A class of multivariate copulas based on products of bivariate copulas

Gildas Mazo, Stephane Girard, Florence Forbes

A class of multivariate copulas based on products of bivariate copulas

Gildas Mazo(⋆), Stéphane Girard and Florence Forbes

Inria and Laboratoire Jean Kuntzmann, Grenoble, France
(⋆)gildas.mazo@free.fr

Abstract

Copulas are a useful tool to model multivariate distributions. While there exist various families of bivariate copulas, much less work has been done when the dimension is higher. We propose a class of multivariate copulas based on products of transformed bivariate copulas. The analytical forms of the copulas within this class allow to naturally associate a graphical structure which helps to visualize the dependencies and to compute the full joint likelihood even in high dimension. Numerical experiments are conducted both on simulated and real data thanks to a dedicated `R` package.

Keywords: maximum-likelihood inference, graphical models, message-passing algorithm, multivariate, copula.

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 copula. 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, $i = 1, \dots, d$. From Sklar's Theorem [26], there exists a unique function C such that

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)), \quad (x_1, \dots, x_d) \in \mathbb{R}^d. \tag{1}$$

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. [24]. 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 [13], Vines [1] or elliptical copulas [5].

Archimedean copulas write

$$C(u_1, \dots, u_d) = \psi(\psi^{-1}(u_1) + \cdots + \psi^{-1}(u_d)),$$

where ψ is a function from $[0, \infty)$ to $[0, 1]$ which has to verify certain properties for the copula to be well defined, see [23]. The generator ψ may be chosen in a given

parametric family of functions. For instance, $\psi_\theta(t) = \exp(-t^{1/\theta})$, $\theta \geq 1$ yields the Gumbel family of copulas, see Example 1 in Section 3. Since there is a single parameter to model a d-dimensional phenomenon, this model is recognised not to be very flexible. Indeed, Archimedean copulas are exchangeable i.e. $C(u_1, \ldots, u_d) = C(u_{\pi(1)}, \ldots, u_{\pi(d)})$ for any permutation π of $\{1, \ldots, d\}$. In particular, all pairs of variables share the same statistical distribution. These properties may not be desirable in practice.

Vines, on the opposite, achieve greater flexibility but at the price of increased complexity. As an illustration, we briefly describe a canonical vine copula – one of the two main types of vine copula models – through a decomposition of its density [1]:

$$c(u_1, \ldots, u_d) = \prod_{j=1}^{d-1} \prod_{i=1}^{d-j} c_{j,j+1|1,\ldots,j-1} \left(F(u_j|u_1, \ldots, u_{j-1}), F(u_{j+i}|u_1, \ldots, u_{j-1}) \right)$$

where $c_{j,j+1|1,\ldots,j-1}(\cdot,\cdot)$ is the (conditional) bivariate density of the j-th and $(j+1)$-th variables and where $F(\cdot|\cdot)$ represents the conditional distribution of the variables at play. When $d = 10$, there are more than one million possible decompositions, and, for each decomposition, there are many choices of parametric families for each conditional bivariate density in the product.

A third class of copulas to be presented in this introduction is the class of elliptical copulas. An elliptical copula is the copula of an elliptical distribution, whose density is given by [5, 22]

$$f(x) = |\Sigma|^{-1/2} g\left((x-\mu)^\top \Sigma^{-1} (x-\mu) \right), \quad x \in \mathbb{R}^d,$$

for some positive definite matrix Σ and vector μ. The function g is called the density generator. This model implies, in particular, that if X has density f as above, then $X - \mu$ is distributed as $\mu - X$. This, in turn, implies that the lower and upper tail dependence coefficients (defined in Section 3) are equal, which is unrealistic in some applications, as, for example, extreme-value statistics. Moreover, elliptical copulas have in general as many as $O(d^2)$ parameters and it is thus difficult to carry out maximum likelihood inference [3] when d is large.

The main contribution of this paper is to propose a new class of multivariate copulas based on a product of bivariate copulas. The product is performed following the edges of a graph which permits to visualize the dependencies and to efficiently compute the likelihood, even in high dimension. The use of bivariate copulas as building blocks allows to take profit of the numerous parametric families proposed in the copula literature.

The rest of this paper is organized as follows. The new copula model is introduced in Section 2. Some links with Liebscher's construction [19] are stressed. Section 3 discusses some properties of the new copulas. The ability to construct new extreme-value models is highlighted. The dependence properties of bivariate marginals of the proposed class are also established. More specifically, some bounds are given on the most popular dependence coefficients (Spearman's rho and Kendall's tau) and on tail dependence coefficients. Section 4 is dedicated to the numerical aspects. A simulation procedure is provided and estimation by maximization of the pseudo-likelihood is discussed. The proposed copula model is applied in Section 5 to simulated and real datasets. The appendix gathers some proofs and technical details about the estimation procedure.

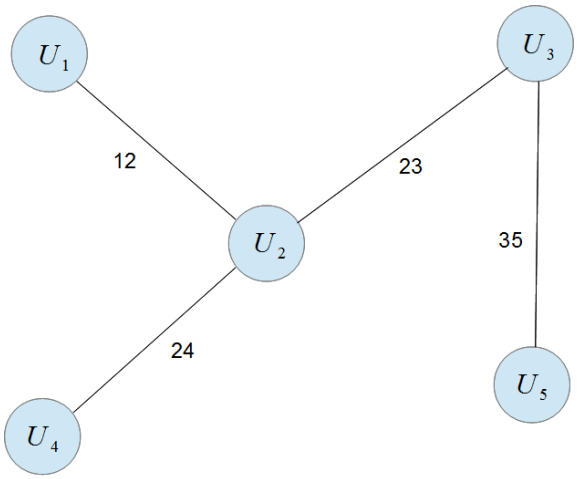


Figure 1: Graphical representation of the set $E = \{\{12\}, \{24\}, \{23\}, \{35\}\}$. The edges which are neighbors of each node are $N(1) = \{\{12\}\}$, $N(2) = \{\{12\}, \{23\}, \{24\}\}$, $N(3) = \{\{23\}, \{35\}\}$, $N(4) = \{\{24\}\}$ and $N(5) = \{\{35\}\}$. The numbers of neighbors are respectively given by $n_1 = 1$, $n_2 = 3$, $n_3 = 2$, $n_4 = 1$ and $n_5 = 1$.

2 Constructing high dimensional copulas by multiplying bivariate ones

In this section, we propose a way to build high-dimensional copulas starting from bivariate ones. This construction allows one to take advantage of the large number of bivariate copulas introduced in the statistical literature. It is well known that a product of copulas is not a copula in general, the margins being no longer uniform. Roughly speaking, the new copula is thus obtained by multiplying bivariate copulas after a suitable transformation of the margins. The main feature of the new copula is that it can be associated with a graph describing the dependencies between the variables. To be more specific, let $U_1, \ldots, U_d$ be d standard uniform random variables and denote by $\{ij\}$ the index of the pair (U_i, U_j). Introduce $E \subset \{\{ij\} : i, j = 1, \ldots, d,\ j > i\}$ a subset of the set of all pair indices. The cardinal of E, denoted by $|E|$, is less or equal to $d(d-1)/2$. The pair index $e \in E$ is said to contain the variable index i if there exists $k \neq i$ such that $e = \{ik\}$ or $e = \{ki\}$. For all $i = 1, \ldots, d$, let $N(i)$ be the set of neighbors of i defined as $N(i) = \{e \in E \text{ such that } e \text{ contains } i\}$ and introduce $n_i := |N(i)|$. It is then natural to associate a graph to the set E as follows: An element $e = \{ij\} \in E$ is an edge linking U_i and U_j in the graph whose nodes are the variables $U_1, \ldots, U_d$. The example $E = \{\{12\}, \{24\}, \{23\}, \{35\}\}$ is illustrated in Figure 1. For $u = (u_1, \ldots, u_d) \in [0,1]^d$, consider the functional

$$C(u_1, \ldots, u_d) = \prod_{\{ij\} \in E} \tilde{C}_{ij}\left(u_i^{1/n_i}, u_j^{1/n_j}\right), \tag{2}$$

where the $\tilde{C}_{ij}$'s are arbitrary bivariate copulas for all $\{ij\} \in E$. Keeping in mind the graphical representation associated with E, the function C defined in (2) is a product

over the edges of the graph. For instance, when $E = \{\{12\}, \{24\}, \{23\}, \{35\}\}$ as in Figure 1, function (2) can be written as

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

In the following, (2) is referred to as the Product of Bivariate Copulas (PBC) copula, or PBC model. The next result establishes that (2) is a copula.

