Chapter 2
Selection of Vine Copulas

Claudia Czado, Eike Christian Brechmann, and Lutz Gruber

Abstract Vine copula models have proven themselves as a very flexible class of multivariate copula models with regard to symmetry and tail dependence for pairs of variables. The full specification of a vine model requires the choice of a vine tree structure, the copula families for each pair copula term and their corresponding parameters. In this survey we discuss the different approaches, both frequentist and Bayesian, for these model choices so far and point to open problems.

2.1 Introduction

The analysis of high-dimensional data sets requires flexible multivariate stochastic models that can capture the inherent dependency patterns. The copula approach, which separates the modeling of the marginal distributions from modeling the dependence characteristics, is a natural one to follow in this context. This development has spawned a tremendous increase in copula-based applications in the last 10 years, especially in the areas of finance, economics, and hydrology.

Considerable efforts have been undertaken to increase the flexibility of multivariate copula models beyond the scope of elliptical and Archimedean copulas. Vine copulas are among the best-received of such efforts. Vine copulas use (conditional) bivariate copulas as the so-called pair copula building blocks to describe a multivariate distribution (see [37]). A set of linked trees—the "vine"—describes a vine copula's factorization of the multivariate copula density function into the density functions of its pair copulas (see [8,9]). The article by [1] illustrates a first application of the vine copula concept using non-Gaussian pair copulas to financial data. The first comprehensive account of vine copulas is found in [45], a

C. Czado (✉) · E.C. Brechmann · L. Gruber
Zentrum Mathematik, Technische Universität München, Munich, Germany
e-mail: cczado@ma.tum.de; brechmann@ma.tum.de; lutz.gruber@ma.tum.de

P. Jaworski et al. (eds.), *Copulae in Mathematical and Quantitative Finance*,
Lecture Notes in Statistics 213, DOI 10.1007/978-3-642-35407-6_2,

recent survey in [20], and the current developments of this active research area in [46].

Elliptical copulas as well as Archimedean copulas have been shown to be inadequate models to describe the dependence characteristics of real data applications (see, for example, [1, 23, 24]). As a pair copula construction, vine copulas allow different structural behaviors of pairs of variables to be modeled suitably, in particular so with regard to their symmetry, or lack thereof, strength of dependence, and tail dependencies. Such flexibility requires well-designed model selection procedures to realize the full potential of vine copulas as dependence models. Successful applications of vines can be found, amongst others, in [10, 12, 17, 22, 24, 32, 48, 51, 54, 61].

A parametric vine copula consists of three components: a set of linked trees identifying the pairs of variables and their conditioning variables, the copula families for each pair copula term given by the tree structure, and the corresponding copula parameters. The three-layered definition leads to three fundamental estimation and selection tasks: (1) Estimation of copula parameters for a chosen vine tree structure and pair copula families, (2) Selection of the parametric copula family for each pair copula term and estimation of the corresponding parameters for a chosen vine tree structure, and (3) Selection and estimation of all three model components. In this survey we address these tasks and give an overview of the statistical approaches taken so far. These range from frequentist to Bayesian methods.

The remainder of this paper is structured as follows. In Sect. 2.2 we provide the necessary methodical background on regular vines and regular vine copulas. We then discuss estimation of parameters of regular vine copulas in Sect. 2.3 and the selection of appropriate pair copulas in Sect. 2.4. Section 2.5 treats the joint selection of the regular vine tree structure, the copula families, and their parameters. Section 2.6 concludes with a discussion of available software and open problems.

2.2 Regular Vine Copulas

Copulas describe a statistical model's dependence behavior separately from its marginal distributions [60]. As such, a copula is a multivariate distribution function with all marginal distributions being uniform: i.e. the copula associated with an n-variate cumulative distribution function $F_{1:n}$ with univariate marginal distribution functions $F_1, \ldots, F_n$ is a distribution function $C : [0,1]^n \to [0,1]$ satisfying

$$F_{1:n}(\mathbf{x}) = C\left(F_1(x_1), \ldots, F_n(x_n)\right), \quad \mathbf{x} = (x_1, \ldots, x_n)' \in \mathbb{R}^n.$$

If C is absolutely continuous, its density is denoted by c.

The factorization of multivariate copula densities into (conditional) bivariate copula densities is due to [8, 37], which were developed independently. The details of these factorizations are represented by the graph theoretical construction called regular vine to organize different decompositions. Graphs are defined in terms of a set of nodes N and a set of edges E connecting these nodes, i.e. $E \subset N \times N$. Vines are based on trees which are particular graphs where there is a unique sequence of edges between each two nodes.

Definition 2.1 (Regular Vine Tree Sequence). A set of linked trees $\mathscr{V} = (T_1, T_2, \ldots, T_{n-1})$ is a regular vine (R-vine) on n elements if

1. T_1 is a tree with nodes $N_1 = \{1, \ldots, n\}$ and a set of edges denoted by E_1.
2. For $i = 2, \ldots, n-1$, T_i is a tree with nodes $N_i = E_{i-1}$ and edge set E_i.
3. For $i = 2, \ldots, n-1$, if $a = \{a_1, a_2\}$ and $b = \{b_1, b_2\}$ are two nodes in N_i connected by an edge, then exactly one of the a_i equals one of the b_i (proximity condition).

In other words, the proximity condition requires that the edges corresponding to two connected nodes in tree T_i share a common node in tree T_{i-1}. This ensures that the decomposition into bivariate copulas which is given below is well defined.

Two sub-classes of regular vines have been studied extensively in the literature: canonical vines (C-vines) and drawable vines (D-vines) (see [1, 45]). C-vines are characterized by a root node in each tree T_i, $i \in \{1, \ldots, n-1\}$, which has degree $n - i$; that means that the root node is connected to all other nodes of the tree. D-vines, on the other hand, are uniquely characterized through their first tree which is, in graph theoretical terms, a path; this means that each node has degree of at most 2. Therefore the order of variables in the first tree defines the complete D-vine tree sequence.

Some more definitions are needed to introduce regular vine copulas: the complete union A_e of an edge $e = \{a, b\} \in E_i$ in tree T_i of a regular vine $\mathscr{V}$ is defined by

$$A_e = \{v \in N_1 : \exists e_m \in E_m,\ m = 1, \ldots, i-1,\ \text{such that } v \in e_1 \in \cdots \in e_{i-1} \in e\}. \tag{2.1}$$

The conditioning set associated with $e = \{a, b\}$ is defined as $D_e := A_a \cap A_b$ and the conditioned sets associated with $e = \{a, b\}$ are defined as $\mathscr{C}_{e,a} := A_a \setminus D_e$ and $\mathscr{C}_{e,b} := A_b \setminus D_e$. Bedford and Cooke [8] showed that the conditioned sets are singletons, and we will therefore refer to edges by their labels $\{\mathscr{C}_{e,a}, \mathscr{C}_{e,b} | D_e\} \mathrel{\hat{=}} \{i(e), j(e) | D(e)\}$. An exemplary regular vine on five elements is shown in Fig. 2.1.

Given these sets, we can specify a regular vine copula by associating a (conditional) bivariate copula with each edge of the regular vine, a so-called pair copula.

Definition 2.2 (Regular Vine Copula). A regular vine copula $C = (\mathscr{V}, \mathscr{B}(\mathscr{V}), \boldsymbol{\theta}(\mathscr{B}(\mathscr{V})))$ in n dimensions is a multivariate distribution function such that for a random vector $\mathbf{U} = (U_1, \ldots, U_n)' \sim C$ with uniform margins

1. $\mathscr{V}$ is a regular vine on n elements,
2. $\mathscr{B}(\mathscr{V}) = \left\{C_{i(e),j(e)|D(e)} | e \in E_m,\ m = 1, \ldots n-1\right\}$ is a set of $n(n-1)/2$ copula families identifying the conditional distributions of $U_{i(e)}, U_{j(e)} | \mathbf{U}_{D(e)}$,
3. $\boldsymbol{\theta}(\mathscr{B}(\mathscr{V})) = \left\{\boldsymbol{\theta}_{i(e),j(e)|D(e)} | e \in E_m,\ m = 1, \ldots n-1\right\}$ is the set of parameter vectors corresponding to the copulas in $\mathscr{B}(\mathscr{V})$.

