Aas, Czado, Frigessi, Bakken:

Pair-copula constructions of multiple dependence

Projektpartner

Pair-copula constructions of multiple dependence

Kjersti Aas†

The Norwegian Computing Center, Oslo, Norway

Claudia Czado

Technische Universität, München, Germany

Arnoldo Frigessi

University of Oslo and The Norwegian Computing Center, Norway

Henrik Bakken

The Norwegian University of Science and Technology, Trondheim,Norway‡

Abstract. Building on the work of Bedford, Cooke and Joe, we show how multivariate data, which exhibit complex patterns of dependence in the tails, can be modelled using a cascade of pair-copulae, acting on two variables at a time. We use the pair-copula decomposition of a general multivariate distribution and propose a method to perform inference. The model construction is hierarchical in nature, the various levels corresponding to the incorporation of more variables in the conditioning sets, using pair-copulae as simple building blocs. Pair-copula decomposed models also represent a very ¤exible way to construct higher-dimensional coplulae. We apply the methodology to a £nancial data set. Our approach represents the £rst step towards developing of an unsupervised algorithm that explores the space of possible pair-copula models, that also can be applied to huge data sets automatically.

1. Introduction

The pioneering work of Bedford and Cooke (2001b, 2002), also based on Joe (1996), which introduces a probabilistic construction of multivariate distributions based on the simple building blocs called *pair-copulae*, has remained completely overseen. It represents a radically new way of constructing complex multivariate highly dependent models, which parallels classical hierarchical modelling (Green et al., 2003). There, the principle is to model dependency using simple local building blocs based on conditional independence, e.g. cliques in random fields. Here, the building blocs are pair-copulae. The modelling scheme is based on a decomposition of a multivariate density into a cascade of pair copulae, applied on original variables and on their conditional and unconditional distribution functions. In this paper, we show that this decomposition can be a central tool in model building, not requiring conditional independence assumptions when these are not natural, but maintaining the logic of building complexity by simple elementary bricks. We present some of the theory of

†*Address for correspondence:* The Norwegian Computing Center, P.O. Box 114 Blindern, N-0314 Oslo, Norway, Kjersti.Aas@nr.no

Bedford and Cooke (2001b, 2002) from a more practical point of view, as a general modelling approach, concentrating on inference based on n variables repeatedly observed, say over time.

Building higher-dimensional copulae is generally recognised as a difficult problem. There is a huge number of parametric bivariate copulas, but the set of higher-dimensional copulae is rather limited. There have been some attempts to construct multivariate extensions of Archimedean bivariate copulae, see e.g. Embrechts et al. (2003) and Savu and Trede (2006). However, it is our opinion that the pair-copula decomposition treated in this paper represents a more flexible and intuitive way of extending bivariate copulae to higher dimensions.

The paper is organised as follows. In Section 2 we introduce the pair-copula decomposition of a general multivariate distribution and illustrate this with some simple examples. In Section 3 we see the effect of conditional independence, if assumed, on the pair-copula construction. Section 4 describes how to simulate from pair-copula decomposed models. In Section 5 we describe our estimation procedure, while Section 6 reviews several basic pair-copulae useful in model constructions. In Section 7 we discuss aspects of the model selection process. In Section 8 we apply the methodology, and discuss its limitations and difficulties in the context of a financial data set. Finally, Section 9 contains some concluding remarks.

2.   A pair-copula decomposition of a general multivariate distribution

Consider n random variables $\boldsymbol{X} = (X_1, \ldots, X_n)$ with a joint density function $f(x_1, \ldots, x_n)$. This density can be factorised as

$$f(x_1, \ldots, x_n) = f(x_n) \cdot f(x_{n-1}|x_n) \cdot f(x_{n-2}|x_{n-1}, x_n) \ldots \cdot f(x_1|x_2, \ldots, x_n), \qquad (1)$$

and this decomposition is unique up to a relabelling of the variables.

In a sense every joint distribution function implicitly contains both a description of the marginal behaviour of individual variables and a description of their dependency structure. Copulae provide a way of isolating the description of their dependency structure. A copula is multivariate distribution, C, with uniformly distributed marginals $U(0,1)$ on [0,1]. Sklar's theorem (Sklar, 1959) states that every multivariate distribution F with marginals F_1, $F_2, \ldots, F_n$ can be written as

$$F(x_1, \ldots, x_n) = C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)), \qquad (2)$$

for some apropriate n-dimensional copula C. In fact, the copula from (2) has the expression

$$C(u_1, \ldots, u_n) = F(F_1^{-1}(u_1), F_2^{-1}(u_2), \ldots, F_n^{-1}(u_n)),$$

where the F_i^{-1}'s are the inverse distribution functions of the marginals.

Passing to the joint density function f, for an absolutely continous F with strictly increasing, continuous marginal densities $F_1, \ldots F_n$ (McNeil et al., 2006), we have

$$f(x_1, \ldots, x_n) = c_{12 \cdots n}(F_1(x_1), \ldots F_n(x_n)) \cdot f_1(x_1) \cdots f_n(x_n) \qquad (3)$$

for some (uniquely identified) n-variate copula density $c_{12 \cdots n}(\cdot)$. In the bivariate case (3) simplifies to

$$f(x_1, x_2) = c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1) \cdot f_2(x_2),$$

where $c_{12}(\cdot, \cdot)$ is the appropriate *pair-copula density* for the pair of transformed variables $F_1(x_1)$ and $F_2(x_2)$. For a conditional density it easily follows that

$$f(x_1|x_2) = c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1),$$

for the same pair-copula. For example, the second factor, $f(x_{n-1}|x_n)$, in the right hand side of (1) can be decomposed into the pair-copula $c_{(n-1)n}(F(x_{n-1}), F(x_n))$ and a marginal density $f_n(x_n)$. For three random variables X_1, X_2 and X_3 we have that

$$f(x_1|x_2, x_3) = c_{12|3}(F_{1|3}(x_1|x_3), F_{2|3}(x_2|x_3)) \cdot f(x_1|x_3), \tag{4}$$

for the appropriate pair-copula $c_{12|3}$, applied to the transformed variables $F(x_1|x_3)$ and $F(x_2|x_3)$. An alternative decomposition is

$$f(x_1|x_2, x_3) = c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \cdot f(x_1|x_2), \tag{5}$$

where $c_{13|2}$ is different from the pair-copula in (4). Decomposing $f(x_1|x_2)$ in (5) further, leads to

$$f(x_1|x_2, x_3) = c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \cdot c_{12}(F_1(x_1), F_2(x_2)) \cdot f_1(x_1),$$

where two pair-copulae are present.

It is now clear that each term in (1) can be decomposed into the appropriate pair-copula times a conditional marginal density, using the general formula

$$f(x|\boldsymbol{v}) = c_{xv_j|\boldsymbol{v}_{-j}}(F(x|\boldsymbol{v}_{-j}), F(v_j|\boldsymbol{v}_{-j})) \cdot f(x|\boldsymbol{v}_{-j}),$$

for a d-dimensional vector $\boldsymbol{v}$. Here v_j is one arbitrarily chosen component of $\boldsymbol{v}$ and $\boldsymbol{v}_{-j}$ denotes the v-vector, excluding this component. In conclusion, under appropriate regularity conditions, a multivariate density can be expressed as a product of pair-copulae, acting on several different conditional probability distributions. It is also clear that the construction is iterative in its nature, and that given a specific factorisation, there are still many different reparameterisations.

The pair-copula construction involves marginal conditional distributions of the form $F(x|\boldsymbol{v})$. For every j, Joe (1996) showed that

$$F(x|\boldsymbol{v}) = \frac{\partial\, C_{x,v_j|\boldsymbol{v}_{-j}}(F(x|\boldsymbol{v}_{-j}), F(v_j|\boldsymbol{v}_{-j}))}{\partial F(v_j|\boldsymbol{v}_{-j})}, \tag{6}$$

where $C_{ij|\boldsymbol{k}}$ is a bivariate copula distribution function. For the special case where v is univariate we have

$$F(x|v) = \frac{\partial\, C_{xv}(F_x(x), F_v(v))}{\partial F_v(v)}.$$

In Sections 4-7 we will use the function $h(x, v, \Theta)$ to represent this conditional distribution function when x and v are uniform, i.e. $f(x) = f(v) = 1$, $F(x) = x$ and $F(v) = v$. That is,

$$h(x, v, \Theta) = F(x|v) = \frac{\partial\, C_{x,v}(x, v, \Theta)}{\partial v}, \tag{7}$$

where the second parameter of $h(\cdot)$ always corresponds to the conditioning variable and Θ denotes the set of parameters for the copula of the joint distribution function of x and v. Further, let $h^{-1}(u, v, \Theta)$ be the inverse of the h-function with respect to the first variable u, or equivalently the inverse of the conditional distribution function.
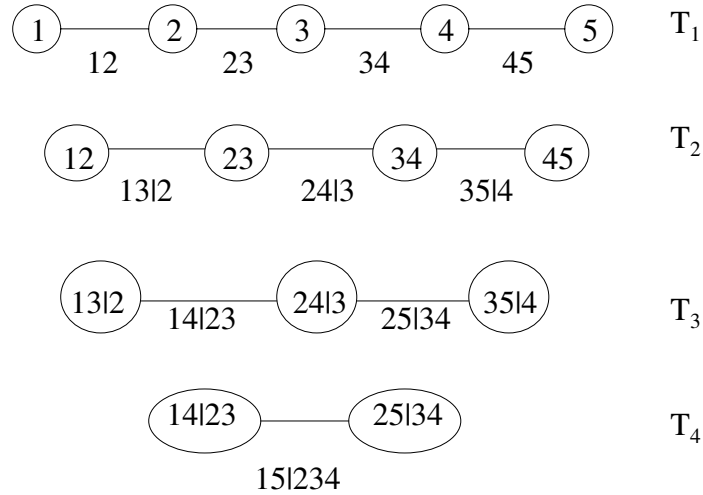
Figure 1. A D-vine with 5 variables, 4 trees and 10 edges. Each edge may be may be associated with a pair-copula.

2.1. Vines

For high-dimensional distributions, there are a significant number of possible pair-copulae constructions. For example, as will be shown in Section 2.4, there are 240 different constructions for a five-dimensional density. To help organising them, Bedford and Cooke (2001b, 2002) have introduced a graphical model denoted *the regular vine*. The class of regular vines is still very general and embraces a large number of possible pair-copula decompositions. Here, we concentrate on two special cases of regular vines; the *canonical vine* and the *D-vine* (Kurowicka and Cooke, 2004). Each model gives a specific way of decomposing the density. The specification may be given in form of a nested set of trees. Figure 1 shows the specification corresponding to a five-dimensional D-vine. It consists of four trees $T_j, j = 1, \ldots 4$. Tree T_j has $6 - j$ nodes and $5 - j$ edges. Each edge corresponds to a pair-copula density and the edge label corresponds to the subscript of the pair-copula density, e.g. edge $14|23$ corresponds to the copula density $c_{14|23}(\cdot)$. The whole decomposition is defined by the $n(n-1)/2$ edges and the marginal densities of each variable. The nodes in tree T_j are only necessary for determining the labels of the edges in tree T_{j+1}. As can be seen from Figure 1, two edges in T_j, which become nodes in T_{j+1}, are joined by an edge in T_{j+1} only if these edges in T_j share a common node. Note that the tree structure is not strictly necessary for applying the pair-copula methodology, but it helps identifying the different pair-copula decompositions.

Bedford and Cooke (2001b) give the density of an n-dimensional distribution in terms of a regular vine, which we specialise to a D-vine and a canonical vine. The density

$f(x_1, \ldots, x_n)$ corresponding to a D-vine may be written as

$$\prod_{k=1}^{n} f(x_k) \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{i,i+j|i+1,\ldots,i+j-1} \left(F(x_i|x_{i+1}, \ldots, x_{i+j-1}), F(x_{i+j}|x_{i+1}, \ldots, x_{i+j-1}) \right), \quad (8)$$

where index j identifies the trees, while i runs over the edges in each tree.

In a D-vine, no node in any tree T_j is connected to more than two edges. In a canonical vine, each tree T_j has a unique node that is connected to $n - j$ edges. Figure 2 shows a canonical vine with 5 variables. The n-dimensional density corresponding to a canonical vine is given by

$$\prod_{k=1}^{n} f(x_k) \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{j,j+i|1,\ldots,j-1} \left(F(x_j|x_1, \ldots, x_{j-1}), F(x_{j+i}|x_1, \ldots, x_{j-1}) \right). \quad (9)$$
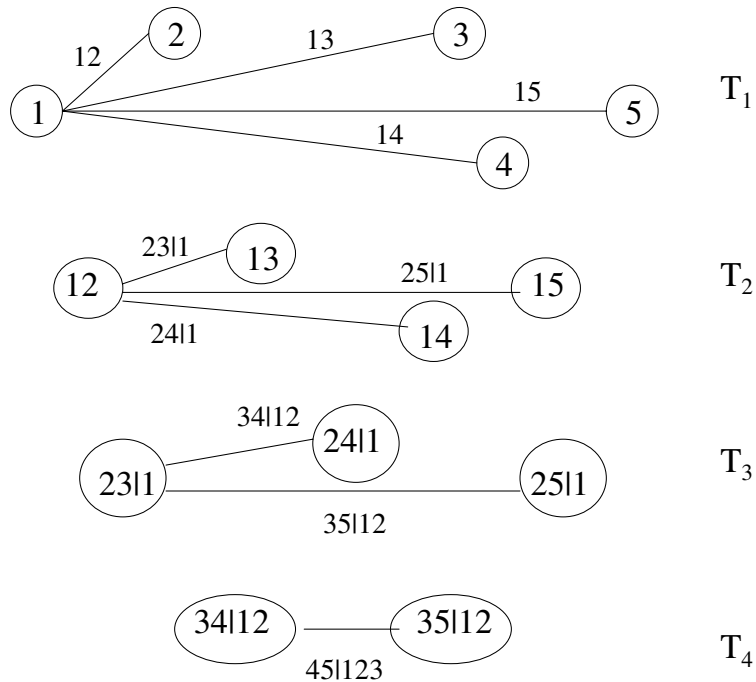

Figure 2. A canonical vine with 5 variables, 4 trees and 10 edges.

