Chapter 3

Vines Arise

Roger M. Cooke, Harry Joe and Kjersti Aas

Resources for the Future, and Department of Mathematics, Delft University of Technology
Department of Statistics, University of British Columbia
Norwegian Computing Centre

An introduction to the main idea of vines as graphical models is presented, with various notation and graphs for representing vines. The early history of vines is summarized, together with motivation for their construction. The relation to compatibility of subsets of marginal distributions is given to provide some intuition. Important properties and applications of vines are included.

Contents

3.1. Introduction

A vine is a graphical tool for labeling constraints in high-dimensional distributions. A regular vine is a special case for which all constraints are two-dimensional or conditional two-dimensional. Regular vines generalize trees, and are themselves specializations of something called Cantor trees[5] . Combined with copulae, regular vines have proven to be a flexible tool in high-dimensional dependence modeling. Copulae[24,42] are multivariate distributions with uniform univariate margins. Representing a joint distribution as univariate margins plus copulae allows us to separate the problems of estimating univariate distributions from problems of estimating dependence. This is handy in as much as univariate distributions in many cases can be adequately estimated from data, whereas dependence information is rough hewn, involving summary indicators and judgment[3,30] . Whereas the number of parametric multivariate copula families with flexible dependence is limited, there are many parametric families of bivariate copulae. Regular vines owe their increasing popularity to the fact that they leverage from bivariate copulae and enable extensions to arbitrary dimensions. Sampling theory and estimation theory for regular vines are well developed[2,36] , and model inference has left the post[2,33,34] . Regular vines have proven useful in other problems such as (constrained) sampling of correlation matrices[26,37,38] , building non-parametric continuous Bayesian belief nets[16,17] , and characterizing the set of rank correlation matrices[27] .

This chapter traces the historical development of vines, and summarizes their most important properties. We focus on formulating the main results and indicating their role in the development; for proofs the reader is referred to the original articles. Section 3.2 gives precise definitions, while Section 3.3 describes different types of vines. Section 3.4 on historical origins gives an informal rendering of the main ideas, and Section 3.5 makes the links to compatibility of marginal distributions. Sections 3.6–3.9 treat respectively sampling, model inference and applications.

3.2. Regular Vines

Graphical models called *vines* were introduced in Cooke[9] , Bedford and Cooke[5] and Kurowicka and Cooke[31] . A vine $\mathcal{V}$ on n variables is a nested set of connected trees $\mathcal{V} = \{T_1, \ldots, T_{n-1}\}$ where the edges of tree j are the nodes of tree $j+1$, $j = 1, \ldots, n-2$. A *regular vine* on n variables is a vine in which two edges in tree j are joined by an edge in tree $j+1$ only if these edges share a common node, $j = 1, \ldots, n-2$. The formal definitions follow (based on Section 4.4.1 of Kurowicka and Cooke[33]).

Definition 3.1 (Regular vine). *$\mathcal{V}$ is a regular vine on n elements with $E(\mathcal{V}) = E_1 \cup \cdots \cup E_{n-1}$ denoting the set of edges of $\mathcal{V}$ if*

1. $\mathcal{V} = \{T_1, \ldots, T_{n-1}\}$,
2. *T_1 is a connected tree with nodes $N_1 = \{1, \ldots, n\}$, and edges E_1;*
for $i = 2, \ldots, n-1$ T_i is a tree with nodes $N_i = E_{i-1}$.
3. (proximity) *for $i = 2, \ldots, n-1$, $\{a, b\} \in E_i, \#(a \triangle b) = 2$ where $\triangle$ denotes the symmetric difference operator and $\#$ denotes the cardinality of a set.*

An edge in tree T_j is an unordered pair of nodes of T_j, or equivalently, an unordered pair of edges of T_{j-1}. By definition, the *order* of an edge in tree T_j is $j-1$, $j = 1, \ldots, n-1$. The degree of a node is the number of edges attached to that node. A regular vine is called a *canonical* or *C-vine* if each tree T_i has a unique node of degree $n-i$, hence has maximum degree. A regular vine is called a *D-vine* if all nodes in T_1 have degree not higher than

2 (see Figs. 3.3, 3.2). [a]

The constraint, conditioning, and the conditioned set of an edge are defined as follows:

Definition 3.2.

1. For $e \in E_i, i \le n-1$, the **constraint set** associated with e is the **complete union** U_e^* of e, that is, the subset of $\{1, \ldots, n\}$ reachable from e by the membership relation.
2. For $i = 1, \ldots, n-1, e \in E_i$, if $e = \{j, k\}$ then the **conditioning set** associated with e is

$$D_e = U_j^* \cap U_k^*$$

and the **conditioned set** associated with e is

$$\{C_{e,j}, C_{e,k}\} = \{U_j^* \setminus D_e,\ U_k^* \setminus D_e\}.$$

Note that for $e \in E_1$, the conditioning set is empty. One can see that the order of an edge is the cardinality of its conditioning set. For $e \in E_i, i \le n-1, e = \{j, k\}$ we have $U_e^* = U_j^* \cup U_k^*$.

Fig. 3.1 shows a regular vine (left) and a non-regular vine (right). Fig. 3.2 shows a D-vine on five variables with the constraint sets added. Conditioning variables are shown to the right of '|', conditioned variables to the left. The trees at each echelon are drawn in a different style. Fig. 3.3 shows similar information for the C-vine. Although the D-vine looks simpler, in many ways the C- (for canonical) vine is simpler mathematically. Compare the algorithms 1 and 2 for maximum likelihood estimation in Section 3.7.

The following propositions of regular vines are proved in:[5,31,35]

[a]The term *canonical vine* first appears in Bedford and Cooke[4] , with abbreviation of C-vine in Kurowicka and Cooke[33] ; the term D-vine first appears in Kurowicka and Cooke[33,35] . The designation 'D' has nothing to recommend it, beyond being the letter to follow 'C' but it is linked to drawable on page 93 of Kurowicka and Cooke[33] (unfounded is the suggestion that D-vine is an irreverent pun).

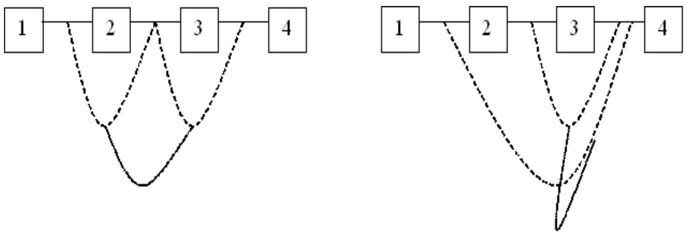

Figure 3.1. A regular (left) and a non regular (right) vine on 4 variables.

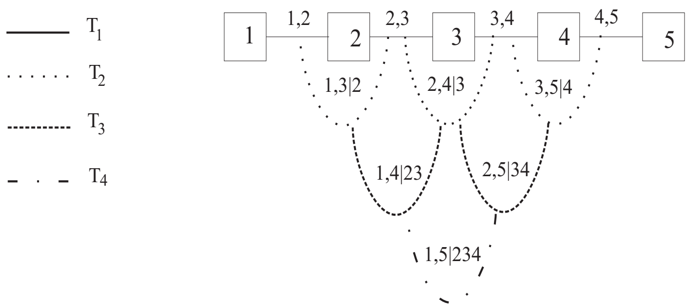

Figure 3.2. A D-vine on 5 variables with constraint sets.

Proposition 3.1. *Let $\mathcal{V} = \{T_1, \ldots, T_{n-1}\}$ be a regular vine, then*

(1) the number of edges is $n(n-1)/2$,
(2) each conditioned set is a doubleton, each pair of variables occurs exactly once as a conditioned set,
(3) if two edges have the same conditioning set, then they are the same edge.

Definition 3.3 (m-child; m-descendent). *If node e is an element of node f, we say that e is an* **m-child** *of f; similarly, if e is reachable from f via the membership relation: $e \in e_1 \in \cdots \in f$, we say that e is an* **m-descendent** *of f.*

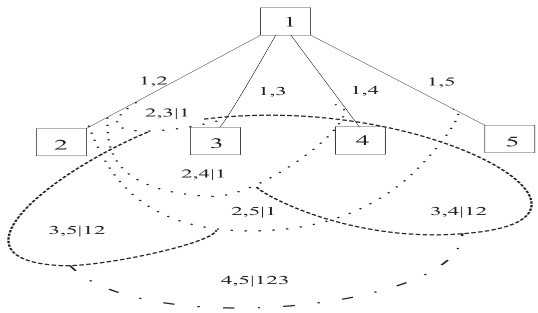

Figure 3.3. A $C-$vine on 5 variables with constraint sets.

Proposition 3.2. *For any node K of order $k > 0$ in a regular vine , if variable i is a member of the conditioned set of K, then i is a member of the conditioned set of exactly one of the m-children of K, and the conditioning set of an m-child of K is a subset of the conditioning set of K.*

The search for an optimal vine requires a method for enumerating and searching all vines. The number of regular vines grows very quickly. A closed formula for the number of regular vines on n elements was found in Morales Napoles et al.[41] :

Theorem 3.1.

(1) For any regular vine on $n-1$ elements, the number of regular n-dimensional vines which extend this vine is 2^{n-3}.
(2) There are $\binom{n}{2} \times (n-2)! \times 2^{(n-2)(n-3)/2}$ labeled regular vines in total.

Note that the number of extensions of a regular vine does not depend on the vine itself.

From Kurowicka and Cooke[33] (see also Chapter **??** in this volume), we have that for $n = 3$, all vines are in the same equivalent class, and for $n = 4$, all regular vines are either C-vines or D-vines. For $n \geq 5$, there are many vines that are neither C-vines nor D-vines. However, the C-vines and D-vines are boundary cases of the possible vines. An extension to non-regular vines is presented in Bedford and Cooke[5] .

We conclude this subsection with some examples to illustrate the notation. The examples consist of a C-vine for $n = 3$, a general C-vine, a general D-vine, a D-vine for $n = 4$, and a vine for $n = 5$ that is neither a C-vine nor a D-vine.

For a C-vine for $n = 3$, in T_1, $N_1 = \{1, 2, 3\}$ and $E_1 = \{\{1, 2\}; \{1, 3\}\} = \{1, 2; 1, 3\} = \{12; 13\}$; then in T_2, $N_2 = E_1$, and $E_2 = \{[\{1, 2\}; \{1, 3\}]\} = \{2, 3|1\} = \{23|1\}$. The shorthand notation with fewer commas and braces is used for simplicity. For the edge $e = 23|1$ in T_2, $U_j^* = \{1, 2\}$, $U_k^* = \{1, 3\}$, the conditioning set is $D_e = \{1, 2\} \cap \{1, 3\} = \{1\}$, $C_{e,j} = \{1, 2\} \setminus \{1\} = \{2\}$, $C_{e,k} = \{1, 3\} \setminus \{1\} = \{3\}$, and the conditioned set is $C_{e,j} \cup C_{e,k} = \{2, 3\}$.

