Technische Universität München

DEPARTMENT OF MATHEMATICS

Regular vine copula based quantile regression

Master’s Thesis

by

Jakob Samuel Herrmann

Supervisor: Prof. Claudia Czado, PhD
Advisor: Prof. Claudia Czado, PhD
Nicole Barthel
Submission date: October 29, 2018
I hereby declare that this thesis is my own work and that no other sources have been used except those clearly indicated and referenced.

Munich, October 29, 2018
Acknowledgments

Firstly, I would like to express my sincere gratitude to my advisors Prof. Dr. Czado and Nicole Barthel for their continuous support of my master’s thesis, for their patience, motivation, and immense knowledge. Their guidance and feedback helped me in all the time of writing this thesis. Thinking of all the inspiring and productive meetings, I could not have imagined having better advisors for my thesis.
Abstract

As the prediction of conditional quantiles plays an important role in various fields of econometrics (e.g. value at risk in finance), quantile regression has steadily gained importance in statistical modeling. Since the introduction of linear quantile regression ([Koenker and Bassett Jr. 1978]), which models the linear conditional quantile function using regression coefficients, multiple methods have been developed aiming to improve the model’s shortfalls, such as the linearity assumption and quantile crossing. D-vine copula based quantile regression, introduced by [Kraus and Czado 2017], was the first quantile regression method that estimates the conditional quantile function using vine copulas. As the method’s name indicates, D-vine copula based quantile regression is restricted to the class of D-vine copulas.

The goal of this thesis is to develop a regular vine (R-vine) copula based quantile regression method starting from the ideas approaches proposed by [Kraus and Czado 2017]. That means we want to sequentially fit an optimal R-vine copula to given copula data and estimate the conditional quantile function of a response given a set of covariates using the conditional distribution described by the estimated R-vine copula. Intuitively one might think of an one to one extension of the D-vine quantile regression approach to the more general class of regular vine copulas. However, since already for small numbers of covariates the class of R-vine copulas exceeds the class of D-vine copulas enormously, it would require too much computational effort to determine the optimal R-vines by conditional likelihood arguments, as suggested in [Kraus and Czado 2017]. Therefore we propose a partial correlation based approach to select the sequence on how the potential covariates enter the R-vine copula and use additional criteria in order to decrease the number of R-vine tree sequence candidates.

In order to enable the reader to understand the terms and notations used in our methods, we present common, as well as method-specific concepts related to R-vine copulas in Part I. After introducing and illustrating our methods in Part II, we compare our approach to the benchmarks D-vine copula based quantile regression and linear quantile regression within a simulation study and apply our method to the abalone data set aiming to predict the weight of female abalones in Part III.
Zusammenfassung


Ziel dieser Masterarbeit ist es nun, basierend auf den Ansätzen von Kraus and Czado [2017] eine Quantils-Regressions-Methode zu entwickeln, welche die Quantils-Funktion durch reguläre Vine (R-vine) Kopulas modelliert. Im Detail bedeutet das, dass, gegeben einem Datensatz auf der Kopula-Ebene, optimale R-vine Kopulas sequentiell geschätzt werden und anschließend die implizierte bedingte Verteilung der zu erklärenden Variablen genutzt wird um die zugehörige Quantils-Funktion herzuleiten. Da die D-vine Kopula basierte Methode die optimale Kopula sequentiell durch Maximierung des bedingten Likelihoods wählt, die Menge an möglichen R-vine Kopulas jedoch bereits für eine geringe Zahl an Kovariablen um ein Vielfaches größer als die Menge der validen D-vine Kopulas ist, wäre eine einfache Ausweitung der D-vine Methodik auf die Klasse der regulären Vine Kopulas mit einem zu hohen Rechenaufwand verbunden. Deshalb wählt die in der Arbeit vorgestellte Methode die Reihenfolge der Kovariaten an Hand von Pearson - und partiellen Korrelationen und beschränkt die Menge der möglichen R-vine Strukturen durch zusätzliche Kriterien.

Contents

I  Theoretical background 5

1 Bivariate copula classes 6
   1.1 Multivariate distributions and copulas 6
   1.2 Dependence measures . 10
   1.3 Bivariate parametric copulas 15
   1.4 Relationship between Kendall's τ and copula parameters 18
   1.5 Variable scales 20
   1.6 Estimation of bivariate parametric copulas 22

2 Regular vine copulas and distributions 25
   2.1 Graph theory 25
   2.2 Regular vine tree sequences 27
   2.3 R-vine distributions and R-vine copulas 34

3 Quantile Regression 40
   3.1 Conditional quantile- and check loss function 40
   3.2 Linear quantile regression 42

II Regular vine based quantile regression 45

4 Regular vine copulas for quantile regression 46

5 Selection of an R-vine copula for quantile regression 52
   5.1 Computational effort in the selection of the global cℓℓ- optimal R-vine copula 52
   5.2 D-vine quantile regression 55
   5.3 Computational complexity in the extension of D-vine quantile regression to the set of all R-vine copulas 56
   5.4 Selection of covariates by partial correlations 58
   5.5 Forward selection of the R-vine tree sequence 64

6 Algorithm - forward selection of an R-vine copula 75
   6.1 Parameters and data input 76
   6.2 Pseudo code 76
   6.3 Illustration 78
## III Simulation Study and data application

### 7 Simulation study

#### 7.1 Setup

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1.1 Simulation of the initial training and evaluation data set</td>
<td>99</td>
</tr>
<tr>
<td>7.1.2 Scenarios</td>
<td>101</td>
</tr>
<tr>
<td>7.1.3 Competitor methods</td>
<td>101</td>
</tr>
</tbody>
</table>

#### 7.2 In-sample conditional log likelihood comparison

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.2.1 Setup and results</td>
<td>103</td>
</tr>
<tr>
<td>7.2.2 Effect of the selection of covariates</td>
<td>106</td>
</tr>
</tbody>
</table>

#### 7.3 Out-of-sample evaluation

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.3.1 Considered quantities</td>
<td>111</td>
</tr>
<tr>
<td>7.3.2 Results in scenarios 1, 5, 7 and 8</td>
<td>113</td>
</tr>
<tr>
<td>7.3.3 Summary</td>
<td>123</td>
</tr>
</tbody>
</table>

#### 7.4 Required computational times

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>125</td>
</tr>
</tbody>
</table>

### 8 Application to the abalone data set

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1 Data preparation</td>
<td>127</td>
</tr>
<tr>
<td>8.2 Quantile regression of Whole</td>
<td>130</td>
</tr>
<tr>
<td>8.2.1 In-sample $\ell$-evaluation of all valid R-vine structures</td>
<td>131</td>
</tr>
<tr>
<td>8.2.2 Out-of-sample quantile prediction</td>
<td>135</td>
</tr>
</tbody>
</table>

## IV Conclusion

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>139</td>
</tr>
</tbody>
</table>

## V Appendix

### A.1 Simulation Study

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.1.1 Seeds</td>
<td>150</td>
</tr>
<tr>
<td>A.1.2 Tree plots of minimal - and maximal - $\ell$-R-vine copulas in scenar-</td>
<td>151</td>
</tr>
<tr>
<td>ios 1 to 4</td>
<td></td>
</tr>
<tr>
<td>A.1.3 Out-of-sample results</td>
<td>160</td>
</tr>
</tbody>
</table>
Part I

Theoretical background
Chapter 1

Bivariate copula classes

In the following chapter, we introduce some basic theory on bivariate copula classes.

1.1 Multivariate distributions and copulas

As copulas belong to the class of multivariate distributions, we start with discussing the main concepts and properties of multivariate distributions. They describe the behavior of multiple random variables among themselves. If a random vector \(X := (X_1, \ldots, X_d)^T\) follows a (joint) distribution \(F\), we denote it by

\[X \sim F.\]

Joint distributions induce marginal and conditional distributions. While marginals characterize a random variable's behavior on its own, the conditional distribution includes effects of a set of conditioning variables.

**Definition 1.1.** (Marginal, joint and conditional distributions)

Let \(X = (X_1, \ldots, X_d)^T\) be a \(d\)-dimensional random vector. Then we use the following notation:

- marginal distribution (density) function of \(X_j\): \(F_j(x_j) (f_j(x_j))\), \(j = 1, \ldots, d\),
- joint distribution (density) function of \(X\): \(F(x) (f(x))\) with \(x = (x_1, \ldots, x_d)^T \in \mathbb{R}^d\),
- conditional distribution (density) function of \(X_j\) given \(X_{j_1}, \ldots, X_{j_s}\):
  \[F_{j|j_1,\ldots,j_s}(x_j|x_{j_1},\ldots,x_{j_s}) (f_{j|j_1,\ldots,j_s}(x_j|x_{j_1},\ldots,x_{j_s}))\] with \(j, j_k \in \{1, \ldots, J\}, j \neq j_k, k = 1, \ldots, s\).

Remember that, according to Bayes Theorem, the conditional density arising from a bivariate (joint) distributions can be calculated by dividing the joint density by the marginal density of the conditioning variable. Thus, if we are given a random vector \((X_1, X_2) \sim F_{1,2}, X_j \sim F_j, j = 1, 2\), the conditional density \(f_{1|2}\) of \(X_1|X_2\) can be expressed in the following way:

\[f_{1|2}(x_1|x_2) = \frac{f_{12}(x_1, x_2)}{f_2(x_2)}. \quad (1.1)\]
Similarly to the univariate case, we can define an empirical distribution function for multivariate distributions for non-parametric estimation.

**Definition 1.2.** (Multivariate, non-parametric empirical distribution function) Let \( x_i = (x_{1i}, \ldots, x_{di}) \in \mathbb{R}^d, i = 1, \ldots, n \) be an i.i.d sample from a d-dimensional random vector \( X \sim F \). Then, the empirical distribution function \( \hat{F} \) is defined by:

\[
\hat{F}(x) := \frac{1}{n+1} \sum_{i=1}^{n} \mathbb{1}_{\{x_{1i} \leq x_1, \ldots, x_{di} \leq x_d\}}, \forall x = (x_1, \ldots, x_d)^T \in \mathbb{R}^d.
\]

In order to be able to analyze the dependence within a random vector \( X = (X_1, \ldots, X_d)^T \) using copula methods, we need to segregate the marginal effects. Thus, we standardize each random variable \( X_j, j = 1, \ldots, d \), using the probability integral transform defined in Definition 1.3.

**Definition 1.3.** (Probability integral transform) Let \( X = (X_1, \ldots, X_d)^T \sim F \) and \( X_j \sim F_j, j = 1, \ldots, d \). Then the transformation

\[
U_j := F_j(X_j)
\]

is called probability integral transform (PIT) of \( X_j \).

**Remark 1.4.** (Distribution of PIT)

1. \( X_j \sim F_j \Rightarrow U_j := F_j(X_j) \sim U(0, 1) \), since:

\[
\mathbb{P}(U_j \leq u) = \mathbb{P}(F_j(X_j) \leq u) = \mathbb{P}(X_j \leq F^{-1}_j(u)) = F_j(F^{-1}_j(u)) = u, \forall u \in [0, 1].
\]

2. If the marginal distribution \( F_j \) is estimated parametrically by \( F_j(\cdot, \hat{\theta}) \) or non-parametrically by \( \hat{F}_j \), (1) holds only approximately.

As already mentioned above, the copula approach separates the dependence structure within a random vector from its marginal distributions. For this the dependence of a marginally standardized multivariate distribution is considered. The corresponding distribution function is called a copula (Definition 1.5). Since the copula density is only defined for absolute continuous distributions, we will assume absolute continuity throughout this thesis.

**Definition 1.5.** (Copula, copula density)

- A d-dimensional copula is a multivariate distribution function

\[
C : [0, 1]^d \rightarrow [0, 1]
\]

(i) uniformly distributed marginals, i.e. \( C_j(u_j) = u_j \) for \( j = 1, \ldots, d \).
If $C$ is absolute continuous, the corresponding **copula density** $c$ is obtained by partial differentiation, i.e. $c(u) := \frac{\partial^d}{\partial u_1 \cdots \partial u_d} C(u)$ for all $u = (u_1, \ldots, u_d)^T \in [0,1]^d$.

The copula arising from an empirical distribution, is defined in **Definition 1.6**

**Definition 1.6. (Empirical copula)**

Let $u_i = (u_{i1}, \ldots, u_{id}) \in [0,1]^d$, $i = 1, \ldots, n$, be an i.i.d sample from a $d$-dimensional random vector $U \sim C$. Then the **empirical copula** $\hat{C}$ is defined by:

$$\hat{C}(u) := \frac{1}{n+1} \sum_{i=1}^n 1\{u_{i1} \leq u_1, \ldots, u_{id} \leq u_d\}, \quad \forall u = (u_1, \ldots, u_d)^T \in [0,1]^d.$$ 

The fundamental representation theorem (Theorem 1.7) for a multivariate distribution in terms of its marginal distributions and the corresponding copula was proved by Sklar [1959]. The proof can be found in Nelsen [2006].

**Theorem 1.7. (Sklar’s theorem)**

Let $X \sim F$ be a $d$-dimensional random vector with joint distribution $F$ and marginal distributions $F_j$ (i.e. $X_j \sim F_j$), $j = 1, \ldots, d$. Then:

(i) $F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d))$,

(ii) $f(x_1, \ldots, x_d) = c(F_1(x_1), \ldots, F_d(x_d)) \cdot \prod_{j=1}^d f_j(x_j)$,

(iii) $C(u_1, \ldots, u_d) = F(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d))$,

(iv) $c(u_1, \ldots, u_d) = f(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)) \cdot \prod_{j=1}^d f_j(F_j^{-1}(u_j))$. 

**Lemma 1.8** uses Theorem 1.7 and assigns its idea of separating the dependence structure from joint distributions to conditional distributions. It provides a formula to express conditional bivariate distributions and densities as a product of a copula (density) and a marginal (density).

**Lemma 1.8. (Conditional distribution (density) functions of bivariate distributions in terms of their copula)**

Let $X = (X_1, X_2)^T \sim F_{1,2}$ and $C_{1,2}$ the corresponding copula. Then, it holds:

(i) $f_{1|2}(x_1|x_2) = c_{1,2}(F_1(x_1), F_2(x_2))f_1(x_1),$

(ii) $F_{1|2}(x_1|x_2) = \frac{\partial}{\partial u_2} C_{1,2}(F_1(x_1), u_2)|_{u_2 = F_2(x_2)} =: \frac{\partial}{\partial F_2(x_2)} C_{1,2}(F_1(x_1), F_2(x_2)).$

**Proof:**
Theorem 1.7

\[ f_{1|2}(x_1|x_2) = \frac{f_{12}(x_1, x_2)}{f_2(x_2)} \]

(i) 

\[ = c_{1,2}(F_1(x_1), F_2(x_2)) f_1(x_1) \]

(ii) Observe that

\[ c_{1,2}(F_1(x_1), F_2(x_2)) f_1(x_1) = \frac{\partial^2 C_{1,2}(u_1, u_2)}{\partial u_1 \partial u_2} \bigg|_{u_1 = F_1(x_1), u_2 = F_2(x_2)} \frac{\partial u_1}{\partial x_1} \]

\[ = \frac{\partial}{\partial u_2} \left( \frac{\partial}{\partial x_1} C_{1,2}(F_1(x_1), u_2) \right) \bigg|_{u_2 = F_2(x_2)} \]

Thus,

\[ F_{1|2}(x_1|x_2) = \int_{-\infty}^{x_1} f_{1|2}(z_1|x_2) dz_1 = \int_{-\infty}^{x_1} \frac{\partial}{\partial u_2} \left( \frac{\partial}{\partial z_1} C_{1,2}(F_1(z_1), u_2) \right) \bigg|_{u_2 = F_2(x_2)} dz_1 \]

\[ = \frac{\partial}{\partial u_2} \left( \int_{-\infty}^{x_1} \frac{\partial}{\partial z_1} C_{1,2}(F_1(z_1), u_2) dz_1 \right) \bigg|_{u_2 = F_2(x_2)} \]

\[ = \frac{\partial}{\partial u_2} C_{1,2}(F_1(x_1), u_2) \bigg|_{u_2 = F_2(x_2)} \]

□

Applying Lemma 1.8 to a bivariate copula distribution \( C_{1,2} \) leads us to the class of conditional copulas, which were denoted as **h-functions** by Aas et al. [2009].

**Corollary 1.9.** (Conditional copula and conditional distribution in the bivariate case)

In the framework of Lemma 1.8 also the following holds:

(i) 

\[ C_{1|2}(u_1|u_2) = \frac{\partial}{\partial u_2} C_{1,2}(u_1, u_2) \text{ for all } u_1 \in [0, 1], \]

(ii) 

\[ F_{1|2}(x_1|x_2) = C_{1|2}(F_1(x_1), F_2(x_2)). \]

**Definition 1.10.** (h-functions of bivariate copulas)

For all \((u_1, u_2)^T \in [0, 1]^2\) and a bivariate copula \( C_{1,2} \), the corresponding h-functions are defined as

- \[ h_{1|2}(u_1|u_2) = \frac{\partial}{\partial u_2} C_{1,2}(u_1, u_2) \]
- \[ h_{2|1}(u_2|u_1) = \frac{\partial}{\partial u_1} C_{1,2}(u_1, u_2). \]
1.2 Dependence measures

In this section, the reader finds a list of tools which can be applied to measure dependence between two (continuous) random variables. Since we are usually interested in the dependence structure within data, we introduce an empirical version for each of the measures. The most common measure is the **Pearson correlation** (Definition 1.11), which is often simply referred to as “correlation”. Its empirical version is defined in Definition 1.12.

**Definition 1.11.** (Pearson correlation)
The Pearson correlation between two random variables \(X_1\) and \(X_2\) is defined as
\[
\rho_{X_1,X_2} := \rho(X_1,X_2) := \frac{\text{Cov}(X_1,X_2)}{\sqrt{\text{Var}(X_1)} \cdot \sqrt{\text{Var}(X_2)}} \in [-1,1].
\]

**Definition 1.12.** (Empirical Pearson correlation)
Let \((X_1, X_2) \sim F\) and \(S_F^n := \{(x_i) := (x_{i1}, x_{i2}); i = 1,\ldots,n\}\) be a sample of \(F\) of size \(n\). Define the empirical means \(\bar{x}_\ell^n := \frac{1}{n} \sum_{i=1}^n x_{i\ell}, \ell \in \{1,2\}\). Then \(\rho_{X_1,X_2}\) is estimated by
\[
\hat{\rho}_{X_1,X_2} := \hat{\rho}(S_F^n) := \frac{\sum_{i=1}^n (x_{i1} - \bar{x}_1^n)(x_{i2} - \bar{x}_2^n)}{\sqrt{\sum_{i=1}^n (x_{i1} - \bar{x}_1^n)^2} \cdot \sqrt{\sum_{i=1}^n (x_{i2} - \bar{x}_2^n)^2}}.
\]

Even though the Pearson correlation is used in a lot of occasions, it involves several shortcomings. As a function of the covariance and the marginal variances, it can only measure linear dependencies between two variables and it is not defined for distributions with non-finite second moments. Further, it is not invariant to monotone transformations of the marginals and we cannot follow independence between two variables \(X_1\) and \(X_2\), if \(\rho_{X_1,X_2} = 0\).

In the context of copulas we often transform variables by the probability integral transform (Definition 1.3). That is why we introduce **Kendall’s \(\tau\)** [Yule and Kendall 1965], which is a rank-based dependence measure and therefore invariant with respect to monotone transformations of the marginals, i.e. \(\tau_{f(X),g(Y)} = \tau_{X,Y}\) for differentiable and monotone functions \(f(\cdot)\) and \(g(\cdot)\).

**Definition 1.13.** (Kendall’s \(\tau\))
Let \(X\) and \(Y\) be random variables and let \(S_F^n := \{(x_i) := (x_{i1}, x_{i2}); i = 1,\ldots,n\}\) be a sample of \(F\) of size \(n\). \(Kendall’s \(\tau\) between \(X\) and \(Y\) is defined as
\[
\tau_{X,Y} := \tau(X,Y) := \mathbb{P}((X_1 - X_2) \cdot (Y_1 - Y_2) > 0) - \mathbb{P}((X_1 - X_2) \cdot (Y_1 - Y_2) < 0).
\]

In order to define an empirical version of Kendall’s \(\tau\) (Definition 1.16), one first needs to introduce the notion of concordant, discordant and extra pairs.

**Definition 1.14.** (Concordant, discordant and extra pairs)
Let \((X_1, X_2) \sim F\) and let \(S_F^n := \{(x_i) := (x_{i1}, x_{i2}); i = 1,\ldots,n\}\) be a sample of \(F\) of size \(n\). A pair \((x_i, x_j) \in S_F^n \times S_F^n, i, j = 1,\ldots,n\) is called
• **concordant**, if either \( x_{i1} < x_{j1} \) and \( x_{i2} < x_{j2} \) or \( x_{i1} > x_{j1} \) and \( x_{i2} > x_{j2} \),

• **discordant**, if either \( x_{i1} < x_{j1} \) and \( x_{i2} > x_{j2} \) or \( x_{i1} > x_{j1} \) and \( x_{i2} < x_{j2} \),

• **extra \( x_1 \) pair**, if \( x_{i1} = x_{j1} \),

• **extra \( x_2 \) pair**, if \( x_{i2} = x_{j2} \).

**Remark 1.15.** (Characterization of concordance and discordance)

\( 1 \) \( (x_i, x_j) \in S_F^n \times S_F^n \) is concordant, iff \( (x_{i1} - x_{j1}) (x_{i2} - x_{j2}) > 0 \), \( i, j = 1, \ldots, n \),

\( 2 \) \( (x_i, x_j) \in S_F^n \times S_F^n \) is discordant, iff \( (x_{i1} - x_{j1}) (x_{i2} - x_{j2}) < 0 \), \( i, j = 1, \ldots, n \).

Since Kendall’s \( \tau \) is determined by the probability of one variable being greater (less) than another, it is intuitive to count the number of concordant, discordant and extra pairs within a sample \( \{x_i := (x_{i1}, x_{i2}); i = 1, \ldots, n\} \) in order to obtain the empirical version.

**Definition 1.16.** (Empirical Kendall’s \( \tau \))

Let \( (X_1, X_2) \sim F \) and \( S_F^n := \{x_i := (x_{i1}, x_{i2}); i = 1, \ldots, n\} \) be a sample of \( F \). Define:

- \( N_c := N_c(S_F^n) := |\{(x_i, x_j) \in S_F^n \times S_F^n : (x_i, x_j) \text{ concordant}, i, j = 1, \ldots, n\}| \)
- \( N_d := N_d(S_F^n) := |\{(x_i, x_j) \in S_F^n \times S_F^n : (x_i, x_j) \text{ discordant}, i, j = 1, \ldots, n\}| \)
- \( N_\ell := N_\ell(S_F^n) := |\{(x_i, x_j) \in S_F^n \times S_F^n : (x_i, x_j) \text{ extra } x_\ell \text{ pair}, i, j = 1, \ldots, n\}|; \ell \in \{1, 2\} \)

Then \( \tau_{X_1,X_2} \) is estimated by

\[
\hat{\tau}^{n}_{X_1,X_2} := \frac{N_c - N_d}{\sqrt{N_c + N_d + N_1 \cdot N_2}}.
\]

**Remark 1.17.** (Empirical Kendall’s \( \tau \) allowing for no extra pairs)

If we assume \( \mathbb{P}(X_1 = X_2) = 0 \) (e.g. \( X_1, X_2 \) continuous), \( \hat{\tau}^{n}_{X_1,X_2} \) simplifies to

\[
\hat{\tau}^{n}_{X_1,X_2} = \frac{N_c - N_d}{\sqrt{N_c + N_d + 0 \cdot \sqrt{N_c + N_d + 0}}} = \frac{N_c - N_d}{\binom{n}{2}}.
\]

Both Pearson correlation and Kendall’s \( \tau \) are dependence measures focusing on the average (global) dependence between two variables. However, we cannot make statements about the probability of joint occurrence of extremely small or large values. Therefore, we also define the **upper and lower tail dependence coefficients** (Definition 1.18), as well as the (maximal) semi-correlation (Definition 1.19), which are tools providing information on extreme (local) events.

**Definition 1.18.** (Upper and lower tail dependence coefficient)

Let \( (X_1, X_2) \sim F, X_1 \sim F_1, X_2 \sim F_2 \) and let \( C \) be the corresponding (bivariate) copula. Then the upper (1) and lower (2) tail dependent coefficients are defined as
Tail dependence coefficients measure the probability of a random variable \( X_1 \) being greater (or less) than a value with very high (or low) probability, given that another variable \( X_2 \) behaves extremely on the same tail.

The **semi-correlation** (Joe [2017]) is defined for uniformly distributed random variables \( U_1 \) and \( U_2 \) only. It splits the domain of \((U_1,U_2)\), which is the unit square \([0,1]^2\), into four sectors \([0,0.5) \times [0,0.5)\) \([0,0.5) \times [0.5,1]\) \([0.5,1] \times [0,0.5)\) and \([0.5,1] \times [0.5,1]\). After transforming the variables by the standard normal distribution function \( \Phi(\cdot) \), it calculates the Pearson correlations for all four sectors. Often it is useful to have a unique measure. Therefore we also introduce the maximal semi-correlation, which is the maximum over all four correlations.

**Definition 1.19.** (Semi-correlations and maximal semi-correlation)

Let \((U_1, U_2) \sim C\) and let \( \Phi(\cdot) \) be the distribution function of a standard normal \( N(0,1) \) distribution. The **upper** and **lower semi-correlations** between \( U_1 \) and \( U_2 \) are defined as

1. \( \rho_{U_1, U_2}^+: = \rho^+(U_1, U_2) := \rho(\Phi^{-1}(U_1), \Phi^{-1}(U_2)| U_1 > 0.5, U_2 > 0.5) \),
2. \( \rho_{U_1, U_2}^-: = \rho^-(U_1, U_2) := \rho(\Phi^{-1}(U_1), \Phi^{-1}(U_2)| U_1 < 0.5, U_2 < 0.5) \).

Additionally we define the **upper-lower** (3) and **lower-upper** (4) **semi-correlations**:

3. \( \rho_{U_1, U_2}^{+-}: = \rho^{+-}(U_1, U_2) := \rho(\Phi^{-1}(U_1), \Phi^{-1}(U_2)| U_1 > 0.5, U_2 < 0.5) \),
4. \( \rho_{U_1, U_2}^{-+}: = \rho^{-+}(U_1, U_2) := \rho(\Phi^{-1}(U_1), \Phi^{-1}(U_2)| U_1 < 0.5, U_2 > 0.5) \).

Finally we can define the **maximal semi-correlation** between \( U_1 \) and \( U_2 \):

5. \( \rho_{U_1, U_2}^{max}: = \max\{ \rho_{U_1, U_2}^+, \rho_{U_1, U_2}^-, \rho_{U_1, U_2}^{+-}, \rho_{U_1, U_2}^{-+} \} \).

Empirical versions of the semi-correlations are defined in **Definition 1.20**. Pearson correlations are estimated, after the uniformly distributed data have been split into the four sectors described above.

**Definition 1.20.** (Empirical semi-correlations and maximal empirical semi-correlation)

Let \((U_1, U_2) \sim C\) and \( S_C^n := \{u_i := (u_{i1}, u_{i2}): i = 1, \ldots, n\} \) be a sample of \( C\). Define \( S_C^{n,+} := \{u_i \in S_C^n : u_{i1} > 0.5, u_{i2} > 0.5\} \), \( S_C^{-} := \{u_i \in S_C^n : u_{i1} < 0.5, u_{i2} < 0.5\} \), \( S_C^{n,-} := \{u_i \in S_C^n : u_{i1} > 0.5, u_{i2} < 0.5\} \), \( S_C^{n,+} := \{u_i \in S_C^n : u_{i1} < 0.5, u_{i2} > 0.5\} \).

Then, \( \rho_{U_1, U_2}^+, \rho_{U_1, U_2}^-, \rho_{U_1, U_2}^{+-} \) and \( \rho_{U_1, U_2}^{-+} \) are estimated by
The 

 maximal empirical semi-correlation
 is defined as

 \[ \hat{\rho}_{U_1,U_2}^{\text{semi}} := \max\{\hat{\rho}_{U_1,U_2}, \hat{\rho}_{U_1,U_2}^+, \hat{\rho}_{U_1,U_2}^-, \hat{\rho}_{U_1,U_2}^{++}, \hat{\rho}_{U_1,U_2}^{+-}, \hat{\rho}_{U_1,U_2}^{--}\}. \]

 If we are given \( d \) uniformly distributed random variables \( U_1, \ldots, U_d \), the presented dependence measures provide information on the dependence of any pair of variables \( U_i, U_j \), \( i \neq j \in \{1, \ldots, d\} \). Additionally, one might be interested in the dependence of \( U_i \) and \( U_j \) after the effect of the remaining variables \( U_k, k \in \{1, \ldots, d\} \setminus \{i,j\} \), is removed. That is why we define a last dependence measure - the partial correlation.

**Definition 1.21.** (Partial correlation)

Let \( Z_1, \ldots, Z_d \) be random variables with \( \mathbb{E}[Z_j] = 0 \) and \( \text{Var}[Z_i] = \sigma_i^2 \), \( i = j, \ldots, d \). Define \( I_{d-(i,j)} := \{1, \ldots, d\} \setminus \{i, j\} \) and \( Z_{I_{d-(i,j)}} := \{Z_k : k \in I_{d-(i,j)}\} \) for \( i \neq j \), \( i, j \in \{1, \ldots, d\} \). Then, the partial correlation between \( Z_i \) and \( Z_j \) given \( Z_{I_{d-(i,j)}} \) is defined as

\[ \rho_{Z_i,Z_j;I_{d-(i,j)}} := \text{sign}(b_{i,j;i_{d-(i,j)}}) \cdot \sqrt{b_{i,j;i_{d-(i,j)}} \cdot b_{j,i;i_{d-(i,j)}}}, \]

where \( \{b_{i,j;i_{d-(i,j)}}, i < j, i, j \in \{1, \ldots, d\}\} \) is chosen such that the quantities \( b_{i,j;i_{d-(i,j)}} \) minimize the functions

\[ Q(a_{i,j;i_{d-(i,j)}}, 1 \leq i < j \leq d) := \mathbb{E}[(Z_i - \sum_{j=2,j \neq i}^d a_{i,j;i_{d-(i,j)}} Z_j)^2]. \]

Observe that, if we define \( Z_j := \Phi(U_j), j = 1, \ldots, d \), the resulting random variables fulfill the assumptions in Definition 1.21. That is why we will transform uniformly distributed data to standard normal data, before the corresponding partial correlations are estimated. By Theorem 1.22, partial correlations are functions of the available Pearson correlations only. Thus, one can estimate empirical versions of the partial correlations starting from estimated Pearson correlations. The recursion formula was proven by Yule and Kendall [1965]. Example 1.23 illustrates the estimation of the partial correlations for \( d = 4 \) variables \( Z_j := \Phi(U_j), j = 1, \ldots, 4 \).

**Theorem 1.22.** (Recursion for partial correlations)

Let \( Z = (Z_1, \ldots, Z_d)^T \) be a \( d \)-dimensional random vector with \( Z_j \) fulfilling the assumptions of Definition 1.21. Let \( i, j, k \in \{1, \ldots, d\}, i \neq j \neq k \) and define \( Z^{-i,j} := \{Z_1, \ldots, Z_d\} \setminus \{Z_i, Z_j\}, Z^{-i,j,k} := \{Z_1, \ldots, Z_d\} \setminus \{Z_i, Z_j, Z_k\} \). Then, the partial correlations defined in Definition 1.21 are obtained by the following recursion formula:

\[ \rho_{Z_i,Z_j;Z^{-i,j}} = \frac{\rho_{Z_i,Z_j;Z^{-i,j,k}} - \rho_{Z_i,Z_k;Z^{-i,j,k}} \rho_{Z_j,Z_k;Z^{-i,j,k}}}{\sqrt{1 - \rho_{Z_i,Z_d;Z^{-i,j,k}}^2} \cdot \sqrt{1 - \rho_{Z_j,Z_d;Z^{-i,j,k}}^2}}, \]
Example 1.23. (Estimation of partial correlations for 4 variables)

Let \( z_j = (z_{1j}, \ldots, z_{nj})^T \) be \( n \) observations of variable \( Z_j := \Phi(U_j), j = 1, \ldots, 4 \). We want to estimate the empirical partial correlations

\[
\hat{\rho}_{Z_j, j}^{\ell} := \Phi(U_j), j_\ell \in \{1, \ldots, 4\}, j_\ell \neq j_k \text{ for all } \ell \neq k.
\]

Using the given observations \( z_j, j = 1, \ldots, 4 \), we start with estimating the empirical Pearson correlations, defined in Definition 1.12:

\[
\hat{\rho}_{Z_j, k}^{\ell} := \Phi(U_j), j, k \in \{1, \ldots, 4\}, j \neq k.
\]

In the next step, we calculate the empirical partial correlations

\[
\hat{\rho}_{Z_j, j_\ell}^{\ell} := \Phi(U_j), j_\ell \in \{1, \ldots, 4\}, j_\ell \neq j_k \text{ for all } \ell \neq k.
\]

Now we illustrate the recursion formula. For example we might be interested in the empirical partial correlation \( \hat{\rho}_{Z_3, Z_2; Z_1, Z_4} \).

By Theorem 1.22 we know that

\[
\hat{\rho}_{Z_3, Z_2; Z_4} = \frac{\hat{\rho}_{Z_3, Z_2} - \hat{\rho}_{Z_3, Z_4} \hat{\rho}_{Z_2, Z_4}}{\sqrt{1 - (\hat{\rho}_{Z_3, Z_4})^2} \sqrt{1 - (\hat{\rho}_{Z_2, Z_4})^2}}, \tag{1.2}
\]

\[
\hat{\rho}_{Z_3, Z_1; Z_4} = \frac{\hat{\rho}_{Z_3, Z_1} - \hat{\rho}_{Z_3, Z_4} \hat{\rho}_{Z_1, Z_4}}{\sqrt{1 - (\hat{\rho}_{Z_3, Z_4})^2} \sqrt{1 - (\hat{\rho}_{Z_1, Z_4})^2}}, \tag{1.3}
\]

and

\[
\hat{\rho}_{Z_2, Z_1; Z_4} = \frac{\hat{\rho}_{Z_2, Z_1} - \hat{\rho}_{Z_2, Z_4} \hat{\rho}_{Z_1, Z_4}}{\sqrt{1 - (\hat{\rho}_{Z_2, Z_4})^2} \sqrt{1 - (\hat{\rho}_{Z_1, Z_4})^2}}. \tag{1.4}
\]

Then \( \hat{\rho}_{Z_3, Z_2; Z_1, Z_4} \) is given by

\[
\hat{\rho}_{Z_3, Z_2; Z_1, Z_4} = \frac{\hat{\rho}_{Z_3, Z_2; Z_1, Z_4} - \hat{\rho}_{Z_3, Z_1; Z_4} \hat{\rho}_{Z_2, Z_1; Z_4}}{\sqrt{1 - (\hat{\rho}_{Z_3, Z_1; Z_4})^2} \sqrt{1 - (\hat{\rho}_{Z_2, Z_1; Z_4})^2}},
\]

where \( \hat{\rho}_{Z_3, Z_2; Z_1, Z_4}, \hat{\rho}_{Z_3, Z_1; Z_4} \) and \( \hat{\rho}_{Z_2, Z_1; Z_4} \) are given in Equations (1.2), (1.3) and (1.4).

The empirical partial correlations \( \hat{\rho}_{Z_j, j_\ell; Z_{j_k} Z_{j_k} \ell}, j_\ell \in \{1, \ldots, 4\} \), can be obtained similarly.
1.3 Bivariate parametric copulas

In this section we introduce the class of bivariate parametric copulas. They are defined by a copula family and a copula parameter vector.

**Definition 1.24.** (Bivariate parametric copulas)
A bivariate copula \( C(\cdot, \cdot) \) is called parametric, if it is uniquely defined by a copula family \( b := b(C) \) and a parameter vector \( \theta := \theta(b(C)) \). Then, we write \( C^b(\cdot, \cdot; \theta) \).

As indicated by the notation \( \theta(b(C)) \), domain and dimension of the parameter vector \( \theta \) are restricted by the copula family \( b(C) \). Since most of the common copula families allow for one or two parameters, we set \( \theta = (\theta_1, \theta_2) \).

\( \theta_2 \) is set to 0, if \( b(C) \) allows univariate parameters only. Then the dimension of \( \theta \) is defined by the number of non-zero entries, i.e.

\[ |\theta| := 1_{\{\theta_1 \neq 0\}} + 1_{\{\theta_2 \neq 0\}}. \]

In the following, we present various examples of bivariate parametric copulas. Two popular classes of parametric copula families are elliptical and Archimedean copula families. While elliptical copulas correspond to elliptical distributions, such as the normal and Student t distribution, Archimedean copulas are defined by a generator \( \phi \), which is a function fulfilling certain assumptions.

**Elliptical copulas**

**Example 1.25.** (Bivariate Gaussian (N) copula)
Let \( \Phi(\cdot) \) be the distribution function of a (univariate) standard normal \( \mathcal{N}(0, 1) \) distribution (with density \( \varphi(\cdot) \)) and let \( \Phi_2(\cdot, \cdot; \rho) \) be the distribution function of a bivariate normal \( \mathcal{N}_2(\mu, \Sigma_\rho) \) distribution with \( \mu = (0, 0) \) and \( \Sigma_\rho = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \) for some \( \rho \in [-1, 1] \).

Applying the inverse statement of Sklar’s theorem (Theorem 1.7) we obtain the bivariate Gaussian copula:

\[ C^N(u_1, u_2; \rho) = \Phi_2(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \rho) \]

The corresponding copula density is given by

\[ c^N(u_1, u_2; \rho) = \frac{1}{\varphi(z_1)\varphi(z_2)} \frac{1}{\sqrt{1-\rho^2}} \exp \left\{ -\frac{\rho^2(z_1^2 + z_2^2) - 2\rho z_1 z_2}{2(1-\rho^2)} \right\}, \]

where \( z_j := \Phi^{-1}(u_j), j \in \{1, 2\} \)

**Example 1.26.** (Bivariate Student t copula)
Let \( T_\nu(\cdot) \) be the distribution function of a (univariate) Student t distribution with \( \nu > 0 \) degrees of freedom (with density \( t_\nu(\cdot) \)) and let \( t_2(\cdot, \cdot; \nu, \rho) \) be the density of a bivariate
Student $t_2(\nu, \mu, \Sigma_\rho)$ distribution with $\nu > 0$ degrees of freedom, $\mu = \left(\begin{array}{c}0 \\ 0\end{array}\right)$ and $\Sigma_\rho = \begin{pmatrix}1 & \rho \\ \rho & 1\end{pmatrix}$ for some $\rho \in [-1, 1]$.

Applying the inverse statement of Sklar’s theorem (Theorem 1.7) we obtain the bivariate Student $t$ copula:

$$C_t(u_1, u_2; \nu, \rho) = \beta_1 \int_{-\infty}^{b_1} \int_{-\infty}^{b_2} t(x_1, x_2; \nu, \rho) dx_1 dx_2,$$

where $b_j := T_\nu(u_j)$, $j \in \{1, 2\}$. The corresponding copula density is given by

$$c_t(u_1, u_2; \nu, \rho) = \frac{t_2(T_\nu^{-1}(u_1), T_\nu^{-1}(u_2); \nu, \rho)}{t_\nu(T_\nu^{-1}(u_1)) t_\nu(T_\nu^{-1}(u_2))}.$$

**Archimedean copulas**

**Definition 1.27.** (Bivariate Archimedean copulas)

$C(\cdot, \cdot)$ is called bivariate Archimedean copula with generator $\phi$, if:

$$C(u_1, u_2) = \phi^{-1}[(\phi(u_1) + \phi(u_2)),$$

where $\phi : [0, 1] \rightarrow [0, \infty]$ is strict monotone increasing and convex with $\phi(1) = 0$. $\phi^{-1}$ denotes the pseudo inverse of $\phi$, i.e.

$$\phi^{-1} : [0, \infty] \rightarrow [0, 1]; \phi^{-1}(t) := 1_{\{0 \leq t \leq \phi(0)\}} \phi^{-1}(t).$$

If $\phi(0) = \lim_{t \rightarrow 0^+} \phi(t) = \infty$, $\phi$ is called strict generator.

**Lemma 1.28.** (Density of Archimedean copulas)

If $C$ is a continuous Archimedean copula, then its density $c$ can be expressed as

$$c(u_1, u_2) = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2} = \frac{\phi''(C(u_1, u_2)) \phi'(u_1) \phi'(u_2)}{[\phi'(C(u_1, u_2))]^3}.$$

**Example 1.29.** (One-parametric bivariate Archimedean copulas)

1. Bivariate Clayton ($C$) copula: let $0 < \delta < \infty$:

   $$C^C(u_1, u_2; \delta) = (u_1^{-\delta} + u_2^{-\delta} - 1)^{-\frac{1}{\delta}}.$$

2. Bivariate Gumbel ($G$) copula: let $\delta \geq 1$:

   $$C^G(u_1, u_2; \delta) = \exp \left\{- [(- \log(u_1))^\delta + (- \log(u_2))^\delta]^{\frac{1}{\delta}}\right\}.$$

3. Bivariate Frank ($F$) copula: let $\delta \in [-\infty, \infty] \setminus \{0\}$:

   $$C^F(u_1, u_2; \delta) = -\frac{1}{\delta} \log \left(\frac{1}{1 - e^{-\delta}} \left[\left(1 - e^{-\delta u_1}\right)\left(1 - e^{-\delta u_2}\right)\right]\right).$$
(4) Bivariate Joe \((J)\) copula: let \(\delta \geq 1\):

\[
C^J(u_1, u_2; \delta) = 1 - \left[ (1 - u_1)^\delta + (1 - u_2)^\delta - (1 - u_1)^\delta (1 - u_2)^\delta \right]^\frac{1}{\delta}.
\]

**Example 1.30.** (Bivariate BB copulas)

(1) Bivariate BB1 copula: let \(\theta > 0, \delta \geq 1\) and \(\eta_1(s) := \eta_1(s; \theta, \delta) := (1 + s^\frac{1}{\delta})^{-\frac{1}{\theta}}\): \(C^{BB1}(u_1, u_2; \theta, \delta) = \eta_1(\eta_1^{-1}(u_1) + \eta_1^{-1}(u_2))\).

(1) Bivariate BB7 copula: let \(\theta \geq 1, \delta > 0\) and \(\eta_7(s) := \eta_7(s; \theta, \delta) := 1 - \left[ 1 - (1 + s)^{-\frac{1}{\delta}} \right]^\frac{1}{\theta}\): \(C^{BB7}(u_1, u_2; \theta, \delta) = \eta_7(\eta_7^{-1}(u_1) + \eta_7^{-1}(u_2))\).

A non-parametric bivariate copula, used to model independence between two uniformly distributed random variables \(U_1\) and \(U_2\) is given by the so-called independence copula defined in Definition 1.31.

**Definition 1.31.** (Bivariate independence copula)

\(C^I(\cdot, \cdot)\) is called independence copula, if

\[
C^I(u_1, u_2) = u_1 \times u_2.
\]

The corresponding copula density is given by

\[
c^I(u_1, u_2) = 1.
\]

Note that the conditional copulas \(C^I_{1|2}(\cdot|u_2)\) and \(C^I_{2|1}(\cdot|u_1)\) corresponding to the independence copula are given by

\[
C^I_{1|2}(u_1|u_2) = \frac{\partial}{\partial u_2} C^I(u_1, u_2) = u_1 \text{ and } \quad C^I_{2|1}(u_2|u_1) = \frac{\partial}{\partial u_1} C^I(u_1, u_2) = u_2,
\]

i.e. the conditional distribution of one variable is independent of the other.

In order to emphasize the difference in the tail behavior between elliptical and Archimedean copulas, we close this section by presenting normalized contour plots of a bivariate Gaussian, a bivariate Student t, a bivariate Clayton and a bivariate Gumbel copula in Figure 1.1.
1.4 Relationship between Kendall’s τ and copula parameters

In this section we analyze the behavior of Kendall’s τ corresponding to a bivariate parametric copula $C^b(\cdot,\cdot;\theta)$ given a certain parameter vector $\theta$. While parameters are not interpretable for different copula families, we can use the corresponding Kendall’s τ in order to compare the dependencies within different copula families.

Theorem 1.32 provides formulas to calculate Kendall’s τ out of the copula parameters for bivariate elliptical and Archimedean copulas. The proof for elliptical copulas is given for example in Embrechts et al. [2003], while the proof for the Archimedean copula can be found in Hurlimann [2003].

**Theorem 1.32.** (Kendall’s τ for bivariate elliptical and Archimedean copulas)

Let $\rho$ be the association parameter of a bivariate elliptical copula and $\phi$ the generator of
a bivariate Archimedean copula. Then

$$\rho = \sin\left(\frac{\pi}{2}\tau\right)$$

for elliptical copulas and

$$\tau = 1 + 4 \int_0^1 \frac{\phi(t)}{\phi'(t)} dt$$

for Archimedean copulas.

Example 1.33 shows Kendall’s $\tau$ corresponding to the bivariate copulas presented in Section 1.3. $\tau^b(\theta_1)$ denotes Kendall’s $\tau$ corresponding to a bivariate parametric copula $C^b(\cdot, \cdot; \theta)$ with copula family $b$ and parameter vector $\theta = (\theta_1, \theta_2)$. Note that this notation is sufficient, since for the considered bivariate parametric copulas, Kendall’s $\tau$ solely depends on the first copula parameter.

**Example 1.33.** (Relationship between Kendall’s $\tau$ and copula parameters for different bivariate parametric copulas)

1. Gaussian copula:
   $$\tau^N(\rho) = \frac{2}{\pi} \arcsin(\rho), \quad \rho \in [-1, 1] \Rightarrow \tau^N(\rho) \in [-1, 1].$$

2. Student t copula:
   $$\tau^t(\rho) = \frac{2}{\pi} \arcsin(\rho), \quad \rho \in [-1, 1] \Rightarrow \tau^t(\rho) \in [-1, 1].$$

3. Clayton copula:
   $$\tau^C(\delta) = \frac{\delta}{\delta + 2}, \quad 0 < \delta < \infty \Rightarrow \tau^C(\delta) \in [0, 1].$$

4. Gumbel copula:
   $$\tau^G(\delta) = 1 - \frac{1}{\delta}, \quad 1 \leq \delta < \infty \Rightarrow \tau^G(\delta) \in [0, 1].$$

5. Frank copula:
   $$\tau^F(\delta) = 1 - \frac{4}{\delta} + 4 \frac{D_1(\delta)}{\delta} \Rightarrow \tau^F(\delta) \in [-1, 1],$$
   with $D_1(\delta) = \int_0^\delta \frac{x}{\delta(e^x - 1)} dx, \quad \delta \in [-\infty, \infty] \setminus \{0\}.$
(6) Joe copula:

$$\tau^J(\delta) = 1 + \gamma + 2 \log(2) \frac{\Psi\left(\frac{1}{2}\right)}{\delta} + \Psi\left(\frac{1+\delta}{2}\right) + \delta \Rightarrow \tau^J(\delta) \in [0, 1],$$

with $\gamma = \lim_{n \to \infty} \left( \frac{1}{n} \sum_{i=1}^{n} -\log(n) \right) \approx 0.57721$

and $\Psi(x) = \frac{d}{dx} \log(\Gamma(x)) = \frac{d}{dx} \frac{\Gamma(x)}{\Gamma(x)}$.

Since $\lim_{\delta \to 0^+} \tau^C(\delta) = 0$ and $\lim_{\delta \to \infty} \tau^C(\delta) = 1$, we can model independent Clayton copula data for small values of $\delta$ and strongly positively dependent data using a high copula parameter. However, we cannot model negative dependencies using Clayton, Gumbel or Joe copulas. The range of Kendall’s $\tau$ can be extended by counterclockwise rotations.

**Definition 1.34.** (Rotated copula densities)
Let $c^b(\cdot; \theta)$ be the density corresponding to a bivariate parametric copula with copula family $b$ and parameter vector $\theta$. Then, we can define the following rotated copula densities:

- 90 degree rotation: $c^{b}_{90}(u_1, u_2; \theta) := c^b(1 - u_1, u_2; \theta)$,
- 180 degree rotation: $c^{b}_{180}(u_1, u_2; \theta) := c^b(1 - u_1, 1 - u_2; \theta)$,
- 270 degree rotation: $c^{b}_{270}(u_1, u_2; \theta) := c^b(u_1, 1 - u_2; \theta)$.

For example, if we extend the density of a bivariate Clayton copula to

$$c^C_{\text{extended}}(u_1, u_2; \delta) := \begin{cases} 
 c^C(u_1, u_2; \delta), & \text{if } \delta > 0 \\
 c^C_{90}(u_1, u_2; \delta), & \text{otherwise}
\end{cases},$$

we extend the range of Kendall’s $\tau$ from $[0, 1]$ to $[-1, 1]$.

### 1.5 Variable scales

In this section, we briefly describe the notation we are using when we work with data sets containing observations

- $\mathbf{y} = (y_1, \ldots, y_n)^T \in \mathbb{R}^n$ and
- $\mathbf{x}_1, \ldots, \mathbf{x}_J \in \mathbb{R}^n, \mathbf{x}_{ij} := (x_{1j}, \ldots, x_{nj})^T, j = 1, \ldots, J$

on

- a response variable $\mathbf{Y} \sim F_Y$ and
• $J$ covariates $X_1, \ldots, X_d, [X_j] \sim F_j$, $j = 1, \ldots, J$.

Since copulas are solely defined on the unit square, it is necessary to transform the original data vectors. Therefore we assume that their marginal distributions $F_Y, F_j$, $j = 1, \ldots, J$ are known.

Then, we can define the PIT-transformed data vectors

- $v = (v_1, \ldots, v_n)^T \in [0, 1]^n$ and
- $u_1, \ldots, u_J \in [0, 1]^n$, $[u_j] := (u_{1j}, \ldots, u_{nj})^T$, $j = 1, \ldots, J$,

where

- $v := (F_Y(y_{1j}), \ldots, F_Y(y_{nj}))^T$ and
- $u_j := (F_j(x_{1j}), \ldots, F_j(x_{nj}))^T$, $j = 1, \ldots, J$.

