mathematics

MDPI

Article

Adaptive Bernstein Copulas and Risk Management

Dietmar Pfeifer [1,*] and Olena Ragulina [2]

[1] Institute of Mathematics, Carl von Ossietzky Universität Oldenburg, D-26111 Oldenburg, Germany
[2] Department of Probability Theory, Statistics and Actuarial Mathematics, Taras Shevchenko National University of Kyiv, Volodymyrska Str. 64, 01601 Kyiv, Ukraine; ragulina.olena@gmail.com
* Correspondence: dietmar.pfeifer@uni-oldenburg.de; Tel.: +49-(0)441-798-3214

Received: 29 October 2020; Accepted: 11 December 2020; Published: 14 December 2020

Abstract: We present a constructive approach to Bernstein copulas with an admissible discrete skeleton in arbitrary dimensions when the underlying marginal grid sizes are smaller than the number of observations. This prevents an overfitting of the estimated dependence model and reduces the simulation effort for Bernstein copulas a lot. In a case study, we compare different approaches of Bernstein and Gaussian copulas regarding the estimation of risk measures in risk management.

Keywords: copulas; partition-of-unity copulas; Monte Carlo methods

MSC: 62H05; 62H12; 62H17; 11K45

1. Introduction

Since the pioneering paper [1] by Serge Bernstein in 1912, Bernstein polynomials have been an indispensable tool in calculus and approximation theory (see, e.g., [2]). Bernstein copulas, which can be considered to be Bernstein polynomials for empirical and other copula functions, have a long tradition in nonparametric modelling of dependence structures in arbitrary dimensions, in particular, with applications in risk management, and have come into a deeper focus in the recent years. There is an extensive list of research papers on this topic, for instance, [3–14]. In particular, the monographs [15,16] have devoted separate chapters to the topic of Bernstein copulas.

A very important aspect of Bernstein copulas lies in Monte Carlo simulation techniques of dependence structures, in particular, in higher dimensions. The structure of such procedures ranges from very complex (see, e.g., [10]) to extremely simple (see, e.g., [5]), such that Monte Carlo simulations could, for instance, be performed easily with ordinary spreadsheets, in particular, in applications concerning quantitative risk management.

From a statistical point of view, the problem of a potential overfitting of the true underlying dependence structure with Bernstein polynomials emerges naturally when the pertaining Bernstein polynomial degree increases, i.e., with an increasing number of observations. This has been discussed extensively, for instance, in [6] (Section 3.1), [7] (Section 4), [17] (Section 3.2.1) or [14] (Remark 4). This leads to the fact that the Bernstein copula density becomes more wiggly the more empirical observations are used in the analysis. In comparison with classical parametric dependence models such as elliptically contoured or Archimedean copulas, this is probably a nondesirable property. In particular, this problem has been tackled seemingly first in [10] by approximating the underlying discrete skeleton for the Bernstein copula by a least-squares approach and recently in the Ph.D. thesis [18], where cluster analytic methods were used.

In the present paper, we propose a simple but yet effective approach to reduce the complexity of Bernstein copulas in a two-step approach, namely first an augmentation step in combination with a second reduction step. The reduction step is also discussed in [12]; however, without a possible application to a general complexity reduction of Bernstein copulas.

2. Some Important Facts about Multivariate Bernstein Polynomials

Let f be an arbitrary bounded real-valued function on the unit cube $\mathbf{C}_d := [0,1]^d$ with dimension $d \in \mathbb{N}$. Furthermore, let $n_1, \ldots, n_d$ be integers. The corresponding multivariate Bernstein polynomial is defined by

$$B_{\mathbf{n}} f(x_1, \ldots, x_d) := \sum_{i_d=0}^{n_d} \cdots \sum_{i_1=0}^{n_1} f\left(\frac{i_1}{n_1}, \cdots, \frac{i_d}{n_d}\right) \prod_{j=1}^{d} \binom{n_j}{i_j} x_j^{i_j} (1-x_j)^{n_j - i_j}, \quad \mathbf{x} = (x_1, \ldots, x_d) \in \mathbf{C}_d, \tag{1}$$

with $\mathbf{n} = (n_1, \ldots, n_d)$ (see, e.g., [2] (p. 51)). It is known that for $\min(n_1, \ldots, n_d) \to \infty$ multivariate Bernstein polynomials converge to f at any point of continuity and approximate f uniformly if f is continuous on $\mathbf{C}_d$.

Another important property of multivariate Bernstein polynomials, which is perhaps less known in the mathematical community, is the fact that the multivariate Bernstein polynomial density given by

$$b_{\mathbf{n}} f(x_1, \ldots, x_d) := \frac{\partial^d}{\partial x_1 \cdots \partial x_d} B_{\mathbf{n}} f(x_1, \ldots, x_d), \quad \mathbf{x} \in \mathbf{C}_d, \tag{2}$$

can be written as a linear combination of product beta densities. For this purpose, consider univariate beta densities

$$f_{\text{beta}}(x; \alpha, \beta) := \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)} \quad \text{for} \quad 0 < x < 1, \;\; \alpha, \beta > 0, \tag{3}$$

where $B(\alpha, \beta)$ denotes the Euler beta function, i.e., $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$. We need the following definition to proceed.

Definition 1. *Let g be a real-valued bounded function on $\mathbb{R}^d$. We call*

$$\Delta_{\mathbf{a}}^{\mathbf{b}} g := \sum_{(\varepsilon_1, \ldots, \varepsilon_d) \in \{0,1\}^d} (-1)^{\sum_{i=1}^d \varepsilon_i} g\big(\varepsilon_1 a_1 + (1-\varepsilon_1) b_1, \ldots, \varepsilon_d a_d + (1-\varepsilon_d) b_d\big) \geqslant 0 \tag{4}$$

the Δ-difference of g over the interval $(\mathbf{a}, \mathbf{b}] := \left(\overset{d}{\underset{i=1}{\times}} (a_i, b_i]\right)$ with $\mathbf{a} = (a_1, \ldots, a_d) \in \mathbb{R}^d$, $\mathbf{b} = (b_1, \ldots, b_d) \in \mathbb{R}^d$ and $a_i < b_i$, $1 \leqslant i \leqslant d$. (We adopt here a notation similar to that in [19] (Definition 2.1), which is slightly different from the notation in [20] (Definition 1.2.10).)

Proposition 1. *With the above notation, the Bernstein polynomial density $b_{\mathbf{n}} f$ can be represented as*

$$b_{\mathbf{n}} f(x_1, \ldots, x_d) = \sum_{i_d=0}^{n_d-1} \cdots \sum_{i_1=0}^{n_1-1} \Delta_{\mathbf{a}_i}^{\mathbf{b}_i} f \prod_{j=1}^{d} f_{\text{beta}}(x_j; i_j + 1, n_j - i_j) \tag{5}$$

with $\mathbf{a}_i := \left(\frac{i_1}{n_1}, \cdots, \frac{i_d}{n_d}\right)$ and $\mathbf{b}_i := \left(\frac{i_1+1}{n_1}, \cdots, \frac{i_d+1}{n_d}\right)$.

Proof. This follows immediately from the arguments in the proof of Theorem 2.2. in [5]; compare also the line of proofs in [21]. □

Example 1. *We consider the polynomial $f(x,y) := 2x(1-y)^3 - 3(1-x)^3 y^4$, $0 \leqslant x, y \leqslant 1$, with $n_1 = 2$, $n_2 = 3$. In this case, the two-dimensional Bernstein polynomial $B_{\mathbf{n}} f$ differs from f due to smaller polynomial degrees. We have*

$$B_{\mathbf{n}} f(x,y) = 2x - \frac{y}{9} - \frac{145}{36}xy - \frac{14}{9}y^2 - \frac{4}{3}y^3 + \frac{97}{18}xy^2 - \frac{x^2 y}{12} + \frac{17}{9}xy^3 - \frac{7}{6}x^2y^2 - x^2y^3 \tag{6}$$

with

$$b_{\mathbf{n}}f(x,y) = -\frac{145}{36} - \frac{x}{6} + \frac{97}{9}y + \frac{17}{3}y^2 - \frac{14}{3}xy - 6xy^2. \tag{7}$$

Please note that here

$$\Delta f_{\mathbf{a}_i}^{\mathbf{b}_i} = f\left(\frac{i_1+1}{n_1}, \frac{i_2+1}{n_2}\right) + f\left(\frac{i_1}{n_1}, \frac{i_2}{n_2}\right) - f\left(\frac{i_1}{n_1}, \frac{i_2+1}{n_2}\right) - f\left(\frac{i_1+1}{n_1}, \frac{i_2}{n_2}\right), \tag{8}$$

or, in tabular form, the values are given in Table 1.

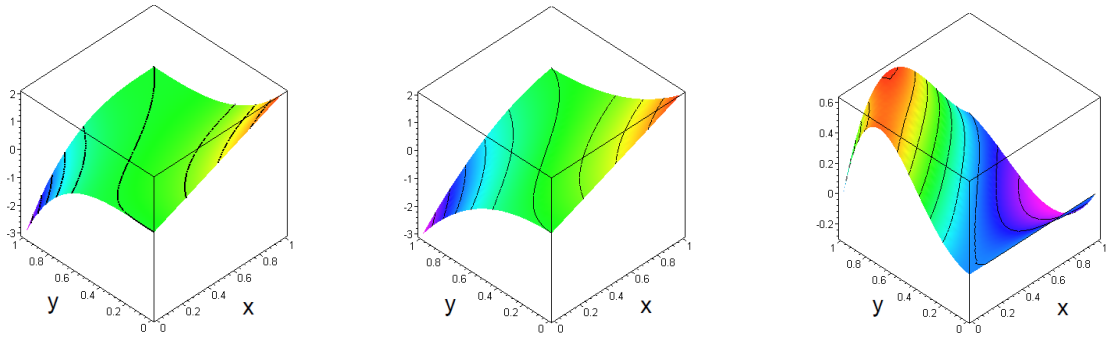
Table 1. Values of $\Delta f_{\mathbf{a}_i}^{\mathbf{b}_i}$ for the polynomial $f(x,y)$ in Example 1.

i_1	0	1	0	1	0	1
i_2	0	0	1	1	2	2
$\Delta f_{\mathbf{a}_i}^{\mathbf{b}_i}$	$-\frac{145}{216}$	$-\frac{151}{216}$	$\frac{49}{216}$	$-\frac{41}{216}$	$\frac{149}{72}$	$\frac{19}{72}$

After a little computation it is easy to see that indeed here

$$b_{\mathbf{n}}f(x,y) = \sum_{i_2=0}^{2}\sum_{i_1=0}^{1} \Delta f_{\mathbf{a}_i}^{\mathbf{b}_i} \frac{x^{i_1}(1-x)^{1-i_1}}{B(i_1+1, 2-i_1)} \frac{y^{i_2}(1-y)^{2-i_2}}{B(i_2+1, 3-i_2)}. \tag{9}$$

The plots of $f(x,y)$, $B_{\mathbf{n}}f(x,y)$ and $f(x,y) - B_{\mathbf{n}}f(x,y)$ are shown in Figure 1.



(**a**) (**b**) (**c**)

Figure 1. Plots of the functions in Example 1. (**a**) Plot of $f(x,y)$. (**b**) Plot of $B_{\mathbf{n}}f(x,y)$. (**c**) Plot of $f(x,y) - B_{\mathbf{n}}f(x,y)$.

A direct consequence of Proposition 1 concerns the monotonicity behavior of multivariate Bernstein polynomials.

Definition 2. *Let g be a real-valued function on $\mathbb{R}^d$. We call g d-monotone iff $\Delta g_{\mathbf{a}}^{\mathbf{b}} \geqslant 0$ for all $\mathbf{a} = (a_1, \ldots, a_d)$, $\mathbf{b} = (b_1, \ldots, b_d) \in \mathbb{R}^d$ with $a_i < b_i$, $1 \leqslant i \leqslant d$.*

It is obvious by the iterated mean value theorem that for a sufficiently smooth function g, d-monotonicity is equivalent to

$$\frac{\partial^d}{\partial x_1 \cdots \partial x_d} g(x_1, \ldots, x_d) \geqslant 0 \quad \text{for all} \quad (x_1, \ldots, x_d) \in \mathbb{R}^d. \tag{10}$$

Please note that in the case that g is a d-dimensional cumulative distribution function of a probability measure $\mathbb{P}$ on the d-dimensional Borel σ-field $\mathfrak{B}^d$, then $\Delta g_{\mathbf{a}}^{\mathbf{b}} = \mathbb{P}\left(\bigtimes_{i=1}^{d} (a_i, b_i]\right)$.

