Truncation of vine copulas using fit indices

Eike C. Brechmann* Harry Joe†

February 2, 2015

Abstract

Vine copulas are flexible multivariate dependence models, which are built up from a set of bivariate copulas in different hierarchical levels. However, vine copulas have a computational complexity that is increasing quadratically in the number of variables. This complexity can be reduced by focusing on the sub-class of truncated vine copulas, which use only a limited number of hierarchical levels. We propose a new approach to select the adequate number of levels, such that the vine copula is still sufficiently flexible to provide a good fit to given data. The approach is based on fit indices, as used for structural equation models, to measure the goodness of a fitted truncated model. To select such truncated models, we propose methods to effectively explore the search space of truncated vine copulas, so that we are able to improve over previous greedy sequential approaches that optimized over one tree of the vine at each step. This new selection approach is evaluated in a simulation study as well as in two applications to data sets of financial returns.

Keywords: vine copula, truncated vine, partial correlation, fit index

1 Introduction

Statistical dependence modeling using copulas is very popular, because it allows the convenient separation of the modeling of the univariate marginal distributions from the modeling of the dependence on a standardized scale. Formally, d-dimensional copulas are multivariate distribution functions on $[0,1]^d$ with uniform marginal distribution functions. Their popularity stems from the famous theorem by Sklar (1959), which states that any multivariate distribution is directly linked to a copula in the following way: Let $\boldsymbol{X} = (X_1, \dots, X_d)' \sim F_{1,\dots,d}$ with marginal distribution functions $F_1, \dots, F_d$, then

$$F_{1,\dots,d}(\boldsymbol{x}) = C(F_1(x_1), \dots, F_d(x_d)), \quad \boldsymbol{x} := (x_1, \dots, x_d)' \in (\mathbb{R} \cup \{-\infty, \infty\})^d,$$

where C is a d-dimensional copula. If $\boldsymbol{X}$ is a continuous random vector, then the copula C is unique and the multivariate density $f_{1,\dots,d}$ of $\boldsymbol{X}$ can be decomposed as

$$f_{1,\dots,d}(\boldsymbol{x}) = c(F_1(x_1), \dots, F_d(x_d)) f_1(x_1) \cdots f_d(x_d),$$

where c is the density of the copula C and $f_1, \dots, f_d$ are the marginal densities of $f_{1,\dots,d}$. More details on copulas and their properties can be found in Joe (1997) and Nelsen (2006). Their use in finance is treated, e.g., in Cherubini et al. (2004) and McNeil et al. (2005), while Salvadori et al. (2007) discuss the application of copulas in the natural sciences.

* *Corresponding author.* Center for Mathematical Sciences, Technische Universität München, Germany. E-mail: eike.brechmann@tum.de. Phone: +49 89 289 17425. Fax: +49 89 289 17435.

† Department of Statistics, University of British Columbia, Canada.

Despite their popularity, most common parametric copula families are however rather inflexible in higher dimensions: Archimedean copulas assume exchangeability and have typically only one or two parameters to describe the dependence among variables. Elliptical copulas are more flexible in the sense that pairwise dependence can be variable. Large correlation matrices are however cumbersome to specify and may lead to over-parameterized models. In addition, elliptical copulas are reflection symmetric, that is, the dependence in the joint lower and joint upper tail are the same. A promising alternative to these copulas is the class of vine copulas, based on the vine structure of Bedford and Cooke (2001, 2002), of which a boundary case was considered in Joe (1996). By combining a cascade of arbitrary bivariate copulas in different hierarchical levels, so-called vine trees, vine copulas constitute a large class of flexible multivariate models that are available also for higher-dimensional applications (see, e.g., Dißmann et al. (2013) and Brechmann and Czado (2013)).

An issue that comes along with the model flexibility is however the model complexity. Vine copulas typically have a quadratic number of bivariate building blocks and hence large numbers of parameters. Hence, the question is if there are sub-classes of vine copulas that provide a sufficiently good fit to given data but are more parsimonious than a full specification. One such sub-class is the class of truncated vine copulas, of which the model complexity is reduced by considering only a limited number of vine trees. Such truncated vine copulas were first studied by Brechmann et al. (2012), who also propose a method to identify an appropriate truncation level, that is, the number of vine trees to be used in the model. Their method is however mainly heuristic and sometimes can fall short of effectively exploring the large search space of possible truncated vine copulas.

The contribution of this paper is twofold. First, we propose a new truncation rule, which exploits optimality results by Brechmann and Joe (2014) on the truncation of vine copulas with Gaussian building blocks. The goodness-of-fit of such vine copulas is measured by so-called fit indices, as used in the literature on structural equation modeling (see, e.g., Bollen (1989) and Mulaik (2009)). Using such fit indices, we are then able to truncate such that a certain closeness to the best possible (non-truncated) case is achieved. Moreover, we derive bounds on what is ignored through truncation, so that a possible lack at fit can be detected.

Second, as our approach separates the selection of the different components of a vine copula, namely the vine trees and the bivariate copulas as building blocks, the selection of the truncation level is computationally very efficient and enables us to better explore the search space of truncated vine copulas. While existing approaches so far relied on locally optimal greedy algorithms (that is, optimizing one tree of the vine in a sequence), we propose to use selection procedures inspired by genetic algorithms (see, e.g., Goldberg and Deb (1991) and Blickle and Thiele (1996)), in order to select flexible vine copulas with yet small truncation levels, that is, with reduced model complexity.

The remainder of the paper is organized as follows. Vine copulas are introduced and defined in Section 2. The sub-class of truncated vine copulas is then treated in Section 3, where the optimality results for vine copulas with Gaussian components are also stated. Based on these optimality results, we propose two new truncation rules in Section 4 and derive bounds on the dependencies that are ignored through truncation. In Section 5, we show how to effectively explore the search space of truncated vine copulas. This new selection approach is evaluated in a simulation study in Section 6 and in two applications to data sets of financial returns in Section 7. Finally, Section 8 provides concluding remarks.

2 Vine copulas

Vine copulas take advantage of the fact that a multitude of parametric copula families is known to flexibly model the dependence between two variables. A d-dimensional vine copula is defined

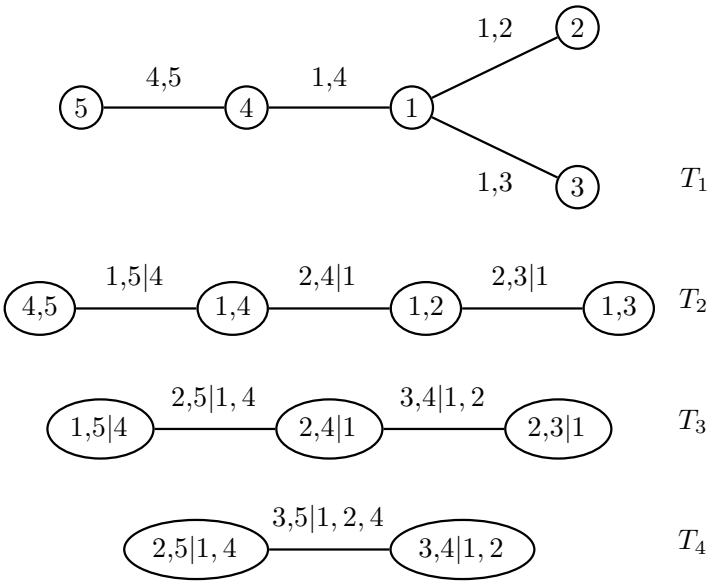

Figure 1: An R-vine $\mathcal{V} = \{T_1, T_2, T_3, T_4\}$ on five elements with edge labels.

in terms of a cascade of $\binom{d}{2} = d(d-1)/2$ such bivariate copulas of arbitrary type. Yet, in order to yield a valid multivariate copula, this cascade has to satisfy certain conditions. *Regular vines*, as introduced by Bedford and Cooke (2001, 2002), are graph theoretical models, which conveniently encode such conditions. A graph is defined in terms of a set of nodes and a set of edges connecting these nodes. Trees are graphs with a unique sequence of edges between each two nodes and are also known as connected acyclic graphs. A regular vine is defined as a linked sequence of trees.

Definition 2.1 (Regular vine)**.** A set of linked trees $\mathcal{V} = \{T_1, T_2, \dots, T_{d-1}\}$ is called a regular vine (R-vine) on d elements if the following three conditions are satisfied.

(i) T_1 is a tree with nodes $N_1 = \{1, \dots, d\}$ and a set of $d-1$ edges denoted by E_1.

(ii) For $i = 2, \dots, d-1$, T_i is a tree with nodes $N_i = E_{i-1}$ and edge set E_i.

(iii) For $i = 2, \dots, d-1$, if $a = \{a_1, a_2\}$ and $b = \{b_1, b_2\}$ are two nodes in N_i, which are connected by an edge, then exactly one of the a_is equals one of the b_is (proximity condition).

Such a d-dimensional R-vine $\mathcal{V}$ has $d(d-1)/2$ edges in $d-1$ trees ($d-1$ edges in T_1, $d-2$ edges in T_2, and so on up to only one edge in T_{d-1}). By associating bivariate copulas with each of these edges, so-called *pair copulas*, we obtain an *R-vine copula*. For a more formal definition, we need to introduce some additional notation. We define the complete union A_e of an edge $e = \{a, b\} \in E_i$ in tree T_i by

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

Then, 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 $\mathcal{C}_{e,a} := A_a \setminus D_e$ and $\mathcal{C}_{e,b} := A_b \setminus D_e$. As shown by Bedford and Cooke (2001), the conditioned sets are singletons, so that we refer to edges by their labels $\{j(e), k(e)|D(e)\} := \{\mathcal{C}_{e,a}, \mathcal{C}_{e,b}|D_e\}$, as illustrated in Figure 1. Using this notation, we define an R-vine copula as follows.

