterize the coefficients  $\Delta_{i,j}(\mathbf{C})$  without any reference to the underlying copula  $\mathbf{C}$ . We have the following useful lemma.

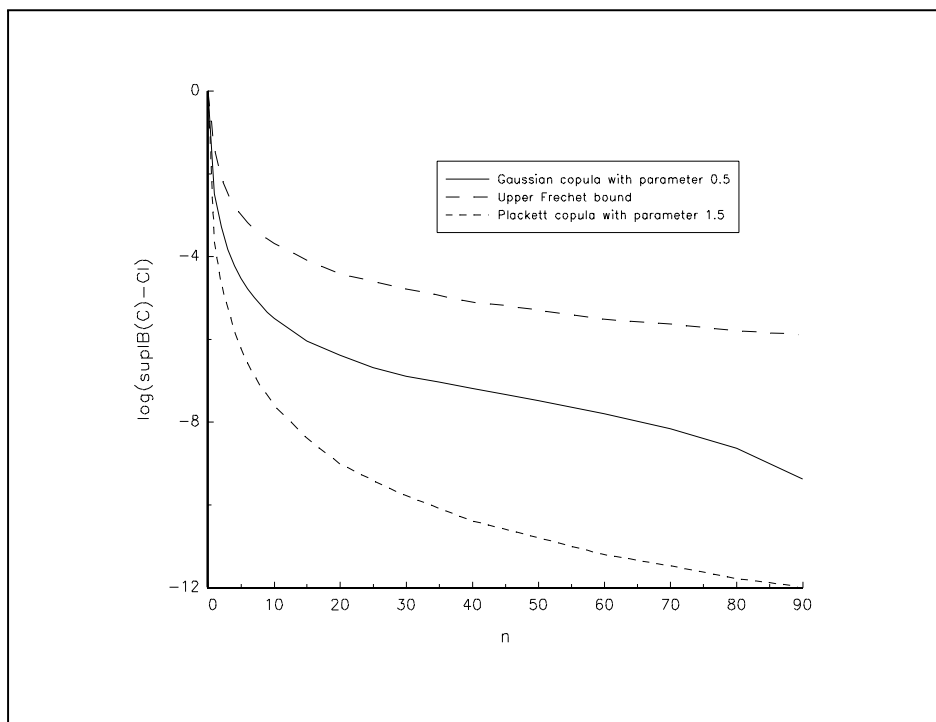


Figure 5: Illustration of the convergence for different copulas

**Lemma 13** Let  $D = (d_{i,j}) \in \mathcal{M}_{n+1}$ . There exists a copula  $\mathbf{C}$  such that  $d_{i,j} = \Delta_{i,j}(\mathbf{C})$  if and only if  $nD \in \mathcal{D}_n$ .

**Proof.** Given a copula  $\mathbf{C}$ , it is obvious from (38) that  $(\Delta_{i,j}(\mathbf{C}))$  is doubly stochastic. On the other hand, given a doubly stochastic matrix  $A$ , the proof of the theorem (2) shows that  $n^{-1}A \in \mathcal{C}_n$ . ■

This lemma justifies the following definition:

**Definition 14** Let  $n$  be positive integer and  $n\Delta = (n\Delta_{i,j})_{1 \leq i,j \leq n} \in \mathcal{D}_n$ . A checkerboard copula associated with  $\Delta$  is defined on  $[0, 1]^2$  by

$$\mathbf{C}(u, v) = n^2 \sum_{i=1}^n \sum_{j=1}^n \Delta_{i,j} \int_0^u \chi_{i,n}(x) dx \int_0^v \chi_{j,n}(y) dy \quad (39)$$

## 4.2 The Kendall's tau concordance measure

Calculations are very similar to these for Bernstein copulas. Let  $\Delta = (\Delta_{i,j}) \in \mathcal{D}_n$  and  $\mathbf{C}$  the associated checkerboard copula.

**Theorem 15** Let  $\Xi \in \mathcal{M}_n$  where

$$\xi_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 2 & \text{if } i > j \\ 0 & \text{if } i < j \end{cases} \quad (40)$$

then the Kendall's tau of the checkerboard copula is

$$\tau = 1 - \text{tr}(\Xi \Delta \Xi \Delta^\top) \quad (41)$$

**Proof.** Here we have

$$\frac{\partial \mathbf{C}}{\partial x}(u, v) = n^2 \sum_{i=1}^n \sum_{j=1}^n \Delta_{i,j} \left( \chi_{i,n}(u) \int_0^v \chi_{j,n}(y) dy \right) \quad (42)$$

such that double integral in (19) equals now to

$$n^4 \sum_{i=1}^n \sum_{j=1}^n \sum_{p=1}^n \sum_{q=1}^n \Delta_{i,j} \Delta_{p,q} \left( \int \int_{[0,1]^2} \left( \left( \chi_{i,n}(u) \int_0^v \chi_{j,n}(y) dy \right) \left( \chi_{q,n}(v) \int_0^u \chi_{p,n}(x) dx \right) \right) du dv \right) \quad (43)$$

Routine calculations give us

$$\int_0^1 \left( \chi_{i,n}(u) \int_0^u \chi_{p,n}(x) dx \right) du = \begin{cases} \frac{1}{2n^2} & \text{if } i = p \\ \frac{1}{n^2} & \text{if } i > p \\ 0 & \text{if } i < p \end{cases} \quad (44)$$

so that the Kendall's tau equals to

$$\tau = 1 - \sum_{i=1}^n \sum_{j=1}^n \sum_{p=1}^n \sum_{q=1}^n \Delta_{i,j} \Delta_{p,q} \xi_{i,p} \xi_{q,j} \quad (45)$$

which can be rewritten as

$$\tau = 1 - \text{tr}(\Xi \Delta \Xi \Delta^\top) \quad (46)$$

■

### 4.3 The Spearman's rho concordance measure

**Theorem 16** Let  $\Omega \in \mathcal{M}_n$  where

$$\omega_{i,j} = n^{-2} (2n - 2i + 1) (2n - 2j + 1) \quad (47)$$

then the Spearman's rho of the checkerboard copula is

$$\varrho_{\mathbf{C}} = 3 \text{tr}(\Omega \Delta) - 3 \quad (48)$$

**Proof.** Routine calculations give us

$$\iint_{[0,1]^2} \left( n^2 \sum_{i=1}^n \sum_{j=1}^n \Delta_{i,j} \int_0^u \chi_{i,n}(x) dx \int_0^v \chi_{j,n}(y) dy \right) du dv = n^2 \sum_{i=1}^n \sum_{j=1}^n \Delta_{i,j} \left( \frac{2n - 2i + 1}{2n^2} \right) \left( \frac{2n - 2j + 1}{2n^2} \right) \quad (49)$$

and the formula is easily derived. ■

### 4.4 The upper tail dependence measure

**Theorem 17** For any checkerboard copula  $\mathbf{C}$ , the upper tail dependence measure  $\lambda$  is

$$\lambda(\mathbf{C}) = 0 \quad (50)$$

**Proof.** We have

$$\frac{\partial \mathbf{C}}{\partial u}(u, u) = n^2 \sum_{i=1}^n \sum_{j=1}^n \Delta_{i,j} \chi_{i,n}(u) \int_0^u \chi_{j,n}(y) dy \quad (51)$$

and

$$\frac{\partial \mathbf{C}}{\partial v}(u, u) + \frac{\partial \mathbf{C}}{\partial u}(u, u) = n^2 \sum_{i=1}^n \sum_{j=1}^n \Delta_{i,j} \left( \chi_{i,n}(u) \int_0^u \chi_{j,n}(y) dy + \chi_{j,n}(u) \int_0^u \chi_{i,n}(x) dx \right) \quad (52)$$

so that

$$\lim_{u \rightarrow 1^-} \frac{\partial \mathbf{C}}{\partial u}(u, u) + \frac{\partial \mathbf{C}}{\partial v}(u, u) = 2 \quad (53)$$

