Available online at www.sciencedirect.com
ScienceDirect
Insurance: Mathematics and Economics 44 (2009) 182–198
INSURANCE MATHEMATICS & ECONOMICS
www.elsevier.com/locate/ime

# Pair-copula constructions of multiple dependence

Kjersti Aas[a,*], Claudia Czado[b], Arnoldo Frigessi[c], Henrik Bakken[d]

[a] The Norwegian Computing Centre, P.O. Box 114, Blindern, N-0314 Oslo, Norway
[b] Technische Universität, München, Germany
[c] Centre for Statistics for Innovation, University of Oslo and the Norwegian Computing Centre, Oslo, Norway
[d] The Norwegian University of Science and Technology, Trondheim, Norway

Received February 2007; accepted 6 February 2007

**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 for performing 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 blocks. Pair-copula decomposed models also represent a very flexible way to construct higher-dimensional copulae. We apply the methodology to a financial data set. Our approach represents the first step towards the development of an unsupervised algorithm that explores the space of possible pair-copula models, that also can be applied to huge data sets automatically.


*Keywords:* Pair-copulae; Vines; Conditional distribution; Decomposition; Multivariate distribution

## 1. Introduction

Inspired by the work of Joe (1996), Bedford and Cooke (2001b, 2002), and Kurowicka and Cooke (2006), we show how multivariate data can be modelled using a cascade of simple building blocks called *pair-copulae*. This probabilistic construction 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 blocks based on conditional independence, e.g., cliques in random fields. Here, the building blocks 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 the pair-copula decomposition of Bedford and Cooke (2002) can be a simple and powerful tool for model building. While it maintains the logic of building complexity using simple elementary bricks, it does not require conditional independence assumptions when these are not natural. We present some of the theory of Bedford and Cooke (2001b, 2002) from a practical point of view, as a general modelling approach, concentrating on likelihood-based inference based on $n$ variables repeatedly observed, say over time.

Kurowicka and Cooke (2006) approach model inference using partial correlations and the determinant of the correlation matrix as a measure of linear dependence. As an alternative, we propose to rely on a maximum pseudo-likelihood approach for parameter estimation of the pair-copula decomposition. An algorithm is given for evaluating the pseudo-likelihood efficiently based on any combination of pair-copulae. This pseudo-likelihood is based on the ranks of the observations. We illustrate this approach for a four-dimensional financial data set for bivariate Student and/or Clayton copulae as building blocks.

Building higher-dimensional copulae is generally recognised as a difficult problem. There are 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 copu-

* Corresponding author. Tel.: +47 22 85 25 00; fax: +47 22 69 76 60.
*E-mail address:* Kjersti.Aas@nr.no (K. Aas).

0167-6687/$ - see front matter 
doi:10.1016/j.insmatheco.2007.02.001

lae; see, e.g., Bandeen-Roche and Liang (1996), Joe (1997), Embrechts et al. (2003), Whelan (2004), Savu and Trede (2006) and McNeil (in press). Meta-elliptical copulae (Fang et al., 2002) also offer some flexibility for multivariate modelling. 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 the conditional independence assumption 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 in Section 6 we discuss aspects of the model selection process. In Section 7 we apply the methodology and discuss limitations and difficulties in the context of a financial data set. Finally, Section 8 contains some concluding remarks.

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

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

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

and this decomposition is unique up to a re-labelling 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 a 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 $F_1(x_1), \ldots, F_n(x_n)$ can be written as

$$F(x_1, \ldots, x_n) = C\{F_1(x_1), \ldots, F_n(x_n)\}, \quad (2)$$

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

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

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

Passing to the joint density function $f$, for an absolutely continuous $F$ with strictly increasing, continuous marginal densities $F_1, \ldots, F_n$ using the chain rule we have

$$f(x_1, \ldots, x_n) = c_{1\cdots n}\{F_1(x_1), \ldots, F_n(x_n)\} \cdot f_1(x_1) \cdots f_n(x_n) \quad (3)$$

for some (uniquely identified) $n$-variate copula density $c_{1\cdots n}(\cdot)$. 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 $c_{(n-1)n}\{F_{n-1}(x_{n-1}), F_n(x_n)\}$ and a marginal density $f_{n-1}(x_{n-1})$. For three random variables $X_1$, $X_2$ and $X_3$ we have that

$$f(x_1|x_2, x_3) = c_{12|3}\{F(x_1|x_3), F(x_2|x_3)\} \cdot f(x_1|x_3), \quad (4)$$

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

$$f(x_1|x_2, x_3) = c_{13|2}\{F(x_1|x_2), F(x_3|x_2)\} \cdot f(x_1|x_2), \quad (5)$$

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

$$f(x_1|x_2, x_3) = c_{13|2}\{F(x_1|x_2), F(x_3|x_2)\} \cdot c_{12}\{F(x_1), F(x_2)\} \cdot f_1(x_1),$$

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 by nature, and that given a specific factorisation, there are still many different re-parametrisations.

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

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

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

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

In Sections 4–6 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,

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

where the second parameter of $h(\cdot)$ always corresponds to the conditioning variable and $\Theta$ denotes the set of parameters for the copula of the joint distribution function of $x$ and $v$. Further,

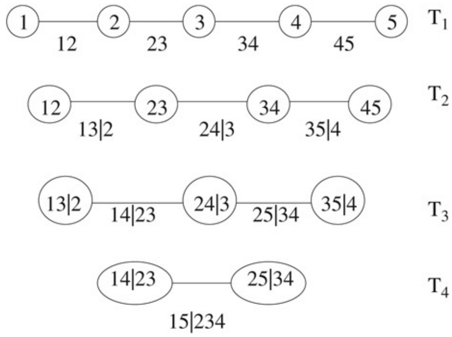


Fig. 1. A D-vine with 5 variables, 4 trees and 10 edges. Each edge may be may be associated with a pair-copula.

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.

### 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 as *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 the form, e.g., of a nested set of trees. Fig. 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 Fig. 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

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

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

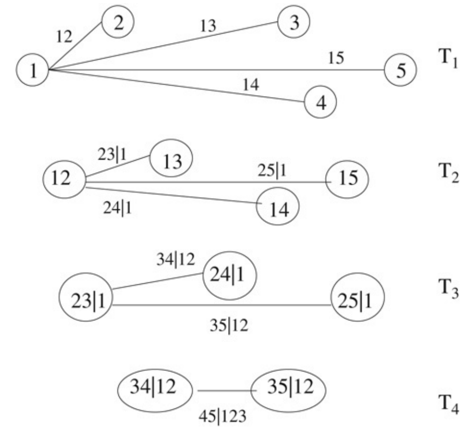


Fig. 2. A canonical vine with 5 variables, 4 trees and 10 edges.

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. Fig. 2 shows a canonical vine with five variables. The $n$-dimensional density corresponding to a canonical vine is given by

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

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 Fig. 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

$$\begin{aligned} f(x_1, x_2, x_3) = {} & f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \\ & \cdot c_{12}\{F_1(x_1), F_2(x_2)\} \cdot c_{23}\{F_2(x_2), F_3(x_3)\} \\ & \cdot c_{13|2}\{F(x_1|x_2), F(x_3|x_2)\}. \end{aligned} \tag{10}$$

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 three decompositions is both a canonical vine and a D-vine.