The transformed vectors contain observations of

- the transformed response variable $V := F_Y(Y) \sim U(0, 1)$ and
- $J$ transformed covariates $U_1, \ldots, U_d$, $[U_j] := F_j(X_j) \sim U(0, 1)$, $j = 1, \ldots, J$.

While the original or untransformed data vectors and variables are said to be on the x-scale, we say that the transformed observations and variables are on the copula scale or u-scale. Additionally, we introduce the so called z-scale. Therefore, copula data are transformed by the inverse of the standard normal distribution function $\Phi^{-1}(\cdot)$ and we obtain observations

- $w = (w_1, \ldots, w_n)^T \in \mathbb{R}^n$ and
- $z_1, \ldots, z_J \in \mathbb{R}^n$, $[z_j] := (z_{1j}, \ldots, z_{nj})^T$, $j = 1, \ldots, J$,

where

- $w := (\Phi^{-1}(v_{1j}), \ldots, \Phi^{-1}(v_{nj}))^T$ and
- $z_j := (\Phi^{-1}(u_{1j}), \ldots, \Phi^{-1}(u_{nj}))^T$, $j = 1, \ldots, J$.

The transformed vectors contain observations of

- the transformed response variable $W := \Phi^{-1}(V) \sim N(0, 1)$ and
- $J$ transformed covariates $Z_1, \ldots, Z_d$, $[Z_j] := \Phi^{-1}(U_j) \sim N(0, 1)$, $j = 1, \ldots, J$. 

21
1.6 Estimation of bivariate parametric copulas

In the following, we assume that we are given data
\[ U^n := (u_1, u_2) \in [0, 1]^{n \times 2} \]
on the copula scale, where
\[ u_j := (u_{1j}, \ldots, u_{nj})^T \in [0, 1]^n, \; j \in \{1, 2\}. \]

Our goal is to find the bivariate parametric copula \( \hat{C}^b(\cdot, \cdot; \hat{\theta}) \), which fits the available data in the “best” way. In other words, we want to estimate the copula family \( b \) and parameters \( \hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2)^T \) from given data.

In statistics, the “best” fit is usually quantified by a goodness of fit measure. Common measures are log likelihood, AIC (Akaike information criterion) and BIC (Bayesian information criterion). For copulas they are defined in the same way as for other distributions. Definition 1.35 lists the according formulas.

**Definition 1.35.** (Log likelihood, AIC and BIC of bivariate parametric copulas)

Let \( C^b(\cdot, \cdot; \theta) \) be a bivariate parametric copula with density \( c^b(\cdot, \cdot; \theta) \). Additionally assume that we are given copula data \( U^n \). Given \( U^n \), the log likelihood (1), AIC (2) and BIC (3) corresponding to \( C^b(\cdot, \cdot; \theta) \) are defined as

1. \[ \ell^{ll}(C^b(\cdot, \cdot; \theta), U^n) := \sum_{i=1}^{n} \log(c^b(u_{i1}, u_{i2}, \theta)), \]
2. \[ \text{AIC}(C^b(\cdot, \cdot; \theta), U^n) := -2 \ell^{ll}(C^b(\cdot, \cdot; \theta), U^n) + 2|\theta|, \]
3. \[ \text{BIC}(C^b(\cdot, \cdot; \theta), U^n) := -2 \ell^{ll}(C^b(\cdot, \cdot; \theta), U^n) + \log(n)|\theta|, \]

where \( |\theta| \) denotes the number of non-zero entries in \( \theta \).

In the course of the estimation, we want to either maximize the log likelihood or minimize one of the two information criteria. Which measure we choose, is defined by a parameter \( \text{sc} \) (selection criterion). If we set \( \text{sc} = \text{AIC} \) or \( \text{sc} = \text{BIC} \) we look for the bivariate parametric copula, which minimizes \( \text{AIC}(C^b(\cdot, \cdot; \theta), U^n) \) or \( \text{BIC}(C^b(\cdot, \cdot; \theta), U^n) \), respectively. We write \( \text{sc} = \ell^{ll} \), if we want to maximize the log likelihood, i.e. we choose

\[ \text{sc} \in \{\ell^{ll}, \text{AIC}, \text{BIC}\}. \]

Further we need to specify a set of possible copula families \( \mathcal{B} \), i.e. a set of those copula families, which we want to consider during the optimization. For example, if we set \( \mathcal{B} = \{N, t, C, C_{90}, C_{180}, C_{270}\} \), we can only optimize the chosen selection criterion with respect to Gaussian, Student t and Clayton copulas, where also rotated Clayton copulas are possible.

After the choice of the parameters \( \text{sc} \) and \( \mathcal{B} \), we can find a unique optimal bivariate parametric copula. In R, we use the function \( \text{BiCopSelect}() \) from the \texttt{VineCopula} package.
Definition 1.36. (Selection criterion optimizing bivariate parametric copula)
Assume we are given bivariate copula data \( U^n := (u_1, u_2) \in [0, 1]^{n \times 2} \) and let \( B \) be the set of possible copula families. Additionally assume that we are given a selection criterion \( \text{sc} \in \{ \ell \ell, \text{AIC}, \text{BIC} \} \). Then the selection criterion optimizing bivariate parametric copula is defined as

\[
\text{optC}(U^n, B, \text{sc}) := \max_{b \in B, \theta(b)} \{ \text{sc}(C_b(\cdot, \cdot; \theta(b)), U^n) \}.
\]

The following remark discusses the effect of the choice of \( \text{sc} \). Since we usually consider copula families that allow for one or two parameters, the penalty terms for the number of parameters in AIC and BIC have no or only a small effect.

Remark 1.37. (Effect of the selection criterion)
Assume that \( C_{b^*}^1(\cdot, \cdot; \theta^*_1) = \text{optC}(U^n, B, \ell \ell) \) and \( C_{b^*}^2(\cdot, \cdot; \theta^*_2) = \text{optC}(U^n, B, \text{AIC}) = \text{optC}(U^n, B, \text{BIC}) \).

For \( j \in \{1, 2\} \) we define

\[
\ell \ell_j := \ell \ell(C_{b^*_j}^j(\cdot, \cdot; \theta^*_j), U^n), \quad \text{AIC}_j := \text{AIC}(C_{b^*_j}^j(\cdot, \cdot; \theta^*_j), U^n), \quad \text{and BIC}_j := \text{BIC}(C_{b^*_j}^j(\cdot, \cdot; \theta^*_j), U^n).
\]

Then we know that the following must hold:

\[
\ell \ell_1 > \ell \ell_2, \quad (1.5)
\]

\[
-2\ell \ell_1 + 2|\theta^*_1| > -2\ell \ell_2 + 2|\theta^*_2|, \quad (1.6)
\]

\[
-2\ell \ell_1 + \log(n)|\theta^*_1| > -2\ell \ell_2 + \log(n)|\theta^*_2|. \quad (1.7)
\]

If we only consider copula parameters with \( |\theta| \leq 2 \), we follow that \( |\theta^*_1| = 2 \) and \( |\theta^*_2| = 1 \) must hold. Equations (1.6) and (1.7) simplify to

\[
-2\ell \ell_1 + 4 > -2\ell \ell_2 + 2 \iff \ell \ell_1 < \ell \ell_2 + 1, \quad (1.8)
\]

\[
-2\ell \ell_1 + 2 \log(n) > -2\ell \ell_2 + \log(n) \iff \ell \ell_1 < \ell \ell_2 + \frac{1}{2} \log(n). \quad (1.9)
\]

Summarizing our results we know that, the log likelihood maximizing copula is only different from the AIC minimizing copula, if the following inequality holds:

\[
\ell \ell_2 < \ell \ell_1 < \ell \ell_2 + 1.
\]

Similarly, if log likelihood maximizing copula differs from the BIC minimizing one:

\[
\ell \ell_2 < \ell \ell_1 < \ell \ell_2 + \frac{1}{2} \log(n).
\]

Thus, the choice of \( \text{sc} \) influences the estimation only in very special cases.
Remark 1.38. (Independence test prior to the optimization of the selection criterion)

The R-function \textit{BiCopSelect()} offers the option of conducting an independence test between \( U_1 \) and \( U_2 \) before the selection criterion is optimized. It has the following structure:

The null-hypothesis

\[ H_0 : U_1 \text{ is independent of } U_2 \]

is rejected at a level \( \alpha = 0.05 \), if

\[ 2(1 - \Phi(T)) < 0.05, \]

where

\[ T := \sqrt{\frac{9n(n-1)}{2(2n + 5)}} \times |\hat{\tau}_{U_1,U_2}^n|, \]

and \( \Phi(\cdot) \) denotes the distribution function of a \( N(0,1) \)-random variable.

If the null-hypothesis \( H_0 \) is not rejected, a bivariate independence copula (Definition 1.31) is chosen to describe the dependence between \( U_1 \) and \( U_2 \). In the following, we will assume that such an independence test is not conducted, unless we explicitly set \texttt{indepTest}, which is a parameter in the R-function \textit{BiCopSelect()}, equal to 1.
Chapter 2

Regular vine copulas and distributions

In the following chapter, we present background on regular vine copulas (Bedford and Cooke [2002a]). The exposition follows closely Chapter 5 in Scherer and Mai [2017] and Kurowicka and Cooke [2006].

2.1 Graph theory

Besides theory on bivariate parametric copulas, graph theoretic concepts are important. The exposition here follows Diestel [2017]. The defined notions are illustrated in Example 2.8.

Definition 2.1. (Graph, node, edge)

• A graph $G$ is a pair $(N, E)$ of sets, such that $E \subseteq \{(x, y) : x, y \in N\}$.
• $e \in E$ is called edge of the graph $G$, while $v \in N$ is called node.

Definition 2.2. (Neighbor, degree, leaf)
Let $G = (N, E)$ be a graph and $v, w \in N$.

• $v$ is called neighbor of $w$, if there exists an edge $e = (v, w) \in E$ connecting $v$ and $w$.

• The degree of a node $v \in N$ is defined as the number of neighbors of $v$ and denoted by $d(v)$, i.e. $d(v) = |\{w \in N \setminus \{v\} : w \text{ neighbor of } v\}|$.

• $v$ is called leaf in $G$, if $d(v) = 1$.

Remark 2.3.
In the regular vine framework we usually consider undirected and unweighted graphs.

(1) $G = (N, E)$ is called undirected if all edges $e = (x, y) \in E$ are unordered pairs of nodes, i.e. for all $x, y \in N$: $(x, y) = (y, x)$. 
(2) If we introduce a function $\omega : E \to \mathbb{R}$, which assigns weights to each edge $e \in E$, a graph is called \textit{weighted}.

**Definition 2.4.** (Complete graphs and subgraphs)

- A graph is called \textit{complete}, if all nodes $v \in N$ are directly connected by edges, i.e. $E = \{(x, y) : x, y \in N\}$.
- A subgraph of a graph $G=(N,E)$ is a graph $G' = (N', E')$ with $N' \subseteq N$ and $E' \subseteq E$.

**Definition 2.5.** (Paths, cycles and connected graphs)

- A \textit{path} is a graph $P = (N,E)$ with node set $N = \{v_0, v_1, \ldots, v_k\}$ and edges $E = \{\{v_0, v_1\}, \{v_1, v_2\}, \ldots, \{v_{k-1}, v_k\}\}$.
- A \textit{cycle} is a path with $v_0 = v_k$.
- A graph is said to be \textit{connected}, if for all $v_1, v_2 \in N$ there exists a linking path between $v_1$ and $v_2$.

**Theorem 2.6.** (Characterization of trees)
The following statements are equivalent for a graph $T = (N, E)$.

(i) $T$ is a \textit{tree}.

(ii) Any two nodes of $T$ are linked by a unique path in $T$.

(iii) $T$ is minimally connected, i.e. $T$ is connected but $T - e$ is disconnected $\forall e \in E$, where $T - e$ denotes the graph $T' = (N, E')$ with $E' = E \setminus \{e\}$.

(iv) $T$ is maximally acyclic, i.e. $T$ contains no cycle but $T + \{x, y\}$ does for any two non-adjacent nodes $x, y \in N$, where $T + \{x, y\}$ denotes the graph $T$ with additional edge $\{x, y\}$ and $x, y$ non-adjacent means that $\{x, y\} \notin E$.

**Definition 2.7.** (Spanning tree, star tree, root node)

- A \textit{spanning tree} of a graph $G = (N,E)$ is a subgraph $T = (N_T, E_T)$, which is a tree with $N_T = N$.
- A tree, which has a node $v_0$ with $d(v_0) = |N| - 1$, is called a \textit{star tree} and $v_0$ the \textit{root node} of the tree. Obviously, in stars it holds that $d(v) = 1 \ \forall v \in N \setminus \{v_0\}$.

The following example shows three different graphs. While nodes are represented by rectangles, we use lines for edges.

**Example 2.8.** (Illustration of graph theoretic notions)
Figure 2.1 shows three graphs, i.e. $G = (N^G, E^G)$, $T = (N^T, E^T)$ and $S = (N^S, E^S)$, all consisting of the same 5 nodes, i.e. $N^G = N^T = N^S = \{1, \ldots, 5\}$.

It is obvious that graph $G$ is not a tree, since node 1 is not connected to the remaining nodes. Also the subgraph $G' = (N', E')$ consisting of nodes $N' = \{2, \ldots, 4\}$
Figure 2.1: Examples of graphs: disconnected graph G, tree T and star tree S.

and the corresponding edge set \( E' \) only is not a tree, since it is not acyclic. For example, it contains the cycle 2 - 3 - 4 - 5 - 2. A possible spanning tree of \( G' \) is given by the subgraph of \( G' = (N'', E'') \) consisting of the node set \( N'' = N' \) and the edge set \( E'' = \{(2, 3), (3, 4), (4, 5)\} \).

\( T \) is a tree because any pair of nodes is connected by a unique path. For example, nodes 1 and 5 are only connected by the path 1 - 2 - 4 - 5. Node 2 has neighbors 1,3 and 4, since edges \((1, 2), (2, 3)\) and \((2, 4)\) are elements of \( E^T \). Therefore the degree of node 2 is given by \( \text{dg}(2) = 3 \). The degrees of the remaining nodes are given by \( \text{dg}(1) = 1, \text{dg}(3) = 1, \text{dg}(4) = 2 \) and \( \text{dg}(5) = 1 \).

\( S \) also fulfills the requirements of a tree. Observe that in \( S \), we have \( \text{dg}(2) = 4 = |N^S| - 1 \). Thus, \( S \) is a spanning tree with root node 2. Nodes 1,3 and 5 are leafs in both \( T \) and \( S \).

\section{2.2 Regular vine tree sequences}

Based on the graph theoretic concepts introduced in the last section, we can now define an \textit{regular (R-vine) tree sequence}.

\textbf{Definition 2.9.} (Regular vine (R-vine) tree sequence) The set of trees \( \mathcal{V} = (T_1, \ldots, T_{d-1}) \) is a regular vine (R-vine) tree sequence on \( d \) elements if:

(i) Each tree \( T_k, k = 1, \ldots, d - 1 \), is connected,

(ii) \( T_1 = (N_1, E_1) \) with \( N_1 = \{1, \ldots, d\} \),

(iii) for \( k \geq 2 \) \( T_k = (N_k, E_k) \) with \( N_k = E_{k-1} \),

(iv) for \( k = 2, \ldots, d - 1 \) and \( e = (v_1, v_2) \in E_k \) it must hold that \( |v_1 \cap v_2| = 1 \).

\textbf{Remark 2.10.} (Proximity condition) Property (iv) is called \textit{proximity condition}. It ensures that if there is an edge \( e \in T_k \) connecting two nodes \( v_1, v_2 \in N_k \) then \( v_1 \) and \( v_2 \) (which are edges in \( T_{k-1} \)) must share a common node in \( T_{k-1} \).

We will now introduce a simplified edge notation for a regular vine tree sequence that will be used later to assign bivariate copulas to the regular vine construction of multivariate distributions.
Definition 2.11. (Complete union and conditioned sets)

Let \( V = (T_1, \ldots, T_{d-1}) \) be an R-vine tree sequence on \( d \) elements. For an edge \( e = (v_1, v_2) \in T_k \) define:

- The **complete union** of \( e \) as \( A_e := \{ v \in N_1 : \exists e_1 \in E_1, \ldots, e_{k-1} \in E_{k-1}, \text{ s.t.} \ v \in e_1 \in \cdots \in e_{k-1} \in e \} \),

- the **conditioning set** of \( e \) as \( D_e := \begin{cases} A_{v_1} \cap A_{v_2}, & \text{if } k > 1 \\ \emptyset, & \text{if } k = 1 \end{cases} \) and

- the **conditioned set** of \( e \) as \( P_e := A_{v_1} \setminus D_e \cup A_{v_2} \setminus D_e \).

Remark 2.12. (Notation of an edge \( e \))

Definition 2.9 ensures that for an edge \( e \in E_k : |A_e| = k+1, |P_e| = 2 \) and \( |D_e| = 1_{k \geq 2}(k-1) \). Therefore, we use the following notation for an edge \( e \in E_k \):

\[
e = (P_e|D_e) = \begin{cases} (P_{e,1}, P_{e,2}), & e \in E_1 \\ (P_{e,1}, P_{e,2}|D_{e,1}, \ldots, D_{e,k-1}), & e \in E_k \text{ for some } k = 2, \ldots, d-1. \end{cases}
\]

Using the presented notation, we can now define R-vine leafs in regular vine tree sequences. A node of the first tree \( T_1 \) is called **R-vine leaf**, if it is part of the conditioned set when ever it is element of the complete union of an edge. This is formalized in the following definition.

Definition 2.13. (R-vine leaf)

Let \( V = (T_1, \ldots, T_{d-1}) \) be an R-vine tree sequence on \( d \) elements. \( \ell \in N_1 \) is called **R-vine leaf** in \( V \), if for all \( k = 1, \ldots, d-1 \) and \( e \in E_k \) it holds that

\[ \ell \in A_e \Rightarrow \ell \in P_e. \]

Corollary 2.14 explains the connection between R-vine leafs and leafs of graphs, which were presented in Definition 2.2. Note that, since an edge set \( E_k \) becomes the node set \( N_{k+1}, \ k = 1, \ldots, d-2 \), we can also assign the notation presented in Definition 2.11 to nodes in trees \( T_2, \ldots, T_{d-1} \). For example, if we speak of the complete union of node \( v \in N_4 \), we refer to the complete union of the corresponding edge \( e \in E_3 \).

Corollary 2.14. (R-vine leafs are leafs in the corresponding trees)

Let \( \ell \in N_1 \) be an R-vine leaf in the regular vine tree sequence \( V = (T_1, \ldots, T_{d-1}) \). Then, the following holds:

(i) \( \ell \) is a leaf in \( T_1 \).

(ii) For \( k = 2, \ldots, d-1 \): if \( \ell \in A_v \), \( v \) is leaf in \( T_k \).

Proof:
(i) Assume that $\ell$ is not a leaf in $T_1$. Then, there are at least two neighbors $n_1, n_2 \in N_k$ of $\ell$, i.e. $\text{dg}(\ell) \geq 2$. Thus, edges $(\ell, n_1)$ and $(\ell, n_2)$ are part of the edge set $E_1$.

Now assume that $\ell \notin D_e$ for all $e \in \bigcup_{k=1}^{d-2} E_k$.

Then, we have two nodes $v_1$ and $v_2$ in $N_{d-1}$, of which the complete unions $A_{v_1}$ and $A_{v_2}$ contain $\ell$. Thus, the conditioning set of the connecting edge $(v_1, v_2) \in T_{d-1}$ is given by $A_{v_1} \cap A_{v_2} = \{\ldots, \ell, \ldots\}$. Therefore $\ell$ is not an R-vine leaf.

(ii) Let $k \in \{2, \ldots, d-1\}$ and let $\ell \in A_e$ for some node $v \in T_k$.

Assume that $v$ is not a leaf in $T_k$. Then, there are at least two neighbors $n_1, n_2 \in T_k$ of $v$, i.e. $\text{dg}(v) \geq 2$. Thus, edges $(v, n_1)$ and $(v, n_2)$ are part of the edge set $E_k$.

If we continue similarly to (i), we can follow that there exists an edge $e \in \bigcup_{s=k}^{d-1} E_s$, s.t. $\ell \in D_e$. Therefore, $\ell$ could not have been an R-vine leaf in the first place. □

Following the proof of Corollary 2.14, we know that in each tree of an R-vine tree sequence there can only be one edge of which the conditioned set contains an R-vine leaf, i.e.

for $k = 1, \ldots, d - 1$: there exists exactly one edge $e \in E_k$, s.t. $\ell \in P_e$.

As mentioned in Remark 2.12 conditioned sets consist of two elements only. Therefore we can define a unique **partner of an R-vine leaf** for each tree $T_k$, $k = 1, \ldots, d - 1$.

**Definition 2.15.** (Partner of an R-vine leaf)
Let $\ell$ be an R-vine leaf in $V = (T_1, \ldots, T_{d-1})$. For $k \in \{1, \ldots, d-1\}$ let $\ell \in P_e \subseteq N_1$ for some $e \in E_k$. Then the **partner of $\ell$ in tree $T_k$** is defined as $\text{pt}_k^\ell := P_e \setminus \{\ell\}$.

The following example shows a first regular vine tree sequence. It illustrates the notions presented in this section.

**Example 2.16.** (R-vine tree sequence on 5 elements)
Figure 2.2 shows an R-vine tree sequence on 5 elements $V = (T_1, \ldots, T_4)$, where $T_k = (N_k, E_k)$, $k = 1, \ldots, 4$, and the first node set is given by $N_1 = \{1, \ldots, 5\}$. 

29
The dotted lines indicate which edges are not possible by the proximity condition. For example in tree $T_2$ we cannot include edge $\{\{1, 2\}, \{3, 4\}\}$, since edges $(1, 2)$ and $(3, 4)$ do not share a common node in $T_1$. Whereas we could also choose an edge $\{\{1, 2\}, \{1, 5\}\}$ in $T_2$.

Contrary to 1,3 and 5, which are part of at least one conditioning set, 2 and 4 are R-vine leafs in $V$. For example, let us consider $\ell = 4$. The nodes of which the complete unions contain 4 are given by $(4)$, $(3, 4)$, $(1, 4|3)$ and $(4, 5|1, 3)$. Thus, 4 is not part of any conditioning set, but only conditioned sets. Further, all of the nodes are leafs in the corresponding tree, i.e. Corollary 2.14 holds. The partners of 4 are given by $pt_4^1 = 3$, $pt_4^2 = 1$, $pt_4^3 = 5$ and $pt_4^4 = 2$. We observe that $pt_{k_1}^4 \neq pt_{k_2}^4$ for all $k_1 \neq k_2 \in \{1, \ldots, d-1\}$. This observation is considered in the following corollary.
Corollary 2.17. (Partners of an R-vine leaf)
Let \( \ell \) be an R-vine leaf in \( V = (T_1, \ldots, T_{d-1}) \). Then:

(i) \( pt_1^\ell \in N_1 \setminus \{ \ell \} \).

(ii) For \( 2 \leq k \leq d - 1 \): \( pt_k^\ell \in N_1 \setminus \{ \ell, pt_1^\ell, \ldots, pt_{k-1}^\ell \} \).

Proof:

- \( k = 1 \):

\[
T_1 \quad \ell \quad pt_1^\ell \quad pt_1^\ell
\]

By Corollary 2.14, \( \ell \) is a leaf in \( T_1 \) and there exists a unique neighbor \( n_1 \in N_1 \setminus \{ \ell \} \) of \( \ell \). The corresponding connecting edge \( (\ell, n_1) \) contains only two elements \( \ell \) and \( n_1 \), i.e. \( pt_1^\ell = n_1 \in N_1 \setminus \{ \ell \} \).

- \( k = 2 \):

\[
v_2 = (\ell, pt_1^\ell), \text{ which is a leaf in } T_2, \text{ is the only node in } N_2 \text{ whose conditioned set contains } \ell. \text{ Let } n_2 \text{ be the unique neighbor of } v_2. \text{ By the proximity condition, we know that } A_{n_2} \cap A_{v_2} \neq \emptyset. \text{ Since } \ell \notin A_{n_2}, \text{ we can follow that } pt_1^\ell \in A_{n_2}. \text{ Therefore, the connecting edge } (v_2, n_2) \text{ equals } (\ell, pt_1^\ell) \text{ and } pt_2^\ell \in N_1 \setminus \{ \ell, pt_1^\ell \}.
\]

- \( 1, \ldots, k - 1 \to k \):

\[
T_k \quad \ell, pt_{k-1}^\ell | pt_1^\ell, \ldots, pt_{k-2}^\ell \quad \ell, pt_k^\ell | pt_1^\ell, \ldots, pt_{k-1}^\ell \quad pt_{k-1}^\ell, pt_k^\ell | pt_1^\ell, \ldots, pt_{k-2}^\ell
\]

Let \( pt_s^\ell \in N_1 \setminus \{ \ell, pt_1^\ell, \ldots, pt_{s-1}^\ell \} \) for all \( s = 1, \ldots, k - 1 \). From \( s = k - 1 \), we know that there is a node \( v_k \in N_k \) of which the conditioned set is given by \( P_{v_k} = \{ \ell, pt_{k-1}^\ell \} \), and which is a leaf in \( T_k \). Let \( n_k \) be the unique neighbor of \( v_k \) and remember that \( \ell \notin A_{n_k} \). Considering the proximity condition, we must have that \( A_{v_k} \cap A_{n_k} = k - 1 \). Therefore, there must be exactly one element of \( A_{n_k} \) which is not in \( A_{v_k} \) and will be part of the conditioned set \( P_e \) of the connecting edge \( e = (v_k, n_k) \). This element is defined to be \( pt_k^\ell \). Thus, we know that \( pt_k^\ell \in N_1 \setminus \{ \ell, pt_1^\ell, \ldots, pt_{k-1}^\ell \} \).
Example 2.18. (Illustration of Corollary 2.17)
We consider again \( V = (T_1, \ldots, T_4) \) shown in Figure 2.2. As already mentioned in Example 2.16, 2 and 4 are R-vine leafs in \( V \).
As expected, Corollary 2.17 holds for Example 2.18.
(Illustration of Corollary 2.17)

We define the R-vine leaf \( \{ 2, 3, 4 \} \). Similarly, we define \( \{ 2, 4 \} \).

Further observe that, if we remove all edges \( e \) for which the complete union \( A_e \) contains the R-vine leaf \( \ell = 4 \), we still obtain a valid R-vine tree sequence on 4 elements. This result holds for all R-vine leafs, as shown Lemma 2.19.

We define \( v_k^\ell \in N_k \) to be the node in tree \( T_k \) for which the complete union \( A_{v_k^\ell} \) contains the R-vine leaf \( \ell \). Similarly, we define \( e_k^\ell \in E_k \) to be the edge in tree \( T_k \) with \( \ell \in A_{v_k^\ell} \).

Obviously, it holds that \( n_k = e_k^{\ell - 1} \) for all \( k = 2, \ldots, d - 1 \).

Lemma 2.19. (Removing R-vine leafs produces again an R-vine tree sequence)
Let \( \ell \) be an R-vine leaf in \( V = (T_1, \ldots, T_{d-1}) \).
Define \( V^{-\ell} = (T_1^{-\ell}, \ldots, T_{d-2}^{-\ell}) \) to be the subvine of \( V \) with all nodes and edges containing \( \ell \) removed, i.e. \( T_k^{-\ell} = (N_k^{-\ell}, E_k^{-\ell}) \) with \( N_k^{-\ell} = N_k \setminus \{ v_k^\ell \} \) and \( E_k^{-\ell} = E_k \setminus \{ e_k^\ell \} \) with \( v_k^\ell \) and \( e_k^\ell \) being the node (edge) in tree \( T_k \) for which \( \ell \in A_{v_k^\ell} \) (\( \ell \in A_{e_k^\ell} \)), \( k = 1, \ldots, d - 2 \).

Then, \( V^{-\ell} \) is a regular vine tree sequence on \( d - 2 \) elements.

Proof:
First, observe that for all \( k = 1, \ldots, d - 2 \): \( T_k^{-\ell} \) is a tree, since removing a leaf in \( T_k \) does not affect the existence or uniqueness of paths between any pair of nodes \( v_1, v_2 \in N_k \), i.e. (ii) in Definition 2.6 is fulfilled for any \( T_k^{-\ell} \).

Now, we prove that \( V^{-\ell} \) satisfies the definition of a regular vine tree sequence:

(i) For all \( k = 1, \ldots, d - 2 \): \( T_k^{-\ell} \) is connected, since removing \( v_k^\ell \) and \( e_k^\ell \) does not affect nodes \( v \in N_k^{-\ell} \) or edges \( e \in E_k^{-\ell} \).
(ii) Without loss of generality we can assume that \( N_1 = \{1, \ldots, d-1, \ell\} \) and it follows that \( N_1^{-\ell} = N_1 \setminus \{\ell\} = \{1, \ldots, d-1\} \).

(iii) Let \( k \geq 2 \): \( N_k^{-\ell} = N_k \setminus \{v_k^\ell\} = E_{k-1} \setminus \{e_{k-1}^\ell\} = E_{k-1}^{-\ell} \).

(iv) Let \( k \in \{2, \ldots, d-1\} \) and let \( e = (v_1, v_2) \in E_k^{-\ell} \subseteq E_k \) for some nodes \( v_1, v_2 \in N_k^{-\ell} \subseteq N_k \). \( \Rightarrow |v_1 \cap v_2| = 1 \), since \( e \in E_k \).

Finally, we introduce two special types of regular vine tree sequences. Namely they are **drawable vine (D-vine)** and **canonical vine (C-vine) tree sequences**. A special characteristic of D-vines is that they are uniquely defined by the first tree \( T_1 \), i.e. if we now \( T_1 \), we can directly follow all other trees in a D-vine tree sequence.

Both D- and C-vine tree sequences are defined in Definition 2.20 and illustrated in Figure 2.3.

**Definition 2.20.** (C-vine and D-vine tree sequences)

An R-vine tree sequence \( \mathcal{V} = (T_1, \ldots, T_{d-1}) \) is called

- **C-vine tree sequence**, if for all \( k \in \{1, \ldots, d-1\} \), \( T_k \) is a star tree.
- **D-vine tree sequence**, if for all \( k \in \{1, \ldots, d-1\} \) and \( v \in N_k \), it holds: \( d(v) \leq 2 \).
2.3 R-vine distributions and R-vine copulas

In the following, we want to model the dependence structure between $d$ variables $U_1, \ldots, U_d$. Therefore, we link parametric bivariate copulas to regular vine tree sequences on $d$ elements $U_1, \ldots, U_d$, making use of the so-called pair copula construction (PCC), which was introduced by Joe [1996] and Bedford and Cooke [2002b].

Example 2.21 illustrates a pair copula construction for three variables $U_1, U_2, U_3$. 

Figure 2.3: Example of an C- and D-vine tree sequence on 5 elements.
Example 2.21. (Pair copula construction of a three dimensional parametric copula)
Assume that we are given three variables \( U_1, U_2, U_3 \) on the \( u \)-scale and the following distributions are known (e.g. by the selection criterion optimizing parametric copulas (Definition 1.36)):

- \( (U_1, U_2) \sim C_{U_1, U_2}(\cdot; \theta_{U_1, U_2}) \),
- \( (U_3, U_2) \sim C_{U_3, U_2}(\cdot; \theta_{U_3, U_2}) \),
- \( (U_1|U_2 = u_2, U_3|U_2 = u_2) \sim C_{U_1|U_2, U_3}(\cdot; \theta_{U_1|U_2, U_3}; u_2) \).

The density corresponding to the three dimensional parametric copula
\[
C_{U_1, U_2, U_3}(\cdot; \cdot; \cdot; \Theta_{U_1, U_2, U_3}),
\]
i.e. the joint distribution of \((U_1, U_2, U_3)\), is given by
\[
c_{U_1, U_2, U_3}(u_1, u_2, u_3; \Theta_{U_1, U_2, U_3}) = c_{U_1, U_2}(u_1, u_2; \theta_{U_1, U_2}) \times c_{U_3, U_2}(u_3, u_2; \theta_{U_3, U_2}) \times c_{U_1|U_2, U_3}(h_{U_1|U_2}(u_1|u_2), h_{U_3|U_2}(u_3|u_2); \theta_{U_1|U_2, U_3}; u_2),
\]
where \( B_{U_1, U_2, U_3} = \{b_{U_1, U_2}, b_{U_2, U_3}, b_{U_1, U_3}\} \) and \( \Theta_{U_1, U_2, U_3} = \{\theta_{U_1, U_2}, \theta_{U_2, U_3}, \theta_{U_1, U_3}\} \) are sets of the bivariate copula families and parameter vectors, respectively, and
\[
h_{U_j|U_2}(u_j|u_2) := \frac{\partial}{\partial u_2} c_{U_j, U_2}(u_j, u_2; \theta_{U_j, U_2}), \quad j = 1, 3,
\]
are the \( h \)-functions of \( U_j \) given \( U_2 = u_2 \), \( j = 1, 3 \).

Note that the pair copula construction of the joint copula density is not unique, since it depends on the available bivariate copulas and the underlying R-vine structure.

Now, we extend the pair copula approach to \( d \) variables \( U_1, \ldots, U_d \) and introduce the so called \textbf{regular vine distribution} (Definition 2.22), which in general is defined for random vectors on the \( x \)-scale \( X_1, \ldots, X_d \sim F \). Note that R-vine distributions use a regular vine tree sequence \( \mathcal{V} \) on \( d \) elements with \( N_1 = \{X_1, \ldots, X_d\} \) to define a unique pair copula decomposition. Thus, conditioned and conditioning sets of an edge \( e = (P_1|D) \in \bigcup_{k=1}^{d-1} E_k \) are now subsets of \( \{X_1, \ldots, X_d\} \).

**Definition 2.22.** (Regular vine distributions and pair copulas)
Let \( X = (X_1, \ldots, X_d) \sim F \). \( F \) is called R-vine distribution, if there is a triplet \( (\mathcal{F}, \mathcal{V}, \mathcal{C}(\mathcal{V})) \) s.t.:

(i) \( \mathcal{F} = (F_1, \ldots, F_d) \), s.t. \( X_j \sim F_j, F_j \) continuous and invertible, \( j = 1, \ldots, d \).

(ii) \( \mathcal{V} \) is an R-vine tree sequence on \( d \) elements \( U_j := F_j(X_j), \ j = 1, \ldots, d \).
(iii) $\mathcal{C}(\mathcal{V}) := \{ C^{\text{be}}(\cdot, \cdot; \theta_e) : e \in E_k; k = 1, \ldots, d - 1 \}$ where $C^{\text{be}}(\cdot, \cdot; \theta_e)$ is a bivariate copula with density.

(iv) For each $e = (P_{e,1}, P_{e,2}) \in E_1$ it holds that $(P_{e,1}, P_{e,2}) \sim C^{\text{be}}(\cdot, \cdot; \theta_e)$.

(v) For each $e = (P_e|D_e) \in E_k$, $k = 2, \ldots, d - 1$, $C^{\text{be}}(\cdot, \cdot; \theta_e; u_{D_e})$ is the bivariate parametric copula corresponding to the conditional distribution of $P_{e,1}$ and $P_{e,2}$ given the random variables $D_e = u_{D_e}$.

$C^{\text{be}}(\cdot, \cdot; \theta_e; u_{D_e})$ is called **parametric pair copula** corresponding to edge $e$.

We speak of a pair copula decomposition, when the bivariate parametric pair copulas associated with the bivariate conditional distribution of $(U_{P_{e,1}}, U_{P_{e,2}})$ is allowed to depend on the specific value of the underlying conditioning variable $u_{D_e}$. Throughout the thesis, we will assume that this dependence can be ignored, i.e. we assume that

$$C^{\text{be}}(\cdot, \cdot; \theta_e; u_{D_e}) \equiv C^{\text{be}}(\cdot, \cdot; \theta_e),$$

for all $u_{D_e} \in \mathbb{R}^{|D_e|}$.

This is referred to the **simplifying assumption** (Bedford and Cooke [2001]). Note that this assumptions was already made in Example 2.21. Even though the simplifying assumption has been criticized in various papers and journals, it is often assumed to be reasonable. Haff et al. [2010] stated that simplified vine copulas are “a rather good solution, even when the simplifying assumption is far from being fulfilled by the actual model”.

As already seen in Example 2.21, pair copulas corresponding to edges $e \in \bigcup_{k=2}^{d-1} E_k$, take conditional distribution functions as arguments. Those conditional distributions are $h$-functions, which were introduced in Definition 1.10. The following definition connects $h$-functions to the concept of pair copulas.

**Definition 2.23.** (h-functions of parametric pair copulas)

Let $C^{\text{be}}(\cdot, \cdot; \theta_e)$ be a parametric pair copula corresponding to an edge $e = (P_e|D_e) \in \bigcup_{k=2}^{d-1} E_k$.

Then we define

$$h_{P_{e,1}|P_{e,2};D_e}(w_1|w_2) := \frac{\partial}{\partial w_2} C^{\text{be}}(w_1, w_2; \theta_e)$$

and

$$h_{P_{e,2}|P_{e,1};D_e}(w_2|w_1) := \frac{\partial}{\partial w_1} C^{\text{be}}(w_1, w_2; \theta_e).$$

Now we consider $d$ variables $U_1, \ldots, U_d$ on the u-scale and want to model the dependence structure within the given variables. One could model it using an R-vine distribution with uniform margins. However, we introduce a more convenient notation in Definition 2.24.

**Definition 2.24.** (Regular vine copula)

Let $(\mathcal{F}, \mathcal{V}, \mathcal{C}(\mathcal{V}))$ be an R-vine distribution corresponding to the variables $U_1, \ldots, U_d$, with
$U_j \sim U(0,1), j = 1, \ldots, d$. Then $C(V)$ can be uniquely specified by the \textit{set of pair copula families}

$$B(V) := \{b_e := b(C_e) : C_e \in C(V)\}$$

and the \textit{set of pair copula parameters}

$$\Theta(B(V)) := \{\theta_e := \theta(b(C_e)) : C_e \in C(V)\}.$$ 

We can define an \textbf{R-vine copula} $\mathcal{R}$ by

$$\mathcal{R} = (V, B(V), \Theta(B(V))).$$

\textbf{Remark 2.25.} (C- and D-vine copula)
If a regular vine copula $\mathcal{R}$ is based on a C-vine tree sequence $V$, we call it C-vine copula. If it is based on a D-vine tree sequence, it is called D-vine copula.

If we want to estimate a regular vine copula out of given observations $u_1, \ldots, u_d \in [0,1]^n$, we find the bivariate pair copulas by optimizing a chosen selection criterion $sc$ (see Definition 1.36). Therefore, we need to specify the copula data $U^n_e$, which are used for the estimation of a pair copula corresponding to edge $e$.

For the estimation of parametric copulas $C^{b_{u_i,v_j}}(\cdot, \cdot; \theta_e)$, corresponding to edges $e = (U_i, U_j) \in E_1$, we can directly set

$$U^n_e := (u_i, u_j) \in [0,1]^{n \times 2}, \; i \neq j \in \{1, \ldots, d\}.$$ 

However, the copula data used to estimate pair copulas corresponding to edges in trees with tree level two or higher, need to be estimated in advance. Those \textit{pseudo observations} are defined in the following.

\textbf{Definition 2.26.} (Pseudo observations)
Let $C^{b_e}(\cdot, \cdot; \hat{\theta}_e)$ be an estimated parametric pair copula corresponding to edge $e = (P_e|D_e)$. Additionally assume that we are given copula data $U^n_e := (u_1, u_2) \in [0,1]^{n \times 2}$, where $u_j := (u_{1j}, \ldots, u_{nj})^T \in [0,1]^n$ for $j \in \{1, 2\}$. The estimated versions of the $h$-functions $h_{P_e,1|P_e,2;D_e}(\cdot ; \hat{\theta}_e)$ and $h_{P_e,1|P_e,2;D_e}(\cdot ; \hat{\theta}_e)$ are called pseudo observations and are defined as follows:

$$h_1(U^n_e, C^{b_e}(\cdot, \cdot; \hat{\theta}_e)) := h_1(U^n_e, C^{b_e}(\cdot, \cdot; \hat{\theta}_e))$$

$$h_2(U^n_e, C^{b_e}(\cdot, \cdot; \hat{\theta}_e)) := h_2(U^n_e, C^{b_e}(\cdot, \cdot; \hat{\theta}_e)),$$

where

$$h_1(U^n_e, C^{b_e}(\cdot, \cdot; \hat{\theta}_e)) := \left(\frac{\partial}{\partial x_2} C^{b_e}(u_{11}, x_2; \hat{\theta}_e)\big|_{x_2 = u_{12}}, \ldots, \frac{\partial}{\partial x_2} C^{b_e}(u_{n1}, x_2; \hat{\theta}_e)\big|_{x_2 = u_{n2}}\right)^T \in [0,1]^n,$$

$$h_2(U^n_e, C^{b_e}(\cdot, \cdot; \hat{\theta}_e)) := \left(\frac{\partial}{\partial x_1} C^{b_e}(x_1, u_{12}; \hat{\theta}_e)\big|_{x_1 = u_{11}}, \ldots, \frac{\partial}{\partial x_1} C^{b_e}(x_1, u_{n2}; \hat{\theta}_e)\big|_{x_1 = u_{n1}}\right)^T \in [0,1]^n.$$
The number of estimated pseudo observation vectors might get quite large in a regular vine copula. In order to keep an overview on those vectors, we group them in sets of **pseudo observations** \( \mathcal{P}_k \) which are assigned to each tree \( T_k, k = 1, \ldots, d - 1 \).

**Definition 2.27.** (Set of pseudo observations corresponding to tree \( T_k \) in \( V \)) Let \( V = (T_1, \ldots, T_{d-1}) \) be an R-vine tree sequence on \( d \) elements. Then, we denote the sets of pseudo observations corresponding to tree \( T_k = (N_k, E_k), k = 1, \ldots, d - 1 \), by

(i) \( \mathcal{P}_1 := \bigcup_{e \in E_1} \{ \hat{h}_{P_{e,1}}|P_{e,2} \cup \hat{h}_{P_{e,2}}|P_{e,1} \} \),

(ii) \( \mathcal{P}_k := \bigcup_{e \in E_k} \{ \hat{h}_{P_{e,1}}|P_{e,2};D_e \cup \hat{h}_{P_{e,2}}|P_{e,1};D_e \}, k = 2, \ldots, d - 1 \).

We close this section by giving a first example of an estimated regular vine copula. It is based on the regular vine tree sequence presented in Example 2.16.

**Example 2.28.** (Regular vine copula on 5 variables) Let \( U_1, \ldots, U_5 \) be uniformly distributed random variables. The goal is to describe the dependence structure within the set of variables using a regular vine copula \( \mathcal{R} = (\mathcal{V}, \mathcal{B}(\mathcal{V}), \Theta(\mathcal{B}(\mathcal{V}))) \).

\( \mathcal{V} = (T_1, \ldots, T_4) \) is given by the regular vine tree sequence shown in Example 2.16 with \( j := U_j, j = 1, \ldots, 5 \). Then, it remains to find the set of copula families \( \mathcal{B}(\mathcal{V}) \) and the corresponding parameter set \( \Theta(\mathcal{B}(\mathcal{V})) \), i.e. we need to find a parametric pair copula for every edge \( e \in \bigcup_{k=1}^{4} E_k \).

- \( k = 1 \):
  The edge set in \( T_1 \) is given by
  \[
  E_1 = \{(U_1, U_2), (U_1, U_3), (U_1, U_5), (U_3, U_4)\}.
  \]

We start with estimating copula family \( \hat{b}_e \) and parameter vector \( \hat{\theta}_e \) for all \( e \in E_1 \). For example, if we are given observations \( u_j \in [0, 1]^n \) of variable \( U_j \) for all \( j \in \{1, \ldots, 5\} \), we can obtain unique estimates by choosing the AIC-optimal bivariate parametric copulas for some fixed set of possible copula families \( B \), i.e.

\[
C^{b_{U_i,U_j}}(\cdot, \cdot; \hat{\theta}_{U_i,U_j}) := \text{optC}((u_i, u_j), B, \text{AIC}), (i, j) \in \{(1, 2), (1, 3), (1, 5), (3, 4)\}.
\]

We use the estimated copulas to estimate two vectors of pseudo observations based on the original observation vectors \( u_j \) for each \( e \in E_1 \), i.e.

\[
\hat{h}_{U_i|U_j} = h_1((u_i, u_j), C^{b_{U_i,U_j}}(\cdot, \cdot; \hat{\theta}_{U_i,U_j})) \in [0, 1]^n,
\]

\[
\hat{h}_{U_j|U_i} = h_2((u_i, u_j), C^{b_{U_i,U_j}}(\cdot, \cdot; \hat{\theta}_{U_i,U_j})) \in [0, 1]^n.
\]

All \( 4 \cdot 2 = 8 \) vectors are added to the set \( \mathcal{P}_1 \).
- $k = 2$:
  We continue with considering all edges in

  \[ E_2 = \{(U_2, U_3|U_1), (U_3, U_5|U_1), (U_1, U_4|U_3)\}. \]

  The main difference to $k = 1$, is that we want to find conditional copulas. In the pair copula construction, the arguments of conditional copulas are given by conditional distributions, i.e. the corresponding $h$-functions. For example, we have

  \[
  c_{U_1, U_2, U_3}^{B_{U_1, U_2, U_3}}(u_1, u_2, u_3; \Theta_{U_1, U_2, U_3}) = c_{U_1, U_2}^{B_{U_1, U_2}}(u_1, u_2; \hat{\Theta}_{U_1, U_2}) \times c_{U_3}^{B_{U_3}}(u_3; \hat{\Theta}_{U_3})
  \]

  therefore we need to base each estimation on the corresponding pseudo observations out of the set $P_1$. For example for $e = (U_2, U_3|U_1)$ we estimate

  \[
  C_{U_2, U_3|U_1}^{\hat{B}_{U_2, U_3|U_1}}(\cdot, \cdot; \hat{\Theta}_{U_2, U_3|U_1}) = \text{optC}(\hat{h}_{U_2|U_1}, \hat{h}_{U_3|U_1}, B, \text{AIC})
  \]

  by minimizing the AIC.

  After we know the pair copulas corresponding to all edges in $E_2$, we can estimate the $3 \times 2 = 6$ pseudo observation vectors in $P_2$. For example, we set

  \[
  \hat{h}_{U_2|U_3; U_1} = h_1(\hat{h}_{U_2|U_1}, \hat{h}_{U_3|U_1}, C_{U_2, U_3|U_1}^{\hat{B}_{U_2, U_3|U_1}}(\cdot, \cdot; \hat{\Theta}_{U_2, U_3|U_1})) \in [0, 1]^n,
  \]

  \[
  \hat{h}_{U_3; U_2; U_1} = h_2(\hat{h}_{U_2|U_1}, \hat{h}_{U_3|U_1}, C_{U_2, U_3|U_1}^{\hat{B}_{U_2, U_3|U_1}}(\cdot, \cdot; \hat{\Theta}_{U_2, U_3|U_1})) \in [0, 1]^n.
  \]

- $k \in \{3, 4\}$:
  We approach similarly for edge sets

  \[
  E_3 = \{(U_2, U_5|U_1, U_3), (U_4, U_5|U_1, U_3)\} \quad \text{and} \quad E_4 = \{(U_2, U_4|U_1, U_3, U_5)\}.
  \]

  After estimating pair copulas for all edges $e \in \bigcup_{k=1}^4 E_k$ we can define the estimated regular vine copula

  \[
  \hat{\mathcal{R}} := (\mathcal{V}, \hat{\mathcal{B}}(\mathcal{V}), \hat{\Theta}(\hat{\mathcal{B}}(\mathcal{V}))), \quad \text{where}
  \]

  \[
  \hat{\mathcal{B}}(\mathcal{V}) := \{\hat{b}_e : e \in \bigcup_{k=1}^4 E_k\} \quad \text{and} \quad \hat{\Theta}(\hat{\mathcal{B}}(\mathcal{V})) := \{\hat{\theta}_e : e \in \bigcup_{k=1}^4 E_k\}.
  \]

  Note that $|\bigcup_{k=1}^4 E_k| = 4 + 3 + 2 + 1 = 10$. Therefore, both $\hat{\mathcal{B}}(\mathcal{V})$ and $\hat{\Theta}(\hat{\mathcal{B}}(\mathcal{V}))$ contain 10 elements.
Chapter 3

Quantile Regression

Quantile regression is used to estimate the quantiles of a response variable’s conditional distribution conditioned on a set of covariates. Since its first appearance in Koenker and Bassett Jr [1978], it has gained importance in various statistical fields. In this chapter we follow Koenker [2005], which is an updated version of Koenker’s first book on quantile regression. We add some theory introduced in Kraus and Czado [2017] and Padellini and Rue [2018].

3.1 Conditional quantile- and check loss function

Before we introduce regression methods, we define the quantile-function or \( \alpha \)-quantiles of a random variable \( X \sim F_X \), which is basically the inverse of a random variable’s distribution function. Since in general the inverse is not unique, we need to take the infimum of the inverse image. Typically, one considers quantiles for \( \alpha \in \{0.05, 0.25, 0.5, 0.75, 0.95\} \).

**Definition 3.1. (\( \alpha \)-Quantile)**

Let \( X \sim F_X \). For all \( \alpha \in (0, 1) \), the \( \alpha \)-Quantile is defined as

\[
q_X(\alpha) := F_X^{-1}(\alpha) := \inf\{x : F_X(x) \geq \alpha\}
\]

Let us now consider

- a response variable \( Y \) and
- a set of \( d \) covariates \( X_1, \ldots, X_d \).

Then, we can define the conditional quantile function of \( Y \) given the covariates \( X_1, \ldots, X_d \) as the quantile function of the conditional distribution of \( Y|X_1, \ldots, X_d \).