Proposition 2. *Let f be a real-valued d-monotone function on $\mathbb{R}^d$. Then the corresponding multivariate Bernstein polynomial $B_{\mathbf{n}}f$ is also d-monotone. In particular, the Bernstein polynomial density $b_{\mathbf{n}}f$ is a positive linear combination of product beta densities.*

Proof. By the arguments above and the notation as in Proposition 1, we have

$$b_{\mathbf{n}}f(x_1,\ldots,x_d) := \frac{\partial^d}{\partial x_1\cdots\partial x_d}B_{\mathbf{n}}f(x_1,\ldots,x_d) = \sum_{i_d=0}^{n_d-1}\cdots\sum_{i_1=0}^{n_1-1}\Delta f_{\mathbf{a}_i}^{\mathbf{b}_i}\prod_{j=1}^{d} f_{\text{beta}}(x_j;i_j+1,n_j-i_j) \geqslant 0, \quad (11)$$

which is a sufficient condition for $B_{\mathbf{n}}f$ to be d-monotone. □

Please note that the polynomial f from Example 1 is not 2-monotone since $\Delta f_{\mathbf{a}}^{\mathbf{b}} = -0.00126\ldots < 0$ with $\mathbf{a} = (0.2, 0.4)$ and $\mathbf{b} = (0.27, 0.45)$. However, the slightly modified polynomial $g(x,y) = f(x,y) + 6xy$ is 2-monotone since $\frac{\partial^2}{\partial x\partial y}g(x,y) = 6 - 6(1-y)^2 + 36(1-x)^2y^3 \geqslant 0$ with the unique global minimum point $(x,y) = (1,1)$ and $\frac{\partial^2}{\partial x\partial y}g(1,1) = 0$. With respect to the corresponding multivariate Bernstein polynomial, we now obtain the values given in Table 2, which also explicitly shows that the Bernstein polynomial $B_{\mathbf{n}}g$ is 2-monotone.

Table 2. Values of $\Delta g_{\mathbf{a}_i}^{\mathbf{b}_i}$ for the modified polynomial $g(x,y)$.

i_1	0	1	0	1	0	1
i_2	0	0	1	1	2	2
$\Delta g_{\mathbf{a}_i}^{\mathbf{b}_i}$	$\frac{71}{216}$	$\frac{65}{216}$	$\frac{265}{216}$	$\frac{175}{216}$	$\frac{221}{72}$	$\frac{91}{72}$

3. From Bernstein Polynomials to Bernstein Copulas

Remark 1. *Seemingly Proposition 2 can be usefully applied to arbitrary d-dimensional cumulative distribution functions F concentrated on the unit cube $\mathbf{C}_d := [0,1]^d$ (continuous or not) such that the corresponding multivariate Bernstein polynomial*

$$B_{\mathbf{n}}F(x_1,\ldots,x_d) = \sum_{i_d=0}^{n_d}\cdots\sum_{i_1=0}^{n_1} F\left(\frac{i_1}{n_1},\cdots,\frac{i_d}{n_d}\right)\prod_{j=1}^{d}\binom{n_j}{i_j}x_j^{i_j}(1-x_j)^{n_j-i_j}, \quad \mathbf{x} = (x_1,\ldots,x_d) \in \mathbf{C}_d, \quad (12)$$

is also a cumulative distribution function since $B_{\mathbf{n}}F$ is non-negative and d-increasing with $B_{\mathbf{n}}F(0,\ldots,0) = F(0,\ldots,0)$ and $B_{\mathbf{n}}F(1,\ldots,1) = F(1,\ldots,1) = 1$. In particular, the Bernstein polynomial density $b_{\mathbf{n}}F$ is always a (probabilistic) mixture of product beta densities, as explicitly noted in [5,12] for Bernstein copulas. Note also that this observation was the motivation for the setup in [22].

Example 2. *We consider a two-dimensional random vector $\mathbf{X} = (X,Y)$ with a discrete distribution concentrated on $\{\mathbf{x},\mathbf{y}\}$ with $\mathbf{x} = (x_1,x_2) = (0.2,0.7)$ and $\mathbf{y} = (y_1,y_2) = (0.3,0.5)$ given by Table 3.*

Table 3. Distribution of the vector $\mathbf{X} = (X,Y)$ in Example 2.

$\mathbb{P}(\boldsymbol{X = x_i, Y = y_j})$	$\boldsymbol{x_1 = 0.2}$	$\boldsymbol{x_2 = 0.7}$
$y_1 = 0.3$	0.3	0.2
$y_2 = 0.5$	0.2	0.3

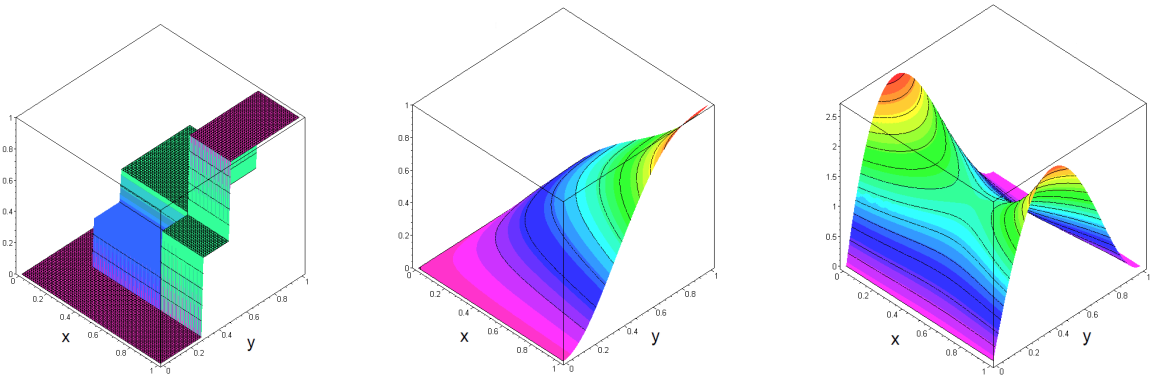
Let H denote the Heaviside unit step function, i.e.,

$$H(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geqslant 0. \end{cases}$$

Then the corresponding cumulative distribution function F for the given discrete distribution is

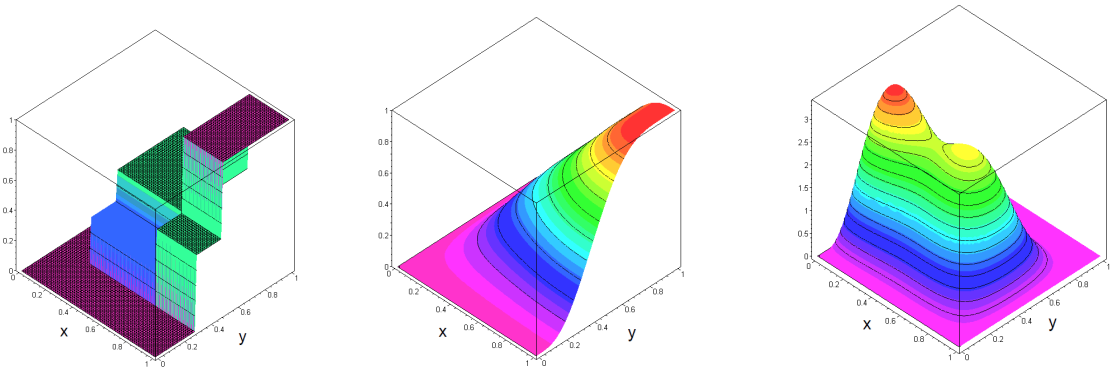
$$F(x,y) = \sum_{j=1}^{2}\sum_{i=1}^{2} \mathbb{P}(X = x_i, Y = y_j)H(x - x_i)H(y - y_j), \quad x, y \in \mathbb{R}. \tag{13}$$

The graphs in Figures 2–4 show the corresponding cumulative distribution function F as well as the corresponding Bernstein polynomials $B_{\mathbf{n}}F$ and densities $b_{\mathbf{n}}F$ for various choices of $\mathbf{n}$ according to relations (11) and (12) above.



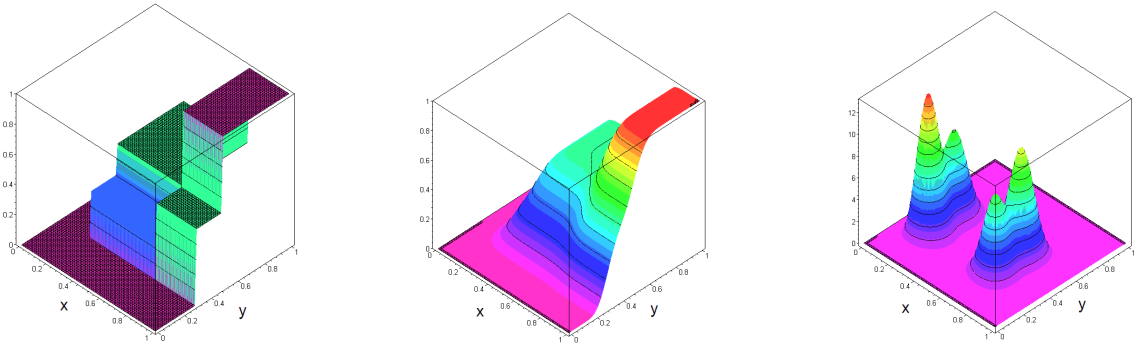
(**a**) (**b**) (**c**)

Figure 2. Plots of the functions in Example 2, $n_1 = 3$ and $n_2 = 5$. (**a**) Plot of $F(x,y)$. (**b**) Plot of $B_{\mathbf{n}}F(x,y)$. (**c**) Plot of $b_{\mathbf{n}}F(x,y)$.



(**a**) (**b**) (**c**)

Figure 3. Plots of the functions in Example 2, $n_1 = 11$ and $n_2 = 7$. (**a**) Plot of $F(x,y)$. (**b**) Plot of $B_{\mathbf{n}}F(x,y)$. (**c**) Plot of $b_{\mathbf{n}}F(x,y)$.



(**a**) (**b**) (**c**)

Figure 4. Plots of the functions in Example 2, $n_1 = 50$ and $n_2 = 50$. (**a**) Plot of $F(x,y)$. (**b**) Plot of $B_{\mathbf{n}}F(x,y)$. (**c**) Plot of $b_{\mathbf{n}}F(x,y)$.

The graphs in Figures 2–4 clearly visualize the approximation effect of multivariate Bernstein polynomials for discrete multivariate distributions if $\min(n_1, \ldots, n_d)$ gets large. In particular, the Bernstein polynomial density has spikes around the support points of the underlying discrete distribution.

To simplify notation, we will use the following convention. Let $d > 1$ be a natural number and $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$ be arbitrary. Then, for $y \in \mathbb{R}$, let

$$\mathbf{x}_{\to k}(y) := \begin{cases} (y, x_2, \ldots, x_d) & \text{if} \quad k = 1, \\ (x_1, \ldots, x_{k-1}, y, x_{k+1}, \ldots, x_d) & \text{if} \quad 1 < k < d, \\ (x_1, \ldots, x_{d-1}, y) & \text{if} \quad k = d, \end{cases} \tag{14}$$

denote the vector $\mathbf{x}$, where the k-th component is replaced by y.