Fitting a canonical vine might be advantageous when a particular variable is known to be a key variable that governs interactions in the data set. In such a situation one may decide to locate this variable at the root of the canonical vine, as we have done with variable 1 in Figure 2. The notation of D-vines resembles independence graphs more than that of canonical vines.

2.2. Three variables

The general expression for both the canonical and the D-vine structures in the three-dimensional case is

$$
\begin{aligned}
f(x_1, x_2, x_3) &= f(x_1) \cdot f(x_2) \cdot f(x_3) \\
&\quad \cdot\ c_{12}\left(F(x_1), F(x_2)\right) \cdot c_{23}\left(F(x_2), F(x_3)\right) \\
&\quad \cdot\ c_{13|2}\left(F(x_1|x_2), F(x_3|x_2)\right).
\end{aligned} \tag{10}
$$

There are six ways of permuting x_1, x_2 and x_3, in (10), but only three give different decompositions. Moreover, each of the 3 decompositions is both a canonical vine and a D-vine.

2.3. Four variables

The four-dimensional canonical vine structure is generally expressed as

$$
\begin{aligned}
f(x_1, x_2, x_3, x_4) &= f(x_1) \cdot f(x_2) \cdot f(x_3) \cdot f(x_4) \\
&\quad \cdot\ c_{12}\left(F(x_1), F(x_2)\right) \cdot c_{13}\left(F(x_1), F(x_3)\right) \cdot c_{14}\left(F(x_1), F(x_4)\right) \\
&\quad \cdot\ c_{23|1}\left(F(x_2|x_1), F(x_3|x_1)\right) \cdot c_{24|1}\left(F(x_2|x_1), F(x_4|x_1)\right) \\
&\quad \cdot\ c_{34|12}\left(F(x_3|x_1, x_2), F(x_4|x_1, x_2)\right),
\end{aligned}
$$

and the D-vine structure as

$$
\begin{aligned}
f(x_1, x_2, x_3, x_4) &= f(x_1) \cdot f(x_2) \cdot f(x_3) \cdot f(x_4) \\
&\quad \cdot\ c_{12}\left(F(x_1), F(x_2)\right) \cdot c_{23}\left(F(x_2), F(x_3)\right) \cdot c_{34}\left(F(x_3), F(x_4)\right) \\
&\quad \cdot\ c_{13|2}\left(F(x_1|x_2), F(x_3|x_2)\right) \cdot c_{24|3}\left(F(x_2|x_3), F(x_4|x_3)\right) \\
&\quad \cdot\ c_{14|23}\left(F(x_1|x_2, x_3), F(x_4|x_2, x_3)\right).
\end{aligned} \tag{11}
$$

In Appendix A we derive these expressions.

In total, there are 12 different D-vine decompositions and 12 different canonical vine decompositions, and none of the D-vine decompositions are equal to any of the canonical vine decompositions. There are no other possible regular vine decompositions. Hence, in the four-dimensional case there are 24 different possible pair-copula decompositions, 12 canonical vines and 12 D-vines.

2.4. Five variables

The general expression for the five-dimensional canonical vine structure is

$$
\begin{aligned}
f(x_1, x_2, x_3, x_4, x_5) &= f(x_1) \cdot f(x_2) \cdot f(x_3) \cdot f(x_4) \cdot f(x_5) \\
&\quad \cdot\ c_{12}\left(F(x_1), F(x_2)\right) \cdot c_{13}\left(F(x_1), F(x_3)\right) \cdot c_{14}\left(F(x_1), F(x_4)\right) \\
&\quad \cdot\ c_{15}\left(F(x_1), F(x_5)\right) \cdot c_{23|1}\left(F(x_2|x_1), F(x_3|x_1)\right) \\
&\quad \cdot\ c_{24|1}\left(F(x_2|x_1), F(x_4|x_1)\right) \cdot c_{25|1}\left(F(x_2|x_1), F(x_5|x_1)\right) \\
&\quad \cdot\ c_{34|12}\left(F(x_3|x_1, x_2), F(x_4|x_1, x_2)\right) \cdot c_{35|12}\left(F(x_3|x_1, x_2), F(x_5|x_1, x_2)\right) \\
&\quad \cdot\ c_{45|123}\left(F(x_4|x_1, x_2, x_3), F(x_5|x_1, x_2, x_3)\right),
\end{aligned}
$$

and the general expression for the D-vine structure is

$$
\begin{aligned}
f(x_1, x_2, x_3, x_4, x_5) =\ & f(x_1) \cdot f(x_2) \cdot f(x_3) \cdot f(x_4) \cdot f(x_5) \\
& \cdot\ c_{12}\left(F(x_1), F(x_2)\right) \cdot c_{23}\left(F(x_2), F(x_3)\right) \cdot c_{34}\left(F(x_3), F(x_4)\right) \\
& \cdot\ c_{45}\left(F(x_4), F(x_5)\right) \cdot c_{13|2}\left(F(x_1|x_2), F(x_3|x_2)\right) \\
& \cdot\ c_{24|3}\left(F(x_2|x_3), F(x_4|x_3)\right) \cdot c_{35|4}\left(F(x_3|x_4), F(x_5|x_4)\right) \\
& \cdot\ c_{14|23}\left(F(x_1|x_2, x_3), F(x_4|x_2, x_3)\right) \cdot c_{25|34}\left(F(x_2|x_3, x_4), F(x_5|x_3, x_4)\right) \\
& \cdot\ c_{15|234}\left(F(x_1|x_2, x_4, x_3), F(x_5|x_2, x_4, x_3)\right).
\end{aligned}
$$

In the five-dimensional case there are regular vines that are neither canonical nor D-vines. One example is the following:

$$
\begin{aligned}
f_{12345}(x_1, x_2, x_3, x_4, x_5) =\ & f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \cdot f_4(x_4) \cdot f_5(x_5) \\
& \cdot\ c_{12}\left(F_1(x_1), F_2(x_2)\right) \cdot c_{23}\left(F_2(x_2), F_3(x_3)\right) \cdot c_{34}(F_3(x_3), F_4(x_4)) \\
& \cdot\ c_{35}(F_3(x_3), F_5(x_5)) \\
& \cdot\ c_{13|2}\left(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)\right) \cdot c_{24|3}(F_{2|3}(x_2|x_3), F_{4|3}(x_4|x_3)) \\
& \cdot\ c_{45|3}(F_{4|3}(x_4|x_3), F_{5|3}(x_5|x_3)) \\
& \cdot\ c_{14|23}(F_{1|23}(x_1|x_2, x_3), F_{4|23}(x_4|x_2, x_3)) \\
& \cdot\ c_{25|34}(F_{2|34}(x_2|x_3, x_4), F_{5|34}(x_5|x_3, x_4)) \\
& \cdot\ c_{15|234}(F_{1|234}(x_1|x_2, x_3, x_4), F_{5|234}(x_5|x_2, x_3, x_4)).
\end{aligned}
$$

The corresponding structure is shown in Figure 3. In tree T_1 node 3 has three neighbours; 2,4 and 5. Hence, this is not a D-vine, for which no node in any tree is connected to more than two edges. Moreover, it is not a canonical vine, since node 3 in T_1 is connected to 3 edges instead of 4.

In total there are 60 different D-vines and 60 different canonical vines in the five-dimensional case, and none of the D-vines are equal to any of the canonical vines. In addition to the canonical and D-vines, there are also 120 other regular vines. Hence, in the five-dimensional case there are 240 different possible pair-copula decompositions, 60 canonical vines, 60 D-vines, and 120 other types of decompositions.

2.5.   *n variables*

Considering Figure 2 we see that the conditioning sets **of the edges** in each of the trees T_2, T_3 and T_4 are the same. For example in T_3 the conditioning set is always $\{12\}$. Extending this idea to n nodes, we see that there are n choices for the conditioning set $\{i_2\}$ in T_2, $n-1$ choices for the conditioning set $\{i_2, i_3\}$ in T_3 once i_2 is chosen in T_2. Finally, we have 3 choices for the conditioning set $\{i_{n-1}, i_{n-2}, \cdots, i_2\}$ when $i_2, \cdots, i_{n-2}$ are chosen before. So all together we have $n(n-1)\cdots 3 = \frac{n!}{2}$ different canonical vines on n nodes.

For an n-dimensional D-vine, there are $n!$ choices to arrange the order in the tree T_1. Since we have undirected edges, i.e. $c_{ij|D} = c_{ji|D}$ for all pairs i, j and arbitrary conditioning sets for D-vines, we can reverse the order in the tree T_1 for a D-vine without changing the corresponding vine. Therefore we have only $\frac{n!}{2}$ different trees on the first level. Given a such a tree T_1, the trees $T_2, T_3, \cdots, T_{n-1}$ are completely determined. This implies that the number of distinct D-vines on n nodes is given by $\frac{n!}{2}$.
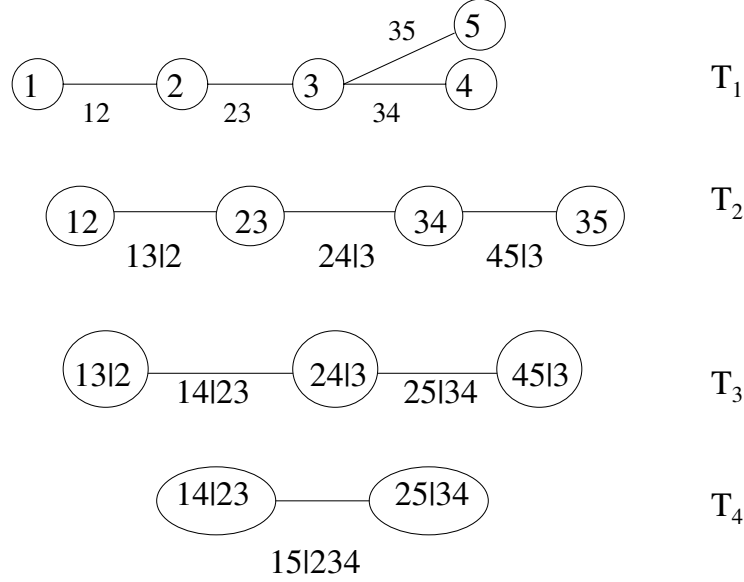
Figure 3. A regular vine with 5 variables, 4 trees and 10 edges.

2.6.   Multivariate Gaussian distribution

If the marginal distributions $f(x_i)$ in (10) are standard normal, and $c_{12}(\cdot)$, $c_{23}(\cdot)$ and $c_{13|2}(\cdot)$ are bivariate Gaussian copula densities (see Appendix B.1) the resulting distribution is trivariate standard normal with the positive definite correlation matrix

$$\begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix}.$$

Here ρ_{12} and ρ_{23} are the correlation parameters of copulae $c_{12}(\cdot)$ and $c_{23}(\cdot)$, respectively, while ρ_{13} is given by

$$\rho_{13} = \rho_{13|2}\sqrt{1-\rho_{12}^2}\sqrt{1-\rho_{23}^2} + \rho_{12}\rho_{23}.$$

The correlation parameter of copula $c_{13|2}(\cdot)$, $\rho_{13|2}$, is called partial correlation, see e.g. Kendall and Stuart (1967) for a definition. Note that it is only for the joint normal distribution that the partial and conditional correlations are equal (Morales et al., 2006). The meaning of partial correlation for non-normal variables is less clear.

3.   Conditional independence and the pair-copula decomposition

Assuming conditional independence may reduce the number of levels of the pair-copula decomposition, and hence simplify the construction. Let us first consider the three-dimensional case again with the pair copula decomposition in (10). If we assume that X_1 and X_3 are
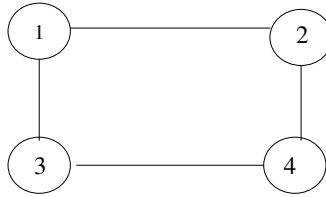
Figure 4. A conditional independence graph with 4 variables.

independent given X_2, we have that $c_{13|2}\left(F(x_1|x_2), F(x_3|x_2)\right) = 1$. Hence, the pair copula decomposition in (10) simplifies to

$$f(x_1, x_2, x_3) = f(x_1) \cdot f(x_2) \cdot f(x_3) \cdot c_{12}\left(F(x_1), F(x_2)\right) \cdot c_{23}\left(F(x_2), F(x_3)\right). \qquad (12)$$

In general, for any vector of variables $\boldsymbol{V}$ and two variables X, Y, it holds that X and Y are conditionally independent given $\boldsymbol{V}$ if and only if

$$c_{xy|\boldsymbol{v}}\left(F_{x|\boldsymbol{v}}(x|\boldsymbol{v}), F_{y|\boldsymbol{v}}(y|\boldsymbol{v})\right) = 1. \qquad (13)$$

As usual in hierarchical modelling, a model simplifies only if the initial factorisation of the joint density takes advantage of assumed conditional independence. For instance, if we use the decomposition $f(x_1, x_2, x_3) = f(x_2|x_1, x_3)f(x_1|x_3)f(x_3)$ in the case when X_1 and X_3 are conditionally independent given X_2, all pair-copulae are needed.

