Technische Universität München
Zentrum Mathematik
Lehrstuhl für Mathematische Statistik

Vine based models for multivariate volatility time-series and time-to-event data

Nicole Barthel

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Prof. Dr. Matthias Scherer
Prüfende/-r der Dissertation:	1.	Prof. Claudia Czado, Ph.D.
	2.	Prof. Dr. Paul Janssen Universiteit Hasselt, Belgien
	3.	Prof. Harry Joe, Ph.D. University of British Columbia, Kanada (schriftliches Gutachten)

Die Dissertation wurde am 14.02.2019 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 04.04.2019 angenommen.

Zusammenfassung

In der vorliegenden Arbeit werden neue Vine basierte Modellierungsmethoden für Daten zweier bedeutender Forschungsgebiete entwickelt: die Modellierung von multivariaten hochfrequenten Volatilitätszeitreihen und die Analyse mehrdimensionaler Überlebenszeitdaten.

Zeitreihenmodelle zur Vorhersage von realisierten Kovarianzmatrizen unterliegen Restriktionen, da positive Definitheit der Prognosen gewährleistet sein muss. Um dies zu umgehen, werden reguläre Vines zur Datentransformation verwendet, wobei der Zusammenhang zwischen einer positiv definiten Korrelationsmatrix und den durch eine reguläre Vine Struktur spezifizierten partiellen Korrelationen genutzt wird. Diese Transformation ist nicht nur interpretierbar, sondern sie ermöglicht auch eine parametersparsame Modellierung der resultierenden Zeitreihen. Abhängigkeitsstrukturen zwischen den realisierten Varianzen sowie den realisierten Pearson und partiellen Korrelationen werden durch flexible reguläre Vine Verteilungen modelliert. Letztere werden entsprechend der zugrundeliegenden regulären Vine Struktur durch beliebige bivariate (bedingte) Copulas und Randverteilungen konstruiert. Die Modellierungs- und Vorhersagegüte der entwickelten Methodik wird anhand einer Datenanwendung evaluiert und mit bekannten Benchmark-Modellen, die auf der Cholesky Faktorisierung beruhen, verglichen.

Nachfolgend werden reguläre Vine Copula Modelle entwickelt, um komplexe Abhängigkeitsstrukturen in mehrdimensionalen Überlebenszeitdaten zu schätzen. Aufgrund eines begrenzten Beobachtungszeitraumes sind Überlebenszeitdaten meist rechtszensiert. Demzufolge bedarf es der Anpassung statistischer Analyseverfahren wie z.B. der Likelihood Schätzung. In einem ersten Projekt werden sogenannte univariat rechtszensierte Daten untersucht, wobei die gleiche Anzahl an Beobachtungen für alle Probeeinheiten der Studie vorliegt. Ein zweistufiges Schätzverfahren wird entwickelt und die Performanz anhand ausführlicher Simulationen untersucht. Für reale Daten wird aufgezeigt, wie ein adäquates reguläres Vine Copula Modell geschätzt werden kann. In einem zweiten Projekt wird berücksichtigt, dass in vielen Überlebenszeitstudien das untersuchte Ereignis wiederkehrend ist. Anders als in der klassischen Überlebenszeitanalyse sind die Zeiten zwischen aufeinanderfolgenden Ereignissen und die Zeit zur Zensierung nicht mehr stochastisch unabhängig. Zudem kann das beobachtete Ereignis für Studienteilnehmer unterschiedlich oft eintreten. Unter Verwendung von D-Vine Copulas werden vier neue Modellierungsstrategien für derart beschaffene Daten vorgestellt: ein einstufiges parametrisches und ein zweistufiges semiparametrisches Schätzverfahren, das jeweils global oder sequentiell durchgeführt werden kann. Ausführliche Simulationen validieren alle Schätzmethoden. Die Analyse realer Daten zeigt, dass D-Vine Copulas entscheidende Einblicke darüber ermöglichen, wie sich die Abhängigkeit wiederkehrender Ereignisse in Form und Stärke über die Zeit entwickelt.

Abstract

In this thesis, novel vine based methodologies for data in two quite different but likewise important research fields are developed: modeling high-frequency volatility time-series and analyzing multivariate time-to-event data.

First, when forecasting realized covariance matrices the requirement of positive definite predictions imposes restrictions on time-series models. To avoid this, regular vines are used to transform the original data relying on the one-to-one relationship between a positive definite correlation matrix and its set of partial correlations specified by any vine. The transformation not only is interpretable, but also allows for parsimonious time-series modeling. Dependence patterns between realized variances and realized standard and partial correlations are modeled by extending regular vines to the flexible class of regular vine distributions. The latter can be constructed from a cascade of arbitrary bivariate (conditional) copulas and marginal distributions according to the underlying regular vine structure. The modeling and forecasting performance of the proposed methodology is evaluated through its application to real life data and is compared to popular Cholesky decomposition based benchmark models.

Second, regular vine copula models to capture the possibly complex dependence pattern in multivariate event time data are developed. Due to a limited follow-up period event time data are often subject to right-censoring. As a consequence, the inferential tools existing for complete data such as likelihood estimation need to be adapted. In a first project, balanced data subject to common right-censoring are investigated. A two-stage estimation approach is established and evaluated through extensive simulations. For real life data it is shown how an appropriate regular vine copula model can be selected for data at hand. In a second project, it is taken into account that in many time-to-event studies, the event of interest is recurrent. In contrast to classical analysis, the times between subsequent events and censoring times can no longer be assumed independent. Also, the number of recurrences typically varies among sample units leading to unbalanced data. Using D-vine copulas, we propose four estimation strategies to tackle these challenges: one-stage parametric and two-stage semiparametric estimation both proceeding either globally or sequentially. Extensive simulations show good finite sample performance of all proposed methods. The analysis of real life data reveals that a D-vine copula detects relevant insights, on how the dependence of recurrent event times changes in strength and type over time.

Acknowledgements

First and foremost, I want to express my sincere gratitude to my supervisor Prof. Claudia Czado. Our joint work started when I was a master's student being asked whether I would be interested in a data set on cow udders – of course, I was. Thank you for your trust, your exceptional guidance and for always having an open door for discussions. It was a pleasure to learn from your statistical expertise and to be inspired by your creativity. You always supported my curiosity and desire to spread my wings, which resulted in the attendance of numerous international conferences, several research stays abroad and the opportunity to teach. Most importantly, I appreciate your compassion and encouragement, which kept me moving forward even with a broken foot.

The close collaboration with Prof. Paul Janssen from the Universiteit Hasselt in Belgium was one of the most rewarding experiences during the past years. The many intense discussions together with Dr. Candida Geerdens and your exceptional striving for accuracy shaped me enormously as a researcher. I feel honored to have you as a member of my PhD committee.

Prof. Harry Joe from the University of British Columbia in Canada was not only a great host during my research stay in Vancouver, but also a challenging and helpful conversation partner, who always gave rise to interesting and relevant research questions. Thank you very much for being willing to act as a referee for my thesis.

I also highly appreciate the collaboration with Prof. Yarema Okhrin from the Universität Augsburg. Our joint DFG-project "Vine copula based modeling and forecasting of multivariate volatility time-series" profited strongly from your econometric expertise and positive spirit.

The enlightening and joyful time I spent with Dr. Candida Geerdens at the Universiteit Hasselt in and outside the office was one of the strongest driving forces for completing my thesis: Bedankt dat je de beste leraar, collega en buddy was die ik had kunnen wensen.

I will preserve many enjoyable memories of the time spent at the Chair of Mathematical Statistics with my colleagues Daniel Kraus, Dominik Müller, Matthias Killiches, Tobias Erhardt, Alexander Kreuzer and Thomas Nagler. Dominik, thank you for being such a great officemate. Alex and Thomas, thank you for the many encouraging and intense discussions especially during the last months of completing my thesis.

Additional thanks are directed towards the German Research Foundation as well as the TUM Graduate School and the International School of Applied Mathematics for financial support.

Finally, and most importantly, I want to sincerely thank my friends and family for their continuous love and encouragement, which always brought me strength, joy and laughter. Jochen, thank you for accompanying me through all ups and downs of this PhD journey. I cannot wait for all the future paths we will take from here.

Contents

Chapter 1

Introduction

Interconnectedness and dependencies are omnipresent navigating phenomena in nature, human interaction, economic development or biomedical processes to name few examples of an endless list. In the recent years, the fast-paced technical progress significantly facilitated the availability of all kind of data, and thus enhanced the opportunities for their statistical analysis. Likewise, the increasing data volume and data complexity request elaborate computer-based mathematical and statistical methodologies, which are able to translate and to transfer the hidden data content into the practical context of relevant application fields.

For modeling the dependence among random variables copulas have become popular alternatives to classical multivariate distribution functions, which often require the cumbersome estimation of many parameters or which are too restrictive to adequately reflect the variables' joint behavior. Sklar (1959) laid the foundations for copula theory. According to his seminal theorem, a copula is a dependence function, which interconnects the univariate marginal distribution functions of random variables and therewith models their joint distribution function. Consequently, a copula model allows to separate the individual behavior of random variables from their joint interaction – the dependence. This is particularly convenient, if for example interest is in separate marginal models or in an explicit dependence model. However, it was not until years later that with improving computational capacities copula theory attracted more and more attention promoted by profound and thorough publications on copula modeling (Joe, 1997; Embrechts et al., 2003; Nelsen, 2006; Joe, 2014).

Ever since, copula models have been investigated and applied in diverse research fields (for example, see Elidan (2013) or Aas (2016) for reviews). Popular and well-studied parametric copula classes are the Archimedean copula family and elliptical copulas with their two well-known representatives: the Gaussian copula and the Student t copula. Elliptical copulas are symmetric and only the Student t copula exhibits tail-dependence. In higher dimensions, they require a large number of parameters. While upper and/or lower tail-dependence can be achieved by Archimedean copulas, they rely on only a small number of parameters which control the dependence among all variables. This results in a lack of flexibility and only restrictive dependence patterns that can be detected by the model.

Motivated by the nevertheless appealing benefits of copula models and the large variety of bivariate copulas, Joe (1996) proposed to construct multivariate copulas from bivariate ones

using conditioning. The resulting pair-copula constructions represent a decomposition of the d-dimensional density function of d random variables into a cascade of $d-1$ unconditional and $d(d-1)/2-(d-1)$ conditional bivariate copulas. Since the type and strength of dependence of these pair-copulas can be arbitrarily chosen and combined, pair-copula constructions stand out through their ability to flexibly model complex asymmetric and nonlinear dependencies even in high dimensions. Clearly, the conditioning underlying a pair-copula construction is not unique. To order and describe the $\frac{d!}{2}2^{\binom{d-2}{2}}$ decompositions possible in d dimensions (Morales Napoles et al., 2010), Bedford and Cooke (2002) introduce regular vines as a graph theoretical object. Therefore, pair-copula constructions are also referred to as regular vine (R-vine) copulas.

R-vine copulas experienced their ultimate kickoff when Aas et al. (2009) established statistical inference such as maximum likelihood estimation. Ever since, R-vine copula theory was constantly enhanced and extensively studied in literature including Bayesian analysis (Min and Czado, 2010; Czado and Min, 2011; Gruber et al., 2015; Gruber and Czado, 2018), nonparametric pair-copula constructions (Nagler and Czado, 2016; Nagler et al., 2017), parsimonious modeling techniques (Brechmann et al., 2012; Brechmann and Joe, 2015), the estimation of standard errors (Stöber and Schepsmeier, 2013) or the development of goodness-of-fit tests (Schepsmeier, 2016) just to name few contributions. In particular, the extensive software provided in the `R`-package `VineCopula` (Schepsmeier et al., 2017) makes R-vine copula based modeling accessible to a broad range of statisticians and practitioners. As a consequence, R-vine copulas have found applications in widespread research fields such as sociology (Cooke et al., 2015), weather forecasting (Möller et al., 2018), biology (Schellhase and Spanhel, 2018), spatial statistics (Erhardt et al., 2015), insurance (Erhardt and Czado, 2012; Shi and Yang, 2018) or finance (Loaiza Maya et al., 2015; Brechmann and Czado, 2015; Fischer et al., 2017; Aas, 2016).

In this thesis, two new research fields are tackled in the context of R-vines and R-vine copulas. In the first part, modeling and forecasting of volatility as one of the most actively discussed topics in financial econometrics is investigated. The availability of high-frequency data allows to obtain the so-called realized volatility as an estimate of the by itself non-observable daily volatility using intra-day returns. Interest is in multivariate data, and thus the goal is to model and forecast time-series of realized covariance matrices. In doing so, symmetry and positive definiteness of the matrix forecasts have to be ensured. A common solution to handle this restraint is to not directly model the components of the realized covariance matrices but to consider transformed data. We propose to use partial correlation vines for this data transformation.

The second part of this thesis discusses R-vine copula models in the context of time-to-event data, also referred to as survival data. The latter are collected whenever primary interest in a study lies in the time until a prespecified event occurs. For example, time until failure of machine parts could be observed or HIV-infected patients could be followed up for time until AIDS diagnosis. Due to a limited follow-up period event time data typically are subject to right-censoring. Thus, for some sample units the true event time is not observed but instead a lower right-censored time is registered. The statistical analysis of univariate right-censored event time data has been studied for decades and therefore is very well established. However, if data appear in clusters the underlying dependence pattern between event times has to be taken

into account as well. Thus, more flexible models are needed. For direct dependence modeling copulas can be used. Copula theory is well established for complete data. However, the presence of right-censoring in clustered event time data complicates the statistical analysis substantially. While the incorporation of right-censoring is indispensable to arrive at a sound statistical data analysis, this modeling aspect is not straightforward. Thus, the application of copulas to right-censored clustered data has been less explored and has been restricted to rather simple copula classes such as elliptical and Archimedean copulas. Against this background, we present two common data settings for clustered right-censored event time data and develop R-vine copula based estimation techniques to model the possibly complex within-cluster dependence.

Outline of the thesis

The content in this thesis is based on three research papers:

- Barthel, N., Czado, C. and Okhrin, Y. (2018a)
  A partial correlation vine based approach for modeling and forecasting multivariate volatility time-series.
  In revision at Computational Statistics & Data Analysis. arXiv: 1802.09585.

- Barthel, N., Geerdens, C., Killiches, M., Janssen, P. and Czado, C. (2018c)
  Vine copula based likelihood estimation of dependence patterns in multivariate event time data.
  Computational Statistics & Data Analysis, 117:109-127.

- Barthel, N., Geerdens, C., Czado, C. and Janssen, P. (2018b)
  Dependence modeling for recurrent event times subject to right-censoring with D-vine copulas.
  To appear in Biometrics. doi:10.1111/biom.13014

For this thesis the content of these papers was revised and extended in various sections including additional methodology, illustrations or explanations.

In Section 2.1, we first introduce regular vines (R-vines) as a graph theoretical object that specifies for d variables a set of $d(d-1)/2$ bivariate (conditional) constraints. We proceed with basic concepts of copula theory in Section 2.2 and combine in Section 2.3 the content of the previous sections to establish the flexible class of R-vine copulas.

Chapter 3 is based on the research paper Barthel et al. (2018a). Motivated by a thorough literature review in Section 3.1 we propose to use partial correlation vine based data transformation to model and forecast time-series of realized covariance matrices. As outlined in Section 3.2, partial correlation vines assign to each edge in an R-vine structure a (partial) correlation coefficient according to the corresponding conditioned and conditioning set. We use that there is a one-to-one relationship between a positive definite correlation matrix and any partial correlation vine (Bedford and Cooke, 2002). Further, the partial correlations specified through an R-vine exhibit the convenient feature to be algebraically independent, i.e. arbitrary values in -1 and 1

can be assigned to the edges of the underlying R-vine structure while positive definiteness of the corresponding correlation matrix is always preserved (Kurowicka and Cooke, 2003).

In Section 3.3, we formulate the general data setting and introduce our main benchmark approach, which relies on the Cholesky decomposition for data transformation. The partial correlation vine data transformation approach is outlined in Section 3.4 along with a real data example. By using the same R-vine structure to transform a series of correlation matrices to a series of partial correlation vines, univariate time-series of realized standard and partial correlations are obtained. Together with the corresponding realized variance time-series the latter are the model components of the partial correlation vine data transformation approach. We start in Section 3.4.1 with the analysis of these model components and find that typical features of volatility data such as long-memory behavior or volatility clustering are less pronounced for certain higher partial correlation time-series. This finding motivates in Section 3.4.2 an R-vine structure selection method, which exclusively relies on historical information of the realized volatility data and which allows for parsimony when modeling the dynamics of the model components. In this second step, we consider multivariate time-series modeling and forecasting based on copulas (Section 3.4.3). Thus, while so far R-vines were exclusively used as a graph theoretical tool to transform the series of realized correlation matrices, R-vine copulas come into play when interconnectedness between the individual time-series of the model components is modeled.

In Section 3.5, we continue the analysis of the real data example. Thereby, appealing benefits of the partial correlation vine data transformation approach such as practical interpretability of the model components, dynamic incorporation of changes on the financial market (Section 3.5.2) and model parsimony for the copula based time-series model (Section 3.5.3) become apparent. For the forecasts of the realized covariance matrices obtained after inverting the predicted partial correlation vines and combining the so-obtained predicted realized correlation matrices with the predicted variances the forecasting performance as compared to the Cholesky decomposition based prediction model is investigated in Section 3.5.4. Our findings both with respect to statistical precision of the forecasts and their mean-variance balance in the context of portfolio optimization strategies give strong justification for the use of the partial correlation vine data transformation approach in practice.

In Chapter 4, we embed R-vine copulas in a completely different context. Here, interest is in modeling dependence patterns in multivariate right-censored event time data. In Section 4.1, we reformulate theoretical background on R-vine copulas in survival terms and provide basics on methodology for univariate right-censored time-to-event data.

In Section 4.2 and Section 4.3, balanced data subject to common right-censoring are discussed. In large parts we refer to the publication Barthel et al. (2018c). After introducing the data setting in Section 4.2.2, we develop an estimation strategy that proceeds in two subsequent steps. In the first step, the marginal survival functions are estimated using standard parametric and nonparametric estimation techniques under the presence of right-censoring. To model, in the second step, the dependence structure we use for right-censored quadruple data results of Barthel (2015) (master's thesis), where the likelihood contributions in terms of R-vine copula components are derived. Right-censoring clearly complicates likelihood optimization and leads to single and

double integrals in the copula likelihood expression. Thus, numerical integration is needed for its evaluation. To lower the computational burden we introduce in Section 4.2.3 a sequential estimation approach. Considering trivariate right-censored event time data generated based on a broad range of simulation scenarios we provide evidence for the good finite sample performance of the presented estimators in Section 4.2.4. Parametric and nonparametric bootstrapping in the presence of right-censoring is discussed in Section 4.3 to obtain standard errors for the likelihood based R-vine copula parameter estimates. In Section 4.3.2, we show for a real data example in four dimensions how to select an appropriate R-vine copula model and how to apply the presented methodology for data at hand.

Section 4.4 is based on the research article Barthel et al. (2018b) and investigates the dependence between recurrent event times subject to right-censoring. We focus on the dependence among the corresponding gap times, i.e. the intervals between two consecutive events. When introducing the general data setting in Section 4.4.2, several modeling challenges become apparent: for example, the presence of induced dependent right-censoring and the unbalanced nature of the data. To capture the serial dependence inherent in gap time data D-vine copulas are the natural choice, since they impose a temporal ordering of the modeled variables. In Section 4.4.4, we suggest a one-stage parametric estimation approach considering both global and sequential likelihood optimization. To increase model flexibility, we propose in Section 4.4.5 two-stage semiparametric estimation, where the marginal survival functions are estimated nonparametrically in a first step. Due to induced dependent right-censoring standard nonparametric estimators for univariate right-censored event time data are no longer consistent. Thus, an alternative nonparametric estimator, which is able to handle induced dependent right-censoring, is introduced. For dependence modeling in the second step both global and sequential likelihood optimization are discussed. The good finite sample performance of the four novel estimation strategies is evaluated through extensive simulations in three and four dimensions (Section 4.4.7). The results allow us to establish in Section 4.4.6 guidelines on the practical use of the four estimation approaches taking into account their sensitivity with respect to specific data characteristics. Methods for standard error estimation and model selection are discussed in Section 4.4.8 and Section 4.4.9, respectively. As a real data example we consider in Section 4.4.10 a study on children suffering from asthma. The data analysis based on D-vine copulas gives new insights on the evolution of the disease. Conditional prediction of the time until relapse given the individual gap time history of a child further demonstrates the flexibility of the proposed D-vine copula based methodology.

Both projects discussed in Chapter 4 stress the need for more flexible copula models in the context of multivariate right-censored time-to-event data.

Chapter 2

R-vines and R-vine copula models

In this chapter, we provide the theoretical basics and mathematical working tools needed throughout the thesis. In large parts, the contents are based on Barthel et al. (2018a) and Barthel et al. (2018b). First, regular vines as a graph theoretical object are introduced in Section 2.1. They will play the key role in Chapter 3 as a data transformation tool for positive definite correlation matrices. In Section 2.2, mathematical background on copulas will be discussed. Focus will be on their usage for dependence modeling. Lastly, Section 2.3 will combine the first two concepts resulting in highly flexible dependence models, namely regular vine copulas. The latter will be the cornerstones for dependence modeling in Chapter 3 and in Chapter 4.

In the remainder of this thesis, capital letters denote random variables and small letters correspond to their realizations. Further, bold capital and bold small letters, denote random vectors and realizations of random vectors, respectively.

2.1 Regular vines

Bedford and Cooke (2002) introduce an R-vine on d elements as a set of $d-1$ linked trees, i.e. undirected and acyclic graphs, $\mathcal{V}_d \coloneqq (\mathcal{T}_1, \dots, \mathcal{T}_{d-1})$ with the set of edges $E(\mathcal{V}_d) \coloneqq E_1 \cup \cdots \cup E_{d-1}$ and the set of nodes $N(\mathcal{V}_d) \coloneqq N_1 \cup \cdots \cup N_{d-1}$ such that

(i) $\mathcal{T}_1$ is a tree with nodes $N_1 = \{1, \dots, d\}$ and edges E_1,

(ii) for $\ell = 2, \dots, d-1$, $\mathcal{T}_\ell$ is a tree with nodes $N_\ell = E_{\ell-1}$ and edges E_ℓ,

(iii) the *proximity condition* holds: For $\ell = 2, \dots, d-1$, whenever two nodes of $\mathcal{T}_\ell$ are connected by an edge, the corresponding edges of $\mathcal{T}_{\ell-1}$ share a node.