### 2.3. Four variables

The four-dimensional canonical vine structure is generally expressed as

$$\begin{aligned} f(x_1, x_2, x_3, x_4) = {} & f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \cdot f_4(x_4) \\ & \cdot c_{12}\{F_1(x_1), F_2(x_2)\} \cdot c_{13}\{F_1(x_1), F_3(x_3)\} \\ & \cdot c_{14}\{F_1(x_1), F_4(x_4)\} \\ & \cdot c_{23|1}\{F(x_2|x_1), F(x_3|x_1)\} \cdot c_{24|1}\{F(x_2|x_1), F(x_4|x_1)\} \\ & \cdot c_{34|12}\{F(x_3|x_1, x_2), F(x_4|x_1, x_2)\}, \end{aligned}$$

and the D-vine structure as

$$f(x_1,x_2,x_3,x_4) = f_1(x_1)\cdot f_2(x_2)\cdot f_3(x_3)\cdot f_4(x_4) \cdot c_{12}\{F_1(x_1),F_2(x_2)\}\cdot c_{23}\{F_2(x_2),F_3(x_3)\} \cdot c_{34}\{F_3(x_3),F_4(x_4)\} \cdot c_{13|2}\{F(x_1|x_2),F(x_3|x_2)\}\cdot c_{24|3}\{F(x_2|x_3),F(x_4|x_3)\} \cdot c_{14|23}\{F(x_1|x_2,x_3),F(x_4|x_2,x_3)\}. \quad (11)$$

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

$$f(x_1,x_2,x_3,x_4,x_5) = f_1(x_1)\cdot f_2(x_2)\cdot f_3(x_3)\cdot f_4(x_4) \cdot f_5(x_5)\cdot c_{12}\{F_1(x_1),F_2(x_2)\}\cdot c_{13}\{F_1(x_1),F_3(x_3)\} \cdot c_{14}\{F_1(x_1),F_4(x_4)\} \cdot c_{15}\{F_1(x_1),F_5(x_5)\}\cdot c_{23|1}\{F(x_2|x_1),F(x_3|x_1)\} \cdot c_{24|1}\{F(x_2|x_1),F(x_4|x_1)\}\cdot c_{25|1}\{F(x_2|x_1),F(x_5|x_1)\} \cdot c_{34|12}\{F(x_3|x_1,x_2),F(x_4|x_1,x_2)\} \cdot c_{35|12}\{F(x_3|x_1,x_2),F(x_5|x_1,x_2)\} \cdot c_{45|123}\{F(x_4|x_1,x_2,x_3),F(x_5|x_1,x_2,x_3)\},$$

and the general expression for the D-vine structure is

$$f(x_1,x_2,x_3,x_4,x_5) = f_1(x_1)\cdot f_2(x_2)\cdot f_3(x_3) \cdot f_4(x_4)\cdot f_5(x_5) \cdot c_{12}\{F_1(x_1),F_2(x_2)\}\cdot c_{23}\{F_2(x_2),F_3(x_3)\} \cdot c_{34}\{F_3(x_3),F_4(x_4)\} \cdot c_{45}\{F_4(x_4),F_5(x_5)\}\cdot c_{13|2}\{F(x_1|x_2),F(x_3|x_2)\} \cdot c_{24|3}\{F(x_2|x_3),F(x_4|x_3)\}\cdot c_{35|4}\{F(x_3|x_4),F(x_5|x_4)\} \cdot c_{14|23}\{F(x_1|x_2,x_3),F(x_4|x_2,x_3)\} \cdot c_{25|34}\{F(x_2|x_3,x_4),F(x_5|x_3,x_4)\} \cdot c_{15|234}\{F(x_1|x_2,x_4,x_3),F(x_5|x_2,x_4,x_3)\}.$$

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

$$f(x_1,x_2,x_3,x_4,x_5) = f_1(x_1)\cdot f_2(x_2)\cdot f_3(x_3) \cdot f_4(x_4)\cdot f_5(x_5) \cdot c_{12}\{F_1(x_1),F_2(x_2)\}\cdot c_{23}\{F_2(x_2),F_3(x_3)\} \cdot c_{34}\{F_3(x_3),F_4(x_4)\}\cdot c_{35}\{F_3(x_3),F_5(x_5)\} \cdot c_{13|2}\{F(x_1|x_2),F(x_3|x_2)\}\cdot c_{24|3}\{F(x_2|x_3),F(x_4|x_3)\} \cdot c_{45|3}\{F(x_4|x_3),F(x_5|x_3)\} \cdot c_{14|23}\{F(x_1|x_2,x_3),F(x_4|x_2,x_3)\} \cdot c_{25|34}\{F(x_2|x_3,x_4),F(x_5|x_3,x_4)\} \cdot c_{15|234}\{F(x_1|x_2,x_3,x_4),F(x_5|x_2,x_3,x_4)\}.$$

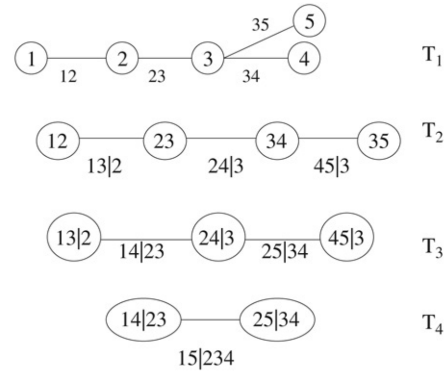

Fig. 3. A regular vine with 5 variables, 4 trees and 10 edges.

The corresponding structure is shown in Fig. 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 three edges instead of four.

In total there are 60 different D-vines and 60 different canonical vines in the five-dimensional case, and none of the D-vines is 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 Fig. 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 three choices for the conditioning set $\{i_{n-1}, i_{n-2}, \dots, i_2\}$ when $i_2, \dots, i_{n-2}$ are chosen before. So altogether we have $n(n-1)\cdots 3 = n!/2$ different canonical vines on $n$ nodes.

For an $n$-dimensional D-vine, there are $n!$ possible ways of ordering the variables 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 $n!/2$ different trees on the first level. Given a such a tree $T_1$, the trees $T_2, T_3, \dots, T_{n-1}$ are completely determined. This implies that the number of distinct D-vines on $n$ nodes is given by $n!/2$.

### 2.6. Multivariate Gaussian distribution

If the marginal distributions $f_i(x_i)$ in (10) are standard normal, and $c_{12}(\cdot)$, $c_{23}(\cdot)$ and $c_{13|2}(\cdot)$ are bivariate Gaussian copula densities (see Appendix C.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 $\rho_{12}$ and $\rho_{23}$ are the correlation parameters of copulae $c_{12}(\cdot)$ and $c_{23}(\cdot)$, respectively, while $\rho_{13}$ 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. In general partial correlation is not equal to conditional correlation, however, for the joint normal distribution the partial and conditional correlations are equal.

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

$$f(x_1, x_2, x_3) = f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \cdot c_{12}\left\{F_1(x_1), F_2(x_2)\right\} \cdot c_{23}\left\{F_2(x_2), F_3(x_3)\right\}.$$

In general, for any vector of variables $\mathbf{V}$ and two variables $X$, $Y$, the latter are conditionally independent given $\mathbf{V}$ if and only if

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

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 $f(x_1, x_2, x_3) = f(x_2|x_1, x_3)f(x_1|x_3)f_3(x_3)$ 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 Fig. 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 1, and the approximation error introduced by the conditional independence assumption is given by the difference $c_{14|23}\left\{F(x_1|x_2, x_3), F(x_4|x_2, x_3)\right\} - 1$.

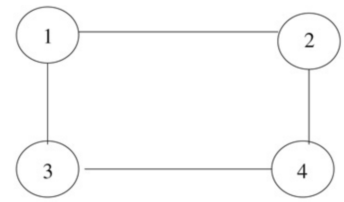

Fig. 4. A conditional independence graph with 4 variables.

