A class of high dimensional copulas based on products of bivariate copulas

Gildas Mazo, Stephane Girard, Florence Forbes

A class of high dimensional copulas based on products of bivariate copulas

Gildas Mazo, Stéphane Girard and Florence Forbes

Inria and Laboratoire Jean Kuntzmann, Grenoble, France

Abstract

Copulas are a useful tool to model multivariate distributions. While there exist various families of bivariate copulas, much fewer has been done when the dimension is higher. To this aim an interesting class of copulas based on products of transformed copulas has been proposed. However the use of this class for practical high dimensional problems remains challenging. Constraints on the parameters and the product form render inference, and in particular the likelihood computation, difficult. In this paper we propose a new class of high dimensional copulas based on a product of transformed bivariate copulas. No constraints on the parameters refrain the applicability of the proposed class which is well suited for applications in high dimension. Furthermore the analytic forms of the copulas within this class allow to associate a natural graphical structure which helps to visualize the dependencies and to compute the likelihood efficiently even in high dimension.

1 Introduction

The modelling of random multivariate events is a central problem in various scientific domains and the construction of multivariate distributions able to properly model the variables at play is challenging. A useful tool to deal with this problem is the concept of copulas. Let $(X_1, \dots, X_d)$ be a random vector with distribution function F. Let F_i be the (continuous) marginal distribution function of X_i, $i = 1, \dots, d$. By Sklar's Theorem [16], there exists a unique function C such that

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)). \quad (1)$$

This function C is called the copula of F and is the d-dimensional distribution function of the random vector $(F_1(X_1), \dots, F_d(X_d))$. For a general account on copulas, see, e.g. [15]. Copulas are interesting since they permit to impose a dependence structure on pre-determined marginal distributions.

While there exist many copulas in the bivariate case, it is less clear how to construct copulas in higher dimension. In the presence of non-Gaussianity and/or tail dependence, various constructions have been adopted, such as, for instance, Archimedean copulas [9], Vines [1] or elliptical copulas [6]. Because Archimedean copulas possess only a few parameters, they lack flexibility in high-dimension. Vines, on the opposite, achieve greater flexibility but at the price of an increase in complexity in the modeling process. The use of elliptical copulas goes together with assuming a similar dependence pattern among all pairs of variables. This may be undesirable in applications. Moreover, they have in general as many as $O(d^2)$ parameters and it is difficult to carry out maximum likelihood inference [4].

Another approach [14] aims at constructing a multivariate copula as a product of transformed bivariate copulas. This approach possesses several advantages. A probabilistic interpretation is available and thus the generation of random vectors is straightforward. The resulting copula is explicit, leading to explicit bounds on dependence coefficients of the bivariate marginals. The class of copulas which can be constructed from this approach is large and can cover a wide range of dependencies. Finally the analysis of extreme values can be performed by constructing extreme-value copulas.

However, although many copulas with different features can be built, the use of this approach for practical applications remains challenging. Indeed, two pitfalls render inference difficult: first constraints on the parameters and second the product form which makes the computation of the density, hence the potential likelihood, complicated to compute even numerically. These issues may explain why, as far as we know, only very simple forms of this approach have been proposed with at most three variables under consideration.

The main contribution of this paper is to revisit the product of transformed copulas in order to propose a new class of high-dimensional copulas well suited for high-dimensional applications. First, there are no constraints on the parameters anymore. Moreover, a graphical structure associated to the copulas within this class permits to visualize the dependencies and to efficiently compute the likelihood, thus opening the door to realistic high dimensional applications.

The rest of this paper is organized as follows. Section 2 reviews the product of transformed copulas and important properties such as random generation and the ability to construct extreme-value models. Section 3 presents the new class and enlightens the link with the product of transformed copulas. A graphical structure is introduced. Section 4 discusses the dependence properties of bivariate marginals of the proposed class by giving bounds on some of the most popular dependence coefficients such as the Spearman's rho, Kendall's tau, and tail dependence coefficients. Proofs are postponed to the Appendix.

2 Product of transformed copulas

It is easily seen that a product of copulas is not a copula in general. Nonetheless the next theorem due to Liebscher [14] shows that up to marginal transformations, a product of copulas can lead to a well defined copula.

Theorem 1. *Assume $C_1, \dots, C_K : [0,1]^d \to [0,1]$ are copulas. Let $g_{si} : [0,1] \to [0,1]$ for $s = 1, \dots, K$, $i = 1, \dots, d$ be functions with the property that each of them is strictly increasing or is identically equal to 1. Suppose that $\prod_{s=1}^{K} g_{si}(v) = v$ for $v \in [0,1], i = 1, \dots, d$, and $\lim_{v \to 0} g_{si}(v) = 0$ for $s = 1, \dots, K$, $i = 1, \dots, d$. Then*

$$\overline{C}(u_1, \dots, u_d) = \prod_{s=1}^{K} C_s \left(g_{s1}(u_1), \dots, g_{sd}(u_d) \right) \tag{2}$$

is also a copula.