According to property (ii), the $d-(\ell-1)$ edges $E_{\ell-1}$ in $\mathcal{T}_{\ell-1}$ become nodes in $\mathcal{T}_\ell$. Based on this linkage, each sequence of trees of an R-vine – from now on referred to as R-vine structure – allows to identify a set of $\binom{d}{2}$ (conditional) bivariate constraints. We refer to Kurowicka and Cooke (2003) and consider an arbitrary edge $e = \{a, b\} \in E_\ell$ of $\mathcal{V}_d$, $2 \leq \ell \leq d-1$, with $a, b \in N_\ell$. Its *complete union* U_e^* is the subset of nodes in $\mathcal{T}_1$, i.e. the subset of $\{1, \dots, d\}$, reachable from

e by the membership relation, i.e.

$$U_e^* \coloneqq \{n \in N_1 : \exists e_1 \in E_1, \dots, e_{\ell-1} \in E_{\ell-1} : n \in e_1 \in \cdots \in e_{\ell-1} \in e\}.$$

The *conditioning set* D_e corresponding to $e = \{a, b\}$ is the intersection of the complete unions U_a^* and U_b^* corresponding to the edges $a, b \in E_{\ell-1}$, i.e.

$$D_e \coloneqq U_a^* \cap U_b^*.$$

The corresponding symmetric difference is referred to as *conditioned set*

$$\{a_e, b_e\} \coloneqq \{U_a^* \backslash D_e, U_b^* \backslash D_e\}.$$

By definition, each conditioned set in $\mathcal{V}_d$ consists of two single elements and in particular forms a unique pair of variables $i, j \in \{1, \dots, d\}$, $i \neq j$. Thus, each pair is modeled by $\mathcal{V}_d$ exactly once either unconditioned, if it forms a conditioned set in the first tree level, or through conditioning, if it forms a conditioning set in tree level $\ell = 2, \dots, d-1$.

Up to four dimensions there are only two possible types of R-vine structures. In all tree levels of a C-vine structure, there is one central node being attached to all edges. This results in a star-like R-vine structure, which reflects an ordering by importance. In a D-vine structure, each node is attached to a maximum of two edges. Thus, a line structure is obtained reflecting a serial ordering. Nodes with only one attached edge are called leafs.

Example 2.1. *Figure 2.1 shows an R-vine structure on six elements labeled with the conditioned set and the conditioning set corresponding to each edge. The latter is indicated by a leading "|". The bold tree segment in $\mathcal{T}_2$ corresponds to the edge $e = \{\{1,2\},\{2,6\}\}$. Reachable from edge $\{1,2\} \in \mathcal{T}_1$ and $\{2,6\} \in \mathcal{T}_1$ are the nodes $1,2 \in N_1$ and $2,6 \in N_1$, respectively. Thus, $D_e = U^*_{\{1,2\}} \cap U^*_{\{2,6\}} = \{1,2\} \cap \{2,6\} = \{2\}$ is the conditioning set corresponding to e and the conditioned set is $\{a_e, b_e\} = \{\{1,2\}\backslash\{2\}, \{2,6\}\backslash\{2\}\} = \{1,6\}$.*

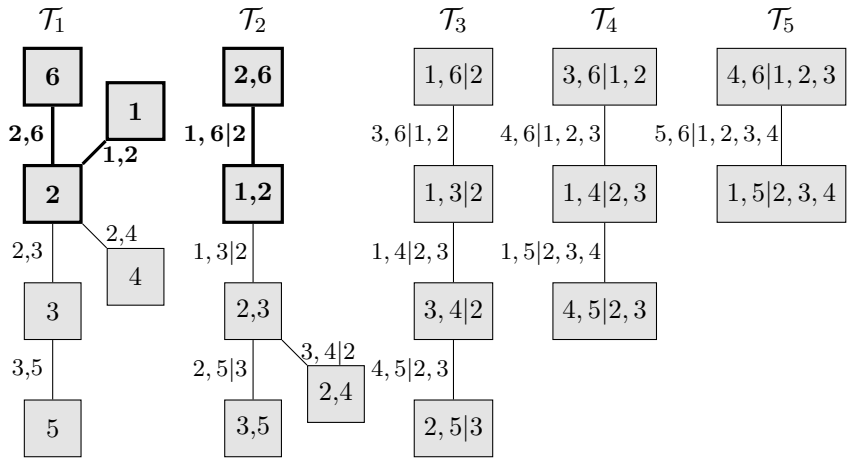


Figure 2.1: Example of a 6-dimensional R-vine structure with conditioning and conditioned sets corresponding to each edge.

2.2 Dependence modeling with copulas

One of the key themes in this thesis is dependence modeling. Consider a d-dimensional vector $\boldsymbol{X} \coloneqq (X_1, \ldots, X_d)$ of continuous random variables and denote by F and f the joint distribution function and the joint density function, respectively. The corresponding marginal distribution functions and marginal density functions are given by F_j and f_j $(j = 1, \ldots, d)$, respectively. The joint distribution function F incorporates information both on the individual behavior of the random variables X_j $(j = 1, \ldots, d)$ described by their univariate marginal distributions F_j and on the dependence between the random variables. Clearly, when interest is in estimating F both data aspects have to be taken into account, which can be cumbersome in higher dimensions.

2.2.1 Sklar's Theorem

For this modeling task, copulas are a useful and flexible tool. A copula $\mathbb{C} : [0,1]^d \to [0,1]$ is defined as a multivariate distribution function with uniform marginal distributions. Thus, the univariate marginal data for a copula are noninformative and the copula itself exclusively describes the dependence between variables. Sklar (1959) provides the fundamental theorem that for each multivariate distribution function F, there exist a copula $\mathbb{C}$, which interconnects the marginal distribution functions F_j $(j = 1, \ldots, d)$ and therewith models the joint distribution function F, i.e.

$$F(x_1, \ldots, x_d) = \mathbb{C}\{F_1(x_1), \ldots, F_d(x_d)\}. \tag{2.1}$$

The copula $\mathbb{C}$ is unique, if all marginal distributions F_j $(j = 1, \ldots, d)$ are continuous. Consequently, the joint density function f expressed in terms of the copula density $\mathbb{c}(u_1, \ldots, u_d) = \frac{\partial^d}{\partial u_1 \cdots \partial u_d} \mathbb{C}(u_1, \ldots, u_d)$ with $(u_1, \ldots, u_d)' \in [0,1]^d$ is given by

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

In the following, we refer to

- the random variables X_j with distribution functions F_j and density functions f_j $(j = 1, \ldots, d)$ as the *original data scale;*
- and to $U_j = F_j(X_j)$ as the *copula data scale*. Note that due to the probability integral transform it holds that $U_j \sim \mathcal{U}[0,1]$ $(j = 1, \ldots, d)$.

It is essential to note that Sklar's Theorem (Sklar, 1959) suggests to split the modeling of F into two subsequent steps. First, the individual behavior, i.e. the marginal distribution functions F_j $(j = 1, \ldots, d)$, are estimated to obtain approximately uniform pseudo copula data using $\hat{U}_j = \hat{F}_j(X_j)$. Second, the dependence is modeled by estimation of the copula $\mathbb{C}$. This proceeding was first proposed by Joe and Xu (1996) and is referred to as inference for margins method. Asymptotic efficiency of this two-stage estimation method is provided in Joe (2005).

2.2.2 Dependence measures

When modeling dependence between random variables measures to quantify the latter are needed.

Pearson correlation

The most popular scalar dependence measure for bivariate data is the Pearson correlation coefficient defined by

$$\rho_{i,j} \coloneqq \rho\left(X_i, X_j\right) = \frac{\operatorname{Cov}\left(X_i, X_j\right)}{\sqrt{\operatorname{Var}\left(X_i\right)}\sqrt{\operatorname{Var}\left(X_j\right)}} \tag{2.2}$$

for two random variables X_i and X_j. An important extension needed later in this thesis are the corresponding partial correlation coefficients. We consider a random vector $\boldsymbol{X}_{\mathcal{I}} \coloneqq (X_1, \ldots, X_d)$, $d \geq 2$, with zero mean, where $\mathcal{I}$ is the index set $\{1, \ldots, d\}$. Further, we consider a subset $L \subseteq \mathcal{I}$ having at least cardinality 2, i.e. $|L| \geq 2$. For a pair (i, j) $(i, j \in L, i \neq j)$ we denote L with the subset $\{i, j\}$ removed by $D_{\{i,j\}} \coloneqq L_{-\{i,j\}} = L \backslash \{i, j\}$ and the corresponding random vector by $\boldsymbol{X}_{D_{\{i,j\}}} \coloneqq \{X_k, k \in D_{\{i,j\}}\}$. The partial regression coefficient $b_{i,j;D_{\{i,j\}}}$ is defined as the quantity that minimizes

$$\mathbb{E}\big[(X_i - \sum_{j \in L_{-\{i\}}} b_{i,j;D_{\{i,j\}}} X_j)^2\big].$$

The corresponding partial correlation coefficient $\rho_{i,j;D_{\{i,j\}}}$ quantifies the dependence between X_i and X_j without the linear effect of $\boldsymbol{X}_{D_{\{i,j\}}}$ and is defined by (Kurowicka and Joe, 2011, p. 47)

$$\rho_{i,j;D_{\{i,j\}}} \coloneqq \operatorname{sgn}(b_{i,j;D_{\{i,j\}}}) \left(b_{i,j;D_{\{i,j\}}} b_{j,i;D_{\{i,j\}}}\right)^{1/2}.$$

While standard and partial correlation coefficients are not invariant with respect to monotone transformations of the data, this is the case for the two subsequent dependence measures.

Kendall's τ

First, Kendall's τ as a global dependence measure is introduced (Kendall, 1938). For independent and identically distributed random vectors (X_i, X_j) and $(\tilde{X}_i, \tilde{X}_j)$ Kendall's τ is defined by

$$\tau_{i,j} \coloneqq \tau\left(X_i, X_j\right) = \mathbb{P}\{(X_i - \tilde{X}_i)(X_j - \tilde{X}_j) > 0\} - \mathbb{P}\{(X_i - \tilde{X}_i)(X_j - \tilde{X}_j) < 0\}.$$

One can show that

$$\tau_{i,j} = 4 \int_{[0,1]^2} \mathbb{C}_{i,j}\left(u_1, u_2\right) \mathrm{d}\mathbb{C}_{i,j}\left(u_1, u_2\right) - 1,$$

where $\mathbb{C}_{i,j}$ is the bivariate copula corresponding to the joint distribution of (X_i, X_j). Thus, Kendall's τ is a property of the underlying copula and therefore does not depend on the marginal

distribution functions. In particular, for the parametric copula families, that will be introduced in Section 2.2.3 and used throughout this thesis, there is a one-to-one relationship between the Kendall's τ value and the specific copula parameter. This is of particular interest when comparing the dependence strength among different parametric copulas fitted to data at hand.

Tail-dependence

Copulas further allow to investigate tail-dependence. The latter is a local dependence measure capturing extremal behavior, i.e. the dependence for joint very small (lower tail-dependence) and joint very large (upper tail-dependence) observations. For random variables $X_i \sim F_i$ and $X_j \sim F_j$ with associated bivariate copula $\mathbb{C}_{i,j}$ we define the upper tail-dependence coefficient by

$$\lambda_{i,j}^{\mathrm{U}} \coloneqq \lim_{u \nearrow 1} \mathbb{P}\{X_i > F_i^{-1}(u) \,|X_j > F_j^{-1}(u)\} = \lim_{u \nearrow 1} \frac{1 - 2u + \mathbb{C}_{i,j}(u,u)}{1-u}$$

given that the limit exists. The lower tail-dependence coefficient is defined by

$$\lambda_{i,j}^{L} \coloneqq \lim_{u \searrow 0} \mathbb{P}\{X_i < F_i^{-1}(u) \,|X_j < F_j^{-1}(u)\} = \lim_{u \searrow 0} \frac{\mathbb{C}_{i,j}(u,u)}{u}$$

given that the limit exists. Clearly, tail-dependence as well is a property of the underlying copula.

2.2.3 Parametric copula families

To adequately model data at hand a wide selection of copula families, which covers a broad range of Kendall's τ values and different tail-dependence behavior, is needed.

The independence copula $\prod$ is a straightforward nonparametric example. It is given by

$$\prod(u_1, \ldots, u_d) = \prod_{j=1}^{d} u_j$$

with constant copula density equal to 1. Except for the independence copula only parametric copula families will be considered in this thesis, i.e. a parametric form of the copula underlying the data will be assumed.

The most popular elliptical representative is the Gaussian copula, which is constructed from the Gaussian distribution using the inversion of Sklar's Theorem in (2.1):

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

With Φ the cumulative distribution function of $\mathcal{N}(0,1)$ and Φ_Σ corresponding to a d-dimensional Gaussian distribution with zero mean, unit variances and correlation matrix Σ the d-dimensional Gaussian copula is defined by

$$\mathbb{C}^{\mathrm{Gauss}}(u_1, \ldots, u_d) = \Phi_\Sigma\{\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)\}.$$

Note that the Gaussian copula exhibits neither lower nor upper tail-dependence.

Table 2.1: Popular bivariate Archimedean copulas with the range of their dependence parameter θ, the formula of ϕ, the corresponding Kendall's τ value and the tail-dependence coefficients λ^{U} and λ^{L}.

	Clayton	Gumbel	Frank
$\theta \in$	$(0, \infty)$	$[1, \infty)$	$(-\infty, \infty) \setminus \{0\}$
$\phi(s)$	$(1+\theta s)^{-1/\theta}$	$e^{-s^{1/\theta}}$	$-\frac{1}{\theta}\ln\{1-(1-e^{-\theta})e^{-s}\}$
$\mathbb{C}(u_1, u_2)$	$(u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}}$	$e^{\left[-\{(-\ln u_1)^\theta + (-\ln u_2)^\theta\}^{\frac{1}{\theta}}\right]}$	$-\frac{1}{\theta}\ln\left\{1 + \frac{(e^{-\theta u_1}-1)(e^{-\theta u_2}-1)}{e^{-\theta}-1}\right\}$
τ	$\tau = \frac{\theta}{\theta+2}$	$\tau = 1 - \frac{1}{\theta}$	$\tau = 1 - \frac{4}{\theta} + 4\frac{D_1(\theta)}{\theta}$ with $D_1(\theta) = \int_0^\theta \frac{t/\theta}{e^t-1}dt$
$\tau \in$	$[0, 1]$	$[0, 1]$	$[-1, 1]$
λ^{U}	$\lambda^{\mathrm{U}} = 0$	$\lambda^{\mathrm{U}} = 2 - 2^{1/\theta}$	$\lambda^{\mathrm{U}} = 0$
λ^{L}	$\lambda^{\mathrm{L}} = 2^{-1/\theta}$	$\lambda^{\mathrm{L}} = 0$	$\lambda^{\mathrm{L}} = 0$

Archimedean copulas constitute another popular copula class. They are given by

$$\mathbb{C}(u_1, \ldots, u_d) = \phi\{\phi^{-1}(u_1) + \ldots + \phi^{-1}(u_d)\}, \tag{2.3}$$

where $\phi : [0, \infty[\to [0, 1]$ is a continuous strictly decreasing function with $\phi(0) = 1$ and $\phi(\infty) = 0$ that satisfies the complete monotonicity condition (Joe, 1997; Nelsen, 2006), i.e. the derivatives of ϕ must alternate in sign. From (2.3) it follows that an Archimedean copula is fully determined by the choice of ϕ. Thus, a restrictive dependence structure is implied. For example, all marginal copulas show exactly the same type and strength of association. Note that for the same global dependence as expressed by Kendall's τ, different Archimedean copulas can exhibit diverse local dependence: a Clayton copula is lower tail-dependent, a Gumbel copula is upper tail-dependent and a Frank copula shows no tail-behavior. For $d = 2$, details are listed in Table 2.1.

Note that for example the bivariate Clayton and Gumbel copula only allow for positive dependence as expressed by Kendall's τ. Model flexibility, however, can be extended by considering reflected forms of these copula families. More precisely, according to Joe (1993) for the counter-clockwise rotated equivalents of a bivariate copula $\mathbb{C}$ with copula density $\mathbb{c}$ we have

- 90 degree: $\mathbb{C}^{90}(u_1, u_2) \coloneqq u_2 - \mathbb{C}(1-u_1, u_2)$
  with $\mathbb{c}^{90}(u_1, u_2) \coloneqq \mathbb{c}(1-u_1, u_2)$,
- 180 degree: $\mathbb{C}^{180}(u_1, u_2) \coloneqq u_1 + u_2 - 1 + \mathbb{C}(1-u_1, 1-u_2)$
  with $\mathbb{c}^{180}(u_1, u_2) \coloneqq \mathbb{c}(1-u_1, 1-u_2)$,
- 270 degree: $\mathbb{C}^{270}(u_1, u_2) \coloneqq u_1 - \mathbb{C}(u_1, 1-u_2)$
  with $\mathbb{c}^{270}(u_1, u_2) \coloneqq \mathbb{c}(u_1, 1-u_2)$.

2.3 R-vine copula models

R-vine distributions combine R-vines as introduced in Section 2.1 and copula theory from Section 2.2. They are also referred to as pair-copula constructions, since they assign to each of the $d\,(d-1)\,/2$ edges of a d-dimensional R-vine structure a bivariate unconditional copula (in tree $\mathcal{T}_1$) or a bivariate conditional copula (in trees $\mathcal{T}_2$ to $\mathcal{T}_{d-1}$). We consider the copula data $(U_1,\ldots,U_d)$ corresponding to the random vector $(X_1,\ldots,X_d)$ with marginal distribution functions F_j $(j=1,\ldots,d)$, i.e. $U_j = F_j\,(X_j)$. Since in this case the marginals of the underlying data are uniform, we speak of an R-vine copula. Following Czado (2010) and using the notation introduced in Section 2.1, the d-dimensional R-vine copula density based on the R-vine structure $\mathcal{V}_d$ with edge set $E\,(\mathcal{V}_d) = E_1 \cup \cdots \cup E_{d-1}$ can be written as

$$\mathbb{c}\,(u_1,\ldots,u_d) = \prod_{\ell=1}^{d-1}\prod_{e\in E_\ell} \mathbb{c}_{a_e,b_e;D_e}\{\mathbb{C}_{a_e|D_e}\,(u_{a_e}|\boldsymbol{u}_{D_e})\,,\mathbb{C}_{b_e|D_e}\,(u_{b_e}|\boldsymbol{u}_{D_e})\,;\boldsymbol{u}_{D_e}\}, \tag{2.4}$$

where

- the indices a_e and b_e correspond to the conditioned variables and D_e represents the conditioning set of edge e.
- $\mathbb{c}_{a_e,b_e;D_e}\,(\cdot,\cdot;\boldsymbol{u}_{D_e})$ denotes the copula density corresponding to the conditional distribution of (U_{a_e},U_{b_e}) given $\boldsymbol{U}_{D_e} = \boldsymbol{u}_{D_e}$ with $\boldsymbol{U}_{D_e}$ the vector containing all variables corresponding to the conditioning set D_e. The corresponding copula will be denoted by $\mathbb{C}_{a_e,b_e;D_e}\,(\cdot,\cdot;\boldsymbol{u}_{D_e})$.
- $\mathbb{C}_{a_e|D_e}\,(\cdot|\boldsymbol{u}_{D_e})$ denotes the conditional distribution of U_{a_e} given $\boldsymbol{U}_{D_e} = \boldsymbol{u}_{D_e}$.

Given the large number of R-vine structures and given that the pair-copulas corresponding to each edge of the underlying R-vine structure can be chosen and combined arbitrarily (for example from the parametric copula families presented in Section 2.2.3) R-vine copulas clearly constitute a highly flexible class of dependence models.

Throughout this thesis, we assume that in (2.4) the conditional pair-copula densities $\mathbb{c}_{a_e,b_e;D_e}$ in trees $\mathcal{T}_\ell$ $(\ell=2,\ldots,d-1)$ do not depend on the conditioning vector $\boldsymbol{u}_{D_e}$. Their arguments $\mathbb{C}_{a_e|D_e}\,(u_{a_e}|\boldsymbol{u}_{D_e})$ and $\mathbb{C}_{b_e|D_e}\,(u_{b_e}|\boldsymbol{u}_{D_e})$ indeed do depend on $\boldsymbol{u}_{D_e}$. For details on this simplifying assumption, see Hobæk Haff et al. (2010) and Stöber et al. (2013).

Joe (1997) provides the important result for pair-copula constructions that the conditional distributions $\mathbb{C}_{a_e|D_e}\,(\cdot|\boldsymbol{u}_{D_e})$ and $\mathbb{C}_{b_e|D_e}\,(\cdot|\boldsymbol{u}_{D_e})$, subsequently abbreviated as $\mathbb{C}_{a|D}\,(\cdot|\boldsymbol{u}_D)$ and $\mathbb{C}_{b|D}\,(\cdot|\boldsymbol{u}_D)$, can be evaluated using only the pair-copulas specified in lower tree levels of the underlying R-vine structure. Define for $i\in\{a,b\}$ the set $D_{+i} \coloneqq D\cup\{i\}$. Then,

$$\mathbb{C}_{a|D_{+b}}\left(u_a|\boldsymbol{u}_{D_{+b}}\right) = h_{a|b;D}\{\mathbb{C}_{a|D}\,(u_a|\boldsymbol{u}_D)\,\big|\mathbb{C}_{b|D}\,(u_b|\boldsymbol{u}_D)\}$$

and

$$\mathbb{C}_{b|D_{+a}}\left(u_b|\boldsymbol{u}_{D_{+a}}\right) = h_{b|a;D}\{\mathbb{C}_{b|D}\,(u_b|\boldsymbol{u}_D)\,\big|\mathbb{C}_{a|D}\,(u_a|\boldsymbol{u}_D)\},$$

where

$$h_{a|b;D}\{\mathbb{C}_{a|D}\left(u_a|\boldsymbol{u}_D\right)|\mathbb{C}_{b|D}\left(u_b|\boldsymbol{u}_D\right)\} \coloneqq \frac{\partial}{\partial u}\mathbb{C}_{a,b;D}\{\mathbb{C}_{a|D}\left(u_a|\boldsymbol{u}_D\right), u\}\Big|_{u=\mathbb{C}_{b|D}(u_b|\boldsymbol{u}_D)} \tag{2.5}$$

and

$$h_{b|a;D}\{\mathbb{C}_{b|D}\left(u_b|\boldsymbol{u}_D\right)|\mathbb{C}_{a|D}\left(u_a|\boldsymbol{u}_D\right)\} \coloneqq \frac{\partial}{\partial u}\mathbb{C}_{a,b;D}\{u, \mathbb{C}_{b|D}\left(u_b|\boldsymbol{u}_D\right)\}\Big|_{u=\mathbb{C}_{a|D}(u_a|\boldsymbol{u}_D)} \tag{2.6}$$

are the so-called h-functions corresponding to the pair-copula $\mathbb{C}_{a,b;D}$. Clearly, the arguments of the h-functions in (2.5) and (2.6) can again be expressed in terms of h-functions such that a recursive representation of $\mathbb{C}_{a|D}\left(u_a|\boldsymbol{u}_D\right)$ and $\mathbb{C}_{b|D}\left(u_b|\boldsymbol{u}_D\right)$ in terms of lower tree pair-copulas is obtained.