## 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 (2007). 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 section 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 $w_1, \ldots, w_n$ independent uniform on [0,1]. Then, set

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

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

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

while for the D-vine we choose

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

### 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 $i$th 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 repeatedly use the $h$-function defined by (7) in Section 2, 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 $c_{j,j+i|1,\ldots,j-1}(\cdot,\cdot)$.

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

```
Sample w_1, ..., w_n independent uniform on [0,1].
x_1 = v_{1,1} = w_1
for i ← 2, ..., n
    v_{i,1} = w_i
    for k ← i − 1, i − 2, ..., 1
        v_{i,1} = h^{-1}(v_{i,1}, v_{k,k}, Θ_{k,i−k})
    end for
    x_i = v_{i,1}
    if i == n then
        Stop
    end if
    for j ← 1, ..., i − 1
        v_{i,j+1} = h(v_{i,j}, v_{j,j}, Θ_{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_1, ..., w_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}, Θ_{1,1})
v_{2,2} = h(v_{1,1}, v_{2,1}, Θ_{1,1})
for i ← 3, ..., n
    v_{i,1} = w_i
    for k ← i − 1, i − 2, ..., 2
        v_{i,1} = h^{-1}(v_{i,1}, v_{i−1,2k−2}, Θ_{k,i−k})
    end for
    v_{i,1} = h^{-1}(v_{i,1}, v_{i−1,1}, Θ_{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}, Θ_{1,i−1})
    v_{i,3} = h(v_{i,1}, v_{i−1,1}, Θ_{1,i−1})
    if i > 3 then
        for j ← 2, ..., i − 2
            v_{i,2j}   = h(v_{i−1,2j−2}, v_{i,2j−1}, Θ_{j,i−j})
            v_{i,2j+1} = h(v_{i,2j−1}, v_{i−1,2j−2}, Θ_{j,i−j})
        end for
    end if
    v_{i,2i−2} = h(v_{i−1,2i−4}, v_{i,2i−3}, Θ_{i−1,1})
end for
```

*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 $c_{i,i+j|i+1,\ldots,i+j-1}(\cdot,\cdot)$.

*4.3. Sampling a three-dimensional vine*

In this section, we describe how to sample from a three-dimensional canonical vine. Since all decompositions in the three-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 $w_1, w_2, w_3$ independent uniform on [0,1]. Then, $x_1 = w_1$. Further, we have $F(x_2|x_1) = h(x_2, x_1, \Theta_{11})$ giving $x_2 = h^{-1}(w_2, x_1, \Theta_{11})$. Finally, $F(x_3|x_1, x_2) = h\{h(x_3, x_1, \Theta_{12}), h(x_2, x_1, \Theta_{11}), \Theta_{21}\}$, meaning that $x_3 = h^{-1}\left[h^{-1}\{w_3, h(x_2, x_1, \Theta_{11}), \Theta_{21}\}, x_1, \Theta_{12}\right]$.

## 5. Inference for a specified 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. Inference for a general regular vine (like the one in Fig. 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. First, we assume for simplicity that the $T$ observations of each variable are independent over time. This is not a limiting assumption, since in the presence of temporal dependence, univariate time-series models can be fitted to the margins and the analysis could henceforth proceed with the residuals. Second, it is pedagogically easier to present the algorithm if each random variable $X_{i,t}$ is assumed to be uniform in [0, 1].

It is important to emphasise that unless the margins are known (which they never are in practice), the estimation method presented below then must rely on the normalised ranks of the data. These are only approximately uniform and independent, meaning that what is being maximised is a pseudo-likelihood. Our proposal extends the method of maximum pseudo-likelihood originally proposed for copulae by Oakes (1994), and later shown to be asymptotically normal and consistent both by Genest et al. (1995) and Shih and Louis (1995). Moreover, recently Kim et al. (2007) have showed by simulation studies that the maximum pseudo-likelihood method performs better than the maximum likelihood method when the marginal distributions are unknown, which is almost always the case in practice. Our model differs from all the above in that a cascade of pair-copulae is modelled instead of one multivariate copula. Hence, the asymptotic properties of the procedure described in this section are yet to be explored.

*5.1. Inference for a canonical vine*

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

$$\sum_{j=1}^{n-1}\sum_{i=1}^{n-j}\sum_{t=1}^{T} \log\left[c_{j,j+i|1,\ldots,j-1}\left\{F(x_{j,t}|x_{1,t},\ldots,x_{j-1,t}),\right.\right.$$
$$\left.\left.F(x_{j+i,t}|x_{1,t},\ldots,x_{j-1,t})\right\}\right]. \quad (12)$$

For each copula in the sum (12) 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 (12). This for-loop consists in turn of two other for-loops. The first of these corresponds to the sum over $i$ in (12). 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 $c_{j,j+i|1,\ldots,j-1}(\cdot,\cdot)$, $h(\cdot)$ 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 $\Theta$ given the data vectors $\boldsymbol{x}$ and $\boldsymbol{v}$. That is,

$$L(\boldsymbol{x},\boldsymbol{v},\Theta) = \sum_{t=1}^{T} \log\left\{c(x_t, v_t, \Theta)\right\}, \tag{13}$$

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

**Algorithm 3** Likelihood evaluation for canonical vine

```
log-likelihood = 0
for i ← 1, ..., n
    v_{0,i} = x_i.
end for
for j ← 1, ..., n − 1
    for i ← 1, ..., n − j
        log-likelihood = log-likelihood
                         + L(v_{j−1,1}, v_{j−1,i+1}, Θ_{j,i})
    end for
    if j == n − 1 then
        Stop
    end if
    for i ← 1, ..., n − j
        v_{j,i} = h(v_{j−1,i+1}, v_{j−1,1}, Θ_{j,i})
    end for
end for
```

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 distribution functions) for tree 2 using the copula parameters from tree 1 and the $h$-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 $h$-function.
(e) Estimate the parameters of the copulae in tree 3 using the observations from (d).
(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

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

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 $c_{i,i+j|i+1,\ldots,i+j-1}(\cdot,\cdot)$.

**Algorithm 4** Likelihood evaluation for a D-vine

```
log-likelihood = 0
for i = 1, ..., n
    v_{0,i} = x_i.
end for
for i = 1, ..., n − 1
    log-likelihood = log-likelihood + L(v_{0,i}, v_{0,i+1}, Θ_{1,i})
end for
v_{1,1} = h(v_{0,1}, v_{0,2}, Θ_{1,1})
for k = 1, ..., n − 3
    v_{1,2k} = h(v_{0,k+2}, v_{0,k+1}, Θ_{1,k+1})
    v_{1,2k+1} = h(v_{0,k+1}, v_{0,k+2}, Θ_{1,k+1})
end for
v_{1,2n−4} = h(v_{0,n}, v_{0,n−1}, Θ_{1,n−1})
for j = 2, ..., n − 1
    for i = 1, ..., n − j
        log-likelihood = log-likelihood
                         + L(v_{j−1,2i−1}, v_{j−1,2i}, Θ_{j,i})
    end for
    if j == n − 1 then
        Stop
    end if
    v_{j,1} = h(v_{j−1,1}, v_{j−1,2}, Θ_{j,1})
    if n > 4 then
        for i = 1, 2, ..., n − j − 2
            v_{j,2i} = h(v_{j−1,2i+2}, v_{j−1,2i+1}, Θ_{j,i+1})
            v_{j,2i+1} = h(v_{j−1,2i+1}, v_{j−1,2i+2}, Θ_{j,i+1})
        end for
    end if
    v_{j,2n−2j−2} = h(v_{j−1,2n−2j}, v_{j−1,2n−2j−1}, Θ_{j,n−j})
end for
```

### 5.3. Inference for a three-variable model

In the special case of a three-dimensional data set with $U[0,1]$ distributed variables, (12) reduces to

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

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 $c_{i,i+j|i+1,\ldots,i+j-1}(\cdot,\cdot)$. Following the procedure described in Section 5.1, we first estimate the parameters of the three copulae involved by a sequential procedure, and then we maximise the full log-likelihood using the parameters obtained from the stepwise procedure as starting values. A numerical example is given in Appendix B.