Proposition 3. *Suppose that for $d > 1$, there is a cumulative distribution function $F : [0,1]^d \to [0,1]$ with $F(0, \ldots, 0) = 0$ and $F(1, \ldots, 1) = 1$ such that for given natural numbers $n_1, \ldots, n_d > 1$ we have $F\left(\mathbf{1}_{\to k}\left(\frac{i}{n_j}\right)\right) = \frac{i}{n_j}$ for $i \in \{0, 1, \ldots, n_j\}$, $j = 1, \ldots, d$, $k = 1, \ldots, d$, where $\mathbf{1} = (1, \ldots, 1) \in \mathbb{R}^d$. Then the d-dimensional Bernstein polynomial $B_{\mathbf{n}}F$ with $\mathbf{n} = (n_1, \ldots, n_d)$ associated with F is a copula.*

Proof. By Remark 1, we know that $B_{\mathbf{n}}F$ is also a cumulative distribution function with $B_{\mathbf{n}}F(0, \ldots, 0) = F(0, \ldots, 0) = 0$ and $B_{\mathbf{n}}F(1, \ldots, 1) = F(1, \ldots, 1) = 1$, and (note that $0^0 = 1$)

$$\begin{aligned} B_{\mathbf{n}}F(\mathbf{1}_{\to k}x) &= \sum_{i_d=0}^{n_d} \cdots \sum_{i_1=0}^{n_1} F\left(\frac{i_1}{n_1}, \cdots, \frac{i_d}{n_d}\right) \prod_{j=1}^{d} \binom{n_j}{i_j} x_j^{i_j} (1 - x_j)^{n_j - i_j} \\ &= \sum_{i=0}^{n_k} \binom{n_k}{i} F\left(\mathbf{1}_{\to k}\left(\frac{i}{n_k}\right)\right) x^i (1-x)^{n_k - i} = \frac{1}{n_k} \sum_{i=0}^{n_k} i \binom{n_k}{i} x^i (1-x)^{n_k - i} = \frac{n_k x}{n_k} = x \end{aligned} \tag{15}$$

for $k = 1, \ldots, d$ and $0 \leqslant x \leqslant 1$ ($n_k x$ is the expectation of the Binomial distribution with n_k trials and success probability x). Hence the marginal distributions induced by B are continuous uniform, which means that B is indeed a copula. □

Please note that Proposition 3 is already implicitly formulated in [5,10] (see also [15] (Chapter 4.1.2)). We reformulate the corresponding statements there in an appropriate way.

Corollary 1. *Let $\mathbf{U} = (U_1, \ldots, U_d)$ be a discrete random vector whose marginal component U_i follows a discrete uniform distribution over $T_i := \{0, 1, \ldots, n_i - 1\}$ with integers $n_i > 1$, $i = 1, \ldots, d$. Then the multivariate Bernstein polynomial $B_{\mathbf{n}}F$ derived from the cumulative distribution function F for the scaled random vector $\mathbf{V} = \left(\frac{U_1 + 1}{n_1}, \ldots, \frac{U_d + 1}{n_d}\right)$ given by $F(x_1, \ldots, x_d) = \mathbb{P}\big(V_1 \leqslant x_1, \ldots, V_d \leqslant x_d\big)$, $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbf{C}_d$, is a copula. The corresponding Bernstein copula density $b_{\mathbf{n}}F$ is given by*

$$b_{\mathbf{n}}F(x_1, \ldots, x_d) = \sum_{i_d=0}^{n_d - 1} \cdots \sum_{i_1=0}^{n_1 - 1} \mathbb{P}\big(\mathbf{U} = (i_1, \ldots, i_d)\big) \prod_{j=1}^{d} f_{\text{beta}}(x_j; i_j + 1, n_j - i_j), \quad (x_1, \ldots, x_d) \in \mathbf{C}_d. \tag{16}$$

Proof. For $i_j \in T_j$, $j = 1, \ldots, d$, we have

$$F\left(\frac{i_1}{n_1}, \ldots, \frac{i_d}{n_d}\right) = \mathbb{P}\left(V_1 \leqslant \frac{i_1}{n_1}, \ldots, V_d \leqslant \frac{i_d}{n_d}\right) = \mathbb{P}\big(U_1 < i_1, \ldots, U_d < i_d\big),$$

and hence

$$\Delta F_{\mathbf{a}_i}^{\mathbf{b}_i} = \mathbb{P}\left(\frac{i_1}{n_1} < V_1 \leqslant \frac{i_1+1}{n_1}, \ldots, \frac{i_d}{n_d} < V_d \leqslant \frac{i_d+1}{n_d}\right) = \mathbb{P}\big(\mathbf{U} = (i_1, \ldots, i_d)\big).$$

□

Remark 2. *We call $B_{\mathbf{n}}F$ the Bernstein copula induced by* **U**. *In coincidence with [10] we also call* **U** *the discrete skeleton of the Bernstein copula $B_{\mathbf{n}}F$ and the number $n_1 \times \cdots \times n_d$ its grid size. If* **V** *is an arbitrary discrete random vector over $\mathbf{T} := \overset{d}{\underset{i=1}{\times}} T_i$, we call* **V** *an admissible discrete skeleton if the marginal distributions are discrete uniform. So every admissible skeleton over* **T** *induces a corresponding Bernstein copula via the multivariate Bernstein polynomial of its rescaled cumulative distribution function. The corresponding Bernstein copula density is a mixture of product beta kernels with weights given by the individual probabilities representing the admissible skeleton.*

4. Empirical Bernstein Copulas

Bernstein copulas can be easily constructed from independent samples $\mathbf{X}_1, \ldots, \mathbf{X}_n$, $n \in \mathbb{N}$, of d-dimensional random vectors with the same intrinsic dependence structure and the same marginal distributions. For simplicity, we assume here that the marginal distributions are continuous in order to avoid ties in the observations. The simplest way to construct an empirical Bernstein copula is on the basis of Deheuvel's empirical copula in the form of a cumulative distribution function (see [20]), which can be represented by an admissible discrete skeleton derived from the individual ranks r_{ij}, $i = 1, \ldots, d$, $j = 1, \ldots, n$, of the observation vectors $\mathbf{x}_j = (x_{1j}, \ldots, x_{dj})$, $j = 1, \ldots, n$, given by the order statistics $x_{i,r_{i1}} < x_{i,r_{i2}} < \ldots < x_{i,r_{in}}$, i.e., $r_{ij} = k$ iff x_{ij} is the k-largest value of the i-th observed component. Here the discrete skeleton **U** is given by a random vector over $\mathbf{T} := \{0, 1, \ldots, n-1\}^d$ with a discrete uniform distribution over the n support points $\mathbf{s}_1, \ldots, \mathbf{s}_n$, where $\mathbf{s}_j = (r_{1j} - 1, \ldots, r_{dj} - 1)$, $j = 1, \ldots, n$. Since the empirical copula converges in distribution to the true underlying copula as $n \to \infty$, it follows that the corresponding empirical Bernstein copula does so likewise (cf. [15] (Chapter 4.1.2)). This provides—in the light of relation (16)—in particular, a simple way to generate samples from an empirical Bernstein copula by Monte Carlo methods in two steps:

- Step 1: Select an index N randomly and uniformly among $1, \ldots, n$.
- Step 2: Generate d independent beta distributed random variables $V_1, \ldots, V_d$ (also independent of N), where V_i follows a beta distribution with parameters r_{iN} and $n + 1 - r_{iN}$, $i = 1, \ldots, d$.

Then $\mathbf{V} := (V_1, \ldots, V_d)$ is a sample point from the empirical Bernstein copula.

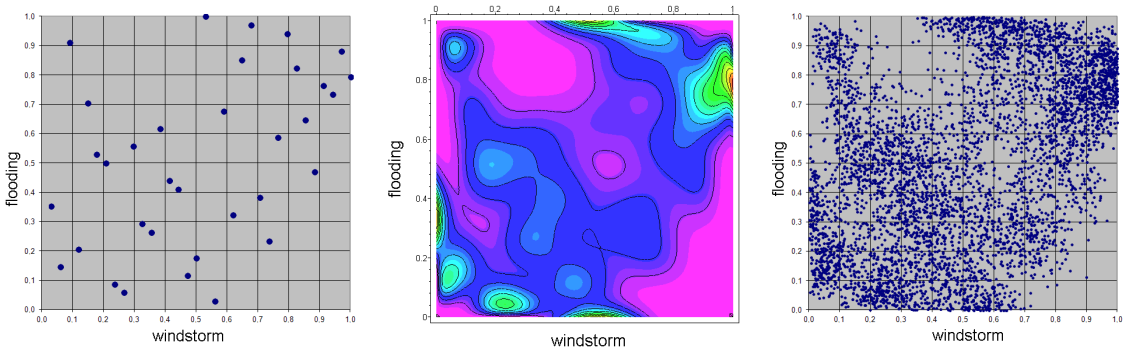
This was also observed in [12], but was known earlier (see, e.g., [5]). In what follows, we will discuss the data set presented in [10] (Section 3) in more detail.

Example 3. *Table 4 contains the ranks r_{ij} for observed insurance data from windstorm ($i = 1$) and flooding ($i = 2$) losses in central Europe for 34 consecutive years discussed in [10].*

Table 4. Ranks r_{ij} for observed insurance data from windstorm ($i = 1$) and flooding ($i = 2$) losses in Example 3.

$i\backslash j$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
2	12	5	31	7	24	18	17	3	2	19	10	9	21	15	14	4	6
$i\backslash j$	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**
1	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
2	34	1	23	11	29	33	13	8	20	32	28	22	16	26	25	30	27

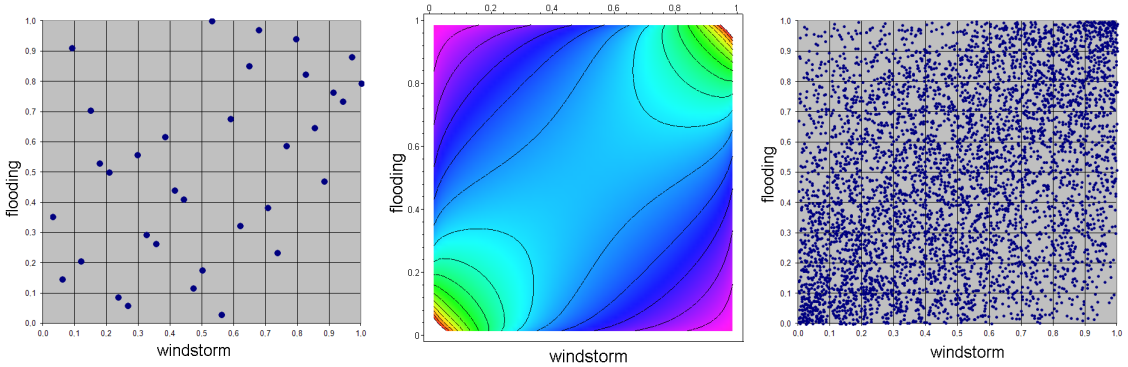
The graphs in Figure 5 show some plots for the empirical Bernstein copula.



(**a**) (**b**) (**c**)

Figure 5. Plots for the empirical Bernstein copula in Example 3. (**a**) Support points of scaled skeleton. (**b**) Empirical Bernstein copula density contour plot. (**c**) Simulation of 5000 empirical Bernstein copula pairs.

It is clearly to be seen from the shape of the contour lines in Figure 5b that the empirical Bernstein copula density is quite bumpy here, e.g., in comparison with the Gaussian copula density fitted to the data set above (see Figure 6).



(**a**) (**b**) (**c**)

Figure 6. Plots for the Gaussian copula in Example 3. (**a**) Support points of scaled skeleton. (**b**) Gaussian copula density contour plot. (**c**) Simulation of 5000 Gaussian copula pairs.

From a practical point of view, it might therefore be desirable to adapt the empirical Bernstein copula to a smaller support set $\mathbf{T}^* := \bigtimes_{i=1}^{d} T_i^* \subset \mathbf{T} := \bigtimes_{i=1}^{d} T_i$ for the underlying discrete skeleton. This is the central idea in [10]. However, the disadvantage of the method proposed in that paper is that the number of support points of $\mathbf{U}^*$ gets dramatically larger and is typically of exponential order with increasing grid sizes. This is due to the fact that the number of support points is usually in the range of $\#(\mathbf{T}^*) = \prod_{i=1}^{d} n_i$ because by the specific method of least squares used there, most support points of $\mathbf{T}^*$ will get a positive weight. Therefore, we propose a simpler way how to find a smaller discrete approximating skeleton $\mathbf{U}^*$ with an arbitrary given grid size in Section 5.