R-vine copulas have been extensively studied in the recent years including the development of comprehensive statistical software. In this thesis, all implementations are done in the programming language `R` (R Core Team, 2017) using and extending methods available in the `VineCopula` package (Schepsmeier et al., 2017).

Chapter 3

R-vine based modeling of multivariate volatility time-series

This chapter is based on the research article Barthel et al. (2018a).

3.1 Introduction

The increasing availability of high-frequency data makes volatility modeling and forecasting to one of the most vividly discussed topics in financial econometrics. Also, the strongly increasing interaction and interconnectedness between financial markets have stimulated the need for reliable modeling and forecasting techniques to capture the cross-sectional and temporal dependencies of financial asset returns. Especially during negative economic phases and periods of financial turmoil, assets become more dependent and linkages between asset market volatility tighten (Cappiello et al., 2006). This affects fields such as asset pricing, portfolio allocation and evaluation of risk.

High-frequency data allow to consistently estimate ex-post realized volatility and realized covariances using the sum of squared intra-day returns (Doléans-Dade and Meyer, 1970; Jacod, 1994). By making naturally latent variables, namely volatilities and covariances, observable and measurable, standard time-series approaches can be applied to model their realized counterparts. Building upon the aforementioned classical estimator first used in the context of high-frequency data by Barndorff-Nielsen and Shephard (2004), many refinements were investigated to improve its overall quality and precision (Zhang, 2011), to reduce market microstructure noise (Gençay et al., 2001; Zhang et al., 2005) and to take into account jumps (Christensen et al., 2010) and asynchronicity (Hayashi et al., 2005).

The main modeling challenge when developing prediction tools for realized covariance matrices are the algebraic restrictions of symmetry and positive definiteness the forecasts need to satisfy. Direct modeling of the components using univariate time-series models does not meet this constraint (Andersen et al., 2006) and neglects for example dynamic volatility spillovers among the series of variances and covariances (Voev, 2008). Several multivariate approaches such as the Wishart Autoregressive (WAR) model (Gouriéroux et al., 2009) and its dynamic counterpart the Conditional Autoregressive Wishart (CAW) model (Golosnoy et al., 2012) have been developed. Andersen et al. (2006) propose a multivariate generalization of the realized GARCH model

(Hansen et al., 2012) by modifying the Dynamic Conditional Correlation (DCC) model of Engle (2002). The basic idea of the latter model is to split up the estimation problem into the two simpler tasks of modeling the conditional volatilities and the correlation dynamics. Halbleib and Voev (2014) adopt this strategy using high-frequency data in the volatility part and daily data in the correlation part at the expense of less flexible correlation specifications. As an alternative, data transformation is one of the most frequently used approaches. Bauer and Vorkink (2011) apply the matrix logarithm function and a factor model approach to the individual components, which, however, leads to a computationally demanding model. First proposed by Andersen et al. (2003) and having evolved to one of the standard ways to proceed, the Cholesky decomposition is a proven tool to guarantee symmetry and positive definiteness of the forecasts. For example, Chiriac and Voev (2011) decompose the series of realized covariance matrices via the Cholesky factorization and model the so-obtained series of Cholesky elements with a vector autoregressive fractionally integrated moving average (VARFIMA) process. Brechmann et al. (2018) build upon this model approach, but pay special attention to the specific dependencies among the Cholesky series induced by the nonlinear data transformation. While Cholesky decomposition based models are straightforward and easy to implement, they also come with drawbacks. There is no clear interpretation of the model components obtained after data transformation and the latter induces an additive bias in the forecasts of the original data due to its nonlinear nature. Also, the Cholesky decomposition depends on the ordering of the data within the realized covariance matrices with no obvious way to fix the order in advance. Complete enumeration leads to a computationally expensive estimation problem. On the other hand, fixing the order upfront ignores a possible changing behavior of the data over time.

Irrespective of the considered data transformation, multivariate approaches for time-series modeling often suffer from lacking flexibility in the parameters. Further, they barely allow for convenient modeling of non-Gaussianity and conditional heteroscedasticity, which, however, are typical features of volatility data. In comparison, univariate time-series models allow for various extensions and refinements to tackle these problems. Besides ARFIMA processes (Andersen et al., 2006), heterogeneous autoregressive (HAR) processes are most commonly applied to (log-transformed) realized volatility time-series capturing their long-memory behavior. They include volatility measured over different time horizons and account for multifractal scaling (Corsi, 2009). Both ARFIMA and HAR models can be extended by GARCH augmentations to account for non-Gaussianity and volatility clustering (Corsi et al., 2008). By considering skewed error distributions for the residuals, typically observed high skewness and kurtosis can be additionally captured (Bai et al., 2003; Fernández and Steel, 1998).

In the light of the above discussion, a tool to transform the realized covariance matrices, which allows for reasonable computational effort, interpretability of the model components obtained after data transformation and to exploit the beneficial features of univariate time-series modeling, is desirable. A promising candidate which meets these requirements are partial correlation vines. The latter assign partial correlations to the edges of an R-vine tree structure. The latter is a graph theoretical object first proposed by Bedford and Cooke (2002), which consists of a set of linked trees specifying bivariate conditional constraints. The set of standard and partial

correlations specified through an R-vine structure has attractive properties. Bedford and Cooke (2002) proof that there is a bijection between the specified (partial) correlations and the set of symmetric and positive definite correlation matrices. Further, Kurowicka and Cooke (2003) find that any partial correlation vine specifies algebraically independent (partial) correlations, i.e. the latter can take arbitrary values in $(-1, 1)$ while still guaranteeing positive definiteness of the corresponding correlation matrix. This result advocates partial correlation vines to be a useful tool in several applications. Kurowicka and Cooke (2006a) use them to solve the completion problem for positive definite matrices, whereas Lewandowski et al. (2009) introduce a method to uniformly generate random correlation matrices from the space of positive definite correlation matrices. Brechmann and Joe (2014) base a parsimonious parameterization of correlation matrices on partial correlation vines in combination with factor analysis and Brechmann and Joe (2015) use these findings to capture the dependence structure in multivariate data. Considering financial data, Poignard (2017) introduces a vine-GARCH approach as flexible multivariate GARCH-type model, which parametrizes the latent correlations appearing in the DCC model of Engle (2002) in terms of a partial correlation vine. Based on the specific nature of an R-vine tree structure, their estimation technique proceeds iteratively by evoking only bivariate GARCH models in each tree level and thus allows for dimension reduction as compared to computationally highly demanding classical multivariate GARCH models.

To our knowledge, data transformation using partial correlation vines has not yet been investigated to model and forecast multivariate realized volatility time-series. We propose a joint estimation and prediction model of the realized variance time-series and a subset of realized standard and partial correlation time-series. The latter are obtained after transforming the series of realized correlation matrices based on an R-vine structure as first step of the model approach. To select among the large number of possible R-vine structures the one used for data transformation, we propose a selection method, which exclusively relies on historical information of the modeled time-series and thus dynamically adapts to changing data behavior over time. We will show that data transformation based on this R-vine structure further allows for parsimony in the resulting multivariate time-series models, which are to be estimated as second step of the model approach. We opt for a copula based time-series model to exploit the beneficial features of elaborate univariate time-series models. By considering flexible copulas for the dependence between the model components possible asymmetry and nonlinearity can be captured. Combining in a third step the predicted realized variances and the predicted realized correlation matrix obtained after back-transformation of the underlying realized partial correlation vine guarantees a symmetric and positive definite realized covariance matrix forecast.

The paper is structured as follows. In Section 3.2, we introduce partial correlation vines combining the notion of partial correlations (Section 2.2.2) and an R-vine structure (Section 2.1). The transformation of a correlation matrix to a partial correlation vine based on a given R-vine structure and vice versa is explained in detail. In Section 3.3, we introduce the general data setting and motivate the choice of Cholesky decomposition based models as our main benchmarks. In Section 3.4, we outline in detail the three main steps of the proposed partial correlation vine data transformation approach including R-vine structure selection in Section 3.4.2 and multi-

variate time-series modeling in Section 3.4.3. Supported by the analysis of high-frequency data for six stocks listed on the NYSE, AMEX and NASDAQ beneficial properties of the proposed modeling strategy will be explored. In Section 3.5, detailed investigation of the real data example will be continued. Section 3.5.4 shows the excellent forecasting performance of the partial correlation vine data transformation approach both with respect to statistical precision and mean-variance balance in portfolio optimization.

3.2 Partial correlation vines

First, we provide necessary background on the two main ingredients of the proposed model approach – partial correlations and regular vines.

Partial correlations

We consider a random vector $\boldsymbol{X}_{\mathcal{I}} \coloneqq (X_1, \ldots, X_d)$, $d \geq 2$, with zero mean, where $\mathcal{I}$ is the index set $\{1, \ldots, d\}$. We denote the $d \times d$ covariance matrix by $\boldsymbol{Y}$ and obtain the corresponding $d \times d$ correlation matrix $\boldsymbol{R}$ as $\boldsymbol{R} = \boldsymbol{D}^{-1/2}\boldsymbol{Y}\boldsymbol{D}^{-1/2}$, where $\boldsymbol{D} = \operatorname{diag}(y_{1,1}, \ldots, y_{d,d})$ is the diagonal matrix of variances. In Section 2.2.2 on page 10, we introduced for a subset $L \subseteq \mathcal{I}$ with $|L| \geq 2$ the partial correlation coefficients $\rho_{i,j;D_{\{i,j\}}}$ ($i, j \in L$, $i \neq j$ and $D_{\{i,j\}} = L_{-\{i,j\}} = L \backslash \{i, j\}$) that quantify the dependence between X_i and X_j with the linear effect of $\boldsymbol{X}_{D_{\{i,j\}}} = \{X_k, k \in D_{\{i,j\}}\}$ removed (Kurowicka and Joe, 2011, p. 47).

In the following, we refer to the cardinality of $D_{\{i,j\}}$ as order of the corresponding partial correlation coefficient. For order zero, i.e. $|L| = |\{i, j\}| = 2$ and thus $D_{\{i,j\}} = \emptyset$, we obtain pairwise standard correlations between X_i and X_j ($i, j \in \mathcal{I}, i \neq j$). Then, as in (2.2) we write $\rho_{i,j;\emptyset} = \rho_{i,j}$. Now, consider for a subset $L \subseteq \mathcal{I}$ of at least cardinality 3, a set of distinct indices $\{i, j, k\} \subseteq L$ ($i \neq j \neq k$). We define $\tilde{D} \coloneqq L_{-\{i,j,k\}}$ such that $D_{\{i,j\}} = \tilde{D} \cup k$. Anderson (1958) derives a formula to recursively calculate the partial correlations of any order $|D_{\{i,j\}}|$ with $|D_{\{i,j\}}| \geq 1$ in terms of (partial) correlations of lower order. With $\rho^2_{i,k;\tilde{D}} < 1$ and $\rho^2_{j,k;\tilde{D}} < 1$ it holds that

$$\rho_{i,j;D_{\{i,j\}}} = \frac{\rho_{i,j;\tilde{D}} - \rho_{i,k;\tilde{D}}\rho_{j,k;\tilde{D}}}{\sqrt{1 - \rho^2_{i,k;\tilde{D}}}\sqrt{1 - \rho^2_{j,k;\tilde{D}}}}. \tag{3.1}$$

Since the evaluation of higher order partial correlations gets too involved when exclusively relying on this recursion formula, in practice typically a more efficient calculation procedure is used (see Whittaker (2009)). Let $\boldsymbol{\Omega}$ be the submatrix of standard correlations with indices $L \subseteq \mathcal{I}$, i.e. $\boldsymbol{\Omega} = (\omega_{k,\ell})_{k,\ell=1,\ldots,|L|} = (\rho_{l_k l_\ell})_{k,\ell=1,\ldots,|L|}$, where l_k is the k-th element in L. Let $\boldsymbol{P}$ be its inverse, i.e. $\boldsymbol{P} = \boldsymbol{\Omega}^{-1} = (p_{k,\ell})_{k,\ell=1,\ldots,|L|}$. Then, it holds

$$\rho_{l_k,l_\ell;D_{\{l_k,l_\ell\}}} = -\frac{p_{k,\ell}}{\sqrt{p_{k,k}p_{\ell,\ell}}}. \tag{3.2}$$

Thus, through inversion of $\boldsymbol{\Omega}$ all partial correlations between X_i and X_j ($i, j \in L$, $i \neq j$) given all other variables $\boldsymbol{X}_{D_{\{i,j\}}}$ are simultaneously calculated. If interest is in a single partial correlation

$\rho_{i,j;D_{\{i,j\}}}$ for $i,j \in L$ with $i \neq j$ fixed, computing complexity can be reduced by assorting $\boldsymbol{\Omega}$ blockwise with indices (i,j) and $D_{\{i,j\}}$, i.e.

$$\boldsymbol{\Omega}^{-1} = \begin{pmatrix} \boldsymbol{\Omega}_{1,1} & \boldsymbol{\Omega}_{1,2} \\ \boldsymbol{\Omega}_{2,1} & \boldsymbol{\Omega}_{2,2} \end{pmatrix}^{-1} = \boldsymbol{P} = \begin{pmatrix} \boldsymbol{P}_{1,1} & \boldsymbol{P}_{1,2} \\ \boldsymbol{P}_{2,1} & \boldsymbol{P}_{2,2} \end{pmatrix},$$

where $\boldsymbol{\Omega}_{1,1}$ is a 2×2 matrix with elements $\omega_{1,1} = \omega_{2,2} = 1$, $\omega_{1,2} = \omega_{2,1} = \rho_{i,j}$ and $\boldsymbol{P}_{1,1}$ its counterpart with elements $p_{1,1}, p_{1,2} = p_{2,1}, p_{2,2}$. Using standard results for block matrix inversion (see Bernstein (2005)) we have $\boldsymbol{P}_{1,1}^{-1} = \boldsymbol{\Omega}_{1,1} - \boldsymbol{\Omega}_{1,2}\boldsymbol{\Omega}_{2,2}^{-1}\boldsymbol{\Omega}_{2,1}$ with elements $\tilde{p}_{1,1}$, $\tilde{p}_{1,2} = \tilde{p}_{2,1}$, $\tilde{p}_{2,2}$. We conclude that

$$\rho_{i,j;D_{\{i,j\}}} \overset{(3.2)}{=} -\frac{p_{1,2}}{\sqrt{p_{1,1}p_{2,2}}} = -\frac{-\frac{1}{\det \boldsymbol{P}_{1,1}}\tilde{p}_{1,2}}{\sqrt{\frac{1}{\det \boldsymbol{P}_{1,1}}\tilde{p}_{1,1}\frac{1}{\det \boldsymbol{P}_{1,1}}\tilde{p}_{2,2}}} = \frac{\tilde{p}_{1,2}}{\sqrt{\tilde{p}_{1,1}\tilde{p}_{2,2}}}. \tag{3.3}$$

From now on, we refer to $\mathcal{C}_d$ as the set of all standard correlations and to $\mathcal{C}_d^{\mathrm{p}}$ as the set of all pairwise standard and partial correlations. The $1 \times \binom{d}{2}$ vector $\boldsymbol{P}_{\mathcal{C}_d}$ and the $1 \times \binom{d}{2}2^{d-2}$ vector $\boldsymbol{P}_{\mathcal{C}_d^{\mathrm{p}}}$ record all standard correlations and all standard and partial correlations, respectively, of the random vector $\boldsymbol{X}_{\mathcal{I}}$ in lexicographical order with increasing subset $L \subseteq \mathcal{I}$, i.e.

$$\begin{aligned}
&\boldsymbol{P}_{\mathcal{C}_d} \coloneqq \left(\rho_{1,2}, \ldots, \rho_{1,d}, \rho_{2,3}, \ldots, \rho_{2,d}, \ldots, \rho_{(d-1),d}\right) \\
&\text{and} \\
&\boldsymbol{P}_{\mathcal{C}_d^{\mathrm{p}}} \coloneqq (\boldsymbol{P}_{\mathcal{C}_d}, \\
&\qquad \rho_{1,2;3}, \ldots, \rho_{1,2;d}, \rho_{1,3;2}, \ldots, \rho_{1,d;(d-1)}, \rho_{2,3;1}, \ldots, \rho_{(d-1),d;(d-2)}, \\
&\qquad \ldots, \\
&\qquad \rho_{1,2;3,\ldots,d}, \ldots, \rho_{1,d;2,\ldots,(d-1)}, \ldots, \rho_{(d-1),d;1,\ldots,(d-2)}\big).
\end{aligned}$$

To conclude and as illustrated in Figure 3.1, from a $d \times d$ covariance matrix $\boldsymbol{Y}$ the $1 \times d$ vector of variances $\boldsymbol{y}$ and the $d \times d$ correlation matrix $\boldsymbol{R}$ can be obtained. The latter fully determines the vector $\boldsymbol{P}_{\mathcal{C}_d^{\mathrm{p}}}$, which takes values in $(-1,1)$ and collects all $\binom{d}{2}2^{d-2}$ standard and partial correlations. In the following, we will show that the other way round the correlation matrix $\boldsymbol{R}$ can be uniquely determined from only a few elements of $\boldsymbol{P}_{\mathcal{C}_d^{\mathrm{p}}}$, which are selected through a regular vine.

Partial correlation vine

Based on the previous section and Section 2.1, building the bridge between partial correlations determined by a $d \times d$ correlation matrix and an R-vine structure $\mathcal{V}_d$ is straightforward: in a partial correlation vine with R-vine structure $\mathcal{V}_d$ each edge $e = \{a,b\} \in E(\mathcal{V}_d)$ is identified with the partial correlation coefficient $\rho_{C_{e,a},C_{e,b};D_e}$ that coincides with the conditioned and conditioning set specified by e. Thus, to each edge in $\mathcal{V}_d$ a value in $(-1,1)$ is assigned. We define the set of the $\binom{d}{2}$ standard and partial correlations specified by $\mathcal{V}_d$ as $\mathcal{C}(\mathcal{V}_d)$ and denote by $\boldsymbol{P}_{\mathcal{C}(\mathcal{V}_d)}$ the $1 \times \binom{d}{2}$ vector that collects the corresponding values specified by the correlation matrix $\boldsymbol{R}$ in

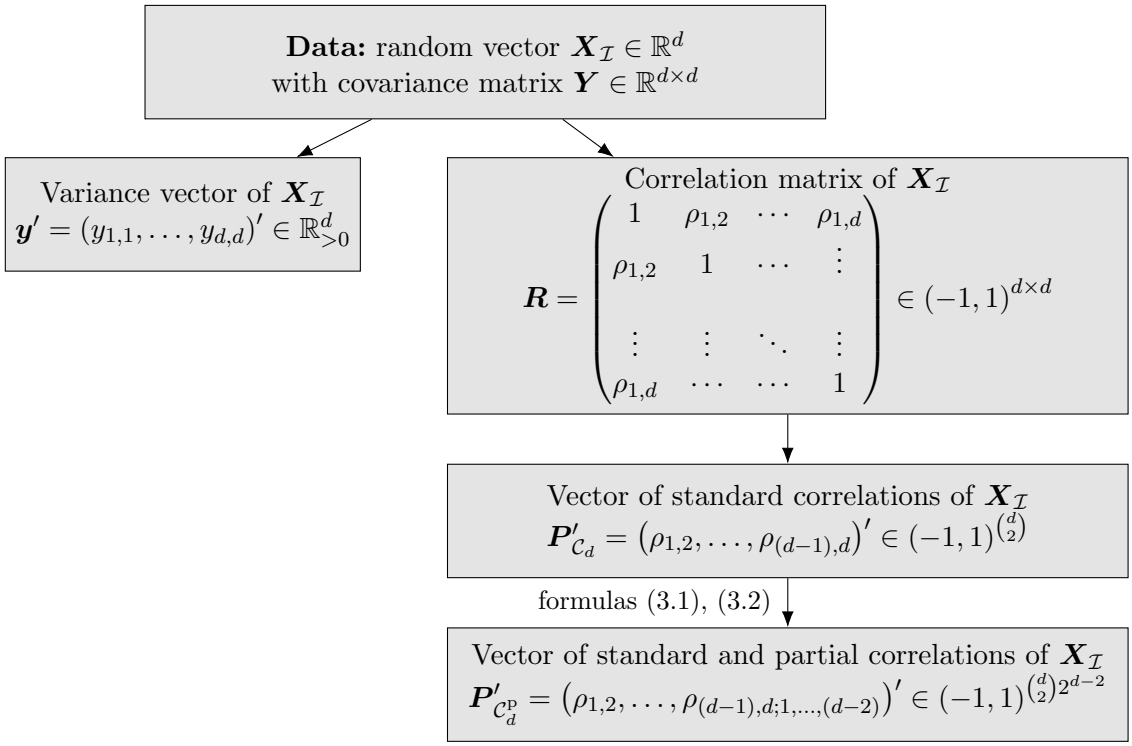


Figure 3.1: Data prespecified by a given covariance matrix $\boldsymbol{Y}$.

lexicographical order.