If the conditional independence assumption is only made to simplify the model construction, we may use the pair-copula decomposition to measure the approximation error introduced by this assumption. For example, take a four variable model, and assume conditional independence as expressed by the four variables in the conditional independence graph given in Figure 4. That is, variables X_1 and X_4 are assumed to be conditionally independent given X_3 and X_2, and variables X_2 and X_3 are assumed to be conditionally independent given X_1 and X_4. If we choose the decomposition in (11), the term $c_{14|23}\left(F(x_1|x_2, x_3), F(x_4|x_2, x_3)\right)$ should be equal to one, and the approximation error introduced by the conditional independence assumption is given by the difference $c_{14|23}\left(F(x_1|x_2, x_3), F(x_4|x_2, x_3)\right) - 1$.

4. Simulation from a pair-copulae decomposed model

Simulation from vines is briefly discussed in Bedford and Cooke (2001a), Bedford and Cooke (2001b), and Kurowicka and Cooke (2005). In this section we show that the simulation algorithms for canonical vines and D-vines are straightforward and simple to implement. In the rest of this paper we assume for simplicity that the margins of the distribution of interest are uniform.

The general algorithm for sampling n dependent uniform [0,1] variables is common for the canonical and the D-vine: First, sample $w_i; i = 1; \ldots n$ independent uniform on [0,1].

Then, set

$$
\begin{aligned}
x_1 &= w_1 \\
x_2 &= F_{2|1}^{-1}(w_2|x_1) \\
x_3 &= F_{3|1,2}^{-1}(w_3|x_1,x_2) \\
.. &= .... \\
x_n &= F_{n|1,2,\ldots,n-1}^{-1}(w_n|x_1,\ldots,x_{n-1}).
\end{aligned} \tag{14}
$$

To determine $F(x_j|x_1,x_2,\ldots,x_{j-1})$ for each j, we use the definition of the h-function in (7) and the relationship in (6), recursively for both vine structures. However, choice of the v_j variable in (6) is different for the canonical vines and D-vines. For the canonical vine we always choose

$$
F(x_j|x_1,x_2,\ldots,x_{j-1}) = \frac{\partial\, C_{j,j-1|1,\ldots j-2}(F(x_j|x_1,\ldots,x_{j-2}),F(x_{j-1}|x_1,\ldots,x_{j-2}))}{\partial F(x_{j-1}|x_1,\ldots,x_{j-2})},
$$

while for the D-vine we choose

$$
F(x_j|x_1,x_2,\ldots,x_{j-1}) = \frac{\partial\, C_{j,1|2,\ldots j-1}(F(x_j|x_2,\ldots,x_{j-1}),F(x_1|x_2,\ldots,x_{j-1}))}{\partial F(x_1|x_2,\ldots,x_{j-1})}.
$$

4.1. Sampling a canonical vine

Algorithm 1 gives the procedure for sampling from a canonical vine. The outer for-loop runs over the variables to be sampled. This loop consists of two other for-loops. In the first, the ith variable is sampled, while in the other, the conditional distribution functions needed for sampling the $(i+1)$th variable are computed. To compute these conditional distribution functions, we use the h-function, defined by (7) in Section 2, repeatedly with previously computed conditional distribution functions, $v_{i,j} = F(x_i|x_1,\ldots x_{j-1})$, as the first two arguments. The last argument of the h-function, $\Theta_{j,i}$, is the set of parameters of the corresponding copula density $c_{j,j+i|1,\ldots,j-1}(\cdot)$.

4.2. Sampling a D-vine

Algorithm 2 gives the procedure for sampling from the D-vine. It also consists of one main for-loop containing one for-loop for sampling the variables and one for-loop for computing the needed conditional distribution functions. However, this algorithm is computationally less efficient than that for the canonical vine, as the number of conditional distribution functions to be computed when simulating n variables is $(n-2)^2$ for the D-vine, while it is $(n-2)(n-1)/2$ for the canonical vine. Again the h-function is defined by (7) in Section 2, but here $\Theta_{j,i}$ is the set of parameters of the copula density $c_{i,i+j|i+1,\ldots,i+j-1}(\cdot)$.

4.3. Sampling a 3-dimensional vine

In this section, we describe how to sample from a 3-dimensional canonical vine. Since all decompositions in the 3-dimensional case are both a canonical vine and a D-vine, the resulting sample will also be a sample from a D-vine.

First, sample $w_i; i = 1,2,3$ independent uniform on [0,1]. Then, $x_1 = w_1$. Further, we have that $F(x_2|x_1) = h(x_2,x_1,\Theta_{11})$. Hence, $x_2 = h^{-1}(w_2,x_1,\Theta_{11})$. Finally,

---

Algorithm 1 Simulation algorithm for a canonical vine.
Generates one sample $x_1, \ldots x_n$ from the vine.

Sample $w_i; i = 1; \ldots n$ independent uniform on [0,1].

$x_1 = v_{1,1} = w_1$

for $i \leftarrow 2, 3, \ldots, n$

    $v_{i,1} = w_i$

    for $k \leftarrow i-1, i-2, \ldots, 1$

        $v_{i,1} = h^{-1}(v_{i,1}, v_{k,k}, \Theta_{k,i-k})$

    end for

    $x_i = v_{i,1}$

    if $i == n$ **then**

        Stop

    end if

    for $j \leftarrow 1, 2, \ldots, i-1$

        $v_{i,j+1} = h(v_{i,j}, v_{j,j}, \Theta_{j,i-j})$

    end for

end for

Algorithm 2 Simulation algorithm for D-vine.
Generates one sample $x_1, \ldots x_n$ from the vine.

Sample $w_i; i = 1; \ldots n$ independent uniform on [0,1].

$x_1 = v_{1,1} = w_1$

$x_2 = v_{2,1} = h^{-1}(w_2, v_{1,1}, \Theta_{1,1})$

$v_{2,2} = h(v_{1,1}, v_{2,1}, \Theta_{1,1})$

for $i \leftarrow 3, 4, \ldots n$

    $v_{i,1} = w_i$

    for $k \leftarrow i-1, i-2, \ldots, 2$

        $v_{i,1} = h^{-1}(v_{i,1}, v_{i-1,2k-2}, \Theta_{k,i-k})$

    end for

    $v_{i,1} = h^{-1}(v_{i,1}, v_{i-1,1}, \Theta_{1,i-1})$

    $x_i = v_{i,1}$

    if $i == n$ **then**

        Stop

    end if

    $v_{i,2} = h(v_{i-1,1}, v_{i,1}, \Theta_{1,i-1})$

    $v_{i,3} = h(v_{i,1}, v_{i-1,1}, \Theta_{1,i-1})$

    if $i > 3$ **then**

        for $j \leftarrow 2, 3, \ldots, i-2$

            $v_{i,2j} \quad = h(v_{i-1,2j-2}, v_{i,2j-1}, \Theta_{j,i-j})$

            $v_{i,2j+1} = h(v_{i,2j-1}, v_{i-1,2j-2}, \Theta_{j,i-j})$

        end for

    end if

    $v_{i,2i-2} = h(v_{i-1,2i-4}, v_{i,2i-3}, \Theta_{i-1,1})$

end for

$F(x_3|x_1, x_2) = h\left(F(x_3|x_1), F(x_2|x_1), \Theta_{21}\right) = h(h(x_3, x_1, \Theta_{12}), h(x_2, x_1, \Theta_{11}), \Theta_{21})$, meaning that $x_3 = h^{-1}\left(h^{-1}\left(w_3, h(x_2, x_1, \Theta_{11}), \Theta_{21}\right), x_1, \Theta_{12}\right)$.

5.   Inference for a speci£ed pair-copula decomposition

In this section we describe how the parameters of the canonical vine density given by (9) or D-vine density given by (8) can be estimated by maximum likelihood. Inference for a general regular vine (like the one in Figure 3) is also feasible, but the algorithm is not as straightforward. Assume that we observe n variables at T time points. Let $\boldsymbol{x}_i = (x_{i,1}, \ldots, x_{i,T}); i = 1, \ldots n$ denote the data set. Here, each random variable $X_{i,t}$ is assumed to be uniform in [0,1]. We assume for simplicity that the T observations of each variable are independent over time. This is not a necessary assumption, as stochastic dependencies and time series dynamics can easily be incorporated.

5.1.   Inference for a canonical vine

For the canonical vine, the log-likelihood is given by

$$\sum_{j=1}^{n-1}\sum_{i=1}^{n-j}\sum_{t=1}^{T}\log\left(c_{j,j+i|1,\ldots,j-1}\left(F(x_{j,t}|x_{1,t},\ldots,x_{j-1,t}), F(x_{j+i,t}|x_{1,t},\ldots,x_{j-1,t})\right)\right). \quad (15)$$

For each copula in the sum (15) there is at least one parameter to be determined. The number depends on which copula type is used. As before, the conditional distributions $F(x_{j,t}|x_{1,t}, \ldots, x_{j-1,t})$ and $F(x_{j+i,t}|x_{1,t}, \ldots, x_{j-1,t})$ are determined using the relationship in (6) and the definition of the h-function in (7). The log-likelihood must be numerically maximised over all parameters.

Algorithm 3 evaluates the likelihood for the canonical vine. The outer for-loop corresponds to the outer sum in (15). This for-loop consists in turn of two other for-loops. The first of these corresponds to the sum over i in (15). In the other, the conditional distribution functions needed for the next run of the outer for-loop are computed. Here $\Theta_{j,i}$ is the set of parameters of the corresponding copula density $c_{j,j+i|1,\ldots,j-1}(\cdot)$, $h(\cdot)$ is given by (7), and element t of $\boldsymbol{v}_{j,i}$ is $v_{j,i,t} = F(x_{i+j,t}|x_{1,t}, \ldots, x_{j,t})$. Further, $L(\boldsymbol{x}, \boldsymbol{v}, \Theta)$ is the log-likelihood of the chosen bivariate copula with parameters Θ given the data vectors $\boldsymbol{x}$ and $\boldsymbol{v}$. That is,

$$L(\boldsymbol{x}, \boldsymbol{v}, \Theta) = \sum_{t=1}^{T}\log\left(c(x_t, v_t, \Theta)\right), \quad (16)$$

where $c(u, v, \Theta)$ is the density of the bivariate copula with parameters Θ.

Starting values of the parameters needed in the numerical maximisation of the log-likelihood may be determined as follows:

(a) Estimate the parameters of the copulae in tree 1 from the original data
(b) Compute observations (i.e. conditional distrubution functions) for tree 2 using the copula parameters from tree 1 and the $h(\cdot)$ function.
(c) Estimate the parameters of the copulae in tree 2 using the observations from (b).
(d) Compute observations for tree 3 using the copula parameters at level 2 and the $h(\cdot)$ function.
(e) Estimate the parameters of the copulae in tree 3 using the observations from (d).

Algorithm 3 Likelihood evaluation for canonical vine

---
    log-likelihood $= 0$
    for $i \leftarrow 1, 2, \ldots, n$
        $\boldsymbol{v}_{0,i} = \boldsymbol{x}_i$.
    end for
    for $j \leftarrow 1, 2, \ldots, n-1$
        for $i \leftarrow 1, 2, \ldots, n-j$
            log-likelihood $=$ log-likelihood $+ L(\boldsymbol{v}_{j-1,1}, \boldsymbol{v}_{j-1,i+1}, \Theta_{j,i})$
        end for
        if $j == n-1$ **then**
            Stop
        end if
        for $i \leftarrow 1, 2, \ldots, n-j$
            $\boldsymbol{v}_{j,i} = h(\boldsymbol{v}_{j-1,i+1}, \boldsymbol{v}_{j-1,1}, \Theta_{j,i})$
        end for
    end for

  (f) etc.

Note that each estimation here is easy to perform, since the data set is only of dimension 2.

5.2. Inference for a D-vine

For the D-vine, the log-likelihood is given by

$$\sum_{j=1}^{n-1} \sum_{i=1}^{n-j} \sum_{t=1}^{T} \log\left(c_{i,i+j|i+1,\ldots,i+j-1}\left(F(x_{i,t}|x_{i+1,t}, \ldots, x_{i+j-1,t}), F(x_{i+j,t}|x_{i+1,t}, \ldots, x_{i+j-1,t})\right)\right).$$
(17)

The D-vine log-likelihood must also be numerically optimised. Algorithm 4 evaluates the likelihood. $\Theta_{j,i}$ is the set of parameters of copula density $c_{i,i+j|i+1,\ldots,i+j-1}(\cdot)$.