We then obtain the desired result. ■

## 5 More results on the dependence measures

In this section, we show that the Kendall's tau and Spearman's rho of Bernstein and checkerboard copulas define new non-parametric estimators of  $\tau$  and  $\varrho$ . Then, we consider the  $\tau - \rho$  problem in the case of the Bernstein copulas. Finally, we introduce a perturbation method in order to specify upper tail dependence.

### 5.1 New non-parametric estimators for $\tau$ and $\varrho$ : the convergence problem

In this paragraph, we use the formulas obtained for Kendall's tau et Spearman's rho to derive non parametric estimators. Indeed, we examine the validity of the following statements:

**We know that  $B_n(\mathbf{C}) \rightarrow \mathbf{C}$  and  $C_n(\mathbf{C}) \rightarrow \mathbf{C}$  when  $n \rightarrow \infty$  in a certain sense. However, we have shown that  $\lambda(\mathfrak{B}_n(\mathbf{C})) = \lambda(\mathfrak{C}_n(\mathbf{C})) = 0$ . The Bernstein and checkerboard approximations does not preserve the upper tail dependence. Do we have a same conclusion for the Kendall's tau and the Spearman's rho concordance measures? The answer is no as we shall see in the sequel.**

**Theorem 18** *Let  $\mathbf{C}$  be a copula, then*

$$\begin{aligned} \varrho(\mathfrak{B}_n(\mathbf{C})) &\rightarrow \varrho(\mathbf{C}) \\ \varrho(\mathfrak{C}_n(\mathbf{C})) &\rightarrow \varrho(\mathbf{C}) \end{aligned} \quad (54)$$

when  $n \rightarrow \infty$ .

**Proof.** In LI, MIKUSINSKI, SHERWOOD and TAYLOR [1997], it is shown that  $\mathfrak{B}_n(\mathbf{C}) \rightarrow \mathbf{C}$  and  $\mathfrak{C}_n(\mathbf{C}) \rightarrow \mathbf{C}$  uniformly on  $[0, 1]^2$ , that is to say (for example)

$$\lim_{n \rightarrow \infty} \sup_{(x,y) \in [0,1]^2} |\mathfrak{B}_n(\mathbf{C})(x, y) - \mathbf{C}(x, y)| = 0 \quad (55)$$

Since properties of Lebesgue's integral, we have

$$\left| \iint_{[0,1]^2} \mathfrak{B}_n(\mathbf{C})(x, y) dx dy - \iint_{[0,1]^2} \mathbf{C}(x, y) dx dy \right| \rightarrow 0 \quad (56)$$

and thus

$$\varrho(\mathfrak{B}_n(\mathbf{C})) \rightarrow \varrho(\mathbf{C}) \quad (57)$$

The argument is the same for the checkerboard approximation  $\mathfrak{C}_n(\mathbf{C})$ . ■

**Theorem 19** *Let  $\mathbf{C}$  be a copula, then*

$$\begin{aligned}\tau(\mathfrak{B}_n(\mathbf{C})) &\longrightarrow \tau(\mathbf{C}) \\ \tau(\mathfrak{C}_n(\mathbf{C})) &\longrightarrow \tau(\mathbf{C})\end{aligned}\tag{58}$$

when  $n \longrightarrow \infty$ .

**Proof.** Recall the uniform convergence of  $\mathfrak{B}_n(\mathbf{C})$  to  $\mathbf{C}$ , it is shown in DEHEUVELS [1981] that this implies the weak convergence of the probability measures induced by  $\mathfrak{B}_n(\mathbf{C})$  to the probability measure induced by  $\mathbf{C}$ . As  $\mathbf{C}$  is continuous and bounded on  $[0, 1]^2$ , we have

$$\iint \mathbf{C} \, d\mu(\mathfrak{B}_n(\mathbf{C})) \longrightarrow \iint \mathbf{C} \, d\mu(\mathbf{C})\tag{59}$$

Let  $\epsilon$  be positive real number. For  $n$  large, we have

$$\left| \iint \mathbf{C} \, d\mu(\mathfrak{B}_n(\mathbf{C})) - \iint \mathbf{C} \, d\mu(\mathbf{C}) \right| \leq \frac{\epsilon}{2}\tag{60}$$

On the other hand, because of the uniform convergence of  $\mathfrak{B}_n(\mathbf{C})$  to  $\mathbf{C}$ , we have also for  $n$  large

$$\sup_{(x,y) \in [0,1]^2} |\mathfrak{B}_n(\mathbf{C})(x,y) - \mathbf{C}(x,y)| \leq \frac{\epsilon}{2}\tag{61}$$

so that

$$\left| \iint (\mathfrak{B}_n(\mathbf{C}) - \mathbf{C}) \, d\mu(\mathfrak{B}_n(\mathbf{C})) \right| \leq \frac{\epsilon}{2} \iint d\mu(\mathfrak{B}_n(\mathbf{C})) = \frac{\epsilon}{2}\tag{62}$$

We could then write

$$\begin{aligned}& \left| \iint \mathfrak{B}_n(\mathbf{C}) \, d\mu(\mathfrak{B}_n(\mathbf{C})) - \iint \mathbf{C} \, d\mu(\mathbf{C}) \right| \\ &= \left| \iint (\mathfrak{B}_n(\mathbf{C}) - \mathbf{C}) \, d\mu(\mathfrak{B}_n(\mathbf{C})) + \left( \iint \mathbf{C} \, d\mu(\mathfrak{B}_n(\mathbf{C})) - \iint \mathbf{C} \, d\mu(\mathbf{C}) \right) \right| \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} \\ &= \epsilon\end{aligned}\tag{63}$$