Bedford and Cooke (2002) provide the fundamental result that for any R-vine structure $\mathcal{V}_d$ there is a one-to-one relationship between the set of $d \times d$ positive definite correlation matrices and its set $\mathcal{C}(\mathcal{V}_d)$, i.e. for each R-vine structure $\mathcal{V}_d$ there exists a bijection

$$F_{\text{Cor2PCor}} : (-1,1)^{\binom{d}{2}} \to (-1,1)^{\binom{d}{2}}, \; F_{\text{Cor2PCor}}\left(\boldsymbol{P}_{\mathcal{C}_d}\right) = \boldsymbol{P}_{\mathcal{C}(\mathcal{V}_d)}. \tag{3.4}$$

In particular, according to Kurowicka and Cooke (2003) the elements in $\boldsymbol{P}_{\mathcal{C}(\mathcal{V}_d)}$ are algebraically independent, i.e. for any arbitrary assignment of values in $(-1,1)$ to the edges of R-vine structure $\mathcal{V}_d$ the correlation matrix calculated from $\boldsymbol{P}_{\mathcal{C}(\mathcal{V}_d)}$ using the 'inverse' of (3.4) is positive definite with correlation values in $(-1,1)$ for all off-diagonal elements. An efficient implementation of the bijection F_{Cor2PCor} and its 'inverse' is available in the R-package `VineCopula` (Schepsmeier et al., 2017). Pseudo-code is provided in Joe (2014). Note that while in the derivation of (3.3) and in the following explanations we assume the submatrix of standard correlations $\boldsymbol{\Omega}$ to be assorted blockwise with indices (i,j) and $D_{\{i,j\}}$, Joe (2014) assorts the indices using the order $D_{\{i,j\}}$ and (i,j).

Example 3.1 (Example 2.1 on page 8 continued)**.** *We illustrate the data transformation based on R-vine structure $\mathcal{V}_6$ in Figure 2.1. As illustrated below, each standard correlation in $\boldsymbol{R}$ on the left-hand side is specified in the partial correlation vine corresponding to $\mathcal{V}_6$ through a (partial) correlation of order $\ell - 1$ modeled in tree $\mathcal{T}_\ell$ ($\ell = 1, \ldots, 5$) , i.e.*

$$\boldsymbol{R} = \begin{pmatrix} \rho_{1,2} & \rho_{1,3} & \rho_{1,4} & \rho_{1,5} & \rho_{1,6} \\ & \rho_{2,3} & \rho_{2,4} & \rho_{2,5} & \rho_{2,6} \\ & & \rho_{3,4} & \rho_{3,5} & \rho_{3,6} \\ & & & \rho_{4,5} & \rho_{4,6} \\ & & & & \rho_{5,6} \end{pmatrix} \rightleftarrows \begin{pmatrix} \rho_{1,2} & \rho_{1,3;2} & \rho_{1,4;2,3} & \rho_{1,5;2,3,4} & \rho_{1,6;2} \\ & \rho_{2,3} & \rho_{2,4} & \rho_{2,5;3} & \rho_{2,6} \\ & & \rho_{3,4;2} & \rho_{3,5} & \rho_{3,6;1,2} \\ & & & \rho_{4,5;2,3} & \rho_{4,6;1,2,3} \\ & & & & \rho_{5,6;1,2,3,4} \end{pmatrix}$$

$\mathcal{T}_1$ $\mathcal{T}_2$ $\mathcal{T}_3$ $\mathcal{T}_4$ $\mathcal{T}_5$

Transformation of correlation matrix. *First, we derive from $\boldsymbol{R}$ the partial correlations corresponding to $\mathcal{V}_6$. Thus, proceeding in the illustration of the above matrices is from left to right. While the standard correlations in $\mathcal{T}_1$ can simply be taken from the correlation matrix $\boldsymbol{R}$, the first order partial correlations in $\mathcal{T}_2$ can be calculated using recursion formula* (3.1)*. For example,*

$$\rho_{1,6;2} = \frac{\rho_{1,6} - \rho_{1,2}\rho_{2,6}}{\sqrt{1-\rho_{1,2}^2}\sqrt{1-\rho_{2,6}^2}}.$$

From tree level $\ell = 3$ on, we rely on formula (3.3) *and elementwise calculate the partial correlations specified by $\mathcal{T}_3$ to $\mathcal{T}_5$. For example, for $\rho_{3,6;1,2}$ we set*

$$\boldsymbol{\Omega}_{1,1} = \begin{pmatrix} 1 & \rho_{3,6} \\ \rho_{3,6} & 1 \end{pmatrix}, \ \boldsymbol{\Omega}_{1,2} = \begin{pmatrix} \rho_{1,3} & \rho_{2,3} \\ \rho_{1,6} & \rho_{2,6} \end{pmatrix}, \ \boldsymbol{\Omega}_{2,1} = \begin{pmatrix} \rho_{1,3} & \rho_{1,6} \\ \rho_{2,3} & \rho_{2,6} \end{pmatrix} \text{ and } \boldsymbol{\Omega}_{2,2} = \begin{pmatrix} 1 & \rho_{1,2} \\ \rho_{1,2} & 1 \end{pmatrix}$$

and evaluate $\begin{pmatrix} \tilde{p}_{1,1} & \tilde{p}_{1,2} \\ \tilde{p}_{1,2} & \tilde{p}_{2,2} \end{pmatrix} = \boldsymbol{\Omega}_{1,1} - \boldsymbol{\Omega}_{1,2}\boldsymbol{\Omega}_{2,2}^{-1}\boldsymbol{\Omega}_{2,1}$*. Then, we calculate*

$$\rho_{3,6;1,2} \overset{(3.3)}{=} \frac{\tilde{p}_{1,2}}{\sqrt{\tilde{p}_{1,1}\tilde{p}_{2,2}}}.$$

Back-transformation to correlation matrix. *Now, proceeding in the illustration of the above matrices is from right to left. We proceed treewise. The standard correlations from $\mathcal{T}_1$ can directly be taken. To calculate the standard correlations that correspond to the conditioned sets of the first order partial correlations available in $\mathcal{T}_2$, we use recursion formula* (3.1)*. For example,*

$$\rho_{1,6} = \rho_{1,6;2}\sqrt{1-\rho_{1,2}^2}\sqrt{1-\rho_{2,6}^2} + \rho_{1,2}\rho_{2,6}.$$

Note that due to the proximity condition of an R-vine structure all standard correlations needed for this evaluation are available from the previous step.

From tree level $\ell = 3$ on, we rely on formula (3.3) *to calculate the standard correlations that correspond to the conditioned sets of the partial correlations specified in $\mathcal{T}_3$ to $\mathcal{T}_5$. For example, to obtain $\rho_{3,6}$ we set $\mathbf{\Omega}_{1,2}$, $\mathbf{\Omega}_{2,2}$ and $\mathbf{\Omega}_{2,1}$ as above. Due to the proximity condition all standard correlations to do so are available from previous steps $\ell = 1, 2$. We calculate*

$$\begin{pmatrix} q_{1,1} & q_{1,2} \\ & \\ q_{1,2} & q_{2,2} \end{pmatrix} = \mathbf{\Omega}_{1,2}\mathbf{\Omega}_{2,2}^{-1}\mathbf{\Omega}_{2,1}$$

such that $\tilde{p}_{1,1} = 1 - q_{1,1}$, $\tilde{p}_{1,2} = \tilde{p}_{2,1} = \rho_{3,6} - q_{1,2}$, $\tilde{p}_{2,2} = 1 - q_{2,2}$ and obtain

$$\rho_{3,6} \overset{(3.3)}{=} \rho_{3,6;1,2}\sqrt{(1-q_{1,1})(1-q_{2,2})} + q_{1,2}.$$

To conclude, the set of all standard correlations $\mathcal{C}_d$ can be determined from any set $\mathcal{C}(\mathcal{V}_d)$ specified by a partial correlation vine with R-vine structure $\mathcal{V}_d$. In particular, positive definiteness of the correlation matrix is always guaranteed. Figure 3.2 provides a summary overview of the relationships between the sets $\mathcal{C}_d$, $\mathcal{C}_d^{\mathrm{p}}$ and $\mathcal{C}(\mathcal{V}_d)$.

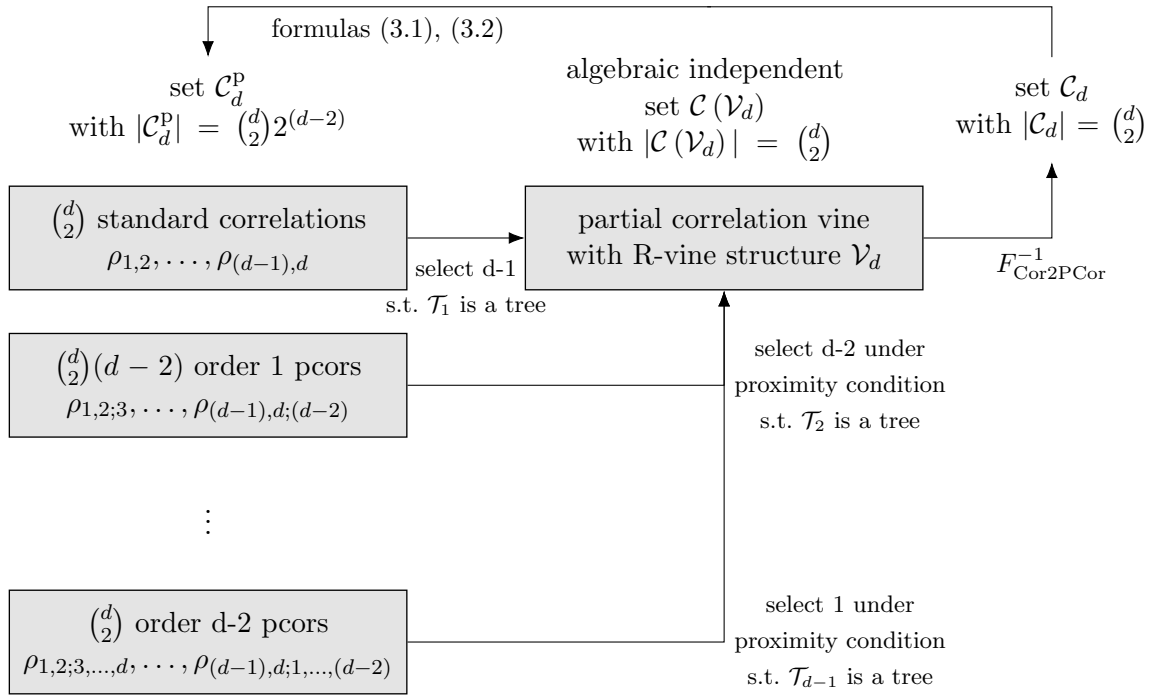


Figure 3.2: Illustration of the transformation of the set of standard correlations $\mathcal{C}_d$ through a partial correlation vine, which consists of a subset of algebraic independent (partial) correlations $\mathcal{C}(\mathcal{V}_d) \subset \mathcal{C}_d^{\mathrm{p}}$ from all standard and partial correlations. The abbreviation "pcor" is used for partial correlation.

3.3 General setting and benchmark models

In the following, partial correlation vine based data transformation will be used to model and forecast multivariate volatility time-series. To do so, we introduce the general data setting first. For the daily price series $\boldsymbol{S}_t \in \mathbb{R}^d$, $t = 1, \ldots, T$, of d assets let $\boldsymbol{r}_t = \log(\boldsymbol{S}_t) - \log(\boldsymbol{S}_{t-1})$ be the $d \times 1$ vector of daily log-returns. The process $\boldsymbol{r}_t$ can be written as

$$\boldsymbol{r}_t = \mathbb{E}[\boldsymbol{r}_t | \mathcal{F}_{t-1}] + \boldsymbol{\epsilon}_t,$$

where $\mathcal{F}_{t-1}$ is the information set containing all information up to and including time point $t-1$. For the innovation term $\boldsymbol{\epsilon}_t$, we suppose that $\boldsymbol{\epsilon}_t = \boldsymbol{\Sigma}_t^{1/2} \boldsymbol{\eta}_t$, where $\boldsymbol{\Sigma}_t = \mathrm{Var}[\boldsymbol{r}_t | \mathcal{F}_{t-1}]$ is the $(d \times d)$-dimensional symmetric and positive definite conditional covariance matrix. For the i.i.d. vector $\boldsymbol{\eta}_t \in \mathbb{R}^d$ it holds that $\mathbb{E}[\boldsymbol{\eta}_t] = 0$ and $\mathrm{Var}[\boldsymbol{\eta}_t] = I_d$. Interest is in modeling and forecasting the series of daily conditional covariance matrices $\boldsymbol{\Sigma}_t$, $t = 1, \ldots, T$, which however are naturally latent variables and therefore are unobservable. Still, as proposed by Barndorff-Nielsen and Shephard (2004) $\boldsymbol{\Sigma}_t$, $t = 1, \ldots, T$, can be specified nonparametrically using the realized covariance matrices as consistent estimates. Considering M intra-day periods per day t, the latter are calculated from high-frequency intra-day log-returns $\boldsymbol{r}_{\ell,t} = \log\left(\boldsymbol{S}_{t-1+\ell/M}\right) - \log\left(\boldsymbol{S}_{t-1+(\ell-1)/M}\right)$ based on the price series $\boldsymbol{S}_{\ell,t} \in \mathbb{R}^d$, $\ell = 1, \ldots, M$. The modeling and forecasting framework is then based on the matrix valued time-series of realized covariance matrices

$$\boldsymbol{Y}_t = \sum_{\ell=1}^{M} \boldsymbol{r}_{\ell,t} \boldsymbol{r}'_{\ell,t}, \quad t = 1, \ldots, T. \tag{3.5}$$

Since for the matrix forecasts symmetry and positive definiteness have to be ensured, algebraic restrictions are imposed on time-series models. Thus, popular modeling strategies avoid direct modeling of the realized covariance matrices considering transformed data instead. Then, the modeling approach basically consists of three consecutive steps: (S1) data transformation of the realized covariance matrices; (S2) multivariate time-series modeling and prediction based on the transformed data; (S3) back-transformation of the transformed data to obtain predictions for the realized covariance matrices, which are proxies for the future conditional covariance matrices.

The novelty in this thesis lies in the use of partial correlation vines for data transformation in steps (S1) and (S3). By modeling and forecasting the time-series of partial correlation vines, we obtain forecasts for the transformed data, which do not have any algebraic restrictions. On the contrary, due to the algebraic independence of the model components positive definiteness of the corresponding predicted correlation matrices is always guaranteed. In the literature, besides the matrix log transformation suggested by Bauer and Vorkink (2011) data transformation based on the Cholesky factorization is one of the most commonly used approaches.

Here, the series of realized covariance matrices $\boldsymbol{Y}_t$, $t = 1, \ldots, T$, is decomposed such that $\boldsymbol{Y}_t = \boldsymbol{C}'_t \boldsymbol{C}_t$, where $\boldsymbol{C}'_t$ is a lower triangular matrix with positive diagonal elements. The Cholesky

elements $c_{i,j;t}$ $(i, j = 1, \ldots, d)$ are recursively calculated by

$$c_{i,j;t} = \begin{cases} \frac{1}{c_{i,i;t}} \left(y_{i,j;t} - \sum_{k=1}^{i-1} c_{k,i;t} c_{k,j;t} \right) & \text{for } i < j, \\ \sqrt{y_{j,j;t} - \sum_{k=1}^{j-1} c_{k,j;t}^2} & \text{for } i = j, \\ 0 & \text{for } i > j. \end{cases} \tag{3.6}$$

By modeling and forecasting the Cholesky elements in step (S2) no parameter restrictions need to be imposed on the multivariate time-series models. Symmetry and positive definiteness of the predicted covariance matrices $\hat{\boldsymbol{Y}}_t$, $t = T+1, T+2, \ldots$, are automatically guaranteed through the back-transformation

$$\hat{y}_{i,j;t} = \sum_{k=1}^{\min\{i,j\}} \hat{c}_{k,i;t} \hat{c}_{k,j;t}. \tag{3.7}$$

Chiriac and Voev (2011) use the Cholesky decomposition in steps (S1) and (S3) and apply a parsimonious VARFIMA model to the multivariate time-series of the Cholesky components in step (S2). In their detailed analysis, they show the superiority of their approach over a variety of competitor models. The comparison includes the above mentioned matrix log transformation used in steps (S1) and (S3) for data transformation combined with VARFIMA and vector HAR models in step (S2). Further, the Wishart autoregressive model of Gouriéroux et al. (2009) as well as the multivariate GARCH model with dynamic conditional correlations of Engle (2002) and its fractionally integrated version proposed by Baillie et al. (1996) are considered. Brechmann et al. (2018) refine the Cholesky-VARFIMA model of Chiriac and Voev (2011) and allow for more flexible modeling of the multivariate time-series in step (S2). They take account of challenging data characteristics in the Cholesky elements by modeling the univariate marginal time-series with elaborate HAR and ARFIMA models including GARCH-augmentations for the residuals. The possibly complex dependence between the Cholesky components is captured by a copula. Given these profound model reviews and comparisons already existing in literature, models based on the Cholesky decomposition will be the main benchmarks in this chapter.

Chiriac and Voev (2011) and Brechmann et al. (2018) both consider high-frequency data from the NYSE TAQ database containing tick-by-tick bid and ask quotes on six stocks listed on the New York Stock Exchange (NYSE), American Stock Exchange (AMEX) and the National Association of Security Dealers Automated Quotation System (NASDAQ). The original raw data were processed by Chiriac and Voev (2011), who provide detailed information on the employed data preparation. Data of the six stocks American Express Inc. (AXP), Citigroup (C), General Electric (GE), Home Depot Inc. (HD), International Business Machines (IBM) and JPMorgan Chase & Co (JPM) were sampled from 9:30 until 16:00 for the period January 1, 2000, until July 30, 2008, i.e. for 2156 trading days. While in (3.5) a single realized covariance matrix is computed from M intra-day log-returns, Chiriac and Voev (2011) obtained for each day a refined subsampled realized covariance matrix, which is more robust to market microstructure noise. For each day t, a 5-minute spaced time grid, i.e. $M = 78$, was shifted by 10 seconds, resulting in 30 distinct sets of realized covariance matrices calculated from 78 intra-day log-returns. By taking

the average of these sets, the subsampled realized covariance matrix for day t was calculated. Although the data are less recent, we consider the same data for comparison reasons. Further, the data cover interesting periods of financial turmoil such as the aftermath of the dotcom bubble in 2000 and the beginning of the financial crisis in 2008. Since the focus in this project is on the novel data transformation defined by partial correlation vines, this will provide new interesting insights about the data.

3.4 Partial correlation vine data transformation approach

In this section, we outline – supported by real data characteristics – steps (S1) to (S3) for the proposed modeling strategy based on partial correlation vines.

3.4.1 Data characteristics

Time-series of realized co(variances) typically exhibit long-memory behavior detectable by high autocorrelations, which decay at a slow rate (see Andersen and Bollerslev (1997); Andersen et al. (2001)). Chiriac and Voev (2011) find that the time-series of Cholesky components obtained through data transformation inherit this data feature. Further, according to Brechmann et al. (2018) appropriate time-series models need to capture non-Gaussianity and volatility clustering of the residuals extracted from the series of Cholesky elements.

In order to also appropriately setup the partial correlation vine data transformation model it is essential to understand the properties of the corresponding model components, namely realized variances and realized (partial) correlations. The latter are specified through the realized covariance matrix via $\boldsymbol{Y}_t = \boldsymbol{D}_t^{1/2}\boldsymbol{R}_t\boldsymbol{D}_t^{1/2}$, $t = 1, \dots, T$. For day t, $\boldsymbol{D}_t = \operatorname{diag}(y_{1,1;t}, \dots, y_{d,d;t})$ contains the realized variances and $\boldsymbol{R}_t$ is the realized correlation matrix. Realized partial correlations can easily be obtained either using recursion formula (3.1) or through simultaneous calculation using (3.2) (recall Figure 3.1). For reasonable time-series modeling later in step (S2), for all model components data on the real line are needed. Thus, we log-transform the all positive realized variance time-series and apply the Fisher z-transformation to the series of (partial) correlations, i.e. for ρ_t being an arbitrary (partial) correlation at day t

$$z(\rho_t) = \frac{1}{2}\log\left(\frac{1+\rho_t}{1-\rho_t}\right), \quad t = 1, \dots, T. \tag{3.8}$$

For the considered real data, Figure 3.3 shows a selection of time-series both on the original (left) and the transformed (right) scale. The first panel illustrates for JPM the daily realized variance series. Striking is the highly volatile behavior particularly during periods of financial turmoil such as the aftermath of the dotcom bubble and the beginning of the financial crisis in August 2007. Panels 2 to 6 show selected daily time-series of realized (partial) correlations with increasing order. For example, the last panel illustrates the time-series of the fourth order realized partial correlation between HD and JPM. For each day, the latter is a proxy of the conditional (with respect to the information set) correlation between the log-returns of IBM and JPM given the four remaining stocks. With increasing order the realized partial correlation time-series

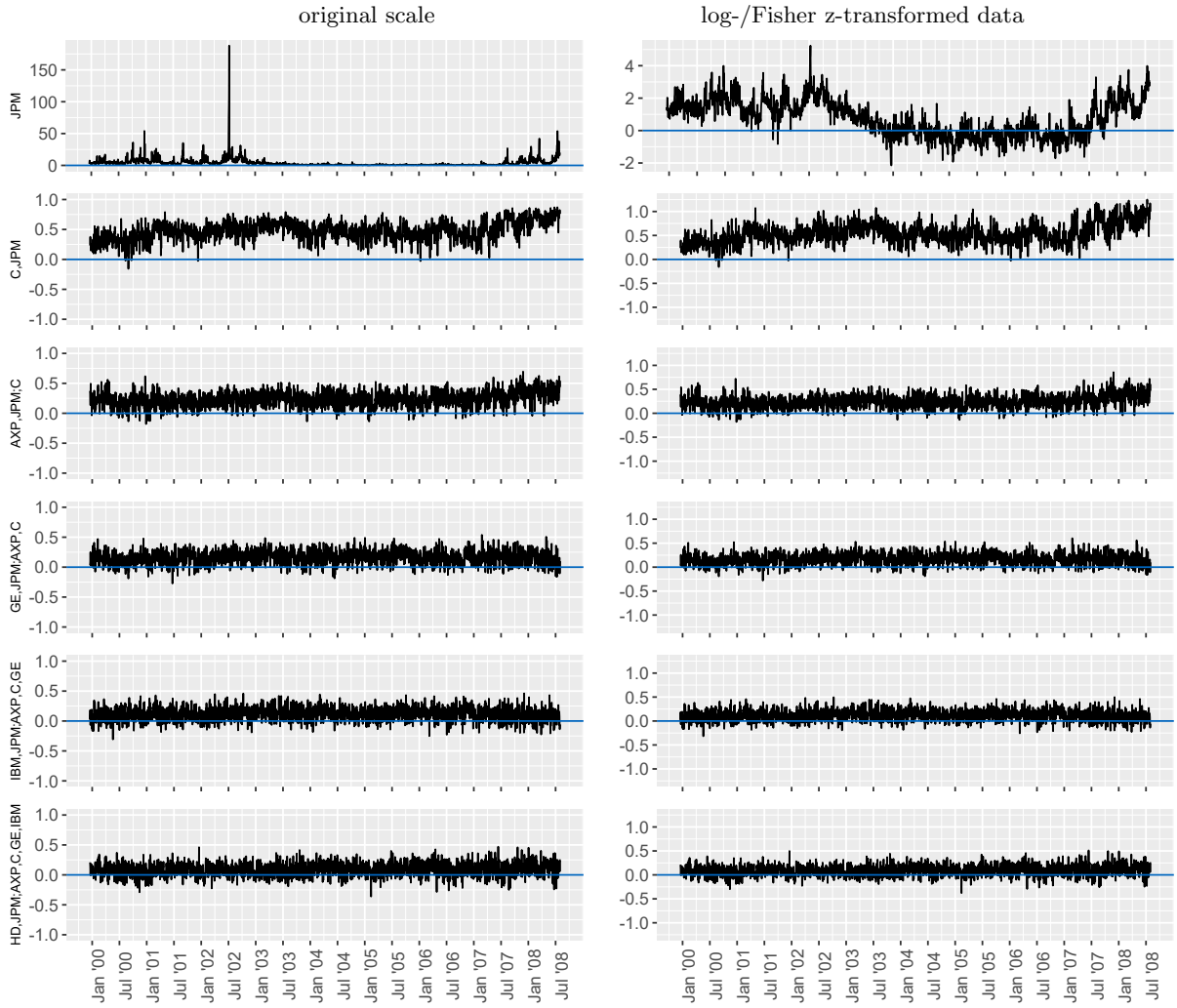


Figure 3.3: Daily realized variance series (1st row) and daily realized (partial) correlation series (2nd – 6th row). Original data are shown in the left panel, log-transformed and Fisher z-transformed data, respectively, are shown in the right panel.

become more stable while still exhibiting highly volatile behavior. Further, in Figure 3.4 data characteristics of four time-series are illustrated. The figures in the top row correspond to the realized variance time-series of JPM, which together with the remaining five realized variance series always will be a model component. The long hyperbolic decay of the autocorrelation function of the squared data on the left confirms the long-memory behavior and the presence of volatility clustering. The log-periodogram shows higher peaks only for short frequencies as expected for self-similar processes. In the second and third row, exemplary time-series, which would appear in tree level $\mathcal{T}_1$ and $\mathcal{T}_4$, respectively, of an R-vine structure are shown. Interestingly, while the realized standard correlation time-series corresponding to tree level $\mathcal{T}_1$ inherits the data characteristics of the realized variance time-series, the latter are less pronounced for the realized third order partial correlation time-series in $\mathcal{T}_4$.