5. Adaptive Bernstein Copula Estimation

We start with the individual ranks r_{ij} of the observation vectors $\mathbf{x}_j = (x_{1j}, \ldots, x_{dj})$, $j = 1, \ldots, n$. Let $\mathbf{U}$ denote the canonical admissible discrete skeleton as described in Section 4, derived from the empirical copula. Our aim is to find a good approximating admissible discrete skeleton $\mathbf{U}_*$ with a given grid size $n_1 \times \cdots \times n_d$, where the n_i are typically smaller than n. We proceed in two steps:

- Step 1: Augmentation

  Select an integer M such that all n_i, $i = 1, \ldots, d$, are divisors of M, for instance, their least common multiple. We construct pseudo-ranks r_{ij}^+ in the following way:

$$r_{ij}^+ := r_{i,\lceil \frac{j}{M} \rceil} M + \left(\left\lceil \frac{j}{M} \right\rceil - 1 \right) M + 1 - j, \quad i = 1, \ldots, d, \quad j = 1, \ldots, Mn. \tag{17}$$

  Here $\lceil x \rceil := \min\{m \in \mathbb{Z} \,|\, x \leqslant m\}$, $x \in \mathbb{R}$, stands for "rounding up". Let $\mathbf{U}^+ = (U_1^+, \ldots, U_d^+)$ be the uniformly discretely distributed random vector over $\{0, 1, \ldots, Mn-1\}^d$ with support points $\mathbf{s}_1$, …, $\mathbf{s}_{Mn}$, where $\mathbf{s}_j = (r_{1j}^+ - 1, \ldots, r_{dj}^+ - 1)$, $j = 1, \ldots, nM$. Please note that the probability mass is $\frac{1}{Mn}$ for each support point, and that $\mathbf{U}^+$ is an admissible discrete skeleton.

- Step 2: Reduction

  Construct the final ranks r_{ij}^* in the following way:

$$r_{ij}^* := \left\lceil \frac{r_{ij}^+ n_i}{nM} \right\rceil, \quad i = 1, \ldots, d, \quad j = 1, \ldots, Mn. \tag{18}$$

  It follows from the above definition that there will be replicates in the final ranks and that r_{ij}^* takes values in the set $T_i^* = \{0, 1, \ldots, n_i - 1\}$. A point $\mathbf{s} = (s_1, \ldots, s_d)$ will be a support point of the final admissible skeleton $\mathbf{U}^*$ if there exist final ranks such that $\mathbf{s} = (r_{1,j_1}, \ldots, r_{d,j_d})$ for some $j_1, \ldots, j_d \in \{1, \ldots, Mn\}$. The probability mass attached to $\mathbf{s}$ is given by the number $\frac{K}{Mn}$, where K is the number of rank combinations $(r_{1,j_1}, \ldots, r_{d,j_d})$ that lead to the same $\mathbf{s}$. This also enables very simple Monte Carlo realizations of the corresponding Bernstein copula as described in Section 4 by first selecting an index N randomly and uniformly among $1, \ldots, Mn$ and then by proceeding as in Step 2 there with all of the r_{ij}^*.

Please note that the above augmentation step creates permutations of the set $\{1, \ldots, Mn\}$ in each component so that the pseudo-ranks r_{ij}^+ actually lead to an admissible discrete skeleton (cf. [5] (Section 4)). The mathematical correctness of the reduction step follows from the proof of Proposition 2.5 in [12].

In the augmentation step, M-wise partial permutations would not change the result but would create a more "random" augmentation of the original ranks.

Example 4. *Consider Table 5 with ranks r_{ij} and probabilities $p(r_1, r_2)$ for $n = 5$.*

Table 5. Ranks r_{ij} and probabilities $p(r_1, r_2)$ in Example 4.

$j \backslash i$	1	2	$p(r_1, r_2)$
1	1	3	0.2
2	2	4	0.2
3	3	1	0.2
4	4	2	0.2
5	5	5	0.2

We want to create approximate final ranks with $n_1 = 3$ and $n_2 = 4$. Both numbers are not a divisor of n, so we choose $M = 3 \cdot 4 = 12$. We show a part of the resulting pseudo-ranks r_{ij}^+ and probabilities $p(r_1^+, r_2^+)$ in Table 6.

Table 6. Resulting pseudo-ranks r_{ij}^{+} and probabilities $p(r_1^{+}, r_2^{+})$ in Example 4.

$j \backslash i$	1	2	$p(r_1^{+}, r_2^{+})$
1	12	36	$0.01\bar{6}$
2	11	35	$0.01\bar{6}$
3	10	34	$0.01\bar{6}$
⋮	⋮	⋮	⋮
13	24	48	$0.01\bar{6}$
14	23	47	$0.01\bar{6}$
15	22	46	$0.01\bar{6}$
⋮	⋮	⋮	⋮
25	36	12	$0.01\bar{6}$
26	35	11	$0.01\bar{6}$
27	34	10	$0.01\bar{6}$
⋮	⋮	⋮	⋮
58	51	51	$0.01\bar{6}$
59	50	50	$0.01\bar{6}$
60	49	49	$0.01\bar{6}$

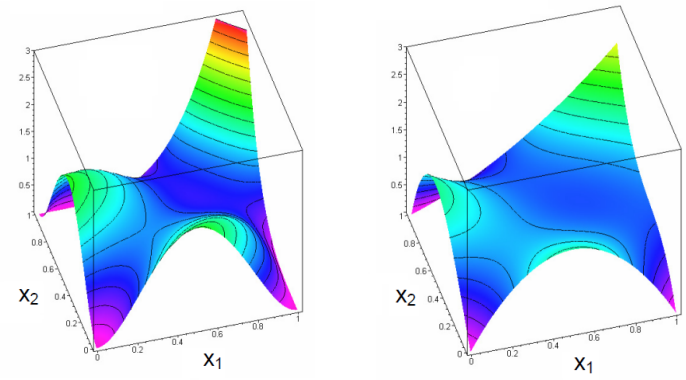
For the final ranks r_{ij}^{*} and probabilities $p(r_1^{*}, r_2^{*})$, we obtain Table 7.

Table 7. Final ranks r_{ij}^{*} and probabilities $p(r_1^{*}, r_2^{*})$ in Example 4.

$j \setminus i$	1	2	$p(r_1^{*}, r_2^{*})$
1	1	2	0.1
2	1	3	$0.2\bar{3}$
3	2	1	0.25
4	2	2	$0.01\bar{6}$
5	2	3	$0.01\bar{6}$
6	2	4	0.05
7	3	2	$0.1\bar{3}$
8	3	4	0.2

From $\sum_{j=1}^{4} p(i,j) = 0.\bar{3}$, $i = 1,2,3$, and $\sum_{i=1}^{3} p(i,j) = 0.25$, $j = 1,2,3,4$, we see that the induced skeleton is indeed admissible.

The graphs in Figure 7 show the corresponding copula densities $c_{\mathbf{U}}(x_1, x_2)$ and $c_{\mathbf{U}^*}(x_1, x_2)$ induced by $\mathbf{U}$ and $\mathbf{U}^*$. Seemingly the shape of both densities is similar, reflecting the structure of the original ranks quite well. However, the density $c_{\mathbf{U}^*}$ is less wiggly than the density $c_{\mathbf{U}}$, as intended.



(**a**) (**b**)

Figure 7. Plots of the copula densities in Example 4. (**a**) Copula density $c_{\mathbf{U}}(x_1, x_2)$. (**b**) Copula density $c_{\mathbf{U}^*}(x_1, x_2)$.

Note also that a reduction of complexity for copulas in the sense discussed here is also an essential topic in [23] (Section 3) (see, in particular, Figure 2 there). However, the underlying problem of a consistent reduction of complexity is not really discussed there.

6. Applications to Risk Management

In this section, we first want to investigate the data set of Example 3 in more detail. It is the basis data set in [10]. In particular, we want to discuss the effect of different adaptive Bernstein copula estimations on the estimation of risk measures such as Value at Risk, which is, for instance, the basis for Solvency II.

In [10], the number $n = 34$ of the original observations is first reduced to $n_1 = n_2 = 10$ by a least squares technique. The resulting optimal discrete skeleton with probabilities p_{ij}, $i \in T_1^*$, $j \in T_2^*$, is presented in Table 8 with $T_1^* = T_2^* = \{0, 1, \ldots, 9\}$. We have highlighted the non-zero entries in Tables 8 and 9 in order to illustrate the effect of a reduction in complexity by an appropriate application of adaptive Bernstein copulas.

Table 8. Probabilities p_{ij} in the resulting optimal discrete skeleton in Example 3.

$j \setminus i$	0	1	2	3	4	5	6	7	8	9
9	0.0032	0.0000	0.0022	0.0000	0.0032	0.0266	0.0320	0.0274	0.0028	0.0028
8	0.0318	0.0000	0.0014	0.0000	0.0024	0.0000	0.0312	0.0000	0.0020	0.0314
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0204	0.0251	0.0545
6	0.0032	0.0275	0.0022	0.0000	0.0032	0.0265	0.0026	0.0000	0.0322	0.0028
5	0.0003	0.0246	0.0287	0.0215	0.0003	0.0000	0.0000	0.0246	0.0000	0.0000
4	0.0034	0.0278	0.0024	0.0246	0.0034	0.0000	0.0029	0.0000	0.0324	0.0030
3	0.0266	0.0000	0.0000	0.0000	0.0266	0.0206	0.0261	0.0000	0.0000	0.0000
2	0.0034	0.0000	0.0025	0.0540	0.0034	0.0000	0.0029	0.0277	0.0031	0.0031
1	0.0252	0.0201	0.0000	0.0000	0.0546	0.0000	0.0000	0.0000	0.0000	0.0000
0	0.0029	0.0000	0.0607	0.0000	0.0029	0.0263	0.0023	0.0000	0.0025	0.0025

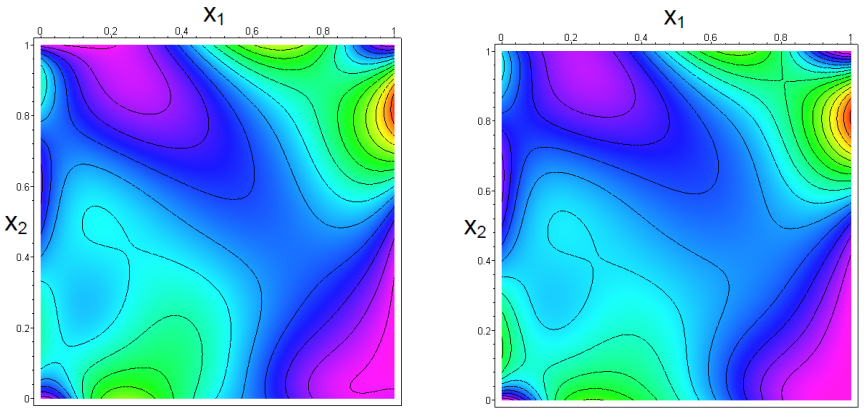
An application of the adaptive strategy described in Section 5 gives alternatively Table 9, which is less complex. Here we have chosen $M = 5$. Seemingly the number of support points for the adaptive probabilities p_{ij}^* are much less than before.

Table 9. Probabilities p_{ij}^* after application of the adaptive strategy in Example 3.

$j \setminus i$	0	1	2	3	4	5	6	7	8	9
9	0.0118	0.0000	0.0000	0.0000	0.0000	0.0294	0.0294	0.0294	0.0000	0.0000
8	0.0176	0.0000	0.0000	0.0000	0.0000	0.0000	0.0294	0.0000	0.0235	0.0294
7	0.0000	0.0059	0.0000	0.0000	0.0000	0.0000	0.0000	0.0059	0.0176	0.0706
6	0.0000	0.0235	0.0000	0.0176	0.0000	0.0294	0.0000	0.0000	0.0294	0.0000
5	0.0000	0.0294	0.0294	0.0118	0.0000	0.0000	0.0000	0.0294	0.0000	0.0000
4	0.0000	0.0235	0.0059	0.0176	0.0235	0.0000	0.0000	0.0000	0.0294	0.0000
3	0.0294	0.0000	0.0000	0.0000	0.0176	0.0059	0.0412	0.0059	0.0000	0.0000
2	0.0000	0.0059	0.0059	0.0529	0.0000	0.0059	0.0000	0.0294	0.0000	0.0000
1	0.0412	0.0118	0.0000	0.0000	0.0471	0.0000	0.0000	0.0000	0.0000	0.0000
0	0.0000	0.0000	0.0588	0.0000	0.0118	0.0294	0.0000	0.0000	0.0000	0.0000

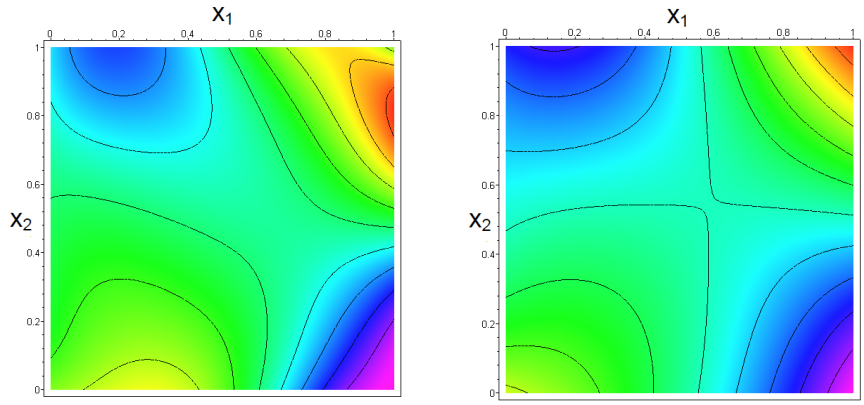
The graphs in Figure 8 show contour plots for the corresponding Bernstein copula densities. Here c_1 denotes the Bernstein copula density derived from Table 8, c_2 denotes the Bernstein copula density derived from Table 9. In the first case we have chosen $M = 5$, in the second case $M = 2$. Seemingly the differences are only marginal. However, in comparison with Figure 5b, the smoothing effect of the adaptive procedure is clearly visible.