For a general C-vine in n variables with standard indexing, $E_1 = \{1, i : i = 2, \ldots, n\}$, $E_2 = \{2, i|1 : i = 3, \ldots, n\}$, ..., $E_\ell = \{\ell, i|1, \ldots, i - 1 : i = \ell + 1, \ldots, n\}$, ..., $E_{n-1} = \{n - 1, n|1, \ldots, n - 2\}$, $T_1 = \{1, 2, \ldots, n\}$ and $T_\ell = E_{\ell-1}$ for $\ell = 2, \ldots, n - 1\}$. For an edge $e = [i_1, i_2|1, \ldots, i_1 - 1]$ with $1 \le i_1 < i_2 \le n$, the conditioning set is $D_e = \{1, \ldots, i_1 - 1\}$ and the conditioned set is $\{i_1, i_2\}$. If the indices $\{1, \ldots, n\}$ are permuted, the result is still a C-vine, since the C-vine is characterized by the degrees of the nodes for $T_1, \ldots, T_{n-1}$.

For a general D-vine in n variables with standard indexing, $E_1 = \{i, i+1 : i = 1, \ldots, n - 1\}$, $E_2 = \{i, i + 2|i + 1 : i = 1, \ldots, n - 2\}$, ..., $E_\ell = \{i, i+\ell|i+1, \ldots, i+\ell-1 : i = 1, \ldots, n-\ell\}$, ..., $E_{n-1} = \{1, n|2, \ldots, n-1\}$, $T_1 = \{1, 2, \ldots, n\}$ and $T_\ell = E_{\ell-1}$ for $\ell = 2, \ldots, n - 1\}$. For an edge $e = [i_1, i_2|i_1 + 1, \ldots, i_2 - 1]$ with $1 \le i_1 < i_2 \le n$, the conditioning set is $D_e = \{i_1 + 1, \ldots, i_2 - 1\}$ and the conditioned set is $\{i_1, i_2\}$. If the indices $\{1, \ldots, n\}$ are permuted, the result is still a D-vine.

Specific details in shorthand notation for the D-vine with $n = 4$ are: $N_1 = \{1, 2, 3, 4\}$, $E_1 = \{12; 23, 34\}$; then in T_2, $N_2 = E_1$ and $E_2 = \{13|2; 24|3\}$; finally, in T_3, $N_3 = E_2$ and $E_3 = \{14|23\}$. For the edge in T_3, the conditioning set is $D_e = \{2, 3\}$ and the conditioned set is $\{1, 4\}$.

An example of a 5-dimensional regular vine that is neither a C-vine nor a D-vine is shown in Fig. 3.4 $E_1 = \{12; 23; 24; 45\}$, $E_2 = \{13|2; 14|2; 25|4\}$, $E_3 = \{34|12; 15|24\}$, $E_4 = \{35|124\}$.

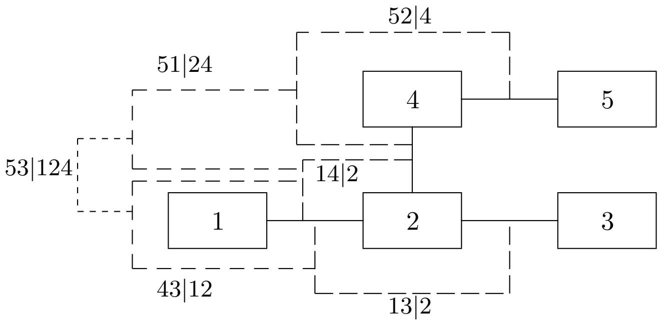

Figure 3.4. A regular vine on 5 variables which is neither a $C-$vine nor a D-vine with constraint sets.

3.3. Vine types

Two main types of regular vines have been treated in the literature; vine copulae and partial correlation vine representations. Vine copulae or pair copula constructions are obtained by assigning a bivariate copula to each edge in the vine. Similarly a partial correlation vine representation of a correlation matrix is obtained by assigning a partial correlation to each edge in the vine. In this section the two types of specifications are discussed.

3.3.1. *Vine copula or pair-copula construction*

A *bivariate copula vine specification* is called a *pair-copula construction*,[1,2] or a *vine copula* (Section 4.4.2 of Kurowicka and Cooke[33]). It is obtained by assigning a bivariate copula C_e for each edge e in the union $E(\mathcal{V}) = E_1 \cup \cdots \cup E_{n-1}$ of the vine defined in the preceding subsection. The set of $\binom{n}{2}$ copulae is denoted by B. The elements of B can be chosen independently of each other (as long as they are bivariate copulae); this follows from Bedford and Cooke[4] .

In general, the form of the joint density of a regular vine copula with margins $F_1, \ldots, F_n$ is given by the following theorem:

Theorem 3.2 (Bedford and Cooke[4]). *Let $\mathcal{V} = (T_1, \ldots, T_{n-1})$ be a regular vine on n elements. For an edge $e \in E(\mathcal{V})$ with conditioned elements e_1, e_2 and conditioning set D_e, let the conditional copula and copula density be $C_{e_1,e_2|D_e}$ and $c_{e_1,e_2|D_e}$, respectively. Let the marginal distributions F_i with densities $f_i, i = 1, \ldots, n$ be given. Then the vine-dependent distribution is uniquely determined, and has a density given by*

$$f_{1\cdots n} = f_1 \cdots f_n \prod_{e \in E(\mathcal{V})} c_{e_1,e_2|D_e}(F_{e_1|D_e}, F_{e_2|D_e}). \tag{3.1}$$

Equation (3.1) shows that vine copulae have closed form densities when $F_1, \ldots, F_n$ and the bivariate copulae in B are differentiable.

Note that C_e is a marginal bivariate copula for edges in T_1 and C_e is a conditional bivariate copula for edges in $T_2, \ldots, T_{n-1}$. For a C-vine, the set of bivariate copulae is denoted as $B = \{C_{i_1 i_2|1,\ldots,i_1-1} : 1 \le i_1 < i_2 \le n\} = \{C_{12}; \ldots; C_{1n}; C_{23|1}; \ldots; C_{2n|1}; \ldots, C_{n-1,n|1,\ldots,n-2}\}$. For a D-vine, the set of bivariate copulae is denoted as $B = \{C_{i_1 i_2|i_1+1,\ldots,i_2-1} : 1 \le i_1 < i_2 \le n\} = \{C_{12}; \ldots; C_{n-1,n}; C_{13|2}; \ldots; C_{n-2,n|n-1}; \ldots, C_{1,n|2,\ldots,n-1}\}$. For the regular vine in Section 3.2 that is not a C-vine or a D-vine, the set of bivariate copulae is: $B = \{C_{12}; C_{23}; C_{24}; C_{45};\ C_{13|2}; C_{14|2}; C_{25|4}; C_{34|12}; C_{15|24}; C_{35|124}\}$.

For applications, univariate margins $F_1, \ldots, F_n$ are specified or estimated, as well as the marginal or conditional copulae in B. The resulting multivariate distribution in the Fréchet class $\mathcal{F}(F_1, \ldots, F_n)$ has a form that can be shown recursively. We show the results for a C-vine with $n = 3$ and and a D-vine with $n = 4$.

First note that, assuming F_1, F_2, C_{12} are differentiable with respective densities f_1, f_2, c_{12}, then $F_{12} = C_{12}(F_1, F_2)$ has density $f_{12} = c_{12}(F_1, F_2)\, f_1 f_2$ and conditional density $f_{2|1} = f_{12}/f_1 = c_{12}(F_1, F_2)\, f_2$.

For the C-vine with $n = 3$, the trivariate distribution comes from the specification $\{F_1, F_2, F_3, C_{12}, C_{13}, C_{23|1}\}$. The $(1, 2)$ and $(2, 3)$ margins are $F_{12} = C_{12}(F_1, F_2)$ and $F_{13} = C_{13}(F_1, F_3)$, from which conditional distribution $F_{2|1}, F_{3|1}$ can be obtained; then

$$F_{123}(x_1, x_2, x_3) = \int_{-\infty}^{x_1} C_{23|1}\big(F_{2|1}(x_2|z), F_{3|1}(x_3|z)\big)dF_1(z). \tag{3.2}$$

If F_i are differentiable with respective densities f_i, $i = 1, 2, 3$, and $C_{12}, C_{13}, C_{23|1}$ have densities $c_{12}, c_{13}, c_{23|1}$ respectively, then the conditional densities $f_{2|1}, f_{3|1}$ exist, and the mixed third order derivative of (3.2), compare Theorem 3.2, is:

$$\begin{aligned}
&f_{123}(x_1, x_2, x_3) = c_{23|1}\big(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)\big)\, f_{2|1}(x_2|x_1)\, f_{3|1}(x_3|x_1)\, f_1(x_1)\\
&= c_{23|1}\big(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)\big)\, c_{12}\big(F_1(x_1), F_2(x_2)\big)\, f_2(x_2)\\
&\quad \cdot c_{13}\big(F_1(x_1), F_3(x_3)\big)\, f_3(x_3)\, f_1(x_1)\\
&= c_{12}\big(F_1(x_1), F_2(x_2)\big)\, c_{13}\big(F_1(x_1), F_3(x_3)\big)\, c_{23|1}\big(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)\big)\\
&\quad \cdot \prod_{i=1}^{3} f_i(x_i)
\end{aligned}$$

For the D-vine with $n = 4$, the 4-variate distribution comes from the specification $\{F_1, F_2, F_3, F_4, C_{12}, C_{23}, C_{34}, C_{13|2}, C_{24|3}, C_{14|23}\}$. The $(i, i+1)$ margins are $F_{i,i+1} = C_{i,i+1}(F_i, F_{i+1})$, F_{123} and F_{234} have expressions like (3.2), and then

$$\begin{aligned}
&F_{1234}\ (x_1, x_2, x_3, x_4) = \\
&\qquad \int_{-\infty}^{x_2}\int_{-\infty}^{x_3} C_{14|23}\big(F_{1|23}(x_1|z_2, z_3), F_{4|23}(x_4|z_2, z_3)\big)dF_{23}(z_2, z_3).
\end{aligned} \tag{3.3}$$

If F_i are differentiable with respective densities f_i, $i = 1, 2, 3, 4$, and C_e have densities c_e for edges e in this vine, then the mixed fourth order derivative of (3.3) is:

$$
\begin{aligned}
&f_{1234}(x_1, x_2, x_3, x_4) \\
&= c_{14|23}\big(F_{1|23}(x_1|x_2, x_3), F_{4|23}(x_4|x_2, x_3)\big)\, f_{1|23}(x_1|x_2, x_3) \\
&\quad \times f_{4|23}(x_4|x_2, x_3)\, f_{23}(x_2, x_3) \\
&= c_{14|23}\big(F_{1|23}(x_1|x_2, x_3), F_{4|23}(x_4|x_2, x_3)\big)\, f_{123}(x_1, x_2, x_3) \\
&\quad \times f_{234}(x_2, x_3, x_4)/f_{23}(x_2, x_3) \\
&= c_{14|23}\big(F_{1|23}(x_1|x_2, x_3), F_{4|23}(x_4|x_2, x_3)\big) \\
&\quad \times c_{13|2}\big(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)\big) c_{24|3}\big(F_{2|3}(x_2|x_3), F_{4|3}(x_4|x_3)\big) \\
&\quad \times c_{12}\big(F_1(x_1), F_2(x_2)\big)\, c_{23}\big(F_2(x_2), F_3(x_3)\big)\, c_{34}\big(F_3(x_3), F_4(x_4)\big) \cdot \prod_{i=1}^{4} f_i(x_i)
\end{aligned}
$$

In applications of vine copulae to date, a parameter (vector) is associated with each $C_e \in B$, and then statistical inference can proceed with maximum likelihood; see Section 3.7.