5.3. Inference for a 3 variable model

In the special case of a 3-dimensional data set with $U[0,1]$ distributed variables, (15) reduces to

$$\sum_{t=1}^{T} \left[\log c_{12}(x_{2,t}, x_{1,t}, \Theta_{11}) + \log c_{23}(x_{2,t}, x_{3,t}, \Theta_{12}) + \log c_{13|2}(v_{1,t}, v_{2,t}, \Theta_{21})\right],$$
(18)

where

$$v_{1,t} = F(x_{1,t}|x_{2,t}) = h(x_{1,t}, x_{2,t}, \Theta_{11})$$

and

$$v_{2,t} = F(x_{3,t}|x_{2,t}) = h(x_{3,t}, x_{2,t}, \Theta_{12})$$

    The parameters to be estimated are $\boldsymbol{\Theta} = (\Theta_{11}, \Theta_{12}, \Theta_{21})$, where $\Theta_{j,i}$ is the set of parameters of the corresponding copula density $c_{i,i+j|i+1,\ldots,i+j-1}(\cdot)$. Following the procedure

---

Algorithm 4 Likelihood evaluation for a D-vine

log-likelihood $= 0$
for $i = 1, 2, \ldots, n$
    $\boldsymbol{v}_{0,i} = \boldsymbol{x}_i.$
end for
for $i = 1, 2, \ldots, n - 1$
    log-likelihood=log-likelihood$+L(\boldsymbol{v}_{0,i}, \boldsymbol{v}_{0,i+1}, \Theta_{1,i})$
end for
$\boldsymbol{v}_{1,1} = h(\boldsymbol{v}_{0,1}, \boldsymbol{v}_{0,2}, \Theta_{1,1})$
for $k = 1, 2, \ldots, n - 3$
    $\boldsymbol{v}_{1,2k} = h(\boldsymbol{v}_{0,k+2}, \boldsymbol{v}_{0,k+1}, \Theta_{1,k+1})$
    $\boldsymbol{v}_{1,2k+1} = h(\boldsymbol{v}_{0,k+1}, \boldsymbol{v}_{0,k+2}, \Theta_{1,k+1})$
end for
$\boldsymbol{v}_{1,2n-4} = h(\boldsymbol{v}_{0,n}, \boldsymbol{v}_{0,n-1}, \Theta_{1,n-1})$
for $j = 2, \ldots, n - 1$
    for $i = 1, 2, \ldots, n - j$
        log-likelihood=log-likelihood$+L(\boldsymbol{v}_{j-1,2i-1}, \boldsymbol{v}_{j-1,2i}, \Theta_{j,i})$
    end for
    if $j == n - 1$ **then**
        Stop
    end if
    $\boldsymbol{v}_{j,1} = h(\boldsymbol{v}_{j-1,1}, \boldsymbol{v}_{j-1,2}, \Theta_{j,1})$
    if $n > 4$ **then**
        for $i = 1, 2, \ldots, n - j - 2$
            $\boldsymbol{v}_{j,2i} = h(\boldsymbol{v}_{j-1,2i+2}, \boldsymbol{v}_{j-1,2i+1}, \Theta_{j,i+1})$
            $\boldsymbol{v}_{j,2i+1} = h(\boldsymbol{v}_{j-1,2i+1}, \boldsymbol{v}_{j-1,2i+2}, \Theta_{j,i+1})$
        end for
    end if
    $\boldsymbol{v}_{j,2n-2j-2} = h(\boldsymbol{v}_{j-1,2n-2j}, \boldsymbol{v}_{j-1,2n-2j-1}, \Theta_{j,n-j})$
end for

---

described in Section 5.1, we first estimate the parameters of the three copulae involved by a sequential procedure, and then we maximize the full log-likelihood using the parameters obtained from the stepwise procedure as starting values.

6. A catalogue of some pair-copulae models

Pair-copulae are basic building blocs for constructing multivariate models. In this section we present four of the most common pair-copulae. We discuss in particular their parameterisation and the type of pair dependence they capture. See Joe (1997) for an overview of other copulae. Tail dependence properties are particularly important in many applications that rely on non-normal multivariate families (Joe, 1996). The four bivariate copulae in question are the Gaussian, the Student's t, the Clayton and the Gumbel copula. The first two are copulae of normal mixture distributions. They are so-called implicit copulae because they do not have a simple closed form. Clayton and Gumbel are Archimedean copulae, for which the distribution function has a simple closed form. These four types of copulae have different strength of dependence in the tails of the bivariate distribution. This can be represented by the probability that the first variable exceeds its q-quantile, given that the other exceeds its own q-quantile. The limiting probability, as q goes to infinity, is called the *upper tail dependence coefficient* (Embrechts et al., 2001) , and a copula is said to be upper tail dependent if this limit is not zero. The *lower tail dependence coefficient* is defined analogously. The Clayton copula is lower tail dependent, but not upper. The Gumbel copula is upper tail dependent, but not lower. The Student's t-copula is both lower and upper tail dependent, while the Gaussian is neither lower nor upper tail dependent.

In Appendix B we give three important formulas for each of these four pair copulae; the density, the h-function and the inverse of the h-function. The Gaussian, Clayton and Gumbel pair-copulae have one parameter, while the Student's-t pair-copula has two. The additional parameter of the latter is the degrees of freedom, that controls the strength of dependence in the tails of the bivariate distribution. The Student's t-copula allows for joint extreme events, either in both bivariate tails or none of them. If one believes that the variables are only lower tail dependent, then the Clayton copula, which is an asymmetric copula, exhibiting greater dependence in the negative tail than in the positive, might be a better choice. The Gumbel copula is also an asymmetric copula, but it is exhibiting greater dependence in the positive tail than in the negative. Figure 5 shows the densities of the four copulae for three different parameter settings.

For all these four pair-copulae the h-function is given by an explicit analytical expression, see Appendix B. This expression can be analytically inverted for all pair-copulae except for the Gumbel, where numerical inversion is necessary. Explicit availability of the h-functions and their inverse is very important for the efficiency of our estimation procedures.
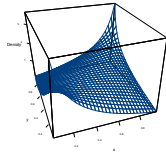
7. Model selection

In Section 5 we described how to do inference for a specific pair-copula decomposition. However, this is only a part of the full estimation problem. Full inference for a pair-copula decomposition should in principle consider (a) the selection of a specific factorisation, (b) the choice of pair-copula types, and (c) the estimation of the copula parameters. For smaller dimensions (say 3 and 4), one may estimate the parameters of all possible decompositions using the procedure described in Section 5 and compare the resulting log-likelihoods. This
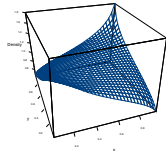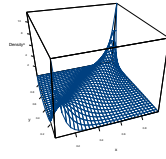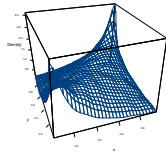
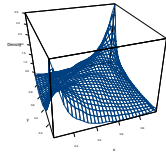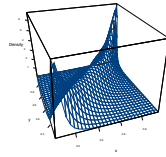Gaussian, $\rho = 0.2$ Gaussian, $\rho = 0.5$ Gaussian, $\rho = 0.9$

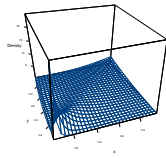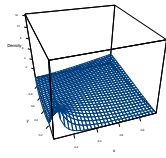Student's t, $\rho = 0.5$, $\nu = 6$ Student's t, $\rho = 0.5$, $\nu = 3$ Student's t, $\rho = 0.9$, $\nu = 3$
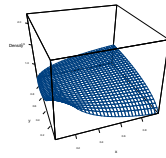
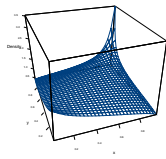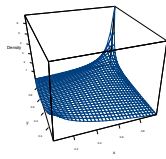Clayton, $\delta = 3$ Clayton, $\delta = 1.5$ Clayton, $\delta = 0.2$

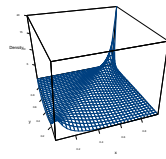Gumbel, $\delta = 1.2$ Gumbel, $\delta = 1.5$ Gumbel, $\delta = 3$

Figure 5. 3D surface plot of bivariate density for Gaussian, Student's t, Clayton and Gumbel copulae, with three different parameter settings.

is in practice infeasible for higher dimensions, since the number of possible decompositions increases very rapidly with the dimension of the data set, as shown in Section 2. One should instead consider which bivariate relationships that are most important to model correctly, and let this determine which decomposition(s) to estimate. D-vines are more flexible than canonical vines, since for the canonical vines we specify the relationships between one specific pilot variable and the others, while in the D-vine structure we can select more freely which pairs to model.

Given data and an assumed pair-copula decomposition, it is necessary to specify the parametric shape of each pair-copula. For example, for the decomposition in Section 5.3 we need to decide which copula type to use for $C_{12}(\cdot,\cdot)$, $C_{23}(\cdot,\cdot)$ and $C_{13|2}(\cdot,\cdot)$ (for instance among the ones described in Section 6). The pair-copulae do not have to belong to the same family. The resulting multivariate distribution will be valid if we choose for each pair of variables the parametric copula that best fits the data. If we choose not to stay in one predefined class, we need a way of determining which copula to use for each pair of (transformed) observations. We propose to use a modified version of the sequential estimation procedure outlined in Section 5.1:

(a) Determine which copula types to use in tree 1 by plotting the original data, or by applying a Goodness-of-Fit (GoF) test, see Section 7.1.
(b) Estimate the parameters of the selected copulae using the original data.
(c) Transform observations as required for tree 2, using the copula parameters from tree 1 and the $h(\cdot)$ function as shown in Sections 5.1 and 5.2.
(d) Determine which copula types to use in tree 2 in the same way as in tree 1.
(e) Iterate.

The observations used to select the copulae at a specific level depend on the specific pair-copulae chosen up-stream in the decomposition. This selection mechanism does not guarantee an globally optimal fit. Having determined the appropriate parametric shapes for each copulae, one may use the procedures in Section 5 to estimate their parameters.

7.1.   Godness-of-£t

To verify whether the dependency structure of a data set is appropriately modelled by a chosen pair-copula decomposition, we need a goodness-of-fit (GOF) test. GOF tests for dependency structures are basically special cases of the more general problem of testing multivariate densities. However, it is more technically complicated as the univariate distribution functions are unknown. Hence, despite an obvious need for such tests in applied work, relatively little is known about their properties, and there is still no recommended method agreed upon.

Of the tests that have been proposed, quite a few are based on the probability integral transform (PIT) of Rosenblatt (1952). The PIT converts a set of dependent variables into a new set of variables which are independent and uniform under the null hypothesis that the data originate from a given multivariate distribution. The technique is somehow the inverse of simulation and is defined as follows.

Let $\boldsymbol{X} = (X_1, \ldots X_n)$ denote a random vector with marginal distributions $F_i(x_i)$ and conditional distributions $F_{i|1,\ldots i-1}(x_i|x_1,\ldots x_{i-1})$, for $i = 1,\ldots n$. The PIT of $\boldsymbol{X}$ is defined

as $T(\boldsymbol{X}) = (T_1(X_1), \dots T_n(X_n))$, where $T_i(X_i)$ is given by:

$$
\begin{aligned}
T_1(X_1) &= F_1(x_1) \\
T_2(X_2) &= F_{2|1}(x_2|x_1) \\
\cdot &= \cdot \\
\cdot &= \cdot \\
T_n(X_n) &= F_{n|1,\dots n-1}(x_n|x_1, \dots x_{n-1}).
\end{aligned}
\tag{19}
$$

$$\tag{20}$$

The random variables $Z_i = T_i(X_i), i = 1, \dots n$ are independent and uniformly distributed on $[0,1]^n$ under the null hypothesis that $\boldsymbol{X}$ comes from the multivariate model used to compute the PIT of $\boldsymbol{X}$. It is relatively easy to specialise the PIT to the pair-copula decomposition. Algorithms 5 and 6 give the procedures for a canonical vine and a D-vine, respectively.

Algorithm 5 PIT algorithm for a canonical vine

$\quad$**for** $t \leftarrow 1, 2, \dots, T$
$\qquad z_{1,t} = x_{1,t}$
$\qquad$**for** $i \leftarrow 2, 3, \dots, n$
$\qquad\quad z_{i,t} = x_{i,t}$
$\qquad\quad$**for** $j \leftarrow 1, 2, \dots i-1$
$\qquad\qquad z_{i,t} = h(z_{i,t}, z_{j,t}, \Theta_{j,i-j})$
$\qquad\quad$**end for**
$\qquad$**end for**
$\quad$**end for**

Having performed the probability integral transform, the next step is to verify whether the resulting variables really are independent and uniform in [0,1]. The most common approach is to compute $S = \sum_{i=1}^{n} \left(\Phi^{-1}(Z_i) \right)^2$, and test whether the observed values of S come from a chi-square distribution with n degrees of freedom. The Anderson-Darling goodness-of-fit test may be applied for the latter, see e.g. Breymann et al. (2003).

8.　Application: Financial returns

8.1.　Data set

In this section, we study four time series of daily data: the Norwegian stock index (TOTX), the MSCI world stock index, the Norwegian bond index (BRIX) and the SSBWG hedged bond index, for the period from 04.01.1999 to 08.07.2003. Figure 6 shows the log-returns of each pair of assets. The four variables will be denoted T, M, B and S.

