

# Data driven partition-of-unity copulas with applications to risk management

CEQURA Conference on

Advances in Financial and Insurance Risk Management

Sept. 25 - 26, 2017, Munich

Dietmar Pfeifer<sup>1</sup>

Andreas Mändle<sup>1</sup>

Olena Ragulina<sup>2</sup>

---

<sup>1</sup> University of Oldenburg, Dept. of Mathematics, Germany

<sup>2</sup> Taras Shevchenko National University of Kyiv, Department of Probability Theory,  
Statistics and Actuarial Mathematics, Ukraine

# Agenda

1. Introduction & formal framework
2. Construction from given data
3. Case studies
4. Extension to arbitrary dimensions
5. Bibliography / References

## 1. Introduction & formal framework

### Motivation:

- Infinite partition-of-unity copulas recently introduced in Pfeifer et al. (2016)
- Construction of new multivariate copulas on the basis of a generalized infinite partition-of-unity approach (extendable to the uncountable infinite case)
- Construction allows for tail-dependence as well as for asymmetry
- Can be easily implemented for risk management purposes
- Particular interest: how to fit such copulas to highly asymmetric data?

## 1. Introduction & formal framework

Formal framework:

Let  $\mathbb{Z}^+ = \{0, 1, 2, 3, \dots\}$  and suppose that  $\{\varphi_i(u)\}_{i \in \mathbb{Z}^+}$  and  $\{\psi_j(v)\}_{j \in \mathbb{Z}^+}$  are non-negative maps defined on  $(0, 1)$  such that:

$$\sum_{i=0}^{\infty} \varphi_i(u) = \sum_{j=0}^{\infty} \psi_j(v) = 1 \quad (1)$$

$$\alpha_i := \int_0^1 \varphi_i(u) du > 0, \quad \beta_j := \int_0^1 \psi_j(v) dv > 0, \quad i, j \in \mathbb{Z}^+. \quad (2)$$

- $\{\varphi_i(u)\}_{i \in \mathbb{Z}^+}$  and  $\{\psi_j(v)\}_{j \in \mathbb{Z}^+}$  can be thought of representing discrete distributions over  $\mathbb{Z}^+$  with parameters  $u$  and  $v$ , resp.
- The sequences  $\{\alpha_i\}_{i \in \mathbb{Z}^+}$  and  $\{\beta_j\}_{j \in \mathbb{Z}^+}$  represent the probabilities of the corresponding mixed distributions.

## 1. Introduction & formal framework

Formal framework:

Let  $\{p_{ij}\}_{i,j \in \mathbb{Z}^+}$  represent the probabilities of an arbitrary discrete bivariate distribution over  $\mathbb{Z}^+ \times \mathbb{Z}^+$  with marginal distributions given by

$$p_{i\cdot} = \sum_{j=0}^{\infty} p_{ij} = \alpha_i \text{ and } p_{\cdot j} = \sum_{i=0}^{\infty} p_{ij} = \beta_j \text{ for } i, j \in \mathbb{Z}^+. \quad (3)$$

Then

$$c(u, v) := \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} p_{ij} \frac{\varphi_i(u)}{\alpha_i} \frac{\psi_j(v)}{\beta_j}, \quad u, v \in (0, 1) \quad (4)$$

defines the density of a bivariate copula, called (infinite) partition-of-unity copula.

## 1. Introduction & formal framework

Formal framework:

From a "dual" point of view, we can rewrite (4) as

$$c(u, v) := \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} p_{ij} \frac{\varphi_i(u)}{\alpha_i} \frac{\psi_j(v)}{\beta_j} = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} p_{ij} f_i(u) g_j(v), \quad u, v \in (0, 1) \quad (5)$$

where

$$f_i(\cdot) = \frac{\varphi_i(\cdot)}{\alpha_i} \quad \text{and} \quad g_j(\cdot) = \frac{\psi_j(\cdot)}{\beta_j}, \quad i, j \in \mathbb{Z}^+ \quad (6)$$

denote the densities induced by  $\{\varphi_i(u)\}_{i \in \mathbb{Z}^+}$  and  $\{\psi_j(v)\}_{j \in \mathbb{Z}^+}$ . This means that the copula density  $c(u, v)$  can also be seen as a mixture of product densities.

## 1. Introduction & formal framework

Formal framework:

Example 1 (Binomial distributions – Bernstein copula):

For fixed integers  $a, b \geq 2$ , consider the family of binomial distributions given by their point masses

$$\varphi_{a,i}(u) = \begin{cases} \binom{a-1}{i} u^i (1-u)^{a-1-i}, & i = 0, \dots, a-1 \\ 0, & i \geq a \end{cases} \quad (7)$$

and  $\psi_{b,j}(v) = \varphi_{b,j}(v)$  for  $i, j \in \mathbb{Z}^+$  and  $(u, v) \in (0, 1)$ .