Normal copulae When each bivariate copula C_e is a bivariate normal copula, then the resulting multivariate copula is a multivariate normal copula. For a multivariate normal copula represented as a vine, there is a correlation or partial correlation parameter associated with each C_e, and the parameters can be summarized into a *partial correlation vine*; see Section 3.3.2. Moreover, since, the multivariate normal has the property that conditional correlations do not depend on the values of the conditioning variables, any multivariate normal copula has many representations as a vine copula. It can be shown that also the multivariate t_ν copulae are special cases of vine copulae.

Dependence properties The following dependence properties for vine copulae are shown in Joe[23] and Joe et al.[28] :

(1) Let edge e be in E_ℓ with $\ell > 1$ and let the conditioned set for e be $\{e_1, e_2\}$. If C_e is more concordant than C'_e, then the margin F_{e_1,e_2} is more concordant than F'_{e_1,e_2}.

(2) If C_e has upper (lower) tail dependence for all $e \in E_1$, and the remaining copulae have support on $[0,1]^2$, all bivariate margins of $F_{1\cdots n}(x_1,\ldots,x_n)$ have upper (lower) tail dependence.
(3) For parametric vine copulae with a parameter θ_e associated with C_e, a wide range of dependence is obtained if each $C_e(\cdot;\theta_e)$ can vary from the bivariate Fréchet lower bound to the Fréchet upper bound. Consider the Kendall tau triple $(\tau_{12},\tau_{13},\tau_{23})$ for $n=3$. It is shown in Joe[23] for a 3-dimensional vine copula that if $C_{23|1}$ is the conditional Fréchet upper (lower) bound copula, then τ_{23} achieves the maximum (minimum) possible, given τ_{12},τ_{13}.

3.3.2. *Partial correlation vine*

In this section we first give the definition of partial correlation. Then, we describe the partial correlation vine structure and finally, we mention two applications of partial correlation vines.

3.3.2.1. *Partial correlation*

A partial correlation can be defined in terms of partial regression coefficients. Consider variables X_i with zero mean and standard deviations $\sigma_i=1$, $i=1,\ldots,n$. Let the numbers $b_{ij;\{1,\ldots,n\}\setminus\{i,j\}}$ minimize

$$E\Big[\Big(X_i-\sum_{j:j\neq i} b_{ij;\{1,\ldots,n\}\setminus\{i,j\}}\Big)^2\Big],\quad i=1,\ldots,n.$$

Definition 3.4 (Partial correlation). *The partial correlation of variables 1 and 2 given the remaining variables is:*

$$\rho_{12;3,\ldots,n}=sgn(b_{12;3,\ldots,n})\left(b_{12;3,\ldots,n}b_{21;3,\ldots,n}\right)^{1/2}.$$

By permuting the indices, other partial correlations on n variables are defined.

Equivalently we could define the partial correlation as

$$\rho_{12;3,\ldots,n}=-\frac{K_{12}}{\sqrt{K_{11}K_{22}}},$$

where K_{ij} denotes the (i, j) cofactor of the correlation matrix. The partial correlation $\rho_{12;3,\ldots,n}$ can be interpreted as the correlation between the orthogonal projections of X_1 and X_2 on the plane orthogonal to the space spanned by $X_3, \ldots, X_n$.

Partial correlations can be computed from correlations with the following recursive formula[49] :

$$\rho_{12;3,\ldots,n} = \frac{\rho_{12;3,\ldots,n-1} - \rho_{1n;3,\ldots,n-1} \cdot \rho_{2n;3,\ldots,n-1}}{\sqrt{1-\rho^2_{1n;3,\ldots,n-1}}\sqrt{1-\rho^2_{2n;3,\ldots,n-1}}}. \tag{3.4}$$

3.3.2.2. *Partial Correlation Vine*

A *partial correlation vine*[5,31,38] , which is a useful parametrization for a multivariate normal or elliptical distribution, is obtained by assigning a partial correlation ρ_e, with a value chosen arbitrarily in the interval $(-1, 1)$, for each edge e in the union $E(\mathcal{V}) = E_1 \cup \cdots \cup E_{n-1}$ of the vine defined in Section 3.2. Note that ρ_e is a correlation for edges in T_1 and ρ_e is a partial correlation for edges in $T_2, \ldots, T_{n-1}$. Theorem 3.3 in Bedford and Cooke[5] shows that a regular vine provides a bijective mapping from $(-1, 1)^{\binom{n}{2}}$ into the set of positive definite matrices with 1's on the diagonal.

Theorem 3.3. *For any regular vine on n elements there is a one-to-one correspondence between the set of $n \times n$ positive definite correlation matrices and the set of partial correlation specifications for the vine.*

All assignments of the numbers between -1 and 1 to the edges of a partial correlation regular vine are consistent, and all correlation matrices can be obtained this way. Specific examples of partial correlation vines are the following: For a C-vine, the set of partial correlations is $\{\rho_{i_1 i_2;1,\ldots,i_1-1} : 1 \leq i_1 < i_2 \leq n\} = \{\rho_{12}, \ldots; \rho_{1n},\ \rho_{23;1}, \ldots, \rho_{2n;1}, \ldots,\ \rho_{n-1,n;1,\ldots,n-2}\}$. For a D-vine, the set of partial correlations is $\{\rho_{i_1 i_2;i_1+1,\ldots,i_2-1} : 1 \leq i_1 < i_2 \leq n\} = \{\rho_{12}, \ldots, \rho_{n-1,n}, \rho_{13;2}, \ldots, \rho_{n-2,n;n-1}, \ldots, \rho_{1,n;2,\ldots,n-1}\}$. For the regular vine in the Section 3.2 that is not a C-vine or a D-vine, the set of partial correlations is: $\{\rho_{12}, \rho_{23}, \rho_{24}, \rho_{45},\ \rho_{13;2}, \rho_{14;2}, \rho_{25;4},\ \rho_{34;12}, \rho_{15;24}, \rho_{35;124}\}$.

One verifies that the correlation between ith and jth variables can be computed from the sub-vine generated by the constraint set of the edge whose conditioned set is $\{i, j\}$ using recursively the formulae (3.4), and the following lemma[31] .

Lemma 3.1. *If $z, x, y \in (-1, 1)$, then also $w \in (-1, 1)$, where*

$$w = z\sqrt{(1 - x^2)(1 - y^2)} + xy.$$

A regular vine may thus be seen as a way of picking out partial correlations which uniquely determine the correlation matrix and which are algebraically independent. The partial correlations in a partial correlation vine need not satisfy any algebraic constraint like positive definiteness. The 'completion problem' for partial correlation vines is therefore trivial. An incomplete specification of a partial correlation vine may be extended to a complete specification by assigning arbitrary numbers in the $(-1, 1)$ interval to the unspecified edges in the vine.

Partial correlation vines have another important property; the product of 1 minus the square partial correlations equals the determinant of the correlation matrix.

Theorem 3.4. *(Kurowicka and Cooke[35]) Let D be the determinant of the n-dimensional correlation matrix ($D > 0$). For any partial correlation vine*

$$D = \prod_{e \in E(\mathcal{V})} (1 - \rho^2_{e_1,e_2;D_e}). \tag{3.5}$$

3.3.2.3. *Applications*

We mention two applications of partial correlation vines. One is the generation of random correlation matrices $\boldsymbol{R}$ that are uniform over the space of correlation matrices. Another is a reparametrization for statistical models where $\boldsymbol{R}$ is a parameter.

Random correlation matrices In Joe[26] for the partial correlation D-vine and in Lewandowski et al.[38] for the general partial correlation vine, results and algorithms are given for generating a random correlation matrix $\boldsymbol{R}$ based on a partial correlation vine. This random correlation matrix generation is based on the property that the partial correlations in a regular vine are algebraically independent. By choosing the distributions of ρ_e to be appropriate Beta distributions on $(-1, 1)$, Lewandowski et al.[38] have developed a method to obtain random correlation matrices that have a uniform density or, more generally, a density proportional to $|\boldsymbol{R}|^{\eta-1}$ where $\eta > 0$.

Numerically, use of the partial correlation C-vine is fastest, but one might want to use a specific regular vine, if there is an indexing of partial correlations of interest.

Reparametrization of statistical models Statistical models such as the multivariate probit model for ordinal data or the t-copula model have an $n \times n$ correlation matrix $\boldsymbol{R} = (\rho_{ij})$ as a parameter. To avoid checking the positive definiteness constraint in the middle of the numerical maximum likelihood iterations, the correlation matrix can be reparametrized via a partial correlation vine. The idea of reparametrizing the correlation matrix to $n-1$ correlations and $(n-1)(n-2)/2$ partial correlations (D-vine) was applied in Xu[48] as a way of allowing the correlation matrix to be a function of covariates. A more common way to deal with the positive definiteness constraint is to reparametrize via the lower triangular Cholesky matrix $\boldsymbol{A} = (a_{ij})$. The partial correlation C-vine might be a more interpretable parametrization. Note that if $\boldsymbol{R} = \boldsymbol{A}\boldsymbol{A}'$, then

$$
\begin{aligned}
a_{i1} &= \rho_{1i}, \quad i = 1, \ldots, n, \\
a_{ij} &= \rho_{ji;1\cdots j-1} \prod_{k=1}^{j-1} \sqrt{1 - \rho^2_{ki;1\cdots k-1}}\,, \quad j = 3, \ldots, n,\ i = j+1, \ldots, n, \\
a_{ii} &= 1 - \sum_{k=1}^{i-1} a^2_{ik}, \quad i = 2, \ldots, n.
\end{aligned}
$$

That is, each element of $\boldsymbol{A}$ that is below the diagonal is a function of partial correlations in the C-vine.