Proposition 1. *PBC* (2) *is a well defined copula.*

The proof is postponed to the Appendix. It is based on the remark that (2) can be interpreted as a particular case of Liebscher's construction [19]:

$$C(u_1, \ldots, u_d) = \prod_{e \in E'} \tilde{C}_e \left(g_{e1}(u_1), \ldots, g_{ed}(u_d)\right). \tag{3}$$

where E' is some finite set. In the general case, $\tilde{C}_e$ is a d-dimensional copula for all $e \in E'$ and g_{ei} is a univariate function for $i = 1, \ldots, d$ and $e \in E'$ which is either strictly increasing or identically equal to 1. Moreover, to ensure that (3) is a well defined copula, the following constraints have to be verified:

$$\prod_{e \in E'} g_{ei}(v) = v \text{ for all } v \in [0, 1] \text{ and } i = 1, \ldots, d. \tag{4}$$

Let us highlight that Liebscher's construction (3) was originally designed to build an asymmetric d-dimensional C starting from a family of symmetric d-dimensional copulas $\tilde{C}_e$, $e \in E'$. In contrast, our re-interpretation of (2) permits to derive d-dimensional copulas starting from bivariate ones $\tilde{C}_{ij}$, $\{ij\} \in E$ and a graph. Another major difference between Liebscher and PBC copulas lies in the tractability of the models. Indeed, Liebscher's model is rather an abstract construction principle and the applicability to data has not been addressed so far. For instance, taking $g_{ei}(v) = v^{\theta_{ei}}$ for $v \in (0, 1]$ and $0 \leq \theta_{ei} \leq 1$, the parameters have to verify $\sum_{e \in E'} \theta_{ei} = 1$ for all $i = 1, \ldots, d$. How to deal with these constraints in the estimation procedure? As a matter of fact, the original paper [19] does not provide any such procedure. In contrast, one can view a PBC copula (2) as a particular case of Liebscher's construction where the constraints are met by construction. The remaining parameters would be those of the $\tilde{C}_{ij}$, but since they are arbitrary, we are left with $|E|$ unconstrained parameter vectors (one for each $\tilde{C}_{ij}$). The use of bivariate copulas as well as the graph structure in the PBC model open the way to an exact computation of the (pseudo-)likelihood. If the graph is a tree, an algorithm based on message passing can then be used to perform the maximum (pseudo-)likelihood estimation of the copula parameters, see Section 4.

The PBC model is a legitimate copula in the sense that it is the only possible model derived from (3) when a few natural assumptions are made about the copula:

Proposition 2. *Suppose that, in* (3)*:*

(i) For all $e \in E'$, $\tilde{C}_e$ takes as arguments only two functions non identically equal to 1,

(ii) For all $i = 1, \ldots, d$ and $e \in E'$, g_{ei} does not depend on e,

then, the only copula which can be constructed from (3) *is the PBC model* (2)*. In particular, necessarily $g_{ei}(v) = v^{1/n_i}$ if $e \in N(i)$ and $g_{ei}(v) = 1$ otherwise.*

Finally, from Sklar's Theorem, the PBC copula (2) is associated with a distribution function F with continuous marginals F_i, $i = 1, \ldots, d$. By substituting (2) into (1), it is easily seen that F writes as a product of bivariate distribution functions:

$$F(x_1, \ldots, x_d) = \prod_{\{ij\} \in E} F_{ij}(x_i, x_j), \quad (x_1, \ldots, x_d) \in \mathbb{R}^d, \tag{5}$$

where F_{ij} is a bivariate distribution function whose first marginal $F_{ij,1}$ only depends on i and second marginal $F_{ij,2}$ only depends on j for all $\{ij\} \in E$. It is interesting to note that the converse is also true as stated in the following proposition.

Proposition 3. *The distribution function corresponding to the PBC copula* (2) *writes as* F *in* (5)*. Conversely, the copula corresponding to any distribution function* F *in* (5) *writes as the PBC copula* (2)*.*

Outside the copula framework, distribution functions which can write as (5) have been studied in the context of graphical modeling in [14], where the authors refer to them as a Cumulative Distribution Networks (CDN). However, the dependence properties were not investigated. Since the PBC copula is a copula associated with a CDN, the study of the dependence properties in Section 3 bridges this gap.

3 Dependence properties and max-stability

The first paragraph is dedicated to the construction of extreme-value copulas using PBC models while the second paragraph presents some dependence properties of the bivariate margins of PBC models.

3.1 Extreme-value copulas

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

$$C_\#(u_1, \ldots, u_d) = \lim_{n \uparrow \infty} C^n(u_1^{1/n}, \ldots, u_d^{1/n}), \tag{6}$$

for all $(u_1, \ldots, u_d) \in [0,1]^d$. In such a case, C is said to be in the maximum domain of attraction of $C_\#$. A copula $C_\#$ is said to be max-stable if for all integer $n \geq 1$ and $(u_1, \ldots, u_d) \in [0,1]^d$

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

Extreme-value copulas exactly correspond to max-stable copulas [11]. It can be seen that the domain of attraction condition (6) is equivalent to

$$\lim_{t \to 0} \frac{1 - C(1 - tx_1, \ldots, 1 - tx_d)}{t} = \ell(x_1, \ldots x_d)$$

where ℓ is the tail dependence function given by $\ell(x_1, \ldots, x_d) = -\log C_\#(e^{-x_1}, \ldots, e^{-x_d})$ for all $(x_1, \ldots, x_d) \in [0, \infty)^d$. It can be shown that the tail dependence function is convex and homogeneous of order one [11]. Extreme-value copulas belonging to the PBC class (2) can easily be constructed thanks to the next result. Its proof is a direct consequence of [4].

Proposition 4. *If, in the PBC copula* (2)*, $\tilde{C}_{ij}$ is an extreme-value copula for all $\{ij\} \in E$, then C is also an extreme-value copula. The associated tail dependence function is given by*

$$\ell(x_1, \ldots, x_d) = \sum_{\{ij\} \in E} \tilde{\ell}_{ij}\left(x_i^{1/n_i}, x_j^{1/n_j}\right)$$

for all $(x_1, \ldots, x_d) \in [0, \infty)^d$, and where $\tilde{\ell}_{ij}$ is the stable tail dependence function of $\tilde{C}_{ij}$.

It thus appears that the PBC model inherits the max-stability from the pairwise copulas used in its construction. The associated tail dependence function has moreover a simple additive and yet flexible form. This is illustrated on the Gumbel copula case, which has been shown to be the only max-stable Archimedean copula [10].