Definition 2.2 (R-vine copula)**.** Let $\boldsymbol{U} = (U_1, \dots, U_d)' \in [0,1]^d$ be a random vector with uniform margins and denote the sub-vector of $\boldsymbol{U}$ with indices in the set D by $\boldsymbol{U}_D = \{U_\ell : \ell \in D\}$. Then $\boldsymbol{U}$ is said to be distributed according to the d-dimensional R-vine copula $C(\cdot; \mathcal{V}, \mathcal{B}, \boldsymbol{\theta})$ if

(i) $\mathcal{V}$ is an R-vine on d elements,

(ii) $\mathcal{B} = \mathcal{B}(\mathcal{V}) = \left\{C_{j(e),k(e);D(e)} : e \in E_i,\ i = 1, \dots, d-1\right\}$ is a set of $d(d-1)/2$ copula families identifying the conditional distributions of $(U_{j(e)}, U_{k(e)})'|\boldsymbol{U}_{D(e)}$, and

(iii) $\boldsymbol{\theta} = \boldsymbol{\theta}(\mathcal{B}(\mathcal{V})) = \left\{\boldsymbol{\theta}_{j(e),k(e);D(e)} : e \in E_i,\ i = 1, \dots, d-1\right\}$ is the set of parameters corresponding to the copulas in $\mathcal{B}(\mathcal{V})$.

One main virtue of R-vine copulas is the availability of a tractable density expression, a result which we reproduce below from Bedford and Cooke (2001).

Theorem 2.3 (R-vine copula density)**.** Let $C(\cdot; \mathcal{V}, \mathcal{B}, \boldsymbol{\theta})$ be a d-dimensional R-vine copula. Then its density is given for $\boldsymbol{u} \in [0,1]^d$ by

$$c(\boldsymbol{u}; \mathcal{V}, \mathcal{B}, \boldsymbol{\theta}) = \prod_{i=1}^{d-1} \prod_{e \in E_i} c_{j(e),k(e);D(e)} \left(C_{j(e)|D(e)}(u_{j(e)}|\boldsymbol{u}_{D(e)}), C_{k(e)|D(e)}(u_{j(e)}|\boldsymbol{u}_{D(e)})\right), \tag{2.1}$$

where the copula $C_{j(e),k(e);D(e)}$ has parameter(s) $\boldsymbol{\theta}_{j(e),k(e);D(e)}$ and $C_{\ell|D(e)}$ is the conditional distribution function of $U_\ell|\boldsymbol{U}_{D(e)},\ \ell \in \{j(e), k(e)\}$.

As it is usually done, we assume that the copula $C_{j(e),k(e);D(e)}$ only depends on the value $\boldsymbol{u}_{D(e)}$ of the conditioning variables $\boldsymbol{U}_{D(e)}$ through its arguments $C_{j(e)|D(e)}(\cdot|\boldsymbol{u}_{D(e)})$ and $C_{k(e)|D(e)}(\cdot|\boldsymbol{u}_{D(e)})$. Recent discussions on this so-called simplifying assumption can be found in Hobæk Haff et al. (2010), Acar et al. (2012) and Stöber et al. (2013). Furthermore, the conditional distribution functions $C_{\ell|D(e)},\ \ell \in \{j(e), k(e)\}$, can be determined recursively tree-by-tree: Let $e \in E_i$, then $C_{\ell|D(e)}$ only depends on pair copulas in trees $T_1, ..., T_i$ (see Dißmann et al. (2013) for more details).

To illustrate Theorem 2.3, we now state the density of an R-vine copula $C(\cdot; \mathcal{V}, \mathcal{B}, \boldsymbol{\theta})$ associated to the R-vine $\mathcal{V}$ shown in Figure 1. For arbitrary pair copulas $\mathcal{B} = \mathcal{B}(\mathcal{V})$ and corresponding parameters $\boldsymbol{\theta} = \boldsymbol{\theta}(\mathcal{B}(\mathcal{V}))$ it is given by

$$\begin{aligned}
&c(\boldsymbol{u}; \mathcal{V}, \mathcal{B}, \boldsymbol{\theta}) = \\
&\quad c_{1,2}(u_1, u_2)\, c_{1,3}(u_1, u_3)\, c_{1,4}(u_1, u_4)\, c_{4,5}(u_4, u_5)\, c_{2,3;1}(C_{2|1}(u_2|u_1), C_{3|1}(u_3|u_1)) \\
&\quad \times c_{2,4;1}(C_{2|1}(u_2|u_1), C_{4|1}(u_4|u_1))\, c_{1,5;4}(C_{1|4}(u_1|u_4), C_{5|4}(u_5|u_4)) \\
&\quad \times c_{3,4;1,2}(C_{3|1,2}(u_3|u_1, u_2), C_{4|1,2}(u_4|u_1, u_2))\, c_{2,5;1,4}(C_{2|1,4}(u_2|u_1, u_4), C_{5|1,4}(u_5|u_1, u_4)) \\
&\quad \times c_{3,5;1,2,4}(C_{3|1,2,4}(u_3|u_1, u_2, u_4), C_{5|1,2,4}(u_5|u_1, u_2, u_4)),
\end{aligned}$$

where $\boldsymbol{u} = (u_1, \dots, u_5)' \in [0,1]^5$.

An R-vine copula gains its flexibility from its three layers: the R-vine tree structure $\mathcal{V}$, the pair copulas $\mathcal{B} = \mathcal{B}(\mathcal{V})$, which can be chosen arbitrarily, and the corresponding parameters $\boldsymbol{\theta} = \boldsymbol{\theta}(\mathcal{B}(\mathcal{V}))$. R-vine copulas with different tree structures or with at least one different pair copula are in general different statistical models. As a result, the class of different R-vine copulas is enormous: Morales-Nápoles (2011) shows that there are $(d!/2) \times 2^{(d-2)(d-3)/2}$ different d-dimensional regular vines, for which $d(d-1)/2$ copula types and parameters can be chosen. A natural question is therefore if there are narrower sub-classes of R-vine copulas, which still yield sufficiently flexible multivariate copulas. Here, we concentrate on the sub-class of *truncated R-vine copulas*.

3 Truncation

Truncated vine copulas have first been considered by Brechmann et al. (2012). They define a truncated R-vine copula as an R-vine copula, of which all pair copulas associated to higher order trees are set to independence copulas. More formally, this can be formulated as follows.

Definition 3.1 (Truncated R-vine copula)**.** Let $\boldsymbol{U} = (U_1, \dots, U_d)'$ be a random vector with uniform margins and let $\ell \in \{0, \dots, d-1\}$ be the truncation level. Then $\boldsymbol{U}$ is said to be distributed according to the d-dimensional ℓ-truncated R-vine copula $C(\cdot; \mathcal{V}, \mathcal{B}, \boldsymbol{\theta})$ if $C(\cdot; \mathcal{V}, \mathcal{B}, \boldsymbol{\theta})$ is a d-dimensional R-vine copula (see Definition 2.2) with

$$C_{j(e),k(e);D(e)} = \Pi \quad \forall e \in E_i,\ i = \ell+1, \dots, d-1,$$

where Π denotes the bivariate independence copula.

Truncation hence greatly reduces the model complexity, since the vine tree structure after T_ℓ does not influence the statistical model and since the number of different pair copulas is reduced to

$$\sum_{i=1}^{\ell}(d-i) = \ell d - \frac{\ell(\ell+1)}{2} = \frac{\ell(2d-(\ell+1))}{2}, \tag{3.1}$$

which is only linear in the dimension d for fixed ℓ. The special case of $\ell = 0$ corresponds to the multivariate independence copula, while $\ell = 1$ corresponds to a Markov tree model and $\ell = d-1$ yields a fully specified, non-truncated R-vine copula.

Since the density of the independence copula is equal to one, the density expression of a truncated R-vine copula is also of simplified form (see Equation (2.1)). It is given by

$$c(\boldsymbol{u}; \mathcal{V}, \mathcal{B}, \boldsymbol{\theta}) = \prod_{i=1}^{\ell}\prod_{e \in E_i} c_{j(e),k(e);D(e)}\left(C_{j(e)|D(e)}(u_{j(e)}|\boldsymbol{u}_{D(e)}), C_{k(e)|D(e)}(u_{j(e)}|\boldsymbol{u}_{D(e)})\right). \tag{3.2}$$

This considerably facilitates parameter estimation using likelihood-based techniques.

Furthermore, Brechmann et al. (2012) and Dißmann et al. (2013) justify the use of truncated R-vine copulas as follows. The most important (and strongest) dependencies among variables can typically be captured best by the pair copulas of the first trees. This is underlined by the fact that, for all variables to be tail dependent, it is sufficient that all pair copulas in the first tree are tail dependent (see Joe et al. (2010)). Moreover, the influence of rounding errors in the tree-by-tree calculation of the conditional distribution functions as arguments of the pair copula densities is minimized. That is, truncated R-vine copulas are not only parsimonious models but also have the potential to specify dependence similarly flexibly as a non-truncated R-vine copula.

For a given data set, the truncation challenge can now be formulated as follows. First, we need to identify an appropriate truncation level ℓ such that an ℓ-truncated R-vine copula fits the data well. Clearly, the smaller ℓ is, the more parsimonious the model is. Second, among all ℓ-truncated R-vine copulas, we need to select the best fitting model. If this is achieved, model parsimony and flexibility can be balanced.

Note that because of the flexibility of R-vine copulas, they should be considered as capable of approximating most multivariate copulas. Hence this means that there can be different R-vine copulas that can equally well approximate a given multivariate copula, and this is also the case for truncated R-vine copulas. With the parsimony principle, we would prefer a truncation at a lower level if possible.

Brechmann et al. (2012) propose a sequential approach to the truncation challenge. They iteratively construct R-vine copulas, which are 1-truncated, 2-truncated, and so on. In each step, they use the Vuong test for model comparison to assess the gain of extending the truncated R-vine copula by an additional tree. If the gain is determined to be negligible, the procedure is stopped and the R-vine copula is truncated at the current level. A disadvantage of this method is that in each step a truncated R-vine copula has to be fully specified, that is, a vine tree has to be determined, pair copulas have to be selected and parameters have to be estimated. Due

to the related computational effort, this hinders an efficient exploration of the search space of truncated R-vine copula models and the authors settle for a locally optimal greedy algorithm in each step of their procedure.