**Definition 3.2. (Conditional quantile function)**

Let \( Y|X_1, \ldots, X_d \sim F_{Y|X_1,\ldots,X_d} \). The conditional quantile function of \( Y \) given \( X_1, \ldots, X_d \) is defined by

\[
Q_{Y|X_1,\ldots,X_d}(\alpha|x_1, \ldots, x_d) := F_{Y|X_1,\ldots,X_d}^{-1}(\alpha|X_1 = x_1, \ldots, X_d = x_d).
\]
Regression methods are designed to explain the behavior of $Y$ as a deterministic function of the covariate values $x_1, \ldots, x_d$, with $X_1 = x_1, \ldots, X_d = x_d$, and some additive random error $\varepsilon$, i.e.

$$Y = r(x_1, \ldots, x_d) + \varepsilon.$$ 

For example, in ordinary linear regression we choose

$$r(x_1, \ldots, x_d) := \beta_0 + \sum_{j=1}^d \beta_j x_j, \quad \beta_j \in \mathbb{R}, \; j = 0, 1, \ldots, d \quad \text{and} \quad \varepsilon \sim N(0,1).$$

The regression function $r(x_1, \ldots, x_d)$ is a summary of the conditional distribution of $Y|X_1 = x_1, \ldots, X_d = x_d$ chosen to minimize the expected loss (or risk) occurring when we neglect the error term $\varepsilon$ to explain $Y$, or, in other words, $r(\cdot)$ must be such that $r(X_1, \ldots, X_d)$ is, on average, “close” to $Y$. In ordinary linear regression the risk is described by the quadratic loss function

$$L(Y, r(X_1, \ldots, X_d)) = (Y - r(X_1, \ldots, X_d))^2.$$ 

The regression function, minimizing the expected quadratic risk, is the conditional mean, i.e.

$$\arg \min_{r(x_1, \ldots, x_d)} \mathbb{E}[(Y - r(x_1, \ldots, x_d))^2] = \mathbb{E}[Y|X_1 = x_1, \ldots, X_d = x_d] = \beta_0 + \sum_{j=1}^d \beta_j x_j.$$ 

If we choose another loss function, we will end up in a regression function different from the conditional mean. Koenker [2005] introduces the so called check (or pinball) loss function (Definition 3.3) and proves that the regression function minimizing the expected check loss is given by the conditional quantile function, i.e.

$$\arg \min_{r(x_1, \ldots, x_d)} \mathbb{E}[\rho_\alpha(Y - r(x_1, \ldots, x_d))] = Q_{Y|X_1, \ldots, X_d}(\alpha|x_1, \ldots, x_d), \; \text{for all} \; \alpha \in (0,1).$$

**Definition 3.3.** (Check loss function $\rho_\alpha(\cdot)$)

For $\alpha \in (0,1)$ we define the loss function as

$$\rho_\alpha(x) := x(\alpha - 1_{\{x<0\}}).$$

Figure 3.1 shows the graphs of the check loss function $\rho_\alpha(x)$ for

$$\alpha \in \{0.05, 0.25, 0.50, 0.75, 0.95\} \quad \text{and} \quad x \in [-1,1].$$
Figure 3.1: Plot of $\{\varrho_\alpha(x) : x \in [-1,1]\}_{\alpha \in \{0.05,0.25,0.50,0.75,0.95\}}$.

### 3.2 Linear quantile regression

Koenker and Bassett Jr [1978] introduced the first quantile regression model. Similarly to ordinary linear regression, the regression function is modeled by a linear function of the covariates and deterministic coefficients depending on the quantile level $\alpha$.

**Definition 3.4.** (Linear quantile regression model)

Let $Y$ be the response variable and let $X_1, \ldots, X_d$ be the set of covariates. For every $\alpha \in (0,1)$ the conditional quantiles of $Y$ given $X_1 = x_1, \ldots, X_d = x_d$ are modeled as

$$Q_{Y|X_1,\ldots,X_d}(\alpha|x_1,\ldots,x_d) := \beta_0(\alpha) + \sum_{j=1}^d \beta_j(\alpha)x_j.$$ 

Given observations $y = (y_1, \ldots, y_n)^T$, $x_j = (x_{1j}, \ldots, x_{nj})^T \in \mathbb{R}^n$ of the response $Y$ and the covariates $X_j$, $j = 1, \ldots, d$, we can estimate the coefficients

$$\hat{\beta}(\alpha) = (\hat{\beta}_0(\alpha), \hat{\beta}_1(\alpha), \ldots, \hat{\beta}_d(\alpha))^T \in \mathbb{R}^{d+1}$$

for every $\alpha \in (0,1)$. We find them by minimizing the empirical risk, which is given by the empirical mean of the check losses.
Definition 3.5. (Parameter estimation in linear quantile regression)
Let \( y = (y_1, \ldots, y_n)^T, x_j = (x_{1j}, \ldots, x_{nj})^T \in \mathbb{R}^n \) be observations of the response \( Y \) and the covariates \( X_j, j = 1, \ldots, d \). For every \( \alpha \in (0, 1) \), the estimated regression coefficients are defined as

\[
\hat{\beta}(\alpha) := \arg\min_{\beta(\alpha)} n \sum_{i=1}^n \rho_\alpha(y_i - x_i \beta(\alpha)),
\]

where \( x_i := (x_{i1}, \ldots, x_{id}) \).

Solving this optimization problem requires linear programming techniques, such as the simplex method or interior point methods. They can be found in Koenker [2005].

Shortfalls of linear quantile regression
Koenker [2005] criticized his method for allowing for quantile crossing (Definition 3.6).

Definition 3.6. (Quantile crossing)
Let \( \hat{Q}_{Y|X_1,\ldots,X_p}(\cdot|\cdot) \) be the predicted quantile function of a response variable \( Y \) given the covariates \( X_1, \ldots, X_p \). We speak of quantile crossing in a point \( x := (1, x_1, \ldots, x_p)^T \in \mathbb{R}^{p+1} \), if

\[
\hat{Q}_{Y|X_1,\ldots,X_p}(\alpha_1|x) > \hat{Q}_{Y|X_1,\ldots,X_p}(\alpha_2|x)
\]

for some \( 0 < \alpha_1 < \alpha_2 < 1 \).

He did prove that the predicted conditional quantile function for linear quantile regression is non-decreasing in the centroid of the design space, i.e.

\[
\bar{x}^T \hat{\beta}(\alpha_1) \leq \bar{x}^T \hat{\beta}(\alpha_2) \text{ for all } 0 < \alpha_1 < \alpha_2 < 1,
\]

where \( \bar{x} := \frac{1}{n} \sum_{i=1}^n x_i = (1, \frac{1}{n} \sum_{i=1}^n x_{i1}, \ldots, \frac{1}{n} \sum_{i=1}^n x_{ip})^T \).

However, it is obvious that there must be quantile crossing sufficiently far away from \( \bar{x} \), since the conditional quantile function is linear in the beta coefficients, which in general differ for different values of \( \alpha \). If there are a lot of points in \( \{ x \in \mathbb{R}^{p+1} : \exists 0 < \alpha_1 < \alpha_2 < 1, \text{ s.t. } x^T \hat{\beta}(\alpha_1) > x^T \hat{\beta}(\alpha_2) \} \), the linear quantile regression model can be considered misspecified.

Figure 3.2 shows an example of quantile crossing when linearly regressing the quantiles of the response variable \( V =: Shuck \) on the covariate \( U =: Whole \) from the abalone data set.

A more detailed description of the data can be found in Chapter 8. While the estimated 0.95-quantile line crosses all other estimated lines out of the range of the covariate \( Whole \), the others cross each other in the range of \( Whole \).
Figure 3.2: Estimated conditional quantiles of Shuck given Whole for $\alpha \in \{0.05, 0.25, 0.50, 0.75, 0.95\}$; the vertical line indicates the minimum of the observations of Whole within the considered data.

Further, classical linear quantile regression has been criticized by Bernard and Czado [2015] for imposing too restrictive assumptions on the shape of the regression quantiles. For example, they have shown that for Gaussian marginals the linearity assumption, i.e. assuming $Q(Y|X_1, \ldots, X_d)(\alpha|x_1, \ldots, x_d)$ to be linear in $x_1, \ldots, x_d$ with deterministic slope, only holds if we can describe the underlying dependence structure by a Gaussian copula.

Additionally, the method suffers from typical shortfalls of linear models such as multicollinearity, selection and significance of covariates and the inclusion of interactions or transformed variables.
Part II

Regular vine based quantile regression
Chapter 4

Regular vine copulas for quantile regression

In contrast to linear quantile regression, which directly models the conditional quantile function, regular vine based quantile regression models aim to estimate the conditional distribution of a response variable \( Y \) given a set of covariates \( X_1, \ldots, X_d \). This conditional distribution is found with the help of regular vine tree sequences and the corresponding parametric pair copulas.

As mentioned before, pair copulas are only defined on the unit square. Therefore, we need to transform all variables to the u-scale by the probability integral transform. For now, we assume that all marginal distributions, \( F_Y, F_{X_j}, j = 1, \ldots, d \), are known and introduce the following notation:

- response variable on the u-scale: \( V := F_Y(Y) \),
- covariates on the u-scale: \( U_j := F_{X_j}(X_j), j = 1, \ldots, d \).

Using those probability integral transforms with the corresponding PIT values \( v := F_Y(y) \) and \( u_j := F_{X_j}(x_j), j = 1, \ldots, d \), it follows that

\[
F_{Y|X_1, \ldots, X_d}(y|x_1, \ldots, x_d) = \mathbb{P}(Y \leq y|X_1 = x_1, \ldots, X_d = x_d) \\
= \mathbb{P}(V \leq v|U_1 = u_1, \ldots, U_d = u_d) \\
= C_{V|U_1, \ldots, U_d}(v|u_1, \ldots, u_d),
\]

where \( C_{V|U_1, \ldots, U_d}(\cdot|u_1, \ldots, u_d) \) denotes the conditional distribution of \( V \) given that \( U_j = u_j \) for all \( j \in \{1, \ldots, d\} \).

Then, we find the conditional \( \alpha \)-quantiles of \( Y \) by inverting the expression, i.e.

\[
F^{-1}_{Y|X_1, \ldots, X_d}(\alpha|x_1, \ldots, x_d) = F^{-1}_Y(C^{-1}_{V|U_1, \ldots, U_d}(\alpha|u_1, \ldots, u_d)).
\]

**Definition 4.1.** (Conditional quantile function in regular vine based quantile regression)
Let $Y \sim F_Y$, $X_j \sim F_{X_j}$, $j = 1, \ldots, d$. Then, the conditional quantile function in regular vine based quantile regression of $Y$ given $X_1, \ldots, X_d$ is defined as
\[
Q_{Y|X_1,\ldots,X_d}(\alpha|x_1,\ldots,x_d) := F_{Y}^{-1}(C_{V_1,\ldots,V_d}^{-1}(\alpha|u_1,\ldots,u_d)),
\]
where $V := F_Y(Y)$, $v := F_Y(y)$, $U_j := F_{X_j}(X_j)$, $u_j := F_{X_j}(x_j)$, $j = 1, \ldots, d$ and $C_{V_1,\ldots,V_d}(\cdot|u_1,\ldots,u_d)$ denotes the conditional distribution of $V$ given $U_1, \ldots, U_d$

If we want to estimate the quantile function using an R-vine copula (Definition 2.24), the main task is to estimate the conditional distribution of $V$ given the covariates $U_1, \ldots, U_d$, i.e. $\hat{C}_{V_1,\ldots,V_d}(\cdot|u_1,\ldots,u_d)$. In order to be able to analytically express conditional distributions in terms of pair copulas, we present an important theorem (Theorem 4.2), proven by Joe [1996].

**Theorem 4.2.** (Recursion for conditional distribution functions)
Let $(Y, X) \sim F$ for some absolute continuous joint distribution $F$, a random variable $Y$ and a random vector $X$. Let $X_j \in X$ and define $X_{-j} := X \setminus X_j$ as the subset of $X$ with $X_j$ removed. Then $F_{Y|X}$, the conditional distribution of $Y$ given $X = x$, satisfies the following recursion:
\[
F_{Y|X}(\cdot|x) = \frac{\partial C_{Y,X_j|X_{-j}}(F_{Y|X_{-j}}(y|x_{-j}), F_{X_j|X_{-j}}(y|x_{-j}))}{\partial F_{X_j|X_{-j}}(x_j|x_{-j})},
\]
where $C_{Y,X_j|X_{-j}}(\cdot, \cdot) := C_{Y,X_j|X_{-j}}^{\theta_{Y,X_j|X_{-j}}}(\cdot, \cdot; \theta_{Y,X_j|X_{-j}})$ is the bivariate parametric copula corresponding to $(Y, X_j)$ given $X_{-j} = x_{-j}$.

The question arises, which R-vine copulas we can use, such that we can describe the conditional distribution of $V$ using the pair copula construction. The idea is to choose an R-vine copula
\[
\mathcal{R} = (V, \mathcal{B}(V), \Theta(\mathcal{B}(V)))
\]
on $d + 1$ elements $V, U_1, \ldots, U_d$, such that $V$ is an R-vine leaf (Definition 2.13) in the corresponding regular vine tree sequence $\mathcal{V} = (T_1, \ldots, T_d)$. Then, we know that $V \notin D_e$ for all $e \in \bigcup_{k=1}^{d} E_k$ and by Corollary 2.17
\[
\text{pt}_1^V \in \{U_1, \ldots, U_d\} \text{ and } \text{pt}_k^V \in \{U_1, \ldots, U_d\} \setminus \{\text{pt}_1^V, \ldots, \text{pt}_{k-1}^V\} \text{ for all } k \in \{2, \ldots, d\}.
\]
Without loss of generality we assume that
\[
\text{pt}_k^V = U_k, \text{ for all } k \in \{1, \ldots, d\}.
\]
Then, a valid regular vine tree sequence $\mathcal{V}$ must contain edges
\[
(V, U_1), (V, U_2|U_1), \ldots, (V, U_d|U_1, \ldots, U_{d-1}).
\]
Thus, we know that $\mathcal{R}$ specifies the parametric pair copulas

$$
C_{V,U_1}^\theta(\cdot, \cdot; \theta_{V,U_1}) \text{ and } C_{V,U_k|U_1,\ldots,U_{k-1}}^\theta(\cdot, \cdot; \theta_{V,U_k|U_1,\ldots,U_{k-1}}) \text{ for all } k \in \{2, \ldots, d\},
$$

which can be used to construct the desired conditional distribution of $V$ given $U_1, \ldots, U_d$.

Accordingly, we define an R-vine copula for quantile regression in the following way:

**Definition 4.3.** (Regular vine copula for quantile regression)

Let $V, U_1, \ldots, U_d$ be uniformly distributed random variables. Then,

$$
R_d = (V^d, \mathcal{B}^d(V^d), \Theta^d(\mathcal{B}^d(V^d)))
$$

is an R-vine copula for quantile regression of the response variable $V$ on covariates $U_1, \ldots, U_d$, if

(i) $R_d$ is a regular vine copula on $d + 1$ elements: $V, U_1, \ldots, U_d$,

(ii) $V$ is an R-vine leaf in $\mathcal{Y}^d = (T^d_1, \ldots, T^d_d)$, $T^d_k = (N^d_k, E^d_k)$, $k = 1, \ldots, d$ and

(iii) $\text{pt}^V_k = U_k$, $k = 1, \ldots, d$.

Example 4.4 shows a pair copula construction of the conditional distribution for three covariates $U_1, \ldots, U_3$ assuming a C-vine tree sequence.

**Example 4.4.** (Conditional distribution of $V$ within a 4-dimensional C-vine)

Let $\mathcal{V} = (T_1, \ldots, T_3)$ be the C-vine tree sequence shown in Figure 4.1 with $N_1 = \{V, U_1, U_2, U_3\}$. We observe that $V$ is an R-vine leaf and $\text{pt}^V_k = U_k$ for all $k \in \{1, 2, 3\}$. 

The corresponding density function can be expressed as

\[ c_{V,\ldots,U_d}(v|u_1,\ldots,u_d) = c_{V,U_1}^{V,U_1}(v,u_1;\theta_{V,U_1}) \times c_{V,U_2}^{V,U_2}(v,u_2;\theta_{V,U_2}) \]

Using the pair copulas corresponding to the edges in \( V \) and the corresponding h-functions, we can express the conditional distribution \( C_{V,U_1,U_2,U_3}(\cdot|u_1,\ldots,u_d) \) recursively as

\[
C_{V,U_1,U_2,U_3}(v|u_1,u_2,u_3) = h_{V,U_1,U_2,U_3}(C_{V,U_1,U_2}(v|u_1,u_2)|h_{V,U_1,U_2}(u_3|u_1,u_2))
\]

Inversion yields the corresponding conditional \( \alpha \)-quantiles:

\[
C^{-1}_{V,U_1,U_2,U_3}(\alpha|u_1,u_2,u_3) = h^{-1}_{V,U_1,U_2,U_3}(\alpha|h_{V,U_1,U_2,U_3}(h_{V,U_1,U_2}(u_3|u_1),h_{U_2,U_1}(u_2|u_1)))|h_{U_2,U_1}(u_2|u_1))].
\]

The corresponding density function can be expressed as

\[
c_{V,\ldots,U_d}(v|u_1,\ldots,u_d) = c_{V,U_1}^{V,U_1}(v,u_1;\theta_{V,U_1}) \times c_{V,U_2}^{V,U_2}(v,u_2;\theta_{V,U_2}) \]

Figure 4.1: C-vine tree sequence on 4 elements \( V, U_1, \ldots, U_4 \) with root node \( U_1 \).
Corollary 4.5 presents a general recursive formulation of $C_{V|U_1,\ldots,U_d}(v|u_1,\ldots,u_d)$, using the pair copulas described by an R-vine copula for quantile regression of $V$ given $U_1, \ldots, U_d$.

**Corollary 4.5.** (Conditional distribution of $V$ given $U_1, \ldots, U_d$ in an R-vine copula for quantile regression)

Let $\mathcal{R}^d$ be an R-vine copula for quantile regression of the response variable $V$ on covariates $U_1, \ldots, U_d$.

Then, we can express the conditional distribution of $V$ given $U_1, \ldots, U_d$ by the recursion:

$$C_{V|U_1}(v|u_1) = h_{V|U_1}(v|u_1),$$

$$C_{V|U_1,\ldots,U_k}(v|u_1,\ldots,u_k) = h_{V|U_1,\ldots,U_k}(v|u_1,\ldots,u_k)\left|C_{U_k|U_1,\ldots,U_{k-1}}(u_k|u_1,\ldots,u_{k-1})\right|, \quad k = 2, \ldots, d.$$  

**Proof:**

Since $U_1 = \text{pt}_1^V$, we know that $(V, U_1) \in E_1^d$ and therefore $b_{V,U_1} \in B^d(\mathcal{Y}^d)$, $\theta_{V,U_1} \in \Theta^d(B^d(\mathcal{Y}^d))$. It follows by the definition of the bivariate conditional distribution that

$$C_{V|U_1}(v|u_1) = \frac{\partial}{\partial u_1} c_{b_{V,U_1}}(v, u_1; \theta_{V,U_1}) = h_{V|U_1}(v|u_1).$$

Similarly, we know that $b_{V,U_{k+1}|U_1,\ldots,U_{k-1}} \in B^d(\mathcal{Y}^d)$, $\theta_{V,U_{k+1}|U_1,\ldots,U_{k-1}} \in \Theta^d(B^d(\mathcal{Y}^d))$ for all $k \in \{2, \ldots, d\}$. Thus, the recursive part follows directly from Theorem 4.2.

Using the recursively defined conditional distributions, we can state a formula for the corresponding conditional density $c_{V|U_1,\ldots,U_d}(v|u_1,\ldots,u_d)$.

**Corollary 4.6.** (Conditional density of $V$ given $U_1, \ldots, U_d$ within an R-vine copula for quantile regression)

In the setup of Corollary 4.5, it holds:

$$c_{V|U_1,\ldots,U_d}(v|u_1,\ldots,u_d) = c_{b_{V,U_1}}(v, u_1; \theta_{V,U_1}) \times c_{b_{V,U_2|U_1}}(h_{V|U_1}(v|u_1), h_{U_2|U_1}(u_2|u_1); \theta_{V,U_2|U_1})$$

$$\times \prod_{k=3}^d c_{b_{V,U_{k+1}|U_1,\ldots,U_{k-1}}}(C_{V|U_1,\ldots,U_{k-1}}(v|u_1,\ldots,u_{k-1}), C_{U_k|U_1,\ldots,U_{k-1}}(u_k|u_1,\ldots,u_{k-1}; \theta_{V,U_k|U_1,\ldots,U_{k-1}})).$$

Note that the pair copula decomposition of $C_{U_k|U_1,\ldots,U_d}(u_k|u_1,\ldots,u_{k-1})$, $k = 2, \ldots, d$, depends on the available pair copulas specified by the R-vine copula $\mathcal{R}^d$.

For example for three covariates $U_1, U_2, U_3$ we can express

$$C_{U_3|U_1,U_2}(u_3|u_1,u_2) = \begin{cases} h_{U_3|U_2,U_1}(h_{U_3|U_1}(u_3|u_1)|h_{U_2|U_1}(u_2|u_1)), & (U_1, U_2) \in E_1^2 \end{cases}$$

$$\begin{cases} h_{U_3|U_2,U_1}(h_{U_3|U_2}(u_3|u_2)|h_{U_1|U_2}(u_1|u_2)), & (U_1, U_2) \in E_1^2 \end{cases}.$$
We close this section by introducing special goodness of fit measures for regular vine copulas. Suppose we are given observations \( \mathbf{v} = (v_1, \ldots, v_n)^T, \mathbf{u}_j = (u_{1j}, \ldots, u_{nj})^T \in [0, 1]^n \) of the response variable \( V \) and the covariates \( U_j, j = 1, \ldots, d \). Since we are interested in predicting conditional quantiles of \( V \), the measures are supposed to refer to the conditional distribution of \( V \) given \( U_1, \ldots, U_d \). Therefore, we introduce conditional log likelihood, conditional AIC and conditional BIC of an R-vine copula.

**Definition 4.7.** (Conditional log likelihood, AIC and BIC of an R-vine copula)

Let \( \mathcal{R}^d \) be an R-vine copula for quantile regression of the response variable \( V \) on covariates \( U_1, \ldots, U_d \). Additionally, assume that we are given observations \( \mathbf{v} = (v_1, \ldots, v_n)^T, \mathbf{u}_j = (u_{1j}, \ldots, u_{nj})^T \in [0, 1]^n \). Then, the number of parameters, conditional log likelihood, AIC, and BIC of \( \mathcal{R}^d \) are defined as follows:

1. \( |\Theta^d| := \sum_{\theta_e \in \Theta^d} |\theta_e| \), where \( \Theta^d := \Theta^d(\mathcal{B}^d(\mathcal{V}^d)) \),

2. \( c\ell\ell(\mathcal{R}^d, \mathbf{v}, (\mathbf{u}_1, \ldots, \mathbf{u}_d)) := \sum_{i=1}^n \log c_V|\mathbf{u}_1, \ldots, \mathbf{u}_d(v_i|u_{i1}, \ldots, u_{id}), \)

3. \( c\text{AIC}(\mathcal{R}^d, \mathbf{v}, (\mathbf{u}_1, \ldots, \mathbf{u}_d)) := -2c\ell\ell(\mathcal{R}^d, \mathbf{v}, (\mathbf{u}_1, \ldots, \mathbf{u}_d)) + 2|\Theta^d|, \)

4. \( c\text{BIC}(\mathcal{R}^d, \mathbf{v}, (\mathbf{u}_1, \ldots, \mathbf{u}_d)) := -2c\ell\ell(\mathcal{R}^d, \mathbf{v}, (\mathbf{u}_1, \ldots, \mathbf{u}_d)) + \log(n)|\Theta^d|. \)
Chapter 5

Selection of an R-vine copula for quantile regression

In the previous section, we have seen how we derive the conditional quantile function, if we are given an (estimated) R-vine copula for quantile regression of $V$ on $U_1, \ldots, U_d$. However, we have not answered the question which R-vine copula we should choose, if we are given observations on the response variable and the covariates.

Consider the case that we are given a data set on $J + 1$ variables, i.e. we are given observations

- $v = (v_1, \ldots, v_n)^T \in [0, 1]^n$ of the response variable $V$ and
- $u_{1}^{p}, \ldots, u_{J}^{p}, \underline{u}_{j}^{p} = (u_{1j}^{p}, \ldots, u_{nj}^{p})^T$ on $J$ potential covariates $U_1^{p}, \ldots, U_J^{p}, \underline{U}_j^{p} \sim U(0, 1)$.

In order to distinguish between potential and selected covariates, potential covariates and their observations are denoted by a superscript “$p$”.

The goal is to estimate an R-vine copula $\hat{R}^d$ for quantile regression of $V$ on $d$ covariates $U_1, \ldots, U_d$. Therefore, we need to choose $d$ covariates out of $U_1^{p}, \ldots, U_J^{p}$, find an appropriate R-vine tree structure $\mathcal{V}^d$ and estimate the corresponding pair copulas, of which the families and parameter vectors will be elements of $\hat{B}^d(\mathcal{V}^d)$ and $\hat{\Theta}^d(\hat{B}^d(\mathcal{V}^d))$.

5.1 Computational effort in the selection of the global $c_{LL}$-optimal R-vine copula

Most intuitively one would like to select the R-vine copula for quantile regression, which is $c_{LL}$-optimal within the set of all possible R-vine copulas for quantile regression.

As indicated by the notation

$$\hat{R}^d = (\mathcal{V}^d, \hat{B}^d(\mathcal{V}^d), \hat{\Theta}^d(\hat{B}^d(\mathcal{V}^d)))$$

an R-vine copula for quantile regression of $V$ on $U_1, \ldots, U_d$ strongly depends on the underlying R-vine tree sequence $\mathcal{V}^d$.

Given a response $V$ and $J$ potential covariates we are therefore interested in identifying
all possible R-vine tree sequences, in which $V$ is an R-vine leaf. As any number $1 \leq d \leq J$ of chosen covariates might be optimal, we have to consider all R-vine tree sequences on $1 + 1, \ldots, J + 1$ elements. However, since $V$ needs to be a leaf in every tree of the corresponding sequence, the task simplifies to determining all possible R-vine tree sequences on $1, \ldots, J$ elements.

The question arises up to which number of potential covariates the approach finds the global $c\ell\ell$-optimal R-vine structure in a feasible time, i.e. how many sequences are possible, if we are given $J$ potential covariates? We first determine the number of possible R-vine sequences for a fixed order within the set of potential covariates and a fixed number of included covariates $d \in \{1, \ldots, J\}$. It is denoted by $\text{NoR}(d)$. The term “fixed order” refers to the order in which the covariates enter an R-vine copula for quantile regression. It will get clearer, once we introduce our covariate selection approach.

**Lemma 5.1. (Number of valid R-vine tree sequences for $d$ ordered elements)**
The number of valid R-vine tree sequences on $d \geq 1$ elements is given by

$$\text{NoR}(d) = 1_{d=1} + 1_{d \geq 2} \times \prod_{i=2}^{d} 2^{i-2}.$$ 

**Proof:**

(i) $d = 1$:
For one covariate $U_1$ there exists a unique D-vine structure:

$$\begin{array}{c}
\text{T}_1 \\
V \\
V, U_1 \longrightarrow U_1
\end{array}$$

(ii) $d \geq 2$ : We will prove this by induction.

- $d = 2$ : For two (preselected) covariates $U_1$ and $U_2$ there exists a unique D-vine tree sequence:

$$\begin{array}{c}
\text{T}_1 \\
V \\
V, U_1 \longrightarrow U_1 \longrightarrow U_1, U_2 \longrightarrow U_2
\end{array}$$

$$\begin{array}{c}
\text{T}_2 \\
V, U_1 \longrightarrow V, U_2 | U_1 \longrightarrow U_1, U_2
\end{array}$$

- $d - 1 \rightarrow d$ : We assume that the formula holds for $d - 1$ and observe that

$$\text{NoR}(d) = 1_{d=1} + 1_{d \geq 2} \times \prod_{d=2}^{d} 2^{i-2} = 2^{d-2} \times \text{NoR}(d - 1).$$
Thus, it remains to prove that the number of possible extensions of an R-vine tree sequence on \( d - 1 \) covariates to \( d \) covariates is given by \( 2^{d-2} \). A theorem proved by [Morales Napoles et al. 2010] states that for any vine on \( d - 1 \) nodes, the number of regular vines on \( d \) nodes, which extend this vine, preserving the so called natural order of the unextended vine is given by \( 2^{d-3} \). Since there exist exactly two different natural orders for every R-vine sequence, we conclude that there are
\[
2 \times 2^{d-3} = 2^{d-2}
\]
possible extensions.

\[\square\]

A technical definition of the term “natural order” can be found in [Morales Napoles et al. 2010]. It answers the question why there are exactly two natural orders for every regular vine tree sequence.

According to Lemma 5.1, \( NoR(d) \) grows exponentially with the number of included covariates \( d \). However, we have not considered

\begin{itemize}
  \item the covariate order and
  \item how many covariates \( d \in \{1, \ldots, J\} \) enter the sequence.
\end{itemize}

It is clear that we have \( J - s + 1 \) possibilities of choosing the \( s \)-th covariate for all \( s \in \{1, \ldots, d\} \). Since all \( d \in \{1, \ldots, J\} \) are possible, the total number of possible R-vine tree sequences for \( J \) potential covariates is given by
\[
NR(J) = \sum_{d=1}^{J} NoR(d) \left( \prod_{s=1}^{d} (J - s + 1) \right).
\]

Table 5.1 lists \( NoR(J) \) and \( NR(J) \) for \( J \in \{1, \ldots, 8\} \). Additionally, we introduce the number
\[
ND(J) = \sum_{d=1}^{J} \left( \prod_{s=1}^{d} (J - s + 1) \right)
\]
which yields the number of possible D-vines for \( J \) potential covariates.

<table>
<thead>
<tr>
<th>( J )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>NoR(J)</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>8</td>
<td>64</td>
<td>1.024</td>
<td>32.768</td>
<td>2.097.152</td>
</tr>
<tr>
<td>NR(J)</td>
<td>1</td>
<td>4</td>
<td>21</td>
<td>256</td>
<td>8.785</td>
<td>786.516</td>
<td>170.480.149</td>
<td>85.899.462.496</td>
</tr>
<tr>
<td>ND(J)</td>
<td>1</td>
<td>4</td>
<td>15</td>
<td>64</td>
<td>325</td>
<td>1.956</td>
<td>13.699</td>
<td>109.600</td>
</tr>
</tbody>
</table>

Table 5.1: NoR(J), NR(J) and ND(J) for \( J \in \{1, \ldots, 8\} \).
According to Table 5.1, the number of possible R-vine tree sequences exceeds 1000 already for $J = 5$. Therefore, selecting the global $c_{\ell\ell}$-optimal R-vine copula might be feasible up to 3 or 4 potential covariates, but cannot be an option once we are given 5 or more potential covariates. If we are given a fixed order of covariates, one might think of selecting the $c_{\ell\ell}$-optimal copula up to 5 potential covariates, which would require the estimation of $\sum_{d=1}^{5} NoR(d) = 76$ R-vine copulas. Selecting the global $c_{\ell\ell}$-optimal D-vine copula is feasible for $J \in \{1, \ldots, 5\}$.

Since each of the numbers is too high for $J \geq 6$, we introduce approaches, that sequentially select an optimal R-vine copula, in the following.

### 5.2 D-vine quantile regression

Kraus and Czado [2017] have developed a first method to determine an appropriate R-vine copula for quantile regression $\mathcal{R}^d$, where the regular vine tree sequence $\mathcal{V}^d$ is restricted to a D-vine tree sequence (Definition 2.20) with $T_1^d$ of the form

$$V - U_1 - \cdots - U_d,$$

and $V$ being an R-vine leaf. It is obvious that

$$pt_k^V = U_k, \ k = 1, \ldots, d,$$

holds within such a D-vine tree sequence.

Assume that we are given $n$ observations of a response variable $V$ and $J$ potential covariates $U_1^p, \ldots, U_J^p$, i.e. we have

$$v = (v_1, \ldots, v_n)^T \in [0, 1]^n$$

and

$$u_j^p = (u_{1j}^p, \ldots, u_{nj}^p)^T \in [0, 1]^n, \ j = 1, \ldots, J.$$

The D-vine tree sequence is found by sequentially adding one of the potential covariates to the sequence, i.e. in iteration $s$ one considers the D-vine tree sequence $\mathcal{V}^s$ with $T_1^s$ of the form

$$V - U_1 - \cdots - U_s.$$

The question arises which of the potential covariates is added in each iteration. Since the goal is to predict $\alpha$-quantiles of $V$ as accurate as possible, the idea is to choose $U_s = U_{js}$ in such a way that the conditional log likelihood

$$c_{\ell\ell}(\mathcal{R}^s, v, (u_1, \ldots, u_s)) := \sum_{i=1}^{n} \log c_{V|U_1^s, \ldots, U_i^s}(v_i|u_{i1}, \ldots, u_{is})$$

is maximized or the corresponding conditional AIC (BIC) is minimized. The algorithm stops, if adding any of the potential covariates does not improve the conditional log likelihood significantly. This is quantified by a specific likelihood ratio test which is defined in the following.
**Definition 5.2.** (Conditional likelihood ratio test between two R-vine copulas for quantile regression)

Let $R_s$ be an R-vine copula for quantile regression of a response $V$ on $s$ covariates $U_1, \ldots, U_s$. Let $R_{s+1}$ be the R-vine copula for quantile regression of $V$ on $U_1, \ldots, U_s, U_{s+1}$ which includes an additional covariate $U_{s+1}$. Additionally, assume that we are given $n$ observations on each of the considered variables, i.e. $v, u_j \in [0, 1]^n$, $j = 1, \ldots, J$.

The null-hypothesis $H_0 : \text{Adding } U_{s+1} \text{ to } R_s \text{ does not improve the copula fit}$ is rejected at level $\alpha_{LRT} \in (0, 1)$, if

$$c(\ell(\mathcal{R}_{s+1}, v, (u_1, \ldots, u_{s+1})) - c(\ell(\mathcal{R}_s, v, (u_1, \ldots, u_s))) > \chi^2_{1-\alpha_{LRT}, |\Theta^{s+1}| - |\Theta^s|},$$

where $\chi^2_{1-\alpha_{LRT}, |\Theta^{s+1}| - |\Theta^s|}$ denotes the $(1-\alpha_{LRT})$-quantile of a $\chi^2$-distribution with $|\Theta^{s+1}| - |\Theta^s|$ degrees of freedom.

### 5.3 Computational complexity in the extension of D-vine quantile regression to the set of all R-vine copulas

One could think about extending the D-vine quantile regression approach to sequentially finding the optimal R-vine copula $\mathcal{R}^d$ within the set of all valid R-vine copulas for quantile regression of $V$ on $U_1, \ldots, U_d$. In each iteration $s$ one would need to consider all possible extensions of a regular vine tree sequence $V_s$ on $s$ elements to an R-vine tree sequence on $s + 1$ elements.

Again, consider the case that we are given $J$ potential covariates $U_1^p, \ldots, U_J^p$. In each iteration $s$, we could extend the R-vine copula by any of the variables which have not been included in a previous iteration $1, \ldots, s-1$. Thus, there are $J - s + 1$ possibilities of selecting the covariate $U_s$. Therefore, $J - s + 1$ copulas are estimated in each iteration $s = 1, 2, \ldots, d$ in the D-vine quantile regression algorithm, i.e.

$$N^{Dvine}(J, d) := \sum_{s=1}^{d} (J - s + 1)$$

D-vine copulas are estimated, if the algorithm stops in some iteration $d \leq J$.

Even though the number of possible selections of $U_s$ would not change in the general case, the total number of R-vine copulas to be estimated would increase significantly.

As mentioned in the previous section, there are

$$1_{s \leq 2} + 1_{s \geq 3} 2^{s-2}$$

possible extensions of an R-vine tree sequence on $s-1$ elements to an R-vine tree sequence consisting of $s$ elements. Since $V$ is an R-vine leaf, we do not have to consider the whole
vine tree sequence in each iteration $s$, but we can focus on the corresponding subvine consisting of $U_1, \ldots, U_s$. Thus, one would need to estimate

$$N^{R_{\text{vine}}}(J, d) := \sum_{s=1}^{d} (J - s + 1)(1_{s \leq 2} + 1_{s \geq 3}2^{s-2})$$

R-vine copulas, if one would like to extend the sequential approach of D-vine quantile regression to general R-vine copulas and the algorithm stops in some iteration $d \leq J$. Note that the D-vine quantile approach might include all $J$ potential covariates in the R-vine copula, i.e. we might have $d = J$. Table 5.2 shows $N^{D_{\text{vine}}}$ and $N^{R_{\text{vine}}}$ for $d = J \in \{1, 2, 3, 5, 8, 10, 15, 20\}$.

<table>
<thead>
<tr>
<th>$d = J$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>5</th>
<th>8</th>
<th>10</th>
<th>15</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N^{D_{\text{vine}}}(J, J)$</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>15</td>
<td>36</td>
<td>55</td>
<td>120</td>
<td>210</td>
</tr>
<tr>
<td>$N^{R_{\text{vine}}}(J, J)$</td>
<td>1</td>
<td>3</td>
<td>7</td>
<td>31</td>
<td>255</td>
<td>1023</td>
<td>32767</td>
<td>1048575</td>
</tr>
</tbody>
</table>

Table 5.2: $N^{D_{\text{vine}}}(J, J)$ and $N^{R_{\text{vine}}}(J, J)$ for $J \in \{1, 2, 3, 5, 8, 10, 15, 20\}$.

Observe that already for $d = J = 5$ we would need to estimate twice as many R-vine copulas compared to the number of D-vine copulas within the D-vine quantile regression algorithm. Therefore, it is not feasible to extend the D-vine quantile regression algorithm to the class of all R-vine copulas and we need to find other methods to select the covariates $U_1, \ldots, U_s$ and the R-vine tree sequence $V^s$, which the estimated R-vine copula for quantile regression $\hat{C}_s$ is based on.

**Example 5.3.** (Dax returns)

The data set *daxreturns* from the R-package *VineCopula* ([Schepsmeier et al. 2016](#p. 90)) will be used to illustrate the algorithms introduced in the thesis.

(i) **Description:** This data set contains transformed standardized residuals of daily log returns of 15 major German stocks represented in the index DAX observed from January 2005 to August 2009. Each time series is filtered using a GARCH(1,1) model with Student t innovations.

(ii) **Format:** A data frame with $n = 1158$ observations on $J + 1 = 15$ variables. Column names correspond to ticker symbols of the stocks.

(iii) **Source:** Yahoo! Finance.

(iv) **Variables:**

- **Response variable:**
  
  $V := BMW$ (Bayerische Motoren Werke AG)

- **Covariates:**

  $U_p^1 := ALV$ (Allianz SE),
  $U_p^2 := BAS$ (BASF SE),
\(U_1^p := \text{BAYN} \text{ (Bayer AG)},\)
\(U_2^p := \text{DAI} \text{ (Daimler AG)},\)
\(U_3^p := \text{DBK} \text{ (Deutsche Bank AG)},\)
\(U_4^p := \text{DTE} \text{ (Deutsche Telekom AG)},\)
\(U_5^p := \text{EOAN} \text{ (E.ON SE)},\)
\(U_6^p := \text{FME} \text{ (Fresenius Medical Care AG & Co. KGAA)},\)
\(U_7^p := \text{LIN} \text{ (Linde AG)},\)
\(U_8^p := \text{MUV2} \text{ (Münchener Rückversicherungs-Gesellschaft AG)},\)
\(U_{10}^p := \text{RWE} \text{ (RWE AG)},\)
\(U_{11}^p := \text{SAP} \text{ (SAP SE)},\)
\(U_{12}^p := \text{SIE} \text{ (Siemens AG)},\)
\(U_{14}^p := \text{VOW3} \text{ (Volkswagen AG)}.\)

\((v)\) Data:
For BMW and \(U_j^p\) we are given observations \(bmw = (bmw_1, \ldots, bmw_n)^T,\)
\(u_j^p = (u_{j1}^p, \ldots, u_{jn}^p)^T \in [0,1]^n, \ j = 1, \ldots, 14.\)

### 5.4 Selection of covariates by partial correlations

In general, we distinguish two different types of covariate selection:

- One selects a fixed set of covariates \(\{U_1, \ldots, U_d\}\) out of a set of potential covariates \(\{U_1^p, \ldots, U_J^p\}\), which will all be sequentially included in the R-vine copula \(R^d\).
- One assigns a certain order \(j_1 < \cdots < j_J\) to the potential covariates \(U_1^p, \ldots, U_J^p\) and includes covariate \(U_{js}^p\) in iteration \(s\), if a certain stop criteria is not fulfilled.

One way of selecting a fixed set of covariates \(U_1, \ldots, U_d\) out of a set of potential covariates \(\{U_1^p, \ldots, U_J^p\}\) is to apply the D-vine quantile regression algorithm until it stops after iteration \(d\) according to the conditional likelihood ratio test. \(U_s\) is then defined to be the variable which was added to the D-vine copula in iteration \(s = 1, \ldots, d\).

In the following, we introduce an approach aiming to assign an order to the potential covariates, which is based on (partial) correlation arguments. Therefore, we assume that we are given observations of the response \(V\) and the potential covariates \(U_1^p, \ldots, U_J^p\), i.e.

\[v = (v_1, \ldots, v_n)^T, u_j^p = (u_{j1}^p, \ldots, u_{jn}^p)^T \in [0,1]^n, \ j = 1, \ldots, J.\]

The first covariate to be included in the R-vine copula is defined to be the variable \(U_{j_1}^p\), which is maximally correlated to the response \(V\). The correlation is measured in terms of the absolute empirical Pearson correlation (Definition 1.11) on the z-scale, i.e.

\[U_1 := U_{j_1}^p \text{ with } j_1 := \arg\max_{j \in \{1, \ldots, J\}} |\hat{\rho}_{W,Z_j}^n|,\]

where \(W := \Phi^{-1}(V)\) and \(Z_j^p := \Phi^{-1}(U_{j}^p), \ j = 1, \ldots, J.\)
For $s \geq 2$ the order is based on the absolute empirical partial correlations (Definition 1.21 and Example 1.23) between each potential covariate and the response $V$ given the covariates $U_1, \ldots, U_{s-1}$, i.e.

$$U_s := U_{j_s}^p \text{ with } j_s := \arg \max_{j \in \{1, \ldots, J\} \setminus \{j_1, \ldots, j_{s-1}\}} |\hat{\rho}_{W,Z_j^p}^n; Z_{j_1}^p, \ldots, Z_{j_{s-1}}^p|,$$

where $W := \Phi^{-1}(V)$ and $Z_j^p := \Phi^{-1}(U_j^p)$, $j = 1, \ldots, J$.

Iterations $s = 1, \ldots, 4$ are illustrated in Example 5.4 using the data set *daxreturns* from the R-package *VineCopula* (Example 5.3).

**Example 5.4.** (Daxreturns - Illustration of selection of covariates)

We choose BMW to be the response variable $V$ and want to find $U_1 := U_{j_1}^p, U_4 := U_{j_4}^p$, which will be the first four covariates entering the R-vine copula for quantile regression of BMW.

- **First**, we transform the available copula data to the z-scale and obtain data points $w := (\Phi^{-1}(bmw_1), \ldots, \Phi^{-1}(bmw_n))^T \in \mathbb{R}^n$ and $z_j^p := (\Phi^{-1}(u_{1j}^p), \ldots, \Phi^{-1}(u_{nj}^p))^T \in \mathbb{R}^n$, $j = 1, \ldots, 14$. The transformed random variables are defined as $W := \Phi^{-1}(BMW)$ and $Z_j^p := \Phi^{-1}(U_j^p)$, $j = 1, \ldots, 14$.

- **$s = 1$**:
  The upper left barplot in Figure 5.1 shows the empirical Pearson correlations $\hat{\rho}_{W,Z_j^p}^n$ between $W$ and $Z_j^p$ for all $j = 1, \ldots, 14$.
  We observe that
  $$\max_{j \in \{1, \ldots, 14\}} \{|\hat{\rho}_{W,Z_j^p}^n|\} \approx 0.67,$$
  which is attained for $j = 4$. Therefore,
  $$U_1 := U_4^p = DAI$$
  will be the first covariate entering the R-vine copula.

- **$s = 2$**:
  For $s \geq 2$, we consider the empirical partial correlations
  $$\hat{\rho}_{W,Z_j^p; Z_{j_1}^p, \ldots, Z_{j_{s-1}}^p}^n, \ j \in \{1, \ldots, 14\} \setminus \{j_1, \ldots, j_{s-1}\}.$$
  As shown in the upper right barplot of Figure 5.1,
  $$\arg \max_{j \in \{1, \ldots, 14\} \setminus \{4\}} \{|\hat{\rho}_{W,Z_j^p; Z_4^p}^n|\} = 5.$$
  We conclude that
  $$U_2 := U_5^p = DBK.$$
Figure 5.1: Barplots of empirical (partial) correlations between $W := \Phi^{-1}(BMW)$ and $Z^p_{js} := \Phi^{-1}(U^p_{js})$ (given $Z_{j1}, \ldots, Z_{js-1}$) for $j_s \in \{1, \ldots, 14\} \setminus \{j_1, \ldots, j_{s-1}\}$ and $s = 1, \ldots, 4$.

- $3 \leq s \leq 4$:
  Similarly, we obtain
  
  $U_3 := U^p_{14} = VOW3$

  and

  $U_4 := U^p_{13} = SIE$.

The pseudo-code of the algorithm is presented in Algorithm [1]
Algorithm 1 Selection of covariates by (partial-) correlations.

1: **Input**: \( d, \mathbf{v} \in [0,1]^n, \mathbf{u}_j^p \in [0,1]^n, j = 1, \ldots, J \)

2: \( W := \Phi^{-1}(V), \mathbf{w} := (\Phi^{-1}(v_1), \ldots, \Phi^{-1}(v_n))^T \in \mathbb{R}^n \)

3: \( Z_j^p := \Phi^{-1}(U_j^p), \mathbf{z}_j^p := (\Phi^{-1}(u_{j1}^p), \ldots, \Phi^{-1}(u_{nj}^p))^T \in \mathbb{R}^n, j = 1, \ldots, J \)

4: \( s = 1 \)

5: \( j_1 := \arg \max_{j \in \{1, \ldots, J\}} |\hat{\rho}_{W, Z_j^p}^n| \)

6: \( U_1 := U_{j_1}^p \)

7: for \( 2 \leq s \leq d \) do:

8: \( j_s := \arg \max_{j \in \{1, \ldots, J\} \setminus \{j_1, \ldots, j_{s-1}\}} \{|\hat{\rho}_{W, Z_j^p, Z_{j_1}^p, \ldots, Z_{j_{s-1}}^p}^n|\} \)

9: \( U_s := U_{j_s}^p \)

10: **Output** \( U_1, \ldots, U_d \)

The approach is motivated by the fact that for normally distributed data, partial correlations and conditional correlations coincide. Then, we select \( U_1 \), such that it explains as much as possible of the behavior of the response \( V \) in the first step \( s = 1 \), and further covariates \( U_2, \ldots, U_J \), such the behavior of \( V \) is explained as well as possible given the covariates which have already been selected. 

In addition to this intuitive justification we present a mathematical justification for the selection of \( U_1 \) and \( U_2 \) in the following.

Consider the **conditional log likelihood per observation** of the response \( W := \Phi^{-1}(V) \) on the z-scale. Given observations

\[ \mathbf{w} := (w_1, \ldots, w_n)^T \] of \( W \)

and setting the covariates

\[ \Phi^{-1}(U_1) := Z_1 \text{ and } Z_2 := \Phi^{-1}(U_2) \]

to some fixed values \( Z_1 = z_1, Z_2 = z_2 \) the **conditional log likelihood per observation** is defined as

\[ c\ell_n(w; Z_1 = z_1) := \frac{1}{n} \sum_{i=1}^{n} \log f(w_i|Z_1 = z_1) \text{ for } Z_1 = z_1 \]

and

\[ c\ell_n(w; Z_1 = z_1, Z_2 = z_2) := \frac{1}{n} \sum_{i=1}^{n} \log f(w_i|Z_1 = z_1, Z_2 = z_2) \text{ for } Z_1 = z_1, Z_2 = z_2. \]
As will be shown in Corollary 5.5, for \( n \to \infty \), both \( \tilde{c} \ell_n (w; Z_1 = z_1) \) and \( \tilde{c} \ell_n (w; Z_1 = z_1, Z_2 = z_2) \) converge to expressions, which are independent of the conditioning values \( z_1 \) and \( z_2 \) and monotonically increasing in \( |\rho_{w,z_1}| \) and \( \rho_{w,z_2;z_1} \), respectively. Therefore, the conditional log likelihood per observation is increasing in the considered Pearson and partial correlations (for \( n \) big enough). Thus, it is reasonable to choose the order of the covariates according to the correlation approach in Algorithm 1.

**Corollary 5.5.**

(i) Let \((W, Z)^T \sim N_2(0, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix})\), and \( Z = z \) for some fixed conditioning value \( z \), then

\[
\tilde{c} \ell_n (w; Z = z) \to g(|\rho|),
\]

where \( g(|\rho|) \) is monotonically increasing in \( |\rho| \).

(ii) Define \( \rho_1 := \rho_{W,Z_1} \), \( \rho_2 := \rho_{W,Z_2} \), \( \rho_{1,2} := \rho_{Z_1,Z_2} \), let \((W, Z_1, Z_2)^T \sim N_3(0, \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_{1,2} \\ \rho_2 & \rho_{1,2} & 1 \end{pmatrix})\), and \( Z_1 = z_1, Z_2 = z_2 \) for some fixed conditioning values \( z_1, z_2 \), then

\[
\tilde{c} \ell_n (w; Z_1 = z_1, Z_2 = z_2) \to h(\rho_{w,z_2;z_1}),
\]

where \( h(\rho_{w,z_2;z_1}) \) is monotonically increasing in \( \rho_{w,z_2;z_1} \) for fixed values of \( \rho_1 \) and \( \rho_2 \).