As shown in Section 5.2 of Kurowicka and Cooke[33] , with the use of expert judgement, it might be convenient to specify the conditional bivariate copulae by first assigning a constant conditional rank correlation to each edge of the vine. For $i = 1, \ldots, n-1$, with $e \in E_i$ having $\{j, k\}$ as the conditioned variables and D_e as the conditioning variables, we associate

$$r_{j,k|D_e}.$$

The resulting structure is called a *conditional rank correlation vine.*

In the elicitation of expert judgement of strengths of dependence, the conditional rank correlation vine avoids the constraints in a matrix of rank correlations. It is shown in Joe[27] , using conditional distributions in the form of D-vines, that in dimensions $n \geq 5$, the possible rank correlation matrices (or correlation matrices of dependent uniform random variables) is smaller than the set of all positive definite matrices with 1 on the diagonal.

3.4. Historical Origins

The first regular vine, avant la lettre, was introduced by Joe[22] . The motive was to extend the bivariate extreme value copula to higher dimensions. Consider a multivariate survival function $G(z_1, \ldots, z_n) = \text{Prob}\,\{Z_1 > z_1, \ldots, Z_n > z_n\}$. If G is 'min-stable', then it satisfies

$$G(tz_1, \ldots, tz_n) = e^{-A(tz_1,\ldots,tz_n)} = e^{-t\,A(z_1,\ldots,z_n)}. \tag{3.6}$$

As shown by Pickands (Galambos[13] , Chapter 5), the family of functions A satisfying this equation is infinite dimensional. Joe's goal was to find finite-dimensional parametric subfamilies that would cover the whole family represented by (3.6). To this end he introduced what would later be called the *D-vine*.

Joe[23] was interested in a class of n-variate distributions with given one-dimensional margins, and $n(n-1)$ dependence parameters, whereby $n-1$ parameters correspond to bivariate margins, and the others correspond to conditional bivariate margins. In the case of multivariate normal distributions, the parameters would be $n-1$ correlations and $(n-1)(n-2)/2$ partial correlations, which were noted to be algebraically independent in $(-1,1)$. Implicit in this remark is the observation that partial correlations on what is now called D-vine provide an algebraically independent parametrization of the set of positive definite correlation matrices.

One main idea[23,24] comes from the Fréchet class $\mathcal{F}(F_{iS}, F_{jS})$ where S is a set of indices of variables that does not contain i and j. That is, $\mathcal{F}(F_{iS}, F_{jS})$ is the class of distributions of cardinality $|S|+2$ with the margin F_S in common.

If $S = \{k_1, \ldots, k_m\}$ with $m \geq 1$, a member of $\mathcal{F}(F_{iS}, F_{jS})$ has the form

$$\int_0^{x_{k_1}} \cdots \int_0^{x_{k_m}} F_{ij|S}(x_i, x_j | \boldsymbol{y}_S)\, dF_S(\boldsymbol{y}_S). \tag{3.7}$$

By Sklar's theorem, there are conditional copulae $\{C_{ij|S}(\cdot|\boldsymbol{y}_S)\}$ such that (3.7) is

$$\int_0^{x_{k_1}} \cdots \int_0^{x_{k_m}} C_{ij|S}\big(F_{i|S}(x_i|\boldsymbol{y}_S), F_{j|S}(x_y|\boldsymbol{y}_S) \mid \boldsymbol{y}_S\big)\, dF_S(\boldsymbol{y}_S). \tag{3.8}$$

By imitating multivariate Gaussian distributions, simpler distributions in $\mathcal{F}(F_{iS}, F_{jS})$ have a constant conditional copula: $C_{ij|S}(\cdot|\boldsymbol{y}_S) \equiv C_{ij|S}$ for all $\boldsymbol{y}_S$. By adding a dependence parameter, one can have a bivariate parametric copula family $C_{ij|S}(\cdot|\cdot;\boldsymbol{\theta})$. A wide range of conditional dependence obtains if $C_{ij|S}(\cdot|\cdot;\boldsymbol{\theta})$ interpolates the Fréchet upper bound, independence and the Fréchet lower bound.

Joe[23,24] applied the above idea of Fréchet classes recursively in a D-vine for $\mathcal{F}(F_{i,i+1,\ldots,j-1}, F_{j,i+1,\ldots,j-1})$ with $1 \leq i < j \leq d$. This was partly motivated by variables which might be indices in time or in a one-dimensional spatial direction. Properties of bivariate tail dependence, ordering by concordance, and range of dependences were obtained. The basic sampling strategy was also outlined.

An entirely different motivation underlay the first formal definition of vines in Cooke[9] . Uncertainty analyses of large risk models, such as those undertaken for the European Union and the US Nuclear Regulatory Commission for accidents at nuclear power plants, involve quantifying and propagating uncertainty over hundreds of variables[15,18] . Dependence information for such studies had been captured with *Markov trees*,[47] which are trees constructed with nodes as univariate random variables and edges as bivariate copulae. For n variables, there are at most $n-1$ edges for which dependence can be specified. New techniques at that time involved obtaining uncertainty distributions on modeling parameters by eliciting experts' uncertainties on other variables which are predicted by the models. These uncertainty distributions are pulled back onto the model's parameters by a process known as probabilistic inversion[14,33] . The resulting distributions often displayed a dependence structure that could not be captured as a Markov tree.

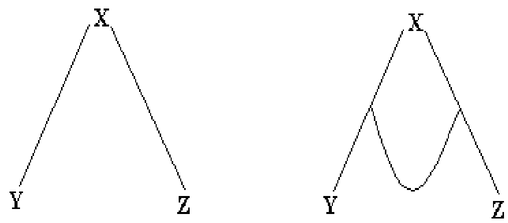

Figure 3.5. A simple Markov tree (left) and vine (right) on 3 variables.

This lead to the invention of regular vines. Regular vines enable an additive decomposition of the mutual information that depends only on the expected mutual information of each edge. Making any conditional copula the conditionally independent copula lowers the mutual information[9] . This remark shows that the minimal information completion of any partially specified regular vine is trivially found by making the unspecified conditional copulae conditionally independent. This situation compares favorably with the problem of completing a partially specified correlation matrix. If a partially specified regular vine has the property that no unspecified edge has specified m-parents, then the partial specification is called *m-saturated*. If we consider the indices in the conditioned sets of a partially specified regular vine, then placing an edge between two indices in the same conditioned set generates a graph. m-saturation is equivalent to the decomposability of this graph, which is equivalent to the graph being chordal and to the existence of a junction tree[35] . Bedford and Cooke[5] extend the result of Joe[23] and show that partial correlations in $(-1, 1)$ on the edges of any regular vine provide an algebraically independent parametrization of the positive definite correlation matrices, and introduce Cantor trees as a generalization of regular vines. Bedford and Cooke[4] give an explicit formula, factorizing any multivariate density in terms of (conditional) copula densities on any regular vine. This generalizes the Hammersley-Clifford theorem applied to Markov trees[6] .

3.5. Compatibility of marginal distributions

n-dimensional vine copulae are based on $\binom{n}{2}$ bivariate copulae which can be specified completely independently of each other. To do this, $n-1$ of the bivariate copulae are bivariate margins and the remaining $(n-1)(n-2)/2$ are conditional copulae. In this section, we provide some results on sets of marginal distributions that can be compatible and hence provide some intuition for the definition of vines and vine copulae (pair-copula construction) in Section 3.2.

The Fréchet class $\mathcal{F}(F_j, 1 \leq j \leq n; F_{jk}, 1 \leq j < k \leq n)$ of given (continuous) univariate and bivariate margins is hard to study; there is no general result on when the set of $\binom{n}{2}$ bivariate margins or copulae are compatible with an n-variate distribution. Assuming that bivariate margins agree on the univariate margins, the maximal number of bivariate margins that can be compatible with no constraints is $n-1$. The maximum is attained if an acyclic condition is satisfied. Consider the Fréchet class $\mathcal{F}(F_j, 1 \leq j \leq n, F_{j_i k_i} : j_i < k_i, i = 1, \ldots, n-1)$ with $n-1$ distinct pairs. This class is non-empty for any choice of the $n-1$ bivariate margins, if the graph with nodes $\{1, \ldots, n\}$ and edges $\{(j_i, k_i) : i = 1, \ldots, n-1\}$ has no cycles. This result follows the compatibility condition in Kellerer,[29] summarized in Section 3.7 of Joe[24] . It also follows from ideas presented below. If any additional bivariate margin $F_{j_n k_n}$ is added, then the graph will definitely have a cycle, and some choices of the $F_{j_i k_i}$ will lead to non-compatibility.

With univariate margins fixed, let's illustrate the above results with bivariate copulae. Because we can permute indices of the variables without changing probabilistic properties, a simple way to get the $n-1$ bivariate margins is to have pairs

$$\{(1,2),(j_2,3),\ldots,(j_{n-1},n)\}, \quad j_i \in \{1,\ldots,i\},\ i=1,\ldots,n-1. \qquad (3.9)$$

The first edge has nodes 1 and 2 and edge $(1,2)$. The second edge adds node 3 and connects to one of the nodes 1 or 2. The ith edge adds node i and connects to one of the nodes between 1 and $i-1$ inclusive. In this way, no cycle is formed, and the result is a tree after $n-1$ steps. If an nth edge is added without adding another node, there will definitely be a cycle (the reader can confirm by drawing some diagrams).

Let's add the nth edge to get a cycle. By relabeling, we can assume that the edges of the cyclical subgraph are $\{(1,2),(2,3),\ldots,(m-1,m),(1,m)\}$ where $3 \le m \le n$. Consider copulae $C_{12}, C_{23}, \ldots, C_{m-1,m}, C_{1m}$. If $C_{12}, C_{23}, \ldots, C_{m-1,m}$ are co-monotonic (Fréchet upper bound) and C_{1m} is counter-monotonic (Fréchet lower bound), then this set of bivariate margins has no compatible n-variate or m-variate distribution.