Because a greedy algorithm with local optimality cannot always be globally optimal, heuristics are desirable for alternative sequential approaches. The heuristic considered in this paper is based on the separation of the selection and the estimation of the pair copulas from the identification of appropriate vine tree structures. We propose a sequential approach that iteratively considers ℓ-truncated models with $\ell = 1, 2, \ldots$ to find the smallest truncation level that meets a desired fit index. For this, we build on a result by Brechmann and Joe (2014): If data comes from a random vector $\boldsymbol{Z} := (Z_1, \ldots, Z_d)'$ distributed according to a multivariate normal distribution with zero mean vector and unit variances (possibly after standardization), then the optimal (with respect to maximum likelihood estimation) ℓ-truncated R-vine copula with vine tree structure $\mathcal{V} = \{T_1, T_2, \ldots, T_{d-1}\}$ and Gaussian pair copulas is given by the trees $T_1, \ldots, T_\ell$ that minimize

$$\sum_{i=1}^{\ell} \sum_{e \in E_i} \log\left(1 - r^2_{j(e),k(e);D(e)}\right), \tag{3.3}$$

where $r_{j(e),k(e);D(e)}$ is the empirical partial correlation coefficient of the variables $Z_{j(e)}$ and $Z_{k(e)}$ given $\boldsymbol{Z}_{D(e)}$. The trees $T_{\ell+1}, \ldots, T_{d-1}$ can be chosen arbitrarily according to the R-vine construction rules (see Definition 2.1), since they do not influence the model likelihood (recall that the density of the independence copula is 1; see Equation (3.2)). In other words, under multivariate normality, finding the optimal truncated R-vine copula with Gaussian pair copulas requires no estimation of pair copula parameters but only the selection of an appropriate vine tree structure $\mathcal{V}$. This selection is based on empirical partial correlation coefficients, which can be computed recursively using the formula

$$r_{j,k;D} = \frac{r_{j,k;D\setminus m} - r_{j,m;D\setminus m} r_{m,k;D\setminus m}}{\sqrt{1 - r^2_{j,m;D\setminus m}}\sqrt{1 - r^2_{m,k;D\setminus m}}},$$

where $m \in D$, if $D \neq \emptyset$, and dependence on the edge e is suppressed for better readability. We therefore call such an R-vine copula a *partial correlation R-vine*.

While in general no optimality results about best truncation levels are available, we exploit this result here and show how it can conveniently be used for vine tree selection. Clearly, data will mostly not come from a multivariate normal random vector. Let $\boldsymbol{X} := (X_1, \ldots, X_d)'$ have marginal distribution functions $F_1, \ldots, F_d$. Then, the normal or van der Waerden score $Z_j := \Phi^{-1}(F_j(X_j))$, $j = 1, \ldots, d$, where Φ is the standard normal distribution function, is standard normally distributed, while the joint distribution of $(Z_1, \ldots, Z_d)'$ will generally be different from the multivariate normal (unless the copula of $\boldsymbol{X}$ is the Gaussian). Nonetheless, the correlation matrix of the normal scores provides a standardized measure of pairwise between-variable dependence in the same spirit as matrices of pairwise Spearman's ρ or Kendall's τ values. As it is a minimum requirement of any multivariate copula to accurately model this general level of dependence, we propose to select the vine tree structure $\mathcal{V} = \{T_1, T_2, \ldots, T_{d-1}\}$ of an ℓ-truncated R-vine copula as the optimal tree structure of an ℓ-truncated partial correlation R-vine for the normal scores. That is, we select $T_1, \ldots, T_\ell$ such that they minimize Equation (3.3), where the partial correlation coefficients are calculated based on the empirical normal scores of the original data. Yet in other words, we select the tree structure such that we best approximate the correlation matrix of the normal scores in terms of a truncated partial correlation R-vine.

Given this optimality result, two open questions remain to solve the truncation challenge: How do we identify an appropriate truncation level ℓ? And how do we actually optimize Equation (3.3) with respect to the tree structure? These questions are now treated one after the other in the next two sections.

4 Fit indices for truncation

Whether an empirical correlation matrix R is well approximated by a correlation matrix $\widehat{R}$ fitted according to a certain structural model is a common question in structural equation modeling (see, e.g., Bollen (1989) and Mulaik (2009)). We assume that the sample size n is large enough to meaningfully compare different d-dimensional structured correlation matrices. To test the goodness-of-fit of the structural model, a likelihood ratio test is often used. Its test statistic is given by

$$D(\widehat{R}|R) := n\left(\log|\widehat{R}| - \log|R| + \operatorname{tr}(\widehat{R}^{-1}R) - d\right). \tag{4.1}$$

If the model is completely unstructured (the saturated model), then $\widehat{R} = R$ and $D(\widehat{R}|R) = 0$. On the other hand, if the model assumes that all variables are unrelated, then all off-diagonal entries of $\widehat{R}$ are zero and $D_0 := D(\widehat{R}|R) = -n\log|R|$. Reasonable models should lie somewhere in between these two extreme cases, since it is clearly essential for any model to improve over the case of unrelated variables. At the same time, a structured correlation model should not be over-specified and have more parameters than correlation coefficients to estimate.

To measure the relative distance of a fitted from the empirical correlation matrix, fit indices can be used. Bentler and Bonett (1980) define the *normed fit index (NFI)* as

$$\mathrm{NFI}(\widehat{R}|R) := \frac{D_0 - D(\widehat{R}|R)}{D_0} = 1 - \frac{D(\widehat{R}|R)}{D_0}. \tag{4.2}$$

It is greater or equal to zero if the (extreme) case of unrelated variables is nested within the models. It follows that the smaller $D(\widehat{R}|R)$ is, the closer $\mathrm{NFI}(\widehat{R}|R)$ is to 1 and the better is the fit.

In our setting, the structural model is an ℓ-truncated partial correlation R-vine and we denote the corresponding fitted correlation matrix by $\widehat{R}_\ell$. According to Brechmann and Joe (2014), assuming Gaussian standardized data, it holds under maximum likelihood estimation that $\operatorname{tr}(\widehat{R}_\ell^{-1}R) = d$ and Equation (4.1) simplifies to

$$D_\ell := D(\widehat{R}_\ell|R) = n\left(\log|\widehat{R}_\ell| - \log|R|\right).$$

For general normal scores from copula models, D_ℓ is viewed as a discrepancy measure. This yields the following simple formula of the NFI (4.2) of ℓ-truncated partial correlation R-vines:

$$\mathrm{NFI}(\widehat{R}_\ell|R) = 1 - \frac{D_\ell}{D_0} = \frac{\log|\widehat{R}_\ell|}{\log|R|} =: \delta_\ell. \tag{4.3}$$

That is, we can use δ_ℓ to measure the closeness to the empirical correlation matrix and truncate the R-vine copula when δ_ℓ is close enough to 1. We therefore define the *optimal NFI truncation level* as

$$\ell_\delta^* = \min\left\{\ell \in \{0,\dots,d-1\} : \delta_\ell \geq (1-\alpha)\right\}, \tag{4.4}$$

where $\alpha \in (0,1)$ is a small positive number.

A disadvantage of the NFI is however that it does not take into account the number of model parameters and hence does not reward model parsimony. An alternative fit index, which does exactly this, is the *comparative fit index (CFI)* by Bentler (1990). The degrees of freedom ν_ℓ of an ℓ-truncated partial correlation R-vine are

$$\nu_\ell = \frac{d(d-1)}{2} - \frac{\ell(2d-\ell-1)}{2} = \frac{(d-\ell)(d-\ell-1)}{2}, \tag{4.5}$$

according to Equation (3.1); in particular, $\nu_0 = d(d-1)/2$ in the case of complete independence ($\ell = 0$). Then the CFI of an ℓ-truncated partial correlation R-vine is defined as

$$\mathrm{CFI}(\widehat{R}_\ell|R) := 1 - \frac{\max(0, D_\ell - \nu_\ell)}{\max(0, D_0 - \nu_0, D_\ell - \nu_\ell)} =: \gamma_\ell, \tag{4.6}$$

which takes on values between 0 and 1. We assume $D_0 - \nu_0 > 0$ or $D_\ell - \nu_\ell > 0$; otherwise the CFI is not well defined. If the true correlation matrix is the identity matrix, then D_0 has a distribution that is close to $\chi^2_{\nu_0}$ and hence $\Pr(D_0 - \nu_0 < 0)$ is close to 0.5. With more dependence, $\Pr(D_0 - \nu_0 < 0)$ is much smaller and is negligible if the sample size is not small or the dependence is strong enough to be detected with a small sample size. In practice, we look for a parsimonious non-trivial dependence model, such as a truncated vine, only when there is non-negligible to strong dependence. It is possible that $D_\ell - \nu_\ell < 0$ for a good-fitting or overfitted truncation level. Hence when $D_0 - \nu_0 > 0$, the CFI will reach a value of 1 at truncation level $d-1$ or before. The corresponding *optimal CFI truncation level* is given by

$$\ell^*_\gamma = \min\left\{\ell \in \{0, \ldots, d-1\} : \gamma_\ell \geq (1-\alpha)\right\}, \tag{4.7}$$

where $\alpha \in (0, 1)$. As a rule of thumb, a model is often already considered to provide a good fit when the CFI exceeds 0.95, that is, when $\alpha = 0.05$ (see Mulaik (2009, Chapter 15) for an overview and a discussion of appropriate levels of α).

Hu and Bentler (1999) also have comparisons with fit indices that can exceed 1. The determination of truncation levels, similar to the above, can be based on these indices or others that have good properties. We next list three of these fit indices with brief comparisons and comments, but subsequently concentrate on the NFI and the CFI.

- The *Tucker-Lewis index (TLI)* or *non-normed fit index*:

$$\mathrm{TLI}(\widehat{R}_\ell|R) := \frac{D_0/\nu_0 - D_\ell/\nu_\ell}{D_0/\nu_0 - 1}.$$

  Because of the ratios instead of the differences, the TFI can behave erratically as ℓ increases. As a result, we do not recommended to use this as a criterion for finding the appropriate truncation level.

- The *relative non-centrality index (RNI)*:

$$\mathrm{RNI}(\widehat{R}_\ell|R) := \frac{(D_0 - \nu_0) - (D_\ell - \nu_\ell)}{D_0 - \nu_0}.$$

  This is the same as the CFI when $D_0 - \nu_0 > D_\ell - \nu_\ell > 0$. If the dependence is strong enough, both inequalities tend to hold for ℓ much smaller than d. The second inequality might fail for ℓ closer to $d-1$ and then the RNI is greater than 1.

- The *incremental fit index (IFI)*:

$$\mathrm{IFI}(\widehat{R}_\ell|R) := \frac{D_0 - D_\ell}{D_0 - \nu_0}. \tag{4.8}$$

  It behaves somewhat like the NFI or the CFI, but it is larger: It is larger than the NFI because of a smaller denominator. It is also larger than the RNI because the numerator is larger by $\nu_0 - \nu_\ell$. The IFI can hence more often exceed 1 for ℓ getting closer to $d-1$ and lead to a lower truncation level than the NFI or the CFI. This is demonstrated in Section 7.