**Proof:**

(i) By the strong law of large numbers we have that

\[
\tilde{c} \ell_n (w; Z = z) \to E_{W|Z=z}[\log f(W|Z = z)].
\]

Further, we know that \( W|Z = z \sim N(\rho z, 1 - \rho^2) \) and the conditional density of \( W \) is given by

\[
f(w|Z = z; \rho) = \frac{1}{\sqrt{2\pi(1 - \rho^2)}} \exp \left\{ -\frac{1}{2(1 - \rho^2)}(w - z\rho)^2 \right\}.
\]

Consequently, \( \log f(w|Z = z) \) is given by

\[
\log f(w|Z = z) = -\frac{1}{2} \log(2\pi(1 - \rho^2)) - \frac{1}{2(1 - \rho^2)}(w - z\rho)^2.
\]

Using the definition of \( E_{W|Z=z}[\log f(W|Z = z)] \), we find that

\[
E_{W|Z=z}[\log f(W|Z = z)] = \int_{-\infty}^{\infty} \log f(w|Z = z) f(w|Z = z) dw
\]

\[
= -\frac{1}{2} \log(2\pi(1 - \rho^2)) \int_{-\infty}^{\infty} f(w|Z = z) dw - \frac{1}{2(1 - \rho^2)} \int_{-\infty}^{\infty} (w - z\rho)^2 f(w|Z = z) dw
\]

62
\[
\int \log(2\pi(1 - \rho^2)) \cdot 1 - \frac{1}{2(1 - \rho^2)} \text{Var}(W|Z = z) \\
= -\frac{1}{2} \log(2\pi(1 - \rho^2)) - \frac{1}{2(1 - \rho^2)}(1 - \rho^2) = -\frac{1}{2} \log(2\pi(1 - \rho^2)) - \frac{1}{2} =: g(|\rho|),
\]

since we integrate the conditional density of \( Z \) over the whole domain of \( Z \) in the first integral and integrate \( (w - \mathbb{E}_{W|Z=z}(W|Z = z))^2 \) times the conditional density in the second integral.

The monotonicity of \( g(|\rho|) \) follows by
\[
g'(|\rho|) = \frac{1}{1 - \rho^2} \geq 0, \text{ for all } |\rho| \in [0, 1].
\]

(ii) Similarly to (i), we use the strong law of large numbers and conclude
\[
\ell_n(w; Z_1 = z_1, Z_2 = z_2) \xrightarrow{n \to \infty} \mathbb{E}_{W|Z_1=z_1,Z_2=z_2}[\log f(W|Z_1 = z_1, Z_2 = z_2)].
\]

Further, we observe
\[
\mathbb{E}_{W|Z_1=z_1,Z_2=z_2}[\log f(W|Z_1 = z_1, Z_2 = z_2)]
= \int_{-\infty}^{\infty} \log f(w|Z_1 = z_1, Z_2 = z_2) f(w|Z_1 = z_1, Z_2 = z_2) \, dw
= \int_{-\infty}^{\infty} \left[ -\frac{1}{2} \log(2\pi \sigma_{W|1,2}^2) - \frac{1}{2\sigma_{W|1,2}^2}(w - \mu_{W|1,2})^2 \right] f(w|Z_1 = z_1, Z_2 = z_2) \, dw
= -\frac{1}{2} \log(2\pi \sigma_{W|1,2}^2) - \frac{1}{2\sigma_{W|1,2}^2} \mu_{W|1,2}^2 - \frac{1}{2} \log(2\pi \sigma_{W|1,2}^2) - \frac{1}{2} \sigma_{W|1,2}^2
\]

where \( \mu_{W|1,2} \) and \( \sigma_{W|1,2}^2 \) denote conditional expectation and variance of \( W \) given \( Z_1 = z_1, Z_2 = z_2 \).

According to theory on multivariate Gaussian distributions, we have
\[
\sigma_{W|1,2}^2 = 1 - \frac{1}{1 - \rho_{1,2}^2} \left[ (\rho_1, \rho_2) \begin{pmatrix} 1 & -\rho_{1,2} \\ -\rho_{1,2} & 1 \end{pmatrix} (\rho_1, \rho_2)^T \right]
= 1 - \frac{1}{1 - \rho_{1,2}^2} \left[ (\rho_1^2 - \rho_1\rho_2^2\rho_{1,2} + \rho_2(\rho_2 - \rho_1\rho_{1,2})) \right].
\]

Theorem \textbf{1.22} enables us to express the \( \rho_2(\rho_2 - \rho_1\rho_{1,2}) \) in terms of \( \rho_{W,Z_2;Z_1} \), i.e.
\[
\rho_{W,Z_2;Z_1} = \frac{\rho_2 - \rho_1\rho_{1,2}}{\sqrt{(1 - \rho_1^2)(1 - \rho_{1,2}^2)}}
\]

63
\[ \Rightarrow \rho_2(\rho_2 - \rho_1 \rho_2 \rho_{1,2}) = \rho_{W,Z_2;Z_1} \sqrt{(1 - \rho_2^2)(1 - \rho_{1,2}^2)}. \]

Since \( \rho_2(\rho_2 - \rho_1 \rho_2 \rho_{1,2}) \) is obviously increasing in \( \rho_{W,Z_2;Z_1} \) for fixed \( \rho_1 \) and \( \rho_2 \), \( \sigma_{W|1,2}^2 \) is decreasing in \( |\rho_{W,Z_2;Z_1}| \). Therefore,

\[ h(|\rho_{W,Z_2;Z_1}|) := -\frac{1}{2} \log(2\pi \sigma_{W|1,2}^2(|\rho_{W,Z_2;Z_1}|)) - \frac{1}{2} \]

is increasing in \( |\rho_{W,Z_2;Z_1}| \) for fixed \( \rho_1 \) and \( \rho_2 \).

\[ \square \]

### 5.5 Forward selection of the R-vine tree sequence

In the previous section, we have seen how we can select the set of covariates \( U_1, \ldots, U_d \) out of a set of given variables \( \{U^p_1, \ldots, U^p_f\} \) and can assign an order of inclusion. Now, assume that we know which covariate \( U_s \) enters the R-vine copula up to iteration \( s = d \). We still have not clarified how \( U_s \) enters the copula in iteration \( s \). In other words, which R-vine tree sequence should we choose after \( U_s \) has entered the copula?

We have seen in Section 5.3 that it is not feasible to try out every possible R-vine tree sequence, since it would require a lot of computational effort. Therefore, we introduce criteria which uniquely define the R-vine structure after each iteration.

By the definition of an R-vine copula for quantile regression (Definition 4.3), we know that in iteration \( s = 1, \ldots, d \) the following must hold:

- \( V \) must be an R-vine leaf in \( V^s \),
- \( pt_s^V = U_s \).

Those requirements already define \( V^s \) for iterations \( s = 1, 2 \) because we can only choose the D-vines \( V - U_1 \) and \( V - U_1 - U_2 \).

In iteration \( s = 3 \), we could either select the D-vine \( V - U_1 - U_2 - U_3 \) or the C-vine with root node \( U_1 \), which we have already seen in Figure 4.1. In other words, we have to choose a partner for \( U_3 \) out of the set \( \{U_1, U_2\} \) in tree \( T^3_1 \), i.e.

\[ pt^U_3 \in N^2_{1} \setminus \{V\} = \{U_1, U_2\}. \]

Note that \( pt^U_3 \) is well-defined, since \( U_3 \) is an R-vine leaf in \( V^3 \). Given observations on each variable, the idea is to choose that covariate as partner which maximizes the corresponding empirical Kendall’s \( \tau \) (Definition 1.16) or the empirical maximal semi-correlation (Definition 1.20), respectively. The idea is formalized in Definition 5.6.

**Definition 5.6.** (Weight criterion maximizer in the first tree)

Let \( T^{s-1}_1 = (N^{s-1}_1, E^{s-1}_1) \) be the first tree of a regular vine tree sequence \( V^{s-1} \) on \( s \) elements, where \( N^{s-1}_1 = \{V, U_1, \ldots, U_{s-1}\} \). Then, the weight criterion (wc) maximizer in \( T^s_1 \) is defined as

\[
\max W_1(U_s, \{U_1, \ldots, U_{s-1}\}, \text{wc}) := \begin{cases} \arg \max_{U \in \{U_1, \ldots, U_{s-1}\}} |\hat{\rho}_{U_s,U}^n|, & \text{if wc} = \tau, \\ \arg \max_{U \in \{U_1, \ldots, U_{s-1}\}} |\hat{\rho}_{U_s,U}^{n,\max}|, & \text{if wc} = \rho^{\max}. \end{cases}
\]

64
Note that the selected partner is based on an estimation out of the given data and should actually be denoted by an “hat”. However, we will stick to the “no-hat” notation for all elements of an R-vine tree sequence, even if their selection or structure is based on one or more estimations, in order to avoid inconsistencies.

Further, we introduce a notation for the set of partner candidates for a new variable. Let $\mathcal{V}^{s-1} = (T_1^{s-1}, \ldots, T_{s-1}^{s-1})$ be an R-vine tree sequence on $s$ elements with node set $N_1^{s-1} = \{V, U_1, \ldots, U_{s-1}\}$. The subset of $N_1^{s-1}$, that contains all elements, which are allowed to be partner of a “new” variable $U_s$ in tree $T_k^s$ after considering

- the fact that $V$ cannot be a partner of $U_s$,  
- the proximity condition holds and  
- previously chosen partners

will be denoted by

$\mathcal{PT}_{U^s_k} \quad k = 1, \ldots, s - 1.$

For $s \geq 4$ it might not be enough to apply a weight criterion in the first tree, since in general $T_1^s$ does not define the whole tree sequence. Iteration $s = 6$ is illustrated in Example 5.7 extending an R-vine tree sequence on 6 elements $(V, U_1, \ldots, U_5)$ by the covariate $U_6$.

**Example 5.7.** (Illustration of weight criteria maximizing variables - extension of an R-vine on 5 elements)

Starting point is the R-vine tree sequence $\mathcal{V}^5 = (T_1^5, \ldots, T_5^5)$ shown in Figure 5.2.
In iteration $s = 6$, we want to add a variable $U_6$ to $\mathcal{V}^5$, i.e. we want to find some new tree sequence $\mathcal{V}^6 = (T_6^1, \ldots, T_6^5)$.

- $k = 1$:

  We need to choose some node $v \in N_5^6$ to be neighbor of $U_6$ in $T_6^1$. This is equivalent to selecting $\text{pt}_{U_6} \in N_5^1$. Given that $V$ is an R-vine leaf, we can only choose $\text{pt}_{U_6} \in \{U_1, \ldots, U_5\}$, i.e. the set of partner candidates is given by

  $\text{PT}_{U_6} = \{U_1, \ldots, U_5\}$.

Now, assume that observations

$\mathbf{v}, \mathbf{u}_1, \ldots, \mathbf{u}_6 \in [0, 1]^n$

of the corresponding random variables are available and choose $\text{wc} := \rho^{\text{max}}$.

Assume that the maximal empirical semi-correlation (Definition 1.20) reaches its maximum for $U_2$, i.e.

$\max W_1(U_6, \text{PT}_{U_6}^{U_6}, \rho^{\text{max}}) = U_2$.

We conclude that $\text{pt}_{U_6} := U_2$ and obtain $T_1^6 = (N_6^1, E_1^6)$ as shown in Figure 5.3. The extended edge set is given by

$E_1^6 := E_1^5 \cup \{e_1^6\}$,
where $e_1^6$ denotes the new or added edge in tree $T_1^6$, i.e. 
\[ e_1^6 := (U_2, U_6). \]

- $k = 2$:

From $k = 1$ we know that node $(U_2, U_6) \in N_2^5$. Now, we need to find 
\[ \text{pt}^{U_6}_{U_2} \in N_2^5 \setminus \{V, \text{pt}^{U_6}_{U_1}\}, \]
which will be in the conditioned set of edge $(\text{pt}^{U_6}_{U_2}, U_6|U_2) \in E_2^6$.

Since nodes $(U_1, U_2), (U_2, U_5)$ are the only elements of $N_2^5 = E_1^5$, containing $\text{pt}^{U_6}_{U_1} = U_2$, we can only choose 
\[ \text{pt}^{U_6}_{U_2} \in \text{PT}_{U_1}^U = \{U_1, U_5\} \]
without violating the proximity condition.

Remember that the pair copula corresponding to edge $(\text{pt}^{U_6}_{U_2}, U_6|U_2) = (\text{pt}^{U_6}_{U_1}, U_6|U_2)$ takes the pseudo observations $\hat{h}_{U_6|U_2} \in [0, 1]^n$ and $\hat{h}_{\text{pt}^{U_6}_{U_2}|U_2} \in [0, 1]^n$ as arguments.

That is why we cannot directly apply Definition 5.6. However, we can proceed similarly applying the weight criterion to the conditional distributions 
\[ U_s|U_2 \text{ for all } s \in \{1, 5, 6\}. \]

We observe that observations $\hat{h}_{U_s|U_2} \in [0, 1]^n$, $s \in \{1, 5, 6\}$, of the corresponding conditional random variables are available, since the edges 
\[ (U_s, U_2), s \in \{1, 5, 6\} \]
are part of $E_1^5$, and define similarly to $k = 1$:

$$\max W_2(U_6, PT_{U_6}^{U_2}, wc) := \begin{cases} 
\arg \max_{U \in PT_{U_6}^{U_2}} |\hat{\tau}_{U_6(U_2,U_1)}|, & \text{if } wc = \tau; \\
\arg \max_{U \in PT_{U_6}^{U_2}} |\hat{\rho}_{U_6(U_2,U_1)}|, & \text{if } wc = \rho_{\max}.
\end{cases}$$

Here, $\hat{\tau}_{U_6(U_2,U_1)}$ and $\hat{\rho}_{U_6(U_2,U_1)}^{\max}$ denote the empirical Kendall’s $\tau$ and the maximal empirical semi-correlation between the conditional variables $U_6|U_2$ and $U|U_2$, $U \in PT_{U_6}^{U_2} = \{U_1, U_3\}$. Again, set $wc := \rho_{\max}$ and assume

$$\max W_2(U_6, PT_{U_6}^{U_2}, \rho_{\max}) = U_1.$$ 

The resulting tree $T_2^6$ is shown in Figure 5.4. The extended edge set is given by

$$E_2^6 := E_2^5 \cup \{e_2^6\},$$

where $e_2^6$ denotes the new or added edge in tree $T_1^6$, i.e.

$$e_2^6 := (U_1, U_6|U_2).$$

![Figure 5.4: 2nd-order tree of the extended R-vine tree sequence $\Psi^6$.](image-url)

- $k = 3$:

From $k = 2$ we know that node $(U_1, U_6|U_2) \in N_3^6$. Now, we need to find

$$pt_{U_6}^U \in N_1^5 \setminus \{V, pt_{U_1}^{U_6}, pt_{U_2}^{U_6}\},$$
which will be in the conditioned set of edge \((pt_3^U, U_0|U_1, U_2) \in E_3^6\).

Since nodes \((U_2, U_3|U_1), (U_2, U_4|U_1)\) and \((U_1, U_5|U_2)\) are elements of \(N_3^5 = E_2^5\), we can choose any element of

\[\text{PT}^U_3 = N_1^5 \setminus \{V, U_1, U_2\} = \{U_3, U_4, U_5\}\]

without violating the proximity condition.

Observing the available pair copulas corresponding to edges in \(E_2^6\), we have to find the right arguments for the chosen weight criterion. The pair copula corresponding to edge \((pt_3^U, U_0|U_1, U_2)\) takes the pseudo observations \(\hat{h}_{U_0|U_1; U_2} \in [0, 1]^n\) and

\[
\begin{cases}
\hat{h}_{pt_3^U|U_2; U_1} \in [0, 1]^n, & \text{if } pt_3^U \in \{U_3, U_4\} \\
\hat{h}_{pt_3^U|U_1; U_2} \in [0, 1]^n, & \text{if } pt_3^U = U_5
\end{cases}
\]

as arguments. In other words, the structure of the conditional variables depends on the choice of \(pt_3^U\), i.e. we consider the conditional variables \(U_0|U_1; U_2\) and

\[
\begin{cases}
U|U_2; U_1, & \text{if } U \in \{U_3, U_4\}; \\
U|U_1; U_2, & \text{if } U = U_5.
\end{cases}
\]

In order to simplify the notation, we introduce permutation vectors

\[\pi^6_3(U) = (\pi^6_{3,1}(U), \pi^6_{3,2}(U))^T \in \{1, 2\}^2, \text{ s.t. } (U, U_{x_3^6(U)}|U_{x_3^6(U)}) \in E_2^6 \text{ for every } U \in \text{PT}^U_3.
\]

Those vectors are given by

\[\pi^6_3(U_3) = \pi^6_3(U_4) = (2, 1)^T \text{ and } \pi^6_3(U_5) = (1, 2)^T.
\]

Then we define similarly to \(k = 1, 2\):

\[
\text{maxW}_3(U_6, \text{PT}^U_3, wc) := \begin{cases}
\arg\max_{U \in \{U_3, U_4, U_5\}} \left| \hat{\tau}^n_{U_0|U_1; U_2, U|U_{x_3^6(U)}; U_{x_3^6(U)}} \right|, & \text{if } wc = \tau, \\
\arg\max_{U \in \{U_3, U_4, U_5\}} \left| \hat{\rho}^{n, \max}_{U_0|U_1; U_2, U|U_{x_3^6(U)}; U_{x_3^6(U)}} \right|, & \text{if } wc = \rho^{\max}.
\end{cases}
\]

Choose \(wc := \rho^{\max}\) and assume \(\text{maxW}_3(U_6, \{U_3, U_4, U_5\}, \rho^{\max}) = U_3\). Then, \(T_3^6\) looks as shown in Figure 5.5. The extended edge set is given by

\[E_3^6 := E_3^5 \cup \{e_3^6\},\]

where \(e_3^6\) denotes the new or added edge in tree \(T_3^6\), i.e.

\[e_3^6 := (U_3, U_0|U_1, U_2)\]
Figure 5.5: 3rd-order tree of the extended R-vine tree sequence $\mathcal{V}^6$.

- $k = 4$

From $k = 3$ we know that node $(U_3, U_6|U_1, U_2) \in N^6_4$. Now, we need to find

$$\text{pt}_{U_6}^4 \in N^5_1 \setminus \{V, \text{pt}_{U_1}^4, \text{pt}_{U_2}^4, \text{pt}_{U_3}^4\},$$

which will be in the conditioned set of edge $(\text{pt}_{U_6}^4, U_6|U_1, U_2, U_3) \in E^6_4$.

Since nodes $(U_3, U_4|U_1, U_2)$ and $(U_3, U_5|U_1, U_2)$ are elements of $N^5_3 = E^3_3$, we can choose any element of

$$\text{PT}_{U_6}^4 = N^5_1 \setminus \{V, U_1, U_2, U_3\} = \{U_4, U_5\}$$

without violating the proximity condition.

We set

$$\mathbf{\pi}_4^6(U_4) = \mathbf{\pi}_3^6(U_5) = (3, 1, 2)^T$$

and define similarly to $k = 1, 2, 3$:

$$\text{maxW}_4(U_6, \text{PT}_{U_6}^4, wc) := \begin{cases} \arg \max_{U \in \{U_4, U_5\}} |\hat{\tau}^6_{U_6|U_3; U_1, U_2, U}|, & \text{if } wc = \tau, \\ \arg \max_{U \in \{U_4, U_5\}} |\hat{\rho}^\text{max}_{U_6|U_3; U_1, U_2, U}|, & \text{if } wc = \rho^\text{max}. \end{cases}$$

Choose $wc := \rho^\text{max}$ and assume $\text{maxW}_4(U_6, \{U_4, U_5\}, \rho^\text{max}) = U_4$. Then, $T^6_4$ looks as shown in Figure 5.6. The extended edge set is given by

$$E^6_4 := E^5_4 \cup \{e^6_4\},$$

where $e^6_4$ denotes the new or added edge in tree $T^6_4$, i.e.

$$e^6_4 := (U_4, U_6|U_1, U_2, U_3).$$
In $k = 5$, we do not have to make use of a weight criteria, since only $U_5$ is left to be partner of $U_6$ in $T^6_5 = (N^6_5, E^6_5)$ with

$$N^6_5 = E^6_4 \quad \text{and} \quad E^6_5 = E^5_5 \cup \{e^6_5\} = E^5_5 \cup \{(U_5, U_6|U_1,\ldots,U_4)\}.$$

In $k = 6$, we define $pt^V_6 := U_6$ and obtain $T^6_6 = (N^6_6, E^6_6)$ with

$$N^6_6 = E^6_5 \quad \text{and} \quad E^6_6 = \{e^6_6\} = \{(V, U_6|U_1,\ldots,U_5)\}.$$

Summarizing our results we obtain the R-vine tree sequence $\mathcal{V}^6 = (T^6_1,\ldots,T^6_6)$ as shown in Figure 5.7.
Using the permutation vectors $\pi_s^k(U)$ which we have introduced in Example 5.7, we can formulate a recursive definition of the weight criterion maximizer in tree $T_k^s$ with respect to the corresponding set of partner candidates $PT_{k}^s$, $k = 1, \ldots, s - 1$.

**Definition 5.8.** (Weight criterion maximizer in the tree $T_k^s$)
Let $V^{s-1} = (T_1^{s-1}, \ldots, T_{s-1}^{s-1})$ be a regular vine tree sequence on $s$ elements, where $N_1^{s-1} = \{V, U_1, \ldots, U_{s-1}\}$. Then, the weight criterion maximizer in $T_k^s$ is defined as

$$\max W_k(U_s, PT_k^s, wc) := \begin{cases} \arg\max_{U \in PT_k^s} \xi_{U_s(k), L_U(k)}^n, & \text{if } wc = \tau, \\ \arg\max_{U \in PT_k^s} \xi_{U_s(k), L_U(k)}^{n, max}, & \text{if } wc = \rho^{max}. \end{cases}$$
with

\[ L^*_s(k) := \begin{cases} \{ U_s \}, & k = 1, \\ \{ U_s \mid pt_{k-1}^U \}, & k = 2, \\ \{ U_s \mid pt_{k-1}^U, \ldots, pt_{k-2}^U \}, & k = 3, \ldots, s - 1, \end{cases} \]

\[ L^U(k) := \begin{cases} \{ U \}, & k = 1, \\ \{ U \mid pt U \}, & k = 2, \\ \{ U \mid U_{\pi_{k,1}^s(U)}; U_{\pi_{k,2}^s(U)}, \ldots, U_{\pi_{k,k-1}^s(U)} \}, & k = 3, \ldots, s - 1, \end{cases} \]

where

\[ pt_k^U := \max_{\ell}(U_s, PT^s_k, wc), \ell = 1, \ldots, k - 1 \]

and

\[ \pi_k^s(U) := (\pi_{k,1}^s(U), \ldots, \pi_{k,k-1}^s(U))^T \in \{1, \ldots, s - 1\}^{k-1}, \text{ s.t.} \]

\[ (U_{\pi_{k,1}^s(U)}, U_{\pi_{k,2}^s(U)}), \ldots, U_{\pi_{k,k-1}^s(U)}) \in E_{k-1}^{s-1}. \]

Remember that only those variables are elements of \( PT_k^U \), of which the selection does not violate the proximity condition. This ensures the existence of the permutation vector \( \pi_k^s(U) \). Corollary 5.9 helps us to identify \( PT_k^U \).

**Corollary 5.9.** (Set of possible partners of \( U_s \) in tree \( T_k^s \))

In the situation of Definition 5.8 the set of partner candidates for \( U_s \) in tree \( T_k^s \) \((s \geq 2, k = 1, \ldots, s - 1)\) is given by

\[ \text{PT}_k^s = \begin{cases} N_{1}^{s-1} \setminus \{ V \} = \{ U_1, \ldots, U_{s-1} \}, & k = 1 \\ \{ n \in N_{1}^{s-1} \setminus \{ V, pt_1^U, \ldots, pt_{k-1}^U \} : \exists e \in E_{k-1}^{s-1}, \text{ s.t.} n, pt_1^U, \ldots, pt_{k-1}^U \in A_{e} \}, & k = 2, \ldots, s - 1. \end{cases} \]

**Proof:**

- \( k = 1 \):
  In the first tree \( T_1^s \), the proximity condition (Definition 2.9) does not restrict the set of valid partners for \( U_s \), since it is only defined in trees of order \( k \geq 2 \). Since \( V \) needs to be an R-vine leaf, it cannot be partner of \( U_s \).

- \( k = 2 \):
  We need to make sure that the proximity condition is fulfilled in edge \( e_2^s = (U_s, pt_2^U \mid pt_1^U) \), i.e. it must hold that \( e_2^s = (v_1, v_2) \) with \( v_1 = (U_s, pt_1^U) \in E_1^s \) and \( v_2 \in E_1^{s-1} \), such that \( |A_{v_1} \triangle A_{v_2}| = 2 \).
Since $U_s$ is an R-vine leaf in $\mathcal{V}^s$, we know that only $pt_1^{U_s}$ can be in the considered intersection, i.e.

$$U_s \notin A_{v_2} \Rightarrow pt_1^{U_s} \in A_{v_2}.$$  

We conclude that the set of partner candidates of $U_s$ in tree $T_2^s$ is given by all elements $n \in N_1^{s-1} \setminus \{V, pt_1^{U_s}\}$, such that there exists an edge $v_2 \in E_1^{s-1}$ with $v_2 = (n, pt_1^{U_s})$, i.e.

$$\text{PT}_{2}^s = \{ n \in N_1^{s-1} \setminus \{V, pt_1^{U_s}\} : \exists e \in E_1^{s-1}, \text{ s.t. } n, pt_1^{U_s} \in A_e \}.$$

- $k \rightarrow k + 1$:

We need to make sure that the proximity condition is fulfilled in edge

$$e_{k+1}^s = (U_s, pt_{k+1}^{U_s} | pt_1^{U_s} \ldots, pt_{k}^{U_s}),$$

i.e. it must hold that $e_{k+1}^s = (v_1, v_2)$ with

$$v_1 = (U_s, pt_k^{U_s} | pt_1^{U_s} \ldots, pt_{k-1}^{U_s}) \in E_k^s$$

and

$$v_2 \in E_k^{s-1}, \text{ such that } |A_{v_1} \triangle A_{v_2}| = 2.$$  

Since $U_s$ is an R-vine leaf in $\mathcal{V}^s$, we know that $U_s \notin A_{v_2}$. Therefore, it must hold that

$$pt_1^{U_s}, \ldots, pt_{k+1}^{U_s} \in A_{v_2}, \text{ i.e. } pt_{k+1}^{U_s} \in \{ n \in N_1^{s-1} \setminus \{V, pt_1^{U_s}, \ldots, pt_{k}^{U_s}\} : \exists e \in E_k^{s-1}, \text{ s.t. } n, pt_1^{U_s}, \ldots, pt_{k}^{U_s} \in A_e \}.$$  

□

In Example 5.7 and the proof of Corollary 5.9 we have used a specific notation for “new” or “added” edges in tree $T_k^s$. We close this section by generalizing this notation.

**Definition 5.10.** (New edge in tree $T_k^s$)

Let $s \geq 2$, $k = 1, \ldots, s - 1$, and let $T_k^s$ be the $k$-th tree of the extended R-vine tree sequence $\mathcal{V}^s$. The new edge $e_k^s$ in tree $T_k^s$ is defined as

$$e_k^s := E_k^s \setminus E_k^{s-1}.$$
Chapter 6

Algorithm - forward selection of an R-vine copula

In the following we summarize the developed theory and present an algorithm used to select an estimated R-vine copula describing the data well enough to predict the conditional quantiles of a response variable $V$ as accurate as possible. Therefore, assume we are given observations

- of a response variable $V$
- of a set of potential covariates $U_1^p, \ldots, U_J^p$.

Therefore, we assume that

- either a fixed and ordered set of covariates $U_1 := U_{j_1}^p, \ldots, U_d := U_{j_d}^p$, $d \leq J$ (FIXED)
- or an order $j_1 < \cdots < j_J$ (ORDER)

has been defined in advance (e.g. by Algorithm 1 or by D-vine quantile regression). While we

- include all variables $U_1, \ldots, U_d$ in (FIXED),
- we stop the inclusion according to a conditional likelihood ratio test (Definition 5.2) in (ORDER).

We will use a dummy variable

$$fC = \begin{cases} 
1, & \text{in (FIXED)}, \\
0, & \text{in (ORDER)} 
\end{cases}$$

in order to distinguish the cases (FIXED) and (ORDER).

We state the pseudo code and illustrate it in an example using the daxreturns data described in Example 5.3. First, we present an overview over the algorithm’s input data and parameters.
6.1 Parameters and data input

The parameters and data input which need to be defined or available prior to Algorithm 2 are given by

- the dummy variable $fC \in \{0, 1\}$ denoting which case of (FIXED) and (ORDER) we are facing,
- a data set containing observations of
  - a response variable $V: v \in [0, 1]^n$,
  - either all covariates $U_1, \ldots, U_d: \boldsymbol{u}_1, \ldots, \boldsymbol{u}_d \in [0, 1]^n$ (FIXED)
  - or all potential covariates $U_1 := U^n_{p1}, \ldots, U_J := U^n_{pj}; \boldsymbol{u}_1, \ldots, \boldsymbol{u}_J \in [0, 1]^n$ (ORDER),
- the set of possible copula families $B$ (Definition 1.36),
- the selection criterion $sc \in \{\ell, \text{AIC, BIC}\}$ (Definition 1.36),
- the weight criterion $wc \in \{\tau, \rho^{\max}\}$ (Definition 5.8),
- the independence test dummy variable $\text{indepTest} \in \{0, 1\}$ (Remark 1.38),
- the significance level $\alpha_{LRT}$ used in the conditional likelihood ratio test (Definition 5.2) (only relevant for case (ORDER)).

6.2 Pseudo code

The goal is to find an estimated R-vine copula

$$\hat{R}^S = (\mathcal{V}^S, \hat{\mathcal{B}}^S(\mathcal{V}^S), \hat{\Theta}^S(\hat{\mathcal{B}}^S(\mathcal{V}^S))), \mathcal{V}^S = (T^S_1, \ldots, T^S_S), T^S_k = (N^S_k, E^S_k), k = 1, \ldots, S,$$

where

$$S = d, \text{ if } fC = 1 \text{ and } S \in \{1, \ldots, J\}, \text{ if } fC = 0.$$

The corresponding pseudo code can be found in Algorithm 2.

**Algorithm 2** Pseudo code - forward selection of regular vine copula.

1: $s = 1$
2: (i) $pt^1_Y := U_1$
3: (ii) $N^1 := \{V, U_1\}, E^1 := \{(V, U_1)\}, \mathcal{V}^1 := T^1 := (N^1, E^1)$
4: (iii) $e = (P_{e,1}, P_{e,2}) := e^1 := (V, U_1)$
5: (a) $U^n_e := (P_{e,1}, P_{e,2}) = (v, u_1)$
(b) $C_{bh}^e(\cdot, \cdot; \hat{\theta}_e) := \text{optC}(U_n^e, B, sc)$, $\hat{b}_{e_1}^e := \hat{b}_e$, $\hat{\theta}_{e_1}^e := \hat{\theta}_e$; 
(c) $\hat{h}_{P_{e_1}|P_{e_2}} := h_1(U_n^e, C_{bh}^e(\cdot, \cdot; \hat{\theta}_e))$, $\hat{h}_{P_{e_2}|P_{e_1}} := h_2(U_n^e, C_{bh}^e(\cdot, \cdot; \hat{\theta}_e))$; 
(d) $\hat{P}_1^e := \{\hat{h}_{P_{e_1}|P_{e_2}}, \hat{h}_{P_{e_2}|P_{e_1}}\}$

(iv) $\hat{B}^1 := \{\hat{b}_e : e = e_1^1\}$, $\hat{\Theta}^1 := \{\hat{\theta}_e : e = e_1^1\}$, $\hat{R}^1 := (V^1, \hat{B}^1, \hat{\Theta}^1)$

(v) $c\ell \ell := c\ell\ell(\hat{R}^1, v, u_1)$, $c\text{AIC}^1 := c\text{AIC}(\hat{R}^1, v, u_1)$, $c\text{BIC}^1 := c\text{AIC}(\hat{R}^1, v, u_1)$

for $2 \leq s \leq J$ do:

- $k = 1$:

(i) $\text{pt}_{U^e}^1 := \max W_1(U_s, \{U_1, \ldots, U_{s-1}\}, \text{wc})$;

(ii) $N_1^s := N_1^{s-1} \cup \{U_s\}$, $E_1^s := E_2^{s-1} \cup \{(U_s, \text{pt}_{U^e}^1)\}$, $T_1^s := (N_1^s, E_1^s)$;

(iii) $e = (P_{e_1}, P_{e_2}) := e_1^s := (U_s, \text{pt}_{U^e}^1)$;

(a) $U_e^s := (P_{e_1}, P_{e_2})$;

(b) $C_{bh}^e(\cdot, \cdot; \hat{\theta}_e) := \text{optC}(U_n^e, B, sc)$, $\hat{b}_{e_1}^e := \hat{b}_e$, $\hat{\theta}_{e_1}^e := \hat{\theta}_e$;

(c) $\hat{h}_{P_{e_1}|P_{e_2}} := h_1(U_n^e, C_{bh}^e(\cdot, \cdot; \hat{\theta}_e))$, $\hat{h}_{P_{e_2}|P_{e_1}} := h_2(U_n^e, C_{bh}^e(\cdot, \cdot; \hat{\theta}_e))$;

(d) $\hat{P}_1^s := \{\hat{h}_{P_{e_1}|P_{e_2}}, \hat{h}_{P_{e_2}|P_{e_1}}\}$

for $2 \leq k \leq s$ do:

(i.1) if $2 \leq k \leq s - 1$ then:

\[ \text{PT}_k^s := \{n \in N_1^{s-1} \setminus \{V, \text{pt}_{U^e}^1, \ldots, \text{pt}_{U^e}^k\} : \exists e \in E_{k-1}^s, \text{ s.t. } n, \text{pt}_{U^e}^1, \ldots, \text{pt}_{U^e}^k \in A_e\}; \]

\[ \text{pt}_{U^e}^k := \max W_k(U_s, \text{PT}_k, \text{wc}); \]

(i.2) if $k = s$ then: $\text{pt}_{U^s}^s := U_s$;

(ii.1) if $2 \leq k \leq s - 1$ then: $e_k^s := (U_s, \text{pt}_{U^e}^k | \text{pt}_{U^e}^1, \ldots, \text{pt}_{U^e}^k)$;

\[ N_k^s := E_{k-1}^s, \quad E_k^s := E_{k-1}^s \cup \{e_k^s\}, \quad T_k^s := (N_k^s, E_k^s); \]

(ii.2) if $k = s$ then: $e_s^s := (V, U_s | U_1, \ldots, U_{s-1})$;

\[ N_s^s := E_{s-1}^s, \quad E_s^s := \{e_s^s\}, \quad T_s^s := (N_s^s, E_s^s), \quad V^s = (T_1^s, \ldots, T_s^s); \]

(iii) Set $e := e_k^s$ and find permutation vectors $\pi, \tilde{\pi} \in \{1, \ldots, k - 1\}^{k-1}$, s.t.

\[ \hat{h}_{P_{e_1}|D_{e_1}D_{e_2} \ldots D_{e_{s-1}}1}, \hat{h}_{P_{e_2}|D_{e_1}D_{e_2} \ldots D_{e_{s-1}}1} \in \hat{P}_{k-1}^s; \]

(a) $U_n^e := (\hat{h}_{P_{e_1}|D_{e_1}D_{e_2} \ldots D_{e_{s-1}}1}, \hat{h}_{P_{e_2}|D_{e_1}D_{e_2} \ldots D_{e_{s-1}}1})$;

(b) $C_{bh}^e(\cdot, \cdot; \hat{\theta}_e) := \text{optC}(U_n^e, B, sc)$, $\hat{b}_{e_1}^e := \hat{b}_e$, $\hat{\theta}_{e_1}^e := \hat{\theta}_e$;

(c) $\hat{h}_{P_{e_1}|P_{e_2}|D_e} := h_1(U_n^e, C_{bh}^e(\cdot, \cdot; \hat{\theta}_e))$, $\hat{h}_{P_{e_2}|P_{e_1}|D_e} := h_2(U_n^e, C_{bh}^e(\cdot, \cdot; \hat{\theta}_e))$;
(d.1) if $2 \leq k \leq s - 1$ then: $\hat{P}_k^s := \hat{P}_k^{s-1} \cup \{\hat{h}_{P_e,1|P_{e,2};D_e}, \hat{h}_{P_e,2|P_{e,1};D_e}\};$

(d.2) if $k = s$ then: $\hat{P}_s^s := \{\hat{h}_{P_e,1|P_{e,2};D_e}, \hat{h}_{P_e,2|P_{e,1};D_e}\};$

(iv) $\mathcal{E}^s := \bigcup_{k=1}^{s} \{e_k\}, \hat{B}^s := \hat{B}^{s-1} \cup \{\hat{b}_e : e \in \mathcal{E}^s\}, \hat{\Theta}^s := \hat{\Theta}^{s-1} \cup \{\hat{\theta}_e : e \in \mathcal{E}^s\},$

$\mathcal{R}^s = (V^s, \hat{B}^s, \hat{\Theta}^s);$

(v) $c\ell^s := c\ell(\mathcal{R}^s, v, (\mathbf{u}_1, \ldots, \mathbf{u}_s)), c\text{AIC}^s := c\text{AIC}(\mathcal{R}^s, v, (\mathbf{u}_1, \ldots, \mathbf{u}_s)),$

$c\text{BIC}^s := c\text{BIC}(\mathcal{R}^s, v, (\mathbf{u}_1, \ldots, \mathbf{u}_s));$

(vi.1) if $f_C = 0$ then: STOP, if $s = d;$

(vi.2) if $f_C = 1$ then: STOP, if $c\ell^s - c\ell^{s-1} \leq X^2_{1-\alpha, \text{LRT}}, |\hat{\Theta}^s| - |\hat{\Theta}^{s-1}|.$

### 6.3 Illustration

**Example 6.1.** (Dax returns - Illustration of forward selection of an R-vine copula)

In the following, we illustrate Algorithm 2 using the data set daxreturns described in Example 5.3. The response variable is chosen to be $V := \text{BMW}.$

Again we distinguish between the following two cases:

(FIXED) All elements in the fixed set of covariates

$$\{U_1 = \text{DAI}, U_2 = \text{DBK}, U_3 = \text{VOW3}, U_4 = \text{SIE}\}$$

chosen by Algorithm 1 are included in the R-vine copula $\mathcal{R}^4.$

ORDER The potential covariate $U_{p_j}^s$ is included in the R-vine copula $\mathcal{R}^s$ in step $s \geq 1$ if the conditional likelihood test between $\mathcal{R}^s$ and $\mathcal{R}^{s-1}$ is not rejected. They are included in the order

$$4 < 5 < 14 < 13 < 1 < 8 < 2 < 7 < 3 < 9 < 10 < 12 < 6 < 11,$$

with

- $U_p^5 = \text{DAI},$
- $U_p^6 = \text{DBK},$
- $U_{14}^p = \text{VOW3},$
- $U_{13}^p = \text{SIE},$
- $U_1^p = \text{ALV},$
- $U_8^p = \text{FME},$
- $U_2^p = \text{BAS},$
- $U_7^p = \text{EOAN},$
- $U_9^p = \text{BAYN},$
- $U_6^p = \text{LIN}.$
\[-U_{10}^p = MIUV2,\]
\[-U_{12}^p = SAP,\]
\[-U_6^p = DTE,\]
\[-U_{11}^p = RWE,\]

selected by Algorithm 1.

According to Example 5.3, we are given \(n = 1158\) observations of the response \(BMW\) and each potential covariate \(U_1^p, \ldots, U_{14}^p\). The vectors containing those observations are denoted by small bold letters.

The parameters were chosen as

- \(B := \{\text{Independence (I), Gaussian (N), Student t (t), Clayton (C), Gumbel (G), Frank (F), Joe (J)}\}\),
- \(sc := \text{AIC}\),
- \(wc := \rho_{\text{max}}^n\),
- \(\text{indepTest} = 0\),
- \(\alpha_{LRT} = 0.05\).

\(s = 1:\)

In both cases (FIXED) and (ORDER), the first two variables which are included in the R-vine copula \(\mathcal{R}_1^1\) are the response \(BMW\) and \(U_1 = U_{j_1}^p = DAI\). For two variables there is one unique R-vine tree sequence \(\mathcal{V}^1 = T_1^1\) as shown in Figure 6.1.

![Figure 6.1: \(\mathcal{V}^1 = (T_1^1)\) in example daxreturns.](image)

In terms of the notation used in Algorithm 2 this translates to:

(i) \(pt_{1, BMW}^1 = U_1 = U_{j_1}^p = U_4^p = DAI\),
(ii) \(N_1^1 = \{V, U_1\} = \{BMW, DAI\}, \ E_1^1 = \{e_1^1\} = \{(V, U_1)\} = \{(BMW, DAI)\}, \ \mathcal{V}^1 = T_1^1 = (N_1^1, E_1^1)\).
(iii) Next, pair copula family, pair copula parameters and pseudo observations corresponding to the “new” edge

\[e = (P_{e,1}, P_{e,2}) := e_1^1 = (BMW, DAI)\]

are estimated.
(a) The copula data matrix which is used for the estimation is given by
\[ U_{e}^{1158} := (P_{e,1}, P_{e,2}) = (bmw, dai) \in [0, 1]^{1158 \times 2}. \]

(b) Estimating \( C_{\theta}^{b}(\cdot, \cdot; \tilde{\theta}_e) := \text{optC}(U_{e}^{1158}, B, \text{AIC}) \) by the R-function \( \text{BiCopSelect()} \), we obtain the following family and parameters:
\[ \hat{b}_{e} \hat{\theta}_{e} := \hat{b}_{e} \hat{\theta}_{e} = \begin{pmatrix} \hat{b}_{e} \\ \hat{\theta}_{e} \end{pmatrix} = \begin{pmatrix} t \\ \begin{pmatrix} 0.68 \\ 8.77 \end{pmatrix} \end{pmatrix}. \]

(c) Using the result from b) we can estimate the pseudo observations
\[ \hat{h}_{P_{e,1}|P_{e,2}} = \hat{h}_{BMW|DAI} := h_{1}(U_{e}^{1158}, C_{\theta}^{b}(\cdot, \cdot; \hat{\theta}_e)) \in [0, 1]^{1158} \]
and
\[ \hat{h}_{P_{e,2}|P_{e,1}} = \hat{h}_{DAI|BMW} := h_{2}(U_{e}^{1158}, C_{\theta}^{b}(\cdot, \cdot; \hat{\theta}_e)) \in [0, 1]^{1158}. \]

(d) Those pseudo observations are added to the set \( P_{1}^{1} \), i.e.
\[ \tilde{P}_{1} := \{ \hat{h}_{BMW|DAI}, \hat{h}_{DAI|BMW} \}. \]

(v) Finally, our first R-vine copula is given by \( \hat{R}_{1} = (V_{1}, B_{1}, \hat{\Theta}_{1}) \) with
\[ B_{1} := \{ \hat{b}_{e} \} = \{ t \}, \hat{\Theta}_{1} := \{ \hat{\theta}_{e} \} = \{ \begin{pmatrix} 0.68 \\ 8.77 \end{pmatrix} \}. \]

(vi) The goodness of fit measures corresponding to \( \hat{R}_{1} \) are given by
\[ -c_{\ell}^{\ell} := c_{\ell}(\hat{R}_{1}, bmw, dai) = 357.3561, \]
\[ -c_{\text{AIC}}^{\ell} := c_{\text{AIC}}(\hat{R}_{1}, bmw, dai) = 710.7122, \]
\[ -c_{\text{BIC}}^{\ell} := c_{\text{BIC}}(\hat{R}_{1}, bmw, dai) = -700.6033. \]

\( s = 2 \):

Now, we add \( U_{2} = U_{j_2}^{p} = DBK \) to the regular vine copula \( \hat{R}_{2} \). The corresponding tree sequence \( V_{2} = (T_{1}, T_{2}) \) needs to be found iteratively.

\[- k = 1: \]

(i) Since the response variable \( BMW \) is supposed to be an R-vine leaf (Definition 2.13), the second covariate \( U_{2} = U_{j_2} = DBK \) can only be neighbor of \( DAI \), i.e.
\[ pt_{1}^{DBK} \in PT_{1}^{2} = N_{1}^{1} \setminus \{ BMW \} = \{ DAI \}, \]
\[ \Rightarrow pt_{1}^{DBK} := \max_{W_{1}}(DBK, \{ DAI \}, \text{AIC}) = DAI. \]
(ii)
\[
N_1^2 := N_1^1 \cup \{DBK\} = \{BMW, DAI, DBK\},
E_1^2 := E_1^1 \cup \{(DBK, DAI)\} = \{(BMW, DAI), (DBK, DAI)\}, \\
T_1^2 := (N_1^1, E_1^1).
\]

(iii) Next, pair copula family, pair copula parameters and pseudo observations corresponding to the “new” edge
\[ e = (P_{e,1}, P_{e,2}) := e_1^2 = (DBK, DAI) \]
are estimated.

(a) The copula data matrix which is used for the estimation is given by
\[
U_{e_1}^{1158} := (P_{e,1}, P_{e,2}) = (\text{dbk, dai}) \in [0, 1]^{1158 \times 2}.
\]

(b) Estimating \( \hat{C}^{\hat{b}^e}_{\hat{\theta}^e}(\cdot, \cdot; \hat{\theta}^e) := \text{optC}(U_{e_1}^{1158}, B, AIC) \) by the R-function BiCopSelect(), we obtain the following family and parameters:
\[
\hat{b}_e := t, \quad \hat{\theta}_e = \left( \frac{\hat{\theta}_1^e}{\hat{\theta}_2^e} \right) = \left( \frac{0.64}{5.64} \right).
\]

(c) Using the result from b) we can estimate the pseudo observations
\[
\hat{h}_{P_{e,1}|P_{e,2}} = \hat{h}_{DBK|DAI} := h_1(U_{e_1}^{1158}, C^{\hat{b}^e}_{\hat{\theta}^e}(\cdot, \cdot; \hat{\theta}^e)) \in [0, 1]^{1158}
\]
and
\[
\hat{h}_{P_{e,2}|P_{e,1}} = \hat{h}_{DAI|DBK} := h_2(U_{e_2}^{1158}, C^{\hat{b}^e}_{\hat{\theta}^e}(\cdot, \cdot; \hat{\theta}^e)) \in [0, 1]^{1158}.
\]

(d) Those pseudo observations are added to the set \( \hat{P}_1^2 \), i.e.
\[
\hat{P}_1^2 := \hat{P}_1^1 \cup \{\hat{h}_{DBK|DAI}, \hat{h}_{DAI|DBK}\}.
\]

- \( k = 2 \):

(i.2) We are in the case \( k = s = 2 \) and set
\[
pt_{2}^{BMW} := U_2 = U_{P_2}^p = DBK.
\]

(ii.2) The “new” edge is given by
\[
e_2^2 = (BMW, DBK|DAI).
\]

It follows that
\[
N_1^2 := E_1^2, \quad E_2^2 = \{(BMW, DBK|DAI)\}, \quad T_2^2 = (N_2^1, E_2^1).
\]
Then, $V^2 = (T_1^2, T_2^2)$ looks as shown in Figure 6.2.

![Diagram](image)

Figure 6.2: $V^2 = (T_1^2, T_2^2)$ in example daxreturns.

(iii) Set 

$e = (P_{e,1}, P_{e,2}|D_{e,1}) := e_2^2 = (BMW, DBK|DAI)$

and remember that 

$\hat{P}_{k-1}^* = \hat{P}_1^2 := \{\hat{h}_{BMW|DAI}, \hat{h}_{DAI|BMW}, \hat{h}_{DBK|DAI}, \hat{h}_{DAI|DBK}\}$.

We observe that in $s = 2$, we have computed the pseudo observations 

$\hat{h}_{P_{e,1}|D_{e,1}} = \hat{h}_{BMW|DAI}, \hat{h}_{P_{e,2}|D_{e,1}} = \hat{h}_{DBK|DAI} \in \hat{P}_{k-1}^* = \hat{P}_1^2$.

Thus, we choose 

$\pi = \pi_1 = \hat{\pi} = \hat{\pi}_1 = 1 \in \{1\}^1$, s.t. 

$\hat{h}_{P_{e,1}|D_{e,1}}, \hat{h}_{P_{e,2}|D_{e,1}} \in \hat{P}_1^2$.

Next, pair copula family, pair copula parameters and pseudo observations corresponding to the “new” edge 

$e = (P_{e,1}, P_{e,2}|D_{e,1}) := e_2^2 = (BMW, DBK|DAI)$

are estimated.

(a) The copula data matrix which is used for the estimation is given by 

$U_{e}^{1158} := (\hat{h}_{P_{e,1}|D_{e,1}}, \hat{h}_{P_{e,2}|D_{e,1}}) = (\hat{h}_{BMW|DAI}, \hat{h}_{DAI|BMW}) \in [0, 1]^{1158\times 2}$.

(b) Estimating $C_{e}^{b_\hat{\theta}_e}(\cdot, \cdot; \hat{\theta}_e) := \hat{optC}(U_{e}^{1158}, B, AIC)$ by the R-function BiCopSelect(), we obtain the following family and parameters:

$\hat{b}_e^1 := \hat{b}_e = F, \hat{\theta}_e^1 := \hat{\theta}_e = \left(\begin{array}{c}
\hat{\theta}_{e1}^1 \\
\hat{\theta}_{e2}^2
\end{array}\right) = \left(\begin{array}{c}
1.9 \\
0
\end{array}\right)$.