To gain a deeper understanding of this last observation, recall that in a partial correlation

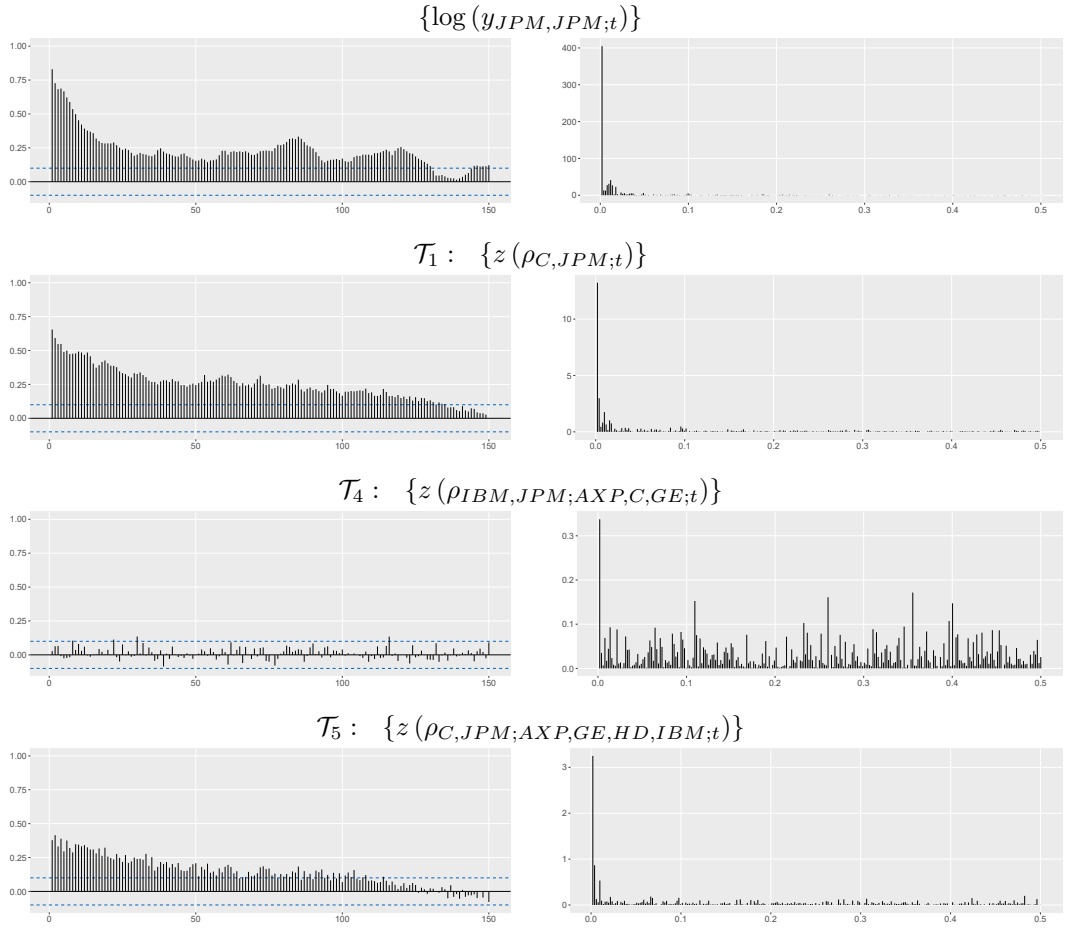


Figure 3.4: Illustration of autocorrelation functions of squared data (left panel) and corresponding log-periodograms (right panel) based on data from July 1, 2006, to June 30, 2008. In the first row, the log-transformed realized variance time-series of JPM is considered. In rows 2 to 4 exemplary Fisher z-transformed realized (partial) correlation time-series of increasing order are considered.

vine each variable pair (i, j) $(i, j \in \{1, \ldots, d\},\ i \neq j)$ forms exactly once the conditioned set of an edge. Thus, depending on the tree level ℓ the proxy for the conditional (with respect to the information set) correlation between the log-returns of stocks i and j is either represented by the realized standard correlation (if (i, j) occurs as conditioned set in $\mathcal{T}_1$) or through a $(\ell-1)$-th order realized partial correlation (if (i, j) occurs as conditioned set in $\mathcal{T}_\ell$ $(\ell = 2, \ldots, d-1)$). In the latter case, the linear effect of the $\ell-1$ stocks forming the conditioning set is removed. Clearly, for some pairs the realized standard correlations might mainly be driven by other variables. Once this influence is removed data features such as long-memory behavior weaken and the corresponding realized partial correlation time-series behave more and more like noise. On the other hand, this effect is not observable for pairs, which truly are strongly correlated such as the log-returns of the two financial stocks C and JPM. The corresponding realized fourth order partial correlation time-series, which would occur in the highest tree level, i.e. $\mathcal{T}_5$, of an R-vine structure, underlies the two figures in the last row of Figure 3.4. It shows similar data characteristics as the realized variance time-series in the top row.

The above analysis clearly stresses the practical interpretability of the model components in the partial correlation vine data transformation approach, namely realized variances and realized (partial) correlations. In the following, the detected inhomogeneous data complexity motivates a specific choice for the R-vine structure used for data transformation in step (S1).

3.4.2 Step (S1): R-vine structure selection for data transformation

In d dimensions there exist $d!/2 \cdot d^{(d-2)(d-3)/2}$ valid R-vines (Morales Napoles et al., 2010). It is important to note that, in general, each of these tree structures allows a valid transformation of the realized correlation matrices. To allow for model parsimony later in step (S2), when modeling the dynamics of the realized variance series and the realized (partial) correlation time-series selected by the R-vine structure, we refer to the data characteristics detected in Section 3.4.1 and propose in this section an algorithm for R-vine structure selection.

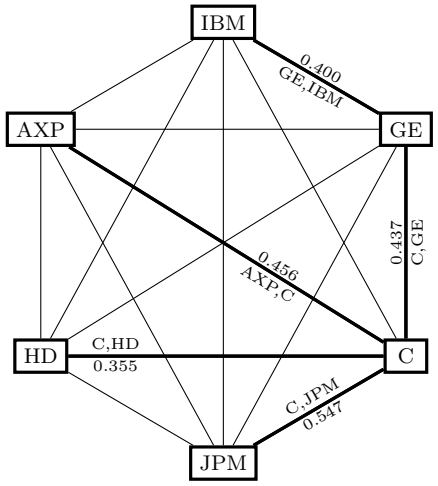
We know that when transforming a series of realized correlation matrices based on the same R-vine structure, to each edge in this R-vine a univariate time-series of realized standard or partial correlations is assigned. Therefore, each edge can be characterized by a weight derived from sample properties of the corresponding time-series. We decide for the average (partial) correlation strength: we consider the average correlation matrix $\bar{\boldsymbol{R}} = (\bar{\rho}_{i,j})_{i,j=1,\dots,d}$ (which is positive definite) calculated from $\boldsymbol{R}_t$, $t = 1, \dots, T$. Then, we find the maximum spanning tree $\mathcal{T}_1$ (Katoh et al., 1981) with edge weights set to $\bar{\rho}_{C_{e,a},C_{e,b}}$. To construct tree $\mathcal{T}_2$, we calculate all average first order partial correlations $\bar{\rho}_{C_{e,a},C_{e,b};D_e}$, i.e. $|D_e| = 1$, where $(C_{e,a}, C_{e,b}; D_e)$ satisfies the proximity condition given $\mathcal{T}_1$. Based on these weights, we find the maximum spanning tree $\mathcal{T}_2$. In general, we construct the R-vine structure $\mathcal{V}_d$ within a top-down procedure and find tree by tree ($\ell = 1, \dots, d-1$) the maximum spanning tree $\mathcal{T}_\ell$ with edge weights set to $\bar{\rho}_{C_{e,a},C_{e,b};D_e}$, where $|D_e| = \ell - 1$ and $(C_{e,a}, C_{e,b}; D_e)$ satisfies the proximity condition given $\mathcal{T}_1$ to $\mathcal{T}_{\ell-1}$. By doing so, we equip based on historical information the R-vine structure with the highest realized (partial) correlation means.

The correlation matrix $\bar{\boldsymbol{R}}$ can be obtained in various ways depending on how the average is calculated. Considering for each pair (i,j) $(i,j \in \{1,\dots,d\},\ i \neq j)$ the empirical mean $\bar{\rho}_{i,j} = \frac{1}{T}\sum_{t=1}^{T} \rho_{i,j;t}$ assigns to each day's value $\rho_{i,j;t}$ the same influence $1/T$ irrespective of how far it lies in the past. For example, by using an exponentially weighted moving average (EWMA) more influence can be assigned to values of more recent days. Here, the exact weights are controlled by the smoothing parameter $\lambda \in]0,1[$ and are defined as $w_t = (1-\lambda)\lambda^{T-t}$, $t = 1, \dots, T$. Thus, for decreasing λ the impact of more recent days increases and therewith the sensitivity of the selected R-vine structure with respect to market changes.

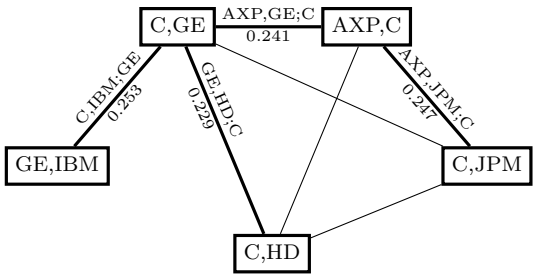
For the real data example, the proposed R-vine structure selection method is illustrated in Table 3.1. Recall that the data include three market participants of financial sectors, namely AXP, C and JPM, IBM as an IT service, HD representing building materials trade and the diversified industrial corporation GE. As edge weights the empirical means of the realized (partial) correlation series based on all data points, i.e. $t = 1, \dots, 2156$ (January 1, 2000 - July 30, 2008), are chosen. In $\mathcal{T}_1$, we start with a full graph, i.e. all edges are allowed to be chosen. Edge by edge a tree, i.e. a connected and acyclic graph, is built adding edges with the highest possible

Table 3.1: Illustration of the R-vine structure selection method for the real data example considering all available data points, i.e. the mean values $\bar{\rho}_{\mathcal{C}_{e,a},\mathcal{C}_{e,b};\mathcal{D}_e}$ are based on $t = 1, \ldots, 2156$.

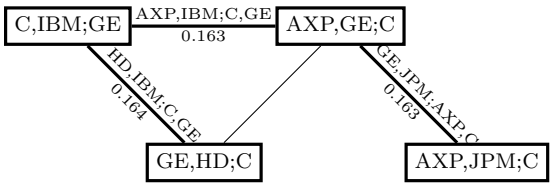
	pairs allowed by proximity condition		selected tree
$\mathcal{D}_e$	$\mathcal{C}_{e,a}, \mathcal{C}_{e,b}$	$\bar{\rho}_{\mathcal{C}_{e,a},\mathcal{C}_{e,b};\mathcal{D}_e}$	
$\emptyset$	C,JPM	0.547	
$\emptyset$	AXP,C	0.456	
$\emptyset$	C,GE	0.437	
$\emptyset$	AXP,JPM	0.433	
$\emptyset$	GE,IBM	0.400	
$\emptyset$	AXP,GE	0.394	
$\emptyset$	GE,JPM	0.393	
$\emptyset$	C,IBM	0.390	
$\emptyset$	IBM,JPM	0.362	
$\emptyset$	AXP,IBM	0.358	
$\emptyset$	C,HD	0.355	
$\emptyset$	GE,HD	0.352	
$\emptyset$	AXP,HD	0.333	
$\emptyset$	HD,JPM	0.333	
$\emptyset$	HD,IBM	0.330	
GE	C,IBM	0.253	
C	AXP,JPM	0.247	
C	AXP,GE	0.241	
C	GE,HD	0.229	
C	GE,JPM	0.214	
C	AXP,HD	0.203	
C	HD,JPM	0.182	
C,GE	HD,IBM	0.164	
C,GE	AXP,IBM	0.163	
AXP,C	GE,JPM	0.163	
C,GE	AXP,HD	0.154	
C,GE,IBM	AXP,HD	0.129	
AXP,C,GE	IBM,JPM	0.121	
AXP,C,GE,IBM	HD,JPM	0.093	



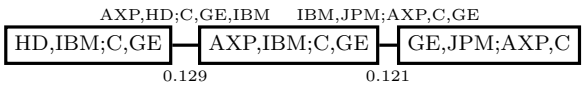
$\mathcal{T}_1$



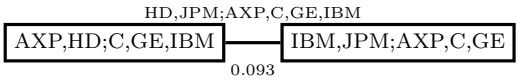
$\mathcal{T}_2$



$\mathcal{T}_3$



$\mathcal{T}_4$



$\mathcal{T}_5$

correlation mean. For example, including the pair (AXP,JPM) would result in a cycle and is thus not allowed. For $\mathcal{T}_2$ only edges satisfying the proximity condition given $\mathcal{T}_1$ are allowed. A tree is constructed by the four edges with the highest mean values of realized first order partial correlations, etc. The resulting R-vine structure captures the strong pairwise realized correlations between the three financial services in trees $\mathcal{T}_1$ and $\mathcal{T}_2$. From $\mathcal{T}_3$ on only realized partial correlations corresponding to stocks from different market sectors are modeled. Note that all time-series illustrated in Figure 3.3 and Figure 3.4 are included in the final R-vine structure. Consequently, using in step (S1) an R-vine structure selected by the proposed algorithm will likely result in higher order realized partial correlation series, which allow for a parsimonious time-series model specification in step (S2).

3.4.3 Step (S2): Multivariate time-series modeling and forecasting

After transforming the series of realized covariance matrices in step (S1), multivariate time-series models in step (S2) can be applied to the transformed data without imposing any parameter restrictions. Except for the considered modeling components this step does not differ from Cholesky decomposition based benchmark models. In both approaches, there are $d(d+1)/2$ model components after data transformation. In particular, the time-series of log-transformed realized variances and Fisher z-transformed realized (partial) correlations could be modeled using a VARFIMA model as suggested for the Cholesky elements in Chiriac and Voev (2011). Compared with this, copula based time-series modeling as applied in Brechmann et al. (2018) showed superior results especially for economic applications.

A $\tilde{d}$-dimensional copula is a multivariate distribution function on $[0,1]^{\tilde{d}}$ with uniformly distributed margins. Since data are required to be approximately i.i.d., the copula model usually is not directly applied to the observed time-series, but to the corresponding standardized residuals $\left(\varepsilon_{1;t},\ldots,\varepsilon_{\tilde{d};t}\right)$, $t=1,\ldots,T$. The latter are extracted after fitting appropriate univariate time-series models to the original marginal data. While no longer being subject to temporal dependence, the residuals inherit the cross-sectional dependence between the time-series components. According to Sklar (1959), their joint distribution function F can be expressed in terms of its marginal distributions F_j $(j=1,\ldots,\tilde{d})$ and its corresponding copula, i.e. $F\left(\varepsilon_1,\ldots,\varepsilon_{\tilde{d}}\right)=\mathbb{C}\{F_1\left(\varepsilon_1\right),\ldots,F_{\tilde{d}}\left(\varepsilon_{\tilde{d}}\right)\}$. Consequently, in a copula based time-series model the individual behavior of the time-series components and their dependence are modeled separately. This allows us to deepen the analysis of the realized variance and (partial) correlation series.

Univariate marginal time-series modeling

As discussed in Section 3.4.1 specific univariate time-series models are needed to reproduce the long-memory property of the Cholesky components as well as of the realized variance and some of the realized (partial) correlation series. HAR (Corsi, 2009) and ARFIMA (Andersen et al., 2003) models are popular models capable of doing so.

Let η_t denote the variable of interest, i.e. a log-transformed realized variance, a Fisher z-transformed realized (partial) correlation or a Cholesky element, at time t. A basic HAR model accounts for different time horizons by incorporating one day $(d=1)$, one week $(w=5)$ and

one month ($m = 22$) averages η_{t-1}, $\eta_{t-1}^{(w)}$ and $\eta_{t-1}^{(m)}$ as regressors for η_t:

$$\eta_t = \alpha_0 + \alpha_1 \eta_{t-1} + \alpha_2 \eta_{t-1}^{(5)} + \alpha_3 \eta_{t-1}^{(22)} + \epsilon_t.$$

The error term ϵ_t is usually assumed to be Gaussian white noise. While showing very good modeling and prediction performance given complex data features, the basic HAR model describes an easy to estimate restricted autoregressive process.

The ARFIMA(p, D, q) model for the time-series η_t, $t = 1, \ldots, T$, is specified by

$$\phi(L)(1-L)^D(\eta_t - \mu) = \psi(L)\epsilon_t,$$

where $\phi(L) = 1 - \phi_1 L - \ldots - \phi_p L^p$ and $\psi(L) = 1 + \psi_1 L + \ldots + \psi_q L^q$ are lag polynomials for $p, q \in \mathbb{N}$. D is the parameter of fractional differencing. We choose $D \in (0, 0.5)$ to guarantee stationarity of the process. Gaussian white noise is usually assumed for the error term ϵ_t.

In these basic models, the volatility h of the error term $\epsilon_t = h\varepsilon_t$ with $\varepsilon_t \sim \mathcal{N}(0,1)$ is assumed to be constant. Given the presence of volatility clustering in the Cholesky series, Brechmann et al. (2018) include a GARCH$(1,1)$ component, i.e. $\epsilon_t = h_t \varepsilon_t$ with $h_t^2 = \omega + \beta_1 \epsilon_{t-1}^2 + \beta_2 h_{t-1}^2$. Usually, the innovation terms ε_t are standard normally distributed, i.e. $\varepsilon_t \sim \mathcal{N}(0,1)$. To additionally capture possible high kurtosis and skewness, Brechmann et al. (2018) further allow the innovations to follow a skewed generalized error distribution, i.e. $\varepsilon_t \sim \text{SGED}(\mu, \sigma, \nu, \xi)$ (Bai et al., 2003; Corsi et al., 2008; Fernández and Steel, 1998). A specification of the skewed generalized error distribution is provided in Section A.1 in Appendix A.

Dependence modeling

After fitting one of the above univariate time-series models to each of the model components, the sample of i.i.d. standardized residuals $\left(\varepsilon_{1;t}, \ldots, \varepsilon_{d(d+1)/2;t}\right)$, $t = 1, \ldots, T$, can be extracted. To them, usually a two-stage proceeding is applied, called inference for margins methods (Joe and Xu, 1996; Joe, 2005). First, the probability integral transform, $\hat{u}_{j;t} = \hat{F}_j(\varepsilon_{j;t})$, is applied to each residual component ($j = 1, \ldots, d(d+1)/2$) to obtain pseudo copula data $\left(\hat{u}_{1;t}, \ldots, \hat{u}_{d(d+1)/2;t}\right)$, $t = 1, \ldots, T$. The marginal estimates $\hat{F}_j$ are specified through the corresponding marginal time-series fit. For example, in case of a basic HAR or ARFIMA model $\hat{F}_j$ is a normal distribution with sample mean (approximately 0) and sample standard deviation (approximately 1).

Second, a copula is fitted to the pseudo copula data. To do so, we consider R-vine copulas as introduced in Section 2.3. Note that while in step (S1) and (S3) of the partial correlation vine data transformation approach, R-vines are exclusively used as a graph theoretical tool for data transformation, they are now the cornerstones for this flexible copula class. Fitting an R-vine copula to the sample $\left(\hat{u}_{1;t}, \ldots, \hat{u}_{d(d+1)/2;t}\right)$, $t = 1, \ldots, T$, finalizes the model specification based on in-sample data.