It comes that

$$\left| \iint \mathfrak{B}_n(\mathbf{C}) \, d\mu(\mathfrak{B}_n(\mathbf{C})) - \iint \mathbf{C} \, d\mu(\mathbf{C}) \right| \rightarrow 0\tag{64}$$

when  $n \longrightarrow \infty$  and this completes the proof. ■

To illustrate these two theorems, we have computed the values of these estimators for the Gaussian copula with  $\rho = 0.5$  and the upper Fréchet bound  $\mathbf{C}^+$ . Moreover, we have reported in the figure ?? the values taken by the Nelsen estimators (see NELSEN [1998] page 177):

$$\hat{\varrho} = \frac{12}{n^2 - 1} \sum_{i=1}^n \sum_{j=1}^n \left( \mathbf{C} \left( \frac{i}{n}, \frac{j}{n} \right) - \frac{i}{n} \frac{j}{n} \right)\tag{65}$$

and

$$\hat{\tau} = \frac{2n}{n-1} \sum_{i=2}^n \sum_{j=2}^n \sum_{p=1}^{i-1} \sum_{q=1}^{j-1} (\Delta_{i,j}(\mathbf{C}) \Delta_{p,q}(\mathbf{C}) - \Delta_{i,q}(\mathbf{C}) \Delta_{p,j}(\mathbf{C}))\tag{66}$$

For these two copulas, the convergence of the Bernstein and checkerboard estimators is more slowly than the Nelsen ones.

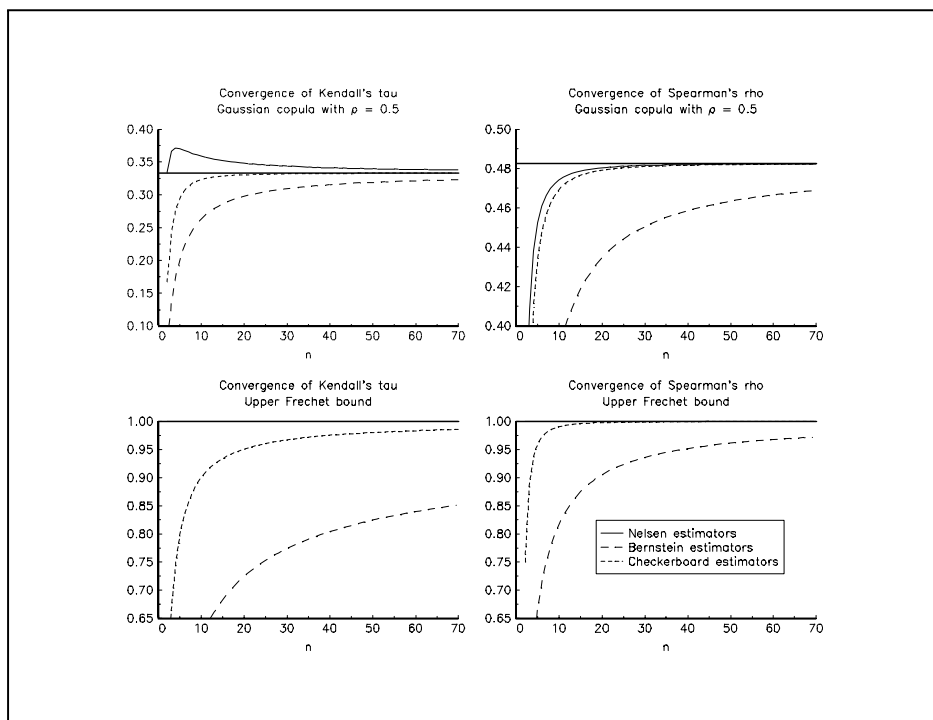


Figure 6: Convergence of the Nelsen, Bernstein and checkerboard estimators

## 5.2 The $\tau - \rho$ problem

We are now trying to obtain a regular copula which have fixed  $\tau$  and  $\varrho$ . We use Bernstein copulas for that purpose. In order to construct a Bernstein copula, we have to define the matrix  $D \in \mathcal{C}_n$  or equivalently the matrix  $A \in \mathcal{D}_n$ . Because  $\mathcal{D}_n$  is a convex set, we could define  $A$  in terms of the convex hull of  $\mathcal{D}_n$ . Let us denote  $\mathbf{P}_i$  the extreme points of  $\mathcal{D}_n$ . Thanks to Birkhoff's theorem (HORN and JOHNSON [1991]), we could write  $A$  as a positive convex combination of  $\mathbf{P}_i$

$$A = \sum_i \alpha_i \mathbf{P}_i \quad (67)$$

with  $\sum \alpha_i = 1$  and  $\alpha_i \geq 0$ . Because the convex hull of  $\mathcal{D}_n$  is finite and corresponds to the set of the permutation matrices  $(\Pi_i)$  for  $1 \leq i \leq n!$ , we have also

$$A = \sum_{i=1}^{n!} \lambda_i \Pi_i \quad (68)$$

with  $\sum_{i=1}^{n!} \lambda_i = 1$  and  $\lambda_i \geq 0$ . It comes that

$$D = \sum_{i=1}^{n!} \lambda_i \mathbf{E} \Pi_i \mathbf{E}^\top \quad (69)$$

with

$$\mathbf{E} = \frac{1}{\sqrt{n}} \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & & & 0 \\ 1 & 1 & \ddots & & 0 \\ \vdots & \vdots & \ddots & 0 & \vdots \\ 1 & 1 & \cdots & 1 & 0 \\ 1 & 1 & \cdots & 1 & 1 \end{bmatrix} \quad (70)$$

For a given pair  $(\tau, \varrho)$ , the vector  $\Lambda = (\lambda_i)$  must also match the following system

$$\begin{cases} \tau = 1 - \sum_{i=1}^{n!} \sum_{j=1}^{n!} \lambda_i \lambda_j \operatorname{tr} \left( \Theta \mathbf{E} \Pi_i \mathbf{E}^\top \Theta \mathbf{E} \Pi_j^\top \mathbf{E}^\top \right) \\ \varrho = 12 \sum_{i=1}^{n!} \lambda_i \operatorname{tr} \left( \Gamma \mathbf{E} \Pi_i \mathbf{E}^\top \right) - 3 \\ \sum_{i=1}^{n!} \lambda_i = 1 \\ \lambda_i \geq 0 \end{cases} \quad (71)$$

Since there are only 3 equations for  $n!$  degrees of freedom, the former system may have a solution. Once we found a solution  $\Lambda$ , the two place function

$$\mathbf{C}(u, v) = \sum_{i=1}^n \sum_{j=1}^n B_{i,n}(u) B_{j,n}(v) d_{i,j} \quad (72)$$

with  $D = \sum_{i=1}^{n!} \lambda_i \mathbf{E} \Pi_i \mathbf{E}^\top$  is the corresponding copula. Because  $\varrho(\mathfrak{B}_n(\mathbf{C})) \rightarrow \varrho(\mathbf{C})$  and  $\tau(\mathfrak{B}_n(\mathbf{C})) \rightarrow \tau(\mathbf{C})$ , it is theoretically possible to attain all the points of the bounding  $\tau - \varrho$  region (NELSEN [1998], page 144) if  $n$  is sufficiently large. Nevertheless, even if the previous problem is quadratic, there are some computational difficulties (for example, with  $n$  equal to 7, the dimension of  $\Lambda$  is 5040).