We want to compare a four-dimensional pair-copula decomposition with Student's t-copulae for all pairs with the four-dimensional Student's t-copula. The n-dimensional Student's t-copula has been used repeatedly for modelling multivariate financial return data. A number of papers, such as Mashal and Zeevi (2002), have shown that the fit of this copula is generally superior to that of other n-dimensional copulae for such data. However, the Students's t-copula has only one parameter for modelling tail dependence, independent of dimension. Hence, if the tail dependence of different pairs of risk factors in a portfolio are very different, we believe the pair-copula decomposition with Student's t-copulae for all

---

Algorithm 6 PIT algorithm for a D-vine

---

for $t \leftarrow 1, 2, \ldots, T$
    $z_{1,t} = x_{1,t}$
    $z_{2,t} = h(x_{2,t}, x_{1,t}, \Theta_{1,1})$
    $v_{2,1} = x_{2,t}$
    $v_{2,2} = h(x_{1,t}, x_{2,t}, \Theta_{1,1})$
    for $i \leftarrow 3, 4, \ldots n$
        $z_{i,t} = h(x_{i,t}, x_{i-1,t}, \Theta_{1,i-1})$
        for $j \leftarrow 2, 3, \ldots i - 1$
            $z_{i,t} = h(z_{i,t}, v_{i-1,2(j-1)}, \Theta_{j,i-j})$
        end for
        if $i == n$ **then**
            Stop
        end if
        $v_{i,1} = x_{i,t}$
        $v_{i,2} = h(v_{i-1,1}, v_{i,1}, \Theta_{1,i-1})$
        $v_{i,3} = h(v_{i,1}, v_{i-1,1}, \Theta_{1,i-1})$
        for $j \leftarrow 1, 2, \ldots, i - 3$
            $v_{i,2j+2} = h(v_{i-1,2j}, v_{i,2j+1}, \Theta_{j+1,i-j-1})$
            $v_{i,2j+3} = h(v_{i,2j+1}, v_{i-1,2j}, \Theta_{j+1,i-j-1})$
        end for
        $v_{i,2i-2} = h(v_{i-1,2i-4}, v_{i,2i-3}, \Theta_{i-1,1})$
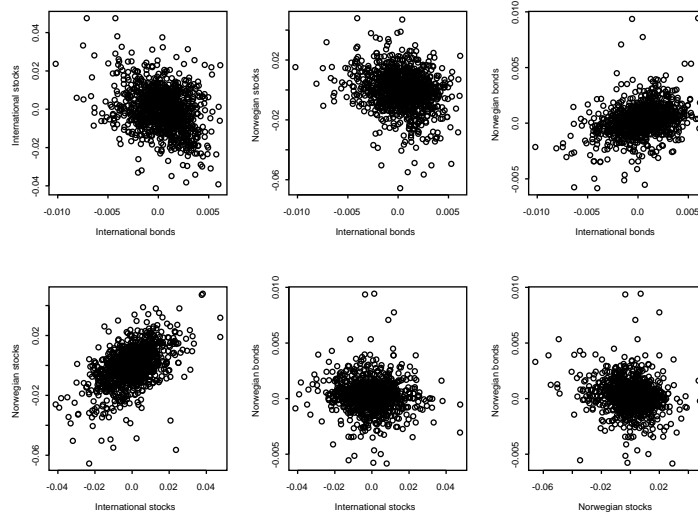   **end for**
end for

Figure 6. Log-returns for pairs of assets during the period from 04.01.1999 to 08.07.2003.
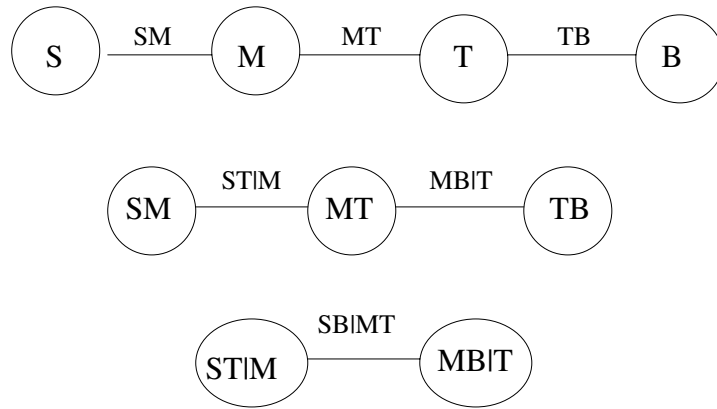
Figure 7. Selected D-vine structure for the data set in Section 8.1.

pairs to describe the dependence structure of the risk factors better. Since we are mainly interested in estimating the dependence structure of the risk factors, the original data vectors are converted to uniform variables using the empirical distribution functions before further modelling.

8.2. Selecting an appropriate pair-copula decomposition

Having decided to use Student's t-copulae for all pairs of the decomposition, the next step is to choose the most appropriate ordering of the risk factors. This is done by first fitting a bivariate Student's t-copula to each pair of risk factors, obtaining an estimated degree of freedom for each pair. The estimation of the Student's t-copula parameters requires numerical optimisation of the log-likelihood function, see for instance Mashal and Zeevi (2002) or Demarta and McNeil (2005).

Having fitted a bivariate Student's t-copula to each pair, the risk factors are ordered such that the three copulae to be fitted in tree 1 in the pair-copula decomposition are those corresponding to the three smallest numbers of degrees of freedom. A low number of degrees of freedom indicates strong dependence. The number of degrees of freedom parameters from our case are shown in Table 1. The dependence is strongest between international bonds and stocks (S and M), international and Norwegian stocks (M and T), and Norwegian stocks and bonds (T and B). Hence, we want to fit the copulae $C_{S,M}$, $C_{M,T}$ and $C_{T,B}$ in tree 1 of the vine. This means that we can not use a canonical vine, since there is no pilot variable. However, using a D-vine specification with the nodes S, M, T, and B in the listed order, gives the three above-mentioned copulae at level 1. See Figure 7 for the whole D-vine structure in this case.

8.3. Inference

The parameters of the D-vine are estimated using the algorithm in Section 5.2. For each pair-copula, the log-likelihood is computed using (16) and the density and the h-function

Table 1. Estimated numbers of degrees of freedom for bivariate Student's t-copula for pairs of variables.

Between	M	T	B
S	4.11	32.74	9.98
M		8.03	13.53
T			7.96

Table 2. Estimated parameters for four-dimensional pair-copula decomposition.

Param	*Start*	*Final*
ρ_{SM}	-0.25	-0.27
ρ_{MT}	0.51	0.52
ρ_{TB}	-0.18	-0.18
$\rho_{ST|M}$	-0.09	-0.08
$\rho_{MB|T}$	0.02	0.02
$\rho_{SB|MT}$	0.33	0.33
ν_{SM}	4.11	4.21
ν_{MT}	8.03	8.32
ν_{TB}	7.96	7.41
$\nu_{ST|M}$	300.00	300.00
$\nu_{MB|T}$	300.00	300.00
$\nu_{SB|MT}$	9.37	9.44
log.likelih.	344.37	344.54

for the Student's t-copula given in Appendix B.2.

Table 2 shows the starting values obtained using the sequential estimation procedure (left column), and the final parameter values together with the corresponding log-likelihood values. In the numerical search for the degrees of freedom parameter we have used 300 as the maximum value. As can be seen from the table, the likelihood slightly increases when estimating all parameters simultaneously. The Akaike's Information Criterion (AIC) for the final model is -665.08. The p-value for the goodness-of-fit test described in Section 7.1 was 0.70, meaning that we cannot reject the null hypothesis of a D-vine. It should be noted that using the empirical distribution functions to convert the original data vectors to uniform variables before fitting the dependency structure, will affect the critical values of this test in a complicated, non-trivial way. This is still an unsolved problem, not only for a pair-copula decomposition, but for copulae in general.

8.4. Validation by simulation

Having estimated the pair-copula decomposition, it is interesting to investigate the bivariate distributions of the pairs of variables which were not explicitly modelled in the decomposition. We sample from the estimated pair-copula decomposed model, with estimated parameters as above, and check if simulated values and observed data have similar features.

We used the simulation procedure described in Section 4.2 and the h-function and its inverse given in Section B.2 to generate a set of 10,000 samples from the estimated pair-copula decomposition. Then, we estimated pairwise Student's t-copulae for all bivariate margins. The results are shown in Table 3. Comparing these to the ones in Table 1 we see

Table 3. Estimated numbers of degrees of freedom for bivariate Student's t-copula for pairs of simulated variables.

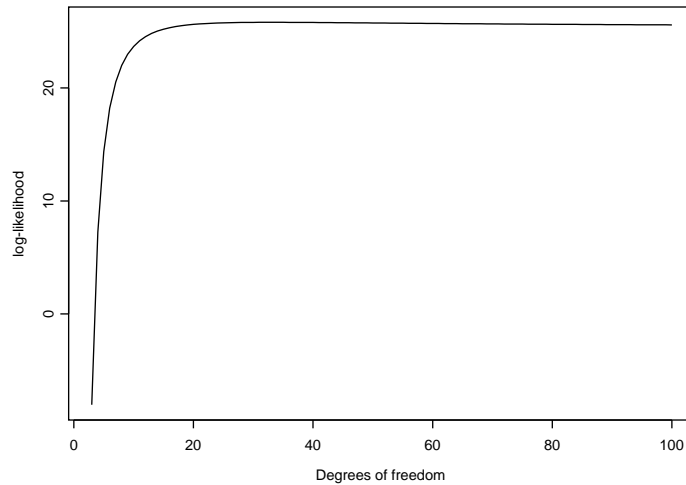| *Between*	M	T	B
S	4.06	12.10	11.44
M		8.10	15.21
T			6.87

Figure 8. The log-likelihood of the copula density for the margin ST as a function of ν_{ST} for ρ_{ST} £xed to -0.21.

that except for the pair S, T, the numbers are similar, indicating that all dependencies are quite well captured, also those that are not directly modelled.

In Figure 8, we show the log-likelihood of the pair-copula density for the pair S, T as a function of ν_{ST}, for ρ_{ST} fixed to -0.21. As can be seen from the figure, the likelihood is very flat for a large range of ν_{ST} values, meaning that the difference between $\nu_{ST} = 12.10$ and $\nu_{ST} = 32.74$ is not as large as it may seem at first sight. This is verified by Figure 9, which shows the contours of the densities of the copulae for the pairs S, T, S, B, and M, B, estimated both from the simulated D-vine and from the original data. The plots in the two columns are very similar.

8.5. Comparison with the four-dimensional Student's t-copula

In this section we compare the results obtained with the pair-copula decomposition from Section 8.3 with those obtained with a four-dimensional Student's t-copula. The parameters of the Student's t-copula are estimated using the method described in Mashal and Zeevi (2002) and (Demarta and McNeil, 2005). They are shown in Table 4. The AIC for this model is -632.18, i.e. higher than that for the pair-copulae decomposition. Further, all condi-
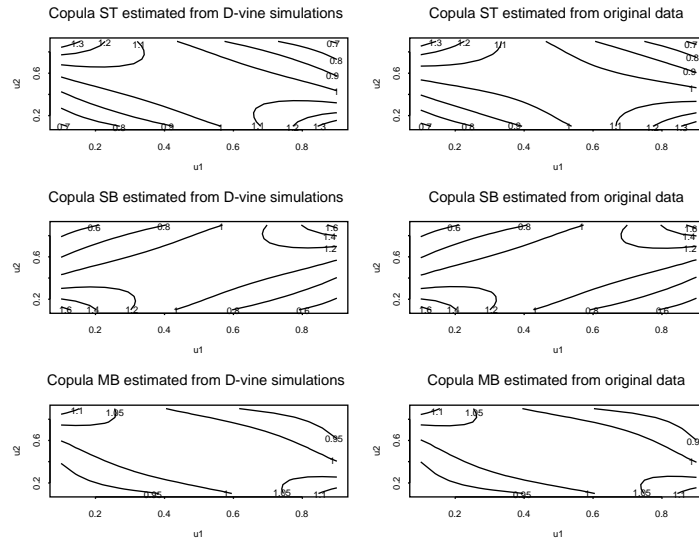
Figure 9. Contours of the densities of the copulae for the bivariate margins ST, SB, and MB, estimated both from the simulated D-vine and from the original data.

tional distributions of a multivariate Students's t-distribution are Students's t-distributions. Hence, the n-dimensional Student's t-copula is a special case of an n-dimensional D-vine with the needed pairwise copulae in the D-vine structure set to the corresponding conditional bivariate distributions of the multivariate Student's t-distribution. Therefore, the four-dimensional Students's t-copula is nested within the considered D-vine structure and the likelihood ratio test statistic is 344.54-323.09=21.45 with 12-7=5 degrees of freedom. This yields a p-value of 0.0007 and shows that the 4 dimensional Student's t-copula can be rejected in favour of the D-vine.

To illustrate the difference between the four-dimensional Students's t-copula and the four-dimensional pair-copula decomposition, we have computed the tail dependence coefficients for the three bivariate margins SM, MT and TB for both structures. See Section 6 for the definition of the upper and lower tail dependence coefficients. For the Student's t-copula, the two coefficients are equal and given by (Embrechts et al., 2001)

$$\lambda_l(X,Y) = \lambda_u(X,Y) = 2\,t_{\nu+1}\left(-\sqrt{\nu+1}\sqrt{\frac{1-\rho}{1+\rho}}\right),$$

where $t_{\nu+1}$ denotes the distribution function of a univariate Student's t-distribution with $\nu+1$ degrees of freedom. Table 5 shows the tail dependency coefficients for the three margins and both structures. For the bivariate margin SM, the value for the pair-copula distribution is 29 times higher than the corresponding one for the Student's t-copula, and it is also significantly higher for the two other bivariate margins. For a trader holding a portfolio of international stocks and bonds, the practical implication of this difference in tail dependence is that the probability of observerving a large portfolio loss is much higher for the four-dimensional pair-copula decomposition that it is for the four-dimensional Student's t-copula.

Table 4. Estimated parameters for four-dimensional Student's t-copula.

Param	Value
ρ_{SM}	-0.25
ρ_{ST}	-0.21
ρ_{SB}	0.34
ρ_{MT}	0.51
ρ_{MB}	-0.07
ρ_{TB}	-0.18
ν_{STMB}	10.05
log.likelih.	323.09

Table 5. Tail dependence coef£cients.