We next show via an example why $n-1$ bivariate margins satisfying the tree condition (3.9) imply that there are no extra constraints for compatibility. Consider the Fréchet class of bivariate copulae $\mathcal{F}(C_{12}, C_{13}, C_{14}, C_{25})$ for $n=5$; the choice of bivariate margins satisfies (3.9). C_{12}, C_{13} can be specified completely independently because this is the same as specifying the univariate conditional distributions $\{C_{2|1}(\cdot|u_1), C_{3|1}(\cdot|u_1) : 0 < u_1 < 1\}$. For each u_1, $C_{2|1}, C_{3|1}$ can be coupled with a conditional copula, and from (3.7–3.8), one can construct all trivariate copulae with C_{12}, C_{13} as bivariate margins. The same statement holds for the pairs $\{C_{13}, C_{14}\}$. Hence one can get distributions C_{123}, C_{134} with bivariate margins $\{C_{12}, C_{13}, C_{14}\}$. For the resulting C_{123}, C_{134}, one can build a 4-variate copula C_{1234} via

$\{C_{2|13}(\cdot; u_1, u_3), C_{4|13}(\cdot; u_1, u_3) : 0 < u_1, u_3 < 1\}$ and (3.7). After adding C_{25}, one can get a trivariate distribution C_{125} with bivariate margins $\{C_{12}, C_{25}\}$. By coupling the appropriate conditional distributions, one can get C_{1235} with C_{123}, C_{125} as trivariate margins. Finally, one can couple $C_{4|123}(\cdot|u_1, u_2, u_3)$ and $C_{5|123}(\cdot|u_1, u_2, u_3)$ in (3.7) to get a 5-variate copula C_{12345}. Hence $\{C_{12}, C_{13}, C_{14}, C_{25}\}$ is a set of compatible copulae with no additional constraints.

The above example extends for any set of $n-1$ bivariate copulae with pairs of marginal indices of the form of (3.9). This explains the first tree of a vine. However vines also provide conditions for bivariate conditional copulae. Working with bivariate conditional copulae is easier than studying conditions for compatibility of trivariate (and higher-dimensional) margins. We next show that conditions for compatible trivariate margins are more complicated.

For trivariate margins, we can consider a subset of $n-2$ to consider compatibility. For $n=5$, $n-2=3$ and there are three possible patterns of 3 trivariate margins from the full set of $10 = \binom{5}{3}$.

(a) Two indexes appear in all three triplets: e.g., $\{(1,2,3),\ (1,2,4),\ (1,2,5)\}$: this is compatible with copulae for the univariate conditional distributions $3|12$, $4|12$ and $5|12$.
(b) Two of the three pairs intersect in two indices and one pair intersect in one index, e.g., $\{(1,2,3), (1,2,5), (1,3,4)\}$. The above construction shows that something like this will always be compatible.
(c) One of the three pairs intersect in two indices and the other two pairs intersect in one index, e.g., $\{(1,2,3), (2,3,4), (1,4,5)\}$. It is shown in Example 3.4 of Joe[24] that this set of trivariate margins does not satisfy the compatibility condition in Kellerer[29] .

In general, the condition to determine which sets of trivariate margins that always are compatible is more complicated than the condition for bivariate margins given above. Vines are a way to specify a set of compatible bivariate margins and bivariate conditional distributions and they lead to compatible marginal distributions of higher dimensions. The condition for compatibility is straightward to check, compared with something like Kellerer's condition. The examples in this section show why the definition of vines involves tree graphs and conditional distributions.

3.6. Sampling

We assume that variables $X_1, X_2, \ldots, X_n$ are uniform on $(0,1)$. Each edge in a regular vine may be associated with a conditional copula, that is, a conditional bivariate distribution with uniform margins. Given a conditional rank correlation vine as defined in Section 3.3.2.3, we choose a class of copulae indexed by correlation coefficients in the interval $(-1,1)$ and select the copulae with correlation corresponding to the conditional rank correlation assigned to the edge of the vine. A joint distribution satisfying the vine-copula specification can be constructed and sampled on the fly, and will preserve maximum entropy properties of the conditional bivariate distributions[4,9] .

The conditional rank correlation vine plus copula determines the whole joint distribution. There are two strategies for sampling such a distribution, which we term the *cumulative* and *density* approaches.

3.6.1. *Sampling a D-vine*

We first illustrate the cumulative approach with the distribution specified by a D-vine on four variables, D(1,2,3,4): Sample four independent variables distributed uniformly on interval $[0,1]$, U_1, U_2, U_3, U_4 and determine the values of correlated variables X_1, X_2, X_3, X_4 as follows:

(1) $x_1 = u_1$,
(2) $x_2 = F^{-1}_{r_{12};x_1}(u_2)$,
(3) $x_3 = F^{-1}_{r_{23};x_2}\left(F^{-1}_{r_{13|2};F_{r_{12};x_2}(x_1)}(u_3)\right)$,

(4)

$$x_4 = F^{-1}_{r_{34};x_3}\left(F^{-1}_{r_{24|3};F_{r_{23};x_3}(x_2)}\left(F^{-1}_{r_{14|23};F_{r_{13|2};F_{r_{23};x_2}(x_3)}(F_{r_{12};x_2}(x_1))}(u_4)\right)\right)$$

where $F_{r_{ij|k};x_i}(X_j)$ denotes the cumulative distribution function for X_j, applied to X_j, given $X_i = x_i$ under the conditional copula with correlation $r_{ij|k}$. Notice that the D-vine sampling procedure uses conditional and inverse conditional distribution functions. A more general form of the above procedure simply refers to conditional cumulative distribution functions:

$$\begin{aligned}
x_1 &= u_1 \\
x_2 &= F^{-1}_{2|1:x_1}(u_2), \\
x_3 &= F^{-1}_{3|2:x_2}\left(F^{-1}_{3|12:F_{1|2}(x_1)}(u_3)\right), \\
x_4 &= F^{-1}_{4|3:x_3}\left(F^{-1}_{4|23:F_{2|3}(x_2)}\left(F^{-1}_{4|123:F_{1|23}(x_1)}(u_4)\right)\right).
\end{aligned} \tag{3.10}$$

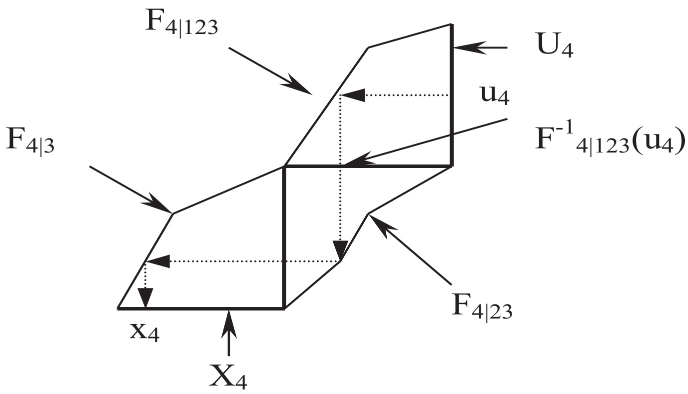

Figure 3.6. Staircase graph representation of D-vine sampling procedure.

Fig. 3.6 depicts the sampling of X_4 in the D-vine with a so-called staircase graph. Following the dotted arrows, we start by sampling U_4 (realization u_4) and use this with the copula for the conditional rank correlation of $\{1,4\}$ given $\{2,3\}$ to find the argument of $F^{-1}_{4|23}$, etc. Notice that for the

D-vine, values of $F_{2|3}$ and $F_{1|23}$ that are used to conditionalize copulae with correlations $r_{24|3}$ and $r_{14|23}$ to obtain $F_{4|23}$ and $F_{4|123}$, respectively, have to be calculated.

The staircase graph shows that if any of the cumulative conditional distributions in Fig. 3.6 is uniform, then the corresponding abscissa and ordinates can be identified. This corresponds to noting that the inverse cumulative function in (3.10) is the identity, and this in turn corresponds to a conditional rank correlation being zero and the corresponding variables being conditionally independent. Notice that the conditional rank correlations can be chosen arbitrarily in the interval $[-1, 1]$; they need not be positive definite or satisfy any further algebraic constraint.

3.6.2. *Sampling an arbitrary Regular Vine*

The content of this section is based on Section 6.4.2 of Kurowicka and Cooke[33] . A regular vine on n nodes will have a single node in tree $n-1$. It suffices to show how to sample one of the conditioned variables in this node, say n, assuming we have sampled all the other variables. We proceed as follows:

(1) By Lemma 3.2, the variable n occurs in trees $1, \ldots, n-1$ exactly once as a conditioned variable. The variable with which it is conditioned in tree j is called its j-partner. We define an ordering for n as follows: index the j-partner of variable n as variable j. We denote the conditional bivariate constraints corresponding to the partners of n as:

$$(n, 1|\emptyset), (n, 2|D_2^n), (n, 3|D_3^n), \ldots, (n, n-1|D_{n-1}^n).$$

Again by Lemma 3.2, variables $1, \ldots, n-1$ appear first as conditioned variables (to the left of '|') before appearing as conditioning variables (to the right of '|'). Also,

$$0 = \#D_1^n < \#D_2^n < \ldots < \#D_{n-1}^n = n-2.$$

(2) Assuming that we have sampled all variables except n, sample one variable uniformly distributed on the interval (0,1), denoted u_n. We use

the general notation $F_{a|b,C}$ to denote $F_{a,b|C:F_{b|C}}$; that is the conditional copula for $\{a, b|C\}$ conditional on a value of the cumulative conditional distribution $F_{b|C}$. Here, $\{a, b|C\}$ is the conditional bivariate constraint corresponding to a node in the vine.

(3) Sample x_n as follows:

$$x_n = F^{-1}_{n|1,D^n_1}\left(F^{-1}_{n|2,D^n_2}\left(\cdots\left(F^{-1}_{n|n-1,D^n_{n-1}}(u_n)\right)\cdots\right)\right). \tag{3.11}$$

The innermost term of (3.11) is:

$$\begin{aligned} F^{-1}_{n|n-1,D^n_{n-1}} &= F^{-1}_{n,n-1|D^n_{n-1}:F_{n-1|D^n_{n-1}}} \\ &= F^{-1}_{n,n-1|D^n_{n-1}:F_{n-1,n-2|D^{n-1}_{n-2}:F_{n-2|D^{n-1}_{n-2}}}}. \end{aligned}$$

See the article of Joe in this volume for pseudocode for the regular vine.

3.6.3. *Density approach sampling*

When the vine-copula distribution is given as a density, the density approach to sampling may be used. Assume that the marginal distributions in (3.1) are uniform [0,1]. Then (3.1) can be rewritten as

$$\prod_{e\in E} c_{ij|D_e}(F_{i|D_e}(x_i), F_{j|D_e}(x_j)), \tag{3.12}$$

where, by uniformity, the density $f_i(x_i) = 1$. Expression (3.12) may be used to sample the vine distribution; namely, draw a large number of samples $(x_1, \ldots, x_n)$ uniformly, and then resample these with probability proportional to (3.12). This is less efficient than the general sampling algorithm given previously; however it may be more convenient for conditionalization.

3.7. Parametric inference for a specific pair-copula construction

Aas et al.[2] develop a maximum likelihood procedure to estimate parameters in copulae for D- and C-vines. The procedure can be extended to arbitrary regular vines, but the algorithms are less transparent.