The probabilistic interpretation of (2) is as follows. Let

$$(U_1^{(1)}, \ldots, U_d^{(1)}), \ldots, (U_1^{(K)}, \ldots, U_d^{(K)})$$

be K independent random vectors having distribution function $C_1, \ldots, C_K$ respectively. Let g_{si}, $s = 1, \ldots, K$, $i = 1, \ldots, d$ be as in Theorem 1 and define $g_{si}^{-1}(v) := 0$ for $v \leq g_{si}(0)$ and $J_i = \{s \in \{1, \ldots, K\} : g_{si} \neq 1\}$. Then $\overline{C}$ is the joint distribution function of the random vector

$$(\max_{s \in J_1}\{g_{s1}^{-1}(U_1^{(s)})\}, \ldots, \max_{s \in J_d}\{g_{sd}^{-1}(U_d^{(s)})\}). \tag{3}$$

If there exists a random generation procedure for C_s, $s = 1, \ldots, K$ then thanks to (3) a random generation procedure for $\overline{C}$ can be derived as well.

The statistical analysis of extreme values should theoretically be carried out with the help of extreme-value copulas [8]. A copula $C_\#$ is an extreme-value copula if there exists a copula C such that

$$C_\#(u_1, \ldots, u_d) = \lim_{n \uparrow \infty} C^n(u_1^{1/n}, \ldots, u_1^{1/n}), \tag{4}$$

for every $(u_1, \ldots, u_d) \in [0,1]^d$. A copula $C_\#$ is said to be max-stable if for every integer $n \geq 1$ and every $(u_1, \ldots, u_d) \in [0,1]^d$

$$C_\#(u_1^{1/n}, \ldots, u_d^{1/n})^n = C_\#(u_1, \ldots, u_d).$$

Extreme-value copulas correspond exactly to max-stable copulas [8]. Extreme-value copulas are the only theoretical well-grounded copulas to model extremal dependence. A typical example in hydrology is the analysis of rainfall annual maxima at different locations [3]. Theorem 1 can be used to construct extreme-value copulas as shown in the next proposition due to [5].

Proposition 1. *In* (2) *let* $g_{si}(v) = v^{\theta_{si}}$, $v \in [0,1]$ *with* $\theta_i \in [0,1]$ *and* $\sum_{s=1}^{K} \theta_{si} = 1$ *for* $i = 1, \ldots, d$. *If* C_s, $s = 1, \ldots, K$ *is max-stable then so is* $\overline{C}$.

Out of the context of extreme values, applications of Theorem 1 can be found, for instance, in the analysis of directional dependence [13] ($K = d = 2$), finance [2] ($K = d = 2$) and hydrology [5] ($K = 2$, $d = 3$).

We are not aware of applications of Theorem 1 in practice when $K > 2$ or $d > 3$. As pointed out in the introduction, the product form (2) renders the density $\frac{\partial^d \overline{C}(u_1,\ldots,u_d)}{\partial u_1 \ldots \partial u_d}$, hence the likelihood, complicated to compute even numerically. Furthermore, the constraints $\prod_{s=1}^{K} g_{si}(v) = v$, $v \in [0,1]$, $i = 1, \ldots, d$ in Theorem 1 are not easy to deal with in practice.

The next section aims at extending the applicability of (2) to high-dimensional problems.

3 The proposed copula class

The product over $s \in \{1, \ldots, K\}$ in (2) can be taken over $s \in S$, where S is an arbitrary finite set, yielding

$$\overline{C}(u_1, \ldots, u_d) = \prod_{s \in S} C_s\left(g_{s1}(u_1), \ldots, g_{sd}(u_d)\right). \tag{5}$$

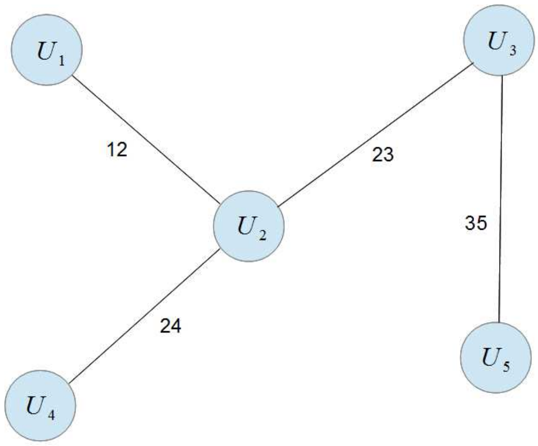

Figure 1: Graphical representation of the set $S = \{\{12\},\{24\},\{23\},\{35\}\}$. $N(1) = \{\{12\}\}$, $N(2) = \{\{12\},\{23\},\{24\}\}$, $N(3) = \{\{23\},\{35\}\}$, $N(4) = \{\{24\}\}$ and $N(5) = \{\{35\}\}$.