Margin	Pair-copula decomp.	Student's t-copula
SM	0.029	0.001
MT	0.119	0.086
TB	0.008	0.002

8.6. *Pair-copula decomposition with copulae from different families*

In this section we investigate whether we would get an even better fit for our data set if we allowed the pair-copulae in the decomposition defined by Figure 7 to come from different families. Figure 10 shows the data sets used to estimate the six pair-copulae in the decomposition described in Sections 8.2 and 8.3. The three scatter plots in the upper row correspond to the three bivariate margins SM, MT and TB. The data clustering in the two opposite corners of these plots is a strong indication of both upper and lower tail dependence, meaning that the Student's t-copula is an appropriate choice. In the two leftmost scatter plots in the lower row, the data seems to have no tail dependence and the two margins also appear to be uncorrelated. This is in accordance with the parameters estimated for these data sets, $\rho_{ST|M}$, $\rho_{MB|T}$, $\nu_{ST|M}$, $\nu_{MB|T}$, shown in Table 2. The correlation parameters are very close to 0 and the degrees of freedom parameters are very large, meaning that the two variables constituting each pair are close to being independent. If so, $c_{ST|M}(\cdot)$ and $c_{MB|T}(\cdot)$ are both 1, which means that the pair-copula construction defined by Figure 7 may be simplified to

$$c_{SM}(x_S, x_M) \, c_{MT}(x_M, x_T) \, c_{TB}(x_T, x_B) \, c_{SB|MT} \left(F(x_S|x_M), F(x_B|x_T) \right).$$

In the scatter plot in the lower right corner of the figure, there seems to be data clustering in the lower left corner, but not in the upper right. This indicates that the Clayton copula might be a better choice than the Student's t-copula, since it has lower tail dependence, but not upper. Hence, we have fitted the Clayton copula to this data set. The parameter was estimated to $\delta = 0.425$. The likelihood of the Clayton copula is lower than that of the Student's t-copula (59.86 vs. 68.62). However, since the two copulae are non-nested we cannot really compare the likelihoods. Instead we have used the procedure suggested by Genest and Rivest (1993) for identifying the appropriate copula. According to this procedure, we examine the degree of closeness of the parametric and non-parametric versions of the distribution function $K(z)$, defined by
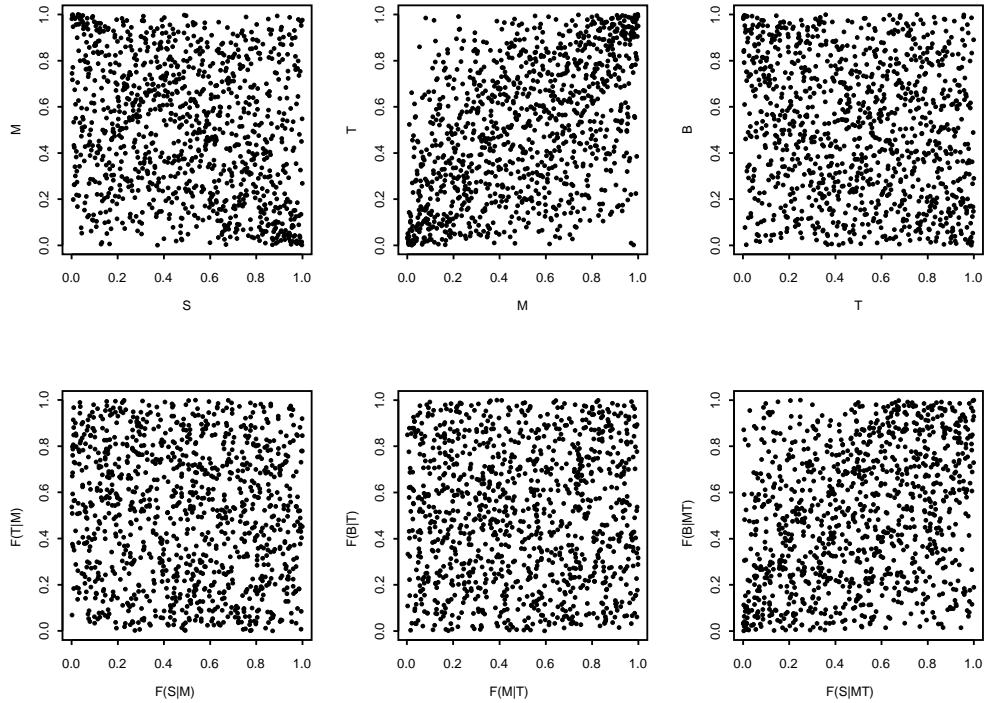
$$K(z) = P(C(u_1, u_2) \le z).$$

Figure 10. The data sets used to estimate the six pair-copulae in the decomposition described in Sections 8.2 and 8.3.

For Archimedean copulae, $K(z)$ is given by an explicit expression, while for the Student's t-copula is has to be numerically derived. In Figure 11 we have plotted the quantiles of the non-parametric and parametric estimates of $K(z)$. The solid line corresponds to the case where the quantiles are equal, the dashed line to the Clayton copula, and the dotted line to the Student's t-copula having parameters $\rho_{SB|MT}$ and $\nu_{SB|MT}$ given in Table 2. Since it is difficult to distinguish the different lines in Figure 11, we have also made a plot of the difference between the parametric and the non-parametric $K(z)$ for the two copulae, shown in Figure 12. The solid line corresponds to the Clayton copula and the dotted line to the Student's t-copula. As can be seen from this figure, the quantiles of the Student's t-copula are closer to the non-parametric quantiles than those of the Clayton copula. Hence, the pair-copula decomposition with Student's t-copulae for all pairs remains the best choice for our data set.

9. Conclusions

We have shown how multivariate data exhibiting complex patterns of dependence in the tails can be modelled using pair-copulae. We have developed algorithms that allow inference on the parameters of the pair-copulae on the various levels of the construction. This
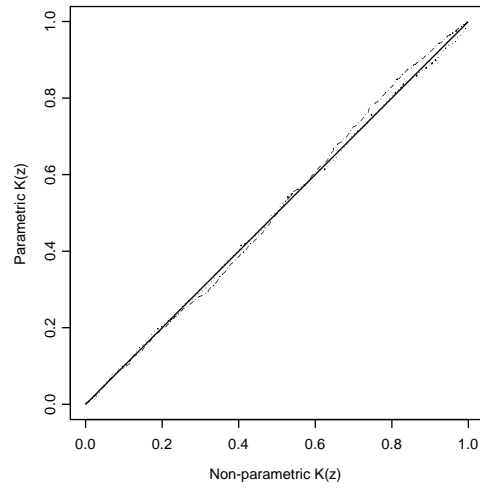
Figure 11. QQ-plot corresponding to the non-parametric and parametric estimates of $K(z)$. The solid line corresponds to the case where the quantiles are equal, the dashed line to the Clayton copula, and the dotted line to the Student's t-copula having parameters $\rho_{SB|MT}$ and $\nu_{SB|MT}$ given in Table 2.
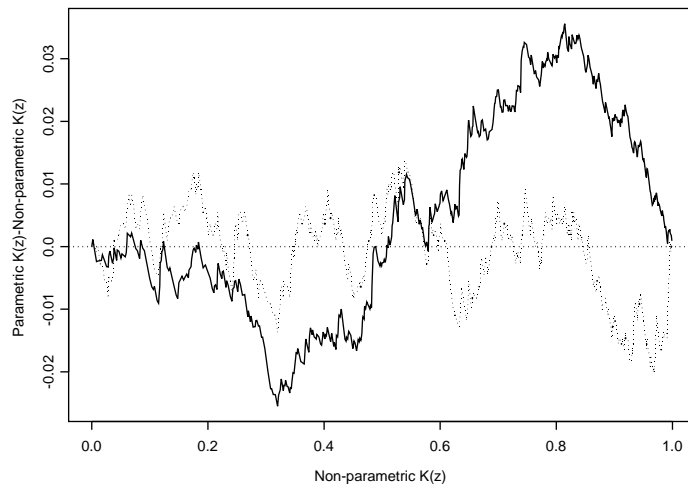

Figure 12. The difference between the non-parametric and parametric estimates of $K(z)$. The solid line corresponds to the Clayton copula, and the dotted line to the Student's t-copula having parameters $\rho_{SB|MT}$ and $\nu_{SB|MT}$ given in Table 2.

construction is hierarchical in nature, the various levels standing for growing conditioning sets, incorporating more variables. This differs from traditional hierarchical models, where levels depict conditional independence. Pair-copulae are simple building blocs, which can be compared to pairwise interaction potentials or cliques in Gibbs fields.

While we have assumed that the observations of the multivariate data are independent, this is not necessary. Pair-copula models can be built and estimated whenever there exists a likelihood function for the data, for example for ARMA or ARCH/GARCH. In fact, various forms of pseudo-maximum likelihood (Bollerslev and Wooldridge, 1992) can also be used instead of full likelihoods. Missing values are acceptable, though likelihoods become more complex, as in any other model. Bayesian versions of inference are easy to imagine, as there is no difficulty in adding priors on the parameters of the pair-copulae. One may also put priors on the choice of pairs to match. Posterior estimates will then substitute maximum likelihood ones.

Further research is needed to produce better comparison methods between alternative pair-copulae and between alternative decompositions. More powerful goodness-of-fit tests for bivariate models are crucial for the construction of an unsupervised algorithm that explores the large space of possible pair-copulae models. This however remains a central aim, since there is an increasing tendency to collect huge quantities of multivariate and dense observations, requiring automatic inferential methods.

Acknowledgements

References

Bedford, T. and R. M. Cooke (2001a). Monte Carlo simulation of vine dependent random variables for applications in uncertainty analysis. In *2001 Proceedings of ESREL2001, Turin, Italy.*

Bedford, T. and R. M. Cooke (2001b). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence 32*, 245–268.

Bedford, T. and R. M. Cooke (2002). Vines - a new graphical model for dependent random variables. *Annals of Statistics 30*(4), 1031–1068.

Bollerslev, T. and J. Wooldridge (1992). Quasi maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Economic Reviews 11*, 143–172.

Breymann, W., A. Dias, and P. Embrechts (2003). Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance 1*, 1–14.

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

Embrechts, P., F. Lindskog, and A. McNeil (2003). Modelling dependence with copulas and applications to risk management. In S.T.Rachev (Ed.), *Handbook of Heavy Tailed Distributions in Finance*. North-Holland: Elsevier.

Embrechts, P., A. J. McNeil, and D. Straumann (2001). Correlation and dependency in risk management: Properties and pitfalls. In *Value at Risk and Beyond*. Cambridge University Press.

Genest, C. and L. Rivest (1993). Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association 88*, 1034–1043.

Green, P., N. L. Hjort, and S. Richardson (2003). *Highly Structured Stochastic Systems.* Oxford: Oxford University Press.

Joe, H. (1996). Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters. In L. Rüschendorf and B. Schweizer and M. D. Taylor (Ed.), *Distributions with Fixed Marginals and Related Topics.*

Joe, H. (1997). *Multivariate Models and Dependence Concepts.* London: Chapman & Hall.

Kendall, M. and A. Stuart (1967). *The Advanced Theory of Statistics, Vol.2. Inference and relationship, 2nd ed.* Griffin, London.

Kurowicka, D. and R. M. Cooke (2004). Distribution - free continuous bayesian belief nets. In *Fourth International Conference on Mathematical Methods in Reliability Methodology and Practice, Santa Fe, New Mexico.*

Kurowicka, D. and R. M. Cooke (2005). Sampling algorithms for generating joint uniform distributions using the vine - copula method. In *3rd IASC world conference on Computational Statistics & Data Analysis, Limassol, Cyprus.*

Mashal, R. and A. Zeevi (2002). Beyond correlation: Extreme co-movements between financial assets. Technical report, Columbia University.

McNeil, A. J., R. Frey, and P. Embrechts (2006). *Quantitative Risk Management: Concepts, Techniques and Tools.* Princeton University Press.

Morales, O., D. Kurowicka, and A. Roelen (2006). Elicitation procedures for conditional and unconditional rank correlations. In *Resources for the future, Expert Judgement Policy Symposium and Technical Workshop, Washington, DC, March 13-14.*

Rosenblatt, M. (1952). Remarks on multivariate transformation. *Ann. Math. Statist. 23*, 1052–1057.

Savu, C. and M. Trede (2006). Hierarchical Archimedean copulas. In *International Conference on High Frequency Finance, Konstanz, Germany, May.*

Sklar, A. (1959). Fonctions dé repartition á n dimensions et leurs marges. *Publ. Inst. Stat. Univ. Paris 8*, 229–231.

Venter, G. G. (2001). Tails of copulas. In *Proceedings ASTIN Washington, USA*, pp. 68–113.