Example 1. *Let $\tilde{C}_{ij}$ in* (2) *be a (max-stable) Gumbel copula with parameter $\theta_{ij} \geq 1$, that is,*

$$\tilde{C}_{ij}(u_i, u_j) = \exp\left\{-\left[(-\log u_i)^{\theta_{ij}} + (-\log u_j)^{\theta_{ij}}\right]^{1/\theta_{ij}}\right\}.$$

Then, the associated PBC model is an extreme-value copula with tail dependence function

$$\ell(x_1, \ldots, x_d) = \sum_{\{ij\} \in E} \left(x_i^{\theta_{ij}/n_i} + x_j^{\theta_{ij}/n_j}\right)^{1/\theta_{ij}}. \tag{7}$$

As a comparison, the tail dependence function associated with a d-dimensional Gumbel copula with parameter $\theta \geq 1$ is

$$\ell(x_1, \ldots, x_d) = \left(\sum_{i=1}^{d} x_i^{\theta}\right)^{1/\theta}. \tag{8}$$

It is clear that (7) offers much more flexibility than (8). The choice of E tunes the weights assigned to each variable through the number of neighbors. Besides, the dependence parameter θ_{ij} may be different for each pair $\{ij\}$. Other examples are provided in the next paragraph.

3.2 Bivariate margins properties

In this paragraph, the dependence properties of the bivariate margins of a PBC copula C defined by (2) are established. To this end, for all $(k, \ell) \in \{1, \ldots, d\}^2$, let $C_{k\ell}$ be the copula associated with the random pair (U_k, U_ℓ) and given by $C_{k\ell}(u_k, u_\ell) = C(1, \ldots, 1, u_k, 1, \ldots, 1, u_\ell, 1, \ldots, 1)$. Our first result shows that $C_{k\ell}$ is closely related to the bivariate copula $\tilde{C}_{k\ell}$ involved in the construction (2).

Proposition 5. *The bivariate marginal $C_{k\ell}$ of the PBC copula* (2) *is given by*

$$C_{k\ell}(u_k, u_l) = \begin{cases} u_k^{(n_k-1)/n_k} u_\ell^{(n_\ell-1)/n_\ell} \tilde{C}_{k\ell}(u_k^{1/n_k}, u_\ell^{1/n_\ell}) & \text{if } \{k\ell\} \in E, \\ u_k u_\ell & \text{otherwise,} \end{cases} \tag{9}$$

for all $(k, \ell) \in \{1, \ldots, d\}^2$.

It thus appears that random pairs (U_k, U_ℓ) which are not connected by an edge i.e. $\{k\ell\} \notin E$ are independent. The dependence between connected random variables is mainly driven by $\tilde{C}_{k\ell}$ through the equation

$$C_{k\ell}(u_k, u_\ell) = u_k^{1-\kappa} u_\ell^{1-\lambda} \tilde{C}_{k\ell}(u_k^\kappa, u_\ell^\lambda), \tag{10}$$

where $\kappa = 1/n_k$ and $\lambda = 1/n_\ell$. Let us highlight that (10) is sometimes referred to as Khoudraji's device; see [7], Proposition 2. As an immediate consequence of Proposition 5, a PBC copula is not exchangeable in the general case. Recall that in the PBC construction (2), the choice for the $\tilde{C}_{k\ell}$ is quite arbitrary. Thus, since the dependence structure of $C_{k\ell}$ is determined by that of $\tilde{C}_{k\ell}$, one can get any desired dependence structure for the pairs of variables that are connected by an edge in the PBC model. Some examples are provided below.

Example 2. *Let $\tilde{C}_{k\ell}$ in (10) be a Marshall-Olkin copula MO(α, β) with parameters $(\alpha, \beta) \in [0,1]^2$:*

$$\tilde{C}_{k\ell}(u_k, u_\ell) = \min(u_k^{1-\alpha} u_\ell, u_\ell^{1-\beta} u_k),$$

see for instance [24], p. 53. Then, $C_{k\ell}$ is MO$(\alpha\kappa, \beta\lambda)$. If, moreoever, $\alpha = \beta$ then $\tilde{C}_{k\ell}$ is a Cuadras-Augé copula and $C_{k\ell}$ is MO$(\alpha\kappa, \alpha\lambda)$. If $\alpha = \beta = 0$ then both $\tilde{C}_{k\ell}$ and $C_{k\ell}$ are the independence copula. If $\alpha = \beta = 1$ then $\tilde{C}_{k\ell}$ is the Fréchet upper bound copula and $C_{k\ell}$ is MO(κ, λ).

The Marshall-Olkin class of copulas is thus stable with respect to the transformation (9). Besides, from Proposition 4, and since the Marshall-Olkin copula is max-stable, it follows that all the PBC models associated with Example 2 are max-stable copulas.

Let us now investigate the dependence properties associated with the bivariate margins of the PBC model. Recall that the dependence between U_k and U_ℓ is positive if, roughly speaking, U_k and U_ℓ tend to be large or small together. Below are recalled a few definitions of statistical concepts about positive dependence. The copula $C_{k\ell}$ associated with (U_k, U_ℓ) has the TP2 (totally positive of order 2) property if and only if

$$C_{k\ell}(u_1, u_2) C_{k\ell}(v_1, v_2) \geq C_{k\ell}(u_1, v_2) C_{k\ell}(v_1, u_2), \text{ for all } u_1 < v_1 \text{ and } u_2 < v_2.$$

Also, $C_{k\ell}$ is said to be PQD (positive quadrant dependent) if $C_{k\ell}(u, v) \geq uv$ for all $(u, v) \in [0,1]^2$. The random variable U_ℓ is said to be LTD (left tail decreasing) in U_k if for all $v \in [0,1]$, the function $u \mapsto \mathbb{P}(U_\ell \leq v | U_k \leq u)$ is decreasing in u. See [24] and [16] for further details about these concepts. Since the PBC copula (2) is a particular form of the Liebscher copula (3) (as shown in Section 2), one can apply the results of [19] to obtain properties for the bivariate margins of the PBC copula.

Proposition 6. *If in (9) $\tilde{C}_{k\ell}$ is TP2, LTD or PQD then $C_{k\ell}$ is also TP2, LTD or PQD respectively.*

The dependence between U_k and U_ℓ can be quantified through dependence measures such as the Kendall's tau or the Spearman's rho respectively given by

$$\tau_{k\ell} = 4 \int_{[0,1]^2} C_{k\ell}(u, v)\, dC_{k\ell}(u, v) - 1, \tag{11}$$

$$\rho_{k\ell} = 12 \int_{[0,1]^2} C_{k\ell}(u, v)\, du\, dv - 3. \tag{12}$$

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

$$\lambda_{k\ell}^{(U)} = \lim_{u\uparrow 1} \frac{1-2u+C_{k\ell}(u,u)}{1-u} \in [0,1], \quad \lambda_{k\ell}^{(L)} = \lim_{u\downarrow 0} \frac{C_{k\ell}(u,u)}{u} \in [0,1]. \tag{13}$$

Liebscher's erratum [20] provides conditions on copulas $\tilde{C}_{k\ell}$ under which Kendall's tau or Spearman's rho are increased or decreased by the transformation (9). Here, we give explicit bounds in terms of the number of neighbors for these dependence coefficients. The behavior of (9) when the number of neighbors tends to infinity is also investigated.