### 5.3 How to introduce an upper tail dependence?

Let us begin with a general result which explain why we often obtain  $\lambda(\mathbf{C}) = 0$ .

**Lemma 20** *Let  $\mathbf{C}$  be a copula. Suppose  $\mathbf{C}$  is  $C^1$  on the neighborhood of  $(1, 1)$ , then  $\lambda(\mathbf{C}) = 0$ .*

**Proof.** By virtue of the hypothesis,  $\frac{\partial \mathbf{C}}{\partial x}(1, 1)$  and  $\frac{\partial \mathbf{C}}{\partial y}(1, 1)$  exist and because of the probabilistic interpretation of the partial derivatives, we have

$$\begin{aligned} \frac{\partial \mathbf{C}}{\partial x}(1, 1) &= \Pr\{V \leq 1 \mid U = 1\} \\ \frac{\partial \mathbf{C}}{\partial y}(1, 1) &= \Pr\{U \leq 1 \mid V = 1\} \end{aligned} \quad (73)$$

where  $U$  and  $V$  are two uniform variables on  $[0, 1]$ . Then, it comes that

$$\frac{\partial \mathbf{C}}{\partial x}(1, 1) = \frac{\partial \mathbf{C}}{\partial y}(1, 1) = 1 \quad (74)$$

Because  $\mathbf{C}$  is  $C^1$ , we have

$$\lim_{u \rightarrow 1^-} \frac{d}{du} \mathbf{C}(u, u) = \frac{\partial \mathbf{C}}{\partial x}(1, 1) + \frac{\partial \mathbf{C}}{\partial y}(1, 1) = 2 \quad (75)$$

We obtain also the desired result

$$\lambda(\mathbf{C}) = 2 - \lim_{u \rightarrow 1^-} \frac{d}{du} \mathbf{C}(u, u) = 0 \quad (76)$$

■

**Remark 21** *We showed that under general regularity conditions  $\lambda(\mathbf{C}) = 0$ . This includes many cases for which the proof is more complicated (for instance the gaussian copula).*

As we saw above, our approximations of copulas (Bernstein, checkerboard) lead to an upper tail dependence which is equal to zero. This can be a real problem in financial applications. In this section, we try to slightly modify the copula  $\mathbf{C}$  in order to obtain a good upper tail dependence. In this direction, we propose to use a perturbation method. The perturbation will not be regular on the neighborhood of  $(1, 1)$  in order to have  $\lambda > 0$ . Let  $\epsilon(x, y)$  be a 2 place function and define a new copula  $\mathbf{C}'$  as follows

$$\mathbf{C}'(u, v) = \mathbf{C}(u, v) + \epsilon(u, v) \quad (77)$$

$\mathbf{C}'$  is then a copula if we verify the two properties:

1. the values of  $\epsilon(x, y)$  on the square  $\partial[0, 1]^2$  are 0;
2. the measure induced by  $\mathbf{C}'$  is positive.

## 6 Generating new families of copulas

### 6.1 Partition of unity

#### 6.1.1 The main result

The former two cases (Bernstein and checkerboard approximations) are in fact only examples of a more general family of copulas. This family is constructed via a *partition of unity*. This generalisation is done in LI, MIKUSINSKI, SHERWOOD and TAYLOR [1997]. Let us recall the following definition:

**Definition 22**  $\{\phi_1, \dots, \phi_n\} \in L^\infty([0, 1])$  is called a *partition of unity* if it satisfies the following statements

1.  $\phi_i(x) \geq 0$ ;
2.  $\int_0^1 \phi_i(x) dx = \frac{1}{n}$ ;
3.  $\sum_{i=1}^n \phi_i(x) = 1$  for all  $x \in [0, 1]$ .

We have then this important theorem:

**Theorem 23** Let  $\{\phi_1, \dots, \phi_n\} \in L^\infty([0, 1])$  be a partition of unity and  $\mathfrak{P}_n(\mathbf{C})$  the two place function defined by

$$\mathfrak{P}_n(\mathbf{C})(u, v) = n^2 \sum_{i=1}^n \sum_{j=1}^n \Delta_{i,j}(\mathbf{C}) \int_0^u \phi_i(x) dx \int_0^v \phi_j(y) dy \quad (78)$$

where  $\Delta_{i,j}(\mathbf{C}) = \mathbf{C}\left(\frac{i}{n}, \frac{j}{n}\right) - \mathbf{C}\left(\frac{i-1}{n}, \frac{j}{n}\right) - \mathbf{C}\left(\frac{i}{n}, \frac{j-1}{n}\right) + \mathbf{C}\left(\frac{i-1}{n}, \frac{j-1}{n}\right)$ . Then  $\mathfrak{P}_n(\mathbf{C})$  is a copula.



**Proof.** See LI, MIKUSINSKI, SHERWOOD and TAYLOR [1997]. ■

We will now use this theorem to generate new parametric families of copulas. For this, it suffices to take  $(\Delta_{i,j})_{1 \leq i,j \leq n}$  a doubly stochastic matrix (recall that  $n\Delta_{i,j}(\mathbf{C})$  is a doubly stochastic matrix) and consider the two place function

$$\mathbf{C}(x, y) = n \sum_{i=1}^n \sum_{j=1}^n \Delta_{i,j} \int_0^x \phi_i(u) du \int_0^y \phi_j(v) dv \quad (79)$$

with  $\{\phi_1, \dots, \phi_n\} \in L^\infty([0, 1])$  a partition of unity.  $\mathbf{C}$  is a copula and the parameters are  $\Delta_{i,j}$ .

**Remark 24** *The Bernstein copula corresponds to the following partition of unity*

$$\phi_i(x) = \binom{n}{i} x^i (1-x)^{n-i} \quad (80)$$

For the checkerboard copula, we have

$$\phi_i(x) = \chi_i(x) \quad (81)$$

where  $\chi_i$  stands for the indicator function of  $[\frac{i-1}{n}, \frac{i}{n}]$ .

### 6.1.2 Applications

Similar expressions of Spearman's rho and Kendall's tau can easily be derived for this general form. Let us define  $\Phi_i(x)$  as follows

$$\Phi_i(x) = \int_0^x \phi_i(u) du \quad (82)$$

Then, we could show this theorem:

**Theorem 25** *Let  $\Gamma \in \mathcal{M}_n$  where*

$$\gamma_{i,j} = \int_0^1 \phi_i(x) \Phi_j(x) dx \quad (83)$$

*then the Kendall's tau and the Spearman's rho associated to the copula induced by a partition of unity are*

$$\tau = 1 - 4n^2 \text{tr}(\Gamma \Delta \Gamma \Delta^\top) \quad (84)$$

and

$$\varrho = 3 - 12n (\mathbf{1}^\top \Gamma \Delta \Gamma \Delta^\top \mathbf{1}) \quad (85)$$

where  $\mathbf{1}$  is a vector of ones.