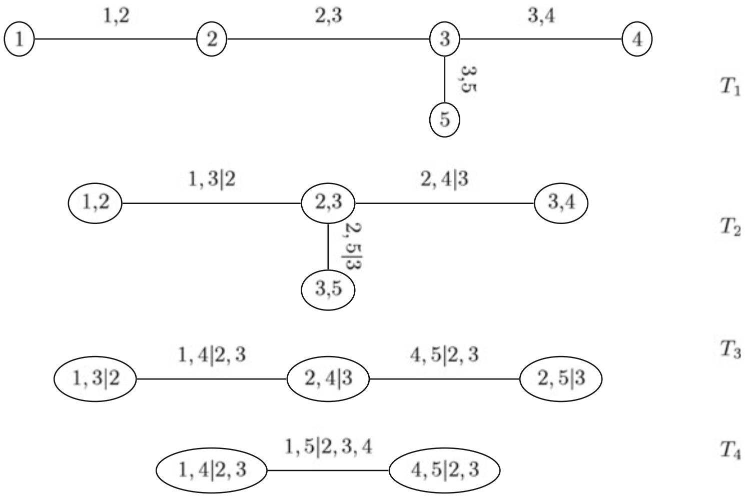

Fig. 2.1 Regular vine on five elements

Therefore the full specification of a regular vine copula consists of three layers: the regular vine tree structure $\mathscr{V}$, the pair copula families $\mathscr{B} = \mathscr{B}(\mathscr{V})$, and the pair copula parameters $\boldsymbol{\theta} = \boldsymbol{\theta}(\mathscr{B}(\mathscr{V}))$. Regular vine copulas which differ in the tree structure or in at least one pair copula family represent in general different statistical models. Notable exceptions from this are the multivariate Gaussian, Student's t or Clayton copulas, which can be decomposed into bivariate Gaussian, Student's t or Clayton copulas, respectively, in multiple ways (see [64]). The probability density function $f_{1:n}$ at point $\mathbf{x} = (x_1, \dots, x_n)' \in \mathbb{R}^n$ of an n-dimensional regular vine-dependent distribution $F_{1:n}$ is easily calculated as

$$f_{1:n}\left(\mathbf{x}|\mathscr{V},\mathscr{B},\boldsymbol{\theta}\right) = \left(\prod_{m=1}^{n-1}\prod_{e=\{a,b\}\in E_m} c_{i(e),j(e)|D(e)}\left(F_{i(e)|D(e)}, F_{j(e)|D(e)}|\boldsymbol{\theta}_{i(e),j(e)|D(e)}\right)\right) \times f_1\left(x_1\right)\cdots f_n\left(x_n\right), \tag{2.2}$$

where $F_{i(e)|D(e)} := F_{i(e)|D(e)}\left(x_{i(e)}|\mathbf{x}_{D(e)}\right)$ and $F_{j(e)|D(e)} := F_{j(e)|D(e)}\left(x_{j(e)}|\mathbf{x}_{D(e)}\right)$ (see [8]). These conditional distribution functions can be determined recursively tree-by-tree using the following relationship

$$\begin{aligned} F_{i(e)|D(e)}\left(x_{i(e)}|\mathbf{x}_{D(e)}\right) &= F_{\mathscr{C}_{e,a}|D_e}\left(x_{\mathscr{C}_{e,a}}|\mathbf{x}_{D_e}\right) \\ &= \frac{\partial C_{\mathscr{C}_{a,a_1},\mathscr{C}_{a,a_2}|D_a}\left(F_{\mathscr{C}_{a,a_1}|D_a}\left(x_{\mathscr{C}_{a,a_1}}|\mathbf{x}_{D_a}\right), F_{\mathscr{C}_{a,a_2}|D_a}\left(x_{\mathscr{C}_{a,a_2}}|\mathbf{x}_{D_a}\right)\right)}{\partial F_{\mathscr{C}_{a,a_2}|D_a}\left(x_{\mathscr{C}_{a,a_2}}|\mathbf{x}_{D_a}\right)}, \end{aligned} \tag{2.3}$$

where $e = \{a, b\}$ with $a = \{a_1, a_2\}$ as before (see [23] for details).

To facilitate inference, it is assumed that the pair copulas $C_{i(e),j(e)|D(e)}$ only depend on the variables with indices in $D(e)$ through the arguments $F_{i(e)|D(e)}$ and $F_{j(e)|D(e)}$. This so-called simplifying assumption has been investigated by [35, 64]. In particular, [35] show examples where the assumption is not severe and [64] show that the multivariate Clayton copula is the only Archimedean copula which can be represented as a simplified vine. A critical look on the subject can be found in [2] who take a first step in building vine copulas of non-simplified nature.

Following expression (2.2), the likelihood L of a regular vine copula $C = (\mathscr{V}, \mathscr{B}, \boldsymbol{\theta})$ given the observed data $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)' \in \mathbb{R}^{N \times n}$ can be calculated as

$$L(\mathscr{V}, \mathscr{B}, \boldsymbol{\theta} | \mathbf{x}) = \prod_{k=1}^{N} f_{1:n}(\mathbf{x}_k | \mathscr{V}, \mathscr{B}, \boldsymbol{\theta}). \tag{2.4}$$

The corresponding log-likelihood is denoted by ℓ.

A regular vine copula is said to be *truncated at level M* if all pair copulas conditioning on M or more variables are set to bivariate independence copulas [14]. As a result, the iteration index m of the first product in (2.2) runs only up to $m = M$ and the distribution is fully specified by $\big(\mathscr{V} = (T_1, \ldots, T_M), \mathscr{B}(\mathscr{V}), \boldsymbol{\theta}(\mathscr{B}(\mathscr{V}))\big)$.

In the following, we discuss in reverse order how the components of a regular vine copula can be selected and estimated. That is, we begin with estimation of the parameters $\boldsymbol{\theta}$, then treat the selection of appropriate copula families $\mathscr{B}$, and finally discuss the selection of vine trees $\mathscr{V}$.

2.3 Parameter Estimation for Given Vine Tree Structure and Pair Copula Families

Given a vine tree structure $\mathscr{V}$ and pair copula families $\mathscr{B} = \mathscr{B}(\mathscr{V})$, the challenge is to estimate the parameters $\boldsymbol{\theta} = \boldsymbol{\theta}(\mathscr{B}(\mathscr{V}))$ of a regular vine copula for observed data $\mathbf{x} \in \mathbb{R}^{N \times n}$. The crucial point here is to evaluate the conditional distribution functions $F_{i(e)|D(e)}$, which depend on the copulas of previous trees [see (2.3)].

2.3.1 Maximum Likelihood Estimation

Classically, parameters of a statistical model are often estimated using maximum likelihood techniques. Here, this means that regular vine copula parameters $\boldsymbol{\theta}$ and parameters of the marginal distributions are estimated by maximizing the likelihood (2.4) in terms of these parameters. For copulas, in particular if $n > 2$, the number of parameters to be estimated may, however, be too large, so that one typically either uses empirical distribution functions for the margins as proposed by [25, 26] or estimates parameters of the marginal distributions in a first step and then fixes these marginal parameters to their estimated values in the estimation of the copula parameters. The latter method is called inference functions for margins (IFM) by [38, 39].

But even when estimation of marginal and dependence parameters is separated, joint maximum likelihood estimation of regular vine copula parameters can be computationally intensive, since the vine decomposition involves $n(n-1)/2$ bivariate copulas with corresponding parameters. Aas et al. [1] therefore proposed a sequential method: starting with the copulas of the first tree, this method proceeds tree-wise and estimates the parameters of the copulas in a tree by fixing the parameters of copulas in all previous trees.

Example 2.1 (Sequential Estimation). Let the 5-dimensional regular vine $\mathscr{V}$ of Fig. 2.1 be given with copulas $\mathscr{B}(\mathscr{V})$. In the first step, we estimate the parameters of the copulas $C_{1,2}$, $C_{2,3}$, $C_{3,4}$, and $C_{3,5}$ using maximum likelihood based on the transformed observations $F_j(x_{kj})$ for x_{kj}, $k = 1,\dots,N$, $j = 1,\dots,5$. In the second tree, we then have to estimate, for instance, the parameter(s) of the copula $C_{1,3|2}$. For this we form pseudo-observations $F_{1|2}(x_{k1}|x_{k2}, \hat{\boldsymbol{\theta}}_{1,2})$ and $F_{3|2}(x_{k3}|x_{k2}, \hat{\boldsymbol{\theta}}_{2,3})$, $k = 1,\dots,N$, according to expression (2.3) and using the estimated parameters of copulas $C_{1,2}$ and $C_{2,3}$, respectively. Based on these pseudo-observations estimation of $\boldsymbol{\theta}_{1,3|2}$ is again straightforward. All other copulas are estimated analogously.

Note that this strategy only involves the estimation of bivariate copulas and therefore is computationally much simpler than joint maximum likelihood estimation of all parameters at once. For more details, see [1] as well as [33] who investigates the asymptotic properties of this sequential approach. A comparison study of estimators for regular vine copulas can be found in [34].

If joint maximum likelihood estimates are desired, the sequential method can be used to obtain starting values for the numerical optimization. A detailed discussion how to compute score functions and the observed information matrix for regular vine copulas is provided by [65].