(c) Using the result from b) we can estimate the pseudo observations 

$\hat{h}_{P_{e,1}|P_{e,2}|D_{e,1}} = \hat{h}_{BMW|DBK|DAI} := h_1(U_{e}^{1158}, C_{e}^{b_\hat{\theta}_e}(\cdot, \cdot; \hat{\theta}_e)) \in [0, 1]^{1158}$

and 

$\hat{h}_{P_{e,2}|P_{e,1}|D_{e,1}} = \hat{h}_{DBK|BMW|DAI} := h_2(U_{e}^{1158}, C_{e}^{b_\hat{\theta}_e}(\cdot, \cdot; \hat{\theta}_e)) \in [0, 1]^{1158}$.
(d) Those pseudo observations are added to the set \( \hat{\mathcal{P}}_2 \), i.e.

\[
\hat{\mathcal{P}}_2 := \{ \hat{h}_{BMW|DBK;DAI}, \hat{h}_{DBK|BMW;DAI} \}.
\]

(v) The second R-vine copula is given by \( \hat{\mathcal{R}}^2 = (\mathcal{Y}^2, \hat{\mathcal{B}}^2, \hat{\Theta}^2) \), with

\[
\hat{\mathcal{B}}^2 := \mathcal{B}^1 \cup \{ \hat{b}_e : e \in e^2_1 \cup e^2_2 \} = \{ t, t, F \}
\]

and

\[
\hat{\Theta}^2 := \hat{\Theta}^1 \cup \{ \hat{\theta}_e : e \in e^2_1 \cup e^2_2 \} = \{ (0.68, 8.77), (0.64, 5.64), (1.9, 0) \}.
\]

(vi) The goodness of fit measures corresponding to \( \hat{\mathcal{R}}^2 \) are given by

- \( c_{\ell\ell}^2 := c_{\ell\ell}(\hat{\mathcal{R}}^2, bmw, (dai, dbk)) = 412.1252 \),
- \( c_{\text{AIC}}^2 := c_{\text{AIC}}(\hat{\mathcal{R}}^2, bmw, (dai, dbk)) = 814.2504 \),
- \( c_{\text{BIC}}^2 := c_{\text{BIC}}(\hat{\mathcal{R}}^2, bmw, (dai, dbk)) = -788.9781 \).

Observe that

\[
|\hat{\Theta}^2| - |\hat{\Theta}^1| = |\{ \hat{\theta}_e : e \in e^2_1 \cup e^2_2 \}| = 2 + 1 = 3
\]

and

\[
c_{\ell\ell}^2 - c_{\ell\ell}^1 = 412.1252 - 357.3561 = 54.7691.
\]

Then, the conditional likelihood test between \( \hat{\mathcal{R}}^2 \) and \( \hat{\mathcal{R}}^1 \) is rejected at a level \( \alpha_{LRT} = 0.05 \), since

\[
c_{\ell\ell}^2 - c_{\ell\ell}^1 = 54.7691 > \chi^2_{1-\alpha_{LRT}, \|\hat{\Theta}^2| - |\hat{\Theta}^1|} = \chi^2_{0.95, 3} = 7.8147.
\]

The corresponding p-value is given by \( p^2 = 7.690893e - 12 \).

Thus, \( U_{j2} = U_2 = DBK \) is included in \( \hat{\mathcal{R}}^2 \) for \( fC \in \{0, 1\} \).

- **s = 3:**

In step \( s = 3 \), we add \( U_3 = U_{j3} = VOW3 \) to the regular vine copula.

- **k = 1:**

  (i) The set of partner candidates for \( VOW3 \) in tree \( T^3_1 \) is given by

\[
\text{PT}^3_1 = N^2_1 \setminus \{BMW\} = \{DAI, DBK\}.
\]

Since \( |\text{PT}^3_1| > 1 \), we need to make use of the weight criterion \( wc = \rho^{max} \).
Table 6.1: $\hat{\rho}_{1158,\max}^{\text{VOW}_3,\text{U}}, \ U \in \{\text{DAI, DBK}\}$, $\hat{\rho}_{\text{SIE},\max}^{\text{U}}, \ U \in \{\text{DAI, DBK, VOW}_3\}$.

<table>
<thead>
<tr>
<th>$\hat{\rho}_{1158,\max}^{\text{VOW}_3,\text{U}}$</th>
<th>DAI</th>
<th>DBK</th>
<th>VOW3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2343</td>
<td>0.2011</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

According to Table 6.1 which contains the maximal empirical semi-correlations $\hat{\rho}_{1158,\max}^{\text{VOW}_3,\text{U}}$ for $\ U \in \{\text{DAI, DBK}\}$, it holds that

$$\text{pt}_{1}^{\text{VOW}_3} = \max W_{1}(\text{VOW}_3, \{\text{DAI, DBK}\}; \rho^{\max}) = \text{DAI},$$

since

$$\hat{\rho}_{\text{VOW}_3,\text{DAI}}^{1158,\max} = 0.2343 > \hat{\rho}_{\text{VOW}_3,\text{DAI}}^{1158,\max} = 0.2011.$$  

(ii) \begin{align*}
N_{3} & := N_{1}^{2} \cup \{\text{VOW}_3\} = \{\text{BMW, DAI, DBK, VOW}_3\}, \\
E_{3} & := E_{1}^{2} \cup \{\text{VOW}_3, \text{DAI}\} \\
 & = \{(\text{BMW, DAI}), (\text{DBK, DAI}), (\text{VOW}_3, \text{DAI})\}, \\
T_{3} & := (N_{3}^{1}, E_{3}^{2}).
\end{align*}

(iii) Next, pair copula family, pair copula parameters and pseudo observations corresponding to the “new” edge
e = (P_{e,1}, P_{e,2}) := e_{1}^{3} = (\text{VOW}_3, \text{DAI})

are estimated.

(a) The copula data matrix, which is used for the estimation, is given by

$$\mathcal{U}_{e}^{1158} := (P_{e,1}, P_{e,2}) = (\text{vow3}, \text{dai}) \in [0,1]^{1158 \times 2}.$$ 

(b) Estimating $C_{k}^{\hat{\theta}_{e}}(\cdot, \cdot; \hat{\theta}_{e}) := \text{optC}(\mathcal{U}_{e}^{1158}, B, \text{AIC})$ by the R-function $\text{BiCopS\text{-}select()}$, we obtain the following family and parameters:

$$\hat{b}_{e_{1}} := \hat{b}_{e} = t, \quad \hat{\theta}_{e_{1}} := \hat{\theta}_{e} = \left(\frac{\hat{\theta}_{1}^{1}}{\hat{\theta}_{2}^{1}}\right) = \left(0.53, 11.27\right).$$

Note that this is the first case, that we would have obtained different estimates for another selection criteria $sc$. For example for $sc = \text{BIC}$, we would have obtained

$$\hat{b}_{e_{1}} := \hat{b}_{e} = F, \quad \hat{\theta}_{e_{1}} := \hat{\theta}_{e} = \left(\frac{\hat{\theta}_{1}^{1}}{\hat{\theta}_{2}^{1}}\right) = \left(3.75, 0\right).$$

However, in the following we will stick to the estimates obtained for $sc = \text{AIC}$. 

84
(c) Using the result from b) we can estimate the pseudo observations
\[ \hat{h}_{P_e,1|P_e,2} = \hat{h}_{VOW3|DAI} := h_1(\mathcal{U}_e^{\text{DAI}}, C^b(\cdot, \cdot, \theta_e)) \in [0, 1]\]
and
\[ \hat{h}_{P_e,2|P_e,1} = \hat{h}_{DAI|VOW3} := h_2(\mathcal{U}_e^{\text{DAI}}, C^b(\cdot, \cdot, \theta_e)) \in [0, 1]\].

(d) Those pseudo observations are added to the set \( \hat{P}_1^3 \), i.e.
\[ \hat{P}_1^3 := \hat{P}_1^2 \cup \{ \hat{h}_{VOW3|DAI}, \hat{h}_{DAI|VOW3} \}. \]

- \( k = 2: \)

(i.1) The set of partner candidates for \( VOW3 \) in tree \( T_2^3 \) is given by
\[ PT_2^3 = \{ n \in N^2_1 \setminus \{ BMW, pt_1^{VOW3} \} : \exists e \in E_1^2, \text{ s.t. } n, pt_1^{VOW3} \in A_e \} = \{ DBK \}, \]
since
\[ N^2_1 \setminus \{ BMW, pt_1^{VOW3} \} = \{ DBK \} \]
and \( (DBK, DAI) = (DBK, pt_1^{VOW3}) \in E_1^2. \)

Thus, the partner of \( VOW3 \) in \( T_2^3 \) is given by
\[ pt_2^{VOW3} := \max W_2(\text{VOW3}, \{ DBK \}, \rho^{\text{max}}) = DBK. \]

(ii.1) \( N^3_2 = E_1^3, E_2^3 = E_1^2 \cup \{ (\text{VOW3}, DBK|DAI) \}, T_2^3 = (N^3_2, E_2^3). \)

(iii) Set
\[ e = (P_{e,1}, P_{e,2}|D_{e,1}) := e_2^3 = (\text{VOW3}, DBK|DAI) \]
and remember that
\[ \hat{P}_{k-1}^3 = \hat{P}_1^3 \]
\[ = \{ \hat{h}_{BMW|DAI}, \hat{h}_{DAI|BMW}, \hat{h}_{DBK|DAI}, \hat{h}_{DAI|DBK}, \hat{h}_{VOW3|DAI}, \hat{h}_{DAI|VOW3} \}. \]

Since \( D_e = D_{e,1} = DAI \), we consider all pseudo observation vectors that are conditioned on \( DAI \) and observe that
\[ \hat{h}_{P_{e,1}|D_{e,1}} = \hat{h}_{VOW3|DAI}, \hat{h}_{P_{e,2}|D_{e,1}} = \hat{h}_{DBK|DAI} \]
are elements of \( \hat{P}_1^3 \). Since there is only one element in \( D_e \), we need to set
\[ \pi = \pi_1 = \tilde{\pi} = \tilde{\pi}_1 = 1 \in \{ 1 \}^1. \]

However, this leads us to the desired pseudo observations corresponding to
\[ P_{e,1}|D_{e,\pi_1} = VOW3|DAI \text{ and } P_{e,2}|D_{e,\tilde{\pi}_1} = DBK|DAI. \]

85
Next, pair copula family, pair copula parameters and pseudo observations corresponding to the “new” edge
e = (P_{e,1}, P_{e,2}|D_{e,1}) := e^2 = (VOW3, DBK|DAI)
are estimated.

(a) The copula data matrix which is used for the estimation is given by
\[ U_{e}^{1158} := (\hat{h}_{P_{e,1}|D_{e,1}}, \hat{h}_{P_{e,2}|D_{e,1}}) = (\hat{h}_{VOW3|DAI}, \hat{h}_{DBK|DAI}) \in [0, 1]^{1158 \times 2}. \]

(b) Estimating \( C^{b_{e}}(\cdot, \cdot; \hat{\theta}_{e}) := \text{optC}(U_{e}^{1158}, B, \text{AIC}) \) by the R-function \text{BiCopS-elect()}, we obtain the following family and parameters:
\[ \hat{b}_{e} := \hat{b}_{e} = F, \quad \hat{\theta}_{e} := \left( \hat{\theta}^{1}_{e}, \hat{\theta}^{2}_{e} \right) = \left( 0.93, 0 \right). \]

(c) Using the result from b) we can estimate the pseudo observations
\[ \hat{h}_{P_{e,1}|P_{e,2};D_{e,1}} = \hat{h}_{VOW3|DBK;DAI} := h_{1}(U_{e}^{1158}, C^{b_{e}}(\cdot, \cdot; \hat{\theta}_{e})) \in [0, 1]^{1158} \]
and
\[ \hat{h}_{P_{e,2}|P_{e,1};D_{e,1}} = \hat{h}_{DBK|VOW3;DAI} := h_{2}(U_{e}^{1158}, C^{b_{e}}(\cdot, \cdot; \hat{\theta}_{e})) \in [0, 1]^{1158}. \]

(d) Those pseudo observations are added to the set \( \hat{P}_{3} \), i.e.
\[ \hat{P}_{2} := \hat{P}_{2} \cup \{ \hat{h}_{VOW3|DBK;DAI}, \hat{h}_{DBK|VOW3;DAI} \}. \]

– \( k = 3 \):

(i.2) We are in the case \( k = s = 3 \) and set
\[ p_{3}^{BMW} := U_{3} = U_{3}^{p} = VOW3. \]

(ii.2) The “new” edge is given by
\[ e_{2}^{3} = (BMW, VOW3|DAI, DBK). \]

It follows that
\[ N_{3}^{3} := E_{3}^{3}, \quad E_{3}^{3} = \{ (BMW, VOW3|DAI, DBK) \}, \quad T_{3}^{3} = (N_{3}^{3}, E_{3}^{3}). \]

Then, \( V^{3} = (T_{1}^{3}, T_{2}^{3}, T_{3}^{3}) \) looks as shown in Figure 6.3.
Figure 6.3: $\mathcal{V}^3 = (T_1^3, T_2^3, T_3^3)$ in example $\text{daxreturns}$.

(iii) Set

$$e = (P_{e,1}, P_{e,2}|D_{e,1}, D_{e,2}) := e_3^3 = (BMW, VOW3|DAI, DBK)$$

and remember that

$$\hat{P}_{k-1}^* = \hat{P}_2^3 = \{\hat{h}_{BMW|DBK;DAI}, \hat{h}_{DBK|BMW;DAI}, \hat{h}_{VOW3|DBK;DAI}, \hat{h}_{DBK|VOW3;DAI}\}.$$

We observe that

$$\hat{h}_{P_{e,1}|D_{e,2};D_{e,1}} = \hat{h}_{BMW|DBK;DAI}, \hat{h}_{P_{e,2}|D_{e,2};D_{e,1}} = \hat{h}_{VOW3|DBK;DAI} \in \hat{P}_{k-1}^* = \hat{P}_2^3.$$

Since now we have two elements in the conditioning set $D_e = \{D_{e,1}, D_{e,2}\} = \{DAI, DBK\}$, we need to order its elements correctly, such that the pseudo observations corresponding to $P_{e,1} = BMW$ and $P_{e,2} = VOW3$ are part of $\hat{P}_2^3$. For example, for $P_{e,2} = VOW3$ we check if either $\hat{h}_{P_{e,2}|D_{e,1};D_{e,2}} = \hat{h}_{P_{e,2}|DAI;DBK}$ or $\hat{h}_{P_{e,2}|D_{e,2};D_{e,1}} = \hat{h}_{P_{e,1}|DBK;DAI}$ is part of $\hat{P}_2^3$.

As $\hat{h}_{P_{e,1}|D_{e,2};D_{e,1}} = \hat{h}_{BMW|DBK;DAI}$ and $\hat{h}_{P_{e,2}|D_{e,2};D_{e,1}} = \hat{h}_{P_{e,1}|DBK;DAI}$ are elements of $\hat{P}_2^3$, we set

$$\pi = (\pi_1, \pi_2)^T = \tilde{\pi} = (\tilde{\pi}_1, \tilde{\pi}_2)^T = (2, 1)^T \in \{1, 2\}^2, \text{ s.t.}$$

$$\hat{h}_{P_{e,1}|D_{e,\pi_1};D_{e,\pi_2}}, \hat{h}_{P_{e,2}|D_{e,\pi_1};D_{e,\pi_2}} \in \hat{P}_2^3.$$

Next, pair copula family, pair copula parameters and pseudo observations corresponding to the “new” edge

$$e = (P_{e,1}, P_{e,2}|D_{e,1}, D_{e,2}) := e_3^3 = (BMW, VOW3|DAI, DBK)$$

are estimated.
(a) The copula data matrix which is used for the estimation is given by
\[ U_{e}^{1158} := (\hat{h}_{P_{e,1}|D_{e,\pi_1};D_{e,\pi_2}}, \hat{h}_{P_{e,2}|D_{e,\pi_1};D_{e,\pi_2}}) = (\hat{h}_{BMW|DBK,DAI}, \hat{h}_{VOW|DBK,DAI}) \in [0,1]^{1158 \times 2}. \]

(b) Estimating \( C_{\hat{b}_e}(\cdot,\cdot;\hat{\theta}_e) := \text{optC}(U_{e}^{1158}, B, AIC) \) by the R-function BiCopSelect(), we obtain the following family and parameters:
\[ \hat{b}_{e} := \hat{b}_e = N, \quad \hat{\theta}_{e} := \hat{\theta}_e = \begin{pmatrix} \hat{\theta}_1^1 \\ \hat{\theta}_2^2 \end{pmatrix} = \begin{pmatrix} 0.17 \\ 0 \end{pmatrix}. \]

(c) Using the result from b) we can estimate the pseudo observations \( \hat{h}_{P_{e,1}|P_{e,2};D_{e,\pi_1},D_{e,\pi_2}} = \hat{h}_{BMW|VOW,DAI,DBK} := h_1(U_{e}^{1158}, C_{\hat{b}_e}(\cdot,\cdot;\hat{\theta}_e)) \in [0,1]^{1158} \) and
\( \hat{h}_{P_{e,2}|P_{e,1};D_{e,\pi_1},D_{e,\pi_2}} = \hat{h}_{VOW|BMW,DAI,DBK} := h_2(U_{e}^{1158}, C_{\hat{b}_e}(\cdot,\cdot;\hat{\theta}_e)) \in [0,1]^{1158}. \)

(d) Those pseudo observations are added to the set \( \hat{\mathcal{P}}_3 \), i.e.
\[ \hat{\mathcal{P}}_3 := \{ \hat{h}_{BMW|VOW,DAI,DBK}, \hat{h}_{VOW|BMW,DAI,DBK} \}. \]

(v) The third R-vine copula is given by \( \hat{\mathcal{R}}^3 = (\mathcal{V}^3, \hat{\mathcal{B}}^3, \hat{\Theta}^3) \), with
\[ \hat{\mathcal{B}}^3 := \hat{\mathcal{B}}^2 \cup \{ \hat{b}_e : e \in \bigcup_{k=1}^{3} e_k^3 \} = \{ t, t, F, t, F, N \} \]
and
\[ \hat{\Theta}^3 := \hat{\Theta}^2 \cup \{ \hat{\theta}_e : e \in \bigcup_{k=1}^{3} e_k^3 \} = \{ (0.68, 8.77), (0.64, 5.64), (1.9, 0), (0.53, 11.27), (0.93, 0), (0.17, 0) \}. \]

(vi) The goodness of fit measures corresponding to \( \hat{\mathcal{R}}^3 \) are given by
\[- \text{c\ell}^3 := \text{c\ell}(\hat{\mathcal{R}}^3, \text{bmw}, (\text{dai, dbk, vow3})) = 429.3538, \]
\[- \text{cAIC}^3 := \text{cAIC}(\hat{\mathcal{R}}^3, \text{bmw}, (\text{dai, dbk, vow3})) = 840.7075, \]
\[- \text{cBIC}^3 := \text{cBIC}(\hat{\mathcal{R}}^3, \text{bmw}, (\text{dai, dbk, vow3})) = -795.2175. \]
Observe that
\[ |\hat{\Theta}^3| - |\hat{\Theta}^2| = |\{ \hat{\theta}_e : e \in \bigcup_{k=1}^{3} e_k^3 \}| = 2 + 1 + 1 = 4. \]
and
\[ c\ell\hat{c} - c\ell^2 = 429.3538 - 412.1252 = 17.2286. \]

Then, the conditional likelihood test between $\hat{R}^3$ and $\hat{R}^2$ is rejected at a level $\alpha_{LRT} = 0.05$, since
\[ c\ell^3 - c\ell^2 = 17.2286 > X[^2_{1-\alpha_{LRT}, |\Theta|}] = X[^2_{0.95,4}] = 9.4877. \]

The corresponding p-value is given by $p^3 = 0.001744941$.
Thus, $U_{j_3} = U_3 = VOW3$ is included in $\hat{R}^3$ for $fC \in \{0, 1\}$.

- **s = 4:**

  In step $s = 4$, we add $U_4 = U_{j_4} = SIE$ to the regular vine copula.

  - $k = 1$:
    
    (i) The set of partner candidates for $SIE$ in tree $T_1^4$ is given by
    \[ PT_1^4 = N_3^1 \setminus \{BMW\} = \{DAI, DBK, VOW3\}. \]
    Since $|PT_1^4| > 1$, we need to make use of the weight criterion $wc = \rho^{max}$.
    According to Table 6.1 which contains the maximal empirical semi-correlations $\hat{\rho}^{1158, max}_{SIE,U}$ for $U \in \{DAI, DBK, VOW3\}$, it holds that
    \[ pt_1^{SIE} = \text{maxW}_1(SIE\{DAI, DBK, VOW3\}, \rho^{max}) = DBK, \]
    since
    \[ \hat{\rho}^{1158, max}_{SIE,DBK} = 0.3369 > \hat{\rho}^{1158, max}_{SIE,DAI} = 0.2716 > \hat{\rho}^{1158, max}_{SIE,VOW3} = 0.2069. \]

    (ii) \[ N_1^4 := N_3^1 \cup \{SIE\} = \{BMW, DAI, DBK, VOW3, SIE\}, \]
    \[ E_1^4 := E_1^3 \cup \{(SIE, DBK)\} \]
    \[ = \{(BMW, DAI), (DBK, DAI), (VOW3, DAI), (SIE, DBK)\}, \]
    \[ T_1^4 := (N_1^4, E_1^4). \]
    (iii) Next, pair copula family, pair copula parameters and pseudo observations corresponding to the “new” edge
    \[ e = (P_{e,1}, P_{e,2}) := e_1^4 = (SIE, DBK) \]
    are estimated.
    (a) The copula data matrix which is used for the estimation is given by
    \[ U_e^{1158} := (P_{e,1}, P_{e,2}) = (\text{sie, dbk}) \in [0, 1]^{1158 \times 2}. \]
(b) Estimating $C_{be}(\cdot, \cdot; \hat{\theta}_e) := \text{opt}(U_{e}^{1158}, B, \text{AIC})$ by the R-function $\text{BiCopSel}()$, we obtain the following family and parameters:

$$
\hat{b}_{e_1} := \hat{b}_e = t, \quad \hat{\theta}_{e_1} := \hat{\theta}_e = \left( \hat{\theta}_1^e, \hat{\theta}_2^e \right) = \left( 0.65, 5.36 \right).
$$

(c) Using the result from b) we can estimate the pseudo observations

$$
\hat{h}_{P_e|P_e,1} = \hat{h}_{SIE|DBK} := h_1(U_{e}^{1158}, C_{be}(\cdot, \cdot; \hat{\theta}_e)) \in [0,1]^{1158}
$$

and

$$
\hat{h}_{P_e|P_e,2} = \hat{h}_{DBK|SIE} := h_2(U_{e}^{1158}, C_{be}(\cdot, \cdot; \hat{\theta}_e)) \in [0,1]^{1158}.
$$

(d) Those pseudo observations are added to the set $\hat{\mathcal{P}}_{1}^{*}$, i.e.

$$
\hat{\mathcal{P}}_{1}^{*} := \hat{\mathcal{P}}_{1}^{\delta} \cup \{ \hat{h}_{SIE|DBK}, \hat{h}_{DBK|SIE} \}.
$$

- $k = 2$:

(i.1) The set of partner candidates for $SIE$ in tree $T_2^{4}$ is given by

$$
\text{PT}^{4}_{2} = \{ n \in N_{3}^{1} \setminus \{ \text{BMW}, \text{pt}_{1}^{SIE} \} : \exists e \in E_{1}^{3}, \text{ s.t. } n, \text{pt}_{1}^{SIE} \in A_{e} \} = \{ \text{DAI} \},
$$

since

$$
N_{3}^{1} \setminus \{ \text{BMW}, \text{pt}_{1}^{SIE} \} = \{ \text{DAI}, \text{VOW3} \},
$$

$$(\text{DAI}, \text{DBK}) = (\text{DAI}, \text{pt}_{1}^{SIE}) \in E_{1}^{3},
$$

and

$$(\text{VOW3}, \text{DBK}) = (\text{VOW3}, \text{pt}_{1}^{SIE}) \notin E_{1}^{3}.
$$

Thus, the partner of $SIE$ in $T_2^{4}$ is given by

$$
\text{pt}_{2}^{SIE} := \text{maxW}_{2}(SIE, \{ \text{DAI} \}, \rho^{\text{max}}) = \text{DAI}.
$$

(ii.1) $N_{2}^{4} = E_{1}^{4}, E_{2}^{4} = E_{2}^{3} \cup \{ (SIE, \text{DAI}|DBK) \}, T_{2}^{4} = (N_{2}^{4}, E_{2}^{4}).$

(iii) Set

$$
e = (P_{e,1}, P_{e,2}|D_{e,1}) := e_{2}^{4} = (SIE, \text{DAI}|\text{DBK})
$$

and remember that

$$
\hat{\mathcal{P}}_{k-1}^{*} = \hat{\mathcal{P}}_{1}^{*} = \{ \hat{h}_{\text{BMW}|\text{DAI}}, \hat{h}_{\text{DAI}|\text{BMW}}, \hat{h}_{\text{DBK}|\text{DAI}}, \hat{h}_{\text{DAI}|\text{DBK}}, \hat{h}_{\text{VOW3}|\text{DAI}} \}
$$

$$
\cup \{ \hat{h}_{\text{DAI}|\text{VOW3}}, \hat{h}_{SIE|DBK}, \hat{h}_{DBK|SIE} \}.
$$

We observe that

$$
\hat{h}_{P_{e,1}|D_{e,1}} = \hat{h}_{SIE|DBK}, \hat{h}_{P_{e,2}|D_{e,1}} = \hat{h}_{\text{DAI}|DBK} \in \hat{\mathcal{P}}_{k-1}^{*} = \hat{\mathcal{P}}_{1}^{*}.
$$
Accordingly, we have to select
\[ \pi = \pi_1 = \hat{\pi} = \tilde{\pi}_1 = 1 \in \{1\}, \text{ s.t.} \]
\[ \hat{h}_{P_e,1|D_{e,1}}, \hat{h}_{P_e,2|D_{e,2}} \in \mathcal{P}_1^3. \]
Next, pair copula family, pair copula parameters and pseudo observations

\[ e = (P_{e,1}, P_{e,2}|D_{e,1}) := e_2^4 = (SIE, DAI|DBK) \]
are estimated.

(a) The copula data matrix, which is used for the estimation, is given by
\[ U_{e}^{1158} := (\hat{h}_{P_e,1|D_{e,1}}, \hat{h}_{P_e,2|D_{e,2}}) = (\hat{h}_{SIE|DBK}, \hat{h}_{DAI|DBK}) \in [0, 1]^{1158 \times 2}. \]

(b) Estimating \( C^b_{\hat{\theta}_e} := \text{optC}(U_{e}^{1158}, B, AIC) \) by the R-function BiCopSelect(), we obtain the following family and parameters:
\[ \hat{b}_{e^1} := \hat{b}_e = t, \hat{\theta}_{e^1} := \hat{\theta}_e = \begin{pmatrix} \hat{\theta}_e^1 \\ \hat{\theta}_e^2 \end{pmatrix} = \begin{pmatrix} 0.34 \\ 8.16 \end{pmatrix}. \]

(c) Using the result from b) we can estimate the pseudo observations
\[ \hat{h}_{P_e,1|D_{e,1}} = \hat{h}_{SIE|DAI;DBK} := h_1(U_{e}^{1158}, C^b_{\hat{\theta}_e}) \in [0, 1]^{1158} \]
and
\[ \hat{h}_{P_e,2|D_{e,2}} = \hat{h}_{DAI|SIE;DBK} := h_2(U_{e}^{1158}, C^b_{\hat{\theta}_e}) \in [0, 1]^{1158}. \]

(d) Those pseudo observations are added to the set \( \hat{\mathcal{P}}_2^4 \), i.e.
\[ \hat{\mathcal{P}}_2^4 := \hat{\mathcal{P}}_2^3 \cup \{ \hat{h}_{SIE|DAI;DBK}, \hat{h}_{DAI|SIE;DBK} \}. \]

- \( k = 3 \):

(i.1) The set of partner candidates for SIE in tree \( T_3^4 \) is given by
\[ PT_3^4 = \{ n \in N^3 \setminus \{BMW, pt_1^{SIE}, pt_2^{SIE} \} : \exists e \in E_2, \text{ s.t. } n, pt_1^{SIE}, pt_2^{SIE} \in A_e \} = \{ VOW3 \}, \]
since
\[ N^3 \setminus \{BMW, pt_1^{SIE}, pt_2^{SIE} \} = \{ VOW3 \} \]
and
\[ (VOW3, DBK|DAI) = (VOW3, pt_1^{SIE}|pt_2^{SIE}) \in E_2^3. \]
Thus, the partner of SIE in \( T_3^4 \) is given by
\[ pt_3^{SIE} := \text{maxW}_2(SIE, \{VOW3\}, \rho^{\text{max}}) = VOW3. \]
(ii.1) \[ N_3^4 = E_1^4, ~ E_3^4 = E_3^3 \cup \{(SIE, VOW3|DAI, DBK)\}, ~ T_3^4 = (N_3^4, E_3^4) \]

(iii) Set

\[ e = (P_{e,1}, P_{e,2}|D_{e,1}, D_{e,2}) := e_3^4 = (SIE, VOW3|DAI, DBK) \]

and remember that

\[ \widehat{P}_{k-1}^s = \widehat{P}_2^4 \]

\[ = \{\hat{P}_{BMW|DBK;DAI}, \hat{P}_{DBK|BMW;DAI}, \hat{P}_{VOW3|DBK;DAI}, \hat{P}_{DBK|VOW3;DAI}\} \]

\[ \cup \{\hat{P}_{SIE|DAI;DBK}, \hat{P}_{DAI|SIE;DBK}\}. \]

We observe that

\[ h_{P_{e,1}|D_{e,1};D_{e,2}} = h_{SIE|DAI;DBK}, h_{P_{e,2}|D_{e,2};D_{e,1}} = h_{VOW3|DBK;DAI} \in \widehat{P}_{k-1}^s = \widehat{P}_2^4. \]

Accordingly we have to select

\[ \pi = (\pi_1, \pi_2)^T = (1, 2)^T \in \{1, 2\}^2 \]

and

\[ \hat{\pi} = (\hat{\pi}_1, \hat{\pi}_2)^T = (2, 1)^T, \text{ s.t.} \]

\[ \hat{h}_{P_{e,1}|D_{e,1};D_{e,2}}, \hat{h}_{P_{e,2}|D_{e,2};D_{e,1}} \in \widehat{P}_2^4. \]

Next, pair copula family, pair copula parameters and pseudo observations corresponding to the “new” edge

\[ e = (P_{e,1}, P_{e,2}|D_{e,1}, D_{e,2}) := e_3^4 = (SIE, VOW3|DAI, DBK) \]

are estimated.

(a) The copula data matrix which is used for the estimation is given by

\[ \mathcal{U}_{c}^{1158} := (\hat{h}_{P_{e,1}|D_{e,1};D_{e,2}}, \hat{h}_{P_{e,2}|D_{e,1};D_{e,2}}) \]

\[ = (\hat{h}_{SIE|DAI;DBK}, \hat{h}_{VOW3|DBK;DAI}) \in [0, 1]^{1158 \times 2}. \]

(b) Estimating \( \hat{C}_c^b(\cdot, \cdot; \hat{\theta}_c) := \text{optC}(\mathcal{U}_c^{1158}, B, AIC) \) by the R-function \( \text{BiCopS}-\text{elect()} \), we obtain the following family and parameters:

\[ \hat{b}_{c_3} := \hat{b}_c = t, ~ \hat{\theta}_{c_3} := \hat{\theta}_c = \left( \begin{array}{c} \hat{\theta}_1^c \\ \hat{\theta}_2^c \end{array} \right) = \left( \begin{array}{c} 0.06 \\ 13.66 \end{array} \right). \]

(c) Using the result from b), we can estimate the pseudo observations

\[ \hat{h}_{P_{e,1}|P_{e,2};D_{e,1},D_{e,2}} = \hat{h}_{SIE|VOW3;DAI, DBK} := h_1(\mathcal{U}_c^{1158}, \hat{C}_c^b(\cdot, \cdot; \hat{\theta}_c)) \in [0, 1]^{1158} \]

and

\[ \hat{h}_{P_{e,2}|P_{e,1};D_{e,1},D_{e,2}} = \hat{h}_{VOW3|BMW;DAI, DBK} := h_2(\mathcal{U}_c^{1158}, \hat{C}_c^b(\cdot, \cdot; \hat{\theta}_c)) \in [0, 1]^{1158}. \]
(d) Those pseudo observations are added to the set \( \hat{\mathcal{P}}_3^4 \), i.e.
\[
\hat{\mathcal{P}}_3^4 := \hat{\mathcal{P}}_3^3 \cup \{ h_{\text{SIE}|\text{VOW3},\text{DAI},\text{DBK}}, h_{\text{VOW3}|\text{SIE},\text{DAI},\text{DBK}} \}.
\]

- \( k = 4 \):

(i.2) We are in the case \( k = s = 4 \) and set
\[
pt_{4}^{\text{BMW}} := U_4 = U_{j_4}^p = \text{SIE}.
\]

(ii.2) The “new” edge is given by
\[
e_4^4 = (\text{BMW}, \text{SIE}| \text{DAI}, \text{DBK}, \text{VOW3}).
\]

It follows that
\[
N_4^4 := E_3^4, \quad E_4^4 = \{(\text{BMW}, \text{SIE}| \text{DAI}, \text{DBK}, \text{VOW3})\}, \quad T_4^4 = (N_4^4, E_4^4).
\]

Then, \( \mathcal{V}_4 = (T_1^4, T_2^4, T_3^4, T_4^4) \) looks as shown in Figure 6.4.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6.4.png}
\caption{\( \mathcal{V}_4 = (T_1^4, T_2^4, T_3^4, T_4^4) \) in example daxreturns.}
\end{figure}
(iii) Set
\[ e = (P_{e,1}, P_{e,2}|D_{e,1}, D_{e,2}, D_{e,3}) := e_4^i = (BMW, SIE|DAI, DBK, VOW3) \]
and remember that
\[ \hat{P}_{k-1}^s = \hat{P}_3^4 = \{ \hat{h}_{BMW|VOW3,DAI, DBK}, \hat{h}_{VOW3|BMW,DAI, DBK} \} \]
\[ \cup \{ \hat{h}_{SIE|VOW3,DAI, DBK}, \hat{h}_{VOW3|SIE, DAI, DBK} \}. \]

We observe that
\[ \hat{h}_{P_e,1}|D_{e,1}; D_{e,2}, \hat{h}_{P_e,2}|D_{e,1}, D_{e,2} = \hat{h}_{BMW|VOW3; DAI, DBK} \in \hat{P}_3^4 \]
and
\[ \hat{h}_{P_e,2}|D_{e,1}; D_{e,2}, \hat{h}_{P_e,2}|D_{e,2}, D_{e,3} = \hat{h}_{SIE|VOW3, DAI, DBK} \in \hat{P}_3^4. \]

Accordingly, we have to select
\[ \pi = (\pi_1, \pi_2, \pi_3)^T = \hat{\pi} = (\hat{\pi}_1, \hat{\pi}_2, \hat{\pi}_3)^T = (3, 1, 2)^T \in \{1, 2, 3\}^3, \text{ s.t.} \]
\[ \hat{h}_{P_e,1}|D_{e,1}; D_{e,2}, \hat{h}_{P_e,2}|D_{e,2}, D_{e,3} \in \hat{P}_3^4. \]

Next, pair copula family, pair copula parameters and pseudo observations corresponding to the “new” edge
\[ e = (P_{e,1}, P_{e,2}|D_{e,1}, D_{e,2}, D_{e,3}) := e_3^i = (BMW, SIE|DAI, DBK, VOW3) \]
are estimated.

(a) The copula data matrix, which is used for the estimation, is given by
\[ U_e^{1158} := (\hat{h}_{P_e,1}|D_{e,1}; D_{e,2}, \hat{h}_{P_e,2}|D_{e,2}, D_{e,3} \in \{0, 1\}^{1158 \times 2}. \]

(b) Estimating \( C^{\hat{b}}(\cdot, \cdot; \hat{\theta}_e) := \text{optC}(U_e^{1158}, B, AIC) \) by the R-function \( BiCopSselect() \), we obtain the following family and parameters:
\[ \hat{b}_e^{\hat{\theta}_e} := \hat{b}_e = F, \quad \hat{\theta}_e = \left( \begin{array}{c} \hat{\theta}_1^1 \\ \hat{\theta}_2^t \\ \end{array} \right) = \left( \begin{array}{c} 0.81 \\ 0 \\ \end{array} \right). \]

(c) Using the result from (b), we can estimate the pseudo observations
\[ \hat{h}_{P_e,1}|P_{e,2}|D_{e,1}, D_{e,2}, D_{e,3} = \hat{h}_{BMW|SIE; DAI, DBK, VOW3} := \hat{h}_1(U_e^{1158}, C^{\hat{b}}(\cdot, \cdot; \hat{\theta}_e)) \in [0, 1]^{1158} \]
and
\[ \hat{h}_{P_e,2}|P_{e,1}|D_{e,1}, D_{e,2}, D_{e,3} = \hat{h}_{SIE|BMW; DAI, DBK, VOW3} := \hat{h}_2(U_e^{1158}, C^{\hat{b}}(\cdot, \cdot; \hat{\theta}_e)) \in [0, 1]^{1158}. \]
(d) Those pseudo observations are added to the set \( \hat{P}_4 \), i.e.

\[
\hat{P}_4 := \{ \hat{h}_{BMW|SIE;DAI, DBK, VOW3}, \hat{h}_{SIE|BMW;DAI, DBK, VOW3} \}.
\]

(v) The fourth R-vine copula is given by \( \hat{R}^4 = (\hat{\mathcal{V}}^4, \hat{\mathcal{B}}^4, \hat{\Theta}^4) \) with

\[
\hat{\mathcal{B}}^4 := \hat{\mathcal{B}}^3 \cup \{ \hat{b}_e : e \in \bigcup_{k=1}^{4} e_k^4 \} = \{ t, t, F, F, N, t, t, t, F \}
\]

and

\[
\hat{\Theta}^4 := \hat{\Theta}^3 \cup \{ \hat{\theta}_e : e \in \bigcup_{k=1}^{4} e_k^4 \}
\]

\[
= \{ (0.68, 8.77), (0.64, 5.64), (1.9, 0), (0.53, 11.27), (0.93, 0), (0.17, 0) \}
\]

\[
\bigcup \{ (0.65, 5.36), (0.34, 8.16), (0.06, 13.66), (0.81, 0) \}.
\]

(vi) The goodness of fit measures corresponding to \( \hat{R}^4 \) are given by

- \( c\ell \ell^4 := c\ell(\hat{R}^4, bmw, (dai, dbk, vow3, sie)) = 439.3335 \),
- \( cAIC^4 := cAIC(\hat{R}^4, (bmw, dai, dbk, vow3, sie)) = 846.6669 \),
- \( cBIC^4 := cBIC(\hat{R}^4, bmw, dai, dbk, vow3, sie) = -765.7957 \).

Observe that

\[
|\hat{\Theta}^4| - |\hat{\Theta}^3| = |\{ \hat{\theta}_e : e \in \bigcup_{k=1}^{4} e_k^4 \}| = 2 + 2 + 2 + 1 = 7
\]

and

\[
c\ell\ell^4 - c\ell\ell^3 = 439.3335 - 429.3538 = 9.9797.
\]

Then, the conditional likelihood test between \( \hat{R}^4 \) and \( \hat{R}^3 \) is not rejected at a level \( \alpha_{LRT} = 0.05 \), since

\[
c\ell\ell^4 - c\ell\ell^3 = 9.9797 < \chi^2_{1-\alpha_{LRT}, |\hat{\Theta}^4| - |\hat{\Theta}^3|} = \chi^2_{0.95, 7} = 14.06714.
\]

The corresponding p-value is given by \( p^4 = 0.1897256 \).

Thus, \( U_{j4} = U_4 = SIE \) is included in \( \hat{R}^4 \) for \( fC = 1 \) and is not included for \( fC = 0 \). However, the algorithm stops in both cases, since \( s = 4 = d \) is reached, and returns \( \hat{R}^4 \) for \( fC = 1 \) and \( \hat{R}^3 \) for \( fC = 0 \).
Table 6.2 summarizes the conditional log likelihoods, AICs and BICs computed in iterations \( s = 1, \ldots, 4 \). The last column contains references to the figure numbers, where the selected R-vine tree sequences are illustrated.

<table>
<thead>
<tr>
<th>( s )</th>
<th>( c\ell^s )</th>
<th>cAIC( s )</th>
<th>cBIC( s )</th>
<th>( V^s )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>357.3561</td>
<td>-710.7122</td>
<td>-700.6033</td>
<td>6.1</td>
</tr>
<tr>
<td>2</td>
<td>412.1252</td>
<td>-814.2504</td>
<td>-788.9781</td>
<td>6.2</td>
</tr>
<tr>
<td>3</td>
<td>429.3538</td>
<td>-840.7075</td>
<td>-795.2175</td>
<td>6.3</td>
</tr>
<tr>
<td>4</td>
<td>439.3335</td>
<td>-846.6669</td>
<td>-765.7957</td>
<td>6.4</td>
</tr>
</tbody>
</table>

Table 6.2: Summary of \( c\ell^s \), cAIC\( s \), cBIC\( s \), \( V^s \) for all \( s = 1,2,3,4 \).

Using the R-vine copula \( \hat{\mathcal{R}}^3 \) returned by the algorithm, one can estimate the \( \alpha \)-quantile of BMW given some evaluation data set, i.e.

\[
DAI = dav_{eval}^{eval}, \ DBK = dbk_{eval}^{eval}, \ VOW3 = vow3_{eval}^{eval},
\]

in the following way:

\[
\hat{Q}_{BMW|DAI, DBK, VOW3}(\alpha|dav_{eval}^{eval}, dbk_{eval}^{eval}, vow3_{eval}^{eval}) = \\
h_{BMW|DAI}^{-1}(h_{BMW|DBK, DAI}^{-1}(h_{BMW|VOW3, DAI, DBK}^{-1}(\alpha) \\
h_{VOW3|DBK, DAI}(h_{VOW3|VOW3, DAI}(\alpha|dav_{eval}^{eval})) \\
h_{DBK|DAI}(dbk_{eval}^{eval}|dav_{eval}^{eval}))h_{DBK|DAI}(dbk_{eval}^{eval}|dav_{eval}^{eval})),|dav_{eval}^{eval}|
\]
Part III

Simulation Study and data application
Chapter 7

Simulation study

In the following simulation study, we compare the R-vine quantile regression algorithm to other benchmark methods and assess the influence of the algorithm’s parameters on the accuracy of the predicted quantiles. We start with describing the setup and present the results afterwards.

7.1 Setup

All considered methods are compared in 100 replications
\[ r = 1, \ldots, 100 \]
and 8 scenarios
\[ S \in \{1, \ldots, 8\} \].

In order to keep the notation as simple as possible, we ignore the belonging to a specific scenario \( S \) while describing the setup.

In each replication \( r \in \{1, \ldots, 100\} \), we simulate an initial R-vine copula for quantile regression
\[
R_{r_0} = (\mathcal{V}_{r_0}, \mathcal{B}_{r_0}(\mathcal{V}_{r_0}), \Theta_{r_0}(\mathcal{B}_{r_0}(\mathcal{V}_{r_0}))), \quad d_0 \in \{8, 20\}.
\]

The copula is used to simulate an initial training and evaluation data set, i.e. we simulate

\[ U_{r_0, \text{train}} := (v_{r, \text{train}}, u_{1, \text{train}}^{r, \text{train}}, \ldots, u_{d_0, \text{train}}^{r, \text{train}}) \in [0,1]^{n_{\text{train}} \times d_0} \]

with \( v_{r, \text{train}}^{r, \text{train}} = (v_{1, \text{train}}^{r, \text{train}}, \ldots, v_{n_{\text{train}}, \text{train}}^{r, \text{train}})^T \), \( u_{j, \text{train}}^{r, \text{train}} = (u_{i_1, \text{train}}^{r, \text{train}}, \ldots, u_{i_{n_{\text{train}}, \text{train}}, \text{train}}^{r, \text{train}})^T, \quad j = 1, \ldots, d_0, \) and

\[ U_{r_0, \text{eval}} := (v_{r, \text{eval}}, u_{1, \text{eval}}^{r, \text{eval}}, \ldots, u_{d_0, \text{eval}}^{r, \text{eval}}) \in [0,1]^{n_{\text{eval}} \times d_0} \]

with \( v_{r, \text{eval}}^{r, \text{eval}} = (v_{1, \text{eval}}^{r, \text{eval}}, \ldots, v_{n_{\text{eval}}, \text{eval}}^{r, \text{eval}})^T \), \( u_{j, \text{eval}}^{r, \text{eval}} = (u_{i_1, \text{eval}}^{r, \text{eval}}, \ldots, u_{i_{n_{\text{eval}}, \text{eval}}, \text{eval}}^{r, \text{eval}})^T, \quad j = 1, \ldots, d_0. \)

The sample sizes are chosen to be

\[ n_{\text{train}} \in \{1000, 2000\} \text{ and } n_{\text{eval}} \in \{500, 1000\}. \]
7.1.1 Simulation of the initial training and evaluation data set

In each replication \( r \in \{1, \ldots, 100\} \), we simulate the initial data set in two separate steps:

1. Simulation of the initial R-vine copula \( \mathcal{R}_{d_0}^d \).

2. Simulation of the initial copula data \( U_{\text{train}} \), \( U_{\text{eval}} \) \( \sim \mathcal{R}_{d_0}^d \).

While we can use the R-function \( \text{RVineSim()} \) from the \text{VineCopula}-package, in order to conduct step (2), the simulation of \( \mathcal{R}_{d_0}^d \) is more complex.

We start with sampling an R-vine tree structure on \((d_0 + 1)\) elements \( U_{1}, \ldots, U_{d_0+1} \) using the R-function \( \text{RVineMatrixSample()} \) from the \text{VineCopula}-package, which is based on an algorithm developed by [Joe et al. 2010]. It samples some R-vine tree structure \( \mathcal{V}_r^{d_0} = (T_1^{d_0, r}, \ldots, T_{d_0}^{d_0, r}) \), \( T_k^{d_0, r} = (N_k^{d_0, r}, E_k^{d_0, r}) \), \( k = 1, \ldots, d_0 \), out of the set of valid R-vine tree sequences with \( T_{d_0}^{d_0, r} \) having the structure shown in Figure 7.1, where

\[
D_{d_0} := \{U_1, \ldots, U_{d_0-1}\} \setminus \{p_{d_0-1} \}.
\]

![Figure 7.1: Highest order tree of an R-vine tree sequence sampled by RVineMatrixSample().](image)

Since both \( U_{d_0+1} \) and \( U_{d_0} \) are R-vine leafs in \( \mathcal{V}_r^{d_0} \), we can choose one of them (each with probability 0.5) to be the response variable \( V \), i.e.

\[
P( V = U_{d_0+1} ) = P( V = U_{d_0} ) = 0.5.
\]

If we define \( V = U_{d_0} \), we set \( U_{d_0} := U_{d_0+1} \) in order to maintain a consistent notation.

Next, we need to choose the sets of copula families \( \mathcal{B}_r^{d_0}(\mathcal{V}_r^{d_0}) \) and parameter vectors \( \Theta_r^{d_0}(\mathcal{B}_r^{d_0}(\mathcal{V}_r^{d_0})) \). Therefore, we sample a copula family \( b_e \in \mathcal{B}_0 \) and a parameter vector \( \theta_e(b_e) \) of the parametric pair copula \( C_{b_e}(\cdot, \cdot; \theta_e) \) for all \( e \in \bigcup_{k=1}^{d_0} E_k^{d_0, r} \). How this sampling works in detail is described in the following.
For every $\varepsilon \in \bigcup_{k=1}^{d_0} E_k^{d_0,r}$, $b_\varepsilon$ is sampled uniformly out of the set of possible initial copula families $B_0$, i.e.

$$\mathbb{P}(b_\varepsilon = b) = \frac{1}{|B_0|}, \text{ for all } b \in B_0$$

with

$$B_0 := \{ \text{"Gaussian", "Student t", "Gumbel", "Gumbel 180"} \}$$

$$\cup \{ \text{"Gumbel 90", "Gumbel 270", "Frank"} \}.$$ 

The first entry of the parameter vector $(\theta^1_\varepsilon)$ is sampled using the relationship to Kendall’s $\tau$ (Section 1.4). Therefore, we first sample $\tau_\varepsilon \in [-1, 1]$, i.e. the Kendall’s $\tau$ corresponding to the pair copula $C^{b_\varepsilon}(\cdot, \cdot; \theta_\varepsilon)$.

For $\varepsilon \in E_k^{r,d_0}$, we sample $\tau^\text{abs}_k \in \{0.2^k, 0.8^k\}$ with

$$\mathbb{P}(\tau^\text{abs}_k = 0.2^k) = \mathbb{P}(\tau^\text{abs}_k = 0.8^k) = 0.5, \ k = 1, \ldots, d_0,$$

and $\tau^\text{sign}_\varepsilon \in \{-1, 1\}$, with

$$\mathbb{P}(\tau^\text{sign}_\varepsilon = 1) = 1 - \mathbb{P}(\tau^\text{sign}_\varepsilon = -1) = \begin{cases} 0.5, & \text{if } b_\varepsilon \in \{ \text{"Gaussian", "Student t"} \}, \\ 1, & \text{if } b_\varepsilon \in \{ \text{"Gumbel", "Gumbel 180"} \}, \\ 0, & \text{if } b_\varepsilon \in \{ \text{"Gumbel 90", "Gumbel 270"} \}. \end{cases}$$

Having samples of $\tau^\text{abs}_k$ and $\tau^\text{sign}_\varepsilon$, we can define

$$\tau_\varepsilon := \tau^\text{sign}_\varepsilon \times \tau^\text{abs}_k \text{ for every } \varepsilon \in E_k^{r,d_0} \text{ and } k = 1, \ldots, d_0.$$ 

Using the R-function $\text{BiCopTau2par()}$ from the $\text{VineCopula}$-package, we finally find the first pair copula parameter $\theta^1_\varepsilon$ as a function of the sampled $\tau_\varepsilon$, i.e. $\theta^1_\varepsilon := \theta^1_\varepsilon(\tau_\varepsilon)$.