**Proof.** The computations are exactly the same as the proofs for Bernstein and checkerboard copulas. Let us define the vector  $\theta$  as follows

$$\theta = \left( \int_0^1 \Phi_i(u) du \right)_{1 \leq i \leq n} \quad (86)$$

To get the expression for  $\varrho$ , it suffices to remark that  $\varrho = \theta^\top \Delta \theta$ ,  $\mathbf{1}^\top \Gamma = \theta^\top$  and  $\Gamma \mathbf{1} = \frac{1}{n} \mathbf{1} - \theta$ . Indeed, we have

$$\begin{aligned} \int_0^1 \Phi_i(u) \, du &= \int_0^1 \left( \sum_{j=1}^n \phi_j(u) \, du \right) \Phi_i(u) \, du \\ &= \sum_{j=1}^n \int_0^1 \phi_j(u) \Phi_i(u) \, du \end{aligned} \quad (87)$$

thus  $\mathbf{1}^\top \Gamma = \theta^\top$ . We also have

$$(\Gamma \mathbf{1})_i = \left( \sum_{j=1}^n \int_0^1 \phi_i(u) \Phi_j(u) \, du \right)_{1 \leq i \leq n} \quad (88)$$

It comes that

$$\begin{aligned} (\Gamma \mathbf{1})_i &= \int_0^1 u \phi_i(u) \, du \\ &= \int_0^1 \phi_i(u) \, du - \int_0^1 \Phi_i(u) \, du \\ &= \frac{1}{n} - \theta_i \end{aligned} \quad (89)$$

Thus  $\Gamma \mathbf{1} = \frac{1}{n} \mathbf{1} - \theta$ . ■

Many other partitions of unity are feasible like

$$\phi_i(x) = \begin{cases} nx + 2 - i & \text{if } x \in \left[ \frac{i-2}{n}, \frac{i-1}{n} \right] \\ -nx + i & \text{if } x \in \left[ \frac{i-1}{n}, \frac{i}{n} \right] \end{cases} \quad \text{for } i \geq 2$$

and

$$\phi_1(x) = \begin{cases} -nx + 1 & \text{if } x \in \left[ 0, \frac{1}{n} \right] \\ 0 & \text{if } x \in \left[ \frac{1}{n}, 1 - \frac{1}{n} \right] \\ nx + 1 - n & \text{if } x \in \left[ 1 - \frac{1}{n}, 1 \right] \end{cases}$$

## 6.2 The special case $n = 2$

When  $n = 2$ , the general form of the doubly stochastic matrix is

$$\begin{bmatrix} \alpha & 1 - \alpha \\ 1 - \alpha & \alpha \end{bmatrix} \quad (90)$$

with  $\alpha \in [0, 1]$ . We define  $\varphi(x)$  such that  $0 \leq \varphi(x) \leq 1$ ,  $\int_0^1 \varphi(x) \, dx = \frac{1}{2}$ . Let  $\phi_1(x) = \varphi(x)$ ,  $\phi_2(x) = 1 - \varphi(x)$  and  $\Phi(x) = \int_0^x \varphi(u) \, du$ . The general form of the copula is then

$$\mathbf{C}(x, y) = 4(2\alpha - 1) \Phi(x) \Phi(y) + 2\alpha xy + 2(1 - 2\alpha)(x\Phi(y) + y\Phi(x)) \quad (91)$$

The expressions of Kendall's tau and the Spearman's rho become

$$\tau = 2(2\alpha - 1)(1 - 4A)^2 \quad (92)$$

and

$$\varrho = 3(2\alpha - 1)(1 - 4A)^2 \quad (93)$$

where  $A$  stands for  $\int_0^1 \Phi(u) du$ . Because of the conditions on  $\varphi$ , we have  $\frac{1}{8} \leq A \leq \frac{3}{8}$ . We remark that

$$\frac{\tau}{\varrho} = \frac{2}{3} \quad (94)$$

for any partition of order 2.

For example, for  $\varphi(x) = \sin^2\left(\frac{\pi}{2}x\right)$ , we obtain the copula

$$\mathbf{C}_\alpha(x, y) = xy + \frac{2\alpha - 1}{\pi^2} \sin(\pi x) \sin(\pi y) \quad (95)$$

for which

$$\varrho = \frac{48}{\pi^2} (2\alpha - 1) \quad (96)$$

and

$$\tau = \frac{32}{\pi^2} (2\alpha - 1) \quad (97)$$

**Remark 26** If  $\varphi(x) = x$ , then we have  $\Phi(x) = \frac{1}{2}x^2$ . It comes that the copula  $\mathbf{C}$  is the Farlie-Gumbel-Morgenstern copula.

### 6.3 The special case $n = 3$

We can apply the previous method to get copulas with more than one parameter. Let's take for the matrix  $\Delta$  the general form

$$\Delta = \begin{bmatrix} 1 - \beta - \gamma - \delta - \varepsilon & \beta + \gamma & \delta + \varepsilon \\ \beta + \varepsilon & 1 - \alpha - \beta - \gamma - \varepsilon & \alpha + \gamma \\ \gamma + \delta & \alpha + \varepsilon & 1 - \alpha - \gamma - \delta - \varepsilon \end{bmatrix} \quad (98)$$

with  $\alpha, \beta, \gamma, \varepsilon, \delta \geq 0$  and  $\alpha + \beta + \gamma + \varepsilon + \delta \leq 1$ . As the partition of unity, we take the functions  $\phi_1, \phi_2$  and  $1 - \phi_1 + \phi_2$ . We denote by  $\Phi_i(x) = \int_0^x \phi_i(u) du$ . By applying the formula (78) we get

$$\begin{aligned} \frac{\mathbf{C}(x, y)}{3} &= (2 - \alpha - \beta - 3\gamma - 4\delta - 3\varepsilon) \Phi_1(x) \Phi_1(y) + (1 - 2\alpha + \beta - 3\gamma - 2\delta) \Phi_1(y) \Phi_2(x) \\ &+ (1 - 2\alpha + \beta - 3\varepsilon - 2\delta) \Phi_1(x) \Phi_2(y) + (2 - 4\alpha - \beta - 3\gamma - \delta - 3\varepsilon) \Phi_2(x) \Phi_2(y) \\ &+ (-1 + \alpha + 2\gamma + 2\delta + \varepsilon) x \Phi_1(y) + (-1 + 2\alpha + \gamma + \delta + 2\varepsilon) x \Phi_2(y) \\ &+ (-1 + \alpha + \gamma + 2\delta + 2\varepsilon) y \Phi_1(x) + (-1 + 2\alpha + 2\gamma + \delta + \varepsilon) y \Phi_2(x) + (1 - \alpha - \gamma - \delta - \varepsilon) xy \end{aligned} \quad (99)$$