The graphs in Figure 9 show contour plots for the adaptive Bernstein copula densities c_3 and c_4 with the choices $n_1 = n_2 = 5$ and $n_1 = n_2 = 4$, respectively.



(**a**) (**b**)

Figure 8. Contour plots for the Bernstein copula densities in Example 3. (**a**) Bernstein copula density $c_1(x_1, x_2)$. (**b**) Bernstein copula density $c_2(x_1, x_2)$.



(**a**) (**b**)

Figure 9. Contour plots for the adaptive Bernstein copula densities in Example 3. (**a**) Adaptive Bernstein copula density $c_3(x_1, x_2)$. (**b**) Adaptive Bernstein copula density $c_4(x_1, x_2)$.

In the next step, we compare estimates for the risk measure Value at Risk VaR_α with the risk level $\alpha = 0.5\%$—corresponding to a return period of 200 years—on the basis of a Monte Carlo study with 1,000,000 repetitions each for the aggregated risk of windstorm and flooding losses. We consider the full Bernstein copula of Example 3 with $n_1 = n_2 = 34$ as well as the adaptive Bernstein copulas with $n_1 = n_2 = 10$, $n_1 = n_2 = 5$ and $n_1 = n_2 = 4$. For the sake of completeness, we also add estimates from the Gaussian copula, the independence as well as the co- and countermonotonicity copulas (see [15] (p. 11) for definitions).

The graphs in Figure 10 show the support points of the underlying adaptive scaled discrete skeletons as well as 5000 simulated pairs of the adapted Bernstein copulas.

Table 10 provides estimated values of the risk measures from the Monte Carlo simulations, which are given in Mio. monetary units. For the marginal distributions, the assumptions in [10] are used.

Table 10. Estimated values of $\text{VaR}_{0.005}$ for different grid types in Example 3.

Grid Type	**34 × 34**	**10 × 10**	**5 × 5**	**4 × 4**	Gaussian	**Independence**	**Comonotonic**	**Countermonotonic**
$\text{VaR}_{0.005}$	1348	1334	1356	1369	1386	1349	1500	1327

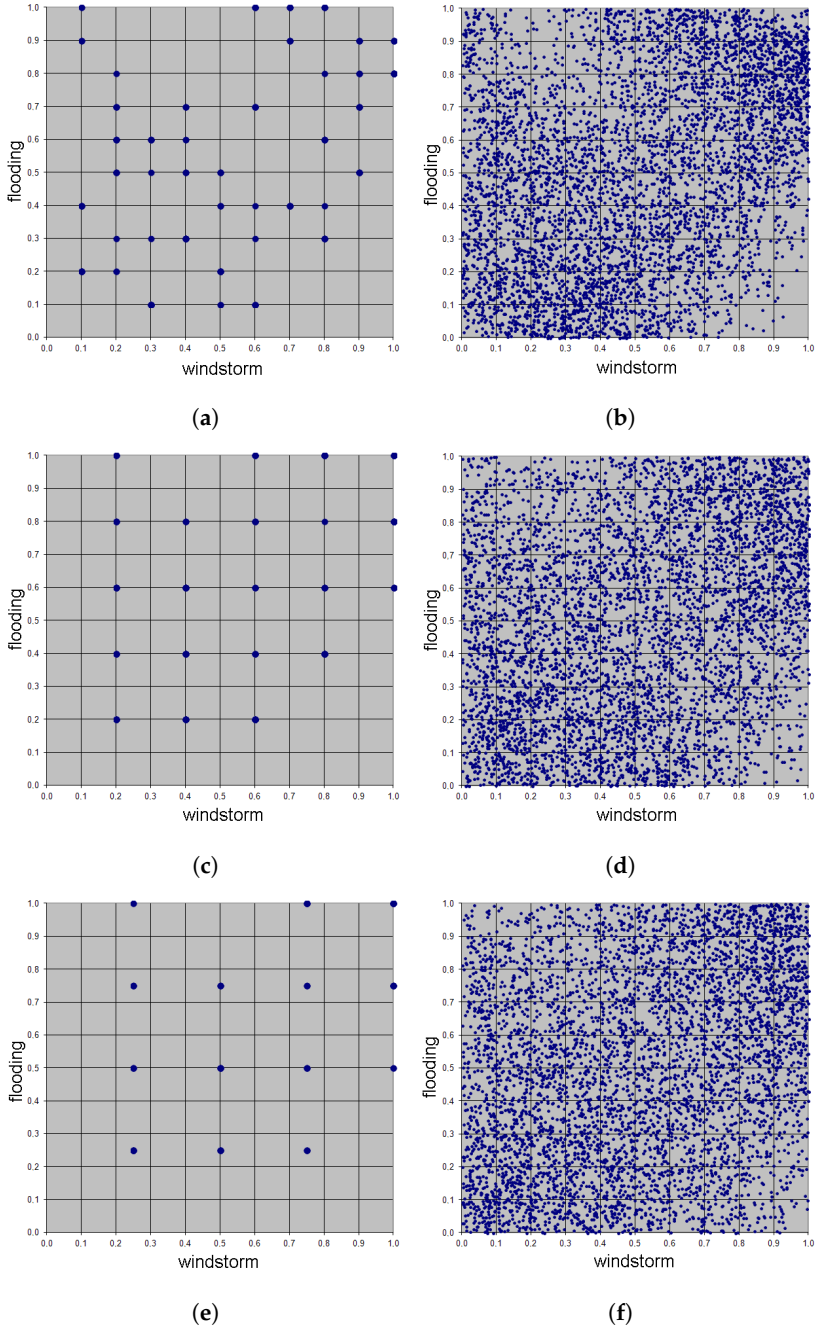


Figure 10. Plots related to the adaptive Bernstein copulas in Example 3. (**a**) Support points of the adaptive scaled discrete skeleton, $n_1 = n_2 = 10$. (**b**) Simulation of 5000 adapted Bernstein copula pairs, $n_1 = n_2 = 10$. (**c**) Support points of the adaptive scaled discrete skeleton, $n_1 = n_2 = 5$. (**d**) Simulation of 5000 adapted Bernstein copula pairs, $n_1 = n_2 = 5$. (**e**) Support points of the adaptive scaled discrete skeleton, $n_1 = n_2 = 5$. (**f**) Simulation of 5000 adapted Bernstein copula pairs, $n_1 = n_2 = 5$.

Seemingly the comonotonicity copula provides the largest $\text{VaR}_{0.005}$-estimate due to an extreme tail dependence, whereas the countermonotonicity copula provides the smallest $\text{VaR}_{0.005}$-estimate. Surprisingly, the $\text{VaR}_{0.005}$-estimates for the adaptive Bernstein copulas do not differ very much from each other (at most 2.5%) and are almost identical to the estimate from the independence copula here. Please note that the $\text{VaR}_{0.005}$-estimate for the Gaussian copula is slightly larger.

However, significant differences are visible if we look at the densities of the aggregated risk. The graphs in Figure 11 show empirical histograms for these densities under the models considered above, from 100,000 simulations each. Please note that the histogram for the full Bernstein copula has two peaks, whereas the other histograms show a more smooth behavior.

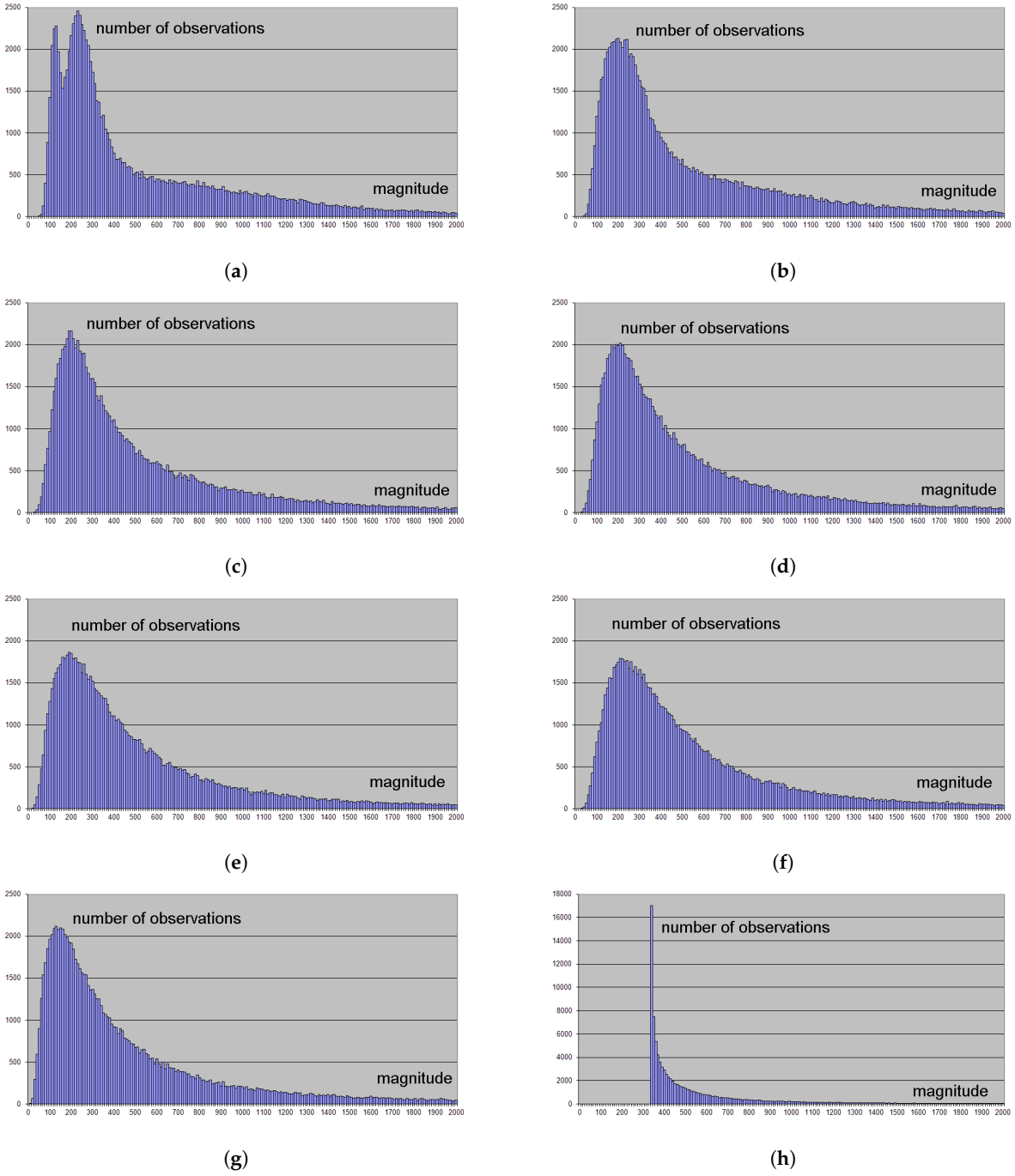


Figure 11. Empirical histograms for the densities of the aggregated risk in Example 3. (**a**) Bernstein copula, grid type 34×34. (**b**) Bernstein copula, grid type 10×10. (**c**) Bernstein copula, grid type 5×5. (**d**) Bernstein copula, grid type 4×4. (**e**) Gaussian copula. (**f**) Independence copula. (**g**) Comonotonicity copula. (**h**) Countermonotonicity copula.