Except the case $b_\varepsilon = \text{"Student t"}$, $\theta^2_\varepsilon$ is set to zero, i.e. we define

$$\theta^2_\varepsilon = \begin{cases} 3, & \text{if } b_\varepsilon = \text{"Student t"}, \\ 0, & \text{otherwise}. \end{cases}$$

Note that we always choose the degrees of freedom of the Student t - copula to be equal to three.

Additionally, we consider the scenario, that a subset of the covariates is independent of the response $V$. Therefore, all pair copulas corresponding to edges $\varepsilon$ with $|\tau_\varepsilon| < 0.01$ and $V \in A_\varepsilon$ are set to the independence copula, i.e.

$$b_\varepsilon = I, \text{ for all } \varepsilon \in \bigcup_{k=1}^{d_0} E_k^{d_0,r} : |\tau_\varepsilon| < 0.01, \ V \in A_\varepsilon.$$

100
Whether this scenario is considered or not, is indicated by a dummy variable

\[ \text{indep}_0 := \begin{cases} 1, & b_e = 1, \text{ for all } e \in \{ \bigcup_{k=1}^{d_0} E_{k}^{d_0,r} : |\tau_e| < 0.01, \ V \in A_e \}, \\ 0, & \text{otherwise.} \end{cases} \]

Those covariates, which are part of the conditioned set \( P_e \) of an edge \( e \), for which the pair copula is set to the independence copula, are elements of the set \( \text{Cov}_{\text{indep}}^r \), i.e.

\[ \text{Cov}_{\text{indep}}^r := \{ U \in \{ U_1, \ldots, U_{d_0} \} : b_e = 1 \text{ for } e \in \bigcup_{k=1}^{d_0} E_{k}^{d_0,r} \text{ and } P_e = \{ V, U \} \}. \]

In other words, \( \text{Cov}_{\text{indep}}^r \) denotes the set of “independent” covariates. Note that \( \text{Cov}_{\text{indep}}^r = \emptyset \), if \( \text{indep}_0 = 0 \).

Before an initial copula data set is simulated in R, a seed is set using the R-function \( \text{set.seed()} \). An overview of the selected seeds can be found in Appendix A.1.1.

### 7.1.2 Scenarios

As already mentioned above, we consider 8 different scenarios \( S \in \{1, \ldots, 8\} \). Each scenario corresponds to a certain parameter set, which is defined by the number of (initial) covariates \( d_0^S \), the sample size of the initial training data sets \( n_{\text{train}}^S \), the sample size of the initial evaluation data sets \( n_{\text{eval}}^S \) and the dummy variable \( \text{indep}_0^S \in \{0, 1\} \).

- \( S = 1 \): \( d_0^1 = 8, n_{\text{train}}^1 = 1000, n_{\text{eval}}^1 = 500, \text{indep}_0^1 = 0 \),
- \( S = 2 \): \( d_0^2 = 8, n_{\text{train}}^2 = 1000, n_{\text{eval}}^2 = 500, \text{indep}_0^2 = 1 \),
- \( S = 3 \): \( d_0^3 = 8, n_{\text{train}}^3 = 2000, n_{\text{eval}}^3 = 1000, \text{indep}_0^3 = 0 \),
- \( S = 4 \): \( d_0^4 = 8, n_{\text{train}}^4 = 2000, n_{\text{eval}}^4 = 1000, \text{indep}_0^4 = 1 \),
- \( S = 5 \): \( d_0^5 = 20, n_{\text{train}}^5 = 1000, n_{\text{eval}}^5 = 500, \text{indep}_0^5 = 0 \),
- \( S = 6 \): \( d_0^6 = 20, n_{\text{train}}^6 = 1000, n_{\text{eval}}^6 = 500, \text{indep}_0^6 = 1 \),
- \( S = 7 \): \( d_0^7 = 20, n_{\text{train}}^7 = 2000, n_{\text{eval}}^7 = 1000, \text{indep}_0^7 = 0 \),
- \( S = 8 \): \( d_0^8 = 20, n_{\text{train}}^8 = 2000, n_{\text{eval}}^8 = 1000, \text{indep}_0^8 = 1 \).

If necessary, we will include the considered scenario \( S \) in our notation.

### 7.1.3 Competitor methods

**Linear quantile regression \([\text{LQR}, \text{LQRWSA}]\)**

Since it was the first method, which was developed, linear quantile regression (Section 3.2) is a commonly used benchmark for simulation studies on quantile regression. Remember
that given an evaluation data set \( U_{r,\text{eval}} \), the corresponding predicted conditional quantile function is given as:

\[
\hat{Q}_{V|U_1,\ldots,U_{d}} \left( \alpha | u_{1,\text{eval}}^{r,\text{eval}}, \ldots, u_{d,\text{eval}}^{r,\text{eval}} \right) := \hat{\beta}_0^r(\alpha) + \sum_{j=1}^{d} \hat{\beta}_j^r(\alpha) u_{ij,\text{eval}}, \quad i = 1, \ldots, n_{\text{eval}}.
\]

We distinguish between two different cases dependent on which of the following assumptions is made:

(i) All covariates are significant for every \( \alpha \), i.e. \( \hat{\beta}_j^r(\alpha) \neq 0 \) for all \( j \in \{0, 1, \ldots, d\} \) and \( r = 1, \ldots, 100 \).

(ii) For every \( \alpha \), only a subset of covariates is significant, i.e. there might exist \( j \in \{1, \ldots, d\} \), s.t. \( \hat{\beta}_j^r(\alpha) = 0 \) for all \( r = 1, \ldots, 100 \).

The significant covariates are found by the R-function \texttt{step()}\texttt{, which automatically selects the subset of covariates such that the corresponding AIC is maximized.}

The methods are denoted by \( m = \text{LQR} \) (“Linear quantile regression”) and \( m = \text{LQRWSA} \) (“Linear quantile regression with step applied”).

**Parametric D-vine quantile regression (DVQR, DVQRPCOR)**

[Kraus and Czado 2017] showed in their simulation study, that D-vine quantile regression (Section 5.2) is usually more accurate than other benchmark methods. That is why we compare it to our R-vine quantile regression algorithm. The method is applied using the R-function \texttt{vinereg()} from the \texttt{vinereg}-package and is denoted by \( m = \text{DVQR} \).

Note that the D-vine quantile regression algorithm automatically selects a subset of significant covariates

\[
U_{1,\text{DVQR}}, \ldots, U_{d,\text{DVQR}}.
\]

One version of the R-vine quantile regression algorithm, will be based on this subset.

Additionally, we analyze a new version of the D-vine quantile algorithm, that includes a fixed set of covariates, which is chosen by (partial) correlation arguments (Section 5.4, Algorithm 1). The method is denoted by \( m = \text{DVQRPCOR} \).

**R-vine quantile regression**

(RVQRDVQR, RVQRPCOR, RVQRkendall, RVQRsemi, RVQRindepTest)

In the R-vine quantile regression algorithm, we can choose whether we want to consider a fixed set of covariates \( U_{1,\text{RVQR}}, \ldots, U_{d,\text{RVQR}} \) (1), or if we want to assign an order \( j_1 < \cdots < j_{d} \) to all covariates (2), i.e. \( fC \in \{0, 1\} \).

(1) \( fC = 1 \):

The fixed set of covariates is either selected by the D-vine quantile regression algorithm, i.e.

\[
U_{j,\text{RVQR}} := U_{j,\text{DVQR}}, \quad j = 1, \ldots, d_{\text{RVQR}}, \quad r = 1, \ldots, 100,
\]

or by the selection of covariates by (partial) correlations. The methods are denoted by \( m = \text{RVQRDVQR} \) and \( m = \text{RVQRPCOR} \), respectively.
The order $j_1 < \cdots < j_{d_0}$ is selected by (partial) correlation arguments (Section 5.4 Algorithm 1). In order to assess the influence of the algorithm’s parameters, we additionally consider the following combination of the parameters $wc$ and indepTest (Remark 1.38).

- $wc = \tau$, indepTest = 0 ($m =$ RVQRkendall),
- $wc = \rho^{max}$, indepTest = 0 ($m =$ RVQRsemi),
- $wc = \tau$, indepTest = 1 ($m =$ RVQRindepTest).

The set of all considered methods is denoted by

$$\mathcal{M} := \{LQR, LQRWSA, DVQR, DVQRPCOR\} \cup \{RVQRPCOR, RVQRDVQR, RVQRkendall, RVQRsemi, RVQRindepTest\}.$$  

$\mathcal{M}$ is summarized in Table 7.1. Column $d_{m,r}$ indicates, if method $m$ chooses exactly $d_0$ covariates ($d_{m,r} = d_0$) or if the corresponding algorithm might stop after less than $d_0$ iterations ($d_{m,r} \leq d_0$). Which method $m$ is used to select the order of the covariates, is specified in column Order of covariates. PCOR refers to the selection of covariates by partial correlations (Algorithm 1, Section 5.4).

<table>
<thead>
<tr>
<th>$m$</th>
<th>$d_{m,r}$</th>
<th>Order of covariates</th>
<th>indepTest</th>
<th>$sc$</th>
<th>$wc$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear quantile regression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LQR</td>
<td>$d_0$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LQRWSA</td>
<td>$\leq d_0$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>D-vine quantile regression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DVQR</td>
<td>$\leq d_0$</td>
<td>DVQR</td>
<td>0</td>
<td>AIC</td>
<td>-</td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>$d_0$</td>
<td>PCOR</td>
<td>0</td>
<td>AIC</td>
<td>-</td>
</tr>
<tr>
<td>R-vine quantile regression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>$\leq d_0$</td>
<td>DVQR</td>
<td>0</td>
<td>AIC</td>
<td>$\tau$</td>
</tr>
<tr>
<td>RVQRPCOR</td>
<td>$d_0$</td>
<td>PCOR</td>
<td>0</td>
<td>AIC</td>
<td>$\tau$</td>
</tr>
<tr>
<td>RVQRkendall</td>
<td>$\leq d_0$</td>
<td>PCOR</td>
<td>0</td>
<td>AIC</td>
<td>$\rho^{max}$</td>
</tr>
<tr>
<td>RVQRsemi</td>
<td>$\leq d_0$</td>
<td>PCOR</td>
<td>0</td>
<td>AIC</td>
<td>$\rho^{max}$</td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>$\leq d_0$</td>
<td>PCOR</td>
<td>1</td>
<td>AIC</td>
<td>$\tau$</td>
</tr>
</tbody>
</table>

Table 7.1: Method overview.

7.2 In-sample conditional log likelihood comparison

7.2.1 Setup and results

In the following, we compare the initially simulated R-vine copulas $R^{d_0}_{\tau}$ and the $RVQRPCOR$-estimated R-vine copulas $R^{d_{RVQRPCOR,r}}_{RVQRPCOR,r}$, $r \in \{1, \ldots, 100\}$, regarding their
in-sample conditional log likelihoods in each of the scenarios.
We additionally consider an estimated initial R-vine copula

\[ \hat{R}^{d_0}_r := (\hat{\gamma}^{d_0}_r, \hat{\nu}^{d_0}_r, \hat{\Theta}^{d_0}_r), \]

which is based on the same R-vine tree sequence as \( R^{d_0}_r \) and of which the set of copula families \( \hat{B}^{d_0}_r \) and copula parameters \( \hat{\Theta}^{d_0}_r \) are estimated by maximizing the corresponding unconditional in-sample likelihoods. To identify \( \hat{R}^{d_0}_r \) in R, we use the R-function `RVineCopSelect()`.

Within this section, we introduce a simplified notation, which enables us to specify the replication \( r \in \{1, \ldots, 100\} \) and scenario \( S \in \{1, \ldots, 8\} \) corresponding to each of the three considered R-vine copulas:

- initially simulated R-vine copula: \( R^S_r := R^{d_0}_r \) in scenario \( S \),
- estimated initial R-vine copula: \( \hat{R}^S_r := \hat{R}^{d_0}_r \) in scenario \( S \),
- RVQRPCOR-estimated R-vine copula: \( \hat{R}^{RVQR,S}_r := \hat{R}^{RVQRPCOR,d_0}_r \) in scenario \( S \).

Denoting the initially simulated training data set in replication \( r \) and scenario \( S \) by

\[ \mathcal{U}^S_{r,train} := (v^S_r, u^S_1, \ldots, u^S_{d_0}) \in [0,1]^{n_{train} \times d_0}, \]

we can finally define the in-sample conditional log likelihood of each R-vine copula in replication \( r \in \{1, \ldots, 100\} \) and scenario \( S \in \{1, \ldots, 8\} \):

- in-sample c\( \ell \ell \) of the initially simulated R-vine copula:
  \[ c\ell\ell^S_r (R) := c\ell\ell(R^S_r, \mathcal{U}^S_{r,train}), \]
- in-sample c\( \ell \ell \) of the estimated initial R-vine copula:
  \[ c\ell\ell^S_r (\hat{R}) := c\ell\ell(\hat{R}^S_r, \mathcal{U}^S_{r,train}), \]
- in-sample c\( \ell \ell \) of the RVQRPCOR-estimated R-vine copula:
  \[ c\ell\ell^{RVQR}_r (\hat{R}) := c\ell\ell(\hat{R}^{RVQR,S}_r, \mathcal{U}^{S,train}_r). \]

In the following, we analyze the in-sample conditional log likelihoods of the initially simulated and the RVQRPCOR-estimated R-vine copulas proportional to the conditional log likelihoods of the estimated initial R-vine copulas, i.e. we compute the quotients

\[ \frac{c\ell\ell^S_r (R)}{c\ell\ell^S_r (\hat{R})} \quad \text{and} \quad \frac{c\ell\ell^{RVQR}_r (\hat{R})}{c\ell\ell^S_r (\hat{R})}. \]
for every $S \in \{1, \ldots, 8\}$ and $r \in \{1, \ldots, 100\}$. Figure 7.2 shows box plots of the sets

$$\left\{ \frac{c \ell^S_r(\mathcal{R})}{c \ell^S_r(\hat{\mathcal{R}})} \right\}_{r=1, \ldots, 100} \quad \text{and} \quad \left\{ \frac{c \ell^S_r(\hat{\mathcal{R}}_{RVQR})}{c \ell^S_r(\mathcal{R})} \right\}_{r=1, \ldots, 100}$$

for every $S \in \{1, \ldots, 8\}$.

Figure 7.2: Box plots of $\left\{ \frac{c \ell^S_r(\mathcal{R})}{c \ell^S_r(\hat{\mathcal{R}})} \right\}_{r=1, \ldots, 100}$ (red) and $\left\{ \frac{c \ell^S_r(\hat{\mathcal{R}}_{RVQR})}{c \ell^S_r(\mathcal{R})} \right\}_{r=1, \ldots, 100}$ (green) for every scenario $S$.

Note that, since $\hat{\mathcal{R}}^S_r$ is estimated by maximizing the unconditional in-sample likelihoods (not the conditional), in theory it may happen that

$$\frac{c \ell^S_r(\mathcal{R})}{c \ell^S_r(\hat{\mathcal{R}})} > 1 \quad \text{or} \quad \frac{c \ell^S_r(\hat{\mathcal{R}}_{RVQR})}{c \ell^S_r(\mathcal{R})} > 1.$$
We conclude that there is no significant deviation in the in-sample conditional log likelihoods corresponding to the initially simulated and the estimated initial R-vine copulas. Thus, we could compare $RVQRP_{COR}$ to any of the initial R-vine copulas. However, we will always refer to $\hat{RV}_R^{S}$ in the following.

Looking at the box plots of $\{c_{\ell\ell}^{S}(\hat{RV}_R^{RVQPR})\}_{r=1,\ldots,100}$, we observe that the medians corresponding to scenarios $S \in \{1, \ldots, 4\}$ lie closer to 1 than for scenarios $S \in \{5, \ldots, 8\}$. We conclude that, proportional to the initial models, the in-sample conditional log likelihoods of the $RVQRP_{COR}$-estimated R-vine copulas decrease with the number of covariates. Further, we observe an effect of $\text{indep}_{0}$, as for all $i \in \{1, 3, 5, 7\}$, the median in scenario $S = i + 1$ is greater than the median in scenario $S = i$. A possible interpretation is that allowing for independence within the initially simulated data, makes it easier for $m = RVQRP_{COR}$ to select an appropriate R-vine copula.

### 7.2.2 Effect of the selection of covariates

In order to understand the effect of the selection of covariates on $\frac{c_{\ell\ell}^{S}(\hat{RV}_R^{RVQPR})}{c_{\ell\ell}^{S}(\hat{R})}$, we first have a closer look at scenario $S = 1$.

We observe that

$$r_{\text{min}} := \arg \min_{r \in \{1, \ldots, 100\}} \frac{c_{\ell\ell}^{1}(\hat{RV}_R^{RVQPR})}{c_{\ell\ell}^{1}(\hat{R})} = 8, \quad \frac{c_{\ell\ell}^{8}(\hat{RV}_R^{RVQPR})}{c_{\ell\ell}^{8}(\hat{R})} = 0.6184$$

and

$$r_{\text{max}} := \arg \max_{r \in \{1, \ldots, 100\}} \frac{c_{\ell\ell}^{1}(\hat{RV}_R^{RVQPR})}{c_{\ell\ell}^{1}(\hat{R})} = 45, \quad \frac{c_{\ell\ell}^{45}(\hat{RV}_R^{RVQPR})}{c_{\ell\ell}^{45}(\hat{R})} = 0.9984.$$  

Thus, the proportional conditional log likelihood gets maximal for $r = 45$ and minimal for $r = 8$. According to Figure 7.2, we observe that $r = 8$ is the only outlying replication in scenario $S = 1$. Therefore, we consider $r = 3$ instead of $r = 8$, for which we obtain the second lowest quotient, i.e.

$$\frac{c_{\ell\ell}^{3}(\hat{RV}_R^{RVQPR})}{c_{\ell\ell}^{3}(\hat{R})} = 0.7118.$$  

Using the R-function `plot.RVineMatrix()`, we plot the first four trees of the initially sampled and the $RVQRP_{COR}$-selected R-vine tree sequences in replications $r = 3$ and $r = 45$. The corresponding trees can be found in Figure 7.3 and Figure 7.4 respectively. Within the plots, covariate $U_j$ is denoted by $j$, $j = 1, \ldots, 8$, and the response $V$ is denoted by 9. Similar plots for scenarios $S \in \{2, 3, 4\}$ are presented in Appendix A.1.2.
Figure 7.3: Trees of order 1-4 corresponding to the initial (true) copula $\hat{R}_{13}$ and the RVQRPCOR-estimated copula $R_{13}^{RVQRPCOR}$ in scenario $S = 1$ and replication $r = 3$ (pair copula families and corresponding Kendall’s $\tau$ values are shown on the edges).
Figure 7.4: Trees of order 1-4 corresponding to the initial (true) copula $R_{14}^1$ and the $RVQRPCOR$-estimated copula $\hat{R}_{145}^{RVQRPCOR}$ in scenario $S = 1$ and replication $r = 45$ (pair copula families and corresponding Kendall’s $\tau$ values are shown on the edges).
We denote the initially sampled (true) and the \( RVQRPCOR \)-selected partner of the response \( V \) in tree \( k \in \{1, \ldots, 8\} \), replication \( r \in \{1, \ldots, 100\} \) and scenario \( S \in \{1, \ldots, 8\} \) by 
\[
pt_k^{V,0}(S, r) \text{ and } pt_k^{V,RVQR}(S, r),
\]
and observe that for \( r = 3 \) the first four true partners of the response \( V \) are sampled as
\[
pt_1^{V,0}(1, 3) = U_1, \ pt_2^{V,0}(1, 3) = U_2, \ pt_3^{V,0}(1, 3) = U_3, \ pt_4^{V,0}(1, 3) = U_7,
\]
whereas \( RVQRPCOR \) selects the covariates (partners of response \( V \)) as
\[
pt_1^{V,RVQR}(1, 3) = U_5, \ pt_2^{V,RVQR}(1, 3) = U_4, \ pt_3^{V,RVQR}(1, 3) = U_2, \ pt_4^{V,RVQR}(1, 3) = U_1.
\]
For \( r = 45 \) the corresponding partners are given by
\[
pt_1^{V,0}(1, 45) = U_1, \ pt_2^{V,0}(1, 45) = U_2, \ pt_3^{V,0}(1, 45) = U_4, \ pt_4^{V,0}(1, 45) = U_3
\]
and
\[
pt_1^{V,RVQR}(1, 45) = U_1, \ pt_2^{V,RVQR}(1, 45) = U_2, \ pt_3^{V,RVQR}(1, 45) = U_4, \ pt_4^{V,RVQR}(1, 45) = U_5.
\]
While in replication \( r = 3 \) the selection of covariates by partial correlation (see Section 5.4) does not even find the first true covariate, it identifies the first 3 covariates in replication \( r = 45 \).

In the following, we analyze if there is an effect of the number of correctly specified covariates on the proportional in-sample conditional log likelihood corresponding to the \( RVQRPCOR \)-estimated R-vine copula.

The number of correctly specified covariates in scenario \( S \) and replication \( r \) is defined as
\[
\text{nCor}_r^S := 1_{\{pt_1^{V,0}(S, r) = pt_1^{V,RVQR}(S, r)\}} \times \max\{k \in \{1, \ldots, d_0\} : pt_k^{V,0}(S, r) = pt_k^{V,RVQR}(S, \nabla), \forall \ell = 1, \ldots, k\}.
\]

In order to illustrate the meaning of \( \text{nCor}_r^S \), we once again consider the example above. As in \( r = 3 \) we have that \( pt_1^{V,0}(1, 3) = U_1 \) and \( pt_1^{V,RVQR}(1, 3) = U_5 \), we already now that \( \text{nCor}_3^1 = 0 \). Whereas in \( r = 45 \) we observe that \( pt_1^{V,0}(1, 45) = pt_1^{V,RVQR}(1, 45) \) for all \( \ell = 1, \ldots, 3 \), but \( pt_4^{V,0}(1, 45) \neq pt_4^{V,RVQR}(1, 45) \). Thus, all covariates up to \( k = \text{nCor}_{45}^1 = 3 \) are specified correctly.

First we have a look at the frequency of \( \text{nCor}_r^S \), i.e. for every scenario \( S \in \{1, \ldots, 8\} \) we answer the question in how many of the replications \( \text{nCor}_r^S \) equals \( i \in \{0, 1, \ldots, d_0\} \). Therefore, we group the replications \( \{1, \ldots, 100\} \) into subsets 
\[
G_i^S := \{r \in \{1, \ldots, 100\} : \text{nCor}_r^S = i\} \subseteq \{1, \ldots, 100\}
\]
for every scenario \( S \in \{1, \ldots, 8\} \). Figure 7.5 shows bar plots of \( |G_i^S| \) for every \( S \in \{1, \ldots, 8\} \) and \( i \in \{1, \ldots, d_0\} : |G_i^S| > 0\).
In most of the replications - more than a half in every scenario - the selection of covariates by partial correlations does not identify the first true covariate. For scenarios \( S = 1, 2 \), there is one replication \( r = 95 \) in which the first 4 covariates are selected correctly. The selection seems to get worse the higher \( d_0 \), as \(|G^S_0|\) for \( S \in \{5, \ldots, 8\} \) is greater than \(|G^S_0|\) for \( S \in \{1, \ldots, 4\} \) and accordingly there are less replications in \( G^S_i \), \( i \in \{1, \ldots, 4\} \) for \( S \in \{5, \ldots, 8\} \) than for \( S \in \{1, \ldots, 4\} \).

Since \(|G^S_i|, i \in \{2, 3, 4\}\), is small in every scenario, we prefer a graphical analysis of the effect of the covariate selection on the proportional in-sample conditional log likelihood corresponding to the \( RVQRPCOR \)-estimated R-vine copula. Figure 7.6 shows box plots of

\[
\left\{ \frac{\ell^S_r(\widehat{R}^{RVQR})}{\ell^S_r(\widehat{R})} \right\}_{r \in G^S_i}
\]

for every \( S \in \{1, \ldots, 8\} \) and \( i \in \{1, \ldots, d_0\} : |G^S_i| > 0 \).
Obviousy, the proportional conditional log likelihood of the RVQRPCOR-estimated R-vine copula is increasing with the number of correctly selected covariates. Thus, the performance of RVQRPCOR is strongly connected to the goodness of the selection of covariates by partial correlations.

7.3  Out-of-sample evaluation

7.3.1  Considered quantities

A reasonable way of quantifying a method’s out-of-sample prediction accuracy is to calculate the so-called averaged tick-loss, which was also considered by Kraus and Czado [2017] in their data application.

Definition 7.1. (Averaged tick loss)
Let \( \mathcal{U}_{r,d_{0}}^{eval} \in [0,1]^{n_{eval} \times (d_{0}+1)} \) be the simulated evaluation data set in iteration \( r = 1, \ldots, R \) and let \( \hat{Q}_{1}^{r,m,r}, \ldots, \hat{Q}_{d_{0}}^{r,m,r} \) be the predicted \( \alpha \)-quantile of method \( m \) in iteration \( r = 1, \ldots, 100 \).

For \( \alpha \in (0,1) \) the averaged tick loss of method \( m \) and replication \( r \in \{1, \ldots, R\} \) is defined as:

\[
\text{avgTickLoss}(m,r) = \frac{1}{n_{eval}} \sum_{i=1}^{n_{eval}} (\hat{Q}_{1}^{r,m,r}(u_{i1}^{r,eval}, \ldots, u_{d_{0}}^{r,eval}) - u_{i1}^{r,eval})^2
\]
as
\[ \text{ATL}_m^r(\alpha) = \frac{1}{n_{\text{eval}}} \sum_{i=1}^{n_{\text{eval}}} \rho_\alpha(v_{r,i}^{\text{eval}} - \hat{Q}_V^{m,r}_{U_{1,i}^{r,\text{eval}}, \ldots, U_{dm,r}^{r,\text{eval}}}(\alpha|u_{r,1}^{r,\text{eval}}, \ldots, u_{r,id_0}^{r,\text{eval}})), \]
where \( \rho_\alpha(x) := x(\alpha - 1_{\{x<0\}}) \) (see also Definition 3.3).

For each scenario \( S \in \{1, \ldots, 8\} \), for all \( m \in M \) and \( \alpha \in A := \{0.05, 0.25, 0.5, 0.75, 0.95\} \),
we visualize \( \{\text{ATL}_m^r(\alpha)\}_{r=1,\ldots,100} \)
using box plot diagrams.
Additionally, we consider the **medians of the computed averaged tick losses**
\[ \text{MEDATL}_m^r(\alpha) := \text{median}(\{\text{ATL}_m^r(\alpha)\}_{r=1,\ldots,100}) \]
and the **mean integrated averaged tick losses**
\[ \text{MIATL}_m^r(\alpha) = \frac{1}{R} \sum_{r=1}^{R} \text{ATL}_m^r(\alpha) \]
for all \( m \in M \) and \( \alpha \in A \). Therefore, we present tables covering the computed values
proportional to the corresponding values induced by \( m = \text{LQR} \), i.e. we analyze
\[ \text{RELMEDATL}_m^r(\alpha) := \frac{\text{MEDATL}_m^r(\alpha)}{\text{MEDATL}_m^r(\alpha)} \]
and
\[ \text{RELMIATL}_m^r(\alpha) := \frac{\text{MIATL}_m^r(\alpha)}{\text{MIATL}_m^r(\alpha)} \]
for every method \( m \in M \) and every \( \alpha \in A \) in each scenario.
Furthermore, we analyze the estimated R-vine copulas considering their vine types, i.e.
for \( m \in M_{\text{vine}} := \{\text{DVQR, DVQRPCOR, RVQRDVQR, RVQRPCOR, RVQRkendall, RVQRsemi, RVQRindepTest}\} \) we use bar plots to visualize the quantities
\[ N_m^{C_{\text{vine}}} := |\{r \in \{1, \ldots, 100\} : \mathcal{V}_{m,r}^{d_{m,r}} \text{ is a C-vine tree sequence}\}|, \]
\[ N_m^{D_{\text{vine}}} := |\{r \in \{1, \ldots, 100\} : \mathcal{V}_{m,r}^{d_{m,r}} \text{ is a D-vine tree sequence}\}| \]
and
\[ N_m^{C/D_{\text{vine}}} := |\{r \in \{1, \ldots, 100\} : \mathcal{V}_{m,r}^{d_{m,r}} \text{ is a C- and D-vine tree sequence}\}|. \]
In general, it holds that a regular vine tree sequence is a C-vine and D-vine at the same
time, if it contains two or three elements. Thus, \( \mathcal{V}_{m,r}^{d_{m,r}} \) is a C- and D-vine tree sequence,
if in replication $r$ method $m \in M_{\text{vine}}$ selects only one or two covariates.

We have chosen $\text{indep}_0^S = 1$ for scenarios $S \in \{2, 4, 6, 8\}$. For those, we additionally analyze what percentage of the “independent” variables have been included in model $m$, i.e. for $m \in M$ we visualize the set

$$\{N_{\text{indep}}^{m,r}\}_{r=1,...,100},$$

in box plot diagrams, where

$$N_{\text{indep}}^{m,r} := 100 \times \frac{|\{U_{d_{1,m}}^r, \ldots, U_{d_{r,m}}^r\} \cap \text{Cov}_{\text{indep}}^r|}{|\text{Cov}_{\text{indep}}^r|}.$$

The corresponding box plots can be found in Appendix A.1.3.

### 7.3.2 Results in scenarios 1, 5, 7 and 8

In order to be able to assess the effects of the scenario parameters, we discuss the results of scenarios $S = 1, 5, 7, 8$. Figure 7.7 shows the change in parameters when we compare scenarios $S = 1$ and $S = 5$, $S = 5$ and $S = 7$, $S = 7$ and $S = 8$.

We can discuss the effect of $d_0$ comparing scenarios $S = 1$ and $S = 5$, the effect of $n_{\text{train}}$, $n_{\text{eval}}$ comparing $S = 5$ and $S = 7$ and the effect of $\text{indep}_0$ comparing $S = 7$ and $S = 8$.

$$n_{\text{train}}^5 = 1000 \rightarrow n_{\text{train}}^7 = 2000$$

$$d_0^1 = 8 \rightarrow d_0^5 = 20$$

$$\text{indep}_0^7 = 0 \rightarrow \text{indep}_0^8 = 1$$

$$n_{\text{eval}}^5 = 500 \rightarrow n_{\text{eval}}^7 = 1000$$

Figure 7.7: Comparison of parameters in scenarios $S = 1, 5, 7, 8$.

The results of the remaining scenarios $S = 2, 3, 4, 6$ can be found in Appendix A.1.3.
Scenario 1

Figure 7.8 shows box plots of $\{ATL_{r}^{m}(\alpha)\}_{r=1,...,100}$ for every method $m \in M$ and $\alpha \in A$ in scenario $S = 1$. Table 7.2 shows $\text{RELMIA}^{m}(\alpha)$ and $\text{RELMED}^{m}(\alpha)$ for all $m \in M$ and $\alpha \in A$. For each $\alpha \in A$ the best performing method’s value is printed in bold letters.

![Box plots of ATL \(_{r}^{m}(\alpha)\) for every method m ∈ M and α ∈ A in scenario S = 1.](image)

Table 7.2: Scenario 1: box plots of $\{ATL_{r}^{m}(\alpha)\}_{r=1,...,100}$ for all $m \in M$ and $\alpha \in A$.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>RELMIATL($\alpha$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LQR</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>LQRWSA</td>
<td>1.0001</td>
<td>1.0001</td>
<td>0.9999</td>
<td>0.9999</td>
<td>1.0005</td>
</tr>
<tr>
<td>DVQR</td>
<td>0.8112</td>
<td><strong>0.8526</strong></td>
<td>0.8670</td>
<td>0.8444</td>
<td><strong>0.8196</strong></td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.8137</td>
<td>0.8545</td>
<td>0.8662</td>
<td>0.8226</td>
<td>0.8227</td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>0.8109</td>
<td>0.8546</td>
<td>0.8675</td>
<td>0.8567</td>
<td>0.8344</td>
</tr>
<tr>
<td>RVQRPCOR</td>
<td>0.8118</td>
<td>0.8534</td>
<td><strong>0.8641</strong></td>
<td>0.8520</td>
<td>0.8278</td>
</tr>
<tr>
<td>RVQRandall</td>
<td>0.8319</td>
<td>0.8743</td>
<td>0.8853</td>
<td>0.8766</td>
<td>0.8506</td>
</tr>
<tr>
<td>RVQRemi</td>
<td>0.8327</td>
<td>0.8760</td>
<td>0.8862</td>
<td>0.8772</td>
<td>0.8494</td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>0.8323</td>
<td>0.8748</td>
<td>0.8866</td>
<td>0.8780</td>
<td>0.8530</td>
</tr>
<tr>
<td>RELMEDATL($\alpha$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LQR</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>LQRWSA</td>
<td>0.9962</td>
<td>1.0005</td>
<td>1.0000</td>
<td>0.9976</td>
<td>0.9992</td>
</tr>
<tr>
<td>DVQR</td>
<td><strong>0.7921</strong></td>
<td>0.8558</td>
<td>0.9010</td>
<td>0.8960</td>
<td><strong>0.7719</strong></td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.7975</td>
<td>0.8880</td>
<td>0.9130</td>
<td>0.8976</td>
<td>0.8044</td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>0.7986</td>
<td>0.8889</td>
<td>0.9130</td>
<td>0.8888</td>
<td>0.7876</td>
</tr>
<tr>
<td>RVQRPCOR</td>
<td>0.7991</td>
<td><strong>0.8791</strong></td>
<td><strong>0.8969</strong></td>
<td><strong>0.8879</strong></td>
<td><strong>0.7904</strong></td>
</tr>
<tr>
<td>RVQRandall</td>
<td>0.7982</td>
<td>0.8975</td>
<td>0.9070</td>
<td>0.9183</td>
<td>0.7891</td>
</tr>
<tr>
<td>RVQRemi</td>
<td>0.8133</td>
<td>0.9024</td>
<td>0.8977</td>
<td>0.9103</td>
<td>0.7966</td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>0.7973</td>
<td>0.8975</td>
<td>0.9070</td>
<td>0.9220</td>
<td>0.7883</td>
</tr>
</tbody>
</table>

Table 114
Obviously, LQR and LQRWSA perform worse than the vine based methods \( m \in \mathcal{M}_{\text{vine}} \), as for all \( \alpha \in \mathcal{A} \) the range of averaged tick losses induced by LQR and LQRWSA lies noticeably above all other ranges. Comparing the box plots corresponding to the vine based methods, we are not able to identify a clear order of their performances. However, we do observe that all averaged tick losses are skewed to the left. Therefore, it is reasonable to assess the (average) performance of a method \( m \in \mathcal{M} \) based on the median of the averaged tick losses \( \{\text{ATL}_r^m(\alpha)\}_{r=1,...,100} \), which is equivalent to comparing \( \text{RELMEDATL}_m^m(\alpha) \) for all \( m \in \mathcal{M} \) and \( \alpha \in \mathcal{A} \). Comparison of the corresponding mean integrated averaged tick losses (\( \text{RELMIATL}_m^m(\alpha) \)) provides additional information on the influence of the outlying averaged tick losses. We conclude that (on average) \( m = \text{DVQR} \) and \( m = \text{RVQRPCOR} \) performed best for \( \alpha \in \{0.05, 0.95\} \) and \( \alpha \in \{0.25, 0.50, 0.75\} \), respectively.

Next, we discuss the selected types of R-vine tree sequences. Figure 7.9 visualizes \( N_{\text{D-vine}}^m \), \( N_{\text{C-vine}}^m \) and \( N_{\text{C/D-vine}}^m \) for \( m \in \mathcal{M}_{\text{vine}} \). Furthermore, it covers the types of the initially sampled R-vine tree sequences \( V_d^0 \) ("R_0").

We observe that already for \( d_0^1 = 8 \) initial covariates, none of the 100 initially sampled R-vine tree sequences is of C- or D-vine structure. As DVQR and DVQRPCOR estimate the corresponding conditional quantiles based on D-vine copulas, they only allow for D-vine tree sequences in each replication \( r = 1, \ldots, 100 \). Not surprisingly RVQRPCOR solely selects general R-vines, as it always includes all \( d_8^1 = 8 \) covariates and C-vine D-vine tree sequences get quite unlikely in set of R-vine tree sequences on 8 elements.
Methods RVQRDVQR, RVQRkendall, RVQRssemi and RVQRindepTest, which other than \( m = \text{RVQRPCOR} \) might stop before the inclusion of all 8 covariates, select general R-vines in most of the replications. While for RVQRDVQR the vine types deviate from general R-vines in less than 20 replications, RVQRkendall, RVQRssemi and RVQRindepTest select C- or D-vines more often. Note that there are also replications in which they choose “C/D”-vines, which means that they do not include more than 3 of the 8 covariates.
Scenario 5

Figure 7.10 shows box plots of $\{\text{ATL}_m^r(\alpha)\}_{r=1,\ldots,100}$ for every method $m \in M$ and $\alpha \in A$ in scenario $S = 5$. Table 7.3 lists $\text{RELMIATL}_m^r(\alpha)$ and $\text{RELMEDATL}_m^r(\alpha)$ for all $m \in M$ and $\alpha \in A$. For each $\alpha \in A$ the best performing method’s value is printed in bold letters.

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RELMIATL(m)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LQR</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>LQRWSA</td>
<td>0.9984</td>
<td>0.9989</td>
<td>1.0069</td>
<td>1.0019</td>
<td>0.9981</td>
</tr>
<tr>
<td>DVQR</td>
<td>0.8451</td>
<td>0.8967</td>
<td>0.9121</td>
<td>0.9014</td>
<td>0.8472</td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.8300</td>
<td>0.8907</td>
<td>0.9059</td>
<td>0.8905</td>
<td>0.8410</td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>0.8284</td>
<td>0.8841</td>
<td>0.8970</td>
<td>0.8835</td>
<td>0.8325</td>
</tr>
<tr>
<td>RVQRkendall</td>
<td>0.8760</td>
<td>0.9305</td>
<td>0.9457</td>
<td>0.9348</td>
<td>0.8998</td>
</tr>
<tr>
<td>RVQRsemi</td>
<td>0.8759</td>
<td>0.9308</td>
<td>0.9452</td>
<td>0.9347</td>
<td>0.8886</td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>0.8766</td>
<td>0.9313</td>
<td>0.9457</td>
<td>0.9356</td>
<td>0.8914</td>
</tr>
<tr>
<td><strong>RELMEDATL(m)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LQR</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>LQRWSA</td>
<td>0.9987</td>
<td>1.0005</td>
<td>1.0021</td>
<td>1.0083</td>
<td>0.9993</td>
</tr>
<tr>
<td>DVQR</td>
<td>0.8270</td>
<td>0.8768</td>
<td>0.9033</td>
<td>0.8884</td>
<td>0.7259</td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.7926</td>
<td>0.8936</td>
<td>0.9048</td>
<td>0.8839</td>
<td>0.7425</td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>0.8146</td>
<td>0.8844</td>
<td>0.9004</td>
<td>0.8809</td>
<td>0.7615</td>
</tr>
<tr>
<td>RVQRkendall</td>
<td>0.7785</td>
<td>0.8645</td>
<td>0.8632</td>
<td>0.8601</td>
<td>0.7256</td>
</tr>
<tr>
<td>RVQRsemi</td>
<td>0.8184</td>
<td>0.9064</td>
<td>0.9189</td>
<td>0.9148</td>
<td>0.7781</td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>0.8184</td>
<td>0.9069</td>
<td>0.9047</td>
<td>0.8956</td>
<td>0.7735</td>
</tr>
</tbody>
</table>

Table 7.3: Scenario 5: $\frac{\text{MIATL}_m^r(\alpha)}{\text{MIAATL}_m^r(\alpha)}$ and $\frac{\text{MEDATL}_m^r(\alpha)}{\text{MEDATL}_m^r(\alpha)}$ for $m \in M$ and $\alpha \in A$.  

117
Figure 7.11 visualizes $N^m_{D-vine}$, $N^m_{C-vine}$ and $N^m_{C/D-vine}$ for $m \in M_{vine}$.

Comparing the outcomes in scenarios $S = 1$ and $S = 5$, we assess the effect of increasing the number of initial covariates from $d^1_0 = 8$ to $d^5_0 = 20$. Looking at the box plots in Figure 7.8 and Figure 7.10, we observe that increasing the number of covariates has no effect on the skewness in the averaged tick losses. However, we find that there are less outlying replications for all methods $m \in M$. According to values listed in Table 7.3, $RVQRPCOR$ is now optimal for all $\alpha \in A$, as it outperforms all other methods regarding their RELMEDATL$^m(\alpha)$- and RELMIATL$^m(\alpha)$-outcomes.

As C- and D-vines get less likely for larger numbers of elements in an R-vine tree sequence, it is not surprising that methods $m = RVQRDVQR$, $m = RVQRkendall$, $m = RVQRsemi$ and $RVQRindepTest$ now select less C-, D- or C/D-vines.
Scenario 7

Figure 7.12 shows box plots of $\{\text{ATL}_m^r(\alpha)\}_{r=1,...,100}$ for every method $m \in \mathcal{M}$ and $\alpha \in \mathcal{A}$ in scenario $S = 7$. Table 7.4 lists $\text{RELMIA}_{\text{ATL}}^m(\alpha)$ and $\text{RELMED}_{\text{ATL}}^m(\alpha)$ for all $m \in \mathcal{M}$ and $\alpha \in \mathcal{A}$. For each $\alpha \in \mathcal{A}$ the best performing method’s value is printed in bold letters.

![Figure 7.12: Scenario 7: box plots of $\{\text{ATL}_m^r(\alpha)\}_{r=1,...,100}$ for all $m \in \mathcal{M}$ and $\alpha \in \mathcal{A}$.](image)

<table>
<thead>
<tr>
<th>$m$</th>
<th>$\alpha$</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>LQR</td>
<td>RELMIA_{ATL}^m(\alpha)</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>LQRWSA</td>
<td>1.0002</td>
<td>1.0005</td>
<td>1.0007</td>
<td>1.0009</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>DVQR</td>
<td>0.8427</td>
<td>0.8857</td>
<td>0.8980</td>
<td>0.8825</td>
<td>0.8343</td>
<td></td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.8359</td>
<td>0.8829</td>
<td>0.8967</td>
<td>0.8815</td>
<td>0.8368</td>
<td></td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>0.8304</td>
<td>0.8829</td>
<td>0.9074</td>
<td>0.8980</td>
<td>0.8412</td>
<td></td>
</tr>
<tr>
<td>RVQRkendall</td>
<td>0.8285</td>
<td>0.8733</td>
<td>0.8882</td>
<td>0.8711</td>
<td>0.8221</td>
<td></td>
</tr>
<tr>
<td>RVQRsemisemidependentTest</td>
<td>0.8806</td>
<td>0.9211</td>
<td>0.9332</td>
<td>0.9199</td>
<td>0.8806</td>
<td></td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>0.8812</td>
<td>0.9221</td>
<td>0.9345</td>
<td>0.9187</td>
<td>0.8793</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$m$</th>
<th>RELMED_{ATL}^m(\alpha)</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>LQR</td>
<td>RELMED_{ATL}^m(\alpha)</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>LQRWSA</td>
<td>1.0010</td>
<td>1.0015</td>
<td>1.0017</td>
<td>1.0019</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>DVQR</td>
<td>0.8161</td>
<td>0.8426</td>
<td>0.8615</td>
<td>0.8542</td>
<td>0.7667</td>
<td></td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.8171</td>
<td>0.8379</td>
<td>0.8614</td>
<td>0.8564</td>
<td>0.7526</td>
<td></td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>0.8261</td>
<td>0.8546</td>
<td>0.8726</td>
<td>0.8749</td>
<td>0.7629</td>
<td></td>
</tr>
<tr>
<td>RVQRkendall</td>
<td>0.8089</td>
<td>0.8466</td>
<td>0.8605</td>
<td>0.8715</td>
<td>0.7308</td>
<td></td>
</tr>
<tr>
<td>RVQRsemisemidependentTest</td>
<td>0.8459</td>
<td>0.8864</td>
<td>0.9139</td>
<td>0.9002</td>
<td>0.8061</td>
<td></td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>0.8589</td>
<td>0.8944</td>
<td>0.9113</td>
<td>0.8930</td>
<td>0.8052</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.4: Scenario 7: $\frac{\text{MIATL}_m^r(\alpha)}{\text{MIATL}_m^r(\alpha)}$ and $\frac{\text{MEATL}_m^r(\alpha)}{\text{MEATL}_m^r(\alpha)}$ for $m \in \mathcal{M}$ and $\alpha \in \mathcal{A}$.
Figure 7.13 visualizes $N^m_{\text{D-vine}}$, $N^m_{\text{C-vine}}$ and $N^m_{\text{C/D-vine}}$ for $m \in \mathcal{M}_{\text{vine}}$.

After increasing the training and evaluation sample sizes from $n^5_{\text{train}} = 1000$ to $n^7_{\text{train}} = 2000$ and from $n^5_{\text{eval}} = 500$ to $n^7_{\text{eval}} = 1000$, we do not observe a clear difference in the structure of the box plots (Figure 7.10, Figure 7.12). Regarding the mean integrated averaged tick losses, $\text{RVQRPCOR}$ still performs optimal for all $\alpha \in \mathcal{A}$. For $\alpha \in \{0.25, 0.75\}$ its RELMEDATL($\alpha$)-values are greater than the corresponding outcomes for the D-vine based methods $m = \text{DVQR}$ and $m = \text{DVQRPCOR}$. However, the performance of $\text{RVQRPCOR}$ became better for $\alpha \in \{0.05, 0.95\}$.

Further, we observe that the number of selected C-, D- and C/D-vines has decreased for methods $m = \text{RVQRDVQR}$, $m = \text{RVQRkendall}$, $m = \text{RVQRsemi}$ and $\text{RVQRindepTest}$. 

Figure 7.13: Scenario 7: bar plot of $N^m_{\text{D-vine}}$, $N^m_{\text{C-vine}}$ and $N^m_{\text{C/D-vine}}$ for $m \in \mathcal{M}_{\text{vine}}$. 

120
Scenario 8

Figure 7.14 shows box plots of $\{\text{ATL}^m_\alpha(\alpha)\}_{r=1,\ldots,100}$ for every method $m \in \mathcal{M}$ and $\alpha \in \mathcal{A}$ in scenario $S = 8$. Table 7.5 lists $\text{RELMIA} \text{TL}_m^\alpha(\alpha)$ and $\text{RELMEDATL}_m^\alpha(\alpha)$ for all $m \in \mathcal{M}$ and $\alpha \in \mathcal{A}$. For each $\alpha \in \mathcal{A}$ the best performing method’s value is printed in bold letters.

Table 7.5: Scenario 8: $\frac{\text{MIATL}_m^\alpha(\alpha)}{\text{MEDATL}_m^\alpha(\alpha)}$ and $\frac{\text{MEDATL}_m^\alpha(\alpha)}{\text{MEDATL}_m^\alpha(\alpha)}$ for $m \in \mathcal{M}$ and $\alpha \in \mathcal{A}$. 

<table>
<thead>
<tr>
<th>$m$</th>
<th>$\alpha$</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>LQR</td>
<td>RELMIATL</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>LQRWSA</td>
<td></td>
<td>1.0005</td>
<td>1.0008</td>
<td>1.0006</td>
<td>1.0003</td>
<td>1.0006</td>
</tr>
<tr>
<td>DVQR</td>
<td>RELMIATL</td>
<td>0.8591</td>
<td>0.8929</td>
<td>0.9017</td>
<td>0.8903</td>
<td>0.8584</td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td></td>
<td>0.8603</td>
<td>0.8910</td>
<td>0.8989</td>
<td>0.8885</td>
<td>0.8579</td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>RELMIATL</td>
<td>0.8577</td>
<td>0.8878</td>
<td>0.8961</td>
<td>0.8837</td>
<td>0.8587</td>
</tr>
<tr>
<td>RVQRPCOR</td>
<td></td>
<td>0.8576</td>
<td>0.8876</td>
<td>0.8961</td>
<td>0.8837</td>
<td>0.8587</td>
</tr>
<tr>
<td>RVQR</td>
<td>RELMIATL</td>
<td>0.8896</td>
<td>0.9284</td>
<td>0.9386</td>
<td>0.9256</td>
<td>0.8860</td>
</tr>
<tr>
<td>RVQRsem</td>
<td></td>
<td>0.8942</td>
<td>0.9354</td>
<td>0.9458</td>
<td>0.9331</td>
<td>0.8930</td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>RELMIATL</td>
<td>0.8889</td>
<td>0.9268</td>
<td>0.9377</td>
<td>0.9251</td>
<td>0.8854</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$m$</th>
<th>$\alpha$</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>LQR</td>
<td>RELMEDATL</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>LQRWSA</td>
<td></td>
<td>0.9946</td>
<td>0.9998</td>
<td>1.0002</td>
<td>0.9993</td>
<td>1.0021</td>
</tr>
<tr>
<td>DVQR</td>
<td>RELMEDATL</td>
<td>0.8082</td>
<td>0.8332</td>
<td>0.8560</td>
<td>0.8485</td>
<td>0.8031</td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td></td>
<td>0.7836</td>
<td>0.8263</td>
<td>0.8620</td>
<td>0.8726</td>
<td>0.7800</td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>RELMEDATL</td>
<td>0.8154</td>
<td>0.8373</td>
<td>0.8665</td>
<td>0.8685</td>
<td>0.7951</td>
</tr>
<tr>
<td>RVQRPCOR</td>
<td></td>
<td>0.8011</td>
<td>0.8302</td>
<td>0.8612</td>
<td>0.8690</td>
<td>0.7573</td>
</tr>
<tr>
<td>RVQR</td>
<td>RELMEDATL</td>
<td>0.8355</td>
<td>0.8673</td>
<td>0.8747</td>
<td>0.8690</td>
<td>0.7811</td>
</tr>
<tr>
<td>RVQRsem</td>
<td></td>
<td>0.8648</td>
<td>0.8847</td>
<td>0.8999</td>
<td>0.8757</td>
<td>0.7777</td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>RELMEDATL</td>
<td>0.8346</td>
<td>0.8716</td>
<td>0.8839</td>
<td>0.8765</td>
<td>0.7863</td>
</tr>
</tbody>
</table>
Figure 7.15 visualizes $N^m_{D\text{-vine}}, N^m_{C\text{-vine}}$ and $N^m_{C/D\text{-vine}}$ for $m \in M_{\text{vine}}$.