Forecasting of the model components

A one-day-ahead out-of-sample forecast $\hat{\boldsymbol{Y}}_{T+1}$ is now generated in multiple steps. First, innovations on the copula scale $\left(\hat{u}_{1;T+1}, \ldots, \hat{u}_{d(d+1)/2;T+1}\right)$ are sampled from the R-vine copula fit. The corresponding innovations on their original scale are obtained using the inverse probability integral transform, i.e. $\hat{\varepsilon}_{j;T+1} = \hat{F}_j^{-1}\left(\hat{u}_{j;T+1}\right)$ $(j = 1, \ldots, d(d+1)/2)$. Then, based on the corresponding time-series fit forecasts for the model components, which involve the corresponding simulated innovations, are calculated. In the Cholesky decomposition based model, this results in a predicted upper triangular matrix $\hat{\boldsymbol{C}}_{T+1}$. In the partial correlation vine data transformation approach, the log-transformation of the realized variances and the Fisher z-transformation of the realized (partial) correlations have to be reversed first. This results in the predicted realized partial correlation vine stored in $\hat{\boldsymbol{P}}_{\mathcal{C}(\mathcal{V}_d);T+1}$ and the corresponding predicted realized variance vector $(\hat{y}_{1,1;T+1}, \ldots, \hat{y}_{d,d;T+1})$.

3.4.4 Step (S3): Back-transformation

Finally, based on $\hat{\boldsymbol{C}}_{T+1}$ back-transformation (3.7) is applied for the Cholesky decomposition based approach. Likewise, $\hat{\boldsymbol{P}}_{\mathcal{C}(\mathcal{V}_d);T+1}$ is back-transformed to a symmetric and positive definite correlation matrix (Section 3.2) based on the R-vine structure selected in step (S1) (Section 3.4.2). Combined with the predicted realized variances a forecast for the realized covariance matrix is obtained. Given that both backward procedures involve nonlinear transformations of the copula-distributed innovation terms, the underlying dependence pattern has an explicit effect on the matrix forecast. Clearly, in practice this simulation based procedure is to be replicated several times. The final point-forecast $\hat{\boldsymbol{Y}}_{T+1}$, which is considered as a proxy for the conditional covariance matrix, is obtained as the mean of the simulation based matrix forecasts.

In both modeling approaches the predictions of $\hat{\boldsymbol{Y}}_{T+1}$ are obtained after inverting a nonlinear data transformation. Consequently, while the prediction errors of the model components have zero mean, the nonlinear back-transformation induces a bias. Even though Chiriac and Voev (2011) derive the theoretical bias correction for the Cholesky decomposition based model, they stress that the theoretical formula crucially depends on the considered time-series model and thus, has to be estimated in practice. However, given that in a copula based time-series model the marginal time-series are estimated independently of each other, consistent estimation of the covariance matrix of the forecast errors in Brechmann et al. (2018) is not feasible. Against this background, Chiriac and Voev (2011) and Brechmann et al. (2018) both advocate a data-driven bias correction. In the partial correlation vine data transformation approach, the forecast bias of the variable pair (i, j) in $\hat{\boldsymbol{Y}}_{T+1}$ depends not only on the underlying time-series model but also on the R-vine structure used for data transformation, making a theoretical correction practically infeasible. We therefore as well opt for the heuristic data-driven bias correction proposed in Chiriac and Voev (2011). The basic idea is to match the level of the observed volatilities by scaling the predicted volatilities $\sqrt{\hat{y}_{j,j;T+1}}$ $(j = 1, \ldots, d)$ by the corresponding mean $\frac{1}{T-s+1}\sum_{t=s}^{T}\frac{\sqrt{y_{j,j;t}}}{\sqrt{\hat{y}_{j,j;t}}}$, where s controls the number of past days included for level matching. Note that this proceeding has no influence on the predicted correlation structure. Thus, in the partial correlation vine data transformation approach only the nonlinear inversion of the log-transformation can be corrected.

3.4.5 Modeling approach at a glance

Figure 3.5 summarizes the partial correlation vine data transformation approach discussed in the previous sections.

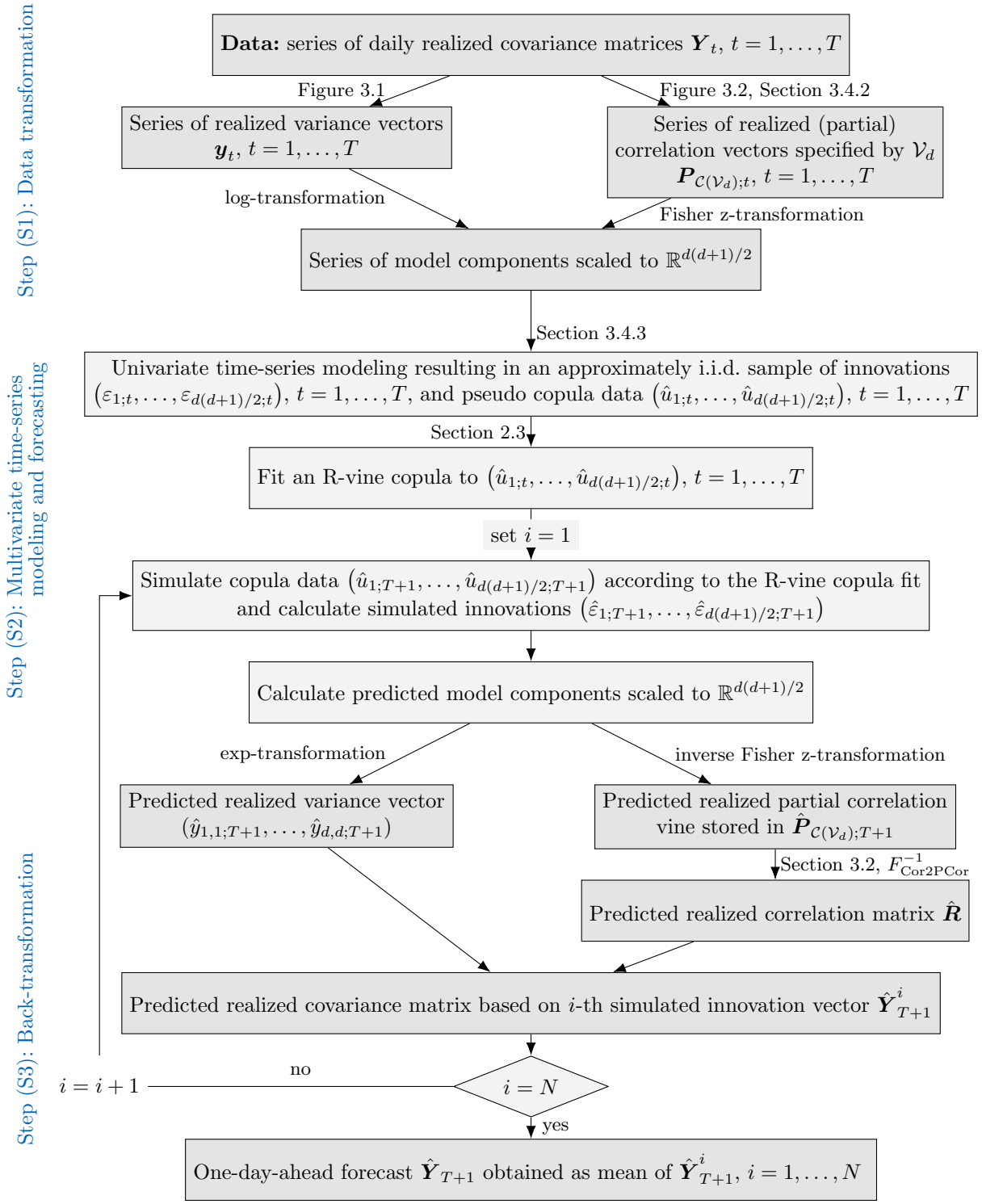


Figure 3.5: Modeling and forecasting approach using partial correlation vine based data transformation of the series of realized covariance matrices in step (S1) and an R-vine copula based time-series model in step (S2). The one-day-ahead forecast $\hat{\boldsymbol{Y}}_{T+1}$ is obtained as mean of N simulation based matrix forecasts.

3.5 Empirical study

The real data example introduced in Section 3.3 and Section 3.4 will now be investigated in more detail. Based on the model specifications in Section 3.5.2 and Section 3.5.3, the out-of-sample forecasting performance of the partial correlation vine data transformation approach and Cholesky decomposition based benchmark models will be evaluated both with respect to statistical precision and mean-variance trade-off in portfolio optimization strategies.

It is crucial to keep in mind that the realized covariance matrices are proxies for the unobservable true conditional covariance matrices, which we aim to predict. As a consequence, when comparing the performance of different forecasting models, loss functions have to satisfy the condition to deliver the same ranking whether the evaluation is based on the unbiased proxy, i.e. the realized covariance matrix, or the true conditional covariance matrix. We will therefore rely on loss functions, which according to Patton (2011) and Laurent et al. (2013) are robust to noise in the volatility proxies. An example for single model components is the root mean squared error.

Further, numerous different models for prediction will be compared. To avoid pairwise comparison of loss functions we apply the model confidence set (MCS) approach developed by Hansen et al. (2011). Starting with an initial set $\mathcal{M}_0$ of m_0 competitor models, it sequentially selects a set of superior models, which contains the best one with a specified level of confidence α. First, for all models k ($k = 1, \dots, m_0$) the loss of the corresponding prediction at time t with respect to the true realization is calculated, i.e.

$$L_{k;t} \coloneqq L\left(X_t, \hat{X}_{k;t}\right), \quad t = 1, \dots, T,$$

where L is a loss function, which satisfies the conditions in Patton (2011) and Laurent et al. (2013). In the following, the series X_t, $t = 1, \dots, T$, can represent either single model components or the realized covariance matrices. Then, for all pairs (k, ℓ) ($k, \ell = 1, \dots, m_0$, $k \neq \ell$) the series of loss differentials

$$d_{k,\ell;t} \coloneqq L_{k;t} - L_{\ell;t}, \quad t = 1, \dots, T,$$

is obtained. Based on the set of competitor models $\mathcal{M}_s \subseteq \mathcal{M}_0$ after step $s > 0$ of the MCS procedure, the null-hypothesis

$$H_{0,\mathcal{M}_s} : \mathbb{E}\left[d_{k,\ell}\right] = 0 \quad \text{for all} \quad k, \ell = 1, \dots, |M_s|$$

is tested based on the test statistic

$$T_{\mathcal{M}_s} = \max_{k,\ell \in \mathcal{M}_s} \frac{|\bar{d}_{k,\ell}|}{\sqrt{\widehat{\mathrm{Var}}\left(\bar{d}_{k,\ell}\right)}},$$

where $\bar{d}_{k,\ell} = \frac{1}{T}\sum_{t=1}^{T} d_{k,\ell;t}$. If $H_{0,\mathcal{M}_s}$ is rejected at the given significance level α, the worst

performing model given by the elimination rule

$$e_{\mathcal{M}_s} = \arg\max_k \left\{ \sup_{\ell \in \mathcal{M}_s} \frac{|\bar{d}_{k,\ell}|}{\sqrt{\widehat{\mathrm{Var}}\left(\bar{d}_{k,\ell}\right)}} \right\}$$

is removed from the set $\mathcal{M}_s$. If $H_{0,\mathcal{M}_s}$ cannot be rejected for the set of remaining models $\mathcal{M}_s$, the MCS procedure stops. More details including implementation aspects are provided in Hansen et al. (2011).

3.5.1 Moving window approach

In the following, we proceed in a moving window approach. Data for the period from January 1, 2000, until June 30, 2008, are available, i.e. for 2156 days. For each time window 502 days (about two years) are used as training set and 22 days (about one month) constitute the test set for which one-day-ahead forecasts are made. Since in case of HAR based time-series models a monthly (22 days) average of the data is involved, the first forecast is obtained for day 525. In total, there are 75 time windows. Figure 3.6 illustrates the moving window approach.

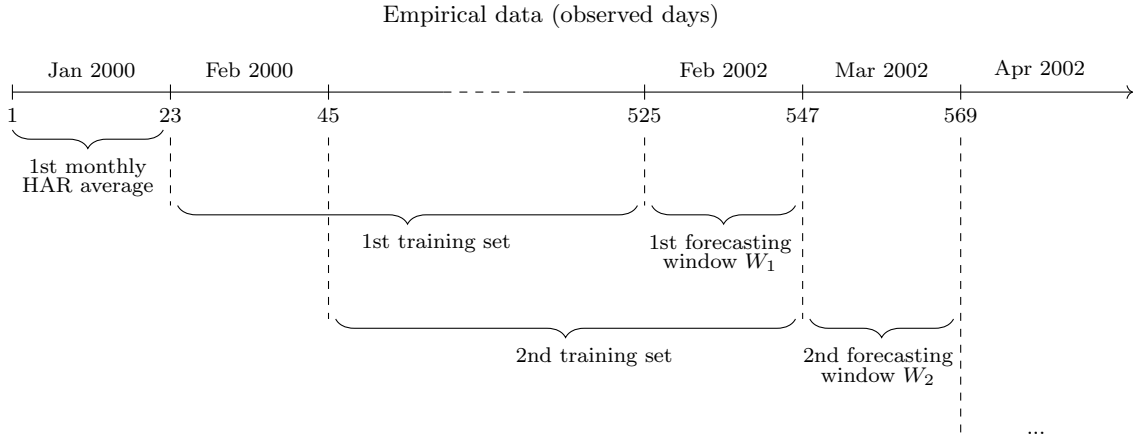


Figure 3.6: Moving window approach illustrated for the considered real data example.

3.5.2 Dynamic data transformation

For each time window W_i $(i = 1, \ldots, 75)$ the realized covariance matrices of the corresponding training set are transformed in step (S1). Clearly, application of the R-vine structure selection algorithm proposed in Section 3.4.2 can lead to varying R-vine structures among time windows. Thus, data transformation in the partial correlation vine data transformation approach may dynamically change over time. Depending on how the average correlation matrix used for R-vine structure selection is calculated, the selected R-vine structure is more or less sensitive to market developments.

In Figure 3.7, the first trees of the R-vine structures selected in each of the 75 time windows are shown indicating the included model components by a black square. In the first row, empirical means of the realized standard correlations are considered. In the second and third row,

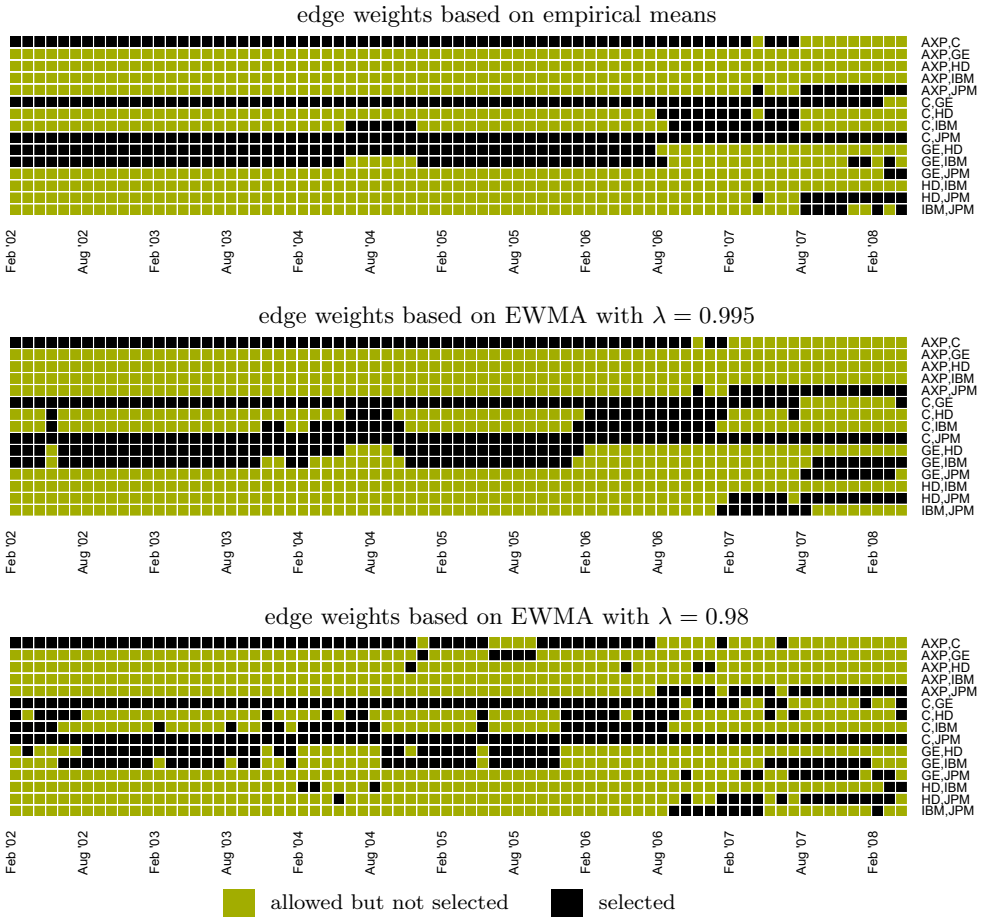


Figure 3.7: Illustration of the dynamically changing R-vine structures used for transforming the series of realized correlation matrices in each of the 75 time windows. The horizontal time axis states the prediction month. Model components selected in tree $\mathcal{T}_1$ are indicated by black squares. Green squares indicate selectable components allowed by the proximity condition (which does not trigger in $\mathcal{T}_1$). In the first panel, the average correlations used for R-vine structure selection are the empirical means of the training set data. In the second and third row, they are exponentially weighted moving averages with $\lambda = 0.995$ and $\lambda = 0.98$, respectively.

exponentially weighted moving averages based on $\lambda = 0.995$ and $\lambda = 0.98$, respectively, are used. While in case of empirical means all days of the two training years are of equal weight, for $\lambda = 0.995$ and $\lambda = 0.98$ the six most recent months and one and a half months, respectively, already contribute half of the information for average calculation. Thus, in the latter case R-vine structure selection is most sensitive to market changes resulting in more frequent variations of the selected model components. For example, changes for the prediction months in mid 2004 or at the beginning of the financial crisis are observed earliest. Nevertheless, for all three setups the selected first tree is quite stable and we may identify three distinct periods: February 2002–August 2006, September 2006–July 2007 and August 2007–July 2008. For these periods, Figure 3.8 illustrates the first tree $\mathcal{T}_1$ of the predominantly chosen R-vine structures. Until August 2006, pairwise correlations between the log-returns including Citigroup (C) and

General Electric (GE) seem to be most pronounced. While C plays a key role within the financial sector, GE as a diversified industrial corporation connects the representatives of the financial sector with the two non-financial stocks. During the period from September 2006 to July 2007 C becomes the first root node in a C-vine, i.e. the node in $\mathcal{T}_1$ with the highest possible number of edges attached to it. At the beginning of the financial crisis in August 2007, the correlations between JP Morgan (JPM) and the other market participants seem to tighten. This results in a predominantly chosen R-vine structure, where except for GE all pairwise correlations including JPM are modeled. Note that in 2007 JPM replaced C as the biggest US-bank in terms of revenues.

To conclude, selecting the R-vine structure for data transformation as proposed in Section 3.4.2 gives interesting insights into market activities over time. In addition, we already know about the resulting inhomogeneous data complexity of the corresponding time-series, which will be further analyzed in the next section. There, the R-vine structures will be selected using EWMA based edge weights with $\lambda = 0.995$. Recall, however, that any R-vine structure could be used for data transformation. To demonstrate the general adequacy of the partial correlation vine data transformation approach irrespective of the R-vine structure used for data transformation, we will consider two alternative ways of R-vine structure selection as well. First, we reverse the idea of inducing model parsimony and select for each time window a C-vine, where in each tree level the root node induces the on average lowest correlation strength. Thus, the effect of decreasing data complexity as for the proposed R-vine structure selection should be eliminated. Second, an R-vine structure on six elements is randomly sampled in each time window (Joe et al., 2011).

For the Cholesky decomposition, the model components depend on the ordering of the assets. However, contrary to the data transformation based on partial correlation vines there is no justifiable rule to decide 'on the fly' for a specific order. Thus, the ordering has to be set upfront. In this sense, the Cholesky decomposition based data transformation is static. Clearly, enumerating all possible permutations of the assets and performing a model analysis for each of them is too time consuming and computationally demanding especially in higher dimensions. Thus, a sensitivity analysis based on several Cholesky decompositions should be performed first. For the considered data, Brechmann et al. (2018) find that the alphabetic ordering of the six stocks performs best. We therefore, choose the latter for all time windows.

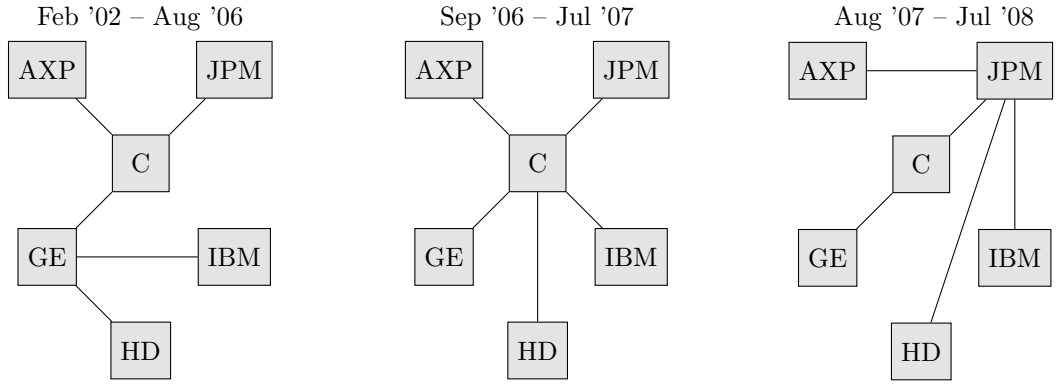


Figure 3.8: Predominantly selected first tree of the R-vine structure used for data transformation during the periods February 2002 to August 2006, September 2006 to July 2007 and August 2007 to July 2008.

3.5.3 Multivariate time-series modeling

As explained in Section 3.4.3, to the sample of model components obtained after transforming the series of realized covariance matrices, marginal time-series models need to be applied first.