In particular, an element $s \in S$ can represent a pair of the variables at play. More precisely, let $U_1,\ldots,U_d$ be d random variables such that U_i, $i = 1,\ldots,d$ has a standard uniform distribution. Denote by $\{ij\}$ the index of the pair (U_i,U_j) and let $S \subset \{\{ij\} : i,j = 1,\ldots,d,\ j > i\}$ be a subset of the set of the pair indices. The cardinal of S, $|S|$, is less or equal to $d(d-1)/2$. The pair index $s \in S$ is said to contain the variable index i if $s = \{ik\}$ for $k \neq i$. Let us introduce $N(i) = \{s \in S : s \text{ contains } i\}$. $N(i)$ is called the set of neighbors of i and has cardinal $|N(i)| = n_i$. It is natural to associate a graph to the set S as follows: an element $s = \{ij\} \in S$ is an edge linking U_i and U_j in the graph whose nodes are the variables $U_1,\ldots,U_d$. The example $S = \{\{12\},\{24\},\{23\},\{35\}\}$ is illustrated in Figure 1. The proposed class of copulas is written below. For $u = (u_1,\ldots,u_d) \in [0,1]^d$, let the functional

$$\overline{C}(u) = \prod_{\{ij\}\in S} C_{ij}(u_i^{1/n_i}, u_j^{1/n_j}), \tag{6}$$

where the C_{ij}'s are bivariate copulas. Keeping in mind the graphical representation, $\overline{C}$ in (6) is a product over the edges. For instance, when $S = \{\{12\},\{24\},\{23\},\{35\}\}$ as in Figure 1, (6) writes

$$\overline{C}(u_1,u_2,u_3,u_4,u_5) = C_{12}(u_1,u_2^{1/3})C_{24}(u_2^{1/3},u_4)C_{23}(u_2^{1/3},u_3^{1/2})C_{35}(u_3^{1/2},u_5).$$

In the following, (6) is referred to as the Product of Bivariate Copulas (PBC) copula, or PBC model. The next theorem establishes that (6) is a copula and makes the link with Theorem 1.

Theorem 2. *If in* (5)*:*

(i) for $s = \{ij\} \in S$, $C_s = C_{\{ij\}}$ takes as argument exactly two functions non identically equal to one, namely, g_{si} and g_{sj}, and

(ii) for $i = 1,\ldots,d$ and $s \in N(i)$, g_{si} do not depend on s;

then the only copula which can be constructed from (5) *is the PBC model* (6)*, where* C_{ij} *is defined by*

$$C_{ij}(u,v) = C_{\{ij\}}(1,\dots,1,u,1,\dots,1,v,1,\dots,1), \quad (u,v) \in [0,1]^2,$$

where in $(1,\dots,1,u,1,\dots,1,v,1,\dots,1)$*,* u *and* v *are at the* i*-th and* j*-th positions respectively.*

Condition (i) in Theorem 2 simply means that only bivariate copulas are allowed in the construction. The simplification (ii) achieves two goals: first to reduce the number of parameters (an important feature in high-dimension), and second to intrinsically satisfy the constraints $\prod_{s\in S} g_{si}(v) = v$, $v \in [0,1]$, $i = 1,\dots,d$ in the assumptions of Theorem 1. If assumption (ii) in Theorem 2 was not made, one could take $g_{si}(v) = v^{\theta_{si}}$, $s \in S$, $i = 1,\dots,d$, $\theta_{si} \in [0,1]$ with the constraints

$$\sum_{s\in N(i)} \theta_{si} = \sum_{k:\{ki\}\in S} \theta_{ki,i} = 1, \quad i = 1,\dots,d. \tag{7}$$

Furthermore if in the PBC copula (6) C_{ij} is governed by a parameter θ_{ij}, that is, $C_{ij}(\cdot,\cdot) = C(\cdot,\cdot;\theta_{ij})$, the copula $\overline{C}$ would write

$$\overline{C}(u) = \prod_{\{ij\}\in S} C(u_i^{\theta_{ij,i}}, u_j^{\theta_{ij,j}}; \theta_{ij}). \tag{8}$$

The constraints (7) would be difficult to handle in practice, and, moreover, the number of parameters of (8) would increase quadratically with the dimension. Indeed, one would have $|S|d - d$ parameters $\theta_{ki,i}$, $k : \{ki\} \in S$, $i = 1,\dots,d$ and $|S|$ parameters θ_{ij}, $\{ij\} \in S$. In the case where the graph associated to S is a tree, as it would be the case in practice then $|S| = d-1$, yielding $O(d^2)$ parameters. As a comparison, in the PBC model (6), there are no constraints and only $O(d)$ parameters in total.

By (1), the PBC copula (6) is associated to a distribution function F with continuous marginals F_i, $i = 1,\dots,d$, such that

$$F(x_1,\dots,x_d) = \overline{C}(F_1(x_1),\dots,F_d(x_d)), \quad (x_1,\dots,x_d) \in \mathbf{R}^d. \tag{9}$$

By substituting (6) into (9), it is easy to see that F writes

$$F(x_1,\dots,x_d) = \prod_{\{ij\}\in S} F_{ij}(x_i,x_j), \quad (x_1,\dots,x_d) \in \mathbf{R}^d, \tag{10}$$

where F_{ij}, $\{ij\} \in S$, is a bivariate distribution function such that the first (respectively the second) marginal $F_{ij,1}$ (respectively $F_{ij,2}$) only depends on i (respectively j). It is interesting to note that the converse is also true as stated in the following proposition.