For notation to cover the C-vine, D-vine and other vines, let $C_{i_1 i_2|m}(u_{i_1}, u_{i_2})$ denote the copula with conditioned set $\{i_1, i_2\}$ and conditioning set m. If $i_1 < i_2$, then $m = \{1, \ldots, i_1 - 1\}$ for the C-vine and $m = \{i_1 + 1, \ldots, i_2 - 1\}$ for the D-vine. For the partial derivatives with respect to u_j and u_{j+i}, we use the notation

$$C_{i_1|i_2:m}(u_{i_1}|u_{i_2}) = \frac{\partial C_{i_1 i_2|m}}{\partial u_{i_2}}, \quad C_{i_2|i_1:m}(u_{i_2}|u_{i_1}) = \frac{\partial C_{i_1 i_2|m}}{\partial u_{i_1}}.$$

The next illustration of notation is for the C-vine in (3.2). If this 3-dimensional distribution is embedded in a C-vine of dimension 4 or more, then the conditional distribution $F_{3|12}$ is needed at the next stage, since (3.2) involves $C_{23|1}$. We use the notation $F_{3|12} = F_{3|2:1}$ to show that it depends on $C_{3|2:1}$. Differentiating (3.2) with respect to x_2, x_3, and then dividing by $f_{12}(x_1, x_2)$ leads to

$$F_{3|2:1}(x_3|x_2, x_1) = C_{3|2:1}\big(F_{3|1}(x_3|x_1)|F_{2|1}(x_2|x_1)\big).$$

Expressions like this must be computed for the likelihood of a C-vine (more generally, a regular vine).

Note that when estimating the parameters here, we assume that the conditional bivariate copulae are constant over the values of the conditioning variables. In the general representation of any multivariate distribution in (3.1) or (3.8), the conditional bivariate copula can vary over the values of the conditioning variables.

Assume that we observe n variables at T time points, or more generally a random sample of size T. Let $\boldsymbol{x}_i = (x_{i,1}, \ldots, x_{i,T})$, $i = 1, \ldots, n$, denote the ith observation vector in 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.

It is important to emphasize that unless the margins are known (which they never are in practice), the estimation method presented below then must rely on the normalized ranks of the data, or on a two-stage procedure where univariate margins have been estimated first and then transformed to uniform. Normalized ranks are only approximately uniform and independent, meaning that what is being maximized is a pseudo-likelihood. A two-stage procedure is better if inferences on tail probabilities are needed; the theory of estimating equations applied for the inference in this case.

3.7.1. *Inference for a C-vine*

In this subsection, we provide an algorithm for computing the log-likelihood of a parametric C-vine where there is a parameter $\theta_{j,i}$ associated with the bivariate copula $C_{j,j+i|1\cdots j-1}(u_j, u_{j+i})$, for $i = 1,\ldots,n-j$, $j = 1,\ldots,n-1$. Here j is the index for the tree level of the vine.

Further, let $\Theta_{j,i}$ be the set of parameters in the copula density $c_{j,j+i|1,\ldots,j-1}(F_{j|1:2\cdots j-1}, F_{j+i|1:2\cdots j-1})$. Note that $F_{j+i|1:2\cdots j-1}$ depends recursively on $\theta_{\ell,k}$, $k = 1,\ldots,j-\ell$ and $j+i-\ell$, $\ell = 1,\ldots,j-1$.

For the canonical vine, the log-likelihood (for the copula parameters, assuming univariate margins have been estimated or transformed to uniform), is given by

$$\begin{aligned}\textstyle\sum_{j=1}^{n-1}\sum_{i=1}^{n-j}\sum_{t=1}^{T} \log \Big[c_{j,j+i|1\cdots j-1}\Big\{&F_{j|j-1:1\cdots j-2}(x_{j,t}|\boldsymbol{x}_t^{(j-1)}),\\ &F_{j+i|j-1:1\cdots j-2}(x_{j+i,t}|\boldsymbol{x}_t^{(j-1)})\Big\}\Big]\end{aligned} \tag{3.13}$$

where $\boldsymbol{x}_t^{(j-1)} = (x_{1,t},\ldots,x_{j-1,t})$. For each copula in the sum (3.13) there is at least one parameter to be determined. The number depends on which copula type is used. The log-likelihood must be numerically maximized over all parameters. If parametric univariate margins are also estimated, say, $f_i(\cdot;\alpha_i)$, $i = 1,\ldots,n$, then the added contribution to (3.13) is

$$\sum_{i=1}^{n}\sum_{t=1}^{T}\log f_i(x_{i,t};\alpha_i).$$

Algorithm 3.1 evaluates the likelihood for the canonical vine. The outer for-loop corresponds to the outer sum in (3.13), corresponding to the tree level of the vine. This for-loop consists in turn of two other for-loops.

The first of these corresponds to the sum over i in (3.13). In the other, the conditional distribution functions needed for the next run of the outer for-loop are computed. In the algorithm,

$$L_{j,j+i}(\boldsymbol{y},\boldsymbol{v},\Theta)=\sum_{t=1}^{T}\log\left\{c_{j,j+i|1\cdots j-1}(y_t,v_t,\Theta)\right\}, \tag{3.14}$$

is the contribution to the log-likelihood from the copula $c_{j,j+i|1\cdots j-1}$.

Algorithm 3.1.

log-likelihood $\leftarrow 0$
for $i \leftarrow 1,\ldots,n$ **do**
  $\boldsymbol{v}_{0,i} \leftarrow \boldsymbol{x}_i$ (vectorized over t)
end for
for $j \leftarrow 1,\ldots,n-1$ **do** (tree level j)
  for $i \leftarrow 1,\ldots,n-j$ **do**
    log-likelihood $\leftarrow$ log-likelihood + $L_{j,j+i}(\boldsymbol{v}_{j-1,1},\boldsymbol{v}_{j-1,i+1},\Theta_{j,i})$
  end for
  if $j == n-1$ **then**
    Stop
  end if
  for $i \leftarrow 1,\ldots,n-j$ **do**
    $\boldsymbol{v}_{j,i} \leftarrow C_{j+i|j:1\cdots j-1}(\boldsymbol{v}_{j-1,i+1}|\boldsymbol{v}_{j-1,1};\Theta_{j,i})$ (vectorized over t)
  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 conditional distributions.
(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 conditional distributions.
(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 in each step.

3.7.2. *Inference for a D-vine*

Similar to the preceding subsection, for the D-vine, the log-likelihood is given by

$$\sum_{j=1}^{n-1}\sum_{i=1}^{n-j}\sum_{t=1}^{T} \log\Big[c_{i,i+j|i+1\cdots i+j-1}\Big\{F_{i|i+j-1:i+1\cdots i+j-2}(x_{i,t}|\boldsymbol{x}_t^{(i,j-1)}),$$
$$F_{i+j|i+1:i+2\cdots i+j-1}(x_{i+j,t}|\boldsymbol{x}_t^{(i,j-1)})\Big\}\Big],$$

where $\boldsymbol{x}_t^{(i,j-1)} = (x_{i+1,t},\ldots,x_{i+j-1,t})$. The D-vine log-likelihood must also be numerically optimised. Algorithm 3.2 evaluates the likelihood. $\Theta_{j,i}$ is the set of parameters of copula density $c_{i,i+j|i+1,\ldots,i+j-1}(F_{i|i+j-1:i+1\cdots i+j-2}, F_{i+j|i+1:i+2\cdots i+j-1})$. Note that the algorithm requires $2(n-j-1)$ conditional distributions at step j for $j=1,\ldots,n-1$. For $j=1$, $C_{i|i+1}, C_{i+1|i}$, $i=1,\ldots,n-1$ are all needed except for $C_{2|1}$ and $C_{n-1|n}$. A similar pattern holds for $j>1$. In the notation in Algorithm 3.2, $v'_{j,i}$ is used in the tree level j when the conditional distribution is $C_{i|i+j:i+1\cdots i+j-1}$ and $v_{j,i}$ is used when the conditional distribution is $C_{i+j|i:i+1\cdots i+j-1}$.

Similar to the C-vine, in the D-vine algorithm,

$$L_{i,i+j}(\boldsymbol{y},\boldsymbol{v},\Theta) = \sum_{t=1}^{T}\log\left\{c_{i,i+j|i+1\cdots i+j-1}(y_t, v_t, \Theta)\right\}$$

is the contribution to the log-likelihood from the copula $c_{i,i+j|i+1\cdots i+j-1}$.

Algorithm 3.2.

log-likelihood $\leftarrow 0$
for $i \leftarrow 1, \ldots, n$ **do**
  $\boldsymbol{v}_{0,i} \leftarrow \boldsymbol{x}_i$ (vectorized over t)
end for
for $i \leftarrow 1, \ldots, n-1$ **do**
  log-likelihood $\leftarrow$ log-likelihood+ $L_{i,1+i}(\boldsymbol{v}_{0,i}, \boldsymbol{v}_{0,i+1}, \Theta_{1,i})$
end for
$\boldsymbol{v}'_{1,1} \leftarrow C_{1|2}(\boldsymbol{v}_{0,1}|\boldsymbol{v}_{0,2}; \Theta_{1,1})$ (vectorized over t; similarly below)
for $k \leftarrow 1, \ldots, n-3$ **do**
  $\boldsymbol{v}_{1,k+1} \leftarrow C_{k+2|k+1}(\boldsymbol{v}_{0,k+2}|\boldsymbol{v}_{0,k+1}; \Theta_{1,k+1})$
  $\boldsymbol{v}'_{1,k+1} \leftarrow C_{k+1|k+2}(\boldsymbol{v}_{0,k+1}|\boldsymbol{v}_{0,k+2}; \Theta_{1,k+1})$
end for
$\boldsymbol{v}_{1,n-1} \leftarrow C_{n|n-1}(\boldsymbol{v}_{0,n}|\boldsymbol{v}_{0,n-1}; \Theta_{1,n-1})$
for $j \leftarrow 2, \ldots, n-1$ **do** (tree level j)
  for $i \leftarrow 1, \ldots, n-j$ **do**
    log-likelihood $\leftarrow$ log-likelihood+ $L_{i,i+j}(\boldsymbol{v}'_{j-1,i}, \boldsymbol{v}_{j-1,i+1}, \Theta_{j,i})$
  end for
  if $j == n-1$ **then**
    Stop
  end if
  $\boldsymbol{v}'_{j,1} \leftarrow C_{1|j+1:2\cdots j}(\boldsymbol{v}'_{j-1,1}|\boldsymbol{v}_{j-1,2}; \Theta_{j,1})$
  if $n > 4$ **then**
    for $i \leftarrow 1, 2, \ldots, n-j-2$ **do**
      $\boldsymbol{v}_{j,i+1} \leftarrow C_{i+j+1|i+1:i+2\cdots i+j}(\boldsymbol{v}_{j-1,i+2}|\boldsymbol{v}'_{j-1,i+1}; \Theta_{j,i+1})$
      $\boldsymbol{v}'_{j,i+1} \leftarrow C_{i+1|i+j+1:i+2\cdots i+j}(\boldsymbol{v}'_{j-1,i+1}|\boldsymbol{v}_{j-1,i+2}; \Theta_{j,i+1})$
    end for
  end if
  $\boldsymbol{v}_{j,n-j} \leftarrow C_{n|n-j:n-j+1\cdots n-1}(\boldsymbol{v}_{j-1,n-j+1}|\boldsymbol{v}'_{j-1,n-j}; \Theta_{j,n-j})$
end for

Note that, similar to other algorithms for C-vines and D-vines, the D-vine algorithm for the likelihood calculation is more complicated than that for the C-vine. Other comments for the C-vine inference also apply to D-vine inference.