Proposition 7. *We have* $\lambda_{k\ell}^{(L)} = 0$ *and* $\lambda_{k\ell}^{(U)} \leq \min(1/n_k, 1/n_\ell)$. *The lower and upper bounds for* $\rho_{k\ell}$ *and* $\tau_{k\ell}$ *are respectively given by*

$$a_\rho(n_k,n_\ell) \leq \rho_{k\ell} \leq b_\rho(n_k,n_\ell),$$
$$a_\tau(n_k,n_\ell) \leq \tau_{k\ell} \leq b_\tau(n_k,n_\ell),$$

with

$$a_\rho(n_k,n_\ell) = \frac{6\beta(2n_k-1,2n_\ell-1)n_k n_\ell}{(2n_k+2n_\ell-1)(n_k+n_\ell-1)} - \frac{3}{(2n_k-1)(2n_\ell-1)},$$
$$b_\rho(n_k,n_\ell) = \frac{3}{2n_k+2n_\ell-1},$$
$$a_\tau(n_k,n_\ell) = \frac{\beta(2n_\ell-1,2n_k-1)}{n_k+n_\ell-1} - \frac{2}{(2n_k-1)(2n_\ell-1)},$$
$$b_\tau(n_k,n_\ell) = \frac{1}{n_k+n_\ell-1},$$

where β *denotes the Beta-function defined by* $\beta(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}dt$ *for* $x>0$ *and* $y>0$. *Furthermore,* $C_{k\ell}(u,v) \to uv$ *as* $\max(n_k,n_\ell) \to \infty$, *for all* $(u,v) \in [0,1]^2$.

To illustrate Proposition 7, numerical values of the bounds are computed in Table 1 for different numbers of neighbors (n_k, n_ℓ).

coefficient / (n_k,n_ℓ)	$\rho_{k\ell}$	$\tau_{k\ell}$	$\lambda_{k\ell}$
$(1,2)$	$[-0.60, 0.60]$	$[-0.50, 0.50]$	$[0.00, 0.50]$
$(2,2)$	$[-0.30, 0.43]$	$[-0.21, 0.33]$	$[0.00, 0.50]$
$(1,3)$	$[-0.43, 0.43]$	$[-0.33, 0.33]$	$[0.00, 0.33]$
$(2,3)$	$[-0.19, 0.33]$	$[-0.13, 0.25]$	$[0.00, 0.33]$
$(3,3)$	$[-0.12, 0.27]$	$[-0.08, 0.20]$	$[0.00, 0.33]$

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

The above results show that we are facing a tradeoff when linking the variables in the graph: on the one hand, the larger the cardinal of E (that is, the more connected the graph, or the more the number of neighbors), the closer to independence the bivariate margins belonging to E are. On the other hand, the smaller the cardinal of E, the more there are independent bivariate margins. These findings suggest also that, unfortunately, while the PBC model may well be adequate to model weakly dependent data, it might not be able to model strongly dependent data.

4 Simulation and estimation

The first subsection gives an algorithm that allows to easily simulate from a PBC copula. The second subsection deals with inference for PBC copulas.

4.1 Simulation

One can use the probabilistic interpretation provided in [19] to simulate from a PBC. The generation procedure is given below.

- For all $\{ij\} \in E$, generate $(U_i^{(ij)}, U_j^{(ij)}) \sim \tilde{C}_{ij}$.
- For all $i = 1, \ldots, d$, compute $U_i = \max_{k \in \{1,\ldots,d\}:\{ki\} \in E} \left\{ \left(U_i^{(ki)} \right)^{n_i} \right\}$.

The resulting vector $(U_1, \ldots, U_d)$ has distribution (2).

4.2 Estimation

Here, we assume that the copulas $\tilde{C}_{ij}$ involved in the construction (2) of the PBC copula C depend on parameters θ_{ij}. The parameter vector is denoted by $\theta = (\theta_{ij})_{\{ij\} \in E}$. The sample of i.i.d. multivariate observations from which θ is to be estimated is $(X_1^{(m)}, \ldots, X_d^{(m)})$ for $m = 1, \ldots, n$, with $(X_1^{(1)}, \ldots, X_d^{(1)}) \sim F$, where F is the cumulative distribution function associated with C. The margins of F are denoted by $F_1, \ldots, F_d$.

In order to model data with a PBC model, one has to assume that there exists a graphical structure underpinning the variables of interest. Some pairs of variables are linked with edges, some are not. In this section, the graph underpinning C is supposed to be known. In other words, we suppose that the set of edges linking the variables of interest is given. Needless to say, this is not the case in practice. While choosing the most appropriate graphical structure is still an open problem in this context, some hints are given in Section 6.

The first estimation procedure considered is a version of the pairwise maximum-likelihood method [21]. This approach consists of maximizing the sum of the pseudo-likelihoods corresponding to all the pairs of variables:

$$\arg\max_{\theta} \sum_{m=1}^{n} \sum_{\{k\ell\}} \log c_{k\ell} \left(\widehat{F}_k(x_k^{(m)}), \widehat{F}_\ell(x_\ell^{(m)}); \theta_{k\ell} \right), \tag{14}$$

where $c_{k\ell}$ is the bivariate density associated with $C_{k\ell}$ and given by

$$c_{k\ell}(u_k, u_\ell; \theta_{k\ell}) = \frac{\partial^2 C_{k\ell}(u_k, u_\ell; \theta_{k\ell})}{\partial u_k \partial u_\ell} = \frac{\partial^2 \left[\tilde{C}_{kl}(u_k^{1/n_k}, u_\ell^{1/n_\ell}) u_k^{1-1/n_k} u_\ell^{1-1/n_\ell} \right]}{\partial u_k \partial u_\ell},$$

see Proposition 5. The pseudo-likelihood term refers to the plugging in (14) of an estimator $\widehat{F}_k$ of the unknown margin distribution F_k. Note that this estimator can be either parametric as in [16], Section 10 or nonparametric as in [6]. In the PBC context, the pairwise approach reduces to maximizing $|E|$ univariate functions independently. Thus, this method has the advantage to be easily implementable. Moreover, it allows to fit different parametric families for different pairs.

However, an estimator based on such a pairwise strategy is not as efficient as an estimator based on the maximization of the full joint maximum (pseudo-)likelihood [21]. It is therefore natural to consider the full optimization problem:

$$\arg\max_{\theta} \sum_{m=1}^{n} \log c \left[\widehat{F}_1(x_1^{(m)}), \ldots, \widehat{F}_d(x_d^{(m)}); \theta \right],$$

where c is the density associated with C and given by

$$c(u_1, \ldots, u_d; \theta) = \frac{\partial^2 C(u_1, \ldots, u_d; \theta)}{\partial u_1 \ldots \partial u_d} = \frac{\partial^2 \left[\prod_{\{ij\} \in E} \tilde{C}_{ij} \left(u_i^{1/n_i}, u_j^{1/n_j} \right) \right]}{\partial u_1 \cdots \partial u_d}. \quad (15)$$

Let us note that, when $\widehat{F}_k$ is a parametric estimator of F_k, this method is referred to as Inference Function for Margins (IFM, [16], Section 10). When $\widehat{F}_k$ is a nonparametric estimator, this amounts to the pseudo-likelihood method [6]. The asymptotic properties in each case are discussed in the aforementioned references. Here, however, the difficulty rather lies in the fact that (15) is hard to compute by standard differentiation procedures because of the product form of the PBC copula. Nonetheless, when the underlying graph is a tree, a recent message-passing algorithm [15] can be adapted to calculate the derivatives (15) in an efficient way. We have provided an implementation of this algorithm in the context of PBC copulas in the `R` package `PBC` [27]. The main ideas of this algorithm are briefly described in Appendix B. Once the density is computed thanks to the algorithm, one can feed any optimization routine to maximize the likelihood. An example using `optim` in `R` (`www.r-project.org`) is provided in the `PBC` package.