2.3.2 Bayesian Posterior Estimation

Bayesian statistics considers parameters $\boldsymbol{\theta}$ as random variables. As such, inference focuses on estimating the entire distribution of the parameters instead of only finding a point estimate. In particular, the so-called posterior distribution $p(\boldsymbol{\theta}) := p(\boldsymbol{\theta}|\mathbf{x})$, the distribution of the parameters $\boldsymbol{\theta}$ given the observed data $\mathbf{x}$, is the main object of interest. The unnormalized posterior density factorizes into the product of the likelihood function $L(\boldsymbol{\theta}|\mathbf{x})$ (2.4) and the prior density function $\pi(\boldsymbol{\theta})$:

$$p(\boldsymbol{\theta}) := p(\boldsymbol{\theta}|\mathbf{x}) = \frac{L(\boldsymbol{\theta}|\mathbf{x}) \cdot \pi(\boldsymbol{\theta})}{f(\mathbf{x})} \propto L(\boldsymbol{\theta}|\mathbf{x}) \cdot \pi(\boldsymbol{\theta}).$$

The prior distribution incorporates a priori beliefs about the distribution of the parameters. It must not depend on, or be conditional upon, the observed data.

Markov chain Monte Carlo (MCMC) procedures clear the remaining obstacle of sampling from a distribution which is known only up to a constant. The trick is to simulate the run of a Markov chain whose equilibrium distribution is the targeted posterior distribution $p(\boldsymbol{\theta})$ of the parameters $\boldsymbol{\theta}$. Upon convergence of the Markov chain, its states represent draws from the desired distribution.

The Metropolis–Hastings algorithm [31, 49] implements the simulation of such a Markov chain through an acceptance/rejection mechanism for the updates of the chain: in the first step, an update of the chain from its current state $\boldsymbol{\theta}$ to $\boldsymbol{\theta}^* \sim q(\cdot|\boldsymbol{\theta})$ is proposed. The proposal distribution q can be chosen almost arbitrarily. In the second step, the proposal is accepted with probability $\alpha := \alpha(\boldsymbol{\theta}, \boldsymbol{\theta}^*)$. The acceptance probability α is chosen such that convergence of the Markov chain to the targeted distribution is guaranteed by Markov chain theory [50]. The Metropolis–Hastings acceptance probability for convergence to the target distribution $p(\cdot)$ is

$$\alpha(\boldsymbol{\theta}, \boldsymbol{\theta}^*) = \frac{p(\boldsymbol{\theta}^*)}{p(\boldsymbol{\theta})} \cdot \frac{q(\boldsymbol{\theta}|\boldsymbol{\theta}^*)}{q(\boldsymbol{\theta}^*|\boldsymbol{\theta})}.$$

The target distribution goes into the Metropolis–Hastings algorithm only at the evaluation of the acceptance probability α. It can be easily seen that it is sufficient to know the density function of the target distribution only up to a constant for this algorithm to work, given that α only depends on the ratio of two densities.

The following Metropolis–Hastings scheme to sample from the posterior distribution of the parameters $\boldsymbol{\theta}$ of a regular vine copula is proposed in [28, 51, 62]. The authors suggest normally distributed random walk proposals be used in the update step.

Algorithm 2.1 (Metropolis–Hastings Sampler for Parameter Estimation).

```
Choose arbitrary starting values $\boldsymbol{\theta}^0 = (\theta_1^0, \ldots, \theta_D^0)'$, where $D := n(n-1)/2$.
for each iteration $r = 1, \ldots, R$ do
    Set $\boldsymbol{\theta}^r = \boldsymbol{\theta}^{r-1}$.
    for $i = 1, \ldots, D$ do
        Draw $\theta_i^r$ from a $N(\theta_i^{r-1}, \sigma_i^2)$ distribution with probability density function
        $\phi_{(\theta_i^{r-1}, \sigma_i^2)}(\cdot)$.
        Evaluate the Metropolis–Hastings acceptance probability of the proposal

        $\alpha := \alpha(\theta_i^{r-1}, \theta_i^r) = \frac{p(\boldsymbol{\theta}^*)}{p(\boldsymbol{\theta})} \cdot \frac{\phi_{(\theta_i^r, \sigma_i^2)}(\theta_i^{r-1})}{\phi_{(\theta_i^{r-1}, \sigma_i^2)}(\theta_i^r)} = \frac{L(\boldsymbol{\theta}^*|\mathbf{x})}{L(\boldsymbol{\theta}|\mathbf{x})} \cdot \frac{\pi(\boldsymbol{\theta}^*)}{\pi(\boldsymbol{\theta})}$,

        where $\boldsymbol{\theta}^* = (\theta_1^r, \ldots, \theta_i^r, \theta_{i+1}^{r-1}, \ldots, \theta_D^{r-1})'$
        and $\boldsymbol{\theta} = (\theta_1^r, \ldots, \theta_{i-1}^r, \theta_i^{r-1}, \ldots, \theta_D^{r-1})'$.
        Accept the proposal $\theta_i^r$ with probability $\alpha$; if rejected, set $\theta_i^r = \theta_i^{r-1}$.
    end for
end for
return $(\boldsymbol{\theta}^1, \ldots, \boldsymbol{\theta}^R)$
```

2.4 Selection of the Pair Copula Families and their Parameters for a Known Vine Tree Structure

An n-dimensional regular vine copula with tree structure $\mathscr{V}$ is based on a set $\mathscr{B}(\mathscr{V})$ of $n(n-1)/2$ bivariate copulas and their corresponding parameters. The copula families can be chosen arbitrarily, e.g., from the popular classes of Archimedean, elliptical or extreme-value copulas. Assuming that an appropriate vine structure $\mathscr{V}$ is chosen, the question is how to select adequate (conditional) pair copulas $C_{i(e),j(e)|D(e)}$ and their parameters for given data $\mathbf{x}$.

2.4.1 *Sequential Selection*

If a regular vine is truncated at level 1, the copulas in trees T_2 to T_{n-1} are set to independence. The corresponding regular vine copula density reduces to the product of the unconditional bivariate copula densities $C_{i(e),j(e)}$, which can be selected based on data $F_{i(e)}(x_{k,i(e)})$ and $F_{j(e)}(x_{k,j(e)})$, $k = 1, \ldots, N$. Typical criteria for copula selection from a given set of families are information criteria such as the AIC as proposed by [11, 47]. The latter compares AIC-based selection to three alternative selection strategies: selection of family with highest p-value of a copula goodness of fit test based on the Cramér–von Mises statistic, with smallest distance between empirical and modeled dependence characteristics (Kendall's τ, tail dependence), or with highest number of wins in pairwise comparisons of families using the test by [66]. In a large-scale Monte Carlo study the AIC turned out to be the most reliable selection criterion.

In a general regular vine, the selection of a pair copula $C_{i(e),j(e)|D(e)}$, however, depends on the choices for copulas in previous trees due to its arguments (2.3). Since a joint selection seems infeasible because of the many possibilities, one typically proceeds tree-by-tree as in the sequential estimation method. That is, instead of estimating the parameters $\boldsymbol{\theta}_{i(e),j(e)|D(e)}$ of $C_{i(e),j(e)|D(e)}$, the copula is selected first and then estimated, which usually coincides for most selection strategies. Given the selected and estimated copulas of previous trees, one then selects the copulas of the next tree.

Example 2.2 (Sequential Selection). Again consider the five-dimensional regular vine $\mathscr{V}$ of Fig. 2.1 but now with unknown copulas $\mathscr{B}(\mathscr{V})$. In the first tree, we select (and then estimate) the copulas $C_{1,2}$, $C_{2,3}$, $C_{3,4}$, and $C_{3,5}$ using our method of choice based on $F_j(x_{kj})$, $k = 1, \ldots, N$, $j = 1, \ldots, 5$. Given these copulas, we then have to select conditional copulas in the second tree. In case of the copula

$C_{1,3|2}$, we therefore again form the pseudo-observations $F_{1|2}(x_{k1}|x_{k2}, \hat{\boldsymbol{\theta}}_{1,2})$ and $F_{3|2}(x_{k3}|x_{k2}, \hat{\boldsymbol{\theta}}_{2,3})$, $k = 1, \ldots, N$, according to expression (2.3) and then select (and estimate) $C_{1,3|2}$ based on them. This can be iterated for all trees.

Clearly this sequential selection strategy accumulates uncertainty in the selection and therefore the final model has to be carefully checked and compared to alternative models. For the latter, the tests for non-nested model comparison by [18, 66] may be used.

2.4.2 Reversible Jump MCMC-Based Bayesian Selection

Bayesian copula family and parameter selection aims at estimating the joint posterior distribution of the pair copula families $\mathscr{B} = \mathscr{B}(\mathscr{V})$ and parameters $\boldsymbol{\theta} = \boldsymbol{\theta}(\mathscr{B}(\mathscr{V}))$. The posterior density function factorizes into the product of the likelihood function L (2.4) and the prior density function π: $p(\mathscr{B}, \boldsymbol{\theta}) \propto L(\mathscr{B}, \boldsymbol{\theta}|\mathbf{x}) \cdot \pi(\mathscr{B}, \boldsymbol{\theta})$.