We have

## 1. Introduction & formal framework

Formal framework:

Example 1 (Binomial distributions – Bernstein copula):

$$\alpha_{a,i} = \int_0^1 \varphi_{a,i}(u) du = \frac{1}{a}, \quad \beta_{b,j} = \int_0^1 \psi_{b,j}(v) dv = \frac{1}{b}, \quad (8)$$

$f_{a,i}$  and  $g_{b,j}$  are densities of a beta distribution with parameters  $(i, a+1-i)$  and  $(j, b+1-j)$  resp.,  $p_{i\cdot} = \frac{1}{a}$  and  $p_{\cdot j} = \frac{1}{b}$ , so

$$c_{a,b}(u,v) = ab \sum_{i=0}^a \sum_{j=0}^b p_{ij} \binom{a-1}{i} \binom{b-1}{j} u^{i-1} (1-u)^{a-i} v^{j-1} (1-v)^{b-j}, \quad u, v \in (0,1) \quad (9)$$

which is the density of a bivariate Bernstein copula.



## 1. Introduction & formal framework

Formal framework:

Example 2 (Negative binomial distributions):

For fixed integers  $a, b \geq 2$ , consider the family of negative binomial distributions given by their point masses

$$\varphi_{a,i}(u) = \binom{a+i-1}{i} u^i (1-u)^a, \quad (10)$$

and  $\psi_{b,j}(v) = \varphi_{b,j}(v)$  for  $i, j \in \mathbb{Z}^+$  and  $(u, v) \in (0, 1)$ .

We have

## 1. Introduction & formal framework

Formal framework:

Example 2 (Negative binomial distributions):

$$\alpha_{a,i} = \int_0^1 \varphi_{a,i}(u) du = \frac{a}{(a+i)(a+i+1)}, \beta_{b,j} = \frac{b}{(b+j)(b+j+1)}, \quad (11)$$

$f_{a,i}$  and  $g_{b,j}$  are densities of a beta distribution with parameters  $(i+1, a+1)$  and  $(j+1, b+1)$ ,  $p_{i\cdot} = \frac{a}{(a+i)(a+1+i)}$ ,  $p_{\cdot j} = \frac{b}{(b+j)(b+1+j)}$ , so

$$c_{a,b}(u,v) = (a+1)(b+1) \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} p_{ij} \binom{a+i+1}{i} \binom{b+j+1}{j} u^i (1-u)^a v^j (1-v)^b, u, v \in (0,1). \quad (12)$$

## 1. Introduction & formal framework

Formal framework:

Example 2 (Negative binomial distributions):

Negative binomial copulas typically show a tail dependence:

$\beta$	1	2	3	4	5	6	7	8	9	10
$\lambda_U(\beta)$	$\frac{1}{2}$	$\frac{5}{8}$	$\frac{11}{16}$	$\frac{93}{128}$	$\frac{193}{256}$	$\frac{793}{1024}$	$\frac{1619}{2048}$	$\frac{26333}{32768}$	$\frac{53381}{65536}$	$\frac{215955}{262144}$

$$\text{with } \lambda_U(\beta) = \lim_{t \uparrow 1} \frac{\int_t^1 \int_t^1 c_\beta(u,v) du dv}{1-t} = \frac{2\Gamma(2\beta)}{\Gamma^2(\beta)} \cdot \int_0^1 \int_0^1 \frac{x^\beta y^\beta}{(x+y)^{2\beta+1}} dx dy = 1 - \frac{\binom{2\beta}{\beta}}{4^\beta} \sim 1 - \frac{1}{\sqrt{\pi\beta}}$$

for large  $\beta$ .

## 1. Introduction & formal framework

Formal framework:

Example 3 (Poisson distributions):

For fixed  $a, b > 0$  consider the family of Poisson distributions given by their point masses

$$\varphi_{a,i}(u) = (1-u)^a \frac{a^i L(u)^i}{i!}, \quad (13)$$

$$L(u) := -\ln(1-u), \quad \psi_{b,j}(v) = \varphi_{b,j}(v), \quad i, j \in \mathbb{Z}^+, \quad (u, v) \in (0, 1).$$

We have

## 1. Introduction & formal framework

Formal framework:

Example 3 (Poisson distributions):

$$\alpha_{a,i} = \int_0^1 \varphi_{a,i}(u) du = \left(\frac{a}{a+1}\right)^i \left(1 - \frac{a}{a+1}\right), \beta_{b,j} = \left(\frac{b}{b+1}\right)^j \left(1 - \frac{b}{b+1}\right) \quad (14)$$

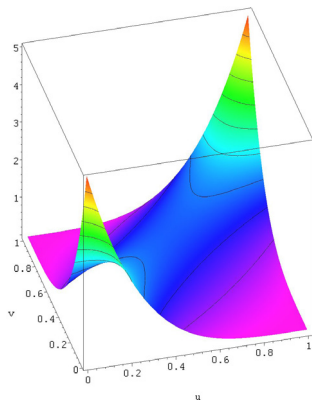
which correspond to geometric distributions over  $\mathbb{Z}^+$  with means  $a$  and  $b$ ,

$$p_{i\cdot} = \left(\frac{a}{a+1}\right)^i \left(1 - \frac{a}{a+1}\right) = \frac{a^i}{(a+1)^{i+1}}, p_{\cdot j} = \frac{b^j}{(b+1)^{j+1}}, i, j \in \mathbb{Z}^+, \quad (15)$$