To assess any of the above fit indices for an ℓ-truncated partial correlation R-vine we need to calculate (the logarithm of) the determinant of $\widehat{R}_\ell$. This is in fact straightforward, as Kurowicka and Cooke (2006) show that it holds for the saturated model and a vine tree structure $\mathcal{V} = \{T_1, \ldots, T_{d-1}\}$ that

$$\log|\widehat{R}_{d-1}| = \log|R| = \sum_{i=1}^{d-1} \sum_{e \in E_i} \log\left(1 - r^2_{j(e),k(e);D(e)}\right).$$

In an ℓ-truncated model, partial correlation coefficients in trees $T_{\ell+1}, \ldots, T_{d-1}$ are set to 0, so that we obtain

$$\log|\widehat{R}_\ell| = \sum_{i=1}^{\ell} \sum_{e \in E_i} \log\left(1 - r^2_{j(e),k(e);D(e)}\right),$$

which is the target quantity to be minimized for maximum likelihood estimation (see Equation (3.3)).

Truncation at level $\ell^* := \ell^*_\delta$ according to the NFI truncation rule given in Equation (4.4) hence implies that

$$1 - \alpha \leq 1 - \alpha^* := \delta_{\ell^*} = \frac{\log|\widehat{R}_{\ell^*}|}{\log|R|} = \frac{\log|\widehat{R}_{\ell^*}|}{\log|\widehat{R}_{\ell^*}| + \sum_{i=\ell^*+1}^{d-1} \sum_{e \in E_i} \log(1 - r^2_{j(e),k(e);D(e)})},$$

where $\alpha^* \leq \alpha$. Rearranging terms and defining $E^c_{\ell^*} := \{E_{\ell^*+1}, \ldots, E_{d-1}\}$ yields

$$-\sum_{e \in E^c_{\ell^*}} \log\left(1 - r^2_{j(e),k(e);D(e)}\right) = -\frac{\alpha^*}{1-\alpha^*} \log|\widehat{R}_{\ell^*}|. \tag{4.9}$$

Using this equality, we can derive some results on the size of the partial correlation coefficients that are set to zero due to the model assumptions (truncation at level ℓ^*). For their maximum value we obtain an upper bound as follows. From Equation (4.9), we have that

$$\max_{e \in E^c_{\ell^*}} \left\{-\log\left(1 - r^2_{j(e),k(e);D(e)}\right)\right\} \leq -\frac{\alpha^*}{1-\alpha^*} \log|\widehat{R}_{\ell^*}|,$$

so that

$$\max_{e \in E^c_{\ell^*}} r^2_{j(e),k(e);D(e)} \leq 1 - |\widehat{R}_{\ell^*}|^{\alpha^*/(1-\alpha^*)}.$$

Taking the square root and substituting $\alpha^* = 1 - \log|\widehat{R}_{\ell^*}|/\log|R|$ thus results in the following upper bound on the maximum absolute partial correlation coefficient in trees $T_{\ell^*+1}, \ldots, T_{d-1}$:

$$\max_{e \in E^c_{\ell^*}} |r_{j(e),k(e);D(e)}| \leq \left(1 - |\widehat{R}_{\ell^*}|^{\log|R|/\log|\widehat{R}_{\ell^*}|-1}\right)^{1/2} = \left(1 - \frac{|R|}{|\widehat{R}_{\ell^*}|}\right)^{1/2}. \tag{4.10}$$

Moreover, we can calculate an upper bound on the average "ignored" partial correlation coefficients in trees $T_{\ell^*+1}, \ldots, T_{d-1}$. For this, note that the number of edges in $E^c_{\ell^*}$ is given by

$$\sum_{i=\ell^*+1}^{d-1} |E_i| = \sum_{i=\ell^*+1}^{d-1} (d-i) = \frac{(d-\ell^*)(d-\ell^*-1)}{2} = \nu_{\ell^*},$$

where ν_{ℓ^*} are the degrees of freedom of an ℓ^*-truncated partial correlation R-vine (see Equation (4.5)). By repeated application of Jensen's inequality, it follows that

$$-\frac{1}{\nu_{\ell^*}} \sum_{e \in E^c_{\ell^*}} \log\left(1 - r^2_{j(e),k(e);D(e)}\right) \geq -\log\left(1 - \frac{1}{\nu_{\ell^*}} \sum_{e \in E^c_{\ell^*}} r^2_{j(e),k(e);D(e)}\right)$$
$$\geq -\log\left(1 - \left(\frac{1}{\nu_{\ell^*}} \sum_{e \in E^c_{\ell^*}} |r_{j(e),k(e);D(e)}|\right)^2\right).$$

Combining this with Equation (4.9) yields

$$-\nu_{\ell^*} \log\left(1 - \left(\frac{1}{\nu_{\ell^*}} \sum_{e \in E^c_{\ell^*}} |r_{j(e),k(e);D(e)}|\right)^2\right) \leq -\frac{\alpha^*}{1-\alpha^*} \log|\widehat{R}_{\ell^*}|,$$

and, after rearrangement of the terms as above,

$$\frac{1}{\nu_{\ell^*}} \sum_{e \in E^c_{\ell^*}} |r_{j(e),k(e);D(e)}| \leq \left(1 - |\widehat{R}_{\ell^*}|^{\frac{\log|R|/\log|\widehat{R}_{\ell^*}|-1}{\nu_{\ell^*}}}\right)^{1/2} = \left(1 - \left(\frac{|R|}{|\widehat{R}_{\ell^*}|}\right)^{1/\nu_{\ell^*}}\right)^{1/2}. \quad (4.11)$$

These upper bounds allow to assess the effect of truncation on the remaining vine trees. While the bound on the maximum value is rather a worst case consideration, the bound on the average value is more realistic and allows to quantify the size of the "ignored" partial correlation coefficients and to detect a potential lack of fit.

Similarly, bounds on the partial correlation coefficients in the tree directly after the truncation level, T_{ℓ^*+1}, can be obtained. This gives an indication what the gain of fitting an additional tree would have been—which is the information that Brechmann et al. (2012) use for their truncation method. In our approach, we however look at all remaining trees and not only the next one.

Finally, such bounds cannot only be derived for the NFI but also for the CFI truncation rule given in (4.7). If we have at truncation at level $\ell^* := \ell^*_\gamma$ that $D_0 - \nu_0 \geq D_{\ell^*} - \nu_{\ell^*} \geq 0$, then

$$1 - \alpha \leq 1 - \alpha^* := \gamma_{\ell^*} = 1 - \frac{D_{\ell^*} - \nu_{\ell^*}}{D_0 - \nu_0} = \frac{\log|R_{\ell^*}| + (\nu_0 - \nu_{\ell^*})/n}{\log|R| + \nu_0/n},$$

where $\alpha^* \leq \alpha$. Using this equation, the above derivations of the bounds on the partial correlation coefficients associated with the edges $E^c_{\ell^*}$ are then straightforward to replicate for the CFI. In fact, it can be shown that the upper bounds on the maximum and on the average absolute partial correlation coefficients in trees $T_{\ell^*+1}, ..., T_{d-1}$ are the same as in Equations (4.10) and (4.11), respectively. If $D_0 - \nu_0 \geq 0 \geq D_{\ell^*} - \nu_{\ell^*}$, similar derivations are possible.

5 Model selection

The final challenge is now to find good ℓ-truncated partial correlation R-vines for $\ell = 1, 2, ...$ until truncation is possible. In other words, we need to optimize Equation (3.3) with respect to the tree structure. For $\ell = 1$ this is trivial: As T_1 has to be a tree on all nodes $N_1 = \{1, \dots, d\}$, a so-called spanning tree, and minimize Equation (3.3), we associate the edge weight $\log(1 - r^2_{j(e),k(e)})$ to each edge $e \in N_1 \times N_1$ and determine the solution as the minimum spanning tree (MST) in terms of these edge weights; by definition, the MST has the minimal sum of edge weights among all spanning trees. Such an MST can be calculated for instance using the algorithm by Prim (1957).

Identification of the best two-truncated partial correlation R-vine is however already significantly more difficult: According to Definition 2.1, it holds that $N_2 = E_1$. In other words, the choice of edges $\widetilde{E}_2 \subset N_2 \times N_2$, which are feasible according to the proximity condition, depends on how the first vine tree is selected. If the first vine tree is given, the best fitting second tree T_2 is obviously again the MST in terms of the edge weights $\log(1 - r^2_{j(e),k(e)|D(e)})$ for all $e \in \widetilde{E}_2$. In total, there are however d^{d-2} different first vine trees by Cayley's theorem. This means that the search space of 2-truncated partial correlation R-vines is extremely large and an exhaustive search is possible only for small d.

Due to the linked nature of the vine trees, this problem becomes worse for subsequent trees and methods to reasonably explore the search space of ℓ-truncated partial correlation R-vines are vital for a good model fitting. As a starting point, we consider the greedy algorithm, where in each step the locally best choice is made. That is, given trees $T_1, \dots, T_{\ell-1}$, we select tree T_ℓ as the MST in terms of the edge weights $\log(1 - r^2_{j(e),k(e)|D(e)})$ for all $e \in \widetilde{E}_\ell$, where $\widetilde{E}_\ell \subset N_\ell \times N_\ell$ is the set of feasible edges according to the proximity condition. Such a greedy approach was already used by Brechmann et al. (2012) for their truncation method, which uses empirical Kendall's τ values as edge weights. Yet, due to the high computational effort of their approach, they do not attempt to further explore the search space of truncated R-vine copulas.

For small truncation levels, Brechmann and Joe (2014) propose to consider the following tree specifications in addition to the greedy algorithm:

- *Neighbors of trees:* As the locally optimal solution, the MST is a reasonable starting point for the search of a better solution. We therefore consider neighbors of the MST, where a 1-neighbor of a tree T is obtained as follows:

  (i) Choose an edge $e \notin T$ and set $T' = T \cup \{e\}$. Then the graph T' is no longer a tree but has a cycle K of edges including e, that is, $e \in K$.

  (ii) Remove an edge $f \in K \setminus \{e\}$ from T' to obtain $T'' = T' \setminus \{f\}$. T'' is a tree and called 1-neighbor of T.