## 6. 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 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 determine which bivariate relationships 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 7.1). 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 6.1.
(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 upstream in the decomposition. This selection mechanism does not guarantee a globally optimal fit. Having determined the appropriate parametric shapes for each copulae, one may use the procedures in Section 5 to estimate their parameters.

### 6.1. Goodness-of-fit

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. For the same reason, there is no demonstrated goodness-of-fit technique currently available for validating our pair-copula decomposition.

Of the tests that have been proposed for copulae (see Genest et al. (2007) for a review of omnibus goodness-of-fit tests for copulae), quite a few are based on the probability integral transform (PIT) of Rosenblatt (1952); see, e.g., Breymann et al. (2003) and Dobrić and Schmid (in press). We propose to use the PIT also for the pair-copula decomposition. The PIT converts a set of dependent variables into a new set of variables that are independent and uniform under the null hypothesis that the data originate from a given multivariate distribution. The technique, which may be viewed as the inverse of simulation, is defined as follows.

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

$$\begin{aligned} T(X_1) &= F(x_1) \\ T(X_2) &= F(x_2|x_1) \\ \cdot &= \cdot \\ \cdot &= \cdot \\ T(X_n) &= F(x_n|x_1, \ldots, x_{n-1}). \end{aligned}$$

The random variables $Z_i = T(X_i), i = 1, \ldots, n$, 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.

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

**Algorithm 5** PIT algorithm for a canonical vine

```
for t ← 1, ..., T
    z_{1,t} = x_{1,t}
    for i ← 2, ..., n
        z_{i,t} = x_{i,t}
        for j ← 1, ..., i − 1
            z_{i,t} = h(z_{i,t}, z_{j,t}, Θ_{j,i−j})
        end for
    end for
end for
```

**Algorithm 6** PIT algorithm for a D-vine

```
for t ← 1, ..., T
    z_{1,t} = x_{1,t}
    z_{2,t} = h(x_{2,t}, x_{1,t}, Θ_{1,1})
    v_{2,1} = x_{2,t}
    v_{2,2} = h(x_{1,t}, x_{2,t}, Θ_{1,1})
    for i ← 3, ..., n
        z_{i,t} = h(x_{i,t}, x_{i−1,t}, Θ_{1,i−1})
        for j ← 2, ..., i − 1
            z_{i,t} = h(z_{i,t}, v_{i−1,2(j−1)}, Θ_{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}, Θ_{1,i−1})
        v_{i,3} = h(v_{i,1}, v_{i−1,1}, Θ_{1,i−1})
        for j ← 1, 2, ..., i − 3
            v_{i,2j+2} = h(v_{i−1,2j}, v_{i,2j+1}, Θ_{j+1,i−j−1})
            v_{i,2j+3} = h(v_{i,2j+1}, v_{i−1,2j}, Θ_{j+1,i−j−1})
        end for
        v_{i,2i−2} = h(v_{i−1,2i−4}, v_{i,2i−3}, Θ_{i−1,1})
    end for
end for
```

with $n$ degrees of freedom. The Anderson–Darling goodness-of-fit test may be applied for the latter. 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 also for copulae in general. We follow the procedure suggested by Dobrić and Schmid (in press) and use parametric bootstrap to determine critical values. This procedure is shown by simulation studies to perform well.

## 7. Application: Financial returns

### 7.1. Tail dependence

Tail dependence properties are particularly important in many applications that rely on non-normal multivariate families (Joe, 1996). This is especially the case for financial applications. Tail dependence in a bivariate distribution 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* (Sibuya, 1960), and a copula is said to be upper-tail dependent if this limit is not zero. In this section we present four pair-copulae that have different strength of dependence in the tails of the bivariate distribution; the Gaussian, the Student, the Clayton and the Gumbel copulae. See Joe (1997) for an overview of other copulae. 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. The Clayton copula is lower-tail dependent, but not upper. The Gumbel copula is upper-tail dependent, but not lower. The Student copula is both lower- and upper-tail dependent, while the Gaussian is neither lower- nor upper-tail dependent.

In Appendix C 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 pair-copula has two. The additional parameter of the latter is the degrees of freedom, controlling the strength of dependence in the tails of the bivariate distribution. The Student 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, a better choice might be the Clayton copula, because it exhibits greater dependence in the negative tail than in the positive. The Gumbel copula is also an asymmetric copula, but it exhibits greater dependence in the positive tail than in the negative. Fig. 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 C. 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.

### 7.2. 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. Fig. 6 shows the log-returns of each pair of assets. The four variables are denoted as $T$, $M$, $B$ and $S$.

We want to compare a four-dimensional pair-copula decomposition with Student copulae for all pairs with the four-dimensional Student copula. The $n$-dimensional Student 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 Student copula has only one parameter for modelling tail dependence, independent of dimension. Hence, if the tail dependences of different pairs of risk factors in a portfolio are very different, we believe that a better description

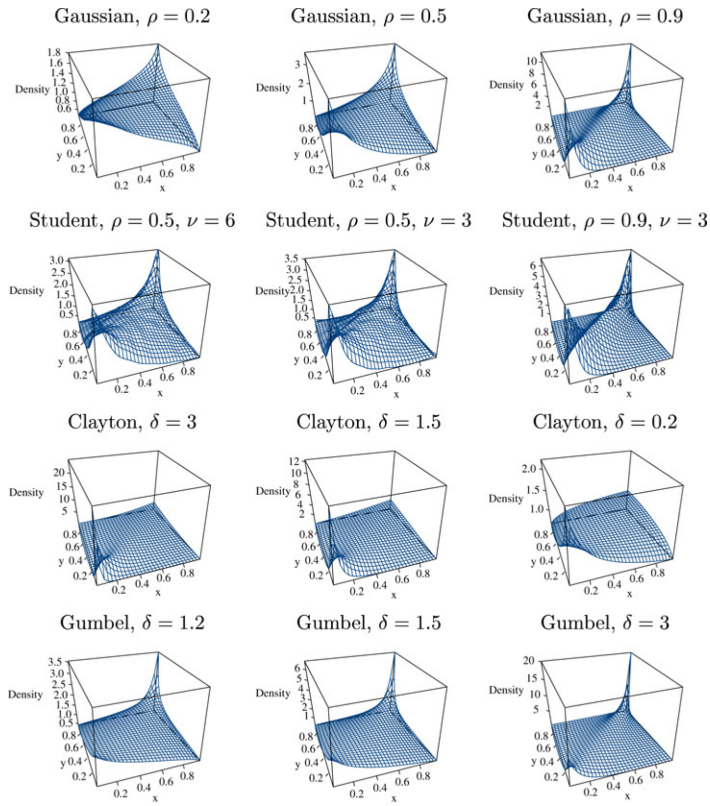


Fig. 5. 3D surface plot of bivariate density for Gaussian, Student, Clayton and Gumbel copulae, with three different parameter settings.