Example 5. *Finally, we discuss a high-dimensional data set, which is also analyzed in [24]. It represents economically adjusted windstorm losses in 19 adjacent areas in Central Europe over a time period of 20 years. In Tables 11 and 12, the losses are given in Mio. monetary units.*

Table 11. Insurance losses in Example 5, part I.

Year	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6	Area 7	Area 8	Area 9	Area 10
1	23.664	154.664	40.569	14.504	10.468	7.464	22.202	17.682	12.395	18.551
2	1.080	59.545	3.297	1.344	1.859	0.477	6.107	7.196	1.436	3.720
3	21.731	31.049	55.973	5.816	14.869	20.771	3.580	14.509	17.175	87.307
4	28.990	31.052	30.328	4.709	0.717	3.530	6.032	6.512	0.682	3.115
5	53.616	62.027	57.639	1.804	2.073	4.361	46.018	22.612	1.581	11.179
6	29.950	41.722	12.964	1.127	1.063	4.873	6.571	11.966	15.676	24.263
7	3.474	14.429	10.869	0.945	2.198	1.484	4.547	2.556	0.456	1.137
8	10.020	31.283	21.116	1.663	2.153	0.932	25.163	3.222	1.581	5.477
9	5.816	14.804	128.072	0.523	0.324	0.477	3.049	7.791	4.079	7.002
10	170.725	576.767	108.361	41.599	20.253	35.412	126.698	71.079	21.762	64.582
11	21.423	50.595	4.360	0.327	1.566	64.621	5.650	1.258	0.626	3.556
12	6.380	28.316	3.740	0.442	0.736	0.470	3.406	7.859	0.894	3.591
13	124.665	33.359	14.712	0.321	0.975	2.005	3.981	4.769	2.006	1.973
14	20.165	49.948	17.658	0.595	0.548	29.350	6.782	4.873	2.921	6.394
15	78.106	41.681	13.753	0.585	0.259	0.765	7.013	9.426	2.180	3.769
16	11.067	444.712	365.351	99.366	8.856	28.654	10.589	13.621	9.589	19.485
17	6.704	81.895	14.266	0.972	0.519	0.644	8.057	18.071	5.515	13.163
18	15.550	277.643	26.564	0.788	0.225	1.230	26.800	64.538	2.637	80.711
19	10.099	18.815	9.352	2.051	1.089	6.102	2.678	4.064	2.373	2.057
20	8.492	138.708	46.708	3.680	1.132	1.698	165.600	7.926	2.972	5.237

Table 12. Insurance losses in Example 5, part II.

Year	Area 11	Area 12	Area 13	Area 14	Area 15	Area 16	Area 17	Area 18	Area 19
1	1.842	4.100	46.135	14.698	44.441	7.981	35.833	10.689	7.299
2	0.429	1.026	7.469	7.058	4.512	0.762	14.474	9.337	0.740
3	0.209	2.344	22.651	4.117	26.586	3.920	13.804	2.683	3.026
4	0.521	0.696	31.126	1.878	29.423	6.394	18.064	1.201	0.894
5	2.715	1.327	40.156	4.655	104.691	28.579	17.832	1.618	3.402
6	4.832	0.701	16.712	11.852	29.234	7.098	17.866	5.206	5.664
7	0.268	0.580	11.851	2.057	11.605	0.282	16.925	2.082	1.008
8	0.741	0.369	3.814	1.869	8.126	1.032	14.985	1.390	1.703
9	0.524	6.554	5.459	3.007	8.528	1.920	5.638	2.149	2.908
10	9.882	6.401	106.197	44.912	191.809	90.559	154.492	36.626	36.276
11	1.052	8.277	22.564	8.961	19.817	16.437	25.990	2.364	6.434
12	0.136	0.364	28.000	7.574	3.213	1.749	12.735	1.744	0.558
13	1.990	15.176	57.235	23.686	110.035	17.373	7.276	2.494	0.525
14	0.630	0.762	25.897	3.439	8.161	3.327	24.733	2.807	1.618
15	0.770	15.024	36.068	1.613	6.127	8.103	12.596	4.894	0.822
16	0.287	0.464	24.211	38.616	51.889	1.316	173.080	3.557	11.627
17	0.590	2.745	16.124	2.398	20.997	2.515	5.161	2.840	3.002
18	0.245	0.217	12.416	4.972	59.417	3.762	24.603	7.404	19.107
19	0.415	0.351	10.707	2.468	10.673	1.743	27.266	1.368	0.644
20	0.566	0.708	22.646	6.652	14.437	63.692	113.231	7.218	2.548

A statistical analysis of the data shows a good fit to lognormal $\mathcal{LN}(\mu, \sigma)$-distributions for the losses per Area k, $k = 1, \ldots, 19$. Thus, the parameters μ_k and σ_k for Area k can be estimated from the log data by calculating means and standard deviations (see Tables 13 and 14).

As expected, insurance losses in locally adjacent areas show a high degree of stochastic dependence, which can also be seen from the correlation Tables 15 and 16. In Table 15, correlations above 0.9 are highlighted.

Table 13. Values of the parameters μ_k and σ_k estimated from the log data in Example 5, part I.

Parameter	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6	Area 7	Area 8	Area 9	Area 10
μ_k	2.8063	4.0717	3.1407	0.6375	0.3984	1.2227	2.3210	2.2123	1.0783	2.1055
σ_k	1.2161	1.0521	1.2110	1.5685	1.2998	1.5987	1.1980	0.9882	1.1445	1.2531

Table 14. Values of the parameters μ_k and σ_k estimated from the log data in Example 5, part II.

Parameter	Area 11	Area 12	Area 13	Area 14	Area 15	Area 16	Area 17	Area 18	Area 19
μ_k	−0.3231	0.3815	3.0198	1.7488	3.0409	1.5501	3.0700	1.2444	0.9378
σ_k	1.0881	1.3353	0.8027	1.0033	1.1221	1.4765	0.9622	0.8577	1.2141

Table 15. Correlations between original losses in adjacent areas in Example 5.

Area	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19
A1	1	0.46	0.03	0.16	0.47	0.20	0.35	0.49	0.41	0.24	0.78	0.64	0.91	0.63	0.85	0.66	0.30	0.67	0.56
A2	0.46	1	0.64	0.78	0.67	0.36	0.51	0.76	0.57	0.51	0.58	−0.04	0.59	0.84	0.68	0.58	0.87	0.77	0.90
A3	0.03	0.64	1	0.93	0.41	0.26	0.11	0.16	0.33	0.16	0.08	−0.09	0.13	0.64	0.25	0.10	0.74	0.14	0.35
A4	0.16	0.78	0.93	1	0.54	0.36	0.16	0.25	0.43	0.19	0.22	−0.10	0.30	0.79	0.36	0.19	0.84	0.32	0.49
A5	0.47	0.67	0.41	0.54	1	0.41	0.35	0.51	0.84	0.63	0.59	0.02	0.64	0.67	0.59	0.50	0.58	0.71	0.67
A6	0.20	0.36	0.26	0.36	0.41	1	0.07	0.11	0.28	0.19	0.28	0.14	0.31	0.42	0.24	0.27	0.39	0.27	0.40
A7	0.35	0.51	0.11	0.16	0.35	0.07	1	0.44	0.27	0.19	0.48	−0.07	0.46	0.35	0.45	0.91	0.64	0.61	0.49
A8	0.49	0.76	0.16	0.25	0.51	0.11	0.44	1	0.50	0.75	0.61	−0.03	0.54	0.47	0.71	0.53	0.40	0.75	0.90
A9	0.41	0.57	0.33	0.43	0.84	0.28	0.27	0.50	1	0.66	0.68	−0.01	0.52	0.60	0.50	0.41	0.46	0.65	0.63
A10	0.24	0.51	0.16	0.19	0.63	0.19	0.19	0.75	0.66	1	0.33	−0.12	0.27	0.28	0.43	0.24	0.23	0.45	0.65
A11	0.78	0.58	0.08	0.22	0.59	0.28	0.48	0.61	0.68	0.33	1	0.19	0.79	0.65	0.80	0.73	0.43	0.84	0.74
A12	0.64	−0.04	−0.09	−0.10	0.02	0.14	−0.07	−0.03	−0.01	−0.12	0.19	1	0.44	0.21	0.28	0.17	−0.12	0.13	0.03
A13	0.91	0.59	0.13	0.30	0.64	0.31	0.46	0.54	0.52	0.27	0.79	0.44	1	0.71	0.86	0.74	0.47	0.76	0.65
A14	0.63	0.84	0.64	0.79	0.67	0.42	0.35	0.47	0.60	0.28	0.65	0.21	0.71	1	0.74	0.54	0.79	0.68	0.72
A15	0.85	0.68	0.25	0.36	0.59	0.24	0.45	0.71	0.50	0.43	0.80	0.28	0.86	0.74	1	0.69	0.47	0.71	0.75
A16	0.66	0.58	0.10	0.19	0.50	0.27	0.91	0.53	0.41	0.24	0.73	0.17	0.74	0.54	0.69	1	0.63	0.77	0.64
A17	0.30	0.87	0.74	0.84	0.58	0.39	0.64	0.40	0.46	0.23	0.43	−0.12	0.47	0.79	0.47	0.63	1	0.59	0.64
A18	0.67	0.77	0.14	0.32	0.71	0.27	0.61	0.75	0.65	0.45	0.84	0.13	0.76	0.68	0.71	0.77	0.59	1	0.86
A19	0.56	0.90	0.35	0.49	0.67	0.40	0.49	0.90	0.63	0.65	0.74	0.03	0.65	0.72	0.75	0.64	0.64	0.86	1

Table 16. Correlations between log losses in adjacent areas in Example 5.

Area	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19
A1	1	0.27	0.30	0.16	0.17	0.45	0.28	0.32	0.32	0.29	0.67	0.51	0.76	0.34	0.67	0.74	0.18	0.21	0.29
A2	0.27	1	0.48	0.66	0.39	0.37	0.71	0.69	0.52	0.64	0.30	−0.02	0.45	0.66	0.58	0.45	0.73	0.74	0.78
A3	0.30	0.48	1	0.70	0.40	0.31	0.42	0.51	0.58	0.53	0.18	0.07	0.21	0.32	0.54	0.26	0.47	0.21	0.57
A4	0.16	0.66	0.70	1	0.77	0.47	0.46	0.47	0.59	0.49	0.18	−0.13	0.33	0.50	0.47	0.18	0.76	0.43	0.54
A5	0.17	0.39	0.40	0.77	1	0.59	0.30	0.20	0.49	0.39	0.28	0.08	0.35	0.56	0.44	0.16	0.55	0.36	0.41
A6	0.45	0.37	0.31	0.47	0.59	1	0.14	0.01	0.36	0.34	0.33	0.12	0.48	0.46	0.48	0.37	0.59	0.17	0.50
A7	0.28	0.71	0.42	0.46	0.30	0.14	1	0.52	0.27	0.40	0.45	−0.07	0.31	0.31	0.46	0.62	0.63	0.58	0.57
A8	0.32	0.69	0.51	0.47	0.20	0.01	0.52	1	0.64	0.81	0.27	−0.02	0.38	0.35	0.56	0.35	0.28	0.62	0.63
A9	0.32	0.52	0.58	0.59	0.49	0.36	0.27	0.64	1	0.78	0.40	0.19	0.27	0.50	0.44	0.30	0.33	0.57	0.61
A10	0.29	0.64	0.53	0.49	0.39	0.34	0.40	0.81	0.78	1	0.21	−0.02	0.21	0.37	0.52	0.30	0.31	0.53	0.81
A11	0.67	0.30	0.18	0.18	0.28	0.33	0.45	0.27	0.40	0.21	1	0.47	0.49	0.45	0.60	0.67	0.20	0.45	0.39
A12	0.51	−0.02	0.07	−0.13	0.08	0.12	−0.07	−0.02	0.19	−0.02	0.47	1	0.44	0.21	0.24	0.46	−0.23	0.25	0.05
A13	0.76	0.45	0.21	0.33	0.35	0.48	0.31	0.38	0.27	0.21	0.49	0.44	1	0.55	0.60	0.71	0.37	0.39	0.24
A14	0.34	0.66	0.32	0.50	0.56	0.46	0.31	0.35	0.50	0.37	0.45	0.21	0.55	1	0.59	0.43	0.57	0.58	0.53
A15	0.67	0.58	0.54	0.47	0.44	0.48	0.46	0.56	0.44	0.52	0.60	0.24	0.60	0.59	1	0.59	0.36	0.35	0.63
A16	0.74	0.45	0.26	0.18	0.16	0.37	0.62	0.35	0.30	0.30	0.67	0.46	0.71	0.43	0.59	1	0.38	0.43	0.39
A17	0.18	0.73	0.47	0.76	0.55	0.59	0.63	0.28	0.33	0.31	0.20	−0.23	0.37	0.57	0.36	0.38	1	0.52	0.56
A18	0.21	0.74	0.21	0.43	0.36	0.17	0.58	0.62	0.57	0.53	0.45	0.25	0.39	0.58	0.35	0.43	0.52	1	0.60
A19	0.29	0.78	0.57	0.54	0.41	0.50	0.57	0.63	0.61	0.81	0.39	0.05	0.24	0.53	0.63	0.39	0.56	0.60	1

The following results have been achieved by Monte Carlo studies of 1,000,000 simulations each, based on various choices of the grid constants $n_i \equiv m$ for fixed numbers m. We first show scatter plots of each 5000 simulated adaptive Bernstein copula points for selected area combinations with high correlations (Area 1 vs. Area 13, Area 3 vs. Area 4, Area 7 vs. Area 16) and a low correlation (Area 3 vs. Area 18).