5 Applications to simulated and real datasets

This section is driven by two goals. First, in Paragraph 5.1, we compare the gain in efficiency when the maximization of the full joint likelihood is preferred to the pairwise likelihood for estimating the parameters of PBC copulas. Second, Paragraph 5.2 shows how PBC copulas can be applied to a real data set and illustrates their key properties compared to other copula models.

The considered families for the bivariate copulas $\tilde{C}_{ij}$ in (2) are the following: the Ali-Mikhail-Haq (AMH), Farlie-Gumbel-Morgenstern (FGM), Frank, Gumbel, and Joe families. See [24] or [16] for details about these families. The corresponding PBC copula models (2) are therefore referred to as PBC AMH, PBC FGM, PBC Frank, PBC Gumbel and PBC Joe respectively. The methods used to simulate and infer the copulas can be found in Section 4.

5.1 A simulation experiment to compare pairwise likelihood and full joint likelihood estimation in PBC copulas

In Paragraph 4.2, two methods were presented to estimate the parameters of a PBC copula model: the (pseudo-) pairwise and (pseudo-) full likelihood methods. While the full likelihood approach is more efficient asymptotically than the pairwise strategy, the efficiency gain needs to be quantified, and, moreover, may not be the same for different choices of parametric families. In the following, a simulation experiment is undertaken to address these issues.

We generated 500 datasets of dimension $d = 9$ and size $n = 100$ according to a PBC copula whose tree graph is depicted on Figure 2. Since the data are generated exactly according to the copulas, the margins need not to be estimated, and we thus focus on the comparison of the true pairwise and true full likelihood methods. The amount of time required to maximize the true full likelihood for five dataset replications was 36, 21, 18, 21 and 21 seconds for PBC AMH, PBC FGM, PBC Frank, PBC Gumbel and PBC Joe respectively with a 8 GiB memory and 3.20 GHz processor computer. The $d - 1 = 8$ coordinates of the parameter vectors were chosen to be regularly spaced within the intervals $[-0.9, 0.9], [-0.9, 0.9], [-9, 11], [2, 20]$ and $[1, 20]$ respectively.

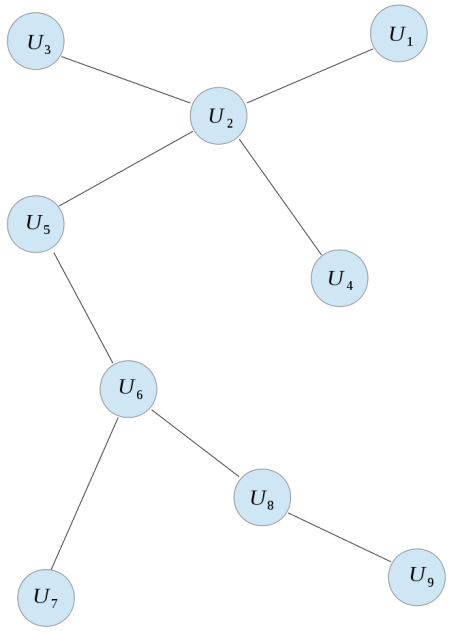


Figure 2: Tree graph associated with the simulated PBC copulas.

The following criteria were calculated in order to assess the results of the experiment. The variance ratio (VR) is defined as

$$VR = \sum_{e=1}^{d-1} \widehat{\text{Var}}\left(\hat{\theta}_e^{FULL}\right) \Big/ \sum_{e=1}^{d-1} \widehat{\text{Var}}\left(\hat{\theta}_e^{PW}\right),$$

where $\hat{\theta}_e^{FULL}$, $\hat{\theta}_e^{PW}$ is the eth coordinate of θ estimated by maximization of the full likelihood, pairwise likelihood, respectively, and where $\widehat{\text{Var}}$ is the empirical variance computed on the 500 replications. For each dataset replication, the mean absolute error associated with the estimated Spearman's rho ρ (MAE_ρ) and Kendall's tau τ

(MAE_τ) are defined as

$$\text{MAE}_\rho = \frac{1}{d-1}\sum_{e=1}^{d-1}|\rho(\theta_e) - \rho(\hat{\theta}_e^{FULL})|, \quad \text{MAE}_\tau = \frac{1}{d-1}\sum_{e=1}^{d-1}|\tau(\theta_e) - \tau(\hat{\theta}_e^{FULL})|.$$

The MAEs were averaged over the 500 replications to get a single value per model.

Copula	VR	MAE_ρ	MAE_τ
PBC AMH	1.01	0.06	0.04
PBC FGM	1.01	0.06	0.04
PBC Frank	0.77	0.04	0.03
PBC Gumbel	0.56	0.00	0.00
PBC Joe	0.63	0.01	0.00

Table 2: Variance ratio (VR) and mean absolute errors (MAEs) for each of the tested PBC models. The MAEs were averaged over the dataset replications.

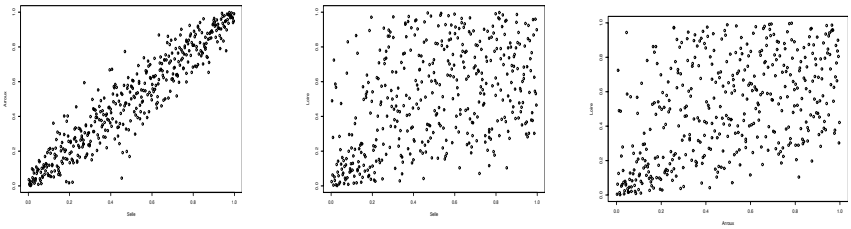
The results are reported in Table 2. It appears that for PBC AMH and PBC FGM, the precision was not improved by maximizing the full joint likelihood relative to the pairwise approach: the variance ratio for those models are close to 1. For the Frank, Gumbel and Joe families, however, the variance decreases by at least 23% in average. These families, in contrast to the AMH and FGM families, are comprehensive, meaning that they include the lower and upper bounds for copulas. Hence, interestingly enough, we observed a gain of efficiency for the most flexible families, and, moreover, this gain has at least a value of 23%. The MAEs are quite low for all the models, indicating that the maximization of the full joint likelihood with the message-passing algorithm of Appendix B performs well.

5.2 Application to an hydrological dataset

In this section, PBC copula models are applied to an hydrological dataset consisting of $d = 3$ stations and $n = 445$ observations, $(X_1^{(m)}, X_2^{(m)}, X_3^{(m)})$, $m = 1, \ldots, n$, of flow rate monthly maxima The sites are located on three French rivers at the following places: La Celle-en-Morvan on the Selle river (S), Rigny-sur-Arroux on the Arroux river (A), and Isclades-et-Rieutord on the Loire river (L). The dependence in the data was confirmed by a statistical test of independence [9], which gave a p-value of 5.10^{-4}. It is implemented in the function `indepTest` of the `R` package `copula` [12]. More surprisingly, the test of max-stability proposed in [17] yields a p-value of 5.10^{-4} suggesting that extreme-value copulas can be ill-adapted to this situation. This test was performed with the function `evTestC` of the `R` package `copula`. For the sake of completeness, we also carried out a test of exchangeability [8, 18] for the three pairs (S,A), (S,L) and (A,L): we found p-values of 0.27, 0.68 and 0.23 respectively. The tests were performed with the function `exchTest` of the `R` package `copula` and suggest exchangeability for all pairs.