Proposition 2. *The copula corresponding to the distribution function* F *in* (10) *writes as the PBC copula* (6)*. Conversely, the distribution function corresponding to the PBC copula* (6) *writes as* F *in* (10)*.*

4 Dependence properties and max-stability

Let $\overline{C}$ be the PBC copula (6). First the dependence properties of a pair (U_k, U_l) whose copula is the bivariate copula $\overline{C}_{kl}(u_k, u_l) = \overline{C}(1, \ldots, 1, u_k, 1, \ldots, 1, u_l, 1, \ldots, 1)$ are studied. The conditions under which the PBC model (6) is an extreme-value copula are given afterwards.

Proposition 3. *The bivariate marginal $\overline{C}_{kl}$ is given by*

$$\overline{C}_{kl}(u_k, u_l) = \begin{cases} u_k^{(n_k-1)/n_k} u_l^{(n_l-1)/n_l} C_{kl}(u_k^{1/n_k}, u_l^{1/n_l}) & \text{if } \{kl\} \in S, \\ u_k u_l & \text{otherwise.} \end{cases} \tag{11}$$

Example 1. *If in* (11) *one puts $\kappa = 1/n_k$ and $\lambda = 1/n_l$, then $\overline{C}_{kl}(u_k, u_l) = u_k^{1-\kappa} u_l^{1-\lambda} C_{kl}(u_k^{\kappa}, u_l^{\lambda})$ corresponds to the mechanism proposed in [7], Proposition 2.*

Example 2. *If in* (11) *C_{kl} is a Marshall-Olkin copula (see for instance [15], p.53) with parameters $0 \leq \alpha, \beta \leq 1$ (denoted by MO(α, β)), that is,*

$$C_{kl}(u_k, u_l) = \min(u_k^{1-\alpha} u_l, u_l^{1-\beta} u_k),$$

then $\overline{C}_{kl}$ is MO$(\alpha/n_k, \beta/n_l)$. If $\alpha = \beta$ then C_{kl} is a Cuadras-Augé copula and $\overline{C}_{kl}$ is MO$(\alpha/n_k, \alpha/n_l)$. If $\alpha = \beta = 0$ then both C_{kl} and $\overline{C}_{kl}$ are the independence copula. If $\alpha = \beta = 1$ then C_{kl} is the Fréchet upper bound copula and $\overline{C}_{kl}$ is MO$(1/n_k, 1/n_l)$.

Let (U, V) be a random vector with copula C. The dependence between U and V is positive if, roughly speaking, U and V tend to be large or small together. Below are recalled a few definitions of statistical concepts about positive dependence. It is said that the copula C has the TP2 (totally positive of order 2) property if and only if

$$C(u_1, u_2)C(v_1, v_2) \geq C(u_1, v_2)C(v_1, u_2), \text{ for all } u_1 < v_1 \text{ and } u_2 < v_2. \tag{12}$$

Also, C is said to be PQD (positive quadrant dependent) if $C(u, v) \geq uv$ for all $(u, v) \in [0, 1]^2$. The random variable V is said to be LTD (left tail decreasing) in U if for all $v \in [0, 1]$, the function $u \mapsto P(V \leq v | U \leq u)$ is decreasing in u. The dependence between U and V can be quantified through dependence measures such as the Kendall's tau or the Spearman's rho respectively given by

$$\tau = 4 \int_{[0,1]^2} C(u, v)\, dC(u, v) - 1, \tag{13}$$

$$\rho = 12 \int_{[0,1]^2} C(u, v)\, du\, dv - 3. \tag{14}$$

The dependence in the upper and lower tails can be respectively measured with

$$\lambda^{(U)} = \lim_{u \uparrow 1} \frac{1 - 2u + C(u, u)}{1 - u} \in [0, 1], \quad \lambda^{(L)} = \lim_{u \downarrow 0} \frac{C(u, u)}{u} \in [0, 1].$$

See [15] and [12] for further details about these concepts. Let us denote by $\overline{\tau}_{kl}$, $\overline{\rho}_{kl}$, $\overline{\lambda}_{kl}^{(U)}$ and $\overline{\lambda}_{kl}^{(L)}$ the Kendall's tau, Spearman's rho, upper tail dependence coefficient and lower tail dependence coefficient of the copula $\overline{C}_{kl}$ in (11) respectively. Similarly denote respectively these dependence measures for C_{kl} in (11) by τ_{kl}, ρ_{kl}, $\lambda_{kl}^{(U)}$ and $\lambda_{kl}^{(L)}$. As shown in Proposition 3, $\overline{C}_{kl}$ is a bivariate marginal of the PBC copula (6) and one may apply the results of [14] to obtain the following.

Proposition 4. *(i) If in* (11) C_{kl} *is TP2, LTD or PQD then* $\overline{C}_{kl}$ *is also TP2, LTD or PQD respectively;*

(ii) The following inequalities hold

$$\overline{\tau}_{kl} \leq \tau_{kl}, \quad \overline{\rho}_{kl} \leq \rho_{kl}, \quad \overline{\lambda}_{kl}^{(U)} \leq \lambda_{kl}^{(U)}, \quad \overline{\lambda}_{kl}^{(L)} \leq \lambda_{kl}^{(L)}.$$

The results of Proposition 4 *(ii)* are precised in the next proposition where explicit bounds in terms of the number of neighbors are given. The behavior of (11) when the number of neighbors tends to infinity is also studied.