  As an example for a 1-neighbor consider T_1 in Figure 1. If one adds the edge $\{2, 3\}$, there is the cycle $K = \{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$. By removing for instance the edge $\{1, 3\}$, one obtains a valid tree again. As neighbors of an MST T, we concentrate on those edges e that have the smallest weights $\log(1 - r^2_{j(e),k(e)|D(e)})$, but are not included in T. To obtain a tree with large overall weight, we then select f as the edge with the largest weight in $K \setminus \{e\}$. There are $d(d-1)/2 - (d-1) = (d-1)(d-2)/2$ such 1-neighbors in the first vine level. The number of 1-neighbors in subsequent levels depends on the proximity condition.

- *Best spanning trees:* Another option to go beyond the MST is to take into account a certain number of best spanning trees, that is, the spanning trees with smallest weight. Identifying these best spanning trees is known to be an NP-hard problem, but there are several algorithms available to optimize the running time (see, e.g., Gabow (1977)). As the number of possible trees depends on the proximity condition, the number of best spanning trees of course also does.

A combination of these alternative tree specifications may improve over the greedy algorithm. For instance, a 3-truncated partial correlation R-vine with third-best spanning tree as T_1 and third 1-neighbor of the MST as T_2 might be better than the greedy algorithm with MSTs for each vine tree. Note that T_ℓ is always chosen as the MST for a given truncation level ℓ.

However, even if only a small number of best spanning trees and neighbors is considered as candidate solutions, the number of potential ℓ-truncated partial correlation R-vines quickly increases: If we consider t candidate solutions at each level, there are up to $t^{\ell-1}$ different ℓ-truncated partial correlation R-vines to choose from (depending on the proximity condition). To

reduce this number but still reasonably explore the search space of ℓ-truncated partial correlation R-vines, we propose the following approach. It is inspired by the selection in genetic algorithms (see, e.g., Goldberg and Deb (1991) and Blickle and Thiele (1996)), where the challenge is also to balance the exploitation (focus on the best solutions) and the exploration of the search space. We proceed as follows:

(i) Calculate the NFI value δ_1 or the CFI value γ_1 of the MST for T_1, which is the best 1-truncated partial correlation R-vine.

(ii) If truncation at level 1 is possible according to the chosen truncation rule ($\ell^* = 1$), truncate and return the MST.

(iii) Otherwise, generate a range of t candidate solutions for T_1 (neighbors, best spanning trees).

(iv) For each of the t 1-truncated candidates, choose the MST as T_2 and calculate the NFI value δ_2 or the CFI value γ_2.

(v) If truncation at level 2 is possible for one of the models according to the chosen truncation rule ($\ell^* = 2$), truncate and return the model with the largest fit index value (best 2-truncated R-vine).

(vi) Otherwise, for each of the t 1-truncated candidates, select a range of t candidates for T_2 (neighbors, best spanning trees). The actual number of candidates may be less than t when the proximity condition only admits a smaller number of spanning trees.

(vii) We now have up to t^2 2-truncated vines as candidates. Since these are too many candidates to continue with, we choose an appropriate number $t_{(2)} < t^2$ of candidates according to their fit index value.

(viii) For each of the $t_{(2)}$ 2-truncated candidates, choose the MST as T_3 and calculate the NFI value δ_3 or the CFI value γ_3.

(ix) If truncation at level 3 is possible for one of the models according to the chosen truncation rule ($\ell^* = 3$), truncate and return the model with the largest fit index value (best 3-truncated R-vine).

(x) Otherwise, for each of the $t_{(2)}$ 2-truncated candidates, select a range of t candidates for T_3 (neighbors, best spanning trees), and so on, until truncation is possible.

That is, in order to prevent the number of candidate solutions to blow up excessively, we select in each step a certain number of the candidates and discard the rest. For $\ell = 2, ..., d-2$ and the $\widetilde{t} := t_{(\ell-1)} \times t$ candidate solutions, this selection is carried out as follows.

First, we compute the *fitness* of each ℓ-truncated candidate solution as

$$f_i = \frac{\delta_{\ell,i}}{1-\alpha} \quad \text{or} \quad f_i = \frac{\gamma_{\ell,i}}{1-\alpha}, \quad i = 1, \dots, \widetilde{t},$$

where $\delta_{\ell,i}$ is the NFI (4.3) and $\gamma_{\ell,i}$ is the CFI (4.6) of the ith candidate and α is given through the truncation rule (4.4) or (4.7), respectively. If any $f_i \geq 1$, then truncation is possible. Otherwise, let $\eta \in [0,1]$ and retain the $\eta \times 100\%$ best solutions $B \subset \{1, \dots, \widetilde{t}\}$ according to their fitness $f_i,\ i \in \{1, \dots, \widetilde{t}\}$. This is called *elitism*. In addition to this elitist selection, we choose $\lambda \in [0,1]$ such that $\lambda \geq \eta$ and randomly pick $(\lambda - \eta) \times 100\%$ of the remaining candidates according to the probabilities

$$p_i = \frac{f_i}{\sum_{j \in \{1,\dots,\widetilde{t}\} \setminus B} f_j}, \quad i \in \{1, \dots, \widetilde{t}\} \setminus B,$$

so that candidates with large fitness value are more likely to be selected.

Hence, in total $\lambda \times 100\%$ of the candidates are chosen after elitism and random selection. All other candidates are discarded. Note that the actual MST can be selected in each step, so that the greedy solution is a possible outcome of our selection strategy in the case that it provides a good fit.

In summary, our truncation approach for R-vine copulas requires the choice of the following control parameters.

- α: the desired degree of closeness of the fitted to the empirical correlation matrix of the normal scores;
- t_{bST}, t_{N}: the number of best spanning trees and 1-neighbors, respectively, to be looked at in each step in order to generate new candidate solutions ($t = t_{\text{bST}} + t_{\text{N}}$);
- λ: the total percentage of solutions to be retained after elitism and random selection;
- η: the percentage of best solutions to be retained (without random selection).

According to these parameters, the algorithm identifies an appropriate truncation level ℓ^* and corresponding trees $T_1, \dots, T_{\ell^*}$. Since trees $T_{\ell^*+1}, \dots, T_{d-1}$ do not impact the model fit and can therefore be chosen arbitrarily, we hence have identified a full vine tree structure $\mathcal{V} = \{T_1, T_2, \dots, T_{d-1}\}$. To specify a full R-vine copula with pair copulas $\mathcal{B}(\mathcal{V})$ and corresponding parameters $\boldsymbol{\theta}(\mathcal{B}(\mathcal{V}))$, the final step is then to select the pair copulas according to a selection criterion such as the AIC and estimate the parameters. These two issues have already been treated extensively in the literature (see Czado et al. (2013) for a recent overview). For instance, pair copula selection according to the AIC proceeds by fitting a range of bivariate copulas to a given pair of variables and then selection the copula with the smallest AIC value (see Section 7). Tail-weighted dependence measures, such as lower and upper semi-correlations, can also be a guide to the choice of bivariate copula families.

To aid in the choice of control parameters for our algorithm, we made comparisons with the optimal truncated partial correlation R-vines at each level $\ell < d-1$ when $d \leq 7$; Joe (2014) has algorithms for enumerating through all R-vines and these are practical for simulations only if $d \leq 7$. By enumeration for $d \leq 7$, it is found in simulations that the minimum CFI value is over 0.99 for 1-truncated vine copulas based on Gaussian, Gumbel and BB1 pair copulas. This suggests using $\alpha = 0.01$ for the CFI truncation rule. For 2-truncated Gaussian and non-Gaussian vines with $\alpha = 0.01$ and $d \leq 7$, the enumeration approach was able to choose the level of 2 in almost 100% of the time; however a 2-truncated partial correlation R-vine can be well approximated by another 2-truncated vine, where some edges of trees T_1 and T_2 have been interchanged. For $\ell \geq 3$, from the enumeration, we find that it is possible that an ℓ-truncated partial correlation R-vine is well approximated by a truncated R-vine at level ℓ or lower. The remaining control parameters t_{bST}, t_N, λ and η were chosen so that our algorithm has a very high proportion of choosing a level of 2 for $d \geq 7$ when the simulated data comes from a 2-truncated vine copula.

6 Simulation study

In this section we have a closer look at two different questions. First, we investigate by how much our newly proposed selection strategy for truncated partial correlation R-vines improves over the greedy algorithm. For this, we generate d-dimensional correlation matrices according to ℓ-truncated partial correlation R-vines with random tree structure and (partial) correlation coefficients drawn randomly from $(-1, 1)$, where $\ell \in \{2, 3, 4\}$ and $d \in \{5, 7, 10, 15\}$. We then select ℓ-truncated partial correlation R-vines according to the selection approach outlined in

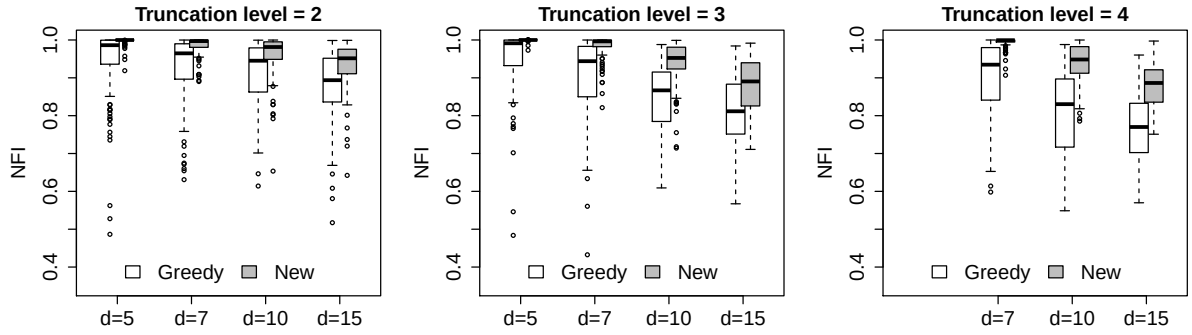

Figure 2: Box plots of 100 NFIs δ_ℓ according to the greedy and our new selection approach. The true truncation levels are $\ell \in \{2, 3, 4\}$ and the dimensions of the randomly generated correlation matrices are $d \in \{5, 7, 10, 15\}$.

Section 5, where we choose $t_{bST} = t_{\mathrm{N}} = 20$, so that $t = 40$ candidate solutions are considered for each extant solution in each step. The selection factor is chosen as $\lambda = 0.2$ and the elitism factor as $\eta = 0.05$. The best obtained NFIs δ_ℓ (4.3) of 100 repetitions are then compared to the corresponding greedy solutions in Figure 2. CFIs γ_ℓ (4.6) are not shown here, since the results are very similar, even for different sample sizes $n \in \{250, 500, 1000, 2000\}$. An NFI or a CFI of 1 corresponds to a perfect fit. The case $\ell = 1$ does not need to be considered here, since, in this case, the greedy solution is optimal, so that the NFI and the CFI are always 1.