Figure 3 displays the pairwise scatterplots of the data after their transformation to uniform margins. It clearly appears that the monthly rates from the Selle and Arroux rivers are strongly positively correlated. To confirm this, quantitative dependence measures, namely empirical Spearman's rho and Kendall's tau coefficients have been computed for all pairs and shown in Table 3 (first line). It follows that the two graph

structures that would be reasonable for our PBC models reduce to S − A − L and A − S − L. The third possible one being discarded as it would imply independence between the S and A variables.



(a) (S,A) (b) (S,L) (c) (A,L)

Figure 3: Pairwise scatterplots for the hydrological dataset.

The three rivers are embedded in the sense that Selle flows into Arroux which flows into Loire. Thus, the graph S − A − L may be easier to interpret but the results are shown for both graphs. The same models as in Section 5.1 were tested, that is, PBC AMH, PBC FGM, PBC Frank, PBC Gumbel, and PBC Joe. The Gumbel copula was also considered, since it is standard in hydrology, see e.g. [28] for an analysis of dependence between rainfall intensity, duration, and depth. Two elliptical copulas were fitted as well: the Gaussian and the Student copulas (see, e.g. [22]). For PBC models, Gumbel and Student copulas, the estimation of the parameters was performed by maximization of the full pseudo-likelihood, as explained in Section 4. The codes that implement these estimation procedures are available in the `R` packages `copula` [12] and `PBC` [27]. In order to assess the fit of the models, the empirical Spearman's rho and Kendall's tau coefficients were compared to their estimated counterparts under the different models. The results are presented in Table 3. As regards these central dependence measures, Gaussian and Student copulas perform clearly better in terms of Spearman's rho and Kendall's tau coefficients, while the PBC models have difficulty to capture such high correlations especially for the AMH and FGM families which are not comprehensive, meaning that they do not allow much dependence (see, e.g., [24]). The standard Gumbel copula that depends on a single parameter cannot model different pairs with different distributions. In particular, the estimated dependence coefficients of the different pairs are equal to each other. The PBC copulas with comprehensive families, i.e. PBC Frank, PBC Gumbel, and PBC Joe, present a better fit. Regarding the choice of the graph structure, we observe that the estimated values were rather consistent from one structure to another except of course for the ones that are zero by construction.

We then assessed the tail behavior of the different copulas. Table 4 shows the estimated upper tail dependence coefficients of all the tested models. Under the assumption of the Gaussian copula, tail dependence is null for every pair of variables, meaning that high levels of flow rates tend to be independent. This is to be put in contrast to the Student, Gumbel, PBC Gumbel or PBC Joe copulas: under the assumption of these, now high levels of certain flow rates tend to be dependent. Since potential dangerous events happen with the co-occurence of extreme flow rates, properly modeling tail dependence is of critical importance. For comparison, the

empirical upper tail coefficients for all pairs were computed by inspecting the empirical version of the function $u \mapsto \lambda^{(U)}(u)$ (13) for high values of u, as in [2]. This obtained empirical values are partly in accordance with the PBC Gumbel and PBC Joe models, i.e. the estimations are close for pairs SA and AL in the $\mathrm{S} - \mathrm{A} - \mathrm{L}$ graph and for pairs SA and SL in the $\mathrm{A} - \mathrm{S} - \mathrm{L}$ graph. The Gaussian copula cannot provide satisfying values due to tail independence while the Student copula seems to capture only the stronger tail dependance (pair SA). As for the Gumbel copula, it suffers again from treating all pairs symmetrically. In this context of tail dependence, the PBC Joe and Gumbel models seem to be more appropriate than the other tested copulas keeping in mind that the choice of the graph structure may impose undesired independence. Finally, although compared to more standard elliptic copulas, the central modelling capability of PBC models may be limited by theoretical bounds as shown in Table 1, they may still be useful to model tail dependence more accurately. Their further advantage is that one can select the most appropriate family of bivariate copulas to allow for more possibilites to model the dependence structure. If estimation is to be performed pair-wise, one may even think of choosing different parametric families for the three pairs of variables.

6 Discussion

In this paper, we have constructed a class of multivariate copulas, called PBC copulas, based on a product of arbitrary bivariate copulas. Therefore, this novel class benefits from the many bivariate families existing in the literature. A natural graph structure helps to visualize the dependencies between the variables and to design an efficient inference algorithm. Full joint multivariate inference can be performed, and shown to perform well, with a message-passing algorithm. An `R` package was developped to promote the use of PBC copulas.

However, PBC copula models still suffer from weaknesses: The more there are edges in the graph, the more the bounds on the dependence coefficients are restrictive. The above point suggests that, perhaps, the use of PBC copulas in practice should stick to weakly dependent data or combined with other models with complementary properties for instance by considering mixture of distributions.

To perform the estimation procedures given in Section 4, the knowledge of the graphical structure underpinning the data has to be assumed. In practice, however, one has to choose which pairs of variables should be linked. A simple method would consist of considering all possible graphical structures and either combine them or select the best ones based on the resulting (pseudo-) likelihoods. Obviously, this approach is unfeasible in practice for high dimensions. Thus, an alternative would be to link the most dependent variables under the constraint that the degree of the graph is below a certain threshold. If one wants to maximize the full (pseudo-)likelihood, the constraint would be that the resulting graph is a tree. In this case, one can run a minimum spanning tree algorithm [25] (implemented in the `R` package `igraph`, see the function `minimum.spanning.tree`) where the weights associated with the edges would decrease as Spearman's rhos (or any other dependence coefficient) increase. This strategy embeds the construction of a tree based on the geographic proximity between the variables at play (the weights would be related to the geographical distances between the sites). Finally, another possible approach would be to rely on application specific expert knowledge.

	$\rho_{\mathrm{S,A}}$	$\rho_{\mathrm{A,L}}$	$\rho_{\mathrm{S,L}}$	$\tau_{\mathrm{S,A}}$	$\tau_{\mathrm{A,L}}$	$\tau_{\mathrm{S,L}}$
Empirical estimates	0.95	0.48	0.45	0.80	0.33	0.31
Gumbel	0.55	0.55	0.55	0.39	0.39	0.39
Gaussian	0.93	0.49	0.45	0.76	0.34	0.31
Student	0.93	0.49	0.45	0.78	0.34	0.31
S - A - L graph						
PBC AMH	0.25	0.24	0	0.17	0.16	0
PBC FGM	0.20	0.20	0	0.13	0.13	0
PBC Frank	0.56	0.35	0	0.38	0.24	0
PBC Gumbel	0.51	0.32	0	0.37	0.22	0
PBC Joe	0.49	0.26	0	0.35	0.17	0
A - S - L graph						
PBC AMH	0.25	0	0.24	0.17	0	0.16
PBC FGM	0.20	0	0.20	0.13	0	0.13
PBC Frank	0.52	0	0.31	0.37	0	0.21
PBC Gumbel	0.51	0	0.24	0.36	0	0.16
PBC Joe	0.50	0	0.13	0.35	0	0.08

Table 3: Empirical pairwise dependence coefficients and their estimations under the tested copulas. The symbols ρ and τ stand for the Spearman's rho and Kendall's tau respectively. For instance, $\rho_{\mathrm{S,A}}$ is the Spearman's rho coefficient between the variables S and A. The PBC models are estimated using two graph structures.