As in Sect. 2.4.1, family selection is understood in the context of choosing from a pre-specified set $\mathbf{B}$ of parametric pair copula families. Model sparsity can be induced through the choice of prior distributions which favor regular vine copulas with fewer parameters over those with more parameters.

Bayesian techniques to select the pair copula families of D-vine copulas are covered in [51, 52, 61]. The reversible jump MCMC algorithm presented in this section follows the latest developments of family selection methods for regular vine copulas laid out in [28].

Reversible jump MCMC is an extension of ordinary MCMC to sample from discrete-continuous posterior distributions. The sampling algorithm and the mathematics underpinning the convergence statements are an immediate generalization of the Metropolis–Hastings algorithm [27]. A reversible jump MCMC algorithm functions like an ordinary MCMC algorithm with one extra step built in: before, or after, updating the parameters, a “jump” to another model is attempted. That is called the “between-models move,” while the updating of the parameters only is called the “within-model move” [16].

Algorithm 2.2 implements a reversible jump MCMC sampler which samples from the joint posterior distribution of the pair copula families $\mathscr{B}$ and the parameters $\boldsymbol{\theta}$, given a regular vine tree structure $\mathscr{V}$. In each iteration, it first updates all parameters as in Algorithm 2.1. Then the pair copula families are updated along with their parameters one-by-one. The pair copula family proposals are drawn from a uniform distribution over $\mathbf{B}$, while the parameter proposals are sampled from a (multivariate) normal distribution centered at the maximum likelihood estimates of the parameters.

Algorithm 2.2 (Reversible Jump MCMC Sampler for Family Selection).

```
Choose arbitrary copula family and parameter starting values $(\mathscr{B}^0, \boldsymbol{\theta}^0) = ((\mathscr{B}_1^0, \boldsymbol{\theta}_1^0), \dots, (\mathscr{B}_D^0, \boldsymbol{\theta}_D^0))$, where $D := n(n-1)/2$.
for each iteration $r = 1, \dots, R$ do
    Set $(\mathscr{B}^r, \boldsymbol{\theta}^r) = (\mathscr{B}^{r-1}, \boldsymbol{\theta}^{r-1})$.
    Perform one update step of Algorithm 2.1 for the parameters $\boldsymbol{\theta}^r$; denote the updated parameter entries by $\boldsymbol{\theta}^+$.
    for $i = 1, \dots, D$ do
        Draw $\mathscr{B}_i^r$ from a Unif$(\mathbf{B})$ distribution.
        Draw $\boldsymbol{\theta}_i^r$ from a multivariate $N(\hat{\boldsymbol{\theta}}_i^r, \hat{\Sigma}_i^r)$ distribution with probability density function $\phi_{(\hat{\boldsymbol{\theta}}_i^r, \hat{\Sigma}_i^r)}(\cdot)$. Here $\hat{\boldsymbol{\theta}}_i^r$ denotes the MLE of the copula parameters of the pair copula $\mathscr{B}_i^r$ and $\hat{\Sigma}_i^r$ denotes the estimated approximative covariance matrix of the parameter estimates $\hat{\boldsymbol{\theta}}_i^r$.
        Evaluate the generalized Metropolis–Hastings acceptance probability of the proposal
            $\alpha := \alpha((\mathscr{B}_i^{r-1}, \boldsymbol{\theta}_i^+), (\mathscr{B}_i^r, \boldsymbol{\theta}_i^r))$
            $\quad = \frac{p(\mathscr{B}^*, \boldsymbol{\theta}^*)}{p(\mathscr{B}, \boldsymbol{\theta})} \cdot \frac{\phi_{(\hat{\boldsymbol{\theta}}_i^{r-1}, \hat{\Sigma}_i^{r-1})}(\boldsymbol{\theta}_i^+)}{\phi_{(\hat{\boldsymbol{\theta}}_i^r, \hat{\Sigma}_i^r)}(\boldsymbol{\theta}_i^r)}$
            $\quad = \frac{\pi(\mathscr{B}^*, \boldsymbol{\theta}^*)}{\pi(\mathscr{B}, \boldsymbol{\theta})} \cdot \frac{L(\mathscr{B}^*, \boldsymbol{\theta}^* | \mathbf{x})}{L(\mathscr{B}, \boldsymbol{\theta} | \mathbf{x})} \cdot \frac{\phi_{(\hat{\boldsymbol{\theta}}_i^{r-1}, \hat{\Sigma}_i^{r-1})}(\boldsymbol{\theta}_i^+)}{\phi_{(\hat{\boldsymbol{\theta}}_i^r, \hat{\Sigma}_i^r)}(\boldsymbol{\theta}_i^r)},$
        where $\mathscr{B}^* = (\mathscr{B}_1^r, \dots, \mathscr{B}_i^r, \mathscr{B}_{i+1}^{r-1}, \dots, \mathscr{B}_D^{r-1})$,
        $\mathscr{B} = (\mathscr{B}_1^r, \dots, \mathscr{B}_{i-1}^r, \mathscr{B}_i^{r-1}, \dots, \mathscr{B}_D^{r-1})$, $\boldsymbol{\theta}^* = (\boldsymbol{\theta}_1^r, \dots, \boldsymbol{\theta}_i^r, \boldsymbol{\theta}_{i+1}^+, \dots, \boldsymbol{\theta}_D^+)$,
        and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^r, \dots, \boldsymbol{\theta}_{i-1}^r, \boldsymbol{\theta}_i^+, \dots, \boldsymbol{\theta}_D^+)$.
        Accept the proposal $(\mathscr{B}_i^r, \theta_i^r)$ with probability $\alpha$; if rejected, set $(\mathscr{B}_i^r, \theta_i^r) = (\mathscr{B}_i^{r-1}, \theta_i^+)$.
    end for
end for
return $((\mathscr{B}^1, \boldsymbol{\theta}^1), \dots, (\mathscr{B}^R, \boldsymbol{\theta}^R))$
```

2.5 Selection of Vine Tree Structure, Pair Copula Families, and Parameters

As pair copula families $\mathscr{B} = \mathscr{B}(\mathscr{V})$ and parameters $\boldsymbol{\theta} = \boldsymbol{\theta}(\mathscr{B}(\mathscr{V}))$ both depend on the vine tree structure $\mathscr{V}$, the identification of adequate trees is crucial to the model selection of vine copulas. As it was already the case for pair copula selection in Sect. 2.4.1, it is again not feasible to simply try and fit all possible regular vine copulas $C = (\mathscr{V}, \mathscr{B}, \boldsymbol{\theta})$ and then choose the "best" one. In particular, the number of possible regular vines on n variables is $\frac{n!}{2} \times 2^{\binom{n-2}{2}}$ as shown by [53]. This means that even if pair copulas and parameters were known, the number of different models would still be enormous.

This remains true even when the selection is restricted to the sub-classes of C- and D-vines, since there are still $n!/2$ different C- and D-vines in n dimensions, respectively (see [1]). It should, however, be noted that C- and D-vine copulas are most appropriate if their structure is explicitly motivated by the data. In particular, C-vine copulas may be used if there is a set of pivotal variables such as stock indices (see [12, 32]) and D-vine copulas are particularly attractive to model variables with temporal order (see [13, 61]). Nevertheless we describe how C- and D-vines can be selected for arbitrary data sets.

2.5.1 Top-Down and Bottom-Up Selection

Due to the proximity condition (see Definition 2.2) regular vine trees are closely linked with each other and have to be constructed carefully. Two construction strategies have been proposed in the literature: a top-down approach by [23] and a bottom-up method by [43]. Both strategies proceed sequentially tree-by-tree and respect the proximity condition in each step. We first describe the top-down, then the bottom-up method.

2.5.1.1 Top-Down Selection

Selecting regular vine trees top-down means that one starts with the selection of the first tree T_1 and continues tree-by-tree up to the last tree T_{n-1}. The first tree T_1 can be selected as an arbitrary spanning tree. Given that a tree T_m, $m \in \{1, \dots, n-2\}$, has been selected, the next tree T_{m+1} is chosen respecting the proximity condition (see Definition 2.2). In other words, T_{m+1} can only be formed by (conditional) pairs $\{i(e), j(e)|D(e)\}$ which fulfill the proximity condition.

Example 2.3 (Top-Down Tree Selection). Assume that we have selected the first tree T_1 as shown in Fig. 2.1. Then the question is which pairs $\{i(e), j(e)|D(e)\}$ are eligible for tree construction in the second tree T_2. According to the proximity condition these are $\{1, 3|2\}$, $\{2, 4|3\}$, $\{2, 5|3\}$, and $\{4, 5|3\}$. Obviously, the last three pairs form a cycle and therefore only two of them can be selected for T_2. One of the three possibilities is shown in Fig. 2.1.