As an example, one can take

$$\begin{aligned} \phi_1(x) &= \frac{4\pi}{3} \frac{1}{1+x^2} \\ \phi_2(x) &= \frac{1}{3} \end{aligned} \quad (100)$$

which gives

$$\begin{aligned} \Phi_1(x) &= \frac{4\pi}{3} \arctan(x) \\ \Phi_2(x) &= \frac{x}{3} \end{aligned} \quad (101)$$

The associated copula has the simple form

$$\begin{aligned} \mathbf{C}(x, y) &= \frac{1}{3\pi^2} [4(-2 + \alpha + \beta + 3\gamma + 4\delta + 3\varepsilon) (\arctan(x) (\pi y - 4 \arctan(y)) + \pi x \arctan(y))] \\ &\quad - \frac{1}{3\pi^2} (-5 + \alpha + \beta + 3\gamma + 4\delta + 3\varepsilon) \pi^2 xy \end{aligned} \quad (102)$$

In fact, we notice that in this special case we have one parameter  $\zeta = \alpha + \beta + 3\gamma + 4\delta + 3\varepsilon$  and the copula we've just obtained can be rewritten as follows

$$\mathbf{C}(x, y) = \frac{1}{3\pi^2} [4(-2 + \zeta) (\arctan(x) (\pi y - 4 \arctan(y)) + \pi x \arctan(y)) + (5 - \zeta) \pi^2 xy]$$

After computation, we get the corresponding Spearman's rho and Kendall's tau

$$\begin{aligned} \varrho &= -\frac{(-2 + \zeta) (\pi - 4 \ln[2])^2}{\pi^2} \\ \tau &= -\frac{2(-2 + \zeta) (\pi - 4 \ln[2])^2}{3\pi^2} \end{aligned} \quad (103)$$

We can notice that once again we have the relationship

$$\tau = \frac{2}{3} \varrho \quad (104)$$

This relationship is striking. It makes us think of Nelsen's open question in DALL'AGLIO, KOTZ and SALINETTI [1991]. He considered a family of copulas  $\{\mathbf{C}_\theta(x, y)\}$  indexed by the parameter  $\theta$  (possibly multidimensional) such that  $\mathbf{C}_{\theta_0} = \mathbf{C}^\perp$  and  $\mathbf{C}_\theta$  is a continuous function of  $\theta$  at  $\theta_0$ . **Do we always have the relationship**

$$\lim_{\theta \rightarrow \theta_0} \frac{d\varrho}{d\tau} = \frac{3}{2} \quad (105)$$

**for such families?** We can weaken the question and try to see whether the above relationship holds for families of copulas for which the copula  $\mathbf{C}_{\theta_0}$  satisfies  $\varrho = \tau = 0$ . For this purpose, let us consider the copula

$$\mathbf{C}(x, y) = 3 \times \Phi(x)^\top \Delta \Phi(y) \quad (106)$$

where  $\Delta$  is the doubly stochastic matrix (98) and  $\Phi$  is the 3 dimensional vector  $\Phi = (\Phi_i)_{1 \leq i \leq 3}$  with

$$\Phi_i(x) = \int_0^x \chi_i(u) du \quad (107)$$

In fact we have used a checkerboard method for  $n = 3$ . After computation, we obtain the following values for the Spearman's rho and the Kendall's tau

$$\begin{aligned} \varrho &= \frac{4}{9} (2 - \alpha - \beta - 3\gamma - 4\delta - 3\varepsilon) \\ \tau &= \frac{2}{9} (3 - 6\delta + \gamma(-5 + \gamma + \delta) - 5\varepsilon + \varepsilon(\gamma + \delta) + \varepsilon^2 + \beta(-2 + \gamma + \delta + \varepsilon) + \alpha(-2 + \beta + \gamma + \delta + \varepsilon)) \end{aligned} \quad (108)$$

We can easily see that  $\varrho \neq \frac{3}{2}\tau$ . We will consider few special cases. First, let us consider the case  $\alpha = \beta = \gamma = \varepsilon = 0$ . We then get a one parameter family of copula. In this case we have ( $0 \leq \delta \leq 1$ )

$$\begin{aligned} \varrho &= \frac{8}{9} - \frac{16}{9}\delta \\ \tau &= \frac{2}{3} - \frac{4}{3}\delta \end{aligned} \quad (109)$$

The range of dependence for this family is  $\varrho \in [-\frac{8}{9}, \frac{8}{9}]$  and  $\tau \in [-\frac{2}{3}, \frac{2}{3}]$ . We also have the relationship

$$\varrho = \frac{4}{3}\tau \quad (110)$$

Moreover, for  $\delta = \frac{1}{2}$ , we have  $\varrho = \tau = 0$ , but the corresponding copula is not the product copula. So we have obtained a family of copula  $\mathbf{C}_\delta(x, y)$  which is a continuous function of  $\delta$  on  $[0, 1]$ , which satisfies  $\varrho = \tau = 0$  for  $\delta = \frac{1}{2}$ , and which does not satisfy the relationship  $\lim_{\delta \rightarrow \frac{1}{2}} \frac{d\varrho}{d\tau} = \frac{3}{2}$ .

Let's take other specifications of the parameters. For example, let us consider the case  $\alpha = \beta = \gamma = \delta = 0$ . We then get a one parameter family for which ( $0 \leq \varepsilon \leq 1$ )

$$\begin{aligned} \varrho &= \frac{8}{9} - \frac{4}{3}\varepsilon \\ \tau &= \frac{2}{3} - \frac{10}{9}\varepsilon + \frac{2}{9}\varepsilon^2 \end{aligned} \quad (111)$$

For  $\varepsilon = \frac{2}{3}$ ,  $\varrho = 0$  and  $\tau = \frac{20}{81}$ . If now we take  $\alpha = \beta = \varepsilon = \delta = 0$ , we have a family with parameter  $\gamma$  and for which we have ( $0 \leq \gamma \leq 1$ )

$$\begin{aligned} \varrho &= \frac{8}{9} - \frac{4}{3}\gamma \\ \tau &= \frac{2}{3} - \frac{10}{9}\gamma \end{aligned} \quad (112)$$

For  $\gamma = \frac{2}{3}$ , we have  $\varrho = 0$  and  $\tau = -\frac{2}{27}$ . For  $\gamma = \frac{3}{5}$   $\tau = 0$ , and  $\varrho = \frac{4}{45}$ .

In the figure 7, we have represented the  $\tau - \varrho$  relation for different families with only one or two parameters.

## 7 Using partitions of unity to construct new Markov sub-algebras

Let's recall the definition of a partition of unity:

**Definition 27**  $\{\phi_1, \dots, \phi_n\} \in L^\infty([0, 1])$  is called a partition of unity if it satisfies the following statements

1.  $\phi_i(x) \geq 0$ ;
2.  $\int_0^1 \phi_i(x) dx = \frac{1}{n}$ ;
3.  $\sum_{i=1}^n \phi_i(x) = 1$  for all  $x \in [0, 1]$ .

The following lemma is the result which gives the intuition of the main result of this section.