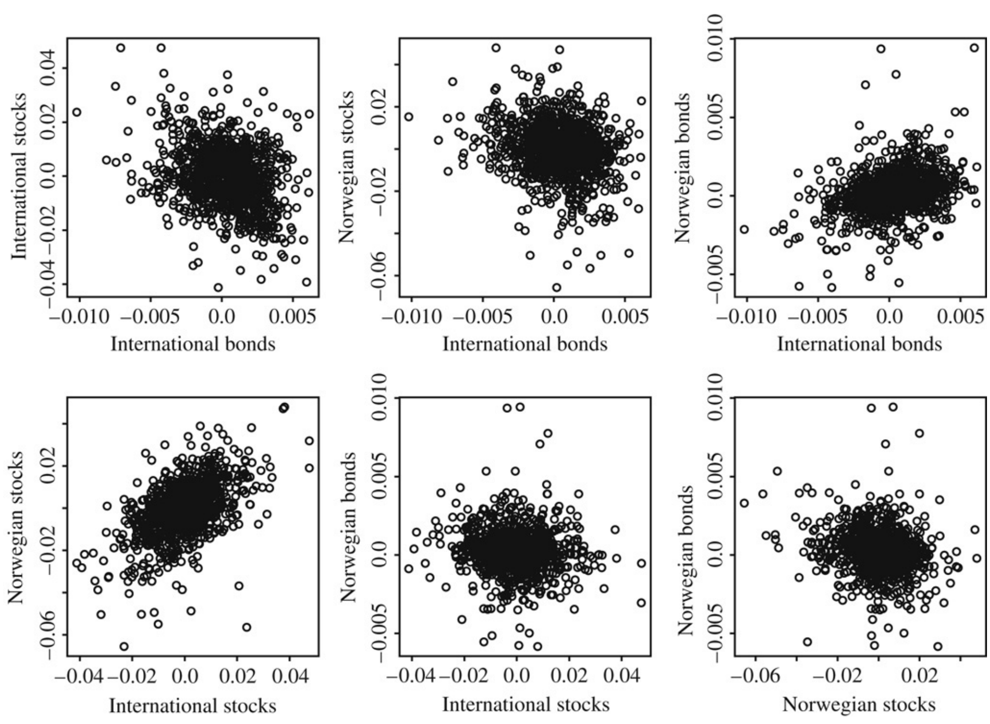


Fig. 6. Log-returns for pairs of assets during the period from 04.01.1999 to 08.07.2003.

of the dependence structure can be achieved with the pair-copula decomposition with Student copulae for each pair.

As stated in Section 5, the observations of each variable must be independent over time. Hence, the serial correlation in the conditional mean and the conditional variance are modelled by an AR(1)- and a GARCH(1, 1)-model (Bollerslev, 1986), respectively. That is, for series $i$, we have the following model for log-return $x_i$:

$$x_{i,t} = c_i + \alpha_i\, x_{i,t-1} + \sigma_{i,t}\, z_{i,t},$$
$$\mathrm{E}[z_{t,i}] = 0 \quad \text{and} \quad \mathrm{Var}[z_{t,i}] = 1, \tag{14}$$
$$\sigma^2_{i,t} = a_{i,0} + a_i\, \epsilon^2_{i,t-1} + b_i\, \sigma^2_{i,t-1},$$

where $\epsilon_{i,t-1} = \sigma_{i,t}\, z_{i,t}$. The further analysis is performed on the standardised residuals $z_i$. If the AR(1)–GARCH(1, 1) models are successful at modelling the serial correlation in the conditional mean and the conditional variance, there should be no autocorrelation left in the standardised residuals and squared standardised residuals. We use the modified Q-statistic (Ljung and Box, 1979) and the Lagrange Multiplier Test (LM) Engle (1982), respectively, to check this. For all series and both tests, the null hypothesis that there is no autocorrelation left cannot be rejected at the 5% level. Since we are mainly interested in estimating the dependence structure of the risk factors, the standardised residual vectors are converted to uniform variables using the empirical distribution functions before further modelling. In the light of recent results due to Chen and Fan (2006), the method of maximum pseudo-likelihood is consistent even when time-series models are fitted to the margins.

### 7.3. Selecting an appropriate pair-copula decomposition

Having decided to use Student 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 copula to each pair of risk factors, obtaining estimated degrees of freedom for each pair. For this we use the two-step maximum likelihood method described in broad terms by Oakes (1994) and later formalised and studied by Genest et al. (1995) and Shih and Louis (1995). The estimation of the Student 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 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 numbers of degrees of freedom 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 cannot 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 Fig. 7 for the whole D-vine structure in this case.

### 7.4. 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 (13) and the density and the $h$-function for the Student copula given in Appendix C.2.

Table 1
Estimated numbers of degrees of freedom for bivariate Student copulae for pairs of variables

| Between | $M$ | $T$ | $B$ |
|---|---|---|---|
| $S$ | 4.21 | 34.16 | 14.47 |
| $M$ | | 8.03 | 15.48 |
| $T$ | | | 12.60 |

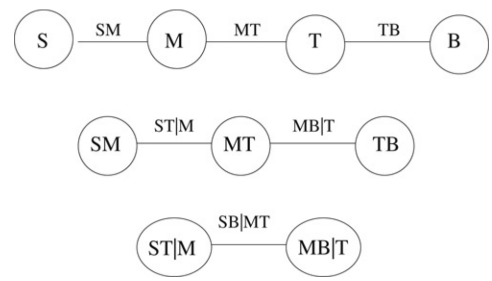


Fig. 7. Selected D-vine structure for the data set in Section 7.2.

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

| Param | Start | Final |
|---|---|---|
| $\rho_{SM}$ | −0.25 | −0.25 |
| $\rho_{MT}$ | 0.47 | 0.47 |
| $\rho_{TB}$ | −0.17 | −0.17 |
| $\rho_{ST\|M}$ | −0.11 | −0.11 |
| $\rho_{MB\|T}$ | 0.02 | 0.03 |
| $\rho_{SB\|MT}$ | 0.29 | 0.28 |
| $\nu_{SM}$ | 4.21 | 4.34 |
| $\nu_{MT}$ | 16.65 | 16.26 |
| $\nu_{TB}$ | 12.60 | 13.17 |
| $\nu_{ST\|M}$ | 300.00 | 300.00 |
| $\nu_{MB\|T}$ | 130.33 | 45.59 |
| $\nu_{SB\|MT}$ | 15.58 | 15.04 |
| log-likelih. | 267.86 | 268.17 |

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 Information Criterion (AIC) for the final model is −512.33. The $p$-value for the goodness-of-fit test described in Section 6.1 was 0.98 (computed by the procedure in Dobrić and Schmid (in press)), meaning that we cannot reject the null hypothesis of a D-vine composed of Student copulae.

### 7.5. 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 whether simulated values and observed data have similar features.

Table 3
Estimated numbers of degrees of freedom for bivariate Student copulae for pairs of simulated variables

| Between | $M$ | $T$ | $B$ |
|---|---|---|---|
| $S$ | 4.19 | 17.29 | 19.49 |
| $M$ | | 14.82 | 19.28 |
| $T$ | | | 11.71 |

We used the simulation procedure described in Section 4.2 and the $h$-function and its inverse given in Appendix C.2 to generate a set of 10,000 samples from the estimated pair-copula decomposition. Then, we estimated pairwise Student copulae for all bivariate margins. The results are shown in Table 3. Comparing these to the ones in Table 1 we see that all dependencies are quite well captured, including those that are not directly modelled.