To perform this iterative selection strategy a criterion is needed to select a spanning tree among the set of eligible edges, where a spanning tree simply denotes a tree on all nodes. Clearly, the log-likelihood ℓ_m of the pair copulas in tree T_m of a regular vine copula [see expression (2.4)] is given by

$$\ell_m\left(T_m, \mathscr{B}_{T_m}, \boldsymbol{\theta}_{T_m}|\mathbf{x}\right) = \sum_{k=1}^{N}\sum_{e\in E_m} \log\big(c_{i(e),j(e)|D(e)} \\ \times \left(F_{i(e)|D(e)}, F_{j(e)|D(e)}|\boldsymbol{\theta}_{i(e),j(e)|D(e)}\right)\big), \quad (2.5)$$

where we write $\mathscr{B}_{T_m} := \mathscr{B}(T_m)$ and $\boldsymbol{\theta}_{T_m} := \boldsymbol{\theta}(\mathscr{B}(T_m))$.

A straightforward solution therefore would be to choose the tree such that (2.5) is maximized after having selected pair copulas with high log-likelihood for each (conditional) pair $\{i(e), j(e)|D(e)\}$ that fulfills the proximity condition. This solution, however, leads to highly over-parameterized models, since models with more parameters in which simpler models are nested will always give a higher likelihood. For instance, the Student's t copula always has a higher likelihood than the Gaussian copula which is a special case of the Student's t as the degrees of freedom tend to infinity.

Therefore we formulate the following algorithm in terms of a general weight ω assuming that we want to maximize it for each tree. The previously discussed strategy corresponds to choosing the pair copula log-likelihoods as weights.

Algorithm 2.3 (Sequential Top-Down Selection Based on Weights).

```
Calculate the weight $\omega_{i,j}$ for all possible variable pairs $\{i, j\}, 1 \le i < j \le n$.
Select the maximum spanning tree, i.e.

    $T_1 = \underset{T=(N,E)\text{ spanning tree}}{\operatorname{argmax}} \sum_{e\in E} \omega_{i(e),j(e)}.$

for each edge $e \in E_1$ do
    Select a copula $C_{i(e),j(e)}$.
    Estimate the corresponding parameter(s) $\boldsymbol{\theta}_{i(e),j(e)}$.
    For $k = 1, \ldots, N$ transform $F_{i(e)|j(e)}(x_{k,i(e)}|x_{k,j(e)}, \hat{\boldsymbol{\theta}}_{i(e),j(e)})$ and
    $F_{j(e)|i(e)}(x_{k,j(e)}|x_{k,i(e)}, \hat{\boldsymbol{\theta}}_{i(e),j(e)})$ using (2.3).
end for
for $m = 2, \ldots, n-1$ do
    Calculate the weights $\omega_{i(e),j(e)|D(e)}$ for all conditional variable pairs
    $\{i(e), j(e)|D(e)\}$ that can be part of tree $T_m$. We denote this set of edges
    which fulfill the proximity condition by $E_P$.
    Among these edges, select the maximum spanning tree, i.e.,

        $T_m = \underset{T=(N,E)\text{spanning tree with } E\subset E_P}{\operatorname{argmax}} \sum_{e\in E} \omega_{i(e),j(e)|D(e)}.$

    for each edge $e \in E_m$ do
        Select a conditional copula $C_{i(e),j(e)|D(e)}$.
        Estimate the corresponding parameter(s) $\boldsymbol{\theta}_{i(e),j(e)|D(e)}$.
```

```
14:         For $k = 1,\dots,N$ transform $F_{i(e)|j(e)\cup D(e)}(x_{k,i(e)}|x_{k,j(e)},\mathbf{x}_{k,D(e)},\hat{\boldsymbol{\theta}}_{i(e),j(e)|D(e)})$ and $F_{j(e)|i(e)\cup D(e)}(x_{k,j(e)}|x_{k,i(e)},\mathbf{x}_{k,D(e)},\hat{\boldsymbol{\theta}}_{i(e),j(e)|D(e)})$ using (2.3).
15:     end for
16: end for
17: return the sequential model estimate $(\hat{\mathscr{V}},\hat{\mathscr{B}},\hat{\boldsymbol{\theta}})$.
```

Clearly, this algorithm only makes a locally optimal selection in each step, since the impact on previous and subsequent trees is ignored. The strategy is, however, reasonable in light of the definition of regular vines (see Definition 2.2). The maximum spanning tree in lines 2 and 10 can be found, e.g., using the classical algorithms by Prim or Kruskal (see [19, Section 23.2]). Possible choices for the weight ω are, for example:

- The absolute empirical Kendall's τ as proposed by [22, 23].
- The AIC of each pair copula corresponding to the discussion in Sect. 2.4.1.
- The (negative) estimated degrees of freedom of Student's t pair copulas (see [48]).
- The p-value of a copula goodness of fit test and variants as proposed by [21].

Some remarks: First, taking the empirical Kendall's τ as weight does not require to select and estimate pair copulas prior to the tree selection step. The other three weights require this, so that lines 4–5 and 12–13 in Algorithm 2.3 may be redundant in this case. Second, AIC weights also maximize the AIC of the entire tree, since the individual AICs sum like the log-likelihood. Third, the strategy proposed by [48] concentrates on tail dependence, since the algorithm will select pairs with estimated small degrees of freedom corresponding to a stronger deviation from Gaussianity with no tail dependence.

Copula goodness of fit tests based on the Cramér–von Mises statistic are considered in [21] to take into account the uncertainty of the pair copula fit. Corresponding p-values are calculated using fast bootstrap methods based on the multiplier approach developed and implemented in [40–42].

Finally, we like to note that the empirical Kendall's τ as weight also approximatively maximizes (2.5), since the log-likelihoods of bivariate copulas tend to increase with increasing absolute values of Kendall's τ. It, however, does not lead to overparameterization, since copula selection and tree selection in a particular tree are independent. The strategy by [23] has therefore already been used successfully in applications with up to 52 variables (see [12]).

Recently, [36] proposed the construction of vines with non-parametric pair copulas based on empirical copula estimators. For vine tree selection [36] also advocate the use of empirical dependence measures such as Spearman's ρ, since it does not require any parametric assumption on the pair copulas. Other dependence measures such as Blomqvist's β could of course be used instead in the approaches by [23, 36].

Especially the strategy based on Kendall's τ leads to regular vine copulas with decreasing dependence in higher order trees. This is advantageous to the vine structure, since numerical errors may add up and estimates become less precise as the number of conditioning variables increases. This has been exploited by [14] to identify appropriate truncation levels of regular vine copulas.

Algorithm 2.3 can easily be modified to select C- or D-vines instead of general regular vines. For C-vines the root node in each tree can simply be identified as the node with maximal sum of weights to all other nodes (see [22]). In the case of D-vines only the order of variables in the first tree has to be chosen. Since D-vine trees are paths on all nodes, so-called Hamiltonian paths, a maximum Hamiltonian path has to be identified. This problem is equivalent to a traveling salesman problem as discussed by [11]. As an NP-hard problem, there is no known efficient algorithm to find a solution.

2.5.1.2 Bottom-Up Selection

Rather than beginning with the first tree T_1, [43] proposed a bottom-up selection strategy which starts with tree T_{n-1} and then sequentially selects trees T_m for $m = n-2, \ldots, 1$. Similar to top-down selection using Kendall's τ by [23], this approach is also motivated by choosing a regular vine with weaker dependence in later trees. Here, dependence is, however, measured in terms of partial correlations. More precisely, each tree is selected such that the sum of absolute partial correlations as edge weights is minimized. This is feasible, since partial correlation estimates for each combination of variables can be obtained from the data without any parametric assumptions (see, e.g., [67]).

The bottom-up strategy by [43] thus selects tree T_{n-1} as the edge corresponding to the pair of variable with lowest absolute partial correlation given all other variables. Kurowicka [43] then provides conditions such that the proximity condition is satisfied when a tree T_m, $m \in \{1, \ldots, n-2\}$, is selected given trees $T_{m+1}, \ldots, T_{n-1}$. As before there may be several choices possible in the set of eligible edges.

Example 2.4 (Bottom-Up Tree Selection). Assume in Fig. 2.1 that we have selected T_4 with edge $e = \{a, b\} = \{1, 5|2, 3, 4\}$ as shown. Then $A_a = \{1, 2, 3, 4\}$ and $A_b = \{2, 3, 4, 5\}$ [see (2.1)]. Any choice of edges from A_a and A_b in tree T_3 leads to a valid tree T_4. For instance, the edges $\{1, 2|3, 4\}$ and $\{4, 5|2, 3\}$ are compatible with A_a and A_b, respectively, and would lead to another vine than the one shown in Fig. 2.1. In fact, it could be constructed as a D-vine with order $1, 4, 3, 2, 5$.