**Lemma 28** Let  $\{\phi_1, \dots, \phi_n\}$  be a partition of unity and let  $\Delta$  be a doubly stochastic matrix. Then

$$\psi = \begin{bmatrix} \psi_1 \\ \vdots \\ \psi_n \end{bmatrix} = \Delta \begin{bmatrix} \phi_1 \\ \vdots \\ \phi_n \end{bmatrix} \quad (113)$$

is another partition of unity.

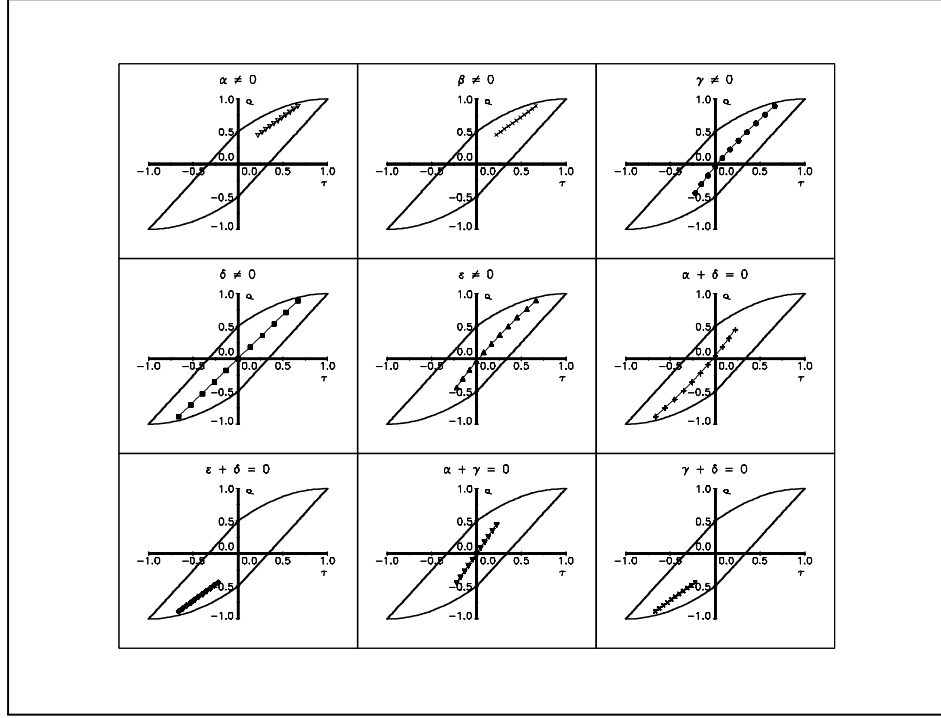


Figure 7:  $\tau - \rho$  diagram for different families

**Proof.** Let us denote  $\Delta = (\Delta_{i,j})_{1 \leq i,j \leq n}$ . We then have for all  $i \in \{1, \dots, n\}$

$$\psi_i(x) = \sum_{j=1}^n \Delta_{i,j} \phi_j(x) \quad (114)$$

It is easy to check that  $\{\psi_1, \dots, \psi_n\} \in L^\infty([0, 1])$ . Moreover, we have  $\psi_i(x) \geq 0$  and

$$\begin{aligned} \int_0^1 \psi_i(x) dx &= \sum_{j=1}^n \Delta_{i,j} \int_0^1 \phi_j(x) dx \\ &= \sum_{j=1}^n \Delta_{i,j} \times \frac{1}{n} \\ &= \frac{1}{n} \end{aligned} \quad (115)$$

because  $\Delta$  is a doubly stochastic matrix. We note also that

$$\begin{aligned} \sum_{i=1}^n \psi_i(x) &= \sum_{i=1}^n \sum_{j=1}^n \Delta_{i,j} \phi_j(x) \\ &= \sum_{j=1}^n \phi_j(x) \left( \sum_{i=1}^n \Delta_{i,j} \right) \\ &= \sum_{j=1}^n \phi_j(x) \\ &= 1 \end{aligned} \quad (116)$$

This completes the proof. ■

This result gives us the idea that if  $\{\phi_1, \dots, \phi_n\}$  and  $\{\psi_1, \dots, \psi_n\}$  are two partitions of unity, then the two place function defined by

$$n \times \begin{bmatrix} \Phi_1(x) & \cdots & \Phi_n(x) \end{bmatrix} \begin{bmatrix} \Psi_1(y) \\ \vdots \\ \Psi_n(y) \end{bmatrix} \quad (117)$$

is a copula, where  $\Phi_i(x) = \int_0^x \phi_i(u) du$  and  $\Psi_i(x) = \int_0^x \psi_i(u) du$ . In fact, this result is always true.

**Proposition 29** *If  $\{\phi_1, \dots, \phi_n\}$  and  $\{\psi_1, \dots, \psi_n\}$  are two partitions of unity, then*

$$\mathbf{C}(x, y) = n \times \begin{bmatrix} \Phi_1(x) & \cdots & \Phi_n(x) \end{bmatrix} \begin{bmatrix} \Psi_1(y) \\ \vdots \\ \Psi_n(y) \end{bmatrix} = n \sum_{i=1}^n \Phi_i(x) \Psi_i(y) \quad (118)$$

is a copula, where  $\Phi_i(x) = \int_0^x \phi_i(u) du$  and  $\Psi_i(x) = \int_0^x \psi_i(u) du$ .

**Proof.** First we have to notice that  $\Phi_i(1) = \Psi_i(1) = \frac{1}{n}$ ,  $\Phi_i(0) = \Psi_i(0) = 0$  and  $\sum_{i=1}^n \Phi_i(x) = \sum_{i=1}^n \Psi_i(x) = x$ . The bounding conditions are satisfied:

$$\begin{aligned} \mathbf{C}(x, 0) &= 0 \\ \mathbf{C}(0, y) &= 0 \\ \mathbf{C}(x, 1) &= n \sum_{i=1}^n \Phi_i(x) \Psi_i(1) = \sum_{i=1}^n \Phi_i(x) = x \\ \mathbf{C}(1, y) &= n \sum_{i=1}^n \Phi_i(1) \Psi_i(y) = \sum_{i=1}^n \Psi_i(y) = y \end{aligned} \quad (119)$$

Moreover, the measure induced by  $\mathbf{C}$  has the following density

$$\frac{\partial^2}{\partial x \partial y} \mathbf{C}(x, y) = n \sum_{i=1}^n \phi_i(x) \psi_i(y) \geq 0 \quad (120)$$

and this completes the proof. ■

Let's denote by  $\mathfrak{A}_n$  the set

$$\begin{aligned} \mathfrak{A}_n &= \{n \times \Phi^\top \Psi\} \\ &= \left\{ n \times \sum_{i=1}^n \Phi_i(x) \Psi_i(y) \right\} \end{aligned} \quad (121)$$

with  $\{\phi_1, \dots, \phi_n\}$  and  $\{\psi_1, \dots, \psi_n\}$  two partitions of unity and  $\Phi_i(x) = \int_0^x \phi_i(u) du$  and  $\Psi_i(x) = \int_0^x \psi_i(u) du$ .  $\mathfrak{A}_n$  is a compact subset of copulas functions. We shall see that  $\mathfrak{A}_n$  is stable with respect to the \*-product introduced by DARSOW, NGUYEN and OLSEN [1992]. They consider the following product of copulas

$$(\mathbf{A} * \mathbf{B})(x, y) = \int_0^1 \frac{\partial \mathbf{A}}{\partial y}(x, s) \frac{\partial \mathbf{B}}{\partial x}(s, y) ds \quad (122)$$