In all dimensions, our newly proposed selection approach clearly improves over the greedy algorithm. While in smaller dimensions ($d \in \{5, 7\}$) the optimal solution is often found by the new approach (NFI/CFI of 1), the average NFI and CFI drop below 0.90 for $d = 15$. In order to obtain better solutions here, we would need to take into account more candidate solutions in each step, which is however computationally more demanding. Nonetheless, the new approach still yields much better results than the greedy approach, so that we can conclude that it works quite well in general.

For further insights, we looked more closely at cases where the greedy approach did well or poorly. In cases where it did poorly, there were large absolute partial correlation coefficients after the first 'optimal' tree (the MST). This explains why the greedy approach can be worse than a more global selection criterion.

The second question of interest is now the choice of α in the truncation rules (4.4) and (4.7). As noted in Section 4, choices of $\alpha = 0.05$ (or even larger) have been proposed in the literature. To investigate this, we extend the previous simulation study as follows. We again randomly generate d-dimensional correlation matrices according to ℓ-truncated partial correlation R-vines. This time, we however do not automatically truncate at level ℓ but let the truncation rules decide about the truncation level ℓ^*. As observed above, the selection approach does not always find the true model. We therefore seek to find the truncation level such that the selected partial correlation R-vine is sufficiently close in terms of the NFI or the CFI. Results for the CFI truncation rule (4.7) and different choices of α are shown in Figure 3, where the control parameters of our new selection algorithm are chosen as above. When the NFI truncation rule (4.4) is used, results are quite similar and therefore not shown here. The main difference is that the mean selected truncation levels according to the NFI truncation level increase more strongly with the dimension.

Clearly, the larger α is chosen, the earlier a partial correlation R-vine is truncated. While the difference in the selected truncation level is small for smaller dimensions, it increases with the dimension: The larger the dimension d is, the more difficult it is to identify a well-fitting partial correlation R-vine with small truncation level, since we chose the number of candidate

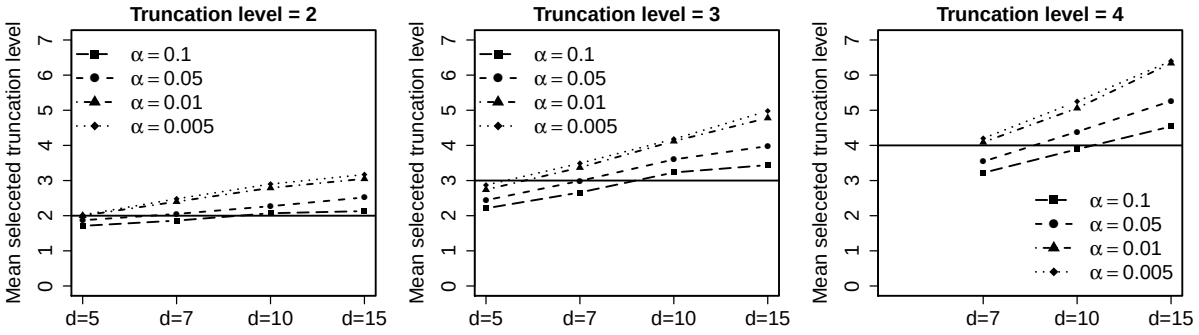

Figure 3: Mean selected truncation levels according to the CFI truncation rule (4.7) with $\alpha \in \{0.005, 0.01, 0.05, 0.1\}$ and our new selection approach. The true truncation levels are $\ell \in \{2, 3, 4\}$ and the dimensions of the randomly generated correlation matrices are $d \in \{5, 7, 10, 15\}$. The sample size is $n = 1000$.

solutions to be independent of d in order to limit the computational complexity. The results further show that, when the data comes from a truncated partial correlation R-vine, our new selection and truncation approach is able to identify models that fit the data well and have only one or two additional vine trees. Based on the discussion in Section 3 of expecting different truncated R-vine copulas to equally well approximate a multivariate copula, this trade-off of getting one or two additional trees without a huge computational time is acceptable. While a choice of $\alpha = 0.05$ may be reasonable in certain cases, we prefer to set $\alpha = 0.01$, which yields a considerable improvement in the fit using, on average, about one additional vine tree; this is also in line with the analysis at the end of Section 5, where $\alpha = 0.01$ is suggested, too.

In addition, we perform a sensitivity analysis with respect to the dependence structure. The derivation of our new selection procedure suggests that it will work best when the dependence is multivariate Gaussian. To investigate the effect of non-Gaussian dependence on the selection approach, we also simulate from R-vine copulas having

(i) Gumbel pair copulas with random non-zero upper tail dependence in all trees, or

(ii) BB1 pair copulas with random non-zero lower and upper tail dependence in the first tree and Frank pair copulas in the remaining trees.

Gumbel pair copulas on all edges lead to a copula model with upper tail dependence and skewness to the upper tail for all margins; BB1/Frank pair copulas lead to a copula model with asymmetric lower and upper tail dependence for all bivariate margins (bivariate margins not from tree T_1 have tail dependence but the marginal copula is not BB1). These types of tail behavior are expected in typical applications where multivariate copulas are used.

The truncation rules are then applied to the empirical correlation matrix of the normal scores based on a sample from these two R-vine copulas. Figure 4 shows the results according to the CFI truncation rule (4.7) and different choices of α; control parameters are chosen as before. Compared to results based on a sample from a partial correlation R-vine, truncation levels are selected to be smaller, especially for $\alpha \in \{0.1, 0.05\}$. However, the choice of $\alpha \in \{0.01, 0.005\}$ yields results not much different to the Gaussian case. We therefore conclude that $\alpha = 0.01$ is reasonable even when the dependence is non-Gaussian. As before, this also holds for the NFI.

Here, we do not report results according to the greedy selection approach. As expected, it yields significantly higher truncation levels than our new approach: In this study, truncation levels were up to 50% higher. This means that with our new selection and truncation approach we are able to identify models that are adequate for the data, but require fewer vine trees and

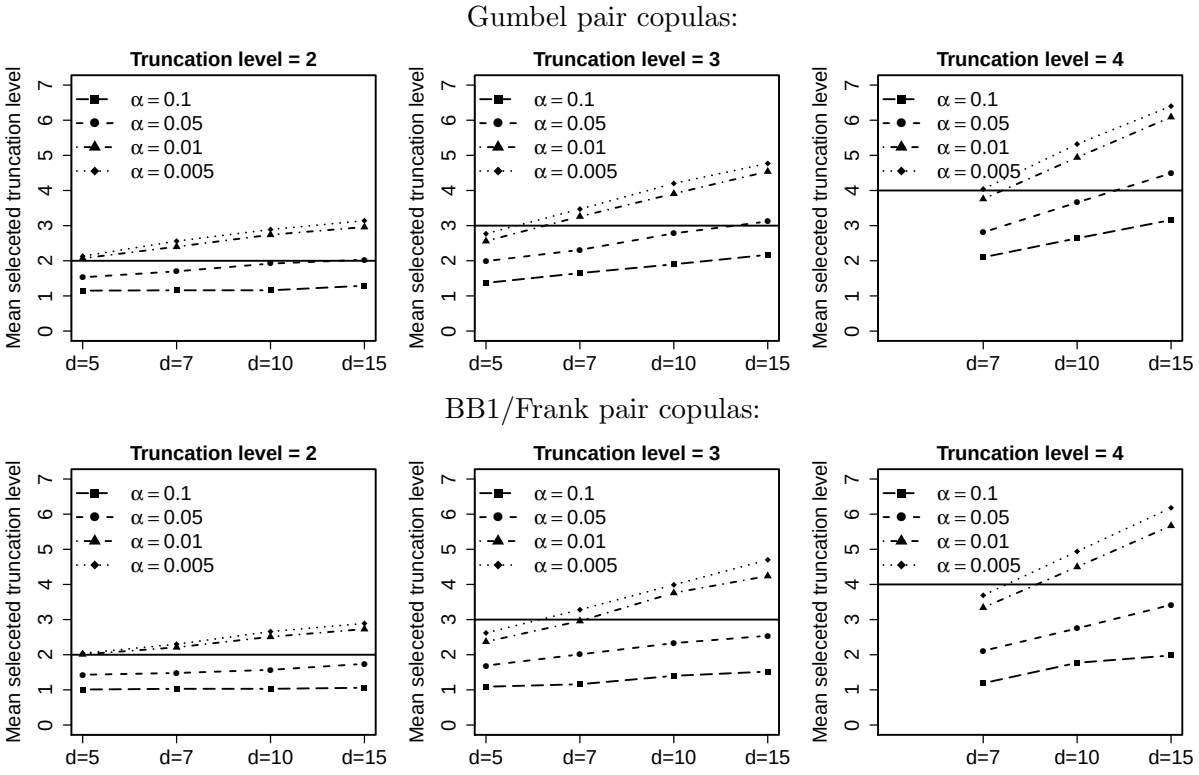

Figure 4: Mean selected truncation levels according to the CFI truncation rule (4.7) with $\alpha \in \{0.005, 0.01, 0.05, 0.1\}$ and our new selection approach for samples of size $n = 1000$ from R-vine copulas with Gumbel pair copulas (first row) and with BB1 and Frank pair copulas (second row). The true truncation levels are $\ell \in \{2, 3, 4\}$ and the dimensions of the randomly generated correlation matrices are $d \in \{5, 7, 10, 15\}$.

are hence are more parsimonious. As before, the reported results do not change much for other sample sizes.

For comparison, we also did some simulations of truncation levels with the RNI and the IFI. For $\alpha = 0.01$, we did not find any example where the RNI led to a different truncation level than the CFI. The IFI on some rare occasions led to smaller truncation levels. Overall, the differences in the truncation level due to the chosen fit index is however minor compared to those in comparing the greedy and the non-greedy selection methods.