Since partial correlations are equal to conditional correlations for elliptical distributions (see [5]), this strategy is particularly appropriate if all pair copula families are elliptical. In contrast to the top-down approach discussed above which requires the selection and estimation of pair copulas in each tree, this strategy chooses the vine tree structure without any assumption on the pair copula families. Having selected the vine tree structure, the selection of pair copulas can therefore proceed as discussed in Sect. 2.4.1.

2.5.2 Sequential Bayesian Tree Selection

Bayesian approaches to estimating the posterior distribution of the tree structure of a regular vine copula have only recently been developed. The sheer number of possible regular vine tree structures poses a relevant challenge to computationally efficient posterior evaluation.

This section presents a reversible jump MCMC-based approach proposed by [29] to obtain a sequential estimate of the posterior distribution of the regular vine tree structure $\mathscr{V}$, the pair copula families $\mathscr{B}$, and the copula parameters $\boldsymbol{\theta}$. The term "sequential estimation" is understood analogously to the notion discussed in Sect. 2.5.1.1 as a tree-by-tree procedure.

As in Sect. 2.4.2, model priors $\pi(\mathscr{V}, \mathscr{B}, \boldsymbol{\theta})$ which favor sparse models can serve to guard against selecting models with runaway complexity. On the other hand, the use of non-informative flat priors allows for tree-by-tree maximum likelihood estimation of the regular vine tree structure $\mathscr{V}$, the pair copula families $\mathscr{B}$, and the copula parameters $\boldsymbol{\theta}$. As a result, the posterior mode estimate of this procedure will agree with the model estimate of Algorithm 2.3, if flat priors are used. Again, the pair copula families are chosen from a set **B** of parametric pair copula families.

For the sake of notational convenience and enhanced readability, the procedure is presented in two algorithms. Algorithm 2.4 gives the general outline of the sampling procedure. Algorithm 2.5 details the reversible jump MCMC algorithm to sample from the posterior distribution of tree T_m of the regular vine tree structure $\mathscr{V} = (T_1, \dots, T_{n-1})$.

Algorithm 2.4 implements the tree-by-tree selection procedure. As such, it calls Algorithm 2.5 for the actual posterior estimation. However, it condenses the posterior sample into the posterior mode estimate and organizes the move from selecting tree T_m to selecting tree T_{m+1}. The posterior mode estimate $(\hat{T}_m, \hat{\mathscr{B}}_{T_m}, \hat{\boldsymbol{\theta}}_{T_m})$ of tree T_m is the most frequently sampled combination of tree structure T_m and pair copula families $\mathscr{B}_{T_m}$ parameterized at the mode of the posterior sample of the parameters $\boldsymbol{\theta}_{T_m}$ of this model.

Algorithm 2.4 (Outline of the Tree-by-Tree Sampling Algorithm).

```
1: Sample from the posterior distribution of the first tree, $(T_1, \mathscr{B}_{T_1}, \boldsymbol{\theta}_{T_1})$: see Algorithm 2.5.
2: Set the tree estimate $(\hat{T}_1, \hat{\mathscr{B}}_{T_1}, \hat{\boldsymbol{\theta}}_{T_1})$ to the posterior mode.
3: for $m = 2, \dots, n-1$ do
4:     Sample from the posterior distribution of the m-th tree, $(T_m, \mathscr{B}_{T_m}, \boldsymbol{\theta}_{T_m})$, given the previous trees' estimates $(\hat{T}_1, \hat{\mathscr{B}}_{T_1}, \hat{\boldsymbol{\theta}}_{T_1}), \dots, (\hat{T}_{m-1}, \hat{\mathscr{B}}_{T_{m-1}}, \hat{\boldsymbol{\theta}}_{T_{m-1}})$: see Algorithm 2.5.
5:     Set the tree estimate $(\hat{T}_m, \hat{\mathscr{B}}_{T_m}, \hat{\boldsymbol{\theta}}_{T_m})$ to the posterior mode.
6: end for
7: return the sequential Bayesian model estimate
       $(\hat{\mathscr{V}}, \hat{\mathscr{B}}, \hat{\boldsymbol{\theta}}) = ((\hat{T}_1, \dots, \hat{T}_{n-1}), (\hat{\mathscr{B}}_{T_1}, \dots, \hat{\mathscr{B}}_{T_{n-1}}), (\hat{\boldsymbol{\theta}}_{T_1}, \dots, \hat{\boldsymbol{\theta}}_{T_{n-1}}))$.
```

Algorithm 2.5 samples from the posterior distribution of $(T_m, \mathscr{B}_{T_m}, \boldsymbol{\theta}_{T_m})$ given the previously selected trees $((\hat{T}_l, \hat{\mathscr{B}}_{T_l}, \hat{\boldsymbol{\theta}}_{T_l}), l = 1, \dots, m-1)$. The within-model move of this algorithm to update the pair copula parameters $\boldsymbol{\theta}$ follows Algorithm 2.1. In the between-models move, the proposal trees are sampled as spanning trees which satisfy the proximity condition. Practitioners may note that this functionality is implemented, e.g., in the C++ `boost` library [59]. Edge weights can be used to fine-tune the variance of the proposal distribution: in our implementation, higher values for $p \in (0,1)$ increase the probability that the proposals for the tree structure, T_m^r, are similar to the current state T_m^{r-1} of the sampling chain. As in Algorithm 2.2, the pair copula family proposals are sampled from a uniform distribution over $\mathbf{B}$, and the parameter proposals are drawn from a (multivariate) normal distribution centered at the maximum likelihood estimates of the parameters.

Algorithm 2.5 (Reversible Jump MCMC Algorithm to Sample $(T_m, \mathscr{B}_{T_m}, \boldsymbol{\theta}_{T_m})$).