Univariate marginal time-series modeling

Given the proposed R-vine structure selection method, we know that with increasing tree level the data complexity decreases such that less elaborate time-series models might already be sufficient for accurate in-sample estimation and out-of-sample forecasting. To support this presumption, for each period within the moving window approach time-series models of different complexity are fitted to the log-transformed realized variance time-series and to the Fisher z-transformed realized (partial) correlations specified by the R-vine structure found in Section 3.4.2. For comparison, also the time-series specified by the C-vine, of which in each tree level the root node induces the on average lowest correlation strength, are investigated.

Besides simply considering the mean value over time, basic univariate HAR and ARFIMA models as well as HAR and ARFIMA models including a GARCH(1,1) component with normal innovations (abbreviated as HN and AN) and with SGED innovations (abbreviated as HSGED and ASGED) are fitted. To evaluate the statistical precision we use the root mean squared error (RMSE), which according to Patton (2011) is robust to noise in the volatility proxies. Table 3.2 shows the out-of-sample RMSE for all time-series model components under consideration. In each row, the set of superior models based on the MCS approach of Hansen et al. (2011) with a confidence level of 10% is highlighted in gray. The model with the lowest RMSE, which is the last one that would be rejected from the model confidence set, is highlighted in bold. In general, ARFIMA based models show a superior prediction performance with respect to the RMSE criterion. However, especially within the variations of the two base models the RMSE values often are very close to each other. For the realized variance time-series and the realized standard correlation time-series in $\mathcal{T}_1$ of the selected R-vine structure, the best model usually includes a GARCH$(1,1)$ augmentation. For tree level $\mathcal{T}_2$ and $\mathcal{T}_3$, there is a shift to basic ARFIMA models, while for tree level $\mathcal{T}_4$ and $\mathcal{T}_5$ even simply using the mean realized partial correlation value as forecast is included in the model confidence set at a confidence level of 10%. This confirms the presumption that given the proposed R-vine structure selection method with increasing tree level more parsimonious time-series models already are sufficient. This hierarchical pattern is not observed for the considered C-vine. Here, base models including a GARCH$(1,1)$ augmentation with normal or SGED innovations most often would be the last ones to be eliminated from the model confidence set. In particular, a simple mean forecast clearly is insufficient even in high tree levels. Similar results are detected for the Cholesky elements and are given in Table A.1 in Appendix A.2. Given the close performance of the different time-series models in terms of the RMSE, from a practical point of view the time-series model which is economically best interpretable should be chosen for univariate marginal modeling.

In the following, we consider two groups of models. One including only HAR based time-series models and the other including only ARFIMA based models. Given the above findings within the

partial correlation vine data transformation approach, we use HN and AN models, respectively, for all components in case that a C-vine or a randomly sampled R-vine structure is taken for data transformation. Likewise, we proceed for the Cholesky decomposition based model. In case of R-vine structure selection according to Section 3.4.2, we stepwise increase model parsimony. For model components corresponding to the realized variance and realized standard correlation time-series in $\mathcal{T}_1$, we use HN and AN models, respectively. For the ones in tree level $\mathcal{T}_2$ and $\mathcal{T}_3$ we apply basic HAR and ARFIMA models, respectively. For components in $\mathcal{T}_4$ and $\mathcal{T}_5$ we consider in one setting basic HAR and ARFIMA models, respectively, and take in another setting simply the mean value of the underlying training set as forecast.

Table 3.2: RMSE with respect to the complete out-of-sample forecasting horizon (1632 days) for the model components in the partial correlation vine data transformation approach. Two different R-vine structures for data transformation are considered. The set of superior models according to the MCS approach at a confidence level of 10% is highlighted in gray. The lowest RMSE is highlighted in bold.

		mean	HAR	HN	HSGED	ARFIMA	AN	ASGED
	AXP	1.0429	0.4711	0.4715	0.4719	0.4684	**0.4668**	0.4680
	C	1.0154	0.4469	0.4479	0.4510	0.4465	**0.4455**	0.4483
	GE	0.8105	0.4634	0.4632	0.4647	0.4627	**0.4625**	0.4627
	HD	0.7766	0.4554	0.4557	0.4568	**0.4540**	0.4543	0.4543
	IBM	0.7242	0.4320	0.4322	0.4323	**0.4317**	0.4331	0.4327
	JPM	1.0137	0.4653	0.4647	0.4671	0.4652	**0.4641**	0.4655
R-vine selection (Section 3.4.2)	AXP,C	0.2183	0.1572	0.1573	0.1576	0.1568	**0.1568**	0.1571
	C,GE	0.1984	0.1531	0.1531	0.1530	0.1526	0.1526	**0.1524**
	C,HD	0.1857	0.1519	0.1520	0.1520	0.1515	0.1516	**0.1512**
	C,JPM	0.2149	0.1619	0.1620	0.1620	0.1615	**0.1615**	**0.1615**
	GE,IBM	0.1914	**0.1489**	0.1490	0.1490	0.1490	0.1490	0.1490
	AXP,GE;C	0.1395	0.1317	0.1317	0.1316	**0.1313**	0.1313	0.1313
	AXP,JPM;C	0.1364	**0.1295**	0.1295	0.1296	0.1297	0.1297	0.1297
	C,IBM;GE	0.1340	0.1271	0.1271	0.1272	**0.1270**	0.1270	0.1270
	GE,HD;C	0.1384	0.1300	0.1301	0.1301	**0.1292**	0.1292	0.1292
	AXP,IBM;C,GE	0.1246	0.1237	0.1237	0.1237	**0.1231**	0.1231	0.1232
	GE,JPM;AXP,C	0.1211	0.1196	0.1196	0.1196	**0.1191**	0.1191	**0.1191**
	HD,IBM;C,GE	0.1260	0.1253	0.1254	0.1254	**0.1250**	0.1250	0.1251
	AXP,HD;C,GE,IBM	0.1175	0.1171	0.1171	0.1172	0.1168	**0.1168**	0.1170
	IBM,JPM;AXP,C,GE	0.1237	0.1227	0.1227	0.1227	0.1225	**0.1223**	0.1225
	HD,JPM;AXP,C,GE,IBM	0.1177	**0.1175**	0.1175	0.1175	0.1178	0.1177	0.1178
C-vine	AXP,HD	0.1873	0.1547	0.1547	0.1549	**0.1538**	0.1538	0.1539
	C,HD	0.1857	0.1519	0.1520	0.1520	0.1515	0.1516	**0.1512**
	GE,HD	0.1856	0.1517	0.1517	0.1517	0.1509	**0.1509**	0.1509
	HD,IBM	0.1731	0.1488	0.1489	0.1489	0.1484	0.1484	**0.1483**
	HD,JPM	0.1804	0.1532	0.1533	0.1533	0.1525	**0.1523**	0.1525
	AXP,IBM;HD	0.1491	0.1349	0.1349	0.1349	0.1345	**0.1344**	0.1345
	C,IBM;HD	0.1502	0.1320	0.1320	0.1320	0.1317	0.1317	**0.1315**
	GE,IBM;HD	0.1574	**0.1352**	**0.1352**	0.1353	0.1353	0.1354	0.1355
	IBM,JPM;HD	0.1486	0.1382	0.1382	0.1383	0.1376	0.1376	**0.1376**
	AXP,GE;HD,IBM	0.1387	0.1303	0.1303	0.1302	0.1299	0.1300	**0.1299**
	C,GE;HD,IBM	0.1371	0.1270	0.1270	0.1268	0.1270	0.1270	**0.1268**
	GE,JPM;HD,IBM	0.1312	0.1251	**0.1251**	0.1251	0.1251	0.1251	0.1252
	AXP,C;GE,HD,IBM	0.1496	0.1283	0.1283	0.1285	0.1281	**0.1281**	0.1282
	AXP,JPM;GE,HD,IBM	0.1526	0.1321	0.1321	0.1321	0.1317	**0.1316**	0.1316
	C,JPM;AXP,GE,HD,IBM	0.1398	0.1247	0.1247	0.1248	0.1241	**0.1241**	0.1242

Dependence modeling

Now, for each time window interest is in the cross-sectional dependence between the model components. Since dependencies between stocks are expected to be most pronounced during financial turmoil, we consider as an example the time window from July 2006 to July 2008. Based on the specified time-series models, the sample of innovations is obtained and transformed to pseudo copula data (Section 3.4.3). Figure 3.9 shows the resulting data based on R-vine structure selection according to Section 3.4.2 and HAR based time-series modeling. It illustrates the corresponding histograms on its diagonal, pairwise contour plots with standard normal margins in the lower left corner and pairs plots with corresponding Kendall's τ values in the upper right corner. Only dependencies between model components corresponding to realized variances (last six components) and realized standard correlations (first five components) are significant with Kendall's τ values ranging from 0.2 to 0.5. Dependencies including components, which correspond to partial correlations, are rather small and close to zero for higher tree levels. Based on these findings, we subsequently consider five different R-vine copula settings. First, independence for all pairs is assumed. Second, a 21-dimensional R-vine copula is fitted to capture dependence between all model components. Third, a reduced structured dependence is imposed, where a 11-dimensional R-vine copula is fitted only to the components corresponding to realized variances and realized standard correlations. The components corresponding to realized partial correlations are assumed to be independent. Both in case of full and reduced structured R-vine copula based dependence modeling, we allow as a first setting the pair-copulas to stem from various copula families such as Clayton, Gumbel, Frank, etc. including their reflected forms. Thus, possible asymmetric and nonlinear dependence patterns can be detected. Given the primarily elliptical shapes in Figure 3.9 we also consider an R-vine copula exclusively built from bivariate (conditional) Gaussian copulas, i.e. a Gaussian vine. Except for the structured dependence, the same settings for the copula models are taken in the Cholesky decomposition based benchmarks. To fit an R-vine copula model we rely on the `R`-package `VineCopula` (Schepsmeier et al., 2017).

3.5.4 Forecasting performance

Based on the above findings, Table 3.3 summarizes the in total 36 data transformation based prediction models considered to obtain one-day-ahead forecasts as described in Section 3.4.4. In addition, we consider three naive benchmarks. First, $\hat{\boldsymbol{Y}}_{T+1}$ is set to the realized covariance matrix at time point T, i.e. $\hat{\boldsymbol{Y}}_{T+1} = \boldsymbol{Y}_T$. Second, $\hat{\boldsymbol{Y}}_{T+1}$ is calculated as the equally weighted average of the realized covariance matrices in the corresponding training set. Third, $\hat{\boldsymbol{Y}}_{T+1}$ is obtained as an exponentially weighted moving average, i.e. in our setup $\hat{\boldsymbol{Y}}_{T+1} = \lambda \hat{\boldsymbol{Y}}_T + (1-\lambda)\,\boldsymbol{Y}_T$, where the smoothing parameter λ is set to 0.94 as commonly suggested in the framework of a RiskMetrics approach (Morgan, 1996).

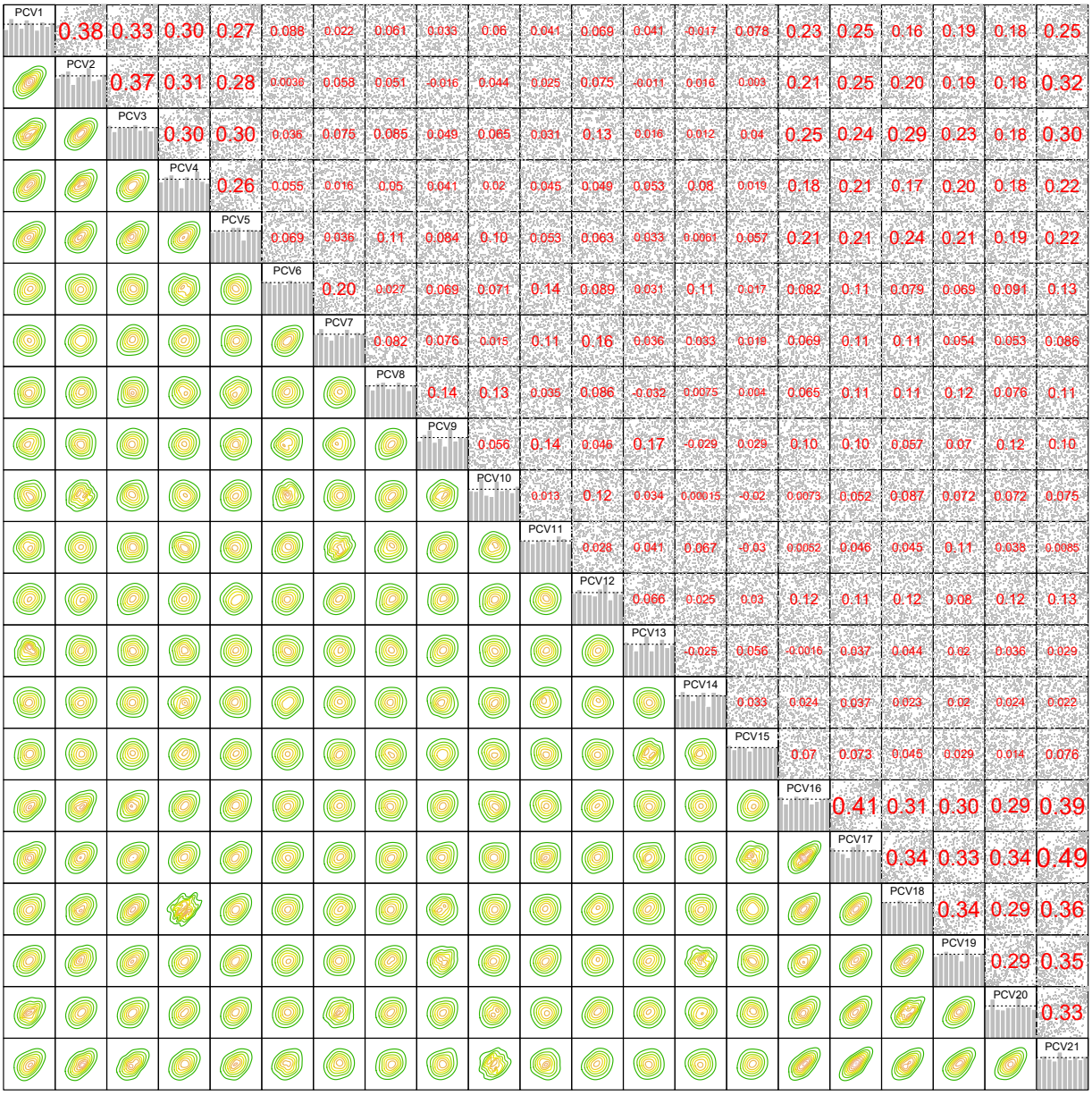


Figure 3.9: Exploratory data analysis for the pairwise dependencies of the 21-dimensional pseudo copula data estimated for the period July 2006 to July 2008 using the proposed method for R-vine structure selection (Section 3.4.2) and HAR based time-series modeling. Pairwise contour plots with normalized margins, histograms and pairs plots with empirical Kendall's τ values are shown. The first five components PCV1 to PCV5 correspond to realized standard correlations in $\mathcal{T}_1$, components PCV6 to PCV9 correspond to realized first order partial correlations in $\mathcal{T}_2$, etc. Variables PCV16 to PCV21 correspond to realized variances.

Table 3.3: Overview of all data transformation based prediction models compared with respect to their forecasting performance in Section 3.5.4. 'dims' and 'PC' are used as abbreviations for 'dimensions' and 'pair-copulas', respectively. For models, which subsequently will be investigated in more detail, short names are introduced in the last column.

Transformation based on	Univariate time-series modeling variances, $\mathcal{T}_1$	$\mathcal{T}_2$ & $\mathcal{T}_3$	$\mathcal{T}_4$ & $\mathcal{T}_5$	R-vine copula assumed for transformed data	In the following referred to as
R-vine selection (Section 3.4.2)	AN/HN	A/H	A/H	21 dims, all PCs 21 dims, Gauss	A-/H-PCV-Sel-full
	AN/HN	A/H	mean	11 dims, all PCs 11 dims, Gauss independence	A-/H-PCV-Sel-struc
C-vine	AN/HN for all components			21 dims, all PCs 21 dims, Gauss 11 dims, all PCs 11 dims, Gauss independence	A-/H-PCV-CVine
random R-vine	AN/HN for all components			21 dims, all PCs 21 dims, Gauss 11 dims, all PCs 11 dims, Gauss independence	A-/H-PCV-random
Cholesky	AN/HN for all components			21 dims, all PCs 21 dims, Gauss independence	A-/H-Chol

Out-of-sample forecasting precision

To illustrate that the proposed forecasting approach is on target, Figure 3.10 shows for the realized variance time-series of JPM (top panel), the realized covariance time-series of C and JPM (mid panel) as well as IBM and JPM (bottom panel) the historical time-series from January 2002 until July 2008 together with the one-day-ahead forecasts based on the R-vine structure selected according to Section 3.4.2, ARFIMA based time-series modeling and a 21-dimensional Gaussian vine for dependence modeling. Results for all other realized variances and covariance pairs are similar and given in Figure A.1 in Section A.2. The trends in all time-series including high short-term peaks are well detected and modeled. Distances between historical extreme peaks and corresponding forecasts are large. This finding holds true for all prediction models and is due to the high volatility of the realized variances and covariances. The predicted time-series incorporate smoothed long-term information of historical data and thus, are more stable.

To evaluate the statistical precision of the matrix forecasts, Table 3.4 summarizes for all considered models the RMSE based on the Frobenius norm between the realized and the predicted covariance matrices. For the real matrix $\boldsymbol{A} = (a_{i,j})_{i,j=1,\dots,d} := \boldsymbol{Y} - \hat{\boldsymbol{Y}}$, the Frobenius norm is defined as $\|\boldsymbol{A}\| = \sum_{i=1}^{d} \sum_{j=1}^{d} a_{i,j}^2$. This loss function satisfies the conditions in Laurent et al. (2013) for consistent model ranking. In the right column, the RMSE based on bias corrected (bc) matrix forecasts are shown. We use historical data over the period of one year for level correction as described in Section 3.4.4. This reduces the out-of-sample forecasting horizon to 1368 days.

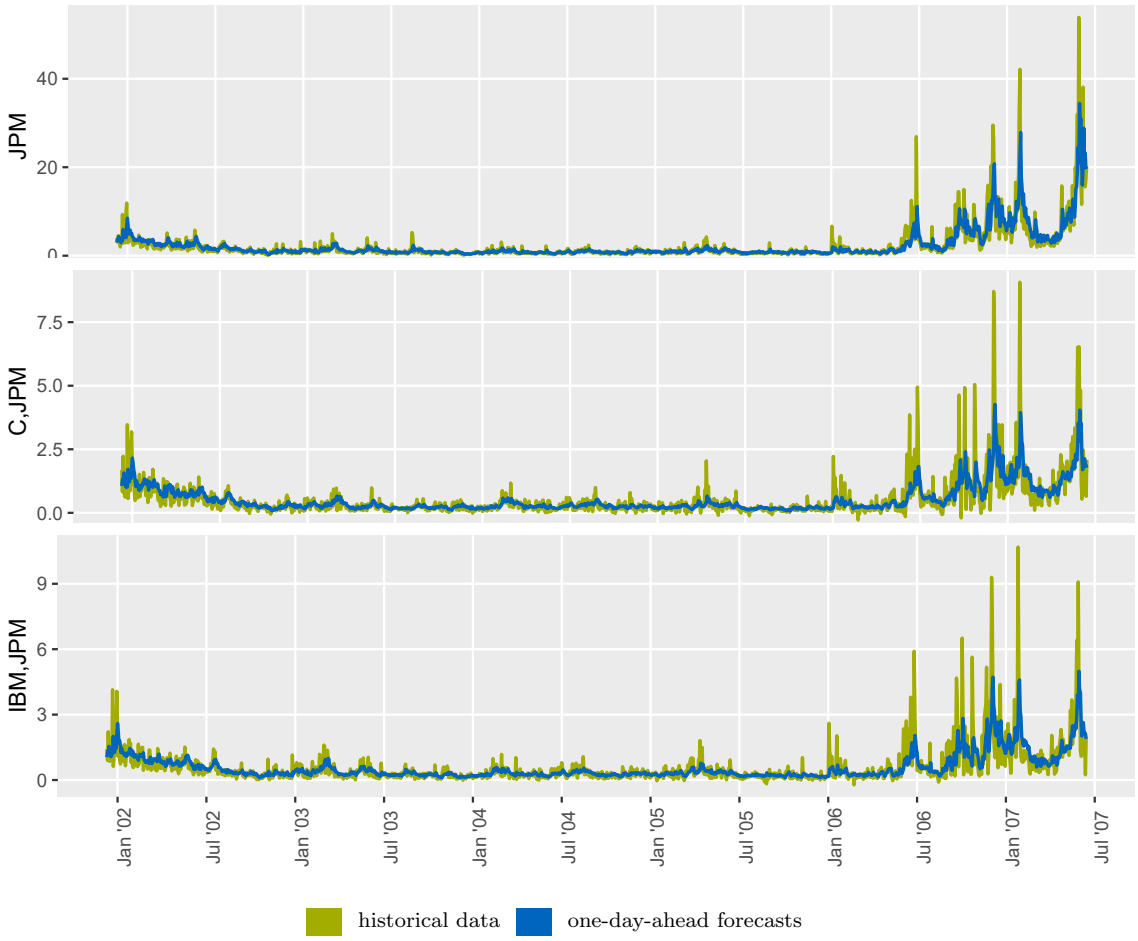


Figure 3.10: Daily realized variance time-series for JPM (1st row) and daily realized covariance time-series (2nd and 3rd row) together with the time-series of the corresponding daily forecasts based on the partial correlation vine data transformation approach with R-vine structure selected according to Section 3.4.2, ARFIMA based time-series modeling and a 21-dimensional Gaussian vine for dependence modeling.

As in the previous analysis of the single model components, ARFIMA based models in general have smaller RMSE values compared to HAR based models. All models using partial correlation vine based data transformation and full dependence modeling exhibit a smaller RMSE than Cholesky decomposition based models and show very similar performance among each other. This confirms that any R-vine structure can be used for data transformation in step (S1) of the model approach. Among the partial correlation vine data transformation based models those with a C-vine structure used for data transformation have the highest RMSE. Recall that by construction more complex data features are induced even for high tree levels. For C-vine and random R-vine structures, time-series modeling in step (S2) with independent components and reduced structured dependence between components is clearly improved by models, which capture dependence between all model components. Here, the decreasing data complexity does not trigger. However, for the R-vine structure selected according to Section 3.4.2 the performance in case of reduced structured dependence is only slightly improved by full dependence model-

ing. Thus, also the dependence between the model components allows for model parsimony. In general, using a Gaussian vine for dependence modeling between the model components shows comparable results as using more elaborate copulas allowing for tail-dependence. All discussed prediction models clearly show superior results as compared to the naive benchmarks. Bias correction in step (S3) slightly improves results while maintaining the above observations among the different models.