### 7.6. *Comparison with the four-dimensional Student copula*

In this section we compare the results obtained with the pair-copula decomposition from Section 7.4 with those obtained with a four-dimensional Student copula. The parameters of the Student copula are shown in Table 4. The AIC for this model is −487.42, i.e., higher than that for the pair-copulae decomposition. Further, all conditional distributions of a multivariate Student distribution are Student distributions. Hence, the $n$-dimensional Student 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 distribution. Therefore, the four-dimensional Student copula is nested within the considered D-vine structure and the likelihood ratio test statistic is $2\,(268.17 - 250.71) = 34.92$ with $12 - 7 = 5$ degrees of freedom. This yields a $p$-value of 1.56e−006 and shows that the four-dimensional Student copula can be rejected in favour of the D-vine.

To illustrate the difference between the four-dimensional Student copula and the four-dimensional pair-copula decomposition, we computed the tail dependence coefficients for the three bivariate margins $SM$, $MT$ and $TB$ for both structures. See Section 7.1 for the definition of the upper- and lower-tail dependence coefficients. For the Student 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 $t$-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 279 times higher than the corresponding one for the Student copula. 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 observing a large portfolio loss is much higher for the four-dimensional pair-copula decomposition that it is for the four-dimensional Student copula.

Table 4
Estimated parameters for four-dimensional Student copulae

| Param | Value |
|---|---|
| $\rho_{SM}$ | −0.25 |
| $\rho_{ST}$ | −0.20 |
| $\rho_{SB}$ | 0.30 |
| $\rho_{MT}$ | 0.47 |
| $\rho_{MB}$ | −0.06 |
| $\rho_{TB}$ | −0.17 |
| $\nu_{STMB}$ | 14.56 |
| log-likelih. | 250.71 |

Table 5
Tail dependence coefficients

| Margin | Pair-copula decomp. | Student copula |
|---|---|---|
| $SM$ | 0.0279 | 0.0001 |
| $MT$ | 0.0229 | 0.0317 |
| $TB$ | 0.0005 | 0.0003 |

### 7.7. *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 Fig. 7 to come from different families. Fig. 8 shows the data sets used to estimate the six pair-copulae in the decomposition described in Sections 7.3 and 7.4. 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 copula is an appropriate choice. In the two leftmost scatter plots in the lower row, the data seem 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 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 this is so, $c_{ST|M}(\cdot)$ and $c_{MB|T}(\cdot)$ are both 1, which means that the pair-copula construction defined by Fig. 7 may be simplified to

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

If we estimate this model instead, the parameters of copula $c_{SB|MT}(\cdot,\cdot)$ are slightly altered to $\rho_{SB|MT} = 0.28$ and $\nu_{SB|MT} = 15.22$. The log-likelihood for this reduced structure is 261.6 compared to 268.17 for the full one. The corresponding AIC values are −507.20 and −512.33, meaning that the full model is slightly better than the reduced one. This is also verified by the likelihood ratio statistic, which is $2\,(261.6 - 268.17) = 13.14$. With $12 - 8 = 4$ degrees of freedom this gives a $p$-value of 0.01 and shows that the reduced structure is rejected in favour of the full one.

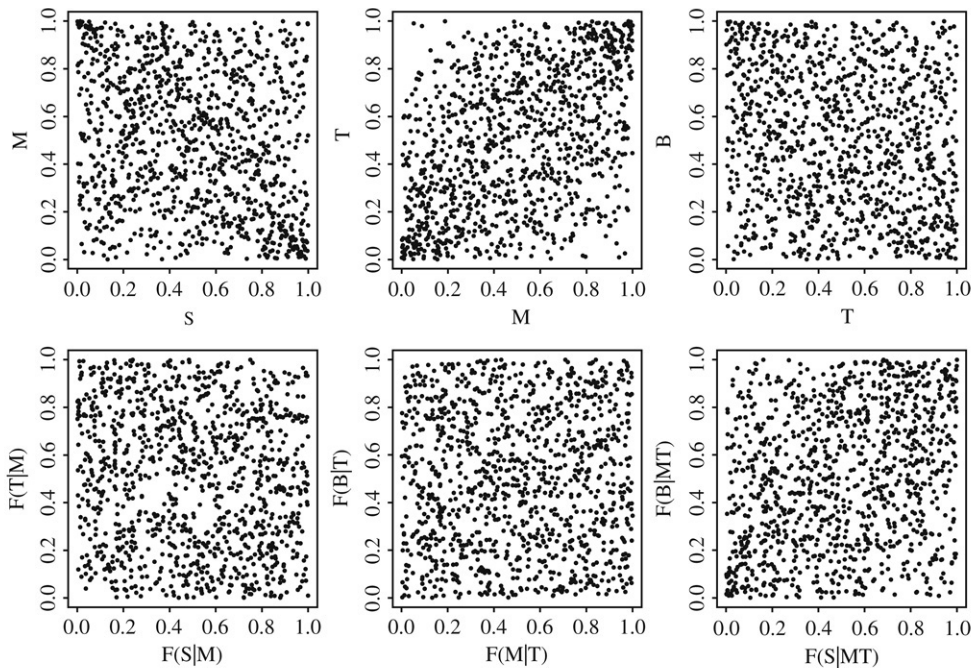


Fig. 8. The data sets used to estimate the six pair-copulae in the decomposition described in Sections 7.3 and 7.4.

Turning to the pair-copula $c_{SB|MT}(\cdot)$, there seems to be data clustering in the lower left corner of the scatter plot to the bottom right of Fig. 8, but not in the upper right. This indicates that the Clayton copula might be a better choice than the Student 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 $\delta = 0.34$. The likelihood of the Clayton copula is lower than that of the Student copula (39.72 versus 47.81). 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 function $\lambda(z)$, given by

$$\lambda(z) = z - K(z).$$

Here $K(z)$ is the copula distribution function $K(z)$, defined by

$$K(z) = P(C(u_1, u_2) \leq z).$$

For Archimedean copulae, $K(z)$ is given by an explicit expression, while for the Student copula it has to be numerically derived. In Fig. 9 the empirical lambda function and its confidence bands, computed as described in Genest and Rivest (1993), are presented together with the fitted lambda functions for the Clayton copula and the Student copula. As can be seen from this figure, the Student copula fits the empirical data remarkably well. This may be more formally verified by a goodness-of-fit test; see, e.g., Chen and Fan (2005).

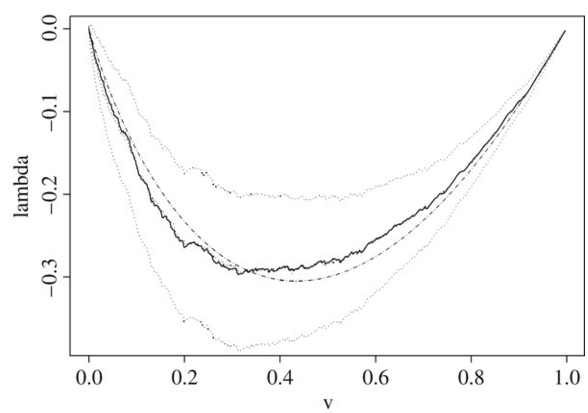


Fig. 9. The empirical lambda function (solid line) and its confidence bands (dotted lines) are presented together with the fitted lambda functions for the Clayton copula (dashed line) and the Student copula (dotted line which hardly can be distinguished from the solid line).

## 8. 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 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 blocks, which can be compared to pairwise interaction potentials or cliques in Gibbs fields.

When presenting the theory we have assumed that the observations of each variable are independent over time. However, this is not a limiting assumption. In our application we have shown that in the presence of temporal dependence, univariate time-series models can be fitted to the margins and the analysis can henceforth proceed with the residuals.

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 could also put priors on the choice of pairs to match. Posterior estimates would 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. However, this remains a central aim, since there is an increasing tendency to collect huge quantities of multivariate and dense observations, requiring automatic inferential methods.