Also choosing some of the initial pair copulas, which correspond to edges with the response variable in the conditioned set, as the independence copula ($\text{indep}_7^0 = 0, \text{indep}_8^0 = 1$), does not have a considerably effect on the structure of the averaged tick losses (Figure 7.12, Figure 7.14). Although in scenario $S = 8$, RVQRPCOR performs worse than in scenario $S = 7$, it is still optimal regarding the mean integrated averaged tick losses. While the D-vine based methods DVQR and DVQRPCOR perform better than RVQRPCOR for $\alpha \in \{0.05, 0.25, 0.5, 0.75\}$, RVQRPCOR beats all other methods for $\alpha = 0.95$.

The number of selected C-, D- or C/D-vines does not differ in scenarios 7 and 8.
7.3.3 Summary

In the following, we summarize the results presented in Section 7.3.2.

Selected vine types

As shown in the bar plots on the vine types, there is no scenario $S \in \{1, \ldots, 8\}$ and replication $r \in \{1, \ldots, 100\}$ in which the initially sampled R-vine tree sequence has C- or D-vine structure.

In every scenario $S \in \{1, \ldots, 8\}$, we have

$$N_{D\text{-vine}}^m = 100, \text{ for } m \in \{\text{DVQR, DVQRPCOR}\},$$

since the D-vine quantile regression algorithm (Section 5.2) only allows the selection of D-vine copulas.

Looking at the vine types chosen by the R-vine quantile regression methods, we observe that RVQRPCOR never chooses a D- or C-vine tree sequence. Whereas for every scenario $S \in \{1, \ldots, 8\}$ there are replications $r \in \{1, \ldots, 100\}$ in which methods $m \in \{\text{RVQRDVQR, RVQRkendall, RVQRsemi, RVQRindepTest}\}$ select D- and C-vines. This is based on the fact, that D- and C-vine structures get more likely in lower dimensions, and while RVQRPCOR includes all $d^S_0$ initial covariates in the copula, the other methods may include less than $d^S_0$ covariates. This also explains why methods $m \in \{\text{RVQRDVQR, RVQRkendall, RVQRsemi, RVQRindepTest}\}$ select less D- and C-vines in scenarios $S = 5, \ldots, 8$ than in scenarios $S = 1, \ldots, 4$, since

$$d^S_0 = \begin{cases} 8, & 1 \leq S \leq 4, \\ 20, & 5 \leq S \leq 8. \end{cases}$$

Prediction accuracy

While the vine-based methods ($m \in M_{\text{vine}}$) outperform the linear quantile regression methods ($m \in \{\text{LQR, LQRWSA}\}$) by a great margin, we do not observe a great difference in the prediction accuracy within $M_{\text{vine}}$. However, we present the best performing method

$$m^*_S(\alpha)$$

for every scenario $S \in \{1, \ldots, 8\}$ and $\alpha \in A$ in the following tables.

We distinguish between the best performing method regarding the mean integrated averaged tick losses (Table 7.6) on the one hand and the median of the corresponding averaged tick losses (Table 7.7) on the other hand. We introduce the following notation.

$$m^\text{MIA TL}_S^*(\alpha) \in M \coloneqq \arg \min_{m \in M} \text{RELMIA TL}_S^m(\alpha)$$

and

$$m^\text{MEDATL}_S^*(\alpha) \in M \coloneqq \arg \min_{m \in M} \text{REL MEDATL}_S^m(\alpha),$$
for all $S \in \{1, \ldots, 8\}$ and $\alpha \in \mathcal{A}$.

Note that we now introduce the notations $\text{RELM}IATL^m_S(\alpha)$ and $\text{RELMED}ATL^m_S(\alpha)$ in order to emphasize which scenario $S \in \{1, \ldots, 8\}$ we refer to.

<table>
<thead>
<tr>
<th>$S$</th>
<th>$\alpha$</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>RVQR</td>
<td>DVQR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>DVQR</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>DVQR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>DVQRP</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>RVQRDVQR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQPCOR</td>
<td>RVQRDVQR</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
</tr>
</tbody>
</table>

Table 7.6: Best performing method according to $\text{RELM}IATL_S(\alpha)$ for scenario $S \in \{1, \ldots, 8\}$, $\alpha \in \mathcal{A}$.

<table>
<thead>
<tr>
<th>$S$</th>
<th>$\alpha$</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>DVQR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>DVQR</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>DVQRP</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>DVQR</td>
<td>RVQRDVQR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRDVQR</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
<td>RVQRPCOR</td>
</tr>
</tbody>
</table>

Table 7.7: Best performing method according to $\text{RELMED}ATL_S(\alpha)$ for scenario $S \in \{1, \ldots, 8\}$, $\alpha \in \mathcal{A}$.

If we only considered $m^\text{MIATL}_S(\alpha)$, we could conclude that RVQRPCOR outperforms all other methods, as in scenarios $S = 4, \ldots, 8$, $m^\text{MIATL}_S(\alpha) = \text{RVQRPCOR}$ for every $\alpha \in \mathcal{A}$. However, considering Table 7.7 we observe that the distribution of the best performing method is not as stable as induced by Table 7.6.

A possible interpretation is that RVQRPCOR, on average, predicts the quantiles as accurate as for example methods $m = \text{DVQR}$, $\text{DVQRPCOR}$, but behaves considerably more stable than the other methods, i.e. there are less outliers in the prediction accuracy of RVQRPCOR than in the accuracy of the other methods. This hypothesis is also supported by the box plots presented in Section 7.3.2. Here, we observe that the box plot diagrams corresponding to ${\text{ATL}^m}_r$ for $r = 1, \ldots, 100$ are mostly less scattered than the others. According to Table 7.7, 26 out of 40 best performing methods (regarding the median of
the averaged tick losses) are based on the partial correlation approach (PCOR). We conclude that the selection of covariates by partial correlations (Section 5.4) performs well compared to the selection by the D-vine quantile regression algorithm (Section 5.2).

7.4 Required computational times

Finally, we have a look at the required computational times (Table 7.8) for every method \( m \in \mathcal{M} \) and \( S \in \{1, \ldots, 8\} \). The computational time measured for method \( m \) in scenario \( S \) and replication \( r \) in minutes is denoted by

\[ CT_{S}^{m,r}. \]

In our analysis, we consider the quantity \( MEDCT_{S}^{m} \), which is the median of all replications \( r = 1, \ldots, 100 \), i.e.

\[ MEDCT_{S}^{m} := \text{median}\{CT_{S}^{m,r}\}_{r=1,\ldots,100} \]

for all \( S \in \{1, \ldots, 8\} \), \( m \in \mathcal{M} \).

<table>
<thead>
<tr>
<th>( m )</th>
<th>( S )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>LQR</td>
<td>0.0002</td>
<td>0.0002</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0012</td>
<td>0.0012</td>
<td></td>
</tr>
<tr>
<td>LQRWSA</td>
<td>0.0075</td>
<td>0.0069</td>
<td>0.0103</td>
<td>0.0106</td>
<td>0.0823</td>
<td>0.0859</td>
<td>0.1301</td>
<td>0.1364</td>
<td></td>
</tr>
<tr>
<td>DVQR</td>
<td>1.2551</td>
<td>1.2365</td>
<td>2.5527</td>
<td>2.5786</td>
<td>6.2163</td>
<td>6.1649</td>
<td>17.5581</td>
<td>15.8775</td>
<td></td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.4080</td>
<td>0.4191</td>
<td>0.8152</td>
<td>0.8398</td>
<td>2.6092</td>
<td>2.6191</td>
<td>5.2633</td>
<td>5.3007</td>
<td></td>
</tr>
<tr>
<td>RVQRPCOR</td>
<td>0.5544</td>
<td>0.5532</td>
<td>1.0937</td>
<td>1.1066</td>
<td>3.5377</td>
<td>3.5471</td>
<td>7.1798</td>
<td>7.2630</td>
<td></td>
</tr>
<tr>
<td>RVQRKendall</td>
<td>0.2994</td>
<td>0.2641</td>
<td>0.8103</td>
<td>0.6531</td>
<td>0.3257</td>
<td>0.3070</td>
<td>0.8258</td>
<td>0.8458</td>
<td></td>
</tr>
<tr>
<td>RVQRsemi</td>
<td>0.3040</td>
<td>0.2821</td>
<td>0.6859</td>
<td>0.6204</td>
<td>0.3141</td>
<td>0.3001</td>
<td>0.7881</td>
<td>0.7830</td>
<td></td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>0.2994</td>
<td>0.2641</td>
<td>0.8103</td>
<td>0.6531</td>
<td>0.3257</td>
<td>0.3070</td>
<td>0.8258</td>
<td>0.8458</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.8: \( MEDCT_{S}^{m} \) for scenario \( S \in \{1, \ldots, 8\} \), \( m \in \mathcal{M} \).

As expected, the computational time increases in the number of (initial) covariates \( (d_{0}) \) and the training data set’s dimension \( (n_{\text{train}}) \). \( LQR \) and \( LQRWSA \) require considerably less computational effort than vine based methods \( (m \in \mathcal{M}_{\text{vine}}) \).

Note that the D-vine tree sequence chosen by \( DVQRPCOR \) is already determined by the order of covariate inclusion selected by the partial correlation approach.

Thus, \( DVQRPCOR \) only needs to estimate the pair copulas corresponding to the edges in the predefined trees. That is why, especially for large values of \( d_{0} \), \( DVQRPCOR \) requires significantly less computational effort than \( RVQRPCOR \), which chooses any of the regular vine tree sequence fulfilling the covariate order defined by the partial correlation approach. In contrast to \( RVQRPCOR \), methods \( RVQRKendall \), \( RVQRsemi \) and \( RVQRindepTest \) stop the covariate inclusion according to conditional log likelihood tests. Thus, they are always faster than \( RVQRPCOR \).

As methods \( DVQR \) and \( RVQRDVQR \) select the covariates by the D-vine quantile regression algorithm (Section 5.2), they require significantly more computational effort than...
the other methods because more pair copulas have to be estimated (see Section 5.3). Thus, the theoretical computational advantage of our forward and covariate selection approach is observed within the simulation study.
Chapter 8

Application to the abalone data set

In the following, we present a small data application to the abalone data set. It is available from the University of California Irvine (UCI) machine learning repository. Metadata can be obtained from [http://archive.ics.uci.edu/ml/datasets/Abalone](http://archive.ics.uci.edu/ml/datasets/Abalone). It is also available in the library [PivotalR](http://archive.ics.uci.edu/ml/datasets/Abalone).

8.1 Data preparation

The data set contains observations (no missing values) on \( n = 4177 \) abalones, \( i \in \{1, \ldots, 4177\} \), which are described by the following 9 characteristics:

- Sex / nominal / M (male), F (female) and I (infant),
- Length / continuous / mm / Longest shell measurement,
- Diameter / continuous / mm / perpendicular to length,
- Height / continuous / mm / with meat in shell,
- Whole weight / continuous / grams / whole abalone,
- Shucked weight / continuous / grams / weight of meat,
- Viscera weight / continuous / grams / gut weight (after bleeding),
- Shell weight / continuous / grams / after being dried,
- Rings / integer / +1.5 gives the age in years.

The observations of Sex are transformed to 1 for M, 0 for F and 2 for I. Thus, we are given the following variables and data vectors on the x-scale:

- \( X_{\text{sex}} = x_{\text{sex}} \in \{1,0,2\}^{4177} \)
- \( X_{\text{len}} = x_{\text{len}} \in [0, \infty)^{4177} \),
- \( X_{\text{dia}} = x_{\text{dia}} \in [0, \infty)^{4177} \),
- \( X_{\text{ring}} = x_{\text{ring}} \in \mathbb{Z}^{4177} \).
• $X_h = x_h \in [0, \infty)^{4177}$,
• $X_{\text{whole}} = x_{\text{whole}} \in [0, \infty)^{4177}$,
• $X_{\text{shuck}} = x_{\text{shuck}} \in [0, \infty)^{4177}$,
• $X_{\text{vis}} = x_{\text{vis}} \in [0, \infty)^{4177}$,
• $X_{\text{shell}} = x_{\text{shell}} \in [0, \infty)^{4177}$,
• $X_{\text{rings}} = x_{\text{rings}} \in \mathbb{N}^{4177}$,

where $x_j := (x_{1j}, \ldots, x_{4177j})^T$ for all $j \in J := \{\text{sex, len, dia, h, whole, shuck, vis, shell, rings}\}$.

Table 8.1 shows the first 6 entries $(x_{1j}, \ldots, x_{6j})^T$ of each data vector $x_j$.

<table>
<thead>
<tr>
<th>i</th>
<th>$x_{\text{sex}}$</th>
<th>$x_{\text{len}}$</th>
<th>$x_{\text{dia}}$</th>
<th>$x_h$</th>
<th>$x_{\text{whole}}$</th>
<th>$x_{\text{shuck}}$</th>
<th>$x_{\text{vis}}$</th>
<th>$x_{\text{shell}}$</th>
<th>$x_{\text{rings}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.4550</td>
<td>0.3650</td>
<td>0.0950</td>
<td>0.5140</td>
<td>0.2245</td>
<td>0.1010</td>
<td>0.1500</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.3500</td>
<td>0.2650</td>
<td>0.0900</td>
<td>0.2255</td>
<td>0.0995</td>
<td>0.0485</td>
<td>0.0700</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0.5300</td>
<td>0.4200</td>
<td>0.1350</td>
<td>0.6770</td>
<td>0.2565</td>
<td>0.1415</td>
<td>0.2100</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0.4400</td>
<td>0.3650</td>
<td>0.1250</td>
<td>0.5160</td>
<td>0.2155</td>
<td>0.1140</td>
<td>0.1550</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>0.3300</td>
<td>0.2550</td>
<td>0.0800</td>
<td>0.2050</td>
<td>0.0895</td>
<td>0.0395</td>
<td>0.0550</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>0.4250</td>
<td>0.3000</td>
<td>0.0950</td>
<td>0.3515</td>
<td>0.1410</td>
<td>0.0775</td>
<td>0.1200</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 8.1: First 6 observations of abalone data set.

In the following, we consider female abalones only, select $Y := X_{\text{whole}}$ to be the response variable and choose $X_1^p := X_{\text{shuck}}, X_2^p := X_{\text{vis}}$ and $X_3^p := X_{\text{shell}}$ to be the set of potential covariates, i.e. we want to predict the total weight of a female abalone given its shucked, viscera and shell weight.

Looking at the observations of $X_h$ (Figure 8.1) for the female abalones, we dismiss one abalone ($i = 2052$), since it has an outlying height.
Thus, denoting the abalone data set by
\[ \mathcal{X} := (\mathcal{X}_{ij}) \in [0, \infty)^{4177 \times 9}, \ i = 1, \ldots, 4177, \ j \in \mathcal{J}, \]
we consider the subset
\[ \mathcal{X}^F \in [0, \infty)^{1306 \times 4} := (\mathcal{X}_{ij}), \ i \in \{\ell \in \{1, \ldots, 4177\} \setminus \{2052\} : x_{\ell,\text{sex}} = 0\}, \ j \in \mathcal{J}^F, \]
where
\[ \mathcal{J}^F = \{\text{whole, shuck, vis, shell}\}. \]
It consists of \( n^F = 1306 \) abalones. The corresponding observation vectors are denoted by \( \mathbf{x}_j^F \) for \( j \in \mathcal{J}^F \).

Using the corresponding empirical distribution functions
\[ \hat{F}_{\mathcal{X}_{\text{whole}}}, \hat{F}_{\mathcal{X}_{\text{shuck}}}, \hat{F}_{\mathcal{X}_{\text{vis}}} \text{ and } \hat{F}_{\mathcal{X}_{\text{shell}}} \]
as marginal distributions, \( \mathcal{X}^F \) is transformed to the \textbf{u-scale}. The resulting copula data set is denoted by
\[ \mathcal{U}^F := (\mathbf{u}_{\text{whole}}, \mathbf{u}_{\text{shuck}}, \mathbf{u}_{\text{vis}}, \mathbf{u}_{\text{shell}}) \in [0, 1]^{1306 \times 4}, \]
where
\[ \mathbf{u}_j = (u_{1j}, \ldots, u_{1306j})^T \text{ and } u_{ij} = \hat{F}_{\mathcal{X}_j}(x_{ij}^F) \text{ for all } i = 1, \ldots, 1306 \text{ and } j \in \mathcal{J}^F. \]
For the response variable and the potential covariates on the u-scale, we introduce the following notation:

\[ \text{Whole} := \hat{F}_{X_{\text{whole}}} (X_{\text{whole}}) \]

and

\[ \text{Shuck} := F_{X_{\text{shuck}}} (X_{\text{shuck}}), \text{ Vis} := F_{X_{\text{vis}}} (X_{\text{vis}}) \text{ and } \text{Shell} := F_{X_{\text{shell}}} (X_{\text{shell}}). \]

According to Figure 8.2, which shows histograms of the data in \( U^F \), it is reasonable to assume that Whole, Shuck, Vis and Shell are marginally \( U(0,1) \)-distributed.

![Histograms of Whole, Shuck, Vis, Shell](image)

Figure 8.2: Histograms on the u-scale - Whole, Shuck, Vis, Shell.

### 8.2 Quantile regression of Whole

As already mentioned, we are interested in predicting the quantiles of a female abalone’s whole weight (Whole) given the potential covariates Shuck, Vis and Shell. First, we analyze all possible R-vine tree sequences for quantile regression regarding their in-sample conditional log likelihoods (Section 8.2.1). Besides identifying the \( c\ell\)-optimal R-vine structure, we investigate which of those sequences are selected by

- the D-vine copula based quantile regression algorithm with covariate selection according to conditional log likelihood tests (Section 5.2), denoted by \( DVQR \), and
• the R-vine copula based quantile regression algorithm with covariate selection by partial correlations (Section 5.4) and forward selection by Kendall’s τ (Section 5.5), denoted by RVQR.

In order to compare both algorithms regarding their out-of-sample prediction accuracy, we perform a cross-validation computing the corresponding out-of-sample averaged tick-losses in a second step. The resulting quantities are compared to the tick-losses induced by the \( c\ell \)-optimal R-vine structure.

8.2.1 In-sample \( c\ell \)-evaluation of all valid R-vine structures

Note that according to Table 5.1 in Section 5.1 there are \( NR(3) = 21 \) valid R-vine sequences for \( J = 3 \) potential covariates. Therefore, it is feasible to consider all possible R-vine tree sequences for quantile regression of Whole given Shuck, Vis and Shell and to identify the global \( c\ell \) optimal R-vine tree sequence as well as the R-vine tree sequences selected by DVQR and RVQR. Besides the conditional log likelihoods (c\( \ell \)), we discuss the number of parameters (\(|\Theta|\)) and the corresponding conditional AICs (cAIC) for all considered R-vine copulas.

The copulas are selected using the R function \texttt{RVineCopSelect()} from the VineCopula-package. Given an R-vine tree sequence \( V_d \), \( d = 1, 2, 3 \), the set of pair copula families \( \hat{\mathcal{B}}(V) \) and the set of pair copula parameters \( \hat{\Theta}(\hat{\mathcal{B}}(V)) \) are estimated, such that the corresponding R-vine copula is optimal in the corresponding log likelihood.

Figure 8.3 covers all R-vine tree sequences which can be used for quantile regression of Shuck given one or two covariates respectively. Since all R-vine tree sequences on two and three elements are D-vines, they solely differ in the order of covariate inclusion.
First, note that all D-vine copulas for \( d = 1 \) require the estimation of \( |\Theta| = 2 \) parameters and that the number of parameters is equal to 6 for all D-vines with \( d = 2 \). Thus, in \( d = 1, 2 \) each of the estimated pair copulas involves two parameters and we can focus on the analysis of the conditional log likelihoods, as the model complexity is the same for all copulas.

Both the D-vine quantile regression algorithm and the partial correlation based covariate selection (Section 5.4, Algorithm 1) result in the following order of covariates

\[ \text{Shuck} > \text{Shell} > \text{Vis}. \]

Thus, both the \( DVQR \) and \( RVQR \) select the D-vine structures

- \( \text{Whole} - \text{Shuck} \) (c\( \ell \ell \)\( ^1 \) = 1740.326) in iteration \( s = 1 \) and
- \( \text{Whole} - \text{Shuck} - \text{Shell} \) (c\( \ell \ell \)\( ^2 \) = 2703.439) in iteration \( s = 2 \).

Note that both algorithms do not stop after the first iteration, since the increase in the conditional log likelihood is great enough. Even if the number of parameters increases
by 4, which would be the maximal number of additional parameters, the conditional log likelihood test is rejected at a level $\alpha_{LRT} = 0.05$, because

$$c\ell_2^2 - c\ell_1^2 = 2703.439 - 1740.326 = 963.113 > 9.4877 = \chi^2_{0.95,4}.$$ 

We observe that both $DVQR$ and $RVQR$ select the $c\ell$-optimal R-vine structure in the first iteration. However, after the inclusion of the second covariate, the R-vine structure with the second highest conditional log likelihood is selected.

After the inclusion of the remaining third covariate, all valid R-vine tree sequences are either D-vines (Figure 8.4) or C-vines (Figure 8.5). The $c\ell$-optimal C-vine beats the $c\ell$-optimal D-vine regarding their conditional log likelihoods (optimal C-vine: $c\ell = 3110.848$, optimal D-vine: $c\ell = 3093.523$).

By identifying the $c\ell$-optimal R-vine sequence, we obtain the following conditional log likelihoods:

- **Whole** - **Shuck** - **Vis** - **Shell**
  - $c\ell = 3091.781$, $|\Theta| = 12$, cAIC = $-6159.562$
- **Whole** - **Shuck** - **Shell** - **Vis**
  - $c\ell = 3093.981$, $|\Theta| = 11$, cAIC = $-6165.962$

- **Whole** - **Vis** - **Shuck** - **Shell**
  - $c\ell = 2986.194$, $|\Theta| = 12$, cAIC = $-5948.389$
- **Whole** - **Vis** - **Shell** - **Shuck**
  - $c\ell = 3058.831$, $|\Theta| = 12$, cAIC = $-6093.662$

- **Whole** - **Shell** - **Shuck** - **Vis**
  - $c\ell = 3073.718$, $|\Theta| = 11$, cAIC = $-6125.435$
- **Whole** - **Shell** - **Vis** - **Shuck**
  - $c\ell = 3093.523$, $|\Theta| = 12$, cAIC = $-6163.047$

Figure 8.4: $c\ell$- evaluation of valid D-vine tree sequences including three covariates - for each sequence the figure covers $T^3_1$. 

133
Figure 8.5: $c_{\ell\ell}$- evaluation of valid C-vine tree sequences including three covariates - for each sequence the figure covers $T_1^3$ and $T_2^3$.

We observe that both $DVQR$ and $RVQR$ do not choose the global $c_{\ell\ell}$-optimal R-vine structure, which reaches the maximal conditional log likelihood of $c_{\ell\ell} = 3110.848$. The main reason is that the sequential approaches optimize the conditional log likelihoods in each iteration, which in general does not lead to the global $c_{\ell\ell}$-optimum.

Since $RVQR$ allows for every possible R-vine structure, the $RVQR$-selected sequence induces a littler higher conditional log likelihood than the sequence selected by $DVQR$ ($RVQR$: $c_{\ell\ell}^3 = 3094.447$, $DVQR$: $c_{\ell\ell}^3 = 3093.981$). However, the D-vine copula chosen by $DVQR$ requires the estimation of $|\Theta| = 11$ parameters, while the number of parameters is equal to 12 for the $RVQR$-selected copula. That is why we observe a little lower in-sample conditional AIC for the $DVQR$-selected D-vine ($RVQR$: cAIC$^3 = −6164.893$, $DVQR$: cAIC$^3 = −6165.962$). Regarding the conditional AICs of all D-vine copulas, $DVQR$ even chooses the cAIC-optimal D-vine structure.

Again, note that both methods do not stop according to the likelihood ratio test (Section 5.2) at level $\alpha_{LRT} = 0.05$, since

- for $DVQR$: $c_{\ell\ell}^3 - c_{\ell\ell}^2 = 3093.981 - 2703.439 = 390.542 > \chi^2_{0.95,6} = 12.592$ and
- for $RVQR$: $c_{\ell\ell}^3 - c_{\ell\ell}^2 = 3094.447 - 2703.439 = 391.008 > \chi^2_{0.95,6} = 12.592$. 

134
In the next section, we only consider the $c\ell$-optimal C-vine structure ($\mathcal{V}^*$), the C-vine chosen by $RVQR$ ($\mathcal{V}^{RVQR}$) and the D-vine selected by $DVQR$ ($\mathcal{V}^{DVQR}$) and analyze their out-of-sample performance in a cross-validation.

8.2.2 Out-of-sample quantile prediction

Given the R-vine structures $\mathcal{V}^*$, $\mathcal{V}^{RVQR}$ and $\mathcal{V}^{DVQR}$ we want to assess the prediction accuracy of the methods $DVQR$ and $RVQR$. Therefore, we perform an out-of-sample cross-validation and compute out-of-sample averaged tick losses, which have been introduced in Definition 7.1 (Section 7.3). Additionally, we discuss the corresponding out-of-sample conditional log likelihoods and AICs.

In order to cross validate the results, we split the female abalone data set $U^F$ into 5 equally sized parts. The abalones entering each of the 5 parts are chosen randomly. We end up with 5 independent data sets $U^F_1, \ldots, U^F_5$ with $\bigcup_{\ell=1}^{5} U^F_\ell = U^F$.

While one of the parts is chosen to be the evaluation data set $U_{eval}$, the remaining 4 parts are used to determine the R-vine copulas for quantile regression corresponding to the sequences $\mathcal{V}^*$, $\mathcal{V}^{RVQR}$ and $\mathcal{V}^{DVQR}$. That means for each $\ell \in \{1, \ldots, 5\}$ and $\mathcal{V} \in \{\mathcal{V}^*, \mathcal{V}^{RVQR}, \mathcal{V}^{DVQR}\}$ we estimate the set of pair copula families $\hat{\mathcal{B}}_\ell(\mathcal{V})$ and parameters $\hat{\Theta}_\ell(\hat{\mathcal{B}}_\ell(\mathcal{V}))$ by optimizing the log likelihood given the selected training data set (R-function $RVineCopSelect()$). Accordingly, we obtain three estimated R-vine copulas

- $\hat{\mathcal{R}}^*_\ell := (\mathcal{V}^*, \hat{\mathcal{B}}_\ell(\mathcal{V}^*), \hat{\Theta}_\ell(\hat{\mathcal{B}}_\ell(\mathcal{V}^*)))$,
- $\hat{\mathcal{R}}^{RVQR}_\ell := (\mathcal{V}^{RVQR}, \hat{\mathcal{B}}_\ell(\mathcal{V}^{RVQR}), \hat{\Theta}_\ell(\hat{\mathcal{B}}_\ell(\mathcal{V}^{RVQR})))$ and
- $\hat{\mathcal{R}}^{DVQR}_\ell := (\mathcal{V}^{DVQR}, \hat{\mathcal{B}}_\ell(\mathcal{V}^{DVQR}), \hat{\Theta}_\ell(\hat{\mathcal{B}}_\ell(\mathcal{V}^{DVQR})))$

for every $\ell \in \{1, \ldots, 5\}$.

For every $\ell \in \{1, \ldots, 5\}$ and $\mathcal{R} \in \{\hat{\mathcal{R}}^*_\ell, \hat{\mathcal{R}}^{RVQR}_\ell, \hat{\mathcal{R}}^{DVQR}_\ell\}$, we first discuss the implied out-of-sample conditional log likelihood denoted by

$\ell \ell^R_\ell := \ell \ell(\mathcal{R}, U^F_\ell)$,

the number of parameters to be estimated in the R-vine copula $\mathcal{R}$, denoted by

$|\Theta|^R_\ell := \begin{cases} |\hat{\Theta}_\ell(\hat{\mathcal{B}}_\ell(\mathcal{V}^{RVQR}))|, & \mathcal{R} = \hat{\mathcal{R}}^{RVQR}_\ell, \\ |\hat{\Theta}_\ell(\hat{\mathcal{B}}_\ell(\mathcal{V}^{DVQR}))|, & \mathcal{R} = \hat{\mathcal{R}}^{DVQR}_\ell, \\ |\hat{\Theta}_\ell(\hat{\mathcal{B}}_\ell(\mathcal{V}^*))|, & \mathcal{R} = \hat{\mathcal{R}}^*_\ell, \end{cases}$

and the corresponding out-of-sample conditional AIC

$cAIC^R_\ell := -2\ell \ell^R_\ell + 2|\Theta|^R_\ell$. 

135
As expected, the range of out-of-sample conditional log likelihoods lies far below the range of the in-sample $c_{\ell\ell}^R$s presented in the previous section. Except for $\ell = 2$ the in-sample $c_{\ell\ell}$-optimal structure $V^*$ induces a higher out-of-sample conditional log likelihood than the R-vines selected by $DVQR$ and $RVQR$. However, for $\ell = 3, 4, 5$, $\hat{R}_{\ell}^{DVQR}$ and $\hat{R}_{\ell}^{RVQR}$ less parameters to be estimated (11 or 10) than $\hat{R}_{\ell}^*$. That is why for $\ell = 3$ $\hat{R}_{\ell}^{RVQR}$ performs better than $\hat{R}_{\ell}^*$ regarding their conditional AICs.

Comparing the conditional log likelihoods induced by $DVQR$ and $RVQR$, we do not observe a great difference for $\ell = 2, \ldots, 5$. However, in $\ell = 1$ $DVQR$ performs considerably better than $RVQR$.

Keeping in mind the results on conditional log likelihoods and AICs, we assess the prediction accuracy of $DVQR$ and $RVQR$ compared to the $c_{\ell\ell}$-optimal copula. Using the
estimated conditional quantile functions corresponding to the vine copulas
\( \hat{R}_\ell^*, \hat{R}_{\ell}^{RVQR}, \hat{R}_{\ell}^{DVQR} \),

we compute the averaged tick losses
\[
ATL_{\ell}(\alpha), \ R \in \{\hat{R}_\ell^*, \hat{R}_{\ell}^{RVQR}, \hat{R}_{\ell}^{DVQR}\}
\]
for \( \ell \in \{1, \ldots, 5\} \) and \( \alpha \in \{0.05, 0.25, 0.5, 0.75, 0.95\} \).

The computed averaged tick losses (\( \times 100 \)) are provided in Table 8.3. The smallest (best) averaged tick losses for each parameter set \((\ell, \alpha)\) are marked bold.

<table>
<thead>
<tr>
<th>R</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{R}_1^* )</td>
<td>0.2210</td>
<td>0.6635</td>
<td>0.9204</td>
<td>0.8756</td>
<td>0.3728</td>
</tr>
<tr>
<td>( \hat{R}_1^{DVQR} )</td>
<td>0.2242</td>
<td>0.7157</td>
<td>0.9654</td>
<td>0.8765</td>
<td>0.3309</td>
</tr>
<tr>
<td>( \hat{R}_1^{RVQR} )</td>
<td>0.2366</td>
<td>0.7403</td>
<td>0.9766</td>
<td>0.8794</td>
<td>0.3351</td>
</tr>
<tr>
<td>( \hat{R}_2^* )</td>
<td>0.4145</td>
<td>0.9713</td>
<td>1.1565</td>
<td>0.9734</td>
<td>0.3662</td>
</tr>
<tr>
<td>( \hat{R}_2^{DVQR} )</td>
<td>0.3970</td>
<td>0.9861</td>
<td>1.3403</td>
<td>1.2777</td>
<td>0.6728</td>
</tr>
<tr>
<td>( \hat{R}_2^{RVQR} )</td>
<td>0.3972</td>
<td>0.9983</td>
<td>1.3541</td>
<td>1.2869</td>
<td>0.6696</td>
</tr>
<tr>
<td>( \hat{R}_3^* )</td>
<td>0.6422</td>
<td>1.0910</td>
<td>1.1872</td>
<td>0.9123</td>
<td>0.3525</td>
</tr>
<tr>
<td>( \hat{R}_3^{DVQR} )</td>
<td>0.3403</td>
<td>0.8152</td>
<td>1.0781</td>
<td>0.9540</td>
<td>0.3455</td>
</tr>
<tr>
<td>( \hat{R}_3^{RVQR} )</td>
<td>0.3429</td>
<td>0.8123</td>
<td>1.0771</td>
<td>0.9516</td>
<td>0.3455</td>
</tr>
<tr>
<td>( \hat{R}_4^* )</td>
<td>0.3642</td>
<td>0.7866</td>
<td>1.0118</td>
<td>0.9664</td>
<td>0.5048</td>
</tr>
<tr>
<td>( \hat{R}_4^{DVQR} )</td>
<td>0.3666</td>
<td>0.8183</td>
<td>1.1128</td>
<td>1.0654</td>
<td>0.5358</td>
</tr>
<tr>
<td>( \hat{R}_4^{RVQR} )</td>
<td>0.3680</td>
<td>0.8231</td>
<td>1.1206</td>
<td>1.0703</td>
<td>0.5407</td>
</tr>
<tr>
<td>( \hat{R}_5^* )</td>
<td>0.2829</td>
<td>0.7546</td>
<td>0.9812</td>
<td>0.9129</td>
<td>0.4410</td>
</tr>
<tr>
<td>( \hat{R}_5^{DVQR} )</td>
<td>0.2702</td>
<td>0.7499</td>
<td>1.0375</td>
<td>0.9664</td>
<td>0.4277</td>
</tr>
<tr>
<td>( \hat{R}_5^{RVQR} )</td>
<td>0.2755</td>
<td>0.7928</td>
<td>1.0743</td>
<td>0.9820</td>
<td>0.4181</td>
</tr>
</tbody>
</table>

Table 8.3: Averaged tick losses \( ATL_{\ell}(\alpha) \times 100 \) for \( \alpha \in \{0.05, 0.25, 0.5, 0.75, 0.95\} \), \( R \in \{\hat{R}_\ell^*, \hat{R}_{\ell}^{RVQR}, \hat{R}_{\ell}^{DVQR}\} \) and \( \ell \in \{1, \ldots, 5\} \).

Except for the outlying averaged tick losses in the parameter sets \((\ell, \alpha) = (2, 0.95)\) and
\((\ell, \alpha) = (3, 0.05)\), the values lie close to each other. This observation is supported by the small differences in the in-sample conditional log likelihoods determined in the previous section. However, we do observe that the \(c_{\ell\ell}\)-optimal R-vine sequence \(V^*\) induces the smallest averaged tick loss for most of the parameter sets.

Summarizing our results, we conclude that none of the R-vine tree sequences \(V^*, V^{RVQR}\) and \(V^{DVQR}\) outperforms the others by a great margin. Therefore, even for small sets of potential covariates, R-vine and D-vine based quantile regression are good alternatives to identifying the global \(c_{\ell\ell}\)-optimal copula.
Part IV

Conclusion
In today’s society, decisions have been increasingly based on collected data. Companies, governments, and individuals have been trying to forecast the behavior of a dependent variable based on potentially influential observations. That is why multiple statistical and data mining approaches have been developed aiming to predict the outcome of a response variable given a set of covariates. Especially since the financial crisis, it was no longer enough to predict the mean behavior of a response. For example, insurance companies are required to predict the 0.995-quantile of their total loss in order to reduce the risk of insolvency. Quantile regression, which estimates the conditional quantile function of a response given observations on a set of covariates, addresses precisely this kind of tasks.

In this thesis we developed a quantile regression method, which derives the conditional quantile function of a response variable from an estimated regular vine copula. As opposed to the class of D-vines, the class of regular vine copulas growth exponentially with the number of considered covariates. Thus, a direct extension of the D-vine based quantile regression algorithm (Kraus and Czado [2017]), which sequentially includes covariates and simultaneously selects the D-vine structure based on conditional likelihood arguments, would be computationally impracticable. Therefore, we proposed an approach, which first determines the order of covariate inclusion by partial correlation arguments and sequentially estimates an optimal R-vine copula afterwards.

The order of covariate inclusion is chosen such that the Pearson correlation between the first covariate and the response is maximal and partial correlations are maximized in further iterations. The approach was justified by considering the conditional likelihood per observation for normally distributed random variables. For fixed partial correlations between the response and the previously chosen covariates, it is monotonically increasing with respect to the corresponding partial correlations.

After selecting the covariate order, the goal was to choose an R-vine copula that allows to construct the conditional density of the response and represents the selected covariate order at the same time. This led us to define the class of R-vine copulas for quantile regression. It contains all R-vine tree sequences, which are restricted by the response variable being an leaf in every tree and sharing the conditioned set with the k-th covariate in tree k. Even though the additional restrictions decrease the number of vine candidates, the set of valid R-vine tree sequences for quantile regression still gets large for small numbers of covariates. Therefore, we introduced a so called weight criteria which enables the algorithm to sequentially select the underlying R-vine structure without exceeding a decent computational effort.
Furthermore, we conducted a simulation study in order to compare our approach to D-vine copula based and linear quantile regression. Therefore, we simulated initial copula data from a sampled R-vine copula and assessed the goodness of fit based on in-sample conditional log likelihoods. Additionally we compared the methods’ prediction accuracy based on out-of-sample averaged tick losses. Outperforming linear quantile regression by a great margin, our method predicted the conditional quantiles as accurate as the D-vine based quantile algorithm. However, it fulfilled the expected computational improvement, as especially for a large number of covariates it required considerably less time than the D-vine based algorithm.

In order to compare our R-vine copula based algorithm and the D-vine copula based method to the R-vine copula inducing the global conditional likelihood maximum, we applied both methods to a small data set on abalones. As we predicted the quantiles of an abalone’s whole weight based on three covariates only, it was feasible to consider all possible R-vine structures and to determine the global optimum. Since none of the two methods was considerably outperformed, it seems that they constitute a good alternative to finding the global conditional likelihood optimum. Admittedly, this might differ for another data set or a larger number of covariate and certainly our approach can be improved. For example, one might try to select the global conditional likelihood optimum up to a decent number (4 or 5) of included covariates and extend it by applying the proposed methods afterwards.
List of Symbols

\[ h_{1|2}(u_1|u_2) \] h-function corresponding to \( U_1|U_2 \)

\[ h_{2|1}(u_2|u_1) \] h-function corresponding to \( U_2|U_1 \)

\( \rho_{X_1,X_2} \) Pearson correlation between \( X_1 \) and \( X_2 \)

\( \hat{\rho}_{X_1,X_2} \) Empirical Pearson correlation between \( X_1 \) and \( X_2 \)

\( \tau_{X_1,X_2} \) Kendall’s \( \tau \) between \( X_1 \) and \( X_2 \)

\( \hat{\tau}_{X_1,X_2} \) Empirical Kendall’s \( \tau \) between \( X_1 \) and \( X_2 \)

\( \rho_{U_1,U_2}^{\text{max}} \) Maximal semi-correlation between \( U_1 \) and \( U_2 \)

\( \hat{\rho}_{U_1,U_2}^{\text{max}} \) Maximal empirical semi-correlation between \( U_1 \) and \( U_2 \)

\( \rho_{Z_i,Z_j;Z_{id}^{\leq (i,j)}} \) Partial correlation between \( Z_1 \) and \( Z_2 \) given \( \rho_{Z_i,Z_j;Z_{id}^{\leq (i,j)}} \)

\( b \) Family of \( C^b(\cdot, \cdot; \theta) \)

\( \theta \) Parameter(s) of \( C^b(\cdot, \cdot; \theta) \)

\( C^b(\cdot, \cdot; \theta) \) Bivariate parametric copula

\( y \) Vector containing observations of \( Y \)

\( x_j \) Vector containing observations of \( X_j \)

\( Y \) Random variable on the x-scale (usually response)

\( X_j \) Random variable on the x-scale (usually covariate)

\( v \) Vector containing observations of \( V \)

\( u_j \) Vector containing observations of \( U_j \)

\( V \) Random variable on the u-scale (usually response)

\( U_j \) Random variable on the u-scale (usually covariate)

\( w \) Vector containing observations of \( W \)
\[ z_j \] Vector containing observations of \( Z_j \)

\[ W \] Random variable on the z-scale (usually response)

\[ Z_j \] Random variable on the z-scale (usually covariate)

\[ \ell(\mathcal{C}^b(\cdot, \cdot; \theta), \mathcal{U}^n) \] Log likelihood of \( \mathcal{C}^b(\cdot, \cdot; \theta) \) given the bivariate u-data \( \mathcal{U}^n \)

\[ \text{AIC}(\mathcal{C}^b(\cdot, \cdot; \theta), \mathcal{U}^n) \] AIC of \( \mathcal{C}^b(\cdot, \cdot; \theta) \) given the bivariate u-data \( \mathcal{U}^n \)

\[ \text{BIC}(\mathcal{C}^b(\cdot, \cdot; \theta), \mathcal{U}^n) \] BIC of \( \mathcal{C}^b(\cdot, \cdot; \theta) \) given the bivariate u-data \( \mathcal{U}^n \)

\[ \text{sc} \] Selection criterion (e.g. Log Likelihood, AIC, BIC)

\[ B \] Set of possible copula families

\[ \text{optC}(\mathcal{U}^n, B, \text{sc}) \] sc-optimal bivariate parametric copula

\[ \text{indepTest} \] Independence test prior to selection of sc-optimal copula?

\[ G \] Graph

\[ e \] Edge in graph or tree

\[ v \] Node in graph or tree

\[ T \] Tree

\[ \mathcal{V} \] Regular vine tree sequence

\[ T_k \] k-th tree in \( \mathcal{V} \)

\[ N_k \] Node set of tree \( T_k \)

\[ E_k \] Edge set of tree \( T_k \)

\[ A_e \] Complete union of edge \( e \)

\[ D_e \] Conditioning set of edge \( e \)

\[ P_e \] Conditioned set of edge \( e \)

\[ \text{pt}_k^\ell \] Partner of \( \ell \) (R-vine leaf) in tree \( T_k \)

\[ (\mathcal{F}, \mathcal{V}, \mathcal{C}(\mathcal{V})) \] R-vine distribution

\[ \mathcal{F} \] Collection of marginals within \( (\mathcal{F}, \mathcal{V}, \mathcal{C}(\mathcal{V})) \)

\[ \mathcal{C}(\mathcal{V}) \] Set of all pair copulas within \( (\mathcal{F}, \mathcal{V}, \mathcal{C}(\mathcal{V})) \)

\[ \mathcal{C}^{b_e}(\cdot, \cdot; \theta_e) \] Parametric pair copula corresponding to edge \( e \)

\[ h_{P_{e,1}\mid P_{e,2};D_e}(w_1|w_2) \] 1st h-function corresponding to \( e = P_e|D_e \)
\( h_{P_e | P_e,1: D_e} (w_2 | w_1) \) 2nd h-function corresponding to \( e = P_e | D_e \)

\( B(V) \) Ordered set of pair copula families within \( R \)

\( \Theta(B(V)) \) Ordered set of pair copula parameters within \( R \)

\( R \) R-vine copula

\( \hat{h}_{P_e | P_e,1: D_e} \) 1st pseudo observation vector corresponding to \( e = P_e | D_e \)

\( \hat{h}_{P_e | P_e,1: D_e} \) 2nd pseudo observation vector corresponding to \( e = P_e | D_e \)

\( \mathcal{P}_k \) Set of pseudo observations corresponding to edges in \( E_k \)

\( q_X(\alpha) \) Quantile function or \( \alpha \)-quantile of \( X \)

\( Q_\mathcal{Y}(\mathcal{X}; \alpha) \) Conditional quantile function of \( Y \) given \( \mathcal{X} \)

\( \rho_\alpha(x) \) Check loss function for \( \alpha \in (0, 1) \)

\( C_{\mathcal{V}|U_1,...,U_d}(v|u_1,...,u_d) \) Conditional distribution of \( V \) given \( U_j = u_j \)

\( \mathcal{R}^d \) R-vine copula for quantile regression of \( V \) given \( U_1,\ldots,U_d \)

\( \mathcal{T}^d_{k} \) \( k \)-th tree within \( \mathcal{V}^d \), \( k = 1,\ldots,d \)

\( N^d_k \) Node set of \( \mathcal{T}^d_{k} \)

\( E^d_k \) Edge set of \( \mathcal{T}^d_{k} \)

\( |\Theta^d| \) Number of parameters within \( \mathcal{R}^d \)

\( \text{c\ell}(\mathcal{R}^d, v, (u_1,\ldots,u_d)) \) Conditional log likelihood of \( \mathcal{R}^d \) given \( v, u_j, j = 1,\ldots,d \)

\( \text{cAIC}(\mathcal{R}^d, v, (u_1,\ldots,u_d)) \) Conditional AIC of \( \mathcal{R}^d \) given \( v, u_j, j = 1,\ldots,d \)

\( \text{cBIC}(\mathcal{R}^d, v, (u_1,\ldots,u_d)) \) Conditional BIC of \( \mathcal{R}^d \) given \( v, u_j, j = 1,\ldots,d \)

\( u_j^p \) Vector containing observations on \( U_j^p \)

\( U_j^p \) \( j \)-th potential covariate, \( j = 1,\ldots,J \)

\( \text{wc} \) Weight criterion - usually \( \text{wc} \in \{\tau, \rho^{\max} \} \)

\( \text{PT}^U_k \) Set of possible partners of \( U_s \) in \( T^s_k \)

\( \pi^k_s(U) \) Permutation vector \( \in \{1,\ldots,s-1\}^{k-1} \)
\textbf{maxW}_k(U, PT_k, wc) \text{ Weight criterion maximizer in tree } T_k^* \text{ \cite{72}}

e^*_k \text{ New edge in the } k\text{-th tree of the extended R-vine } \mathcal{V}^* \text{ \cite{74}}

r \text{ One of 100 replications in the simulation study } \text{ \cite{98}}

S \text{ One of 5 scenarios in the simulation study } \text{ \cite{98}}

d^S_0 \text{ Number of covariates in the initial data set simulated in scenario } S \text{ \cite{101}}

n_{\text{train}}^S \text{ Sample size of the initial training data sets simulated in scenario } S \text{ \cite{101}}

n_{\text{eval}}^S \text{ Sample size of the initial evaluation data sets simulated in scenario } S \text{ \cite{101}}

\textbf{indep}^S \text{ Dummy variable indicating if independence copulas were considered in scenario } S \text{ \cite{101}}

\textbf{LQR} \text{ Method: linear quantile regression } \text{ \cite{101}}

\textbf{LQRWSA} \text{ Method: linear quantile regression with step applied } \text{ \cite{101}}

m \text{ Method considered in the simulation study } \text{ \cite{102}}

\textbf{DVQR} \text{ Method: D-vine copula based quantile regression } \text{ \cite{102}}

\textbf{DVQRPCOR} \text{ Method: DVQR with fixed covariates selected by (partial) correlations } \text{ \cite{102}}

\textbf{RVQRDVQR} \text{ Method: RVQR with fixed covariates selected by DVQR } \text{ \cite{102}}

\textbf{RVQRPCOR} \text{ Method: RVQR with fixed covariates selected by (partial) correlations } \text{ \cite{102}}

\textbf{RVQRkendall} \text{ Method: RVQRPCOR stopped according to } c\ell \text{ tests and } wc = \tau \text{ \cite{102}}

\textbf{RVQRsemi} \text{ Method: RVQRPCOR stopped according to } c\ell \text{ tests and } wc = \rho^{\text{max}} \text{ \cite{102}}

\textbf{RVQRindepTest} \text{ Method: RVQRkendall with indepTest } = 1 \text{ \cite{102}}

\textbf{ATL}_m^r(\alpha) \text{ Averaged tick loss for } \alpha \in (0, 1) \text{ of method } m \text{ in replication } r \text{ \cite{112}}

\textbf{MEDATL}_m^r(\alpha) \text{ Median (of all replications) of the averaged tick losses for } \alpha \in (0, 1) \text{ of method } m \text{ \cite{112}}

\textbf{MIATL}_m^r(\alpha) \text{ Mean (of all replications) of the averaged tick losses for } \alpha \in (0, 1) \text{ of method } m \text{ \cite{112}}

\textbf{REL}_m \text{ \cite{112}}

\textbf{MEDATL}_m^r(\alpha) \text{ \cite{112}}

\textbf{RELMIATL}_m^r(\alpha) \text{ \cite{112}}

N_m^{C\text{-vine}} \text{ Number of replications in which method } m \text{ chooses a C-vine } \text{ \cite{112}}
$N_{D\text{-vine}}^m$  Number of replications in which method $m$ chooses a D-vine

$N_{C/D\text{-vine}}^m$  Number of replications in which method $m$ chooses C/D-vine
Bibliography


Oswaldo Morales Napoles, Roger M Cooke, and Dorota Kurowicka. About the number of vines and regular vines on n nodes. 2010.


Part V

Appendix
A.1 Simulation Study

A.1.1 Seeds

Before an initial R-vine copula was sampled in R, a seed was set using the R-function `set.seed()`. Initially, they were chosen equal to \( r \) for all replications \( r \in \{1, \ldots, 100\} \). Since some of the seeds generated errors, other seeds were chosen in this cases. For those replications the seeds are listed in Table 4.

<table>
<thead>
<tr>
<th>initial seed (= ( r ))</th>
<th>selected seed</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>101</td>
</tr>
<tr>
<td>37</td>
<td>102</td>
</tr>
<tr>
<td>44</td>
<td>103</td>
</tr>
<tr>
<td>51</td>
<td>104</td>
</tr>
<tr>
<td>60</td>
<td>105</td>
</tr>
<tr>
<td>90</td>
<td>106</td>
</tr>
<tr>
<td>91</td>
<td>107</td>
</tr>
</tbody>
</table>

Table 4: Overview of exchanged seeds.

Thus, the replications are based on the seeds as shown in Table 5.