Proposition 5. *We have* $\overline{\lambda}_{kl}^{(L)} = 0$ *and* $\overline{\lambda}_{kl}^{(U)} \leq \min(1/n_k, 1/n_l)$. *The lower and upper bounds for* $\overline{\rho}_{kl}$ *and* $\overline{\tau}_{kl}$ *are respectively given by*

$$a_\rho(n_k, n_l) \leq \overline{\rho}_{kl} \leq b_\rho(n_k, n_l),$$
$$a_\tau(n_k, n_l) \leq \overline{\tau}_{kl} \leq b_\tau(n_k, n_l),$$

with

$$\begin{aligned}
a_\rho(n_k, n_l) &= \frac{6\beta(2n_k - 1, 2n_l - 1)n_k n_l}{(2n_k + 2n_l - 1)(n_k + n_l - 1)} - \frac{3}{(2n_k - 1)(2n_l - 1)}, \\
b_\rho(n_k, n_l) &= \frac{3}{2n_k + 2n_l - 1}, \\
a_\tau(n_k, n_l) &= \frac{\beta(2n_l - 1, 2n_k - 1)}{n_k + n_l - 1} - \frac{2}{(2n_k - 1)(2n_l - 1)}, \\
b_\tau(n_k, n_l) &= \frac{1}{n_k + n_l - 1},
\end{aligned}$$

where β *denotes the* β*-function,* $\beta(x, y) = \int_0^1 t^{x-1}(1-t)^{b-1}dt$. *Furthermore, as* $\max(n_k, n_l) \to \infty$, *we have* $\overline{C}_{kl}(u, v) \to uv$ *for all* $(u, v) \in [0, 1]^2$.

The above results show that we are facing a tradeoff: on the one hand, the larger the cardinal of S (or the more connected the graph associated to S), the less the pairs in S are able to model strong dependencies. On the other hand, the smaller the cardinal of S, the more there are independent pairs (since there are less pairs in S). To illustrate Proposition 5, numerical values of the bounds are computed in Table 1 for different numbers of neighbors (n_k, n_l).

It is easy to construct extreme-value copulas belonging to the PBC class (6).

Proposition 6. *If in the PBC copula* (6), C_{ij} *is an extreme-value copula for* $\{ij\} \in S$, *then* $\overline{C}$ *is also an extreme-value copula.*

Eventually, below are given examples where the PBC model (6) is an extreme-value copula.

Example 3. *All copulas* $\overline{C}_{kl}$ *in Example 2 are max-stable since Marshall-Olkin copulas are max-stable. Thus* $\overline{C}$ *is an extreme-value copula.*

Table 1: Lower and upper bounds [lower, upper] for the Spearman'rho $\overline{\rho}_{kl}$, Kendall's tau $\overline{\tau}_{kl}$ and upper tail dependence coefficient $\overline{\lambda}_{kl}$ depending on the number of neighbors (n_k, n_l).

coefficient	$\overline{\rho}_{kl}$	$\overline{\tau}_{kl}$	$\overline{\lambda}_{kl}$
(n_k, n_l)			
$(1,2)$	$[-0.6, 0.6]$	$[-0.13, 0.25]$	$[0, 0.5]$
$(2,2)$	$[-0.3, 0.43]$	$[-0.21, 0.33]$	$[0, 0.5]$
$(1,3)$	$[-0.43, 0.43]$	$[-0.33, 0.33]$	$[0, 0.33]$
$(2,3)$	$[-0.19, 0.33]$	$[-0.13, 0.25]$	$[0, 0.33]$
$(3,3)$	$[-0.12, 0.27]$	$[-0.08, 0.2]$	$[0, 0.33]$

Example 4. *Let C_{kl} in (11) be a (max-stable) Gumbel copula, that is,*

$$C_{kl}(u_k, u_l) = \exp - \left[(-\log u_k)^\theta + (-\log u_l)^\theta\right]^{1/\theta}, \quad \theta \geq 1.$$

Then $\overline{C}_{kl}$ is also max-stable, hence, $\overline{C}$ is an extreme-value copula.

5 Simulation and inference

5.1 Simulation

Thanks to the probabilistic interpretation given in (3), data simulation from the PBC copula (6) is easy. The generation procedure is given below.

- For all $\{ij\} \in S$, generate $(U_i^{(ij)}, U_j^{(ij)}) \sim C_{ij}$.
- For all $i = 1, \ldots, d$, compute $\overline{U}_i = \max_{k \in \{1,\ldots,d\}:\{ki\} \in S} \left\{ \left(U_i^{(ki)}\right)^{n_i} \right\}$.

The resulting vector $(\overline{U}_1, \ldots, \overline{U}_d)$ has distribution (6).