**Acknowledgements**

This work was sponsored by the Norwegian fund Finansmarkedsfondet and the Norwegian Research Council. Claudia Czado is also supported by the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 386, Statistical Analysis of discrete structures. The authors are very grateful to Professor Christian Genest, Université Laval, for his significant and valuable comments. They also want to thank Daniel Berg for fruitful discussions, and Håvard Rue for telling them about the work of Bedford and Cooke.

Henrik Bakken's part of this work was conducted while he was a master student at the Norwegian Comuputing Centre.

## Appendix A. Four-dimensional pair-copula decompositions

Assume that we decompose a given four-dimensional density $f(x_1, x_2, x_3, x_4)$ as follows:

$$f(x_1, x_2, x_3, x_4) = f_1(x_1) \cdot f(x_2|x_1) \cdot f(x_3|x_1, x_2) \cdot f(x_4|x_1, x_3, x_2). \tag{15}$$

We have

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

and

$$\begin{aligned} f(x_3|x_1, x_2) &= \frac{f(x_2, x_3|x_1)}{f(x_2|x_1)} \\ &= \frac{c_{23|1}\,(F(x_2|x_1), F(x_3|x_1)) \cdot f(x_3|x_1) \cdot f(x_2|x_1)}{f(x_2|x_1)} \\ &= c_{23|1}\{F(x_2|x_1), F(x_3|x_1)\} \cdot f(x_3|x_1) \\ &= c_{23|1}\{F(x_2|x_1), F(x_3|x_1)\} \cdot c_{13}\{F_1(x_1), F_3(x_3)\} \cdot f_3(x_3). \end{aligned}$$

Further,

$$\begin{aligned} f(x_4|x_1, x_3, x_2) &= \frac{f(x_3, x_4|x_1, x_2)}{f(x_3|x_1, x_2)} \\ &= \frac{c_{34|12}\{F(x_3|x_1, x_2), F(x_4|x_1, x_2)\} \cdot f(x_3|x_1, x_2) \cdot f(x_4|x_1, x_2)}{f(x_3|x_1, x_2)} \\ &= c_{34|12}\{F(x_3|x_1, x_2), F(x_4|x_1, x_2)\} \cdot f(x_4|x_1, x_2) \\ &= c_{34|12}\{F(x_3|x_1, x_2), F(x_4|x_1, x_2)\} \cdot \frac{f(x_2, x_4|x_1)}{f(x_2|x_1)} \\ &= c_{34|12}\{F(x_3|x_1, x_2), F(x_4|x_1, x_2)\} \\ &\quad \cdot \frac{c_{24|1}\{F(x_2|x_1), F(x_4|x_1)\} \cdot f(x_2|x_1) \cdot f(x_4|x_1)}{f(x_2|x_1)} \\ &= c_{34|12}\{F(x_3|x_1, x_2), F(x_4|x_1, x_2)\} \cdot c_{24|1}\{F(x_2|x_1), F(x_4|x_1)\} \\ &\quad \cdot f(x_4|x_1) \\ &= c_{34|12}\{F(x_3|x_1, x_2), F(x_4|x_1, x_2)\} \cdot c_{24|1}\{F(x_2|x_1), F(x_4|x_1)\} \\ &\quad \cdot c_{14}\{F_1(x_1), F_4(x_4)\} \cdot f_4(x_4). \end{aligned}$$

Inserting these expressions into (15) gives

$$\begin{aligned} f(x_1, x_2, x_3, x_4) &= f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \cdot f_4(x_4) \\ &\cdot c_{12}\{F_1(x_1), F_2(x_2)\} \cdot c_{13}\{F_1(x_1), F_3(x_3)\} \\ &\cdot c_{14}\{F_1(x_1), F_4(x_4)\} \\ &\cdot c_{23|1}\{F(x_2|x_1), F(x_3|x_1)\} \cdot c_{24|1}\{F(x_2|x_1), F(x_4|x_1)\} \\ &\cdot c_{34|12}\{F(x_3|x_1, x_2), F(x_4|x_1, x_2)\}, \end{aligned}$$

which may be recognised as a canonical vine decomposition. 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 (15). Assume for example

$$f(x_1, x_2, x_3, x_4) = f_2(x_2) \cdot f(x_3|x_2) \cdot f(x_1|x_3, x_2) \cdot f(x_4|x_1, x_3, x_2).$$

We have

$$f(x_3|x_2) = c_{23}\{F_2(x_2), F_3(x_3)\} \cdot f_3(x_3),$$

and

$$\begin{aligned} f(x_1|x_3, x_2) &= \frac{f(x_1, x_3|x_2)}{f(x_3|x_2)} \\ &= \frac{c_{13|2}\{F(x_1|x_2), F(x_3|x_2)\} \cdot f(x_1|x_2) \cdot f(x_3|x_2)}{f(x_3|x_2)} \\ &= c_{13|2}\{F(x_1|x_2), F(x_3|x_2)\} \cdot f(x_1|x_2) \\ &= c_{13|2}\{F(x_1|x_2), F(x_3|x_2)\} \cdot c_{12}\{F_1(x_1), F_2(x_2)\} \cdot f_1(x_1). \end{aligned}$$

Further, we decompose $f(x_4|x_1, x_3, x_2)$ in another order to the above and obtain

$$\begin{aligned} f(x_4|x_1, x_3, x_2) &= \frac{f(x_1, x_4|x_2, x_3)}{f(x_1|x_2, x_3)} \\ &= \frac{c_{14|23}\{F(x_1|x_2, x_3), F(x_4|x_2, x_3)\} \cdot f(x_1|x_2, x_3) \cdot f(x_4|x_2, x_3)}{f(x_1|x_2, x_3)} \\ &= c_{14|23}\{F(x_1|x_2, x_3), F(x_4|x_2, x_3)\} \cdot f(x_4|x_2, x_3) \\ &= c_{14|23}\{F(x_1|x_2, x_3), F(x_4|x_2, x_3)\} \cdot \frac{f(x_4, x_2|x_3)}{f(x_2|x_3)} \\ &= c_{14|23}\{F(x_1|x_2, x_3), F(x_4|x_2, x_3)\} \\ &\quad \cdot \frac{c_{24|3}\{F(x_2|x_3), F(x_4|x_3)\} \cdot f(x_2|x_3) \cdot f(x_4|x_3)}{f(x_2|x_3)} \\ &= c_{14|23}\{F(x_1|x_2, x_3), F(x_4|x_2, x_3)\} \\ &\quad \cdot c_{24|3}\{F(x_2|x_3), F(x_4|x_3)\} \cdot f(x_4|x_3) \\ &= c_{14|23}\{F(x_1|x_2, x_3), F(x_4|x_2, x_3)\} \cdot c_{24|3}\{F(x_2|x_3), F(x_4|x_3)\} \\ &\quad \cdot c_{34}\{F_3(x_3), F_4(x_4)\} \cdot f_4(x_4). \end{aligned}$$

Finally we obtain

$$\begin{aligned} f(x_1,x_2,x_3,x_4) &= f_1(x_1)\cdot f_2(x_2)\cdot f_3(x_3)\cdot f_4(x_4) \\ &\cdot c_{12}\{F_1(x_1),F_2(x_2)\}\cdot c_{23}\{F_2(x_2),F_3(x_3)\} \\ &\cdot c_{34}\{F_3(x_3),F_4(x_4)\} \\ &\cdot c_{13|2}\{F(x_1|x_2),F(x_3|x_2)\}\cdot c_{24|3}\{F(x_2|x_3),F(x_4|x_3)\} \\ &\cdot c_{14|23}\{F(x_1|x_2,x_3),F(x_4|x_2,x_3)\}, \end{aligned}$$

which can be recognised as a D-vine.