7 Applications

To evaluate our newly proposed selection strategy for truncated R-vine copulas, we now consider two data sets from finance, which is a common field where multivariate copulas are used (see Cherubini et al. (2004) and McNeil et al. (2005)). Due to the potentially complex relationships among financial assets, the flexible R-vine copulas are attractive models to model returns of portfolios (see, amongst others, Berg and Aas (2009), Nikoloulopoulos et al. (2012) and Brechmann and Czado (2013)). A major issue of such R-vine copulas for financial data is however typically the model complexity, especially when large numbers of assets are modeled. The truncation of R-vine copulas may therefore be useful to select more parsimonious models.

First, we revisit the data set investigated by Brechmann et al. (2012). It consists of return time series of $d = 19$ financial assets and indices, which constitute the market portfolio of a major Norwegian financial institution. The time series are observed from March 2003 to March

2008 with $n = 1107$ daily observations. Second, we consider the stock returns of the $d = 15$ largest German companies represented in the most important German stock index DAX. The observed time period is January 2005 to July 2011 with $n = 1158$ daily observations.

Since the data sets show the typical features of financial time series, we apply GARCH filters to remove the serial dependence. This approach is commonly referred to copula-GARCH modeling (see, e.g., Liu and Luger (2009)). The copula models are then based on the independent and identically distributed residuals after filtering. If this modeling strategy is carried out in two separate steps—as we do it here—, the estimation approach is called inference functions for margins (IFM; see Joe and Xu (1996)). For the marginal return time series of the Norwegian market portfolio, we use the same models as in Brechmann et al. (2012). For the DAX return time series, GARCH(1,1)-models with Student's t innovations turned out to appropriately fit the data.

In the next step, the residuals are transformed to normal scores using the innovations distributions and the inverse standard normal distribution function, as explained in Section 3. To select the tree structure and the truncation level of an appropriate truncated R-vine copula for the data, we then use the new approach outlined in Section 5. That is, we select a truncated partial correlation R-vine for the empirical correlation matrix of the normal scores. Finally, pair copulas for the identified tree structure are selected according to the AIC from the following list: Gaussian, Student's, Clayton, Gumbel, Frank and BB1 as well as appropriate survival versions of the tail asymmetric copulas. If a (conditional) pair of variables can be modeled as independent (according to an independence test), we select the independence copula. This additionally reduces the model complexity.

As in the simulation study in Section 6, we set $t_{bST} = t_{\mathrm{N}} = 20$ and choose the selection factor as $\lambda = 0.2$ and the elitism factor as $\eta = 0.05$. Further, we choose $\alpha = 0.01$ for the NFI and CFI truncation rules (4.4) and (4.7), respectively. For comparison, we also truncate according to the IFI (4.8). Using this setting of control parameters, we ran our newly proposed selection algorithm ten times for each data set and for each fit index in order to display the variability in the results due to the (partly) random selection of the candidate models. The fitted models are compared to a non-truncated R-vine copula fitted using the selection algorithm by Dißmann et al. (2013), which uses pairwise empirical Kendall's τ values as edge weights for a greedy tree selection. The models are also compared to an R-vine copula truncated according to the method by Brechmann et al. (2012), which is based on the Vuong test (here: using the test statistic without correction for the number of parameters and the 5% significance level).

Table 1 shows the results for the Norwegian market portfolio. First of all, they show that the CFI truncation rule identifies smaller truncation levels than the NFI. This implies larger bounds on the partial correlation coefficients. Especially the worst case bounds on the maximum partial correlation coefficients are rather high. The bounds on the average partial correlation coefficients are however so small that no obvious lack of fit can be detected. Compared to the greedy algorithm, the new selection algorithm consistently identifies smaller truncation levels. In fact, the identified truncation levels are consistent over the ten runs of the selection algorithm. Nevertheless, the underlying vine trees are chosen differently, as reflected in different model fits.

In addition, Table 1 reports the log likelihoods of the partial correlation vines from the first step of our algorithm in finding a near-optimal truncated vine structure. In other words, the log likelihoods of vine copulas with only Gaussian pair copulas are shown, which can easily be calculated using the density expression given in Equation (2.1); note that these are pseudo log likelihoods according to the popular method of inference functions for margins (IFM) by Joe and Xu (1996), since the margins are estimated preliminarily. Next to this column are the log likelihoods when Gaussian copulas in the truncated R-vine are replaced by other pair copulas in our list. The log likelihood improvement is an indication of different tail behavior than Gaussian. Included in the pair copulas are about 75% with tail dependence. Bivariate

Method	Trunc. level	Gauss log lik.	Log lik.	No. par.	BIC	(4.10) max.	(4.11) av.	Abs. τ max.	Abs. τ av.
Dißmann et al. (2013)			6407	91	-12170			0.043	0.013
Brechmann et al. (2012)	6		6290	77	-12036			0.159	0.018
Best 1-truncated (MST)	1	5337	5627	31	-11034	0.843	0.090	0.210	0.045
Greedy selection (NFI)	7	6024	6370	83	-12153	0.286	0.036	0.063	0.016
New selection (NFI)	6	6024	6379	84	-12162	0.287	0.033	0.062	0.015
	6	6023	6379	84	-12163	0.287	0.033	0.067	0.016
	6	6028	6387	87	-12157	0.275	0.032	0.064	0.014
	6	6024	6379	84	-12162	0.287	0.033	0.067	0.017
	6	6028	6372	88	-12120	0.275	0.032	0.052	0.012
	6	6024	6379	84	-12162	0.287	0.033	0.058	0.016
	6	6025	6366	89	-12103	0.284	0.033	0.055	0.013
	6	6024	6379	84	-12162	0.287	0.033	0.072	0.015
	6	6024	6360	85	-12117	0.286	0.033	0.065	0.015
	6	6024	6366	82	-12151	0.285	0.033	0.056	0.016
Greedy selection (CFI)	6	5972	6323	78	-12094	0.398	0.047	0.070	0.017
New selection (CFI)	4	5953	6306	79	-12052	0.431	0.044	0.062	0.015
	4	5957	6304	77	-12064	0.424	0.043	0.085	0.016
	4	5956	6304	78	-12056	0.425	0.044	0.072	0.014
	4	5949	6310	74	-12097	0.437	0.045	0.077	0.017
	4	5949	6315	73	-12113	0.437	0.045	0.058	0.017
	4	5944	6293	69	-12097	0.444	0.046	0.077	0.019
	4	5953	6306	79	-12052	0.431	0.044	0.059	0.015
	4	5948	6310	72	-12110	0.438	0.045	0.061	0.015
	4	5948	6310	72	-12110	0.438	0.045	0.056	0.016
	4	5938	6305	73	-12094	0.454	0.047	0.079	0.017
Greedy selection (IFI)	4	5886	6242	69	-11995			0.130	0.019
New selection (IFI)	3	5848	6204	61	-11977			0.155	0.021
	3	5850	6204	61	-11976			0.155	0.021
	3	5848	6200	60	-11974			0.159	0.019
	3	5848	6204	61	-11977			0.156	0.020
	3	5850	6204	61	-11976			0.163	0.022
	3	5850	6208	62	-11978			0.157	0.021
	3	5854	6219	62	-11999			0.156	0.020
	3	5848	6200	60	-11974			0.156	0.020
	3	5850	6204	61	-11976			0.153	0.019
	3	5854	6214	61	-11996			0.155	0.021

Table 1: Truncation results for the Norwegian market portfolio. The columns headed by (4.10) and (4.11) show the upper bounds on the maximum and the average absolute partial correlation coefficients in the vine trees after truncation.

copulas with tail dependence have densities that quickly increase to infinity in the corners of the unit square. Hence there can be a big difference in the model log likelihood when the data exhibits tail dependence and tail dependent pair copulas are used in tree T_1. The main result in Joe et al. (2010) implies that if all of the edges of the subtree connecting variables j and k in T_1 are fitted with upper (lower) tail dependent pair copulas, then the bivariate (j, k) margin of the vine copula has upper (lower) tail dependence.

While the approach by Brechmann et al. (2012) also yields a truncation level of 6, the 6-truncated R-vine copulas according to the NFI truncation rule improve over it due to the better exploration of the search space of truncated R-vine copulas. Because of that even the 4-truncated R-vine copulas according to the CFI truncation rule are superior (in terms of the BIC) to the model selected according to the Vuong test. For instance, the selected 4-truncated vine tree structure in the second run of the algorithm is as follows: The first vine tree is selected as the second 1-neighbor of the MST, the second vine tree as the tenth-best spanning tree, and the third vine tree as the second-best spanning tree. The vine tree at the truncation level is always selected best as the actual MST. This is clearly different from a simple greedy algorithm, which yields a truncation level of 6.

Truncation according to the IFI results in yet smaller truncation levels than the CFI truncation rule, but also much weaker model fits in terms of the BIC. Hence, it appears that the IFI yields overly parsimonious models due to rather early truncation. In contrast, the newly proposed NFI and CFI truncation rules work well for this data set and, combined with the better model selection approach, are able to identify parsimonious but still flexible truncated R-vine copulas.

The results for the DAX returns, as shown in Table 2, are similar. The CFI truncation rule leads to earlier truncation than the NFI truncation rule and the truncation level according to greedy selection can be reduced through our new selection procedure. For this data set, the truncation level according to the Vuong test is however much larger than according to our new truncation rules, which select more parsimonious models with good BIC values. For instance, the 6-truncated R-vine copula selected in the sixth run of the algorithm has a smaller BIC than the 9-truncated R-vine copula according to the approach by Brechmann et al. (2012). This indicates that the truncation at level 9 is too conservative, since appropriately selected 5- or 6-truncated R-vine copula also fit the data well. Even the 4-truncated R-vine copulas according to the CFI truncation rule provide a good fit. Truncation according to the IFI here only makes a difference for the simple greedy selection, where a smaller truncation level than according to the CFI greedy selection is detected; results according to the new selection procedure are not different. Yet, it is interesting to note that, although the truncation levels according to IFI greedy and non-greedy selection are the same, the new selection procedure is able to select a more appropriate vine tree structure, which yields a better overall model fit.