3.8. Model Inference

Model inference relates to the problem of choosing a regular vine to model a multivariate data set. If the conditional copulae are not constant, then any regular vine can be used to describe any multivariate distribution. Following Joe[23] , the motive underlying the vine copula approach to modeling is to have a flexible low parameter set of models. In the first instance, this has led to the restriction to constant conditional copulae. When a joint distribution is defined by one particular regular vine with constant conditional copulae, these copulae will not in general remain constant when a different regular vine is used.

In Section 3.7 we described how to do inference for some specific pair-copula decompositions. 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 regular vine, (b) the choice of (conditional) 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 factorizations using the procedure described in Section 3.7 and compare the resulting log-likelihoods, Akaike information criterion (AIC) values, or out-of-sample predictions. This is in practice infeasible for higher dimensions, in view of Theorem 3.1. Heuristic strategies are required to choose which decompositions to investigate. In this section, we review two approaches that have been suggested for choosing the 'best' regular vine; the first is a modified version of the sequential estimation procedure outlined in Section 3.7, while the other is based on the mutual information.

3.8.1. *Sequential selection*

In this approach, one first has to decide whether to use a C- or D-vine. D-vines may be more appropriate than C-vines in situations where a distinguished variable of maximal degree at each echelon cannot readily be identified. The next step is to decide the order of the variables. One possibility that has turned out to be promising in practice is to base this decision on the strength of dependence between the variables, ordering the variables such that the copulae to be fitted in tree 1 in the decomposition are those associated with the strongest dependence.

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 3.7 we need to decide which copula type to use for $C_{12}(\cdot,\cdot)$, $C_{23}(\cdot,\cdot)$ and $C_{13|2}(\cdot,\cdot)$. 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 3.7:

(1) Determine which copula types to use in tree 1 by plotting the original data, and checking for tail dependence or asymmetries (these are the patterns that make the multivariate normal copula inadequate).
(2) Estimate the parameters of the selected copulae using the original data.
(3) Transform observations as required for tree 2, using the copula parameters from tree 1 and the conditional functions in Section 3.7.
(4) Determine which copula types to use in tree 2 in the same way as in tree 1.
(5) 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 a globally optimal fit. Having determined the appropriate parametric shapes for each copulae, one may use the procedures in Section 3.7 to estimate their parameters.

3.8.2. *Information Based Model Inference*

A different approach to model learning inspired by Whittaker[47] was developed in Kurowicka and Cooke[33] , based on the factorization of the determinant in Theorem 3.4. We sketch here a more general approach based on the mutual information. Following Joe[21,25] , the mutual information is taken as a general measure of dependence. The strategy is to choose a regular vine which captures the mutual information in a small number of conditional bivariate terms, and to find a copula which renders these mutual information values. Before describing the approach, we give some definitions.

3.8.2.1. *Definitions and Theorems*

Definition 3.5 (Relative information, Mutual information). *Let f and g be densities on $\mathbb{R}^n$ with f absolutely continuous with respect to g;*

- *the* **relative information** *of f with respect to g is*
$$I(f|g) = \int_1 \cdots \int_n f(x_1, \ldots, x_n) \ \ln\Big(\frac{f(x_1, \ldots, x_n)}{g(x_1, \ldots, x_n)}\Big) dx_1 \ldots dx_n$$
- *the* **mutual information** *of f is*
$$MI(f) = I(f \mid \Pi_{i=i}^n f_i)$$
where f_i is the ith univariate marginal density of f and $\Pi_{i=1}^n f(x_1, \ldots, x_n)$ is the independent distribution with univariate margins $\{f_i\}$.

Relative information is also called the Kullback-Leibler information and the directed divergence. The mutual information is also called the information proper. The mutual information will be used to capture general dependence in a set of multivariate data. We do not possess something like an 'empirical mutual information'. It must rather be estimated with kernel estimators, as suggested in Joe[21] . For some copulae, the mutual information can be expressed in closed form[33] :

Theorem 3.5. *Let g be the elliptical copula with correlation ρ, then the mutual information of g is*

$$1 + \ln 2 + \ln(\pi\sqrt{1-\rho^2}\,).$$

Let h be the diagonal band copula with vertical bandwidth parameter $1-\alpha$, then the mutual information of h is

$$-\ln(2^{|\alpha|}(1-|\alpha|)).$$

Note that the mutual information of the elliptical copula with zero correlation is not zero, reflecting the fact that zero correlation in this case does not entail independence.

Theorem 3.6 (Whittaker[47]). *Let f be a joint normal density with mean vector zero, then*

$$MI(f) = -\tfrac{1}{2}\ln(D),$$

where D is the determinant of the correlation matrix.

For a bivariate normal, Theorem 3.6 says that $MI(f) = -(1/2)\ln(1-\rho^2)$. Substituting the appropriate conditional bivariate normal distributions in the right hand side of (3.15) we find $MI(f) = -1/2\sum_{e\in E(\mathcal{V})}\ln(1-\rho^2_{e_1,e_2;D_e})$, which agrees with Theorem 3.4.

The determinant of a correlation matrix indicates the 'amount of linearity' in a joint distribution. It takes the value 1 if the variables are uncorrelated, and the value zero if there is a linear dependence. Theorem 3.6 suggest that

$$e^{-2MI(f)},$$

is the appropriate generalization of the determinant to capture general dependence.

Proposition 3.3. *$e^{-2MI(f)} = 1$ if and only if $f = \Pi f_i$ and $e^{-2MI(f)} = 0$ if f has positive mass on a set of Πf_i measure zero.*

Theorem 3.7 (Cooke[9] , Bedford and Cooke[5]). *Let g be an n-dimensional density satisfying the bivariate vine specification $(F, \mathcal{V}, B)$ with density g and one-dimensional marginal densities $g_1, \ldots, g_n$; then*

$$I\Big(g \mid \prod_{i=1}^{n} g_i\Big) = \sum_{e \in E(\mathcal{V})} E_{D_e} I(g_{e_1,e_2|D_e} \mid g_{e_1|D_e} \cdot g_{e_2|D_e}). \tag{3.15}$$

If D_e is vacuous, then by definition

$$E_{D_e} I(g_{e_1,e_2|D_e} \mid g_{e_1|D_e} \cdot g_{e_2|D_e}) = I(g_{e_1,e_2} \mid g_{e_1} \cdot g_{e_2}).$$

3.8.2.2. *Strategy for model inference*

Theorem 3.7 may be rewritten as

$$MI(f) = \sum_{\{i,j|K(ij)\} \in \mathcal{V}} b_{ij;K(ij)}, \tag{3.16}$$

where $K(ij)$ is conditioning set for the node in $\mathcal{V}$ with conditioned set $\{i, j\}$. The terms $b_{ij;K(ij)}$ will depend on the regular vine which we choose to represent the dependence structure, however the sum of these terms must satisfy (3.16). We seek a regular vine for which the terms $b_{ij;K(ij)}$ in (3.16) are as 'spread out' as possible. In other words, we wish to capture the total dependence $MI(f)$ in a small number of terms, with the remaining terms being close to zero. This concept is made precise with the notion of *majorization*[39] .

Definition 3.6. Let $\boldsymbol{x}$,$\boldsymbol{y} \in \mathbb{R}^n$ be such that $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$; then $\boldsymbol{x}$ majorizes $\boldsymbol{y}$ if for all k; $k = 1, \ldots, n$

$$\sum_{j=1}^{k} x_{(j)} \leq \sum_{j=1}^{k} y_{(j)}, \tag{3.17}$$

where $x_{(1)} \leq \cdots \leq x_{(j)} \leq \cdots \leq x_{(n)}$ are the increasing arrangement of the components of $\boldsymbol{x}$, and similarly for $\{y_{(j)}\}$ and $\boldsymbol{y}$.

In view of (3.16) the model inference problem may be cast as the problem of finding a regular vine whose terms $b_{ij;K(ij)}$ are non-dominated in the sense of majorization. In that case, setting the smallest mutual informations equal to zero will change the overall mutual information as little as possible. Pairs of variables whose (conditional) mutual information is zero, are (conditionally) independent. Finding non-dominated solutions may be difficult, but a necessary condition for non-dominance can be found by maximizing any Schur convex function.

Definition 3.7. A function $\phi : \mathbb{R}^n \to \mathbb{R}$ is Schur convex if $\phi(\boldsymbol{x}) \geq \phi(\boldsymbol{y})$ whenever $\boldsymbol{x}$ majorizes $\boldsymbol{y}$.

Schur convex functions have been studied extensively. A sufficient condition for Schur convexity is given by Marshall and Olkin[39] .

Proposition 3.4. *If* $\phi : \mathbb{R}^k \to \mathbb{R}$ *may be written as* $\phi(\boldsymbol{x}) = \sum \varphi(x_i)$ *with* φ *convex, then* ϕ *is Schur convex.*

Vine Inference Strategy: The following strategy for model inference suggests itself:

(1) Choose a Schur convex function $\phi : \mathbb{R}^{n(n-1)/2} \to \mathbb{R}$;
(2) Find a regular vine $\mathcal{V}(n)$ whose vector $b_{ij;K(ij)}$ maximizes ϕ;
(3) Set the mutual informations in $\mathcal{V}(n)$ equal to zero for which the terms $b_{ij;K(ij)}$ are smallest;
(4) Associate copulae with the nodes in the vine, such that the non-zero mutual information values are preserved.

A different strategy for model inference is proposed in Chapter **??**.

3.9. Applications

This section references applications in the wide sense, including uses of vine-copula representations of multivariate distributions for mathematical and modeling purposes, as well as applications to analysis of multivariate data.

3.9.1. *Multivariate data analysis*

Due to their high flexibility, yet simple structure, pair-copula constructions/vines are becoming increasingly popular for constructing continuous multivariate distributions. While built exclusively from pair-copulae, they can model a wide range of complex dependencies. Lately, a number of publications on applications of pair-copula constructions have appeared in the literature. Most of the publications treat financial applications[1,2,8,11,12,19,40,45] , while Kolbjornsen and Stien[43] present a non-parametric petroleum related application of pair-copula constructions. For more details on some of these applications, see the application chapters (beginning with Chapter **??**) of this book.