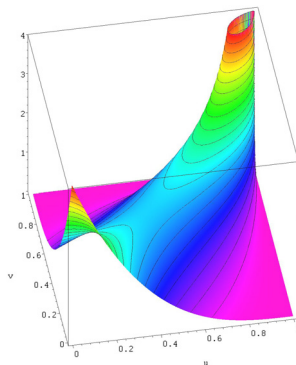
$$c_{a,b}(u,v) = (a+1)(b+1) \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} p_{ij} \frac{(a+1)^i (b+1)^j}{i! j!} L^i(u) (1-u)^a L^j(v) (1-v)^b, u, v \in (0,1). \quad (16)$$

## 1. Introduction & formal framework

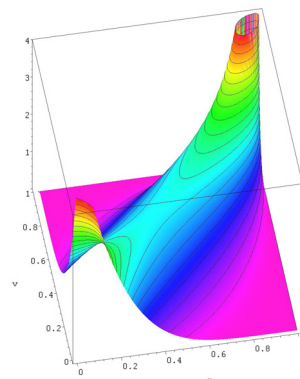
Formal framework:



Bernstein copula,  $m = 3$ ;  
no tail dependence



Negative binomial copula,  $\beta = 3$ ;  
 $\lambda_{\nu}(\beta) = 0.6875$



Poisson copula,  $\gamma = 5$   
no tail dependence

## 1. Introduction & formal framework

Formal framework:

Remark: Sklar's theorem provides a general method to construct pairs of discrete r.v.'s  $(X, Y)$  with joint probabilities  $p_{ij} = P(X = i, Y = j)$  and marginal probabilities  $\{\alpha_i\}_{i \in \mathbb{Z}^+}$  and  $\{\beta_j\}_{j \in \mathbb{Z}^+}$  :

Assume quantile functions  $Q_X, Q_Y$  of  $X, Y$  and a pair of rv's  $(U, V)$  with a given copula  $\tilde{C}$ . Then  $(X, Y) = (Q_X(U), Q_Y(V))$  has joint probabilities

$$\begin{aligned}
 p_{ij} &= P(X = i, Y = j) = P\left(\sum_{k=0}^{i-1} \alpha_k < U \leq \sum_{k=0}^i \alpha_k, \sum_{k=0}^{j-1} \beta_k < V \leq \sum_{k=0}^j \beta_k\right) \\
 &= \tilde{C}\left(\sum_{k=0}^i \alpha_k, \sum_{k=0}^j \beta_k\right) + \tilde{C}\left(\sum_{k=0}^{i-1} \alpha_k, \sum_{k=0}^{j-1} \beta_k\right) - \tilde{C}\left(\sum_{k=0}^{i-1} \alpha_k, \sum_{k=0}^j \beta_k\right) - \tilde{C}\left(\sum_{k=0}^i \alpha_k, \sum_{k=0}^{j-1} \beta_k\right).
 \end{aligned} \tag{17}$$

## 1. Introduction & formal framework

Formal framework:

Idea: use appropriate continuous extensions  $\tilde{C}$  of the empirical copula for modeling the  $\{p_{ij}\}_{i,j \in \mathbb{Z}^+}$  (cf. Bernstein approach).

**Lemma 1:** Let  $(U, V)$  be a pair of rv's with given copula  $\tilde{C}$ . Then the  $(X, Y)$  with  $\{p_{ij}\}_{i,j \in \mathbb{Z}^+}$  as joint probabilities from Examples 1, 2 and 3 can be constructed as follows (note:  $\lceil z \rceil = \min\{x \in \mathbb{R} \mid x \geq z\}, \lfloor z \rfloor = \max\{x \in \mathbb{R} \mid x \leq z\}$ ):

Example 1:  $X = \lceil aU \rceil, Y = \lceil bV \rceil,$

Example 2:  $X = \left\lfloor \frac{aU}{1-U} \right\rfloor, Y = \left\lfloor \frac{bV}{1-V} \right\rfloor,$

Example 3:  $X = \left\lfloor \frac{-\ln(1-U)}{\ln(a+1) - \ln a} \right\rfloor, Y = \left\lfloor \frac{-\ln(1-V)}{\ln(b+1) - \ln b} \right\rfloor.$



## 2. Construction from given data

Assumptions:

- rv's  $(X_i, Y_i), i = 1, \dots, n$  iid pairs with pairwise copula  $C$
- continuous marginal distributions (no ties!)
- $\mathbf{R}_X = (R_{11}, \dots, R_{1n})^T$  and  $\mathbf{R}_Y = (R_{21}, \dots, R_{2n})^T$  being the ranks of the vectors  $\mathbf{X} = (X_1, \dots, X_n)$  and  $\mathbf{Y} = (Y_1, \dots, Y_n)$ , resp.