## Appendix B. Numerical example of likelihood evaluation

In this section we illustrate the inference for a three-dimensional data set with a numerical example. Assume that we have the following data set:

$$\begin{aligned} \boldsymbol{y}_1 &= (4.9, 4.4, 5.1, 4.3, 4.7), \\ \boldsymbol{y}_2 &= (9.9, 8.5, 9.6, 8.8, 9.1), \\ \boldsymbol{y}_3 &= (7.0, 6.2, 6.9, 6.0, 7.7), \end{aligned}$$

and that we want to fit a three-dimensional D-vine specification with Clayton copulae for all pairs. In Section 5.3 we assume that the three vectors have been transformed into normalised ranks, i.e.,

$$\begin{aligned} \boldsymbol{x}_1 &= (0.8, 0.4, 1.0, 0.2, 0.6), \\ \boldsymbol{x}_2 &= (1.0, 0.2, 0.8, 0.4, 0.6), \\ \boldsymbol{x}_3 &= (0.8, 0.4, 0.6, 0.2, 1.0), \end{aligned}$$

before the parameters of the D-vine specification are estimated. In our numerical maximisation of the log-likelihood, the log-likelihood

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

is computed for every possible value of the triplet $(\Theta_{11},\Theta_{12},\Theta_{21})$, and the value that maximises the log-likelihood is chosen as the maximum likelihood estimate. Here, we show how to compute the log-likelihood for the triplet (3.0, 1.3, 0.5). Using the $h$-function for the Clayton copula (see Appendix C), we first compute the observations $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ as

$$v_{1,t} = h(x_{1,t},x_{2,t},3.0)$$

and

$$v_{2,t} = h(x_{3,t},x_{2,t},1.3).$$

The result is

$$\begin{aligned} \boldsymbol{v}_1 &= (0.41, 0.86, 1.00, 0.05, 0.46) \\ \boldsymbol{v}_2 &= (0.60, 0.64, 0.39, 0.13, 1.00). \end{aligned}$$

Then, using the density for the Clayton copula (see Appendix C), and plugging in the values of $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$, $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\Theta_{11}$, $\Theta_{12}$ and $\Theta_{21}$, we compute the log-likelihood of the D-vine specification as the sum of the log-likelihoods of the three Clayton copulae, i.e., $1.91+1.27+0.79=3.97$.

## Appendix C. Pair-copulae

### C.1. The bivariate Gaussian copula

The density of the bivariate Gaussian copula is given by

$$c(u_1,u_2)=\frac{1}{\sqrt{1-\rho_{12}^2}}\exp\left\{-\frac{\rho_{12}^2(x_1^2+x_2^2)-2\rho_{12}x_1x_2}{2(1-\rho_{12}^2)}\right\}$$

where $\rho_{12}$ 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

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

and the inverse of the $h$-function is given by

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

### C.2. The bivariate Student copula

The density of the bivariate Student copula is given by

$$c(u_1,u_2)=\frac{\Gamma(\frac{\nu_{12}+2}{2})/\Gamma(\frac{\nu_{12}}{2})}{\nu_{12}\,\pi\,\mathrm{d}t(x_1,\nu_{12})\,\mathrm{d}t(x_2,\nu_{12})\sqrt{1-\rho_{12}^2}} \times\left\{1+\frac{x_1^2+x_2^2-2\rho_{12}x_1x_2}{\nu_{12}(1-\rho_{12}^2)}\right\}^{-\frac{\nu_{12}+1}{2}}$$

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

For this copula the $h$-function is given by

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

and the inverse of the $h$-function is given by

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

### C.3. The bivariate Clayton copula

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

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

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

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

and the inverse of the $h$-function is given by

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

### C.4. The bivariate Gumbel copula

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

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

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

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

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

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

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.

## References


Bandeen-Roche, K.J., Liang, K.-Y., 1996. Modelling failure-time associations in data with multiple levels of clustering. Biometrika 83, 29–39.

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

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

Bedford, T., Cooke, R.M., 2002. Vines — a new graphical model for dependent random variables. Annals of Statistics 30, 1031–1068.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327.

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

Chen, X., Fan, Y., 2005. Pseudo-likelihood ratio tests for semiparametric multivariate copula model selection. The Canadian Journal of Statistics 33, 389–414.

Chen, X., Fan, Y., 2006. Estimation and model selection of semi-parametric copula-based multivariate dynamic models under copula misspecification. Journal of Econometrics 135, 125–154.

Demarta, S., McNeil, A.J., 2005. The $t$-copula and related copulas. International Statistical Review 73, 111–129.

Dobrić, J., Schmid, F., 2007. A goodness of fit test for copulas based on Rosenblatt's transformation. Computational Statistics and Data Analysis (in press).

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

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

Engle, R.F., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of united kingdom inflation. Econometrica 50, 987–1007.

Fang, H.-B., Fang, K., Kotz, S., 2002. The meta-elliptical distributions with given marginals. Journal of Multivariate Analysis 82, 1–16.

Genest, C., Ghoudi, K., Rivest, L.-P., 1995. A semi-parametric estimation procedure of dependence parameters in multivariate families of distributions. Biometrika 82, 543–552.

Genest, C., et al., 2007. Omnibus goodness-of-fit tests for copulas: A review and a power study, Working paper, Université Laval.

Genest, C., Rivest, L.-P., 1993. Statistical inference procedures for bivariate Archimedean copulas. Journal of the American Statistical Association 88, 1034–1043.

Green, P., Hjort, N.L., Richardson, S., 2003. Highly Structured Stochastic Systems. Oxford University Press, Oxford.

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

Joe, H., 1997. Multivariate Models and Dependence Concepts. Chapman & Hall, London.

Kendall, M., Stuart, A., 1967. Inference and Relationship, 2nd ed. In: The Advanced Theory of Statistics, vol. 2. Griffin, London.

Kim, G., et al., 2007. Comparison of semiparametric and parametric methods for estimating copulas. Computational Statistics and Data Analysis 51, 2836–2850.

Kurowicka, D., Cooke, R.M., 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., Cooke, R.M., 2007. Sampling algorithms for generating joint uniform distributions using the vine - copula method. Computational Statistics and Data Analysis 51, 2889–2906.

Kurowicka, D., Cooke, R.M., 2006. Uncertainty Analysis with High Dimensional Dependence Modelling. Wiley, New York.

Ljung, T., Box, G.E.P., 1979. The likelihood function for a stationary autoregressive moving average process. Biometrika 66, 265–270.

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

McNeil, A.J., 2007. Sampling nested Archimedean copulas. Journal of Statistical Computation and Simulation (in press).

Oakes, D., 1994. Multivariate survival distributions. Journal of Nonparametric Statistics 3, 343–354.

Rosenblatt, M., 1952. Remarks on a multivariate transformation. Annals of Mathematical Statistics 27, 832–837.

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

Shih, J.H., Louis, T.A., 1995. Inferences on the association parameter in copula models for bivariate survival data. Biometrics 51, 1384–1399.

Sibuya, M., 1960. Bivariate extreme statistics. Annals of the Institute of Statistical Mathematics 11, 195–210.

Sklar, A., 1959. Fonctions de répartition à *n* dimensions et leurs marges. Publications de l'Institut de Statistique de L'Université de Paris 8, 229–231.

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

Whelan, N., 2004. Sampling from Archimedean copulas. Quantitative Finance 4, 339–352.