Tables 1 and 2 also show how well the truncated R-vine copulas match the data in terms of pairwise model-based versus empirical Kendall's τ values, where the model-based Kendall's τ values were computed based on simulated samples of size 10 000. For a sample of size n from a bivariate copula C with Kendall's τ value τ_C, let $\widehat{\tau}$ be the sample version of Kendall's τ. In Hoeffding (1948), it is proved that

$$n \operatorname{Var}(\widehat{\tau}) \xrightarrow[n\to\infty]{} 16 \int \left(C + \overline{C}\right)^2 dC - 4(\tau_C + 1)^2, \tag{7.1}$$

where $\overline{C}$ is the survival function of C. The right-hand side of (7.1) can be evaluated for some commonly used parametric bivariate copula families: It is $4/9 = 0.44$ for the independence copula, 0 for the comonotonicity copula and tends to be slowly decreasing when the parameter of the copula is such that the corresponding Kendall's τ increases from 0.1 to 1; see Section 2.12 of Joe (2014). For some copula families, (7.1) increases a little for small positive dependence

Method	Trunc. level	Gauss log lik.	Log lik.	No. par.	BIC	(4.10) max.	(4.11) av.	Abs. τ max.	Abs. τ av.
Dißmann et al. (2013)			4980	96	-9283			0.053	0.014
Brechmann et al. (2012)	9		4962	93	-9268			0.073	0.014
Best 1-truncated (MST)	1	3709	4041	27	-7892	0.874	0.126	0.260	0.144
Greedy selection (NFI)	8	4515	4954	91	-9266	0.228	0.050	0.074	0.017
New selection (NFI)	5	4501	4926	85	-9252	0.274	0.042	0.070	0.014
	5	4501	4926	85	-9252	0.274	0.042	0.064	0.015
	6	4522	4951	90	-9267	0.205	0.034	0.069	0.014
	5	4501	4924	85	-9248	0.274	0.042	0.075	0.016
	6	4522	4965	91	-9289	0.202	0.034	0.062	0.012
	5	4501	4924	85	-9248	0.274	0.042	0.071	0.015
	5	4501	4926	85	-9252	0.274	0.042	0.085	0.018
	5	4501	4926	85	-9252	0.274	0.042	0.074	0.017
	5	4501	4926	85	-9252	0.274	0.042	0.080	0.020
	5	4501	4924	85	-9248	0.274	0.042	0.065	0.017
Greedy selection (CFI)	5	4463	4891	80	-9217	0.367	0.057	0.139	0.024
New selection (CFI)	4	4452	4865	72	-9221	0.387	0.054	0.074	0.024
	4	4453	4878	72	-9249	0.386	0.054	0.083	0.026
	4	4459	4873	76	-9210	0.374	0.052	0.091	0.024
	4	4459	4873	76	-9210	0.374	0.052	0.089	0.024
	4	4452	4875	71	-9248	0.387	0.054	0.079	0.028
	4	4459	4872	76	-9208	0.374	0.052	0.087	0.021
	4	4459	4873	76	-9210	0.374	0.052	0.087	0.023
	4	4459	4873	76	-9210	0.374	0.052	0.095	0.027
	4	4459	4873	76	-9210	0.374	0.052	0.088	0.022
	4	4452	4889	74	-9256	0.388	0.054	0.082	0.023
Greedy selection (IFI)	4	4417	4841	73	-9167			0.124	0.026
New selection (IFI)	4	4453	4843	71	-9185			0.084	0.021
	4	4452	4881	71	-9262			0.095	0.032
	4	4453	4878	72	-9249			0.097	0.026
	4	4454	4862	74	-9202			0.071	0.025
	4	4452	4865	72	-9221			0.069	0.021
	4	4452	4871	71	-9241			0.079	0.024
	4	4459	4873	76	-9210			0.079	0.024
	4	4452	4887	74	-9252			0.085	0.025
	4	4459	4872	76	-9208			0.088	0.023
	4	4459	4872	76	-9208			0.093	0.023

Table 2: Truncation results for the DAX returns. The columns headed by (4.10) and (4.11) show the upper bounds on the maximum and the average absolute partial correlation coefficients in the vine trees after truncation.

before decreasing. When the Kendall's τ value of the copula is in the range 0.4 to 0.5, the typical range of (7.1) is 0.20 to 0.35. For the two examples in this section, the sample sizes are $n = 1107$ and $n = 1158$. For these two values of n, $\sqrt{0.44/n}$, $\sqrt{0.35/n}$ and $\sqrt{0.20/n}$ round to 0.020, 0.018 and 0.013 respectively. Hence the summaries of the maximum absolute difference and average absolute difference of model-based and empirical Kendall's τ values indicate that the fitted truncated vine copula models with the new selection algorithms are capturing the dependence structure within sampling variability. Comparing these results to the case of the best 1-truncated R-vine copula, which is given through the MST, shows that truncation at level 1 is clearly insufficient, since it fails to appropriately model the dependence structure of the data.

8 Conclusion

We propose a new approach to the truncation of R-vine copulas. Since it separates the challenging steps of the vine tree selection and the selection and the estimation of the pair copulas, it is possible to better explore the search space of truncated R-vine copulas using an appropriate selection algorithm. By truncating according to fit indices, it is possible to directly control the goodness of a fitted truncated model relative to the non-truncated case. Furthermore, it is possible to calculate bounds on what is ignored after truncation.

In a simulation study, we show that the newly proposed selection approach improves over the simple greedy algorithm, which has so far been used in the literature to limit the computational effort. These findings are also underlined by the selection results for two data sets of financial returns.

In our approach, the truncation level is determined based on the general dependence of the normal scores of the variables of interest, for which we can benefit from optimality results on the modeling of correlation matrices in terms of partial correlation R-vines. An open question is how an appropriate truncation rule can be defined that focuses on other data characteristics such as the tail dependence of the variables. In this case, it is no longer clear how the goodness-of-fit can be measured adequately (and accurately) such that a truncation level can be detected using a reasonable computational effort.

Acknowledgment

The first author gratefully acknowledges support from the TUM Graduate School's International School of Applied Mathematics as well as from Allianz Deutschland AG. The numerical computations were performed on a Linux cluster supported by DFG grant INST 95/919-1 FUGG. The second author is supported by an NSERC Canada Discovery Grant. We are grateful to the referees for the constructive comments.

References

Acar, E. F., C. Genest, and J. Nešlehová (2012). Beyond simplified pair-copula constructions. *Journal of Multivariate Analysis 110*, 74–90.

Bedford, T. and R. M. Cooke (2001). Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial intelligence 32*(1–4), 245–268.

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

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin 107*, 238–246.

Bentler, P. M. and D. G. Bonett (1980). Significance tests and goodness of fit in the analysis of covariances structures. *Psychological Bulletin 88*, 588–606.

Berg, D. and K. Aas (2009). Models for construction of higher-dimensional dependence: A comparison study. *European Journal of Finance 15*(7–8), 639–659.

Blickle, T. and L. Thiele (1996). A comparison of selection schemes used in evolutionary algorithms. *Evolutionary Computation 4*(4), 361–394.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. New York: Wiley.

Brechmann, E. C. and C. Czado (2013). Risk management with high-dimensional vine copulas: An analysis of the Euro Stoxx 50. *Statistics & Risk Modeling 30*(4), 307–342.

Brechmann, E. C., C. Czado, and K. Aas (2012). Truncated regular vines in high dimensions with applications to financial data. *Canadian Journal of Statistics 40*(1), 68–85.

Brechmann, E. C. and H. Joe (2014). Parsimonious parameterization of correlation matrices using truncated vines and factor analysis. *Computational Statistics & Data Analysis 77*, 233–251.

Cherubini, U., E. Luciano, and W. Vecchiato (2004). *Copula Methods in Finance*. Chichester: John Wiley & Sons.

Czado, C., E. C. Brechmann, and L. Gruber (2013). Selection of vine copulas. In P. Jaworski, F. Durante, and W. Härdle (Eds.), *Copulae in Mathematical and Quantitative Finance*. Berlin: Springer.

Dißmann, J., E. C. Brechmann, C. Czado, and D. Kurowicka (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics & Data Analysis 59*(1), 52–69.

Gabow, H. N. (1977). Two algorithms for generating weighted spanning trees in order. *SIAM Journal on Computing 6*(1), 139–150.

Goldberg, D. E. and K. Deb (1991). A comparative analysis of selection schemes used in genetic algorithms. In *Foundations of Genetic Algorithms*, San Francisco, pp. 69–93. Morgan Kaufmann.

Hobæk Haff, I., K. Aas, and A. Frigessi (2010). On the simplified pair-copula construction - simply useful or too simplistic? *Journal of Multivariate Analysis 101*(5), 1296–1310.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Annals of Mathematical Statistics 19*, 293–325.

Hu, L. and P. M. Bentler (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling 6*(1), 1–55.

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 (Eds.), *Distributions with fixed marginals and related topics*. Hayward: Institute of Mathematical Statistics.

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

Joe, H. (2014). *Dependence Modeling with Copulas*. Boca Raton, FL: Chapman & Hall/CRC.

Joe, H., H. Li, and A. K. Nikoloulopoulos (2010). Tail dependence functions and vine copulas. *Journal of Multivariate Analysis 101*(1), 252–270.

Joe, H. and J. Xu (1996). The estimation method of inference functions for margins for multivariate models. Technical Report 166, Department of Statistics, University of British Columbia.

Kurowicka, D. and R. M. Cooke (2006). Completion problem with partial correlation vines. *Linear Algebra and its Applications 418*, 188–200.

Liu, Y. and R. Luger (2009). Efficient estimation of copula-GARCH models. *Computational Statistics & Data Analysis 53*(6), 2284–2297.

McNeil, A. J., R. Frey, and P. Embrechts (2005). *Quantitative Risk Management: Concepts Techniques and Tools*. Princeton: Princeton University Press.

Morales-Nápoles, O. (2011). Counting vines. In D. Kurowicka and H. Joe (Eds.), *Dependence Modeling: Vine Copula Handbook*. Singapore: World Scientific Publishing Co.

Mulaik, S. A. (2009). *Linear Causal Modeling with Structural Equations*. Boca Raton: Chapman & Hall/CRC.

Nelsen, R. B. (2006). *An Introduction to Copulas* (2nd ed.). Berlin: Springer.

Nikoloulopoulos, A. K., H. Joe, and H. Li (2012). Vine copulas with asymmetric tail dependence and applications to financial return data. *Computational Statistics & Data Analysis 56*(11), 3659–3673.

Prim, R. C. (1957). Shortest connection networks and some generalizations. *Bell System Technical Journal 36*, 1389–1401.

Salvadori, G., C. D. Michele, N. T. Kottegoda, and R. Rosso (2007). *Extremes in Nature: An Approach Using Copulas*. Berlin: Springer.

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

Stöber, J., H. Joe, and C. Czado (2013). Simplified pair copula constructions - Limits and extensions. *Journal of Multivariate Analysis 119*, 101–118.