The empirical copula is usually identified with the point set of relative ranks, i.e.  $\left\{ \left( \frac{r_{11}}{n+1}, \frac{r_{21}}{n+1} \right), \dots, \left( \frac{r_{1n}}{n+1}, \frac{r_{2n}}{n+1} \right) \right\}$ .

For the construction of appropriate  $\{p_{ij}\}_{i,j \in \mathbb{Z}^+}$  we need . . .

## 2. Construction from given data

**Lemma 2:** Let  $C_1, \dots, C_n$  be arbitrary bivariate copulas with densities  $c_1, \dots, c_n$  and  $(U_i, V_i)$  independent random vectors with the copula  $C_i$  for each pair  $(U_i, V_i)$ ,  $i = 1, \dots, n$ . Let further  $\mathbf{r}_1 = (r_{11}, \dots, r_{1n})^T$  and  $\mathbf{r}_2 = (r_{21}, \dots, r_{2n})^T$  be arbitrary permutations of  $(1, 2, \dots, n)^T$  and the random variable  $I$  follow a discrete uniform distribution over the set  $\{1, 2, \dots, n\}$ , independent of the  $(U_i, V_i)$  for  $i = 1, \dots, n$ . Then the random vector  $(U, V)$  defined by

$$U := \frac{r_{1I} - 1 + U_I}{n}, \quad V := \frac{r_{2I} - 1 + V_I}{n} \quad (18)$$

has continuous marginal uniform distributions over  $(0, 1)$  and density

$$c(u, v) = n \sum_{k=1}^n \mathbb{1}_{\left(\frac{r_{1k}-1}{n}, \frac{r_{1k}}{n}\right]}(u) \cdot \mathbb{1}_{\left(\frac{r_{2k}-1}{n}, \frac{r_{2k}}{n}\right]}(v) \cdot c_k(nu - r_{1k} + 1, nv - r_{2k} + 1), \quad u, v \in (0, 1). \quad (19)$$

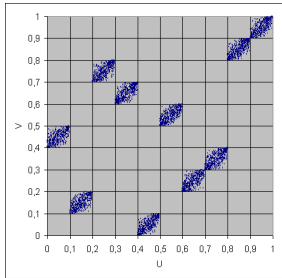
## 2. Construction from given data

To obtain a realization of  $(U, V)$  first select a pair  $(r_{1i}, r_{2i})$  from the set of all permutation pairs by a discrete uniform distribution over  $\{1, 2, \dots, n\}$  and then draw a sample from  $C_i$  rescaled to the interval  $\left(\frac{r_{1i}-1}{n}, \frac{r_{1i}}{n}\right] \times \left(\frac{r_{2i}-1}{n}, \frac{r_{2i}}{n}\right]$ .

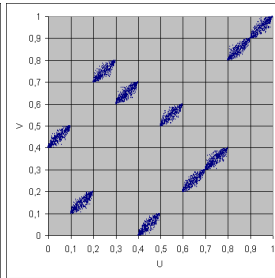
This corresponds to a particular patchwork copula construction, see e.g. Durante et al. (2013).

The following graphs show different realizations of such a construction for  $n = 10$  and  $\mathbf{r}_1 = (3, 1, 4, 2, 8, 6, 5, 7, 9, 10)^T$  and  $\mathbf{r}_2 = (8, 5, 7, 2, 4, 6, 1, 3, 9, 10)^T$ , with local Gaussian copulas for given fixed pairwise correlation  $\rho$ :

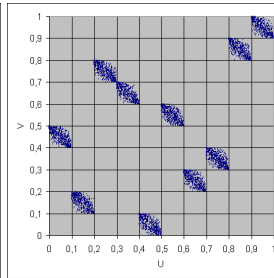
## 2. Construction from given data



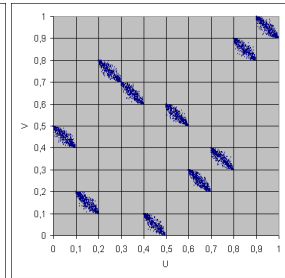
$$\rho = 0.75$$



$$\rho = 0.90$$



$$\rho = -0.75$$

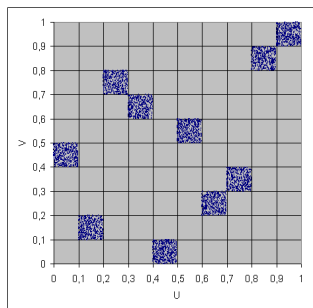


$$\rho = -0.90$$

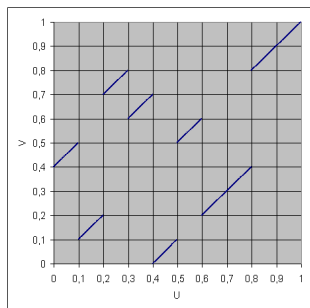
## 2. Construction from given data

Models of particular interest:

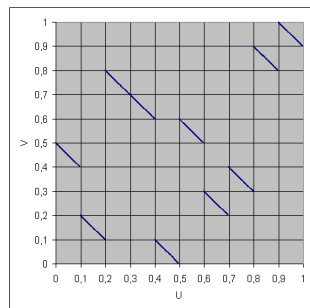
For the rook copula see Cottin and Pfeifer (2014); for the so-called shuffles of  $M$  (Fréchet shuffles) see e.g. Nelsen (2007), chapter 3.2.3.