They show that for any copulas  $\mathbf{A}$  and  $\mathbf{B}$ ,  $\mathbf{A} * \mathbf{B}$  is still a copula. Now let us examine the effect of the  $*$ -product on the subset  $\mathfrak{A}_n$ . We take  $\mathbf{A}(x, y) = nF^\top(x)G(y)$  and  $\mathbf{B}(x, y) = nH^\top(x)K(y)$ . Then we have

$$(\mathbf{A} * \mathbf{B})(x, y) = n^2 \times \int_0^1 F^\top(x)g(s)h^\top(s)K(y) ds \quad (123)$$

where  $g(s) = \frac{d}{ds}G(s)$  and  $h(s) = \frac{d}{ds}H(s)$ . Thus, it comes that

$$\begin{aligned} (\mathbf{A} * \mathbf{B})(x, y) &= n^2 \times F^\top(x) \left( \int_0^1 g(s)h^\top(s) ds \right) K(y) \\ &= n \times F^\top(x) \left( n \times \int_0^1 g(s)h^\top(s) ds \right) K(y) \end{aligned} \quad (124)$$

We can show that  $n \times \int_0^1 g(s)h^\top(s) ds$  is a doubly stochastic matrix. Indeed, we have

$$\begin{aligned} \left[ n \times \int_0^1 g(s)h^\top(s) ds \right] &= \left[ n \times \int_0^1 g_i(s)h_j(s) ds \right]_{1 \leq i, j \leq n} \\ &= \Delta = [\Delta_{i,j}]_{1 \leq i, j \leq n} \end{aligned} \quad (125)$$

if we denote the two vectors  $g(s)$  and  $h(s)$  by  $g(s) = \begin{bmatrix} g_1(s) \\ \vdots \\ g_n(s) \end{bmatrix}$  and  $h(s) = \begin{bmatrix} h_1(s) \\ \vdots \\ h_n(s) \end{bmatrix}$ . By definition,  $g(s)$  and  $h(s)$  are two partitions of unity. Moreover, we have

$$\begin{aligned} \sum_{j=1}^n \Delta_{i,j} &= n \times \sum_{j=1}^n \int_0^1 g_i(s)h_j(s) ds = n \times \int_0^1 g_i(s) \sum_{j=1}^n h_j(s) ds = n \times \int_0^1 g_i(s) ds = 1 \\ \sum_{i=1}^n \Delta_{i,j} &= n \times \sum_{i=1}^n \int_0^1 g_i(s)h_j(s) ds = n \times \int_0^1 h_j(s) \sum_{i=1}^n g_i(s) ds = n \times \int_0^1 h_j(s) ds = 1 \end{aligned} \quad (126)$$

So  $\Delta = n \times \int_0^1 g(s)h^\top(s) ds$  is a doubly stochastic matrix. Note that

$$\begin{aligned} (\mathbf{A} * \mathbf{B})(x, y) &= n \times F^\top(x) \left( n \times \int_0^1 g(s)h^\top(s) ds \right) K(y) \\ &= n \times F^\top(x) \Delta K(y) \end{aligned} \quad (127)$$

By applying the lemma 28, the proposition 29 and by noting  $L(y) = \Delta K(y)$ , then we deduce that  $n \times F^\top(x)L(y)$  is a copula. So  $(\mathbf{A} * \mathbf{B})$  is still an element of  $\mathfrak{A}_n$ . Thus  $\mathfrak{A}_n$  is stable by the  $*$ -product. So  $\mathfrak{A}_n$  is a Markov sub-algebra of the algebra of all copulas with respect to the  $*$ -product.

## 8 Conclusion

In this paper, we have investigated the problem of copulas approximation. We have shown in particular that concordances measures of the Bernstein and the checkerboard copulas converge to the true one. However, these approximations do not preserve the upper tail dependence.

These two approximations are a special case of *partition of unity*. In the second part of the paper, we have studied the associated families. Moreover, we have shown that the set of copulas induced by *partition of unity* is a Markov sub-algebra with respect to the  $*$ -product of DARSOW, NGUYEN and OLSEN [1992]. In a coming paper, we study this sub-algebra in a more detailed way (in particular, we have investigated the characterization of some special elements such as the idempotent elements, the invertible elements, and so on) and show how *partitions of unity* could be used to construct appropriate Markov processes.



## References

- [1] BENEŠ, V. and J. ŠTĚPÁN [1997], Distributions with Given Marginals and Moment Problems, Kluwer Academic Publishers, Dordrecht
- [2] DALL'AGLIO, G., S. KOTZ and G. SALINETTI [1991], Advances in Probability Distributions with Given Marginals (Beyond the Copulas), Kluwer Academic Publishers, Dordrecht
- [3] DARSOW, W.F., B. NGUYEN and E.T. OLSEN [1992], Copulas and markov processes, *Illinois Journal of Mathematics*, **36-4**, 600-642
- [4] DEHEUVELS, P. [1981], A non parametric test for independence, *Publications de l'Institut de Statistique de l'Université de Paris*, **26**, 29-50
- [5] DURRLEMAN, V., A. NIKEGHBALI, and T. RONCALLI [2000], Which copula is the right one?, Groupe de Recherche Opérationnelle, Crédit Lyonnais, *Working Paper*
- [6] LI, X., P. MIKUSIŃSKI, H. SHERWOOD and M.D. TAYLOR [1997], On approximation of copulas, in V. Beneš and J. Štěpán (Eds.), *Distributions with Given Marginals and Moment Problems*, Kluwer Academic Publishers, Dordrecht
- [7] LI, D.X., P. MIKUSIŃSKI and M.D. TAYLOR [1998], Strong approximation of copulas, *Journal of Mathematical Analysis and Applications*, **225**, 608-623
- [8] KULPA, T. [1999], On approximation of copulas, *International Journal of Mathematics and Mathematical Sciences*, **22**, 259-269
- [9] NELSEN, R.B. [1991], Copulas and association, in G. Dall'Aglio, S. Kotz and G. Salinetti (Eds.), *Advances in Probability Distributions with Given Marginals (Beyond the Copulas)*, Kluwer Academic Publishers, Dordrecht
- [10] NELSEN, R.B. [1998], An Introduction to Copulas, *Lectures Notes in Statistics*, **139**, Springer Verlag, New York