	$\hat{\lambda}^{(U)}_{\mathrm{S,A}}$	$\hat{\lambda}^{(U)}_{\mathrm{A,L}}$	$\hat{\lambda}^{(U)}_{\mathrm{S,L}}$
Empirical estimates	0.43	0.20	0.20
Gumbel	0.48	0.48	0.48
Gaussian	0	0	0
Student	0.45	0.02	0.02
S - A - L graph			
PBC AMH	0.01	0.01	0
PBC FGM	0.01	0.01	0
PBC Frank	0.06	0.03	0
PBC Gumbel	0.45	0.29	0
PBC Joe	0.47	0.30	0
A - S - L graph			
PBC AMH	0.01	0	0.01
PBC FGM	0.01	0	0.01
PBC Frank	0.05	0	0.02
PBC Gumbel	0.45	0	0.17
PBC Joe	0.47	0	0.22

Table 4: Pairwise empirical and estimated upper tail dependence coefficients for the tested copulas.

Acknowledgment. The authors thank "Banque HYDRO du Ministère de l'Écologie, du Développement durable et de l'Énergie" for providing the data and Benjamin Renard for fruitful dicussions about statistical issues in hydrological science. They also thank two anonymous referees and the associate editor for helpful comments.

Appendix

A Proofs

Proof of Proposition 1 In the copula (3), put $E' = E$, and, for $v \in (0,1]$,

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

It can be checked that (4) is true and that (3) leads to (2). Therefore, (2) is a well-defined copula.

Proof of Proposition 2 Let us now prove that (2) is the only copula arising from (3). Condition (i) implies that if $e \notin N(i)$ then $g_{ei} = 1$, $i = 1, \ldots, d$. Hence, the constraint over the functions reduces to $\prod_{e \in N(i)} g_{ei}(v) = v$, $v \in [0,1]$. In view of condition (ii), one has $g_{ei} = g_i$ for $e \in N(i)$, hence $(g_i(v))^{n_i} = v$. Therefore

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

To conclude, it suffices to rewrite the product in (3) as

$$\prod_{e \in E} \tilde{C}_e(1, \ldots, 1, u_i^{1/n_i}, 1, \ldots, 1, u_j^{1/n_j}, 1, \ldots, 1) = \prod_{\{ij\} \in E} \tilde{C}_{ij}(u_i^{1/n_i}, u_j^{1/n_j})$$

which corresponds to (2).

Proof of Proposition 3 Let us first prove that (5) is the distribution function associated with the copula (2). From (1) we have

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

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. A similar result holds for the second margin $\Phi_{ij,2}$. Conversely, let us prove that (2) is the copula associated with (5). Let $\Phi_{ij,k}$, $k = 1, 2$ be the k-th univariate marginal of $\Phi_{ij}, \{ij\} \in E$. The copula associated with F is given by

$$C_F(u_1, \ldots, u_d) = F\left(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)\right) = \prod_{\{ij\} \in E} \Phi_{ij}\left(F_i^{-1}(u_i), F_j^{-1}(u_j)\right).$$

For all $\{ij\} \in E$, let $\tilde{C}_{ij}$ be the copula associated with Φ_{ij}. It follows that

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

so that $\Phi_{ij}\left(F_i^{-1}(u_i), F_j^{-1}(u_j)\right) = \tilde{C}_{ij}\left(\Phi_{ij,1} \circ F_i^{-1}(u_i), \Phi_{ij,2} \circ F_j^{-1}(u_j)\right)$ and

$$C_F(u_1, \ldots, u_d) = \prod_{\{ij\} \in E} \tilde{C}_{ij}\left(\Phi_{ij,1} \circ F_i^{-1}(u_i), \Phi_{ij,2} \circ F_j^{-1}(u_j)\right). \tag{16}$$

Moreover, since $\tilde{C}_F$ is a copula, one has for all $k = 1, \ldots, d$:

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

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 $\Phi_k(z) = F_k(z)^{1/n_k}$ for all $z \in \mathbb{R}$ and $k = 1, \ldots, d$. The result is obtained by by pluging Φ_k into (16).

Proof of Proposition 5 Let C be a PBC copula. For all $(k, \ell) \in \{1, \ldots, d\}^2$, the copula associated with the random pair (U_k, U_ℓ) is defined by

$$C_{k\ell}(u_k, u_\ell) = C(1, \ldots, 1, u_k, 1, \ldots, 1, u_\ell, 1, \ldots, 1).$$

Two cases arise: If $\{kl\} \in E$, then

$$\begin{aligned} C_{k\ell}(u_k, u_\ell) =& \left(\prod_{e \in N(k) \setminus \{k\ell\}} \tilde{C}_e(u_k^{1/n_k}, 1)\right)\left(\prod_{e \in N(\ell) \setminus \{k\ell\}} \tilde{C}_e(u_\ell^{1/n_\ell}, 1)\right) \tilde{C}_{k\ell}(u_k^{1/n_k}, u_l^{1/n_\ell}) \\ =& u_k^{(n_k-1)/n_k} u_\ell^{(n_\ell-1)/n_l} \tilde{C}_{k\ell}(u_k^{1/n_k}, u_\ell^{1/n_\ell}) \end{aligned}$$

otherwise

$$\begin{aligned} C_{kl}(u_k, u_\ell) =& \left(\prod_{e \in N(k)} \tilde{C}_e(u_k^{1/n_k}, 1)\right)\left(\prod_{e \in N(\ell)} \tilde{C}_e(u_\ell^{1/n_\ell}, 1)\right) \\ =& u_k u_\ell. \end{aligned}$$

Proof of Proposition 7 The Fréchet-Hoeffding bounds for copulas (see e.g. [24], p. 11) applied to $\tilde{C}_{k\ell}$ in (9) yield

$$W_{k\ell}(u_k, u_\ell) \leq C_{k\ell}(u_k, u_\ell) \leq M_{k\ell}(u_k, u_\ell), \tag{17}$$

where

$$\begin{aligned} W_{k\ell}(u_k, u_\ell) &:= u_k^{1-1/n_k} u_\ell^{1-1/n_\ell} \max(u_k^{1/n_k} + u_l^{1/n_\ell} - 1, 0), \\ M_{k\ell}(u_k, u_\ell) &:= u_k^{1-1/n_k} u_\ell^{1-1/n_\ell} \min(u_k^{1/n_k}, u_l^{1/n_\ell}). \end{aligned}$$

Clearly, $M_{k\ell}(u, u)/u \to 0$ as $u \downarrow 0$. It is easily seen that $W_{k\ell}(u, u)/u \to 0$ as $u \downarrow 0$ which implies $C_{k\ell}(u, u)/u \to 0$ and thus $\lambda_{k\ell}^{(L)} = 0$. It is straightforward that $(1 - 2u + M_{k\ell}(u, u))/(1 - u) \to 1/\max(n_k, n_\ell)$ as $u \uparrow 1$ which entails $\lambda_{k\ell}^{(U)} \leq \min(1/n_k, 1/n_\ell)$. To compute the lower and upper bounds for $\rho_{k\ell}$ and $\tau_{k\ell}$, it suffices to substitute $W_{k\ell}$ and $M_{k\ell}$ into (12) and (11). Lengthy but elementary computations lead to the results. Finally, letting n_k or n_ℓ going to infinity in (17) yields that $C_{k\ell}$ tends to independence.