$\rho = 0$   
rook copula



$\rho = 1$   
upper Fréchet shuffle

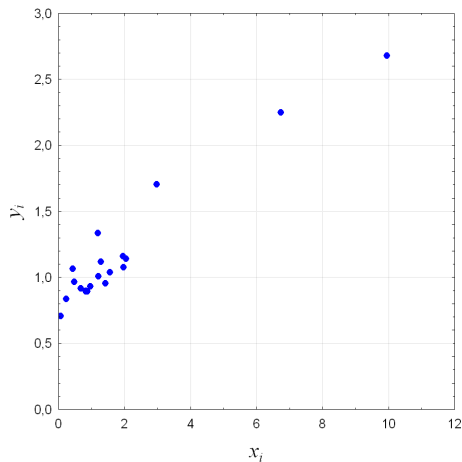


$\rho = -1$   
lower Fréchet shuffle

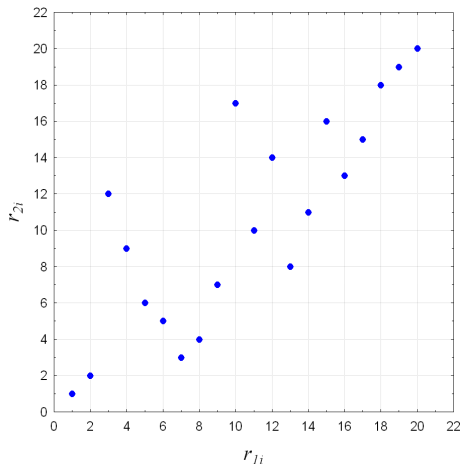
### 3. Case studies

- Data set treated in Cottin and Pfeifer (2014), Example 4.2 and Pfeifer et al. (2016), Section 4.
- Effects of the kind of dependence modeling (w/ or w/o upper tail dependence) on the V@R for the aggregated portfolio with various risk levels; similarly to Maciag et al. (2016)

### 3. Case studies



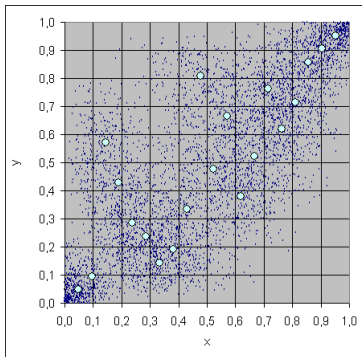
scatterplot of original data



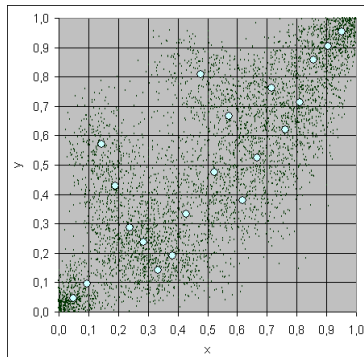
scatterplot of ranks

### 3. Case studies

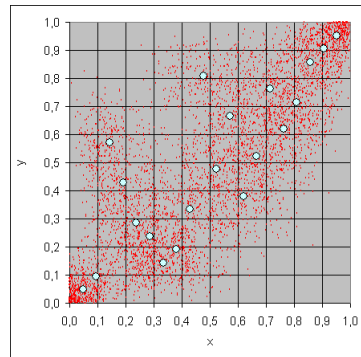
5,000 simulated pairs of the data-driven copulas and empirical copula  
(large points):



upper Fréchet shuffle



rook copula



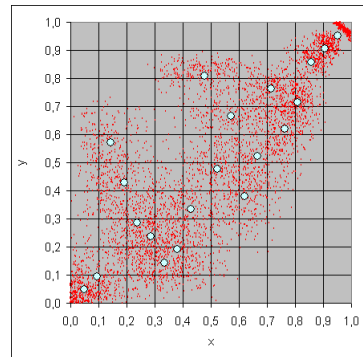
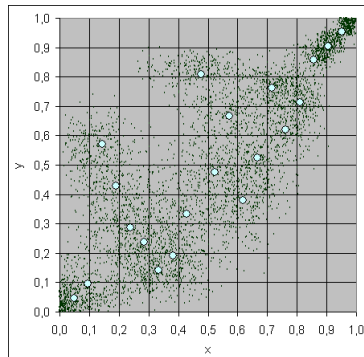
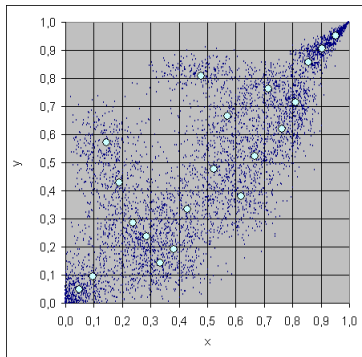
lower Fréchet shuffle