Denote the pair copula families of tree $T_m = (N_m, E_m)$ by $\mathscr{B}_{T_m} = (\mathscr{B}_e, e \in E_m)$, and the corresponding pair copula parameters by $\boldsymbol{\theta}_{T_m} = (\boldsymbol{\theta}_e, e \in E_m)$.

```
Choose arbitrary, but valid, tree structure, copula family and parameter starting values $(T_m^0, \mathscr{B}_{T_m}^0, \boldsymbol{\theta}_{T_m}^0)$. If $m \geq 2$, observe the proximity condition imposed on $T_m^0$.
for each iteration $r = 1, \dots, R$ do
    Set $(T_m^r, \mathscr{B}_{T_m}^r, \boldsymbol{\theta}_{T_m}^r) = (T_m^{r-1}, \mathscr{B}_{T_m}^{r-1}, \boldsymbol{\theta}_{T_m}^{r-1})$.
    Perform one update step of Algorithm 2.1 for the pair copula parameters of tree $T_m$, $\boldsymbol{\theta}_{T_m}^r$; denote the updated parameter entries by $\boldsymbol{\theta}_{T_m}^+$.
    Draw a spanning tree $T_m^r$ that satisfies the proximity condition from the proposal distribution

        $q(T_m^{r-1} = (N_m, E_m^{r-1}) \to T_m^r = (N_m, E_m^r)) \propto p^{|E_m^r \cap E_m^{r-1}|}(1-p)^{|E_m^r \setminus E_m^{r-1}|}.$

    for each edge $e \in E_m^r$ do
        Draw $\mathscr{B}_e^r$ from a Unif$(\mathbf{B})$ distribution.
        Draw $\boldsymbol{\theta}_e^r$ from a multivariate $N(\hat{\boldsymbol{\theta}}_e^r, \hat{\Sigma}_e^r)$ distribution with probability density function $\phi_{(\hat{\boldsymbol{\theta}}_e^r, \hat{\Sigma}_e^r)}(\cdot)$. Here $\hat{\boldsymbol{\theta}}_e^r$ denotes the MLE of the copula parameters of the pair copula $\mathscr{B}_e^r$ and $\hat{\Sigma}_e^r$ denotes the estimated approximative covariance matrix of the parameter estimates $\hat{\boldsymbol{\theta}}_e^r$.
    end for
    Set $\mathscr{B}_{T_m}^r = (\mathscr{B}_e^r, e \in E_m^r)$, $\boldsymbol{\theta}_{T_m}^r = (\boldsymbol{\theta}_e^r, e \in E_m^r)$.
    Evaluate the generalized Metropolis–Hastings acceptance probability of the proposal

        $\alpha := \alpha((T_m^{r-1}, \mathscr{B}_{T_m}^{r-1}, \boldsymbol{\theta}_{T_m}^+), (T_m^r, \mathscr{B}_{T_m}^r, \boldsymbol{\theta}_{T_m}^r))$

        $= \frac{p(\mathscr{V}^*, \mathscr{B}^*, \boldsymbol{\theta}^*)}{p(\mathscr{V}, \mathscr{B}, \boldsymbol{\theta})} \cdot \frac{\prod_{e \in E_m^{r-1}} \phi_{(\hat{\boldsymbol{\theta}}_e^{r-1}, \hat{\Sigma}_e^{r-1})}(\boldsymbol{\theta}_e^+)}{\prod_{e \in E_m^r} \phi_{(\hat{\boldsymbol{\theta}}_e^r, \hat{\Sigma}_e^r)}(\boldsymbol{\theta}_e^r)}$
```

```
            $= \frac{\pi(\mathscr{V}^*, \mathscr{B}^*, \boldsymbol{\theta}^*)}{\pi(\mathscr{V}, \mathscr{B}, \boldsymbol{\theta})} \cdot \frac{L(\mathscr{V}^*, \mathscr{B}^*, \boldsymbol{\theta}^* | \mathbf{x})}{L(\mathscr{V}, \mathscr{B}, \boldsymbol{\theta} | \mathbf{x})} \cdot \frac{\prod_{e \in E_m^{r-1}} \phi_{(\hat{\boldsymbol{\theta}}_e^{r-1}, \hat{\Sigma}_e^{r-1})}(\boldsymbol{\theta}_e^+)}{\prod_{e \in E_m^r} \phi_{(\hat{\boldsymbol{\theta}}_e^r, \hat{\Sigma}_e^r)}(\boldsymbol{\theta}_e^r)},$
        where $\mathscr{V}^* = (\hat{T}_1, \ldots, \hat{T}_{m-1}, T_m^r)$, $\mathscr{V} = (\hat{T}_1, \ldots, \hat{T}_{m-1}, T_m^{r-1})$,
        $\mathscr{B}^* = (\hat{\mathscr{B}}_{T_1}, \ldots, \hat{\mathscr{B}}_{T_{m-1}}, \mathscr{B}_{T_m}^r)$, $\mathscr{B} = (\hat{\mathscr{B}}_{T_1}, \ldots, \hat{\mathscr{B}}_{T_{m-1}}, \mathscr{B}_{T_m}^{r-1})$,
        $\boldsymbol{\theta}^* = (\hat{\boldsymbol{\theta}}_{T_1}, \ldots, \hat{\boldsymbol{\theta}}_{T_{m-1}}, \boldsymbol{\theta}_{T_m}^r)$, and $\boldsymbol{\theta} = (\hat{\boldsymbol{\theta}}_{T_1}, \ldots, \hat{\boldsymbol{\theta}}_{T_{m-1}}, \boldsymbol{\theta}_{T_m}^+)$.
        Accept the proposal $(T_m^r, \mathscr{B}_{T_m}^r, \boldsymbol{\theta}_{T_m}^r)$ with probability $\alpha$; if rejected,
        set $(T_m^r, \mathscr{B}_{T_m}^r, \boldsymbol{\theta}_{T_m}^r) = (T_m^{r-1}, \mathscr{B}_{T_m}^{r-1}, \boldsymbol{\theta}_{T_m}^+)$.
    end for
    return $((T_m^1, \mathscr{B}_{T_m}^1, \boldsymbol{\theta}_{T_m}^1), \ldots, (T_m^R, \mathscr{B}_{T_m}^R, \boldsymbol{\theta}_{T_m}^R))$
```

2.6 Conclusions and Outlook

We have focused on the various model selection and estimation methods for regular vine copulas. Since a regular vine model is specified by three linked components, this results in three fundamental tasks with increasing complexity. We discussed frequentist and Bayesian approaches for each of these tasks. In particular this involved sequential approaches starting from the top tree until the last tree. The frequentist approaches are implemented in the R-packages `CDVine` [15, 57] for D- and C-vines and `VineCopula` [58] for regular vines, respectively.

In view of the linked nature of the vine tree structure the sequential approach is a natural approach. However as in the case of covariate selection in linear models, this might not yield the best fit to the data. In addition the approaches so far primarily considered in-sample fit and model comparisons. More empirical work is needed to validate the models in an out-of-sample performance study. However this is now feasible and tractable and is the subject of current investigations.

It has been recently recognized by [2] that even the flexible class of simplified regular vines might be insufficient for some data sets. This might be the result that the underlying joint density is not well approximated by a simplified regular vine density, where the conditional copula family terms are chosen to be independent of the conditioning value. Currently the proposed solution of [2] using non-parametric two-dimensional smoothing methods is limited to three dimensions and it will be a major challenge to extend the model and the selection methods to higher dimensions.

Other non-standard vine models occur when the pair copula terms depend on the conditioning time point, thus yielding models with time-varying copula parameters. These can also be seen as non-simplified vines in the special case that the conditioning variables follow a functional relationship to time. First parameter-driven time dependence was considered using an AR(1) dynamics in the copula parameters in the papers by [3, 4], while [62] follow a regime switching approach. Here only parameter estimation and assessment of the uncertainty of the parameters are considered so far. It has to be investigated if the additional flexibility of the copula families and the different vine tree structures is needed here.

This survey primarily focused on the selection and estimation problem of regular vines with parametric pair copula families. The approach of [36] can be compared to ones based on kernel methods as proposed in [7, 56].

Other data structures than multivariate continuous data have also gained by models based on pair copula constructions. In particular network structures were considered by [30, 44] in a primarily Gaussian setup and by [7] in a non-Gaussian setting. Here the network is modeled by a directed acyclic graph (DAG) model. While learning the network structure from data is a very complex task even in Gaussian DAG models, non-Gaussian learning algorithms are currently developed and investigated in [6].

Another very interesting multivariate data structure are discrete and mixed discrete-continuous outcomes which occur most often in the life sciences. Pair copula constructions using D-vines for discrete outcomes were recently developed in [55] together with highly efficient parameter estimation techniques. Current research is conducted to allow for mixed outcomes [63].

While these extensions are still concerned primarily with parameter estimation, it remains an important open challenge to find non-sequential solutions to the selection of regular vines when all three components have to be selected.

Acknowledgments The second and third author gratefully acknowledge the support of the TUM Graduate School's International School of Applied Mathematics. The second author is also supported by a research stipend from Allianz Deutschland AG.

References

1. Aas, K., Czado, C., Frigessi, A., Bakken, H.: Pair-copula constructions of multiple dependence. Insur. Math. Econ. **44**, 182–198 (2009)
2. Acar, E.F., Genest, C., Nešlehová, J.: Beyond simplified pair-copula constructions. J. Multivar. Anal. **110**, 74–90 (2012)
3. Almeida, C., Czado, C.: Efficient Bayesian inference for stochastic time-varying copula models. Comput. Stat. Data Anal. **56**(6), 1511–1527 (2012)
4. Almeida, C., Czado, C., Manner, H.: Modeling high dimensional time-varying dependence using D-vine SCAR models. http://arxiv.org/abs/1202.2008 (2012, preprint)
5. Baba, K., Shibata, R., Sibuya, M.: Partial correlation and conditional correlation as measures of conditional independence. Aust. N. Z. J. Stat. **46(4)**, 657–664 (2004)
6. Bauer, A., Czado, C.: Pair copula Bayesian networks. http://arxiv.org/abs/1211.5620 (2012, preprint)
7. Bauer, A., Czado, C., Klein, T.: Pair-copula constructions for non-Gaussian DAG models. Can. J. Stat. **40**(1), 86–109 (2012)
8. Bedford, T., Cooke, R.M.: Probability density decomposition for conditionally dependent random variables modeled by vines. Ann. Math. Artif. Intell. **32**, 245–268 (2001)
9. Bedford, T., Cooke, R.M.: Vines: a new graphical model for dependent random variables. Ann. Stat. **30**(4), 1031–1068 (2002)