The studies of Berg and Aas[1] and Fischer et al.[12] also compare pair-copula constructions with other multivariate models, e.g., n-dimensional parametric copulae and hierarchical Archimedean constructions[24] , and conclude with the superiority of the pair-copula constructions. In Chapter **??** of this book, a short version of the first-mentioned paper is given.

Biller[7] uses vine copulae for copula-based multivariate time-series input models, and compares with Vector-Autoregressive-To-Anything (VARTA).

3.9.2. *Non-parametric Bayesian Belief Nets*

Bayesian Belief Nets[10,20,44,46] *(BBNs)* are directed acyclic graphs. The nodes of the graph represent univariate random variables, which can be discrete or continuous, and the arcs represent directed influences. BBNs provides a compact representation of high-dimensional uncertainty distributions over a set of variables $(X_1, \ldots, X_n)$ and encode the probability density of these variables by specifying a set of conditional independence statements in a form of an acyclic directed graph and a set of probability functions. In their most popular form, BBNs were introduced in the 1980's as a knowledge representation formalism to encode and use the information acquired from human experts in automated reasoning systems to perform diagnostic and prediction[44] .

Until recently, most BBNs were discrete. Moreover, there were only two ways of dealing with continuous BBNs. One was to discretize the continuous variables and work with the corresponding discrete model, while the other was to assume joint normality. Both these methods have serious drawbacks. Therefore, non-parametrice BBNs (NPBBNs) were introduced in Kurowicka and Cooke[32] and extended in Hanea[16] . In the NPBBNs, nodes are associated with arbitrary continuous invertible distributions and arcs with (conditional) rank correlations that are realized by a chosen copula. No joint distribution is assumed, which makes this BBN non-parametric. Non-parametric BBNs have seen several applications to date, the most notable is a very large model of civil aviation transport safety[3] . There is a close relationship between regular vines and NPBBNs. Chapter **??** in this book provide some insights into the differences and similarities between the two types of models.

In a BBN, the arcs of a directed graph can be associated with conditional copula, where the conditioned variables are the source and sink of the arc, and the conditioning variables are a subset of the other parents of the sink node. These conditional copulae, together with the one-dimensional marginal distributions and the conditional independence statements implied by the BBN graph uniquely determine the joint distribution, and every such specification is consistent[32,33] . This requires a copula type for which zero correlation implies independence. The proof pivots on representing the parents of a child node as a D-vine. When the number of non-independent conditional copulae is not too large, BBNs provide a much more perspicuous representation of the dependence structure than regular vines. In a regular vine all edges must be drawn, even if the conditional copula is independent.

References

1. Aas K. and Berg D., (2009), Models for construction of multivariate dependence — a comparison study. *European J. Finance*, 15:639–659.

2. Aas K., Czado C., Frigessi A. and Bakken H., (2009), Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, 44(2):182–198.
3. Ale B.J.M., Bellamy L.J., van der Boom R., Cooper J., Cooke R.M., Goossens L.H.J., Hale A.R., Kurowicka D., Morales O., Roelen A.L.C. and Spouge J., (2009), Further development of a causal model for air transport safety (cats); building the mathematical heart. *Reliability Engineering and System Safety Journal*.
4. Bedford T.J. and Cooke R.M., (2001), Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence*, 32:245–268.
5. Bedford T.J. and Cooke R.M., (2002), Vines — a new graphical model for dependent random variables. *Ann. of Stat.*, 30(4):1031–1068.
6. Besag J., (1974), Spatial interaction and the statistical analysis of lattice systems. *J. Royal. Stat. Soc. B*, 34:192–236.
7. Biller B.(2009), Copula-based multivariate input models for stochastic simulation. *Operations Research*, 57:878–892.
8. Chollete L., Heinen A. and Valdesogo A., (2009), Modeling international financial returns with a multivariate regime switching copula, Journal of Financial Econometrics, 2009, Vol. 7, No. 4, 437–480.
9. Cooke R.M., (1997), Markov and entropy properties of tree and vines-dependent variables. In *Proceedings of the ASA Section of Bayesian Statistical Science*.
10. Cowell R.G., Dawid A.P., Lauritzen S.L. and Spiegelhalter D.J., (1999), *Probabilistic Networks and Expert Systems*. Statistics for Engineering and Information Sciences. Springer- Verlag, New York.
11. Czado C., Min A., Baumann T. and Dakovic R., (2009), Pair-copula constructions for modeling exchange rate dependence. Technical report, Technische Universität München.
12. Fischer M., Köck C., Schlüter S. and Weigert F., (2009), Multivariate copula models at work. Quantitative Finance, 9(7): 839–854.
13. Galambos J., (1987), *The Asymptotic Theory of Extreme Order Statistics*. Kreiger, Malabor, Fla.
14. Goossens L.H.J., Kraan B.C., Cooke R.M., Jones J.A., Brown J., Ehrhardt J., Fischer F. and Hasemann I., (2001), Methodology and processing techniques. Directorate-General for Research EUR 18827 EN, European Commission, Luxembourg.

15. Goossens L.H.J., Harper F.T., Kraan B.C.P. and Métivier H., (2000), Expert judegement for a probabilistic accident consequence uncertainty analsis. *Radiation Protection Dosimetry*, 90(3):295–301, 2000.
16. Hanea A.M., (2008), *Algorithms for Non-parameteric Bayesian Belief Nets.* PhD thesis, Delft Institute of Applied Mathematics, Delft University of Technology.
17. Hanea A.M., Kurowicka D., Cooke R.M. and Ababei D.A.,(2010), Mining and visualising ordinal data with non-parametric continuous BBNs, *Computational Statistics and Data Analysis,* 54: 668–687.
18. Harper F., Goossens L.H.J., Cooke R.M., Hora S., Young M., Pasler-Ssauer J., Miller L., Kraan B.C.P., Lui C., McKay M., Helton J. and Jones A., (1994), Joint USNRC CEC consequence uncertainty study: Summary of objectives, approach, application, and results for the dispersion and deposition uncertainty assessment. Technical Report VOL. III, NUREG/CR-6244, EUR 15755 EN, SAND94-1453.
19. Heinen A. and Valdesogo A., (2008), Canonical vine autoregressive model for large dimensions. Technical report.
20. Jensen F.V., (2001), *Bayesian Networks and Decision Graphs.* Springer-Verlag, New York.
21. Joe H., (1993), Multivariate dependence measures and data analysis. *Comp. Stat. and Data Analysis*, 16:279–297.
22. Joe H., (1994), Multivariate extreme-value distributions with applications in environmental data. *The Canadian Journal of Statistics*, 22:47–64.
23. Joe H., (1996), Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters. In L. Rüschendorf, B. Schweizer and M. D. Taylor, editor, *Distributions with Fixed Marginals and Related Topics*, volume 28, pages 120–141. IMS Lecture Notes.
24. Joe H., (1997), *Multivariate Models and Dependence Concepts.* Chapman & Hall, London.
25. Joe H., (1999), Relative entropy measures of multivariate dependence. *J. Amer. Stat. Assoc*, 84(405):157–164.
26. Joe H., (2005), Generating random correlation matrices based on partial correlations. *J. of Multivariate Analysis*, 97:2177–2189.
27. Joe H., (2006), Range of correlation matrices for dependent random variables with given marginal distributions. In N. Balakrishnan, E. Castillo and J. M. Sarabia, editor, *Advances in Distribution Theory, Order Statistics and Inference, in honor of Barry Arnold*, pages 125–142. Birkhauser, Boston.

28. Joe H., Li H. and Nikoloulopoulos A.K., (2010), Tail dependence functions and vine copulas. *J. of Multivariate Analysis*, 101: 252–270.
29. Kellerer H.G., (1964), Verteilungsfunktionen mit gegebenen Marginalverteilungen. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 3:247–270.
30. Kraan B.C.P. and Cooke R.M., (2000), Processing expert judgements in accident consequence modeling. *Radiation Protection Dosimetry*, 90(3).
31. Kurowicka D. and Cooke R.M., (2003), A parametrization of positive definite matrices in terms of partial correlation vines. *Linear Algebra and its Applications*, 372:225–251.
32. Kurowicka D. and Cooke R.M., (2004), Distribution-free continuous Bayesian belief nets. In *Mathematical Methods in Reliability.*
33. Kurowicka D. and Cooke R.M., (2006), *Uncertainty Analysis with High Dimensional Dependence Modelling.* Wiley.
34. Kurowicka D., Cooke R.M. and Callies U., (2007), Vines inference. *Brazilian Journal of Praobablitity and Statistics.*
35. Kurowicka D. and Cooke R.M., (2006), Completion problem with partial correlation vines. *Linear Algebra and Its Applications*, 418(1):188–200.
36. Kurowicka D. and Cooke R.M., (2007), Sampling algorithms for generating joint uniform distributions using the vine-copula method. *Computational Statistics and Data Analysis*, 51:2889–2906.
37. Lewandowski D., (2008), *High Dimensional Dependence. Copulae, Sensitivity, Sampling.* PhD thesis, Delft Institute of Applied Mathematics, Delft University of Technology.
38. Lewandowski D., Kurowicka D. and Joe H., (2009), Generating random correlation matrices based on vines and extended onion method, *J. Mult. Anal.*, 100:1989–2001.
39. Marshall A.W. and Olkin I., (1979), *Inequalities: Theory of Majorization and its Applications.* Academic Press, San Diego.
40. Min A. and Czado C., (2008), Bayesian inference for multivariate copulas using pair copula constructions. *Submitted for publication.*
41. Morales Napoles O., Cooke R.M. and Kurowicka D., (2008), The number of vines and regular vines on n nodes. Technical report, Delft Institute of Applied Mathematics, Delft University of Technology.
42. Nelsen R.B., (2006), *An Introduction to Copulas*, 2nd ed. Springer, New York.
43. Kolbjornsen O. and Stien M., (2008), The D-vine creation of non-Gaussian random fields. In *GEOSTATS.*
44. Pearl J., (1988), *Probabilistic Reasoning in Intelligent Systems: Networks of*

Plausible Inference. Morgan Kaufman Publishers, San Mateo.

45. Schirmacher D. and Schirmacher E., (2008), Multivariate dependence modeling using pair-copulas. Technical report, Presented at The 2008 ERM Symposium, Chicago.
46. Shachter R.D. and Kenley C.R., (1989), Gaussian influence diagrams. *Menagement Science*, 35(5):527–550.
47. Whittaker J., (1990), *Graphical Models in Applied Multivariate Statistics*. Wiley, Chichester.
48. Xu J. J., (1996), *Statistical Modelling and Inference for Multivariate and Longitudinal Discrete Response Data*. Ph.D. thesis, Department of Statistics, University of British Columbia.
49. Yule G.U. and Kendall M.G., (1965), *An Introduction to the Theory of Statistics*. Charles Griffin & Co. 14th edition, Belmont, California.