B Main principle of the message-passing algorithm

Let us denote the parameter vector as $\theta = (\theta_{ij})_{\{ij\}\in E}$. Recall that the graph is assumed to be a tree, that is, there is no cycles in the graph (then $|E| = d-1$). Let $V = \{1,\dots,d\}$ and $u = (u_1,\dots,u_d)$ a vector in $[0,1]^d$. For a subset $A \subset V$, the notation $\partial_{u_A} C(u;\theta)$ stands for the derivative of C with respect to all the variables in A. For instance the density writes

$$\frac{\partial^d C(u;\theta)}{\partial_{u_1}\dots\partial_{u_d}} = \partial_{u_V} C(u;\theta) = c(u;\theta),$$

and the gradient with respect to the parameter vector is

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

To keep the notation simple, the dependence on the parameter vector θ is dropped in the remaining of this section. The purpose here is not to give the algorithm, but rather to provide an intuitive idea of it. Let us write

$$C(u_1,\dots,u_d) = \prod_{\{ij\}\in E} \tilde{C}_{ij}(u_i^{1/n_i}, u_j^{1/n_j}) =: \prod_{\{ij\}\in E} \Phi_{ij}(u_i,u_j).$$

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 edge indexed by e (see Figure 4). The idea is to note that, since the graph is a tree, the copula C can be decomposed over the subtrees rooted at i:

$$C(u) = \prod_{e\in E} \Phi_e(u) =: \prod_{e\in N(i)} T_{\tau_e^i}(u), \quad u = (u_1,\dots,u_d),$$

where $T_{\tau_e^i}(u)$ corresponds to the product of all edges located in the subtree τ_e^i. Since the $T_{\tau_e^i}(u)$'s do not share any variables (except the root), the derivative and the product operations commute, more precisely,

$$\begin{aligned}\partial_{u_V} C(u) &= \partial_{u_i, u_{V\setminus i}} \left[\prod_{e\in N(i)} T_{\tau_e^i}(u)\right] = \partial_{u_i}\left[\prod_{e\in N(i)} \partial_{u_{\tau_e^i\setminus i}} T_{\tau_e^i}(u)\right] \\ &= \partial_{u_i}\left[\prod_{e\in N(i)} \mu_{e\to i}(u)\right]. \end{aligned} \tag{18}$$

The quantity $\mu_{e\to i}(u) := \partial_{u_{\tau_e^i\setminus i}} T_{\tau_e^i}(u)$ is called a message from the edge indexed by e to the variable indexed by i. Now consider $T_{\tau_e^i}(u)$ and let j be the neighbor variable index of e. One can go deeper into the tree, that is, we have

$$T_{\tau_e^i}(u) = \Phi_e(u_i,u_j) T_{\tau_j^e}(u)$$

where τ_j^e is the subtree rooted at the edge indexed by e and containing the variable indexed by j (see Figure 4). Hence,

$$\partial_{u_{\tau_e^i\setminus i}} T_{\tau_e^i}(u) = \partial_{u_j}\left[\phi_e(u_i,u_j)\partial_{u_{\tau_j^e\setminus j}} T_{\tau_j^e}(u)\right] = \partial_{u_j}\left[\phi_e(u_i,u_j)\mu_{j\to e}(u)\right].$$

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

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

hence,

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

where the message $\mu_{e'\to j}(u)$ has been already defined in (18). To summarize, the calculation of $\mu_{e\to i}(u)$ requires the calculation of $\mu_{j\to e}(u)$, which, in turn, requires the calculation of $\mu_{e'\to j}(u)$, where $e = \{ij\}$ and e' is an edge index attached to j. The algorithm presented above 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 (18).

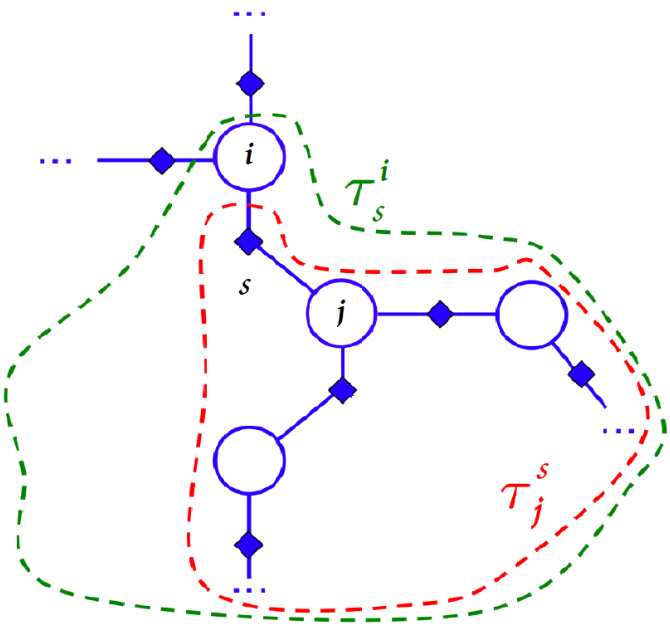


Figure 4: Examples of subtrees. This figure is partly drawn from [14].

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] S. Coles, J. Heffernan, and J. Tawn. Dependence measures for extreme value analyses. *Extremes*, 2:339–365, 1999.

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

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

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

[6] C. Genest, K. Ghoudi, and L.-P. Rivest. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82(3):543–552, 1995.

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

[8] C. Genest, J. Nešlehová, and J.-F. Quessy. Tests of symmetry for bivariate copulas. *Annals of the Institute of Statistical Mathematics*, 64(4):811–834, 2012.

[9] C. Genest and B. Rémillard. Tests of independence and randomness based on the empirical copula process. *Test*, 13:335–369, 2004.

[10] C. Genest and L.-P. Rivest. A characterization of Gumbel's family of extreme value distributions. *Statistics & Probability Letters*, 8(3):207–211, 1989.

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

[12] M. Hofert, I. Kojadinovic, M. Maechler, and J. Yan. *copula: Multivariate Dependence with Copulas*, 2014. R package version 0.999-12.

[13] 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.

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

[15] 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.

[16] H. Joe. *Multivariate models and dependence concepts*. Chapman & Hall/CRC, 1997.

[17] I. Kojadinovic, J. Segers, and J. Yan. Large-sample tests of extreme-value dependence for multivariate copulas. *Canadian Journal of Statistics*, 39(4):703–720, 2011.

[18] I. Kojadinovic and J. Yan. A non-parametric test of exchangeability for extreme-value and left-tail decreasing bivariate copulas. *Scandinavian Journal of Statistics*, 39(3):480–496, 2012.

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

[20] E. Liebscher. Erratum to "construction of asymmetric multivariate copulas"[j. multivariate anal. 99 (2008) 2234–2250]. *Journal of Multivariate Analysis*, 102(4):869–870, 2011.

[21] B. G. Lindsay. Composite likelihood methods. *Contemporary Mathematics*, 80(1):221–39, 1988.

[22] A. J. McNeil, R. Frey, and P. Embrechts. *Quantitative risk management: concepts, techniques, and tools*. Princeton university press, 2010.

[23] A. J. McNeil and J. Nešlehová. Multivariate Archimedean copulas, d-monotone functions and ℓ_1-norm symmetric distributions. *The Annals of Statistics*, 37(5B):3059–3097, 2009.

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

[25] R.C. Prim. Shortest connection networks and some generalizations. *Bell System Technical Journal*, 37:1389–1401, 1957.

[26] 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.

[27] T. Van Pham and G. Mazo. *PBC: product of bivariate copulas*. http://cran.r-project.org, 2014. R package version 1.2.

[28] L. Zhang and V. P. Singh. Gumbel–Hougaard copula for trivariate rainfall frequency analysis. *Journal of Hydrologic Engineering*, 12(4):409–419, 2007.