10. Berg, D., Aas, K.: Models for construction of higher-dimensional dependence: a comparison study. Eur. J. Financ. **15**, 639–659 (2009)
11. Brechmann, E.C.: Truncated and simplified regular vines and their applications. Master's thesis, Technische Universität München (2010)
12. Brechmann, E.C., Czado, C.: Risk management with high-dimensional vine copulas: An analysis of the Euro Stoxx 50. Stat. Risk Model. (2013, in press)
13. Brechmann, E.C., Czado, C.: COPAR - Multivariate time series modeling using the COPula AutoRegressive model. http://arxiv.org/abs/1203.3328 (2012, preprint)
14. Brechmann, E.C., Czado, C., Aas, K.: Truncated regular vines in high dimensions with application to financial data. Can. J. Stat. **40**(1), 68–85 (2012)
15. Brechmann, E.C., Schepsmeier, U.: Modeling dependence with C- and D-vine copulas: the R-package CDVine. J. Stat. Softw. **52**(3), 1–27 (2012)
16. Brooks, S., Gelman, A., Jones, G.L., Meng, X.L.: Handbook of Markov Chain Monte Carlo. Chapman & Hall/CRC, Boca Raton (2011)
17. Chollete, L., Heinen, A., Valdesogo, A.: Modeling international financial returns with a multivariate regime switching copula. J. Financ. Econom. **7**, 437–480 (2009)
18. Clarke, K.A.: A simple distribution-free test for nonnested model selection. Polit. Anal. **15**(3), 347–363 (2007)
19. Cormen, T., Leiserson, E., Rivest, R., Stein, C.: Introduction to Algorithms. The MIT Press, Cambridge (2009)
20. Czado, C.: Pair copula constructions of multivariate copulas. In: Jaworki, P., Durante, F., Härdle, W., Rychlik, W. (eds.) Workshop on Copula Theory and its Applications. Springer, Berlin (2010)
21. Czado, C., Jeske, S., Hofmann, M.: Selection strategies for regular vine copulae. (2012, preprint)
22. Czado, C., Schepsmeier, U., Min, A.: Maximum likelihood estimation of mixed C-vines with application to exchange rates. Stat. Modelling **12**(3), 229–255 (2012)
23. Dißmann, J., Brechmann, E.C., Czado, C., Kurowicka, D.: Selecting and estimating regular vine copulae and application to financial returns. Comput. Stat. Data Anal. **59**(1), 52–69 (2013)
24. Fischer, M., Köck, C., Schlüter, S., Weigert, F.: An empirical analysis of multivariate copula models. Quant. Financ. **9**(7), 839–854 (2009)
25. Genest, C., Ghoudi, K., Rivest, L.P.: A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. Biometrika **82**, 543–552 (1995)
26. Genest, C., Werker, B.J.M.: Conditions for the asymptotic semiparametric efficiency of an omnibus estimator of dependence parameters in copula models. In: Cuadras, C.M., Fortiana, J., Rodríguez Lallena, J.A. (eds.) Distributions with Given Marginals and Statistical Modelling, pp. 103–112. Kluwer, Dordrecht (2002)
27. Green, P.J.: Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika **82**, 711–732 (1995)
28. Gruber, L.F.: Bayesian analysis of R-vine copulas. Master's thesis. Technische Universität München (2011)
29. Gruber, L.F., Czado C.: Sequential Bayesian model selection of regular vine copulas. (2013, preprint)
30. Hanea, A., Kurowicka, D., Cooke, R.: Hybrid method for quantifying and analyzing Bayesian belief nets. Qual. Reliab. Eng. Int. **22**(6), 613–729 (2006)
31. Hastings, W.: Monte Carlo sampling methods using Markov chains and their applications. Biometrika **57**, 97–109 (1970)
32. Heinen, A., Valdesogo, A.: Asymmetric CAPM dependence for large dimensions: The Canonical Vine Autoregressive Model. CORE discussion papers 2009069. Université catholique de Louvain, Center for Operations Research and Econometrics (CORE) (2009)
33. Hobæk Haff, I.: Parameter estimation for pair-copula constructions. Bernoulli **19**(2), 462–491 (2013)
34. Hobæk Haff, I.: Comparison of estimators for pair-copula constructions. J. Multivar. Anal. **110**, 91–105 (2012)

35. Hobæk Haff, I., Aas, K., Frigessi, A.: On the simplified pair-copula construction - simply useful or too simplistic? J. Multivar. Anal. **101**(5), 1296–1310 (2010)
36. Hobæk Haff, I., Segers, J.: Nonparametric estimation of pair-copula constructions with the empirical pair-copula. http://arxiv.org/abs/1201.5133 (2012, preprint)
37. Joe, H.: 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, pp. 120–141. Institute of Mathematical Statistics, Hayward (1996)
38. Joe, H.: Asymptotic efficiency of the two-stage estimation method for copula-based models. J. Multivar. Anal. **94**, 401–419 (2005)
39. Joe, H., Xu, J.: The estimation method of inference functions for margins for multivariate models. Technical Report 166, Department of Statistics, University of British Columbia (1996)
40. Kojadinovic, I., Yan, J.: Modeling multivariate distributions with continuous margins using the copula R package. J. Stat. Softw. **34(9)**, 1–20 (2010)
41. Kojadinovic, I., Yan, J.: A goodness-of-fit test for multivariate multiparameter copulas based on multiplier central limit theorems. Stat. Comput. **21**, 17–30 (2011)
42. Kojadinovic, I., Yan, J., Holmes, M.: Fast large-sample goodness-of-fit test for copulas. Statistica Sinica **21**, 841–871 (2011)
43. Kurowicka, D.: Optimal truncation of vines. In: Kurowicka, D., Joe, H. (eds.) Dependence Modeling: Vine Copula Handbook. World Scientific, Singapore (2011)
44. Kurowicka, D., Cooke, R.: Distribution-free continuous Bayesian belief nets. In: Wilson, A., Limnios, N., Keller-McNulty, S., Armojo, Y. (eds.) Modern Statistical and Mathematical Methods in Reliability, pp. 309–323. World Scientific, Singapore (2005)
45. Kurowicka, D., Cooke, R.M.: Uncertainty Analysis with High Dimensional Dependence Modelling. Wiley, Chichester (2006)
46. Kurowicka, D., Joe, H.: Dependence Modeling: Vine Copula Handbook. World Scientific, Singapore (2011)
47. Manner, H.: Estimation and model selection of copulas with an application to exchange rates. METEOR research memorandum 07/056. Maastricht University (2007)
48. Mendes, B.V.d.M., Semeraro, M.M., Leal, R.P.C.: Pair-copulas modeling in finance. Financ. Markets Portfolio Manag. **24**(2), 193–213 (2010)
49. Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E.: Equation of State Calculations by Fast Computing Machines. J. Chem. Phys. **21**, 1087–1092 (1953)
50. Meyn, S., Tweedie, R.L.: Markov Chain and Stochastic Stability, 2nd edn. Cambridge University Press, Cambridge (2009)
51. Min, A., Czado, C.: Bayesian inference for multivariate copulas using pair-copula constructions. J. Financ. Econom. **8**(4), 511–546 (2010)
52. Min, A., Czado, C.: Bayesian model selection for D-vine pair-copula constructions. Can. J. Stat. **39**, 239–258 (2011)
53. Morales-Nápoles, O.: Counting vines. In: Kurowicka, D., Joe, H. (eds.) Dependence Modeling: Vine Copula Handbook. World Scientific, Singapore (2011)
54. Nikoloulopoulos, A.K., Joe, H., Li, H.: Vine copulas with asymmetric tail dependence and applications to financial return data. Comput. Stat. Data Anal. **56**(11), 3659–3673 (2012)
55. Panagiotelis, A., Czado, C., Joe, H.: Pair copula constructions for multivariate discrete data. J. Am. Stat. Assoc. **107**(499), 1063–1072 (2012)
56. Schellhase, C., Kauermann, G.: Flexible pair-copula estimation in D-vines with penalized splines. (2012, preprint)
57. Schepsmeier, U., Brechmann, E.C.: CDVine: Statistical Inference of C- and D-Vine Copulas. R package version 1.1-12. http://CRAN.R-project.org/package=CDVine (2013)
58. Schepsmeier, U., Stoeber, J., Brechmann, E.C.: VineCopula: Statistical Inference of Vine Copulas. R package version 1.1-1. http://CRAN.R-project.org/package=VineCopula (2013)
59. Siek, J., Lee, L.Q., Lumsdaine, A.: Boost Graph Library: Random Spanning Tree - 1.49.0. http://www.boost.org/doc/libs/1_49_0/libs/graph/doc/random_spanning_tree.html (2012)

60. Sklar, A.: Fonctions de répartition à n dimensions et leurs marges. Publ. Inst. Statist. Univ. Paris **8**, 229–231 (1959)
61. Smith, M., Min, A., Almeida, C., Czado, C.: Modeling Longitudinal Data Using a Pair-Copula Decomposition of Serial Dependence. J. Am. Stat. Assoc. **105**(492), 1467–1479 (2010)
62. Stöber, J., Czado, C.: Detecting regime switches in the dependence structure of high dimensional financial data. http://arxiv.org/abs/1202.2009 (2011, preprint)
63. Stöber, J., Czado, C., Hong, H., Ghosh, P.: Interrelation among chronic diseases in the elderly. (2012, preprint)
64. Stöber, J., Czado, C., Joe, H.: Simplified pair copula constructions - limits and extensions. http://arxiv.org/abs/1205.4844 (2012, preprint)
65. Stöber, J., Schepsmeier, U.: Is there significant time-variation in multivariate copulas? http://arxiv.org/abs/1205.4841 (2012, preprint)
66. Vuong, Q.H.: Ratio tests for model selection and non-nested hypotheses. Econometrica **57**(2), 307–333 (1989)
67. Whittaker, J.: Graphical Models in Applied Multivariate Statistics. Wiley, New York (1990)