binomial copula,  $a = 22$ ,  $b = 27$



### 3. Case studies

5,000 simulated pairs of the data-driven copulas and empirical copula  
(large points):



upper Fréchet shuffle

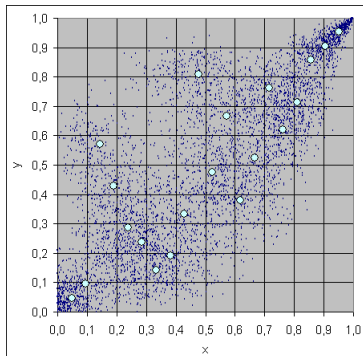
rook copula

lower Fréchet shuffle

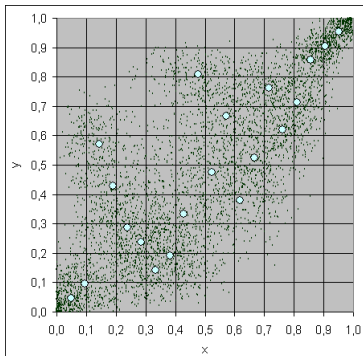
negative binomial copula,  $a = 17$ ,  $b = 22$

### 3. Case studies

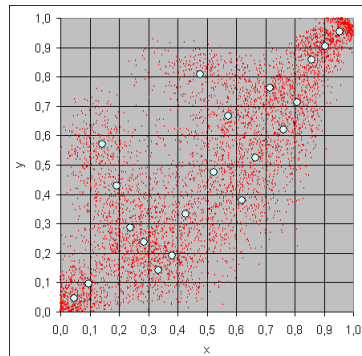
5,000 simulated pairs of the data-driven copulas and empirical copula  
(large points):



upper Fréchet shuffle



rook copula

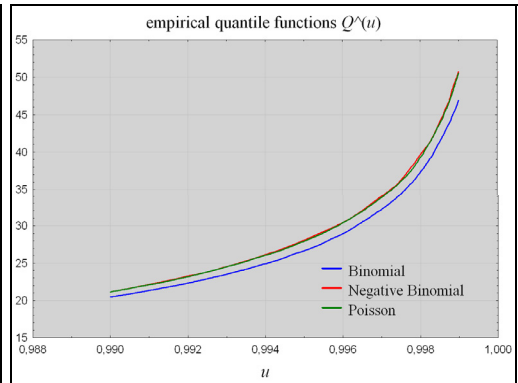
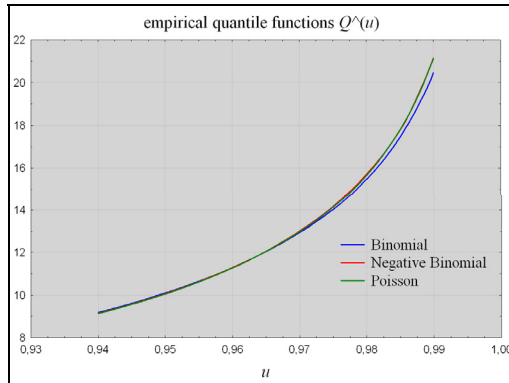


lower Fréchet shuffle

Poisson copula,  $a = 17$ ,  $b = 22$

### 3. Case studies

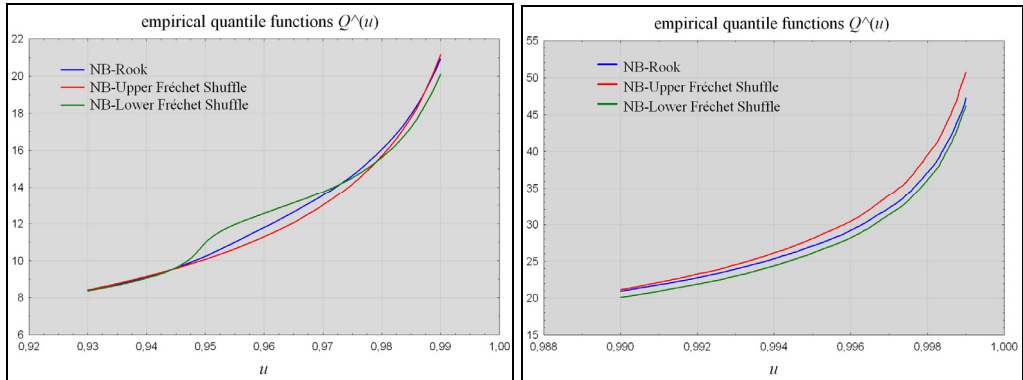
$Q^{\wedge}(u)$  based on the largest 100,000 observations from a total of  $10^6$  simulations:



empirical quantile functions  $Q^{\wedge}(u)$ , upper Fréchet shuffle

### 3. Case studies

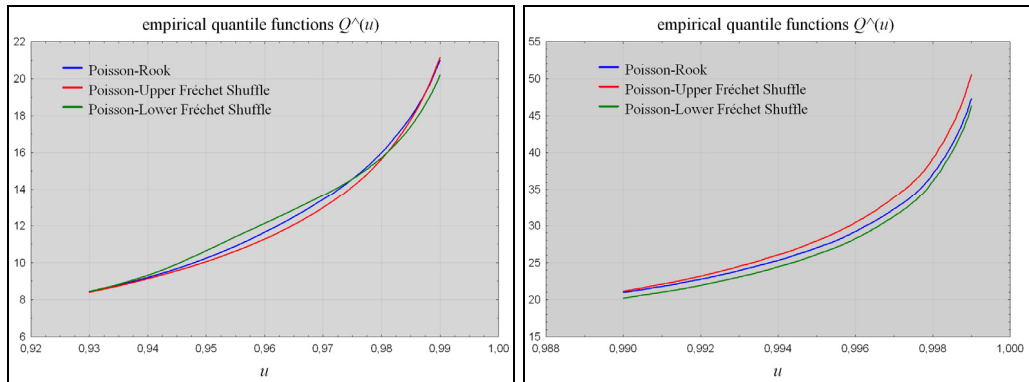
$Q^{\wedge}(u)$  based on the largest 100,000 observations from a total of  $10^6$  simulations:



empirical quantile functions  $Q^{\wedge}(u)$ , negative binomial copula

### 3. Case studies

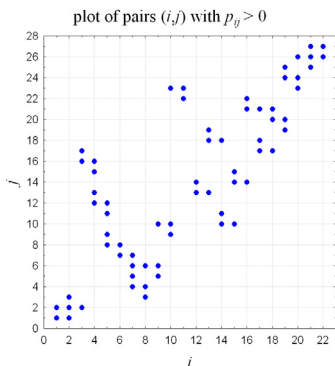
$Q^{\wedge}(u)$  based on the largest 100,000 observations from a total of  $10^6$  simulations:



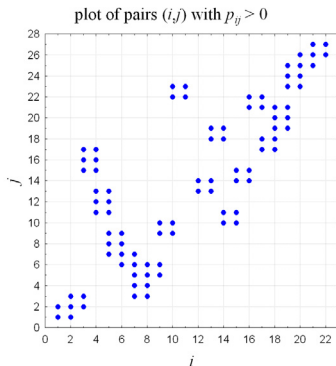
empirical quantile functions  $Q^{\wedge}(u)$ , Poisson copula

### 3. Case studies

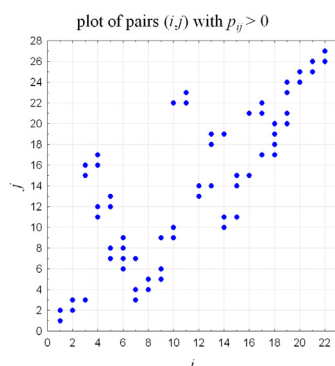
The position of the pairs  $(i, j)$  for which the  $p_{ij}$  are positive follows the graph of rank vectors (empirical copula) very closely:



lower Fréchet shuffle



rook copula



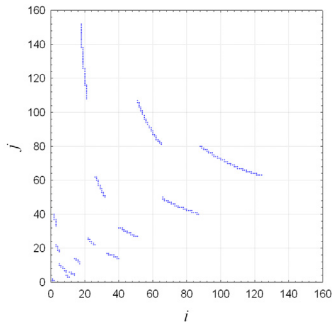
upper Fréchet shuffle

binomial copula,  $a = 22$ ,  $b = 27$

### 3. Case studies

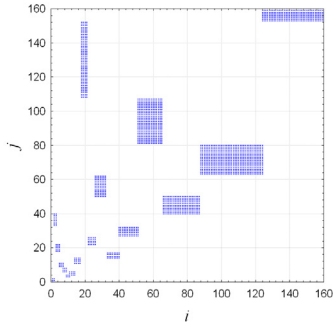
The position of the pairs  $(i, j)$  for which the  $p_{ij}$  are positive follows the graph of rank vectors (empirical copula) very closely:

plot of pairs  $(i, j)$  with  $p_{ij} > 0$



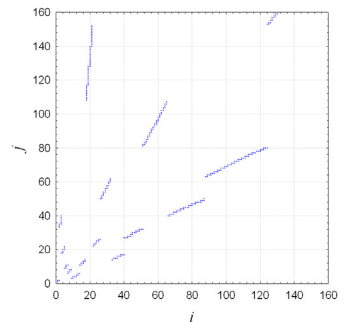
lower Fréchet shuffle

plot of pairs  $(i, j)$  with  $p_{ij} > 0$



rook copula

plot of pairs  $(i, j)$  with  $p_{ij} > 0$



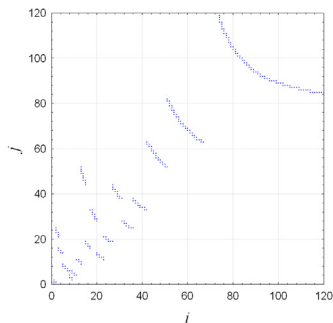
upper Fréchet shuffle

negative binomial copula,  $a = 17$ ,  $b = 22$

### 3. Case studies

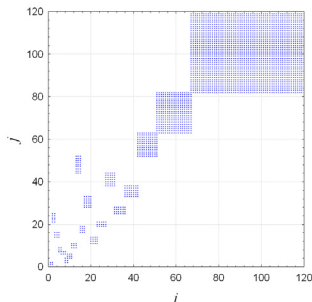
The position of the pairs  $(i, j)$  for which the  $p_{ij}$  are positive follows the graph of rank vectors (empirical copula) very closely:

plot of pairs  $(i, j)$  with  $p_{ij} > 0$



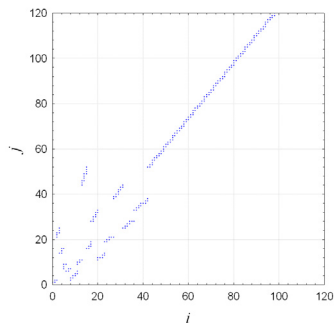
lower Fréchet shuffle

plot of pairs  $(i, j)$  with  $p_{ij} > 0$



rook copula

plot of pairs  $(i, j)$  with  $p_{ij} > 0$



upper Fréchet shuffle

Poisson copula,  $a = 17$ ,  $b = 22$





## 4. Extension to arbitrary dimensions

Assumptions:

- $\{\varphi_{ki}(\mathbf{u})\}_{i \in \mathbb{Z}^+}$  for  $k = 1, \dots, d$  discrete probabilities with

$$\sum_{i=0}^{\infty} \varphi_{ki}(\mathbf{u}) = 1 \text{ for } \mathbf{u} \in (0, 1) \quad (20)$$

$$\int_0^1 \varphi_{ki}(\mathbf{u}) d\mathbf{u} = \alpha_{ki} > 0 \text{ for } i \in \mathbb{Z}^+. \quad (21)$$

- $\{\mathbf{p}_i\}_{i \in \mathbb{Z}^{+d}}$  is a distribution of an arbitrary discrete  $d$ -dimensional random vector  $\mathbf{Z}$  over  $\mathbb{Z}^{+d}$  where, with  $\mathbf{i} = (i_1, \dots, i_d)$ ,

$$P(\mathbf{Z} = \mathbf{i}) = \mathbf{p}_i, \mathbf{i} \in \mathbb{Z}^{+d}. \quad (22)$$

- marginal distributions with

$$P(Z_k = i) = \alpha_{ki}, i \in \mathbb{Z}^+, k = 1, \dots, d. \quad (23)$$

## 4. Extension to arbitrary dimensions

Then

$$c(\mathbf{u}) := \sum_{\mathbf{i} \in \mathbb{Z}^{+d}} \frac{p_{\mathbf{i}}}{\prod_{k=1}^d \alpha_{k,i_k}} \prod_{k=1}^d \varphi_{k,i_k}(u_k), \quad \mathbf{u} = (u_1, \dots, u_d) \in (0,1)^d \quad (24)$$

defines the density of a  $d$ -variate copula, which is again called *generalized partition-of-unity copula*. Alternatively, we can rewrite (24) again as

$$c(\mathbf{u}) = \sum_{\mathbf{i} \in \mathbb{Z}^{+d}} p_{\mathbf{i}} \prod_{k=1}^d f_{k,i_k}(u_k), \quad \mathbf{u} = (u_1, \dots, u_d) \in (0,1)^d \quad (25)$$

where the  $f_{ki}(\cdot) = \frac{\varphi_{ki}(\cdot)}{\alpha_{ki}}$ ,  $i \in \mathbb{Z}^+$ ,  $k = 1, \dots, d$  denote the Lebesgue densities induced by the  $\{\varphi_{ki}(\mathbf{u})\}_{i \in \mathbb{Z}^+}$ .

## 5. Bibliography / References

C. Cottin and D. Pfeifer (2014): *From Bernstein polynomials to Bernstein copulas*. J. Appl. Funct. Anal, 9(3-4):277-288.

F. Durante, J. Fernández-Sánchez, and C. Sempi (2013): *Multivariate patchwork copulas: A unified approach with applications to partial comonotonicity*. Insurance: Mathematics and Economics, 53(3):897 -905.

J. Maciag, F. Hesse, R. Boeve, and A. Pfingsten (2014): *Comparing risk measures when aggregating market risk and credit risk using different copulas*. J. Risk, 18(5):101-136.

R. B. Nelsen (2007). *An Introduction to Copulas*. 2<sup>nd</sup> Ed., Springer, N.Y.

D. Pfeifer, H.A. Tsatedem, A. Mändle, and C. Girschig (2016): *New copulas based on general partitions-of-unity and their applications to risk management*. Depend. Model. 4, 123–140.