For comparison purposes, we start with $m = 100$, which corresponds to an extreme overfitting of the given data (see Figure 12). The scatter plots in Figure 13 correspond to the choice $m = 20$, which represents the ordinary Bernstein copula approach. The scatter plots in Figure 14 correspond to the choice $m = 17$, which represents a slightly adapted Bernstein copula approach. The scatter

plots in Figure 15 correspond to the choice $m = 13$, which represents a moderately adapted Bernstein copula approach. The final scatter plots in Figure 16 correspond to the choice $m = 7$, which represents a strongly adapted Bernstein copula approach.

As can clearly be seen, the choice of m influences significantly the shape of the adapted Bernstein copula. With decreasing magnitude of m, we see a more uniform distribution of adapted Bernstein copula points, as expected. Table 17 shows $\text{VaR}_{0.005}$-estimates depending on the choice of m.

Table 17. $\text{VaR}_{0.005}$-estimates for different values of m in Example 5.

m	100	20	17	13	7
$\text{VaR}_{0.005}$	2842	2247	2204	2105	1878

In contrast to the analysis of Example 3 (cf. Table 10), we see here that the choice of grid constants has a major influence on the estimated risk measure. The overfitted $\text{VaR}_{0.005}$ for $m = 100$ is more than 50% higher than the $\text{VaR}_{0.005}$ for $m = 7$.

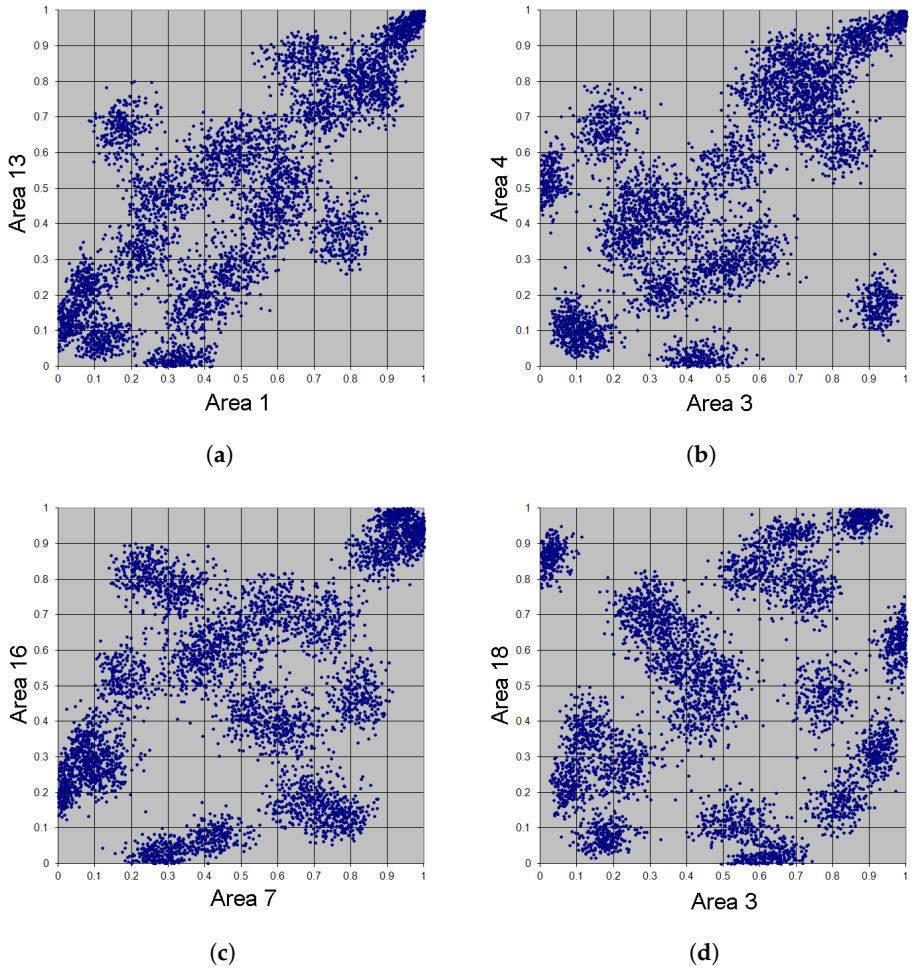


Figure 12. Simulation of 5000 adaptive Bernstein copula points in Example 5, $m = 100$. (**a**) Area 1 vs. Area 13. (**b**) Area 3 vs. Area 4. (**c**) Area 7 vs. Area 16. (**d**) Area 3 vs. Area 18.

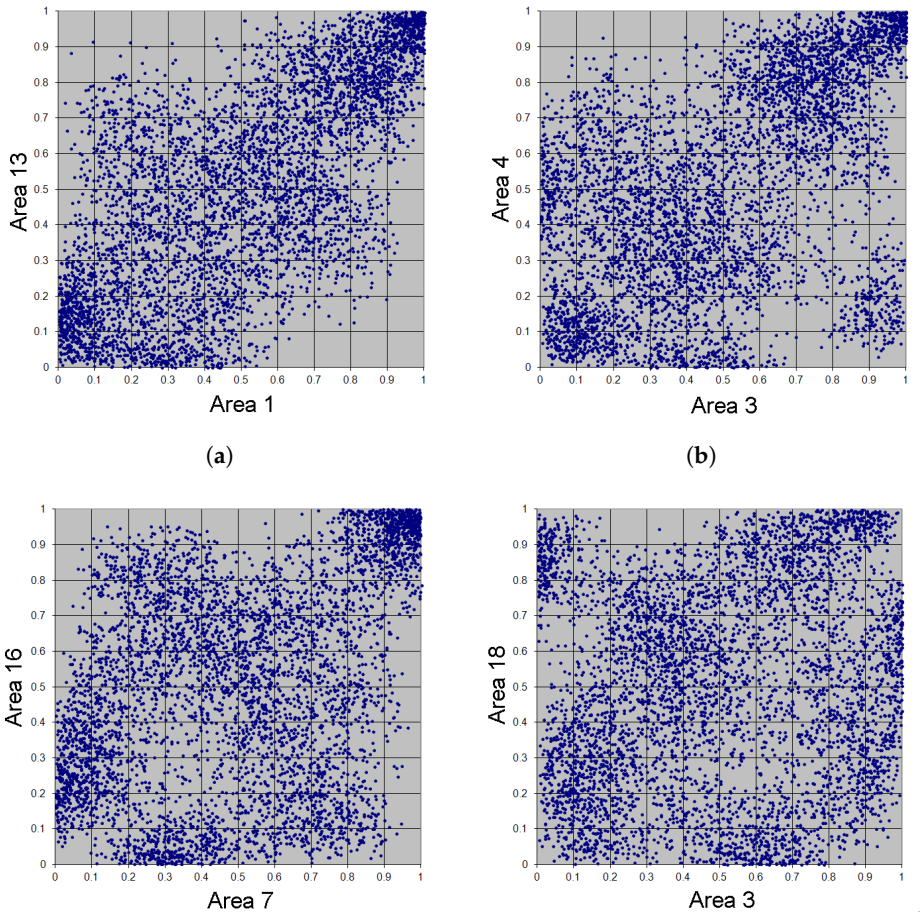


Figure 13. Simulation of 5000 adaptive Bernstein copula points in Example 5, $m = 20$. (**a**) Area 1 vs. Area 13. (**b**) Area 3 vs. Area 4. (**c**) Area 7 vs. Area 16. (d) Area 3 vs. Area 18.

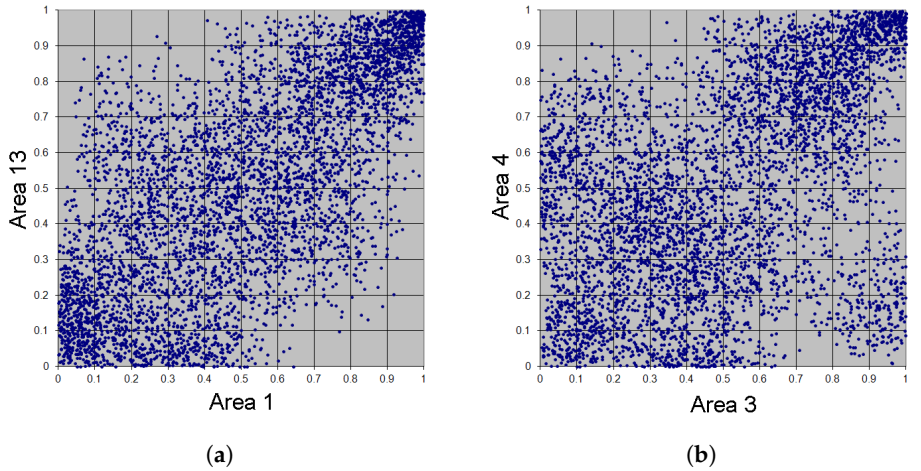
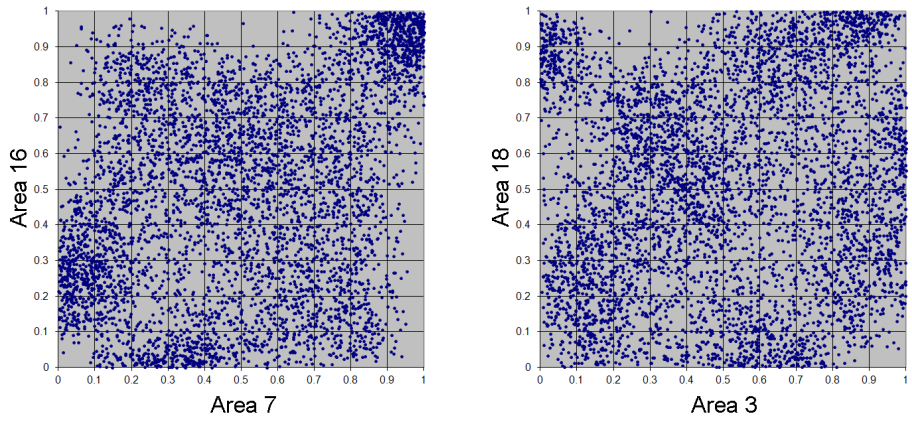
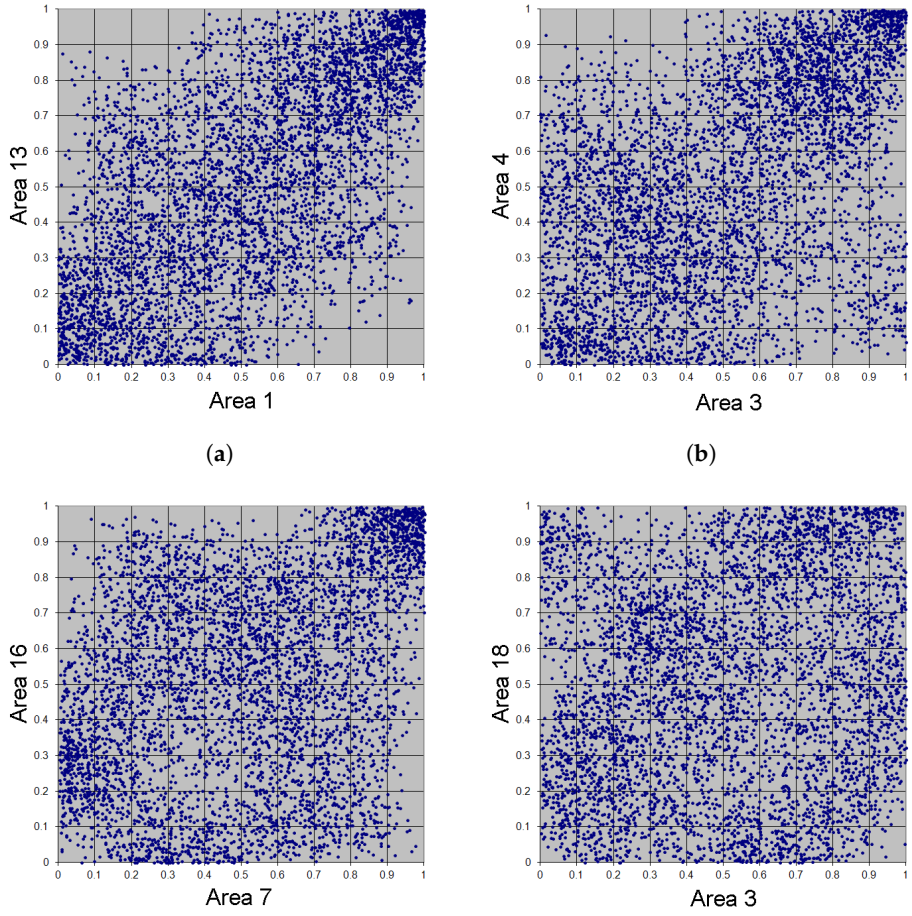