5.2 Inference

In this section it is assumed that the copulas C_{ij} of the PBC model (6) depend on parameters θ_{ij}'s. The parameter vector is denoted by $\boldsymbol{\theta} = (\theta_{ij})_{\{ij\} \in S}$. It is also assumed that we are given a sample of i.i.d data vectors with marginals transformed to standard uniform random variables and joint distribution given by (6). In Section 3 a graph has been associated to the model. Here the graph is assumed to be a tree, that is, there is no cycles in the graph (then $|S| = d - 1$). Let $V = \{1, \ldots, d\}$ and $u = (u_1, \ldots, u_d)$ a vector in $[0,1]^d$. For a subset $A \subset V$, the notation $\partial_{u_A} \overline{C}(u; \boldsymbol{\theta})$ stands for the derivative of $\overline{C}$ with respect to all the variables in A. For instance the density (hence the likelihood) writes

$$\frac{\partial^d \overline{C}(u; \boldsymbol{\theta})}{\partial_{u_1} \ldots \partial_{u_d}} = \partial_{u_V} \overline{C}(u; \boldsymbol{\theta}) = \overline{c}(u; \boldsymbol{\theta}). \tag{15}$$

It has been shown in [11] how to use a message-passing algorithm on the graphical structure to compute the likelihood (15) and the gradient with respect to the parameter vector,

$$\left(\frac{\partial \overline{c}(u;\boldsymbol{\theta})}{\theta_{ij}}\right)_{\{ij\in S\}}.$$

The purpose here is not to give the algorithm, but rather to provide an intuitive idea of it. The reader is referred to [11] for the complete algorithm and [10] for a more detailed explanation. To keep the notation simple, the dependence on the parameter vector $\boldsymbol{\theta}$ is dropped in the remaining of the section.

Let us write

$$\overline{C}(u) = \prod_{\{ij\}\in S} C_{ij}(u_i^{1/n_i}, u_j^{1/n_j}) =: \prod_{\{ij\}\in S} \Phi_{ij}(u_i, u_j).$$

and let an arbitrary variable index i (the root) be given. Let τ_s^i denote the subtree rooted at the variable indexed by i and containing the function indexed by s (see Figure 2). The idea is to note that, since the graph is a tree, the copula $\overline{C}$ can be decomposed over the subtrees rooted at i. We can write

$$\overline{C}(u) = \prod_{s\in S} \Phi_s(u) =: \prod_{s\in N(i)} T_{\tau_s^i}(u)$$

where $T_{\tau_s^i}(u)$ corresponds to the product of all functions located in the subtree τ_s^i. Since the $T_{\tau_s^i}(u)$'s do not share any variables (except the root) we have

$$\begin{aligned}
\partial_{u_V}\overline{C}(u) &= \partial_{u_i,u_{V\setminus i}}\left[\prod_{s\in N(i)} T_{\tau_s^i}(u)\right] \\
&= \partial_{u_i}\left[\prod_{s\in N(i)} \partial_{u_{\tau_s^i\setminus i}}\left[T_{\tau_s^i}(u)\right]\right] \\
&= \partial_{u_i}\left[\prod_{s\in N(i)} \mu_{s\to i}(u)\right]. \qquad (16)
\end{aligned}$$

The quantity $\mu_{s\to i}(u) := \partial_{u_{\tau_s^i\setminus i}}\left[T_{\tau_s^i}(u)\right]$ is called a message from the function indexed by s to the variable indexed by i. Now consider $T_{\tau_s^i}(u)$ and let j the neighbor variable index of the function index s which is not i. One can go deeper in the tree, that is, we have

$$T_{\tau_s^i}(u) = \Phi_s(u_i, u_j) T_{\tau_j^s}(u)$$

where τ_j^s is the subtree rooted at the function indexed by s and containing the variable indexed by j (see Figure 2). Hence,

$$\begin{aligned}
\partial_{u_{\tau_s^i\setminus i}} T_{\tau_s^i}(u) &= \partial_{u_j}\left[\phi_s(u_i,u_j)\partial_{u_{\tau_j^s\setminus j}}\left[T_{\tau_j^s}(u)\right]\right] \\
&= \partial_{u_j}\left[\phi_s(u_i,u_j)\mu_{j\to s}(u)\right].
\end{aligned}$$

A second type of message has been defined: $\mu_{j\to s}(u) := \partial_{u_{\tau_j^s\setminus j}}\left[T_{\tau_j^s}(u)\right]$ is called a message from the variable index j to the function index s. Again,

$$T_{\tau_j^s}(u) = \prod_{s'\in N(j)\setminus s} T_{\tau_{s'}^j}(u),$$

hence

$$\partial_{u_{\tau_j^s \setminus j}} T_{\tau_j^s}(u) = \prod_{s' \in N(j)\setminus s} \partial_{u_{\tau_{s'}^j \setminus j}} T_{\tau_{s'}^j}(u) = \prod_{s' \in N(j)\setminus s} \mu_{s' \to j}(u).$$

The message $\mu_{s' \to j}(u)$ has been already defined in (16). This recursive formulation allows to compute recursively all the messages from the leaves to the root. Once all the messages have been computed, the density is given by the derivative with respect to the root of the product of all the messages (16). The likelihood, in turn, can be computed.