A. Four-dimensional pair-copula decompositions

Assume that we decompose a given four-dimensional density $f_{1234}(x_1, x_2, x_3, x_4)$ as follows

$$f_{1234}(x_1, x_2, x_3, x_4) = f_1(x_1) \cdot f_{2|1}(x_2|x_1) \cdot f_{3|12}(x_3|x_1, x_2) \cdot f_{4|132}(x_4|x_1, x_3, x_2). \quad (21)$$

We have that

$$f_{2|1}(x_2|x_1) = c_{12}(F_1(x_1), F_2(x_2)) \cdot f_2(x_2),$$

and

$$
\begin{aligned}
f_{3|12}(x_3|x_1, x_2) &= \frac{f_{23|1}(x_2, x_3|x_1)}{f_{2|1}(x_2|x_1)} \\
&= \frac{c_{23|1}\left(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)\right) \cdot f_{3|1}(x_3|x_1) \cdot f_{2|1}(x_2|x_1)}{f_{2|1}(x_2|x_1)} \\
&= c_{23|1}\left(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)\right) \cdot f_{3|1}(x_3|x_1) \\
&= c_{23|1}\left(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)\right) \cdot c_{13}(F_1(x_1), F_3(x_3))\, f_3(x_3).
\end{aligned}
$$

Further,

$$
\begin{aligned}
f_{4|132}(x_4|x_1, x_3, x_2) &= \frac{f_{34|12}(x_3, x_4|x_1, x_2)}{f_{3|12}(x_3|x_1, x_2)} \\
&= \frac{c_{34|12}(F_{3|12}(x_3|x_1, x_2), F_{4|12}(x_4|x_1, x_2)) \cdot f_{3|12}(x_3|x_1, x_2) \cdot f_{4|12}(x_4|x_1, x_2)}{f_{3|12}(x_3|x_1, x_2)} \\
&= c_{34|12}(F_{3|12}(x_3|x_1, x_2), F_{4|12}(x_4|x_1, x_2)) \cdot f_{4|12}(x_4|x_1, x_2) \\
&= c_{34|12}(F_{3|12}(x_3|x_1, x_2), F_{4|12}(x_4|x_1, x_2)) \cdot \frac{f_{24|1}(x_2, x_4|x_1)}{f_{2|1}(x_2|x_1)} \\
&= c_{34|12}(F_{3|12}(x_3|x_1, x_2), F_{4|12}(x_4|x_1, x_2)) \\
&\qquad \cdot \frac{c_{24|1}(F_{2|1}(x_2|x_1), F_{4|1}(x_4|x_1)) \cdot f_{2|1}(x_2|x_1) \cdot f_{4|1}(x_4|x_1)}{f_{2|1}(x_2|x_1)} \\
&= c_{34|12}(F_{3|12}(x_3|x_1, x_2), F_{4|12}(x_4|x_1, x_2)) \cdot c_{24|1}(F_{2|1}(x_2|x_1), F_{4|1}(x_4|x_1)) \\
&\qquad \cdot f_{4|1}(x_4|x_1) \\
&= c_{34|12}(F_{3|12}(x_3|x_1, x_2), F_{4|12}(x_4|x_1, x_2)) \cdot c_{24|1}(F_{2|1}(x_2|x_1), F_{4|1}(x_4|x_1)) \\
&\qquad \cdot c_{14}(F_1(x_1), F_4(x_4)) \cdot f_4(x_4).
\end{aligned}
$$

Inserting these expressions into (21) gives

$$
\begin{aligned}
f_{1234}(x_1, x_2, x_3, x_4) &= f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \cdot f_4(x_4) \\
&\quad \cdot c_{12}(F_1(x_1), F_2(x_2)) \cdot c_{13}(F_1(x_1), F_3(x_3)) \cdot c_{14}(F_1(x_1), F_4(x_4)) \\
&\quad \cdot c_{23|1}\left(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)\right) \cdot c_{24|1}(F_{2|1}(x_2|x_1), F_{4|1}(x_4|x_1)) \\
&\quad \cdot c_{34|12}(F_{3|12}(x_3|x_1, x_2), F_{4|12}(x_4|x_1, x_2)),
\end{aligned}
$$

which may be recognised as a canonical vine decomposition. Notice that the decomposition includes three pair-copulae acting on marginal univariate distributions, two pair-copulae acting on conditional distribution functions with only one conditioning variable, and one pair-copula acting on conditional distribution functions with two conditioning variables.

We can obtain different pair-copula decompositions by changing the conditioning cascade in (21). Assume for example

$$f_{1234}(x_1, x_2, x_3, x_4) = f_2(x_2) \cdot f_{3|2}(x_3|x_2) \cdot f_{1|32}(x_1|x_3, x_2) \cdot f_{4|132}(x_4|x_1, x_3, x_2).$$

We have that

$$f_{3|2}(x_3|x_2) = c_{23}(F_2(x_2), F_3(x_3)) \, f_3(x_3),$$

and

$$
\begin{aligned}
f_{1|32}(x_1|x_3, x_2) &= \frac{f_{13|2}(x_1, x_3|x_2)}{f_{3|2}(x_3|x_2)} \\
&= \frac{c_{13|2}\left(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)\right) \cdot f_{1|2}(x_1|x_2) \cdot f_{3|2}(x_3|x_2)}{f_{3|2}(x_3|x_2)} \\
&= c_{13|2}\left(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)\right) \cdot f_{1|2}(x_1|x_2) \\
&= c_{13|2}\left(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)\right) \cdot c_{12}(F_1(x_1), F_2(x_2)) \, f_1(x_1)
\end{aligned}
$$

Further, we decompose $f_{4|132}(x_4|x_1, x_3, x_2)$ in another order than above and obtain

$$
\begin{aligned}
f_{4|132}(x_4|x_1, x_3, x_2) &= \frac{f_{14|23}(x_1, x_4|x_2, x_3)}{f_{1|23}(x_1|x_2, x_3)} \\
&= \frac{c_{14|23}(F_{1|23}(x_1|x_2, x_3), F_{4|23}(x_4|x_2, x_3)) \cdot f_{1|23}(x_1|x_2, x_3) \cdot f_{4|23}(x_4|x_2, x_3)}{f_{1|23}(x_1|x_2, x_3)} \\
&= c_{14|23}(F_{1|23}(x_1|x_2, x_3), F_{4|23}(x_4|x_2, x_3)) \cdot f_{4|23}(x_4|x_2, x_3) \\
&= c_{14|23}(F_{1|23}(x_1|x_2, x_3), F_{4|23}(x_4|x_2, x_3)) \cdot \frac{f_{42|3}(x_4, x_2|x_3)}{f_{2|3}(x_2|x_3)} \\
&= c_{14|23}(F_{1|23}(x_1|x_2, x_3), F_{4|23}(x_4|x_2, x_3)) \\
&\quad \cdot \frac{c_{24|3}(F_{2|3}(x_2|x_3), F_{4|3}(x_4|x_3)) \cdot f_{2|3}(x_2|x_3) \cdot f_{4|3}(x_4|x_3)}{f_{2|3}(x_2|x_3)} \\
&= c_{14|23}(F_{1|23}(x_1|x_2, x_3), F_{4|23}(x_4|x_2, x_3)) \cdot c_{24|3}(F_{2|3}(x_2|x_3), F_{4|3}(x_4|x_3)) \\
&\quad \cdot f_{4|3}(x_4|x_3) \\
&= c_{14|23}(F_{1|23}(x_1|x_2, x_3), F_{4|23}(x_4|x_2, x_3)) \cdot c_{24|3}(F_{2|3}(x_2|x_3), F_{4|3}(x_4|x_3)) \\
&\quad \cdot c_{34}(F_3(x_3), F_4(x_4)) \cdot f_4(x_4)
\end{aligned}
$$

Finally we obtain

$$
\begin{aligned}
f_{1234}(x_1, x_2, x_3, x_4) &= f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \cdot f_4(x_4) \\
&\quad \cdot c_{12}\left(F_1(x_1), F_2(x_2)\right) \cdot c_{23}\left(F_2(x_2), F_3(x_3)\right) \cdot c_{34}(F_3(x_3), F_4(x_4)) \\
&\quad \cdot c_{13|2}\left(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)\right) \cdot c_{24|3}(F_{2|3}(x_2|x_3), F_{4|3}(x_4|x_3)) \\
&\quad \cdot c_{14|23}(F_{1|23}(x_1|x_2, x_3), F_{4|23}(x_4|x_2, x_3)),
\end{aligned}
$$

which can be recognised as a D-vine.

B. Pair-copulae

B.1. The bivariate Gaussian copula

The density of the bivariate Gaussian copula is given by

$$c(u_1, u_2) = \frac{1}{\sqrt{1 - \rho_{12}^2}} \exp\left\{ -\frac{\rho_{12}^2(x_1^2 + x_2^2) - 2\rho_{12}\, x_1\, x_2}{2(1 - \rho_{12}^2)} \right\}$$

where ρ_{12} is the parameter of the copula, $x_1 = \Phi^{-1}(u_1)$, $x_2 = \Phi^{-1}(u_2)$ and $\Phi^{-1}(\cdot)$ is the inverse of the standard univariate Gaussian distribution function.

For this copula the h-function is given by (see Appendix C for how it is derived)

$$h(u_1, u_2, \rho_{12}) = \Phi\left(\frac{\Phi^{-1}(u_1) - \rho_{12}\, \Phi^{-1}(u_2)}{\sqrt{1 - \rho_{12}^2}} \right).$$

and the inverse of the h-function is given by

$$h_{12}^{-1}(u_1, u_2, \rho_{12}) = \Phi\left(\Phi^{-1}(u_1)\sqrt{1 - \rho_{12}^2} + \rho_{12}\, \Phi^{-1}(u_2) \right).$$

B.2. The bivariate Student's t-copula

The density of the bivariate Student's t-copula is given by

$$c(u_1, u_2) = \frac{\Gamma(\frac{\nu_{12}+2}{2})/\Gamma(\frac{\nu_{12}}{2})}{\nu_{12}\,\pi\, dt(x_1, \nu_{12})\, dt(x_2, \nu_{12})\sqrt{1 - \rho_{12}^2}} \left(1 + \frac{x_1^2 + x_2^2 - 2\rho_{12}\, x_1\, x_2}{\nu_{12}(1 - \rho_{12}^2)} \right)^{-\frac{\nu_{12}+1}{2}}$$

where ν_{12} and ρ_{12} are the parameters of the copula, $x_1 = t_{\nu_{12}}^{-1}(u_1)$, $x_2 = t_{\nu_{12}}^{-1}(u_2)$, and $dt(\cdot, \nu_{12})$ and $t_{\nu_{12}}^{-1}(\cdot)$ are the probability density and the quantile function, respectively, for the standard univariate student-t-distribution with ν_{12} degrees of freedom, expectation 0 and variance $\frac{\nu_{12}}{\nu_{12}-2}$.

For this copula the h-function is given by (see Appendix C for how it is derived)

$$h(u_1, u_2, \rho_{12}, \nu_{12}) = t_{\nu_{12}+1}\left(\frac{t_{\nu_{12}}^{-1}(u_1) - \rho_{12}\, t_{\nu_{12}}^{-1}(u_2)}{\sqrt{\frac{\left(\nu_{12} + \left(t_{\nu_{12}}^{-1}(u_2)\right)^2\right)\left(1 - \rho_{12}^2\right)}{\nu_{12}+1}}} \right)$$

and the inverse of the h-fuction is given by

$$h_{12}^{-1}(u_1, u_2, \rho_{12}, \nu_{12}) = t_{\nu_{12}}\left(t_{\nu_{12}+1}^{-1}(u_1)\sqrt{\frac{\left(\nu_{12} + \left(t_{\nu_{12}}^{-1}(u_2)\right)^2\right)\left(1 - \rho_{12}^2\right)}{\nu_{12}+1}} + \rho_{12}\, t_{\nu_{12}}^{-1}(u_2) \right).$$

B.3. The bivariate Clayton copula

The density of the bivariate Clayton copula is given by (Venter, 2001)

$$c(u_1, u_2) = (1 + \delta_{12})(u_1 \cdot u_2)^{-1-\delta_{12}} \left(u_1^{-\delta_{12}} + u_2^{-\delta_{12}} - 1 \right)^{-1/\delta_{12}-2},$$

where $0 < \delta_{12} < \infty$ is a parameter controlling the dependence. Perfect dependence is obtained when $\delta_{12} \to \infty$, while $\delta_{12} \to 0$ implies independence.

For this copula the h-function is given by (see Appendix C for how it is derived)

$$h(u_1, u_2, \delta_{12}) = u_2^{-\delta_{12}-1} \left(u_1^{-\delta_{12}} + u_2^{-\delta_{12}} - 1 \right)^{-1-1/\delta_{12}}$$

and the inverse of the h-function is given by

$$h_{12}^{-1}(u_1, u_2, \delta_{12}) = \left[\left(u_1 \cdot u_2^{\delta_{12}+1} \right)^{-\frac{\delta_{12}}{\delta_{12}+1}} + 1 - v^{-\delta_{12}} \right]^{-1/\delta_{12}}.$$

B.4. The bivariate Gumbel copula

The density of the bivariate Gumbel copula is given by (Venter, 2001)

$$
\begin{aligned}
c(u_1, u_2) &= C_{12}(u_1, u_2)\,(u_1\,u_2)^{-1}((-\log u_1)^{\delta_{12}} + (-\log u_1)^{\delta_{12}})^{-2+2/\delta_{12}}(\log u_1 \log u_2)^{\delta_{12}-1} \\
&\times \left\{ 1 + (\delta_{12}-1)((-\log u_1)^{\delta_{12}} + (-\log u_2)^{\delta_{12}})^{-1/\delta_{12}} \right\},
\end{aligned}
$$

where $C_{12}(u_1, u_2)$ is the copula given by

$$C_{12}(u_1, u_2) = \exp\left(-[(-\log u_1)^{\delta_{12}} + (-\log u_2)^{\delta_{12}}]^{1/\delta_{12}} \right),$$

and $\delta_{12} \geq 1$ is a parameter controlling the dependence. Perfect dependence is obtained when $\delta_{12} \to \infty$, while $\delta_{12} = 1$ implies independence.

For this copula the h-function is given by (see Appendix C for how it is derived)

$$h(u_1, u_2, \delta_{12}) = C_{12}(u_1, u_2) \cdot \frac{1}{u_2} \cdot (-\log u_2)^{\delta_{12}-1} \left[(-\log u_1)^{\delta_{12}} + (-\log u_2)^{\delta_{12}} \right]^{1/\delta_{12}-1}.$$

In this case, the inverse of the h-function must be obtained numerically using for instance the Newton Raphson method. Hence, for large-dimensional problems, it might be better to use the Clayton survival copula, see e.g. Joe (1997), which also is a heavy right tail copula.