Figure 14. *Cont.*



(**c**) (**d**)

Figure 14. Simulation of 5000 adaptive Bernstein copula points in Example 5, $m = 17$. (**a**) Area 1 vs. Area 13. (**b**) Area 3 vs. Area 4. (**c**) Area 7 vs. Area 16. (**d**) Area 3 vs. Area 18.



(**a**) (**b**)

(**c**) (**d**)

Figure 15. Simulation of 5000 adaptive Bernstein copula points in Example 5, $m = 13$. (**a**) Area 1 vs. Area 13. (**b**) Area 3 vs. Area 4. (**c**) Area 7 vs. Area 16. (**d**) Area 3 vs. Area 18.

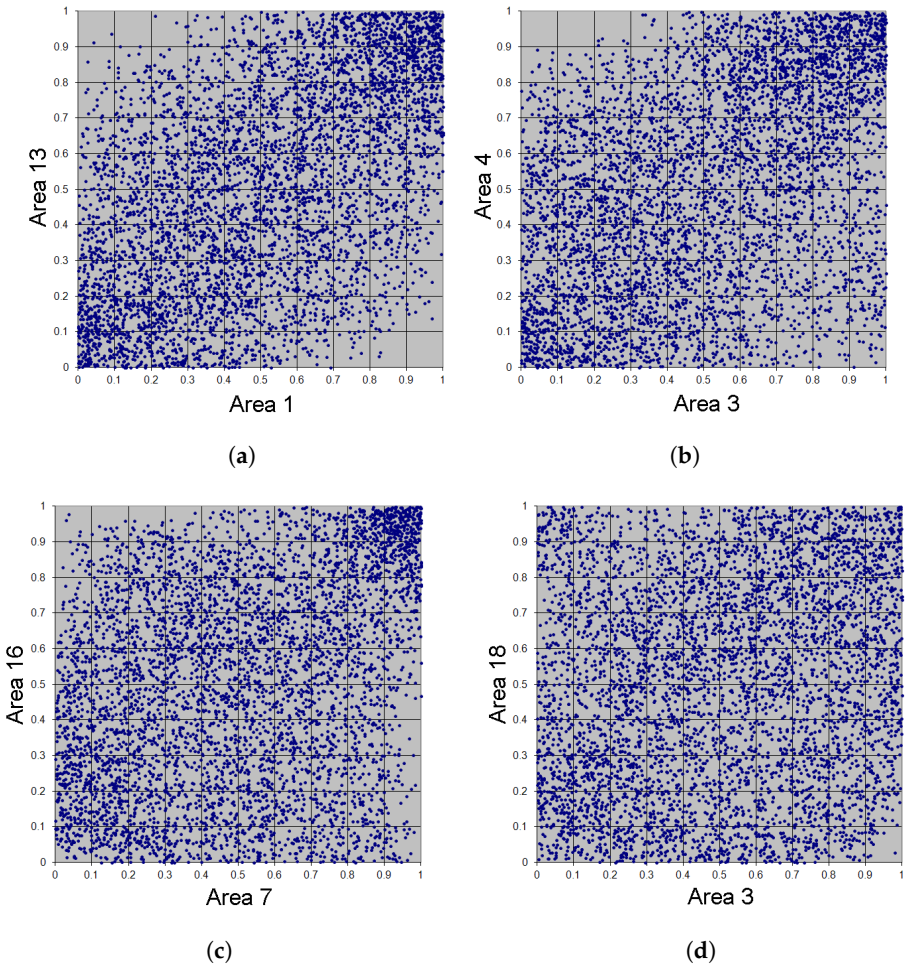


Figure 16. Simulation of 5000 adaptive Bernstein copula points in Example 5, $m = 7$. (**a**) Area 1 vs. Area 13. (**b**) Area 3 vs. Area 4. (**c**) Area 7 vs. Area 16. (**d**) Area 3 vs. Area 18.

7. Conclusions

Adaptive Bernstein copulas are an interesting tool for smoothing or, if desired, also sharpening the empirical dependence structure, in particular, in risk management applications when the number of observations and dimensions is moderate to large. The possibility of a smoothing of the dependence structure prevents in particular a kind of overfitting to copula models. In particular, the choice of the grid constants in the reduction procedure is arbitrary, the selected grid constants need not to be divisors of the number of observations. The method presented here also enables Monte Carlo studies for the comparison of different estimates of risk measures or the shape of the aggregate risk distribution. If the various estimates for the risk measure do not differ much for several adaptive strategies, this could make a sensitivity analysis, for instance, under Solvency II more profound. In other cases, when significant differences in the estimation of risk measures become apparent under various adaptive strategies, one should be cautious with a mere statistical risk assessment. Anyway, a kind of a worst case analysis derived from different approaches could be helpful here.

The method of reducing (or, if desired, sharpening) the complexity in the rank structures of the data might also be applied to partition-of-unity copulas (see [24–26]). With such copulas, tail dependence can be introduced to the dependence models, which cannot be obtained by Bernstein copulas alone due to the boundedness of the corresponding densities.

Author Contributions: The authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: The publication of this paper was funded by the University of Oldenburg.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bernstein, S. Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités. *Commun. Kharkov Math. Soc.* **1912**, *13*, 1–2.
2. Lorentz, G.G. *Bernstein Polynomials*, 2nd ed.; Chelsea Publishing Company: New York, NY, USA, 1986.
3. Babu, G.J.; Canty, A.J.; Chaubey, Y.P. Application of Bernstein polynomials for smooth estimation of a distribution and density function. *J. Statist. Plann. Inference* **2002**, *105*, 377–392. [CrossRef]
4. Cheng, C. The Bernstein polynomial estimator of a smooth quantile function. *Statist. Probab. Lett.* **1995**, *24*, 321–330. [CrossRef]
5. Cottin, C.; Pfeifer, D. From Bernstein polynomials to Bernstein copulas. *J. Appl. Funct. Anal.* **2014**, *9*, 277–288.
6. Guan, Z. Efficient and robust density estimation using Bernstein type polynomials. *J. Nonparametr. Stat.* **2016**, *28*, 250–271. [CrossRef]
7. Guan, Z. Bernstein polynomial model for grouped continuous data. *J. Nonparametr. Stat.* **2017**, *29*, 831–848. [CrossRef]
8. Leblanc, A. On estimating distribution functions using Bernstein polynomials. *Ann. Inst. Statist. Math.* **2012**, *64*, 919–943. [CrossRef]
9. Pfeifer, D.; Nešlehová, J. Modeling dependence in finance and insurance: The copula approach. *Bl. DGVFM* **2003**, *26*, 177–191. [CrossRef]
10. Pfeifer, D.; Straßburger, D.; Philipps, J. Modelling and simulation of dependence structures in nonlife insurance with Bernstein copulas. In Proceedings of the Paper presented on the occasion of the International ASTIN Colloquium, Helsinki, Finland, 1–4 June 2009.
11. Sancetta, A.; Satchell, S. The Bernstein copula and its applications to modeling and approximations of multivariate distributions. *Econom. Theory* **2004**, *20*, 535–562. [CrossRef]
12. Segers, J.; Sibuya, M.; Tsukahara, H. The empirical beta copula. *J. Multivar. Anal.* **2017**, *155*, 35–51. [CrossRef]
13. Vil'chevskii, N.O.; Shevlyakov, G.L. On the Bernstein polynomial estimators of distribution and quantile functions. *J. Math. Sci.* **2001**, *105*, 2626–2629. [CrossRef]
14. Wang, T.; Guan, Z. Bernstein polynomial model for nonparametric multivariate density. *Statistics* **2018**, *53*, 321–338. [CrossRef]
15. Durante, F.; Sempi, C. *Principles of Copula Theory*; CRC Press: Boca Raton, FL, USA, 2018.
16. Ibragimov, R.; Prokhorov, A. *Heavy Tails and Copulas. Topics in Dependence Modelling in Economics and Finance*; World Scientific: Singapore, 2017.
17. Rose, D. Modeling and Estimating Multivariate Dependence Structures with the Bernstein Copula. Ph.D. Thesis, Ludwig-Maximilians-Universität, München, Germany, 8 October 2015; Available online: https://edoc.ub.uni-muenchen.de/18757/ (accessed on 8 December 2020).
18. Avomo Ngomo, J.S.-L. Entwicklung und Implementierung eines Verfahrens zur Optimierung des Speicheraufwands bei Bernstein-und Verwandten Copulas. Ph.D. Thesis, Carl von Ossietzky Universität, Oldenburg, Germany, 18 December 2017; Available online: http://oops.uni-oldenburg.de/3457/ (accessed on 8 December 2020).
19. Masuhr, A.; Trede, M. Bayesian estimation of generalized partition of unity copulas. *Depend. Model.* **2020**, *8*, 119–131. [CrossRef]
20. Deheuvels, P. La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance. *Acad. R. Belg. Bull. Sci.* **1979**, *65*, 274–292. [CrossRef]
21. Butzer, P.L. On two-dimensional Bernstein polynomials. *Canad. J. Math.* **1953**, *5*, 107–113. [CrossRef]
22. Pfeifer, D.; Ragulina, O. Generating VaR scenarios under Solvency II with product beta distributions. *Risks* **2018**, *6*, 122. [CrossRef]
23. Junker, R.R.; Griessenberger, F.; Trutschnig, W. Estimating scale-invariant directed dependence of bivariate distributions. *Comput. Statist. Data Anal.* **2021**, *153*, 107058. [CrossRef]

24. Pfeifer, D.; Mändle, A.; Ragulina, O.; Girschig, C. New copulas based on general partitions-of-unity (part III)—the continuous case. *Depend. Model.* **2019**, *7*, 181–201. [CrossRef]
25. Pfeifer, D.; Tsatedem, H.A.; Mändle, A.; Girschig, C. New copulas based on general partitions-of-unity and their applications to risk management. *Depend. Model.* **2016**, *4*, 123–140. [CrossRef]
26. Pfeifer, D.; Mändle, A.; Ragulina, O. New copulas based on general partitions-of-unity and their applications to risk management (part II). *Depend. Model.* **2017**, *5*, 246–255. [CrossRef]