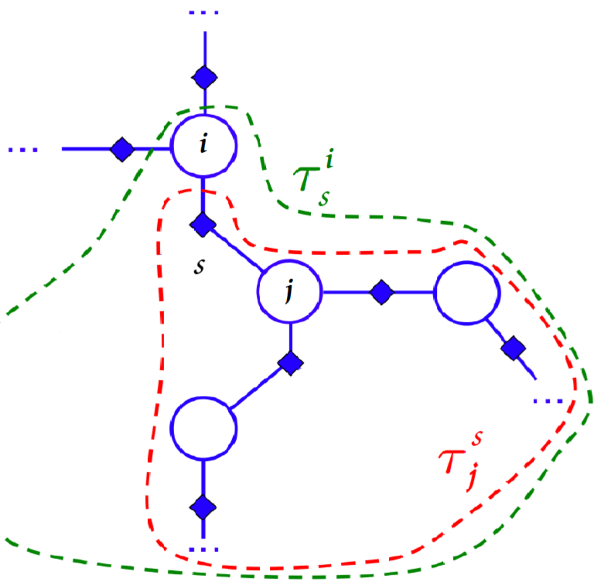

Figure 2: Examples of subtrees. This figure is partly drawn from [10].

A Appendix

Proof of Theorem 2

From Theorem 1, it is straightforward to see that (6) is a copula. Let us now prove that (6) is the only copula arising from (5). Because of condition (i), for $i = 1, \ldots, d$, we have $g_{si} = 1$ if $s \notin N(i)$ and the constraint over the functions reduces to $\prod_{s \in N(i)} g_{si}(v) = v$, $v \in [0,1]$. In view of condition (ii), one has $g_{si} = g_i$ for $s \in N(i)$, hence $(g_i(v))^{n_i} = v$. Therefore

$$g_{si}(v) = \begin{cases} v^{1/n_i} & \text{if } s \in N(i), \\ 1 & \text{otherwise.} \end{cases}$$

To conclude it suffices to rewrite the product in (5) as

$$\prod_{s \in S} C_s(1, \ldots, 1, u_i^{1/n_i}, 1, \ldots, 1, u_j^{1/n_j}, 1, \ldots, 1) = \prod_{\{ij\} \in S} C_{ij}(u_i^{1/n_i}, u_j^{1/n_j})$$

which corresponds to (6).

Proof of Proposition 2

Let us first prove that (10) is the distribution function of (6). By (1) we have

$$\begin{aligned} F(x_1,\dots,x_d) =& \overline{C}(F_1(x_1),\dots,F_d(x_d)) \\ =& \prod_{\{ij\}\in S} C_{ij}(F_i(x_i)^{1/n_i}, F_j(x_j)^{1/n_j}) \\ =:& \prod_{\{ij\}\in S} \Phi_{ij}(x_i,x_j). \end{aligned}$$

The first margin of Φ_{ij} is given by $\Phi_{ij,1}(x) = \Phi_{ij}(x,\infty) = F_i(x_i)^{1/n_i}$ which depends only on i. The same holds for the second margin $\Phi_{ij,2}$.

Let us prove that (6) is the copula associated to (10). Let $\Phi_{ij,k}$, $k = 1,2$ the k-th univariate marginal of $\Phi_{ij}, \{ij\} \in S$. The copula associated to F is given by

$$\begin{aligned} C_F(u_1,\dots,u_d) =& F\left(F_1^{-1}(u_1),\dots,F_d^{-1}(u_d)\right) \\ =& \prod_{\{ij\}\in S} \Phi_{ij}\left(F_i^{-1}(u_i), F_j^{-1}(u_j)\right). \end{aligned}$$

Let C_{ij} be the copula associated to Φ_{ij}. We have

$$\Phi_{ij}(x_i,x_j) = C_{ij}\left(\Phi_{ij,1}(x_i), \Phi_{ij,2}(x_j)\right)$$

so that $\Phi_{ij}\left(F_i^{-1}(u_i), F_j^{-1}(u_j)\right) = C_{ij}\left(\Phi_{ij,1}\circ F_i^{-1}(u_i), \Phi_{ij,2}\circ F_j^{-1}(u_j)\right)$. Thus

$$C_F(u_1,\dots,u_d) = \prod_{\{ij\}\in S} C_{ij}\left(\Phi_{ij,1}\circ F_i^{-1}(u_i), \Phi_{ij,2}\circ F_j^{-1}(u_j)\right). \qquad (17)$$

Moreover, since C_F is a copula we have that

$$\begin{aligned} u_k =& C_F(1,\dots,1,u_k,1,\dots,1) \\ =& \prod_{j>k:\{kj\}\in S} C_{kj}\left(\Phi_{kj,1}\circ F_k^{-1}(u_k),1\right) \prod_{j<k:\{jk\}\in S} C_{jk}\left(1, \Phi_{jk,2}\circ F_k^{-1}(u_k)\right) \\ =& \prod_{j:\{kj\}\in S} \Phi_{kj,1}\circ F_k^{-1}(u_k). \end{aligned}$$

Now by assumption $\Phi_{kj,1} = \Phi_{jk,2} = \Phi_k$ only depends on k and therefore $u_k^{1/n_k} = \Phi_k \circ F_k^{-1}(u_k)$ which implies $\Phi_k(z) = F_k(z)^{1/n_k}$, $z \in \mathbf{R}$. By pluging Φ_k into (17) the result follows.