<table>
<thead>
<tr>
<th>( r )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>seed</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>( r )</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>seed</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
</tr>
<tr>
<td>( r )</td>
<td>31</td>
<td>32</td>
<td>33</td>
<td>34</td>
<td>35</td>
<td>36</td>
<td>37</td>
<td>38</td>
<td>39</td>
<td>40</td>
</tr>
<tr>
<td>seed</td>
<td>41</td>
<td>42</td>
<td>43</td>
<td>44</td>
<td>45</td>
<td>46</td>
<td>47</td>
<td>48</td>
<td>49</td>
<td>50</td>
</tr>
<tr>
<td>( r )</td>
<td>51</td>
<td>52</td>
<td>53</td>
<td>54</td>
<td>55</td>
<td>56</td>
<td>57</td>
<td>58</td>
<td>59</td>
<td>60</td>
</tr>
<tr>
<td>seed</td>
<td>104</td>
<td>52</td>
<td>53</td>
<td>54</td>
<td>55</td>
<td>56</td>
<td>57</td>
<td>58</td>
<td>59</td>
<td>105</td>
</tr>
<tr>
<td>( r )</td>
<td>61</td>
<td>62</td>
<td>63</td>
<td>64</td>
<td>65</td>
<td>66</td>
<td>67</td>
<td>68</td>
<td>69</td>
<td>70</td>
</tr>
<tr>
<td>seed</td>
<td>61</td>
<td>62</td>
<td>63</td>
<td>64</td>
<td>65</td>
<td>66</td>
<td>67</td>
<td>68</td>
<td>69</td>
<td>70</td>
</tr>
<tr>
<td>( r )</td>
<td>71</td>
<td>72</td>
<td>73</td>
<td>74</td>
<td>75</td>
<td>76</td>
<td>77</td>
<td>78</td>
<td>79</td>
<td>80</td>
</tr>
<tr>
<td>seed</td>
<td>71</td>
<td>72</td>
<td>73</td>
<td>74</td>
<td>75</td>
<td>76</td>
<td>77</td>
<td>78</td>
<td>79</td>
<td>80</td>
</tr>
<tr>
<td>( r )</td>
<td>81</td>
<td>82</td>
<td>83</td>
<td>84</td>
<td>85</td>
<td>86</td>
<td>87</td>
<td>88</td>
<td>89</td>
<td>90</td>
</tr>
<tr>
<td>seed</td>
<td>81</td>
<td>82</td>
<td>83</td>
<td>84</td>
<td>85</td>
<td>86</td>
<td>87</td>
<td>88</td>
<td>89</td>
<td>106</td>
</tr>
<tr>
<td>( r )</td>
<td>91</td>
<td>92</td>
<td>93</td>
<td>94</td>
<td>95</td>
<td>96</td>
<td>97</td>
<td>98</td>
<td>99</td>
<td>100</td>
</tr>
<tr>
<td>seed</td>
<td>107</td>
<td>92</td>
<td>93</td>
<td>94</td>
<td>95</td>
<td>96</td>
<td>97</td>
<td>98</td>
<td>99</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 5: Overview of seeds for replications \( r = 1, \ldots, 100 \).
A.1.2 Tree plots of minimal - and maximal - $c_{ll}$ R-vine copulas in scenarios 1 to 4
Figure 6: Trees of order 1-4 corresponding to the initial (true) copula $R_{13}$ and the $RVQRPCOR$-estimated copula $R_{13}^{RVQRPCOR}$ in scenario $S = 1$ and replication $r = 3$ (pair copula families and corresponding Kendall’s $\tau$ values are shown on the edges).
Figure 7: Trees of order 1-4 corresponding to the initial (true) copula $R_{45}^1$ and the $RVQRPCOR$-estimated copula $\hat{R}_{45}^{RVQRPCOR}$ in scenario $S = 1$ and replication $r = 45$ (pair copula families and corresponding Kendall’s $\tau$ values are shown on the edges).
Figure 8: Trees of order 1-4 corresponding to the initial (true) copula \( \hat{R}_{236} \) and the RVQPCOR-estimated copula \( \hat{R}_{236} \) in scenario \( S = 2 \) and replication \( r = 36 \). (Pair copula families and corresponding Kendall's \( \tau \) values are shown on the edges).
Figure 9: Trees of order 1-4 corresponding to the initial (true) copula $\mathcal{R}^2_{54}$ and the $RVQRPCOR$-estimated copula $\hat{\mathcal{R}}^2_{RVQRPCOR}$ in scenario $S = 2$ and replication $r = 54$ (pair copula families and corresponding Kendall's $\tau$ values are shown on the edges).
Figure 10: Trees of order 1-4 corresponding to the initial (true) copula $\mathcal{R}^3_8$ and the $RVQRPCOR$-estimated copula $\mathcal{R}^3_{8RVQRPCOR}$ in scenario $S = 3$ and replication $r = 8$ (pair copula families and corresponding Kendall’s $\tau$ values are shown on the edges).
Figure 11: Trees of order 1-4 corresponding to the initial (true) copula $\mathcal{K}_3^{345}$ and the $RVQRPCOR$-estimated copula $\hat{\mathcal{K}}_3^{345}$ in scenario $S = 3$ and replication $r = 45$ (pair copula families and corresponding Kendall’s $\tau$ values are shown on the edges).
Figure 12: Trees of order 1-4 corresponding to the initial (true) copula $\mathcal{R}^4_{36}$ and the RVQRPCOR-estimated copula $\hat{\mathcal{R}}^4_{36}$ in scenario $S = 4$ and replication $r = 36$ (pair copula families and corresponding Kendall’s $\tau$ values are shown on the edges).
Figure 13: Trees of order 1-4 corresponding to the initial (true) copula $R_{45}^4$ and the RVQRPCOR-estimated copula $\hat{R}_{45}^4$ in scenario $S = 4$ and replication $r = 45$ (pair copula families and corresponding Kendall’s $\tau$ values are shown on the edges).
A.1.3 Out-of-sample results

Results for scenarios 2, 3, 4 and 6

Scenario 2

Figure 14: Scenario 2: box plots of \( \{\text{ATL}^m_r(\alpha)\}_{r=1,\ldots,100} \) for all \( m \in M \) and \( \alpha \in A \).

\[
\begin{array}{cccccc}
\text{\( m \)} & \text{0.05} & \text{0.25} & \text{0.50} & \text{0.75} & \text{0.95} \\
\hline
\text{RELMIATL}(\alpha) \\
LQR & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
LQRWSA & 1.0010 & 1.0003 & 0.9996 & 1.0005 & 1.0004 \\
DVQR & 0.8262 & 0.8607 & 0.8725 & 0.8638 & 0.8365 \\
DVQRPCOR & 0.8312 & 0.8598 & 0.8692 & 0.8587 & 0.8356 \\
RVQR & 0.8288 & 0.8568 & 0.8650 & 0.8548 & 0.8382 \\
RVQRPCOR & 0.8474 & 0.8862 & 0.8966 & 0.8889 & 0.8630 \\
RVQRemi & 0.8461 & 0.8882 & 0.9007 & 0.8931 & 0.8641 \\
RVQRindepTest & 0.8454 & 0.8846 & 0.8954 & 0.8872 & 0.8615 \\
\text{RELMEDATL}(\alpha) \\
LQR & 1.0000 & 1.0000 & 1.0000 & 1.0000 & 1.0000 \\
LQRWSA & 0.9983 & 1.0005 & 1.0006 & 1.0000 & 0.9977 \\
DVQR & 0.8215 & 0.9222 & 0.9120 & 0.9209 & 0.7816 \\
DVQRPCOR & 0.8144 & 0.9007 & 0.9150 & 0.8900 & 0.8071 \\
RVQR & 0.8499 & 0.9093 & 0.9166 & 0.8930 & 0.8055 \\
RVQRPCOR & 0.8057 & 0.8861 & 0.9060 & 0.8854 & 0.7860 \\
RVQRRemi & 0.8248 & 0.9120 & 0.9088 & 0.8893 & 0.8224 \\
RVQRindepTest & 0.8276 & 0.9131 & 0.9037 & 0.8823 & 0.8118 \\
\end{array}
\]

Table 6: Scenario 2: \( \frac{\text{MIATL}^m(\alpha)}{\text{MIATL}^{\text{LQR}}(\alpha)} \) and \( \frac{\text{MEDATL}^m(\alpha)}{\text{MEDATL}^{\text{LQR}}(\alpha)} \) for \( m \in M \) and \( \alpha \in A \).
Figure 15: Scenario 2: bar plot of $N_{D-vine}^m$, $N_{C-vine}^m$ and $N_{C/D-vine}^m$ for $m \in \mathcal{M}_{vine}$. 
Scenario 3

Figure 16: Scenario 3: box plots of $\{\text{ATL}_r^m(\alpha)\}_{r=1,\ldots,100}$ for all $m \in \mathcal{M}$ and $\alpha \in \mathcal{A}$.

Table 7: Scenario 3: $\frac{\text{MIATL}^m(\alpha)}{\text{MIATL}^\text{LQR}(\alpha)}$ and $\frac{\text{MEDATL}^m(\alpha)}{\text{MEDATL}^\text{LQR}(\alpha)}$ for $m \in \mathcal{M}$ and $\alpha \in \mathcal{A}$. 

<table>
<thead>
<tr>
<th>$m$</th>
<th>$\alpha$</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>LQR</td>
<td>RELMIATL(\alpha)</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>LQRWSA</td>
<td>1.0003</td>
<td>1.0002</td>
<td>1.0003</td>
<td>1.0002</td>
<td>0.9997</td>
<td></td>
</tr>
<tr>
<td>DVQR</td>
<td>0.8215</td>
<td>0.8183</td>
<td>0.8632</td>
<td>0.8481</td>
<td>0.8205</td>
<td></td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.8246</td>
<td>0.8509</td>
<td>0.8622</td>
<td>0.8476</td>
<td>0.8213</td>
<td></td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>0.8183</td>
<td>0.8479</td>
<td>0.8595</td>
<td>0.8445</td>
<td>0.8176</td>
<td></td>
</tr>
<tr>
<td>RVQRPCOR</td>
<td>0.8196</td>
<td>0.8483</td>
<td>0.8588</td>
<td>0.8447</td>
<td>0.8173</td>
<td></td>
</tr>
<tr>
<td>RVQRkendall</td>
<td>0.8324</td>
<td>0.8591</td>
<td>0.8696</td>
<td>0.8578</td>
<td>0.8325</td>
<td></td>
</tr>
<tr>
<td>RVQRsemi</td>
<td>0.8360</td>
<td>0.8655</td>
<td>0.8774</td>
<td>0.8644</td>
<td>0.8363</td>
<td></td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>0.8348</td>
<td>0.8601</td>
<td>0.8699</td>
<td>0.8581</td>
<td>0.8341</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$m$</th>
<th>RELMEDATL(\alpha)</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>LQR</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>LQRWSA</td>
<td>0.9989</td>
<td>1.0004</td>
<td>0.9990</td>
<td>1.0006</td>
<td>1.0031</td>
<td></td>
</tr>
<tr>
<td>DVQR</td>
<td>0.7980</td>
<td>0.8633</td>
<td>0.8664</td>
<td>0.8681</td>
<td>0.8146</td>
<td></td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.8067</td>
<td>0.8435</td>
<td>0.8571</td>
<td>0.8722</td>
<td>0.8297</td>
<td></td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>0.8101</td>
<td>0.8399</td>
<td>0.8510</td>
<td>0.8623</td>
<td>0.7953</td>
<td></td>
</tr>
<tr>
<td>RVQRPCOR</td>
<td>0.8196</td>
<td>0.8429</td>
<td>0.8502</td>
<td>0.8648</td>
<td>0.8107</td>
<td></td>
</tr>
<tr>
<td>RVQRkendall</td>
<td>0.8263</td>
<td>0.8431</td>
<td>0.8694</td>
<td>0.8680</td>
<td>0.8161</td>
<td></td>
</tr>
<tr>
<td>RVQRsemi</td>
<td>0.8445</td>
<td>0.8460</td>
<td>0.8605</td>
<td>0.8719</td>
<td>0.8219</td>
<td></td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>0.8421</td>
<td>0.8437</td>
<td>0.8656</td>
<td>0.8687</td>
<td>0.8143</td>
<td></td>
</tr>
</tbody>
</table>
Figure 17: Scenario 3: bar plot of $N_{D\text{-vine}}^m$, $N_{C\text{-vine}}^m$ and $N_{C/D\text{-vine}}^m$ for $m \in \mathcal{M}_{\text{vine}}$. 
Figure 18: Scenario 4: box plots of $\{\text{ATL}^m_r(\alpha)\}_{r=1,\ldots,100}$ for all $m \in \mathcal{M}$ and $\alpha \in \mathcal{A}$.

Table 8: Scenario 4: $\frac{\text{MIATL}^m_r(\alpha)}{\text{MIATL}^{LQR}_r(\alpha)}$ and $\frac{\text{MEDATL}^m_r(\alpha)}{\text{MEDATL}^{LQR}_r(\alpha)}$ for $m \in \mathcal{M}$ and $\alpha \in \mathcal{A}$. 

<table>
<thead>
<tr>
<th>$\alpha$</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>RELMIATL $^m_r(\alpha)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LQR</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>LQRWSA</td>
<td>1.0000</td>
<td>1.0002</td>
<td>1.0001</td>
<td>1.0002</td>
<td>0.9997</td>
</tr>
<tr>
<td>DVQR</td>
<td>0.8382</td>
<td>0.8599</td>
<td>0.8676</td>
<td>0.8543</td>
<td>0.8358</td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.8380</td>
<td>0.8546</td>
<td>0.8628</td>
<td>0.8508</td>
<td>0.8349</td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>0.8359</td>
<td>0.8584</td>
<td>0.8647</td>
<td>0.8521</td>
<td>0.8318</td>
</tr>
<tr>
<td>RVQRPOR</td>
<td>0.8351</td>
<td>0.8509</td>
<td>0.8587</td>
<td>0.8475</td>
<td>0.8315</td>
</tr>
<tr>
<td>RVQRKendall</td>
<td>0.8500</td>
<td>0.8679</td>
<td>0.8722</td>
<td>0.8665</td>
<td>0.8466</td>
</tr>
<tr>
<td>RVQRsemi</td>
<td>0.8519</td>
<td>0.8474</td>
<td>0.8334</td>
<td>0.8708</td>
<td>0.8499</td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>0.8465</td>
<td>0.8629</td>
<td>0.8708</td>
<td>0.8606</td>
<td>0.8447</td>
</tr>
<tr>
<td>RELMEDATL $^m_r(\alpha)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LQR</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>LQRWSA</td>
<td>0.9992</td>
<td>0.9995</td>
<td>0.9994</td>
<td>1.0001</td>
<td>0.9996</td>
</tr>
<tr>
<td>DVQR</td>
<td>0.8974</td>
<td>0.8387</td>
<td>0.8574</td>
<td>0.8608</td>
<td>0.8184</td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.8062</td>
<td>0.8433</td>
<td>0.8566</td>
<td>0.8687</td>
<td>0.8354</td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>0.8276</td>
<td>0.8518</td>
<td>0.8543</td>
<td>0.8658</td>
<td>0.8191</td>
</tr>
<tr>
<td>RVQRPOR</td>
<td>0.8279</td>
<td>0.8514</td>
<td>0.8526</td>
<td>0.8667</td>
<td>0.8139</td>
</tr>
<tr>
<td>RVQRKendall</td>
<td>0.8409</td>
<td>0.8898</td>
<td>0.9055</td>
<td>0.8905</td>
<td>0.8230</td>
</tr>
<tr>
<td>RVQRsemi</td>
<td>0.8419</td>
<td>0.8524</td>
<td>0.8716</td>
<td>0.8713</td>
<td>0.8160</td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>0.8398</td>
<td>0.8531</td>
<td>0.8737</td>
<td>0.8729</td>
<td>0.8229</td>
</tr>
</tbody>
</table>
Figure 19: Scenario 4: bar plot of $N_{D\text{-vine}}^m$, $N_{C\text{-vine}}^m$ and $N_{C/D\text{-vine}}^m$ for $m \in M_{\text{vine}}$. 
Scenario 6

Figure 20: Scenario 6: box plots of $\{\text{ATL}_r^m(\alpha)\}_{r=1,\ldots,100}$ for all $m \in \mathcal{M}$ and $\alpha \in \mathcal{A}$.

<table>
<thead>
<tr>
<th>$m$</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>LQR</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>LQRWSA</td>
<td>1.0001</td>
<td>1.0006</td>
<td>1.0010</td>
<td>1.0001</td>
<td>0.9988</td>
</tr>
<tr>
<td>DVQR</td>
<td>0.8591</td>
<td>0.8999</td>
<td>0.9129</td>
<td>0.9012</td>
<td>0.8532</td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.8564</td>
<td>0.8985</td>
<td>0.9068</td>
<td>0.8981</td>
<td>0.8563</td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>0.8503</td>
<td>0.8984</td>
<td>0.9034</td>
<td>0.9004</td>
<td>0.8581</td>
</tr>
<tr>
<td>RVQRPCOR</td>
<td>0.8529</td>
<td>0.8889</td>
<td>0.8951</td>
<td>0.8854</td>
<td>0.8482</td>
</tr>
<tr>
<td>RVQRkendall</td>
<td>0.8896</td>
<td>0.9372</td>
<td>0.9496</td>
<td>0.9374</td>
<td>0.8864</td>
</tr>
<tr>
<td>RVQRsemi</td>
<td>0.8905</td>
<td>0.9414</td>
<td>0.9540</td>
<td>0.9425</td>
<td>0.8902</td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>0.8889</td>
<td>0.9348</td>
<td>0.9459</td>
<td>0.9346</td>
<td>0.8841</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$m$</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>LQR</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>LQRWSA</td>
<td>1.0003</td>
<td>1.0054</td>
<td>1.0004</td>
<td>1.0029</td>
<td>1.0018</td>
</tr>
<tr>
<td>DVQR</td>
<td>0.7927</td>
<td>0.8648</td>
<td>0.8997</td>
<td>0.8448</td>
<td>0.7657</td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.8657</td>
<td>0.8803</td>
<td>0.8799</td>
<td>0.8548</td>
<td>0.7669</td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>0.8623</td>
<td>0.8638</td>
<td>0.8861</td>
<td>0.8528</td>
<td>0.7887</td>
</tr>
<tr>
<td>RVQRPCOR</td>
<td>0.7942</td>
<td>0.8604</td>
<td>0.8819</td>
<td>0.8549</td>
<td>0.8068</td>
</tr>
<tr>
<td>RVQRkendall</td>
<td>0.8285</td>
<td>0.8794</td>
<td>0.8949</td>
<td>0.8736</td>
<td>0.8190</td>
</tr>
<tr>
<td>RVQRsemi</td>
<td>0.8365</td>
<td>0.8897</td>
<td>0.9005</td>
<td>0.8869</td>
<td>0.8198</td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>0.8242</td>
<td>0.8794</td>
<td>0.8933</td>
<td>0.8840</td>
<td>0.8307</td>
</tr>
</tbody>
</table>

Table 9: Scenario 6: $\text{MIATL}^m(\alpha)$ and $\text{MEDATL}^m(\alpha)$ for $m \in \mathcal{M}$ and $\alpha \in \mathcal{A}$. 

166
Figure 21: Scenario 6: bar plot of $N^m_{D\text{-vine}}$, $N^m_{C\text{-vine}}$ and $N^m_{C/D\text{-vine}}$ for $m \in \mathcal{M}_{\text{vine}}$. 
Box plots - percentage of independence variables

Since in scenarios $S = 2, 4, 6, 8$, it holds that $\text{indep}_0 = 1$ we additionally analyze how many “independent” covariates have been included by method $m$.

Figure 22: Scenario 2: box plots of $\{N_{\text{indep}}^{m,r}\}_{r=1,...,100}$ for all $m \in \mathcal{M}$.

Figure 23: Scenario 4: box plots of $\{N_{\text{indep}}^{m,r}\}_{r=1,...,100}$ for all $m \in \mathcal{M}$.
Figure 24: Scenario 6: box plots of $\{N_{\text{indep}}^{m,r}\}_{r=1,\ldots,100}$ for all $m \in \mathcal{M}$.

Figure 25: Scenario 8: box plots of $\{N_{\text{indep}}^{m,r}\}_{r=1,\ldots,100}$ for all $m \in \mathcal{M}$.
### Absolute values of RELMEDATL and RELMIATL

<table>
<thead>
<tr>
<th>$m$</th>
<th>$\alpha$</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIATL$^m(\alpha) \times 100$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LQR</td>
<td>1.2423</td>
<td>3.6782</td>
<td>4.5568</td>
<td>3.6750</td>
<td>1.2343</td>
<td></td>
</tr>
<tr>
<td>LQRWSA</td>
<td>1.2424</td>
<td>3.6786</td>
<td>4.5562</td>
<td>3.6748</td>
<td>1.2350</td>
<td></td>
</tr>
<tr>
<td>DVQR</td>
<td>1.0060</td>
<td>3.1360</td>
<td>3.9505</td>
<td>3.1390</td>
<td>1.0117</td>
<td></td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>1.0109</td>
<td>3.1430</td>
<td>3.9471</td>
<td>3.1334</td>
<td>1.0155</td>
<td></td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>1.0062</td>
<td>3.1435</td>
<td>3.9530</td>
<td>3.1447</td>
<td>1.0175</td>
<td></td>
</tr>
<tr>
<td>RVQRPCOR</td>
<td>1.0084</td>
<td>3.1390</td>
<td>3.9377</td>
<td>3.1310</td>
<td>1.0217</td>
<td></td>
</tr>
<tr>
<td>RVQRkendall</td>
<td>1.0335</td>
<td>3.2158</td>
<td>4.0340</td>
<td>3.2214</td>
<td>1.0499</td>
<td></td>
</tr>
<tr>
<td>RVQRsemi</td>
<td>1.0344</td>
<td>3.2222</td>
<td>4.0380</td>
<td>3.2236</td>
<td>1.0485</td>
<td></td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>1.0340</td>
<td>3.2176</td>
<td>4.0398</td>
<td>3.2266</td>
<td>1.0528</td>
<td></td>
</tr>
<tr>
<td>MEDATL$^m(\alpha) \times 100$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LQR</td>
<td>1.1048</td>
<td>2.8787</td>
<td>3.5276</td>
<td>2.9121</td>
<td>1.1308</td>
<td></td>
</tr>
<tr>
<td>LQRWSA</td>
<td>1.1006</td>
<td>2.8802</td>
<td>3.5275</td>
<td>2.9050</td>
<td>1.1299</td>
<td></td>
</tr>
<tr>
<td>DVQR</td>
<td>0.8641</td>
<td>2.5501</td>
<td>3.1921</td>
<td>2.5918</td>
<td>0.8750</td>
<td></td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.8811</td>
<td>2.5562</td>
<td>3.2206</td>
<td>2.6140</td>
<td>0.9095</td>
<td></td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>0.8823</td>
<td>2.5587</td>
<td>3.2208</td>
<td>2.5882</td>
<td>0.8906</td>
<td></td>
</tr>
<tr>
<td>RVQRPCOR</td>
<td>0.8718</td>
<td>2.5307</td>
<td>3.1640</td>
<td>2.5856</td>
<td>0.8937</td>
<td></td>
</tr>
<tr>
<td>RVQRkendall</td>
<td>0.8818</td>
<td>2.5837</td>
<td>3.1996</td>
<td>2.6741</td>
<td>0.8923</td>
<td></td>
</tr>
<tr>
<td>RVQRsemi</td>
<td>0.8985</td>
<td>2.5978</td>
<td>3.1667</td>
<td>2.6510</td>
<td>0.9007</td>
<td></td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>0.8808</td>
<td>2.5837</td>
<td>3.1996</td>
<td>2.6849</td>
<td>0.8914</td>
<td></td>
</tr>
</tbody>
</table>

Table 10: Scenario 1: MIATL$^m(\alpha)$ and MEDATL$^m(\alpha)$ for $m \in \mathcal{M}$ and $\alpha \in \mathcal{A}$.

<table>
<thead>
<tr>
<th>$m$</th>
<th>$\alpha$</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIATL$^m(\alpha) \times 100$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LQR</td>
<td>1.2423</td>
<td>3.6782</td>
<td>4.5568</td>
<td>3.6750</td>
<td>1.2343</td>
<td></td>
</tr>
<tr>
<td>LQRWSA</td>
<td>1.2424</td>
<td>3.6786</td>
<td>4.5562</td>
<td>3.6748</td>
<td>1.2350</td>
<td></td>
</tr>
<tr>
<td>DVQR</td>
<td>1.0060</td>
<td>3.1360</td>
<td>3.9505</td>
<td>3.1390</td>
<td>1.0117</td>
<td></td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>1.0109</td>
<td>3.1430</td>
<td>3.9471</td>
<td>3.1334</td>
<td>1.0155</td>
<td></td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>1.0062</td>
<td>3.1435</td>
<td>3.9530</td>
<td>3.1447</td>
<td>1.0175</td>
<td></td>
</tr>
<tr>
<td>RVQRPCOR</td>
<td>1.0084</td>
<td>3.1390</td>
<td>3.9377</td>
<td>3.1310</td>
<td>1.0217</td>
<td></td>
</tr>
<tr>
<td>RVQRkendall</td>
<td>1.0335</td>
<td>3.2158</td>
<td>4.0340</td>
<td>3.2214</td>
<td>1.0499</td>
<td></td>
</tr>
<tr>
<td>RVQRsemi</td>
<td>1.0344</td>
<td>3.2222</td>
<td>4.0380</td>
<td>3.2236</td>
<td>1.0485</td>
<td></td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>1.0340</td>
<td>3.2176</td>
<td>4.0398</td>
<td>3.2266</td>
<td>1.0528</td>
<td></td>
</tr>
<tr>
<td>MEDATL$^m(\alpha) \times 100$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LQR</td>
<td>1.1048</td>
<td>2.8787</td>
<td>3.5276</td>
<td>2.9121</td>
<td>1.1308</td>
<td></td>
</tr>
<tr>
<td>LQRWSA</td>
<td>1.1006</td>
<td>2.8802</td>
<td>3.5275</td>
<td>2.9050</td>
<td>1.1299</td>
<td></td>
</tr>
<tr>
<td>DVQR</td>
<td>0.8641</td>
<td>2.5501</td>
<td>3.1921</td>
<td>2.5918</td>
<td>0.8750</td>
<td></td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.8811</td>
<td>2.5562</td>
<td>3.2206</td>
<td>2.6140</td>
<td>0.9095</td>
<td></td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>0.8823</td>
<td>2.5587</td>
<td>3.2208</td>
<td>2.5882</td>
<td>0.8906</td>
<td></td>
</tr>
<tr>
<td>RVQRPCOR</td>
<td>0.8718</td>
<td>2.5307</td>
<td>3.1640</td>
<td>2.5856</td>
<td>0.8937</td>
<td></td>
</tr>
<tr>
<td>RVQRkendall</td>
<td>0.8818</td>
<td>2.5837</td>
<td>3.1996</td>
<td>2.6741</td>
<td>0.8923</td>
<td></td>
</tr>
<tr>
<td>RVQRsemi</td>
<td>0.8985</td>
<td>2.5978</td>
<td>3.1667</td>
<td>2.6510</td>
<td>0.9007</td>
<td></td>
</tr>
<tr>
<td>RVQRindepTest</td>
<td>0.8808</td>
<td>2.5837</td>
<td>3.1996</td>
<td>2.6849</td>
<td>0.8914</td>
<td></td>
</tr>
</tbody>
</table>

Table 11: Scenario 2: MIATL$^m(\alpha)$ and MEDATL$^m(\alpha)$ for $m \in \mathcal{M}$ and $\alpha \in \mathcal{A}$. 

170
### Table 12: Scenario 3: MIATLₘ(α) and MEDATLₘ(α) for \( m \in \mathcal{M} \) and \( \alpha \in \mathcal{A} \).

<table>
<thead>
<tr>
<th>( m )</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>( \text{MIATL} \times 100 )</td>
<td>( \text{LQR} )</td>
<td>( 1.2182 )</td>
<td>3.6528</td>
<td>4.5205</td>
</tr>
<tr>
<td></td>
<td>( \text{LQRWSA} )</td>
<td>( 1.2186 )</td>
<td>3.6535</td>
<td>4.5219</td>
<td>3.6379</td>
</tr>
<tr>
<td></td>
<td>( \text{DVQR} )</td>
<td>( 1.0008 )</td>
<td>3.1115</td>
<td>3.9024</td>
<td>3.0847</td>
</tr>
<tr>
<td></td>
<td>( \text{DVQRPCOR} )</td>
<td>( 1.0045 )</td>
<td>3.1083</td>
<td>3.8977</td>
<td>3.0829</td>
</tr>
<tr>
<td></td>
<td>( \text{RVQRDVQR} )</td>
<td>( 0.9969 )</td>
<td>3.1039</td>
<td>3.8852</td>
<td>3.0715</td>
</tr>
<tr>
<td></td>
<td>( \text{RVQRPCOR} )</td>
<td>( 0.9984 )</td>
<td>3.1417</td>
<td>3.9661</td>
<td>3.1439</td>
</tr>
<tr>
<td></td>
<td>( \text{RVQRkendall} )</td>
<td>( 1.0140 )</td>
<td>3.1382</td>
<td>3.9313</td>
<td>3.0707</td>
</tr>
<tr>
<td></td>
<td>( \text{RVQRsemi} )</td>
<td>( 1.0184 )</td>
<td>3.1615</td>
<td>3.9661</td>
<td>3.1439</td>
</tr>
<tr>
<td></td>
<td>( \text{RVQRindepTest} )</td>
<td>( 1.0169 )</td>
<td>3.1417</td>
<td>3.9252</td>
<td>3.1208</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( m )</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>( \text{MEDATL} \times 100 )</td>
<td>( \text{LQR} )</td>
<td>( 1.0549 )</td>
<td>2.9118</td>
<td>3.5878</td>
</tr>
<tr>
<td></td>
<td>( \text{LQRWSA} )</td>
<td>( 1.0537 )</td>
<td>2.9131</td>
<td>3.5844</td>
<td>2.9381</td>
</tr>
<tr>
<td></td>
<td>( \text{DVQR} )</td>
<td>( 0.8419 )</td>
<td>2.5156</td>
<td>3.1038</td>
<td>2.5490</td>
</tr>
<tr>
<td></td>
<td>( \text{DVQRPCOR} )</td>
<td>( 0.8510 )</td>
<td>2.4571</td>
<td>3.0752</td>
<td>2.5610</td>
</tr>
<tr>
<td></td>
<td>( \text{RVQRDVQR} )</td>
<td>( 0.8545 )</td>
<td>2.4457</td>
<td>3.0631</td>
<td>2.5320</td>
</tr>
<tr>
<td></td>
<td>( \text{RVQRPCOR} )</td>
<td>( 0.8635 )</td>
<td>2.4543</td>
<td>3.0503</td>
<td>2.5393</td>
</tr>
<tr>
<td></td>
<td>( \text{RVQRkendall} )</td>
<td>( 0.8888 )</td>
<td>2.4551</td>
<td>3.0752</td>
<td>2.5489</td>
</tr>
<tr>
<td></td>
<td>( \text{RVQRsemi} )</td>
<td>( 0.8909 )</td>
<td>2.4634</td>
<td>3.0873</td>
<td>2.5602</td>
</tr>
<tr>
<td></td>
<td>( \text{RVQRindepTest} )</td>
<td>( 0.8883 )</td>
<td>2.4566</td>
<td>3.1057</td>
<td>2.5508</td>
</tr>
</tbody>
</table>

### Table 13: Scenario 4: MIATLₘ(α) and MEDATLₘ(α) for \( m \in \mathcal{M} \) and \( \alpha \in \mathcal{A} \).

<table>
<thead>
<tr>
<th>( m )</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>( \text{MIATL} \times 100 )</td>
<td>( \text{LQR} )</td>
<td>( 1.0735 )</td>
<td>2.9227</td>
<td>3.5829</td>
</tr>
<tr>
<td></td>
<td>( \text{LQRWSA} )</td>
<td>( 1.0727 )</td>
<td>2.9212</td>
<td>3.5809</td>
<td>2.9386</td>
</tr>
<tr>
<td></td>
<td>( \text{DVQR} )</td>
<td>( 0.8667 )</td>
<td>( 2.4512 )</td>
<td>3.0719</td>
<td>2.5293</td>
</tr>
<tr>
<td></td>
<td>( \text{DVQRPCOR} )</td>
<td>( 0.8654 )</td>
<td>2.4647</td>
<td>3.0690</td>
<td>2.5524</td>
</tr>
<tr>
<td></td>
<td>( \text{RVQRDVQR} )</td>
<td>( 0.8884 )</td>
<td>2.4897</td>
<td>3.0610</td>
<td>2.5440</td>
</tr>
<tr>
<td></td>
<td>( \text{RVQRPCOR} )</td>
<td>( 0.8888 )</td>
<td>2.4885</td>
<td>3.0548</td>
<td>2.5466</td>
</tr>
<tr>
<td></td>
<td>( \text{RVQRkendall} )</td>
<td>( 0.9028 )</td>
<td>2.6008</td>
<td>3.2442</td>
<td>2.6163</td>
</tr>
<tr>
<td></td>
<td>( \text{RVQRsemi} )</td>
<td>( 0.9038 )</td>
<td>2.4912</td>
<td>3.1230</td>
<td>2.5602</td>
</tr>
<tr>
<td></td>
<td>( \text{RVQRindepTest} )</td>
<td>( 0.9015 )</td>
<td>2.4933</td>
<td>3.1302</td>
<td>2.5647</td>
</tr>
<tr>
<td>m</td>
<td>α</td>
<td>0.05</td>
<td>0.25</td>
<td>0.50</td>
<td>0.75</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>MIATL$_m$ ($\alpha$) $\times 100$</td>
<td>LQR</td>
<td>1.2040</td>
<td>3.4964</td>
<td>4.3243</td>
<td>3.4852</td>
</tr>
<tr>
<td></td>
<td>LQRWSA</td>
<td>1.2020</td>
<td>3.4923</td>
<td>4.3282</td>
<td>3.4918</td>
</tr>
<tr>
<td></td>
<td>DVQR</td>
<td>1.0174</td>
<td>3.1353</td>
<td>3.9443</td>
<td>3.1416</td>
</tr>
<tr>
<td></td>
<td>DVQRPCOR</td>
<td>0.9994</td>
<td>3.1141</td>
<td>3.9174</td>
<td>3.1035</td>
</tr>
<tr>
<td></td>
<td>RVQRDVQR</td>
<td>1.0290</td>
<td>3.1648</td>
<td>3.9791</td>
<td>3.1723</td>
</tr>
<tr>
<td></td>
<td>RVQRPCOR</td>
<td>0.9974</td>
<td>3.0911</td>
<td>3.8789</td>
<td>3.0792</td>
</tr>
<tr>
<td></td>
<td>RVQRkendall</td>
<td>1.0647</td>
<td>3.2535</td>
<td>4.0896</td>
<td>3.2580</td>
</tr>
<tr>
<td></td>
<td>RVQRsemi</td>
<td>1.0545</td>
<td>3.2544</td>
<td>4.0872</td>
<td>3.2575</td>
</tr>
<tr>
<td></td>
<td>RVQRindepTest</td>
<td>1.0554</td>
<td>3.2561</td>
<td>4.0897</td>
<td>3.2609</td>
</tr>
<tr>
<td>MEDATL$_m$ ($\alpha$) $\times 100$</td>
<td>LQR</td>
<td>1.1031</td>
<td>2.7942</td>
<td>3.4496</td>
<td>2.8777</td>
</tr>
<tr>
<td></td>
<td>LQRWSA</td>
<td>1.0918</td>
<td>2.8039</td>
<td>3.4567</td>
<td>2.9015</td>
</tr>
<tr>
<td></td>
<td>DVQR</td>
<td>0.9903</td>
<td>2.4499</td>
<td>3.1090</td>
<td>2.5567</td>
</tr>
<tr>
<td></td>
<td>DVQRPCOR</td>
<td>0.9878</td>
<td>2.4968</td>
<td>3.1211</td>
<td>2.5437</td>
</tr>
<tr>
<td></td>
<td>RVQRDVQR</td>
<td>0.9886</td>
<td>2.4711</td>
<td>3.1051</td>
<td>2.5350</td>
</tr>
<tr>
<td></td>
<td>RVQRPCOR</td>
<td>0.8588</td>
<td>2.4155</td>
<td>2.9777</td>
<td>2.4752</td>
</tr>
<tr>
<td></td>
<td>RVQRkendall</td>
<td>0.9028</td>
<td>2.5172</td>
<td>3.1289</td>
<td>2.5877</td>
</tr>
<tr>
<td></td>
<td>RVQRsemi</td>
<td>0.9028</td>
<td>2.5327</td>
<td>3.1697</td>
<td>2.6326</td>
</tr>
<tr>
<td></td>
<td>RVQRindepTest</td>
<td>0.9028</td>
<td>2.5172</td>
<td>3.1208</td>
<td>2.5800</td>
</tr>
</tbody>
</table>

Table 14: Scenario 5: MIATL$_m$ ($\alpha$) and MEDATL$_m$ ($\alpha$) for $m \in \mathcal{M}$ and $\alpha \in \mathcal{A}$.

<table>
<thead>
<tr>
<th>m</th>
<th>α</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIATL$_m$ ($\alpha$) $\times 100$</td>
<td>LQR</td>
<td>1.2000</td>
<td>3.5667</td>
<td>4.3513</td>
<td>3.5012</td>
<td>1.1948</td>
</tr>
<tr>
<td></td>
<td>LQRWSA</td>
<td>1.2002</td>
<td>3.5690</td>
<td>4.3556</td>
<td>3.5017</td>
<td>1.1933</td>
</tr>
<tr>
<td></td>
<td>DVQR</td>
<td>1.0309</td>
<td>3.1556</td>
<td>3.9725</td>
<td>3.1554</td>
<td>1.0194</td>
</tr>
<tr>
<td></td>
<td>DVQRPCOR</td>
<td>1.0277</td>
<td>3.1509</td>
<td>3.9460</td>
<td>3.1444</td>
<td>1.0231</td>
</tr>
<tr>
<td></td>
<td>RVQRDVQR</td>
<td>1.0324</td>
<td>3.1736</td>
<td>3.9834</td>
<td>3.1698</td>
<td>1.0255</td>
</tr>
<tr>
<td></td>
<td>RVQRPCOR</td>
<td>1.0235</td>
<td>3.1171</td>
<td>3.8947</td>
<td>3.0998</td>
<td>1.0134</td>
</tr>
<tr>
<td></td>
<td>RVQRkendall</td>
<td>1.0675</td>
<td>3.2866</td>
<td>4.1230</td>
<td>3.2819</td>
<td>1.0590</td>
</tr>
<tr>
<td></td>
<td>RVQRsemi</td>
<td>1.0686</td>
<td>3.3013</td>
<td>4.1510</td>
<td>3.3000</td>
<td>1.0636</td>
</tr>
<tr>
<td></td>
<td>RVQRindepTest</td>
<td>1.0667</td>
<td>3.2781</td>
<td>4.1161</td>
<td>3.2722</td>
<td>1.0565</td>
</tr>
<tr>
<td>MEDATL$_m$ ($\alpha$) $\times 100$</td>
<td>LQR</td>
<td>1.1908</td>
<td>2.8574</td>
<td>3.5144</td>
<td>2.9527</td>
<td>1.1013</td>
</tr>
<tr>
<td></td>
<td>LQRWSA</td>
<td>1.0911</td>
<td>2.8728</td>
<td>3.5159</td>
<td>2.9612</td>
<td>1.1032</td>
</tr>
<tr>
<td></td>
<td>DVQR</td>
<td>0.8646</td>
<td>2.4713</td>
<td>3.1302</td>
<td>2.5240</td>
<td>0.8455</td>
</tr>
<tr>
<td></td>
<td>DVQRPCOR</td>
<td>0.8788</td>
<td>2.5152</td>
<td>3.0925</td>
<td>2.5240</td>
<td>0.8445</td>
</tr>
<tr>
<td></td>
<td>RVQRDVQR</td>
<td>0.8752</td>
<td>2.4682</td>
<td>3.1142</td>
<td>2.5181</td>
<td>0.8686</td>
</tr>
<tr>
<td></td>
<td>RVQRPCOR</td>
<td>0.8663</td>
<td>2.4585</td>
<td>3.0993</td>
<td>2.5242</td>
<td>0.8885</td>
</tr>
<tr>
<td></td>
<td>RVQRkendall</td>
<td>0.9037</td>
<td>2.5119</td>
<td>3.1452</td>
<td>2.5796</td>
<td>0.9029</td>
</tr>
<tr>
<td></td>
<td>RVQRsemi</td>
<td>0.9125</td>
<td>2.5422</td>
<td>3.1647</td>
<td>2.6187</td>
<td>0.9028</td>
</tr>
<tr>
<td></td>
<td>RVQRindepTest</td>
<td>0.8990</td>
<td>2.5128</td>
<td>3.1395</td>
<td>2.6102</td>
<td>0.9148</td>
</tr>
</tbody>
</table>

Table 15: Scenario 6: MIATL$_m$ ($\alpha$) and MEDATL$_m$ ($\alpha$) for $m \in \mathcal{M}$ and $\alpha \in \mathcal{A}$. 
Table 16: Scenario 7: MIATL$^m(\alpha)$ and MEDATL$^m(\alpha)$ for $m \in \mathcal{M}$ and $\alpha \in \mathcal{A}$.

<table>
<thead>
<tr>
<th>$m$ (\alpha)</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>LQR</td>
<td>1.1554</td>
<td>3.4302</td>
<td>4.2611</td>
<td>3.4408</td>
<td>1.1606</td>
</tr>
<tr>
<td>LQRWSA</td>
<td>1.1557</td>
<td>3.4319</td>
<td>4.2640</td>
<td>3.4440</td>
<td>1.1605</td>
</tr>
<tr>
<td>DVQR</td>
<td>0.9736</td>
<td>3.0382</td>
<td>3.8264</td>
<td>3.0364</td>
<td>0.9682</td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.9657</td>
<td>3.0285</td>
<td>3.8210</td>
<td>3.0329</td>
<td>0.9712</td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>0.9825</td>
<td>3.0663</td>
<td>3.8665</td>
<td>3.0653</td>
<td>0.9762</td>
</tr>
<tr>
<td>RVQRPCOR</td>
<td><strong>0.9573</strong></td>
<td><strong>2.9955</strong></td>
<td><strong>3.7848</strong></td>
<td><strong>2.9974</strong></td>
<td><strong>0.9541</strong></td>
</tr>
<tr>
<td>RVQRrendall</td>
<td>1.0203</td>
<td>3.1706</td>
<td>3.9942</td>
<td>3.1709</td>
<td>1.0221</td>
</tr>
<tr>
<td>RVQreemi</td>
<td>1.0175</td>
<td>3.1594</td>
<td>3.9763</td>
<td>3.1652</td>
<td>1.0220</td>
</tr>
<tr>
<td>RVQRindeptest</td>
<td>1.0182</td>
<td>3.1628</td>
<td>3.9819</td>
<td>3.1612</td>
<td>1.0204</td>
</tr>
</tbody>
</table>

Table 17: Scenario 8: MIATL$^m(\alpha)$ and MEDATL$^m(\alpha)$ for $m \in \mathcal{M}$ and $\alpha \in \mathcal{A}$.

<table>
<thead>
<tr>
<th>$m$ (\alpha)</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>LQR</td>
<td>1.0405</td>
<td>2.8297</td>
<td>3.4495</td>
<td>2.8202</td>
<td>1.0906</td>
</tr>
<tr>
<td>LQRWSA</td>
<td>1.0415</td>
<td>2.8325</td>
<td>3.4509</td>
<td>2.8147</td>
<td>1.0911</td>
</tr>
<tr>
<td>DVQR</td>
<td>0.8492</td>
<td>2.3843</td>
<td>2.9719</td>
<td>2.4092</td>
<td>0.8361</td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.8502</td>
<td><strong>2.3709</strong></td>
<td>2.9714</td>
<td>2.4152</td>
<td>0.8208</td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td>0.8602</td>
<td>2.4182</td>
<td>3.0104</td>
<td>2.3674</td>
<td>0.8201</td>
</tr>
<tr>
<td>RVQRPCOR</td>
<td><strong>0.8417</strong></td>
<td>2.3957</td>
<td><strong>2.9682</strong></td>
<td>2.4578</td>
<td><strong>0.8068</strong></td>
</tr>
<tr>
<td>RVQRrendall</td>
<td>0.8802</td>
<td>2.5084</td>
<td>3.1527</td>
<td>2.5388</td>
<td>0.8791</td>
</tr>
<tr>
<td>RVQreemi</td>
<td>0.8937</td>
<td>2.5310</td>
<td>3.1436</td>
<td>2.5184</td>
<td>0.8781</td>
</tr>
<tr>
<td>RVQRindeptest</td>
<td>0.8808</td>
<td>2.4221</td>
<td>3.0299</td>
<td>2.4751</td>
<td>0.8551</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$m$ (\alpha)</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>LQR</td>
<td>1.0748</td>
<td>2.9119</td>
<td>3.5123</td>
<td>2.8466</td>
<td>1.0771</td>
</tr>
<tr>
<td>LQRWSA</td>
<td>1.0690</td>
<td>2.9114</td>
<td>3.5125</td>
<td>2.8445</td>
<td>1.0793</td>
</tr>
<tr>
<td>DVQR</td>
<td>0.9887</td>
<td>3.0786</td>
<td>3.8744</td>
<td>3.0839</td>
<td>0.9955</td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.9925</td>
<td>3.0608</td>
<td>3.8429</td>
<td>3.0630</td>
<td>0.9935</td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td><strong>0.9895</strong></td>
<td><strong>3.0497</strong></td>
<td><strong>3.8308</strong></td>
<td><strong>3.0466</strong></td>
<td><strong>0.9855</strong></td>
</tr>
<tr>
<td>RVQRPCOR</td>
<td>1.0263</td>
<td>3.1893</td>
<td>4.0127</td>
<td>3.1910</td>
<td>1.0271</td>
</tr>
<tr>
<td>RVQRrendall</td>
<td>1.0316</td>
<td>3.2135</td>
<td>4.0434</td>
<td>3.2169</td>
<td>1.0352</td>
</tr>
<tr>
<td>RVQreemi</td>
<td>1.0255</td>
<td>3.1839</td>
<td>4.0089</td>
<td>3.1892</td>
<td>1.0264</td>
</tr>
<tr>
<td>RVQRindeptest</td>
<td>1.0255</td>
<td>3.1839</td>
<td>4.0089</td>
<td>3.1892</td>
<td>1.0264</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$m$ (\alpha)</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>LQR</td>
<td>1.0748</td>
<td>2.9119</td>
<td>3.5123</td>
<td>2.8466</td>
<td>1.0771</td>
</tr>
<tr>
<td>LQRWSA</td>
<td>1.0690</td>
<td>2.9114</td>
<td>3.5125</td>
<td>2.8445</td>
<td>1.0793</td>
</tr>
<tr>
<td>DVQR</td>
<td>0.9887</td>
<td>3.0786</td>
<td>3.8744</td>
<td>3.0839</td>
<td>0.9955</td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.9925</td>
<td>3.0608</td>
<td>3.8429</td>
<td>3.0630</td>
<td>0.9935</td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td><strong>0.9895</strong></td>
<td><strong>3.0497</strong></td>
<td><strong>3.8308</strong></td>
<td><strong>3.0466</strong></td>
<td><strong>0.9855</strong></td>
</tr>
<tr>
<td>RVQRPCOR</td>
<td>1.0263</td>
<td>3.1893</td>
<td>4.0127</td>
<td>3.1910</td>
<td>1.0271</td>
</tr>
<tr>
<td>RVQRrendall</td>
<td>1.0316</td>
<td>3.2135</td>
<td>4.0434</td>
<td>3.2169</td>
<td>1.0352</td>
</tr>
<tr>
<td>RVQreemi</td>
<td>1.0255</td>
<td>3.1839</td>
<td>4.0089</td>
<td>3.1892</td>
<td>1.0264</td>
</tr>
<tr>
<td>RVQRindeptest</td>
<td>1.0255</td>
<td>3.1839</td>
<td>4.0089</td>
<td>3.1892</td>
<td>1.0264</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$m$ (\alpha)</th>
<th>0.05</th>
<th>0.25</th>
<th>0.50</th>
<th>0.75</th>
<th>0.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>LQR</td>
<td>1.0748</td>
<td>2.9119</td>
<td>3.5123</td>
<td>2.8466</td>
<td>1.0771</td>
</tr>
<tr>
<td>LQRWSA</td>
<td>1.0690</td>
<td>2.9114</td>
<td>3.5125</td>
<td>2.8445</td>
<td>1.0793</td>
</tr>
<tr>
<td>DVQR</td>
<td>0.9887</td>
<td>3.0786</td>
<td>3.8744</td>
<td>3.0839</td>
<td>0.9955</td>
</tr>
<tr>
<td>DVQRPCOR</td>
<td>0.9925</td>
<td>3.0608</td>
<td>3.8429</td>
<td>3.0630</td>
<td>0.9935</td>
</tr>
<tr>
<td>RVQRDVQR</td>
<td><strong>0.9895</strong></td>
<td><strong>3.0497</strong></td>
<td><strong>3.8308</strong></td>
<td><strong>3.0466</strong></td>
<td><strong>0.9855</strong></td>
</tr>
<tr>
<td>RVQRPCOR</td>
<td>1.0263</td>
<td>3.1893</td>
<td>4.0127</td>
<td>3.1910</td>
<td>1.0271</td>
</tr>
<tr>
<td>RVQRrendall</td>
<td>1.0316</td>
<td>3.2135</td>
<td>4.0434</td>
<td>3.2169</td>
<td>1.0352</td>
</tr>
<tr>
<td>RVQreemi</td>
<td>1.0255</td>
<td>3.1839</td>
<td>4.0089</td>
<td>3.1892</td>
<td>1.0264</td>
</tr>
<tr>
<td>RVQRindeptest</td>
<td>1.0255</td>
<td>3.1839</td>
<td>4.0089</td>
<td>3.1892</td>
<td>1.0264</td>
</tr>
</tbody>
</table>

173