C. Derivation of h-functions for different copulae

C.1. Gaussian copula

First, we have that the distribution function for the bivariate Gaussian copula is given by

$$C_{12}(u_1, u_2) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\,\pi\,(1-\rho_{12}^2)^{1/2}}\,\exp\left\{ -\frac{x^2 - 2\,\rho_{12}\,x\,y + y^2}{2\,(1-\rho_{12}^2)} \right\} dx\,dy.$$

Set

$$g(x, y) = \frac{1}{2\,\pi\,(1-\rho_{12}^2)^{1/2}}\,\exp\left\{ -\frac{x^2 - 2\,\rho_{12}\,x\,y + y^2}{2\,(1-\rho_{12}^2)} \right\}$$

and

$$b_1 = \Phi^{-1}(u_1) \quad b_2 = \Phi^{-1}(u_2).$$

Then,

$$
\begin{aligned}
h_{12}(u_1, u_2) &= F_{1|2}(u_1|u_2) \\
&= \frac{\partial}{\partial u_2} C_{12}(u_1, u_2) \\
&= \frac{\partial}{\partial u_2} \int_{-\infty}^{b_1} \int_{-\infty}^{b_2} g(x, y) \mathrm{d}x\, \mathrm{d}y \\
&= \frac{\partial b_2}{\partial u_2} \frac{\partial}{\partial b_2} \int_{-\infty}^{b_1} \int_{-\infty}^{b_2} g(x, y) \mathrm{d}x\, \mathrm{d}y \\
&= \frac{1}{\phi(b_2)} \frac{\partial}{\partial b_2} \int_{-\infty}^{b_1} \int_{-\infty}^{b_2} g(x, y) \mathrm{d}x\, \mathrm{d}y \\
&= \frac{1}{\phi(b_2)} \int_{-\infty}^{b_1} \left[\frac{\partial}{\partial b_2} \int_{-\infty}^{b_2} g(x, y) \mathrm{d}x \right] \mathrm{d}y \\
&= \frac{1}{\phi(b_2)} \int_{-\infty}^{b_1} g(x, b_2) \mathrm{d}x \\
&= \frac{1}{\phi(b_2)} \int_{-\infty}^{b_1} \frac{1}{2\pi (1 - \rho_{12}^2)^{1/2}} \exp\left\{ -\frac{x^2 - 2\rho_{12}\, x\, b_2 + b_2^2}{2(1 - \rho_{12}^2)} \right\} \mathrm{d}x \\
&= \frac{1}{\phi(b_2)} \int_{-\infty}^{b_1} \frac{1}{2\pi (1 - \rho_{12}^2)^{1/2}} \exp\left\{ -\frac{(x - \rho_{12} b_2)^2 + (b_2^2 - \rho_{12}^2 b_2^2)}{2(1 - \rho_{12}^2)} \right\} \mathrm{d}x \\
&= \frac{1}{\phi(b_2)} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-b_2^2}{2} \right) \int_{-\infty}^{b_1} \frac{1}{\sqrt{2\pi}\,(1 - \rho_{12}^2)^{1/2}} \exp\left(\frac{-(x - \rho_{12} b_2)^2}{2(1 - \rho_{12}^2)} \right) \mathrm{d}x \\
&= \frac{1}{\phi(b_2)} \frac{\phi(b_2)}{1} \Phi\left(\frac{b_1 - \rho_{12} b_2}{\sqrt{1 - \rho_{12}^2}} \right) \\
&= \Phi\left(\frac{b_1 - \rho_{12} b_2}{\sqrt{1 - \rho_{12}^2}} \right).
\end{aligned}
$$

C.2. Student's t copula

First, we have that the distribution function for the bivariate Student's t-copula is given by

$$
C_{12}(u_1, u_2) = \int_{-\infty}^{t_{\nu_{12}}^{-1}(u_1)} \int_{-\infty}^{t_{\nu_{12}}^{-1}(u_2)} \frac{\Gamma\left(\frac{\nu_{12}+2}{2}\right)}{\Gamma\left(\frac{\nu_{12}}{2}\right) \sqrt{(\pi\nu_{12})^2 (1 - \rho_{12}^2)}} \left\{ 1 + \frac{x^2 - 2\rho_{12}\, x\, y + y^2}{\nu_{12}(1 - \rho_{12}^2)} \right\}^{-(\nu_{12}+2)/2} \mathrm{d}x\, \mathrm{d}y,
$$

Set

$$
g(x, y) = \frac{\Gamma\left(\frac{\nu_{12}+2}{2}\right)}{\Gamma\left(\frac{\nu_{12}}{2}\right) \sqrt{(\pi\nu_{12})^2 (1 - \rho_{12}^2)}} \left\{ 1 + \frac{x^2 - 2\rho_{12}\, x\, y + y^2}{\nu_{12}(1 - \rho_{12}^2)} \right\}^{-(\nu_{12}+2)/2},
$$

$$
f_\nu(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \sqrt{(\pi\nu)}} \left(1 + \frac{x^2}{\nu} \right)^{-\frac{\nu+1}{2}},
$$

and

$$
b_1 = t_{\nu_{12}}^{-1}(u_1) \quad b_2 = t_{\nu_{12}}^{-1}(u_2).
$$

Then,

$$
\begin{aligned}
h_{12}(u_1, u_2) &= F_{1|2}(u_1|u_2) \\
&= \frac{\partial}{\partial u_2} C_{12}(u_1, u_2) \\
&= \frac{\partial}{\partial u_2} \int_{-\infty}^{b_1} \int_{-\infty}^{b_2} g(x, y) \mathrm{dx}\, \mathrm{dy} \\
&= \frac{\partial b_2}{\partial u_2} \frac{\partial}{\partial v_2} \int_{-\infty}^{b_1} \int_{-\infty}^{b_2} g(x, y) \mathrm{dx}\, \mathrm{dy} \\
&= \frac{1}{f_{\nu_{12}}(b_2)} \frac{\partial}{\partial b_2} \int_{-\infty}^{b_1} \int_{-\infty}^{b_2} g(x, y) \mathrm{dx}\, \mathrm{dy} \\
&= \frac{1}{f_{\nu_{12}}(b_2)} \int_{-\infty}^{b_1} \left[\frac{\partial}{\partial b_2} \int_{-\infty}^{b_2} g(x, y) \mathrm{dx} \right] \mathrm{dy} \\
&= \frac{1}{f_{\nu_{12}}(b_2)} \int_{-\infty}^{b_1} g(x, b_2)\, \mathrm{dx} \\
&= \frac{1}{f_{\nu_{12}}(b_2)} \int_{-\infty}^{b_1} \frac{\Gamma\left(\frac{\nu_{12}+2}{2}\right)}{\Gamma\left(\frac{\nu_{12}}{2}\right) \sqrt{(\pi\nu_{12})^2(1-\rho_{12}^2)}} \left\{ 1 + \frac{x^2 - 2\,\rho_{12}\,x\,b_2 + b_2^2}{\nu_{12}\,(1-\rho_{12}^2)} \right\}^{-(\nu_{12}+2)/2} \mathrm{dx} \\
&= \frac{1}{f_{\nu_{12}}(b_2)} \int_{-\infty}^{b_1} \frac{\Gamma\left(\frac{\nu_{12}+2}{2}\right)}{\Gamma\left(\frac{\nu_{12}}{2}\right) \sqrt{(\pi\nu_{12})^2(1-\rho_{12}^2)}} \left[1 + \frac{(x - \rho_{12}b_2)^2}{(\nu_{12}+b_2^2)(1-\rho_{12}^2)} \right]^{-\frac{\nu_{12}+2}{2}} \left[1 + \frac{b_2^2}{\nu_{12}} \right]^{-\frac{\nu_{12}+2}{2}} \mathrm{dx} \\
&= \frac{1}{f_{\nu_{12}}(b_2)} \frac{\Gamma\left(\frac{\nu_{12}+1}{2}\right) \sqrt{\pi(\nu_{12}+b_2^2)(1-\rho_{12}^2)}}{\Gamma\left(\frac{\nu_{12}}{2}\right) \sqrt{(\pi\nu_{12})^2(1-\rho_{12}^2)}} \left[1 + \frac{b_2^2}{\nu_{12}} \right]^{-\frac{\nu_{12}+1}{2}} \left[1 + \frac{b_2^2}{\nu_{12}} \right]^{-\frac{1}{2}} \\
&\; \cdot \int_{-\infty}^{b_1} \frac{\Gamma\left(\frac{\nu_{12}+2}{2}\right)}{\Gamma\left(\frac{\nu_{12}+1}{2}\right) \sqrt{\pi(\nu_{12}+b_2^2)(1-\rho_{12}^2)}} \left[1 + \frac{(x - \rho_{12}b_2)^2}{(\nu_{12}+b_2)^2(1-\rho_{12}^2)} \right]^{-\frac{\nu_{12}+2}{2}} \mathrm{dx} \\
&= \int_{-\infty}^{b_1} \frac{\Gamma\left(\frac{\nu_{12}+2}{2}\right)}{\Gamma\left(\frac{\nu_{12}+1}{2}\right) \sqrt{\pi(\nu_{12}+b_2^2)(1-\rho_{12}^2)}} \left[1 + \frac{(x - \rho_{12}b_2)^2}{(\nu_{12}+b_2)^2(1-\rho_{12}^2)} \right]^{-\frac{\nu_{12}+2}{2}} \mathrm{dx}.
\end{aligned}
$$

Now, set

$$
\begin{aligned}
\nu &= \nu_{12} + 1 \\
\mu &= \rho_{12}b_2 \\
\sigma^2 &= \frac{\nu_{12} + b_2^2}{\nu_{12+1}}(1 - \rho_{12}^2)
\end{aligned}
$$

We then have

$$
\begin{aligned}
h_{12}(u_1, u_2) &= \int_{-\infty}^{b_1} \frac{\Gamma\left(\frac{\nu_{12}+2}{2}\right)}{\Gamma\left(\frac{\nu_{12}+1}{2}\right)\sqrt{\pi(\nu_{12}+b_2^2)(1-\rho_{12}^2)}} \left[1 + \frac{(x-\rho_{12}b_2)^2}{(\nu_{12}+b_2)^2(1-\rho_{12}^2)}\right]^{-\frac{\nu_{12}+2}{2}} dx \\
&= \int_{-\infty}^{b_1} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}\sigma} \left[1 + \frac{1}{\nu}\left(\frac{x-\mu}{\sigma}\right)^2\right]^{-\frac{\nu+1}{2}} \\
&= \int_{-\infty}^{b_1} \frac{1}{\sigma} f_\nu\left(\frac{x-\mu}{\sigma}\right) dx \\
&= \int_{-\infty}^{(b_1-\mu)/\sigma} f_\nu(z) dz \\
&= t_\nu\left(\frac{b_1-\mu}{\sigma}\right).
\end{aligned}
$$

Here $t_\nu(x) = \int_{-\infty}^{x} f_\nu(y)\,dy$. Finally, if we insert expressions for b_1, μ, ν and σ, we have

$$
h_{12}(u_1, u_2) = t_{\nu_{12}+1}\left(\frac{t_{\nu_{12}}^{-1}(u_1) - \rho_{12}\, t_{\nu_{12}}^{-1}(u_2)}{\sqrt{\frac{\left(\nu_{12}+\left(t_{\nu_{12}}^{-1}(u_2)\right)^2\right)\left(1-\rho_{12}^2\right)}{\nu_{12}+1}}}\right).
$$

C.3. Clayton copula

First, we have that the distribution function for the bivariate Clayton copula is given by

$$
C_{12}(u_1, u_2) = (u_1^{-\delta_{12}} + u_2^{-\delta_{12}} - 1)^{-1/\delta_{12}}.
$$

For this copula we have

$$
\begin{aligned}
h_{12}(u_1, u_2) &= F_{1|2}(u_1|u_2) \\
&= \frac{\partial}{\partial u_2} C_{12}(u_1, u_2) \\
&= \frac{\partial}{\partial u_2} (u_1^{-\delta_{12}} + u_2^{-\delta_{12}} - 1)^{-1/\delta_{12}} \\
&= u_2^{-\delta_{12}-1} \left[\left(u_1^{-\delta_{12}} + u_2^{-\delta_{12}} - 1\right)^{-1-1/\delta_{12}}\right]
\end{aligned}
$$

C.4. Gumbel copula

First, we have that the distribution function for the bivariate Gumbel copula is given by

$$
C_{12}(u_1, u_2) = \exp\left(-[(-\log u_1)^{\delta_{12}} + (-\log u_2)^{\delta_{12}}]^{1/\delta_{12}}\right).
$$

For this copula we have

$$
\begin{aligned}
h_{12}(u_1, u_2) &= F_{1|2}(u_1|u_2) \\
&= \frac{\partial}{\partial u_2} C_{12}(u_1, u_2) \\
&= \frac{\partial}{\partial u_2} \exp\left(-[(-\log u_1)^{\delta_{12}} + (-\log u_2)^{\delta_{12}}]^{1/\delta_{12}}\right) \\
&= C_{12}(u_1, u_2) \cdot \frac{1}{u_2} \cdot (-\log u_2)^{\delta_{12}-1} \left[(-\log u_1)^{\delta_{12}} + (-\log u_2)^{\delta_{12}}\right]^{1/\delta_{12}-1}.
\end{aligned}
$$