Proof of Proposition 3

If $\{kl\} \in S$, then

$$\begin{aligned} \overline{C}_{kl}(u_k,u_l) =& \overline{C}(1,\dots,1,u_k,1,\dots,1,u_l,1,\dots,1) \\ =& \left(\prod_{s\in N(k)\setminus\{kl\}} C_s(u_k^{1/n_k},1)\right)\left(\prod_{s\in N(l)\setminus\{kl\}} C_s(u_l^{1/n_l},1)\right) \times \\ & C_{kl}(u_k^{1/n_k}, u_l^{1/n_l}) \\ =& u_k^{(n_k-1)/n_k} u_l^{(n_l-1)/n_l} C_{kl}(u_k^{1/n_k}, u_l^{1/n_l}). \end{aligned}$$

Otherwise,

$$\begin{aligned}\overline{C}_{kl}(u_k,u_l) &= \left(\prod_{s\in N(k)} C_s(u_k^{1/n_k},1)\right)\left(\prod_{s\in N(l)} C_s(u_l^{1/n_l},1)\right)\\ &= u_k^{n_k/n_k}u_l^{n_l/n_l}\\ &= u_k u_l.\end{aligned}$$

Proof of Proposition 5

By applying the Fréchet-Hoeffding bounds for copulas (see e.g. [15], p. 11) we obtain:

$$W_{kl}(u_k,u_l) \le \overline{C}_{kl}(u_k,u_l) \le M_{kl}(u_k,u_l), \tag{18}$$

where

$$\begin{aligned}W_{kl}(u_k,u_l) &= u_k^{1-1/n_k}u_l^{1-1/n_l}\max(u_k^{1/n_k}+u_l^{1/n_l}-1,0),\\ M_{kl}(u_k,u_l) &= u_k^{1-1/n_k}u_l^{1-1/n_l}\min(u_k^{1/n_k},u_l^{1/n_l}).\end{aligned}$$

We have $M_{kl}(u,u)/u \to 0$ as $u \downarrow 0$. It is easily seen that $W_{kl}(u,u)/u \to 0$ as $u \downarrow 0$ which implies $\overline{C}_{kl}(u,u)/u \to 0$. It is straightforward to see that $(1-2u+M_{kl}(u,u))/(1-u) \to 1/\max(n_k,n_l)$ as $u \uparrow 1$. To compute the lower and upper bounds for $\overline{\rho}_{kl}$ and $\overline{\tau}_{kl}$, it suffices to substitute W_{kl} and M_{kl} into (14) and (13). Lengthy but elementary computations lead to the results. Finally, letting n_k or n_l going to infinity in (18) yields that $\overline{C}_{kl}$ tends to independence.

References

[1] K. Aas, C. Czado, A. Frigessi, and H. Bakken. Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, 44(2):182–198, 2009.

[2] L.B.G Andersen and V.V. Piterbarg. *Interest Rate Modeling.* Atlantic Financial Press, 2010.

[3] A. C. Davison, S. A. Padoan, and M. Ribatet. Statistical modeling of spatial extremes. *Statistical Science*, 27(2):161–186, 2012.

[4] S. Demarta and A. J. McNeil. The t copula and related copulas. *International Statistical Review*, 73(1):111–129, 2005.

[5] F. Durante and G. Salvadori. On the construction of multivariate extreme value models via copulas. *Environmetrics*, 21(2):143–161, 2010.

[6] G. Frahm, M. Junker, and A. Szimayer. Elliptical copulas: applicability and limitations. *Statistics & Probability Letters*, 63(3):275–286, 2003.

[7] C. Genest, K. Ghoudi, and L-P. Rivest. Understanding relationships using copulas. *North American Actuarial Journal*, 2(3):143–149, 1998.

[8] G. Gudendorf and J. Segers. Extreme-value copulas. In *Copula Theory and Its Applications*, page 127–145. Springer, 2010.

[9] M. Hofert, M. Mächler, and A. J. McNeil. Archimedean copulas in high dimensions: Estimators and numerical challenges motivated by financial applications. *Journal de la Société Française de Statistique*, 154(1):25–63, 2012.

[10] J. C. Huang. *Cumulative distribution networks: Inference, estimation and applications of graphical models for cumulative distribution functions*. PhD thesis, University of Toronto, 2009.

[11] J.C. Huang and N. Jojic. Maximum-likelihood learning of cumulative distribution functions on graphs. *Journal of Machine Learning Research W&CP Series*, 9:342–349, 2010.

[12] H. Joe. *Multivariate models and dependence concepts*. Chapman & Hall/CRC, 2001.

[13] D. Kim and J-M. Kim. Analysis of directional dependence using asymmetric copula-based regression models. *Journal of Statistical Computation and Simulation*, published online, 2013.

[14] E. Liebscher. Construction of asymmetric multivariate copulas. *Journal of Multivariate Analysis*, 99(10):2234–2250, 2008.

[15] R.B. Nelsen. *An introduction to copulas*. Springer, 2006.

[16] M. Sklar. *Fonctions de répartition à n dimensions et leurs marges*. Publications de l'Institut de Statistique de l'Université de Paris, 8:229-231, 1959.