Table 3.4: RMSE based on the Frobenius norm between the realized and predicted correlation matrices with respect to the complete out-of-sample forecasting horizon (1368 days) for all models. In the last column, results for bias corrected (bc) forecasts are shown.

Marginals	Data transformation based on		R-vine copula assumed for transformed data	RMSE	RMSE bc
ARFIMA based	PCV	R-vine selection (Section 3.4.2)	independence	6.6314	6.6045
			full all	6.5725	6.5632
			full Gauss	6.5685	6.5619
			structured all	6.5858	6.5740
			structured Gauss	6.5864	6.5742
		C-vine	independence	6.7076	6.6733
			full all	6.5968	6.5854
			full Gauss	6.5967	6.5860
			structured all	6.6436	6.6189
			structured Gauss	6.6410	6.6166
		random R-vine	independence	6.6694	6.6393
			full all	6.5826	6.5746
			full Gauss	6.5886	6.5810
			structured all	6.6155	6.5968
			structured Gauss	6.6138	6.5954
		Cholesky	independence	6.6732	6.6437
			all	6.6121	6.6001
			Gauss	6.6193	6.6075
HAR based	PCV	R-vine selection (Section 3.4.2)	independence	6.7218	6.6566
			full all	6.6332	6.5998
			full Gauss	6.6313	6.5962
			structured all	6.6544	6.6146
			structured Gauss	6.6522	6.6122
		C-vine	independence	6.7900	6.7085
			full all	6.6574	6.6153
			full Gauss	6.6575	6.6158
			structured all	6.7117	6.6480
			structured Gauss	6.7094	6.6453
		random R-vine	independence	6.7527	6.6830
			full all	6.6432	6.6117
			full Gauss	6.6474	6.6158
			structured all	6.6821	6.6334
			structured Gauss	6.6796	6.6310
		Cholesky	independence	6.7400	6.6866
			all	6.6863	6.6621
			Gauss	6.6841	6.6603
mean over training set				12.0894	
previous day				7.2937	
EWMA with $\lambda = 0.94$				7.8790	

To test the statistical significance of the results, we apply the MCS approach of Hansen et al. (2011). Based on the above findings, we restrict the analysis to models using a Gaussian vine for dependence modeling between the model components. Only in case of R-vine structure selection according to Section 3.4.2 we consider reduced structured dependence modeling in addition to the full one. In the following, we refer to these models using the short names introduced in Table 3.3. Bias corrected forecasts are taken. Figure 3.11 shows for each half-year period of the out-of-sample horizon the set of superior models (indicated by a gray dot), which contains the best model at a confidence level of 10%. A blue triangle and an orange cross indicate the last and the next model, respectively, that would be eliminated. For almost all periods, all models are selected at the given confidence level showing very close performance of all models. Most often, HAR based models would be eliminated next, while ARFIMA based models usually would be the last ones to be eliminated from the set of superior models. In three out of eleven periods, the ARFIMA-Cholesky model has the smallest RMSE based on the Frobenius norm and therefore automatically would be the last model to be eliminated. All ARFIMA and partial correlation vine data transformation based models show rather robust performance over the out-of-sample forecasting horizon. Especially, the models based on R-vine structure selection according to Section 3.4.2 usually are the ones to be eliminated last from the MCS, i.e. having the smallest loss.

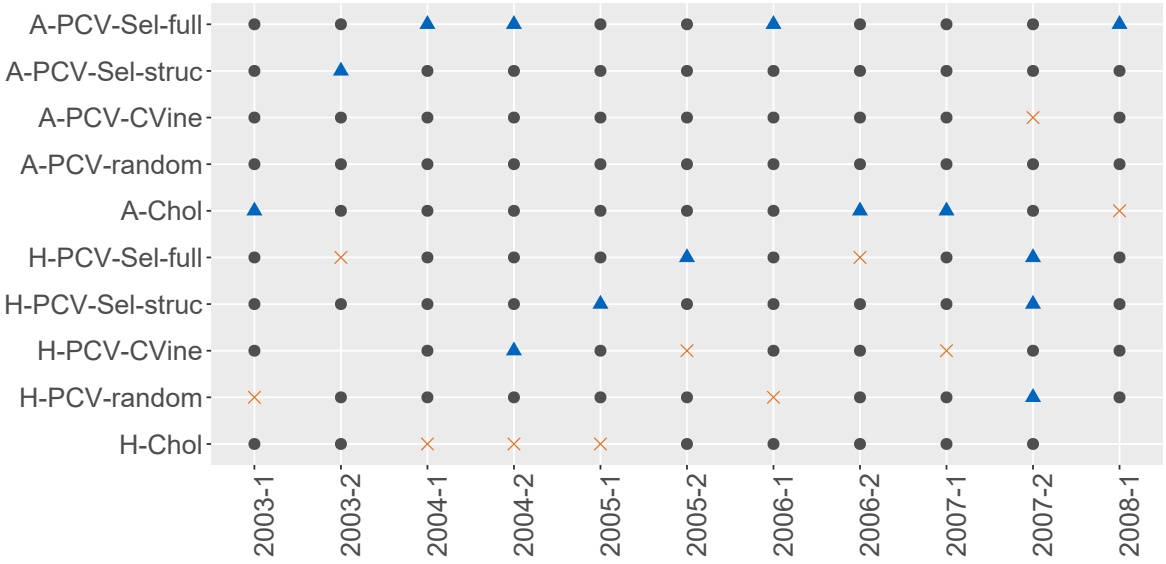


Figure 3.11: Model confidence sets of Hansen et al. (2011) with confidence level 10% for all half-year periods of the out-of-sample forecasting horizon. Gray dots indicate selected models, blue triangles and orange crosses indicate the last and the next model, respectively, that would be eliminated from the set of superior models.

Mean-variance trade-off in portfolio optimization

For additional economic evaluation of the forecasts, we construct portfolios based on each prediction model, which are mean-variance efficient. For a risk-averse investor we assume a quadratic utility function. Then, the problem to maximize the utility is reduced to finding the asset weights $\boldsymbol{w}$, which minimize the portfolio volatility σ_{p} based on a fixed target expected return μ_{p} (Markowitz, 1952). The optimal portfolio is obtained by solving the quadratic problem

$$\min_{\boldsymbol{w}_{t+1}} \boldsymbol{w}_{t+1}' \hat{\boldsymbol{\Sigma}}_{t+1} \boldsymbol{w}_{t+1} \quad \text{s.t.} \quad \boldsymbol{w}_{t+1}' \mathbb{E}\left[\boldsymbol{r}_{t+1} | \mathcal{F}_t\right] = \mu_{\mathrm{p}} \quad \text{and} \quad \boldsymbol{w}_{t+1}' \mathbf{1}_d = 1,$$

where $\boldsymbol{w}_{t+1}$ is the $d \times 1$ vector of portfolio weights chosen at day t for $t+1$, $\mathbf{1}_d$ is a $d \times 1$ vector of ones, μ_{p} is the daily target expected return and $\hat{\boldsymbol{\Sigma}}_{t+1}$ is the conditional (with respect to the information set) covariance forecast at day t for $t+1$. The latter corresponds to the realized covariance forecasts $\hat{\boldsymbol{Y}}_{t+1}$.

For each prediction model, we solve the above optimization problem for a daily target return μ_{p} for all 1368 days in the out-of-sample horizon. Based on the optimal portfolio weights $\boldsymbol{w}_t$ for day t ($t = 1, \ldots, 1368$) the expected risk in terms of standard deviation, $\sqrt{\boldsymbol{w}_t' \hat{\boldsymbol{Y}}_t \boldsymbol{w}_t}$, corresponding to the target expected return μ_{p} can be calculated. Taking the averages over the forecasting horizon and repeating the procedure for a grid of target returns, results in an average efficient frontier for each prediction model. To obtain an average oracle efficient frontier, the true realized covariance matrices for each day t are used. Figure 3.12 shows the efficient frontiers for the considered HAR

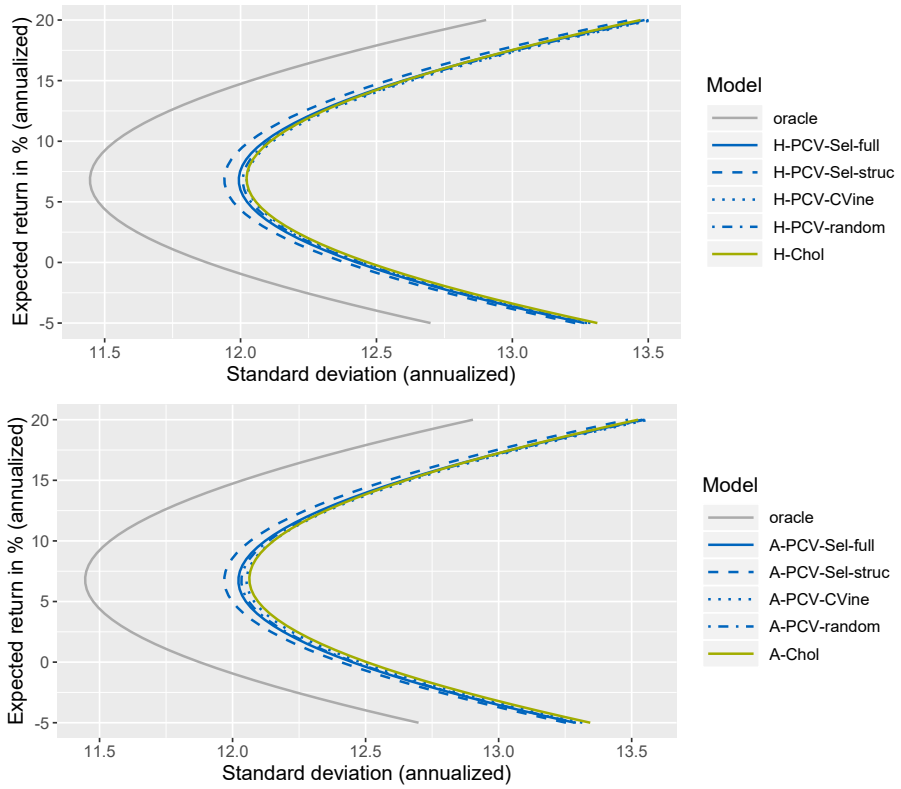


Figure 3.12: Efficient frontier for each HAR (top) and ARFIMA (bottom) based prediction model plotting the expected return versus its corresponding risk in terms of standard deviation. The curves are averages over the out-of-sample horizon (1368 days).

and ARFIMA based prediction models. All partial correlation vine data transformation based models show a clear improvement in terms of the expected mean-variance trade-off compared to the two Cholesky decomposition based prediction models.

To validate this observation in an out-of-sample setting we calculate for each prediction model based on the corresponding optimal portfolio weights $\boldsymbol{w}_{t+1}$ estimated at day t for $t+1$ ($t = 0, \ldots, 1367$) the ex-post realized portfolio return $r_{\mathrm{p},t+1} = \boldsymbol{w}_{t+1}'\boldsymbol{r}_{t+1}$ and the ex-post realized portfolio volatility $\sigma_{\mathrm{p},t+1} = \sqrt{\boldsymbol{w}_{t+1}'\boldsymbol{Y}_{t+1}\boldsymbol{w}_{t+1}}$. Here, $\boldsymbol{r}_{t+1}$ and $\boldsymbol{Y}_{t+1}$ are the true returns and the true covariance matrix, respectively, realized at day $t+1$. Given a small enough grid of target returns, we are able to obtain for each prediction model the series of ex-post portfolio standard deviation $\sigma_{\mathrm{p},t+1}$, $t = 0, \ldots, 1367$, corresponding to a certain average ex-post realized return. For an average annualized ex-post realized portfolio return of approximately 7.5%, 10%, 12% and 15%, Table 3.5 shows the average annualized ex-post realized portfolio standard deviation for each prediction model. The set of models, which includes the model with the lowest standard deviation at a confidence level of 10% based on the MCS approach of Hansen et al. (2011), is highlighted in gray. The model with the lowest loss (deviation from zero) is highlighted in bold. In general, HAR based prediction models perform better than their ARFIMA based counterparts. In the ex-post analysis, all ARFIMA based partial correlation vine data transformation based models have the highest average standard deviation. This confirms the often seen phenomenon that models with the lowest statistical loss do not necessarily show superior results in economical applications (Laurent et al., 2013). The HAR based model with R-vine structure selected according to Section 3.4.2 and with reduced structured dependence among the model components is the best model at a confidence level of 10% for all considered annualized ex-post realized portfolio returns. Comparing the average ex-post realized standard deviations of the HAR based prediction models, further demonstrates the strength of the proposed methodology irrespective of the R-vine structure used for data transformation.

Table 3.5: Annualized average ex-post standard deviation corresponding to four levels of annualized ex-post realized return. The sets of models, which include the one with the smallest standard deviation at a confidence level of 10%, are highlighted in gray. The last model to be eliminated is highlighted bold.

Model	Realized return in % (annualized)			
	7.5	10	12.5	15
A-PCV-Sel-full	12.5217	12.9545	13.5492	14.2832
A-PCV-Sel-struc	12.5055	12.9254	13.5086	14.2317
A-PCV-CVine	12.4981	12.9269	13.5252	14.2722
A-PCV-random	12.4949	12.9151	13.5027	14.2366
A-Chol	12.4616	12.8641	13.4352	14.1510
H-PCV-Sel-full	12.4754	12.8748	13.4359	14.1363
H-PCV-Sel-struc	**12.4588**	**12.8447**	**13.3937**	**14.0796**
H-PCV-CVine	12.4595	12.8606	13.4293	14.1429
H-PCV-random	12.4680	12.8680	13.4349	14.1413
H-Chol	12.4718	12.8729	13.4396	14.1475

3.6 Discussion

In this chapter, we introduced a novel approach to model and forecast time-series of realized covariance matrices. According to Barndorff-Nielsen and Shephard (2004), the latter are consistent nonparametric estimates for the latent and thus non-observable daily conditional covariance matrices underlying a process of daily log-returns.

In Section 3.1, existing literature was reviewed emphasizing the challenge to obtain symmetric and positive definite forecasts for the realized covariance matrices. To avoid restrictions on the time-series models imposed by this requirement, we proposed to not directly model the realized covariance matrices, but to jointly model the realized variances and a subset of realized standard and partial correlations specified by an R-vine structure. In Section 3.2, we therefore introduced partial correlation vines as a graph theoretical object and explained in detail the data transformation to determine from a realized correlation matrix the standard and partial correlations corresponding to a partial correlation vine with given R-vine structure. Since the standard and partial correlations specified in a partial correlation vine are algebraically independent, positive definiteness of the correlation matrix obtained after inverting the data transformation is always guaranteed.

In Section 3.3, we introduced the general data setting and the Cholesky decomposition as a popular and commonly used alternative data transformation. Along with a real data example, we outlined in Section 3.4 the proposed modeling and forecasting approach focusing on three main steps. Specific data characteristics detected in Section 3.4.1 motivated for data transformation in step (S1) the R-vine structure selection method proposed in Section 3.4.2. The selection algorithm built upon the practical interpretation of the model components, namely realized variances and realized (partial) correlations. High average correlation strengths were captured in lower tree levels of the R-vine structure leaving higher order realized partial correlation time-series for which typical and challenging properties of volatility data such as long-memory behavior or volatility clustering were no longer observed. Thus, an inhomogeneous data complexity of the model components was obtained giving hope for possible parsimonious time-series modeling in step (S2) of the model approach. Copula based multivariate time-series modeling and forecasting for the transformed data and back-transformation of the model components in step (S3) were discussed in Section 3.4.3 and Section 3.4.4, respectively.

In Section 3.5, the detailed analysis of the real data example was continued further exploring the beneficial features of the proposed partial correlation vine data transformation approach within a moving window approach (Section 3.5.1). In Section 3.5.2, the proposed R-vine structure selection method allowed the R-vine structure used for data transformation to dynamically change over time providing interesting insights into market activities. Analyzing the univariate time-series of the model components obtained after data transformation in Section 3.5.3, confirmed that data complexity decreases for time-series in higher tree levels when transforming the realized correlation matrices based on the R-vine structure selected as proposed in Section 3.4.2. In addition, cross-sectional dependence between these higher order partial correlation series was negligible allowing for dimension reduction in the considered multivariate time-series

models. In Section 3.5.4, the forecasting performance both in terms of statistical precision and in an economic evaluation, where ex-post realizations of mean-variance efficient portfolios were investigated, showed very good and in several settings even statistically significant superior prediction capability compared to the Cholesky decomposition based benchmark models. In particular, these findings also held true for partial correlation vine data transformation based prediction models, where the R-vine structure was either randomly sampled or constructed such that higher data complexity was intentionally induced for higher tree levels.

Given the excellent prediction power of the Cholesky decomposition based benchmark models often demonstrated in literature, these findings combined with other beneficial properties of the partial correlation vine data transformation approach such as interpretability of the model components and model parsimony provide strong justification for its use in practice.

Chapter 4

Modeling time-to-event data using R-vine copulas

The material in this chapter is very similar to the publications Barthel et al. (2018c) and Barthel et al. (2018b).

4.1 R-vine copulas for time-to-event data

Before the two main projects considering time-to-event data will be discussed in Section 4.2, Section 4.3 and Section 4.4, some additional notation and theoretical background for R-vine copula modeling need to be provided. This includes a general introduction of survival copulas in Section 4.1.1 and the notation of pair-copula constructions in terms of survival components in Section 4.1.2. Further, Section 4.1.3 discusses standard techniques for the estimation of univariate right-censored event time data.

4.1.1 Sklar's Theorem for survival functions

Copula models for time-to-event data typically are formulated in terms of the so-called survival copula. Let $(T_1, \ldots, T_d)$ with $T_j \geq 0$ $(j = 1, \ldots, d)$ be a d-dimensional positive valued random vector of event times with marginal distribution functions F_j, marginal density functions f_j and marginal survival functions S_j, i.e.

$$S_j(t_j) = 1 - F_j(t_j) = \mathbb{P}(T_j > t_j) = \int_{t_j}^{+\infty} f_j(s)\,\mathrm{d}s.$$

Further, let f be the joint density function and S be the joint survival function

$$S(t_1, \ldots, t_d) = \mathbb{P}(T_1 > t_1, \ldots, T_d > t_d).$$

Similar to Sklar's Theorem (Sklar, 1959) as given in Section 2.2 the d-dimensional survival copula $\mathbb{C}^{\mathrm{S}}$ corresponding to S is a dependence function that interconnects the marginal survival functions, and thereby models the joint survival function of event times, i.e.

$$S(t_1, \ldots, t_d) = \mathbb{C}^{\mathrm{S}}\{S_1(t_1), \ldots, S_d(t_d)\}.$$

If all T_j $(j = 1, \ldots, d)$ are continuous, $\mathbb{C}^{\mathrm{S}}$ is unique. Further, if the survival copula density

$$\mathbb{c}^{\mathrm{S}}(u_1, \ldots, u_d) = \frac{\partial^d}{\partial u_1 \cdots \partial u_d} \mathbb{C}^{\mathrm{S}}(u_1, \ldots, u_d)$$

exists, it holds that

$$f(t_1, \ldots, t_d) = (-1)^d \frac{\partial^d}{\partial t_1 \cdots \partial t_d} S(t_1, \ldots, t_d) = \mathbb{c}^{\mathrm{S}}\{S_1(t_1), \ldots, S_d(t_d)\} \prod_{j=1}^{d} f_j(t_j).$$

To provide the connection between the copula $\mathbb{C}$ corresponding to F and its survival copula $\mathbb{C}^{\mathrm{S}}$, let $V_{\ell,\boldsymbol{u}_{1:d}}$ be the set of all d-dimensional vectors $\boldsymbol{v}_{1:d} \coloneqq (v_1, \ldots, v_d)' \in [0,1]^d$, where exactly ℓ elements are set to 0. The other $d - \ell$ elements are set to their corresponding value in $\boldsymbol{u}_{1:d}$, i.e.

$$V_{\ell,\boldsymbol{u}_{1:d}} \coloneqq \{\boldsymbol{v}_{1:d} \in [0,1]^d : v_j \in \{0, u_j\}, \sum_{j=1}^{d} \mathbb{1}(v_j = 0) = \ell\}.$$

Then, according to Georges et al. (2001) the following equalities for the copula $\mathbb{C}$ and the corresponding survival copula $\mathbb{C}^{\mathrm{S}}$ hold:

$$\mathbb{C}(u_1, \ldots, u_d) = \sum_{j=0}^{d} (-1)^j \sum_{\boldsymbol{v}_{1:d} \in V_{d-j,\boldsymbol{u}_{1:d}}} \mathbb{C}^{\mathrm{S}}(1 - v_1, \ldots, 1 - v_d) \tag{4.1}$$

and vice versa

$$\mathbb{C}^{\mathrm{S}}(u_1, \ldots, u_d) = \sum_{j=0}^{d} (-1)^j \sum_{\boldsymbol{v}_{1:d} \in V_{d-j,\boldsymbol{u}_{1:d}}} \mathbb{C}(1 - v_1, \ldots, 1 - v_d). \tag{4.2}$$

Consequently, if we identify $u_j = S_j(t_j)$ $(j = 1, \ldots, d)$ in (4.2) and take the partial derivatives with respect to all arguments, for the copula densities it follows that

$$\mathbb{c}^{\mathrm{S}}\{S_1(t_1), \ldots, S_d(t_d)\} = \mathbb{c}\{1 - S_1(t_1), \ldots, 1 - S_d(t_d)\} = \mathbb{c}\{F_1(t_1), \ldots, F_d(t_d)\}. \tag{4.3}$$

Example 4.1. *We derive* (4.2) *and* (4.3) *for* $d = 3$ *as an example. Let* F_1*,* F_2*,* F_3*,* $F_{1,2}$*,* $F_{1,3}$ *and* $F_{2,3}$ *denote the marginal distribution functions corresponding to the joint distribution function* F *of the event times* (T_1, T_2, T_3)*. Likewise, let* $\mathbb{C}_{1,2}$*,* $\mathbb{C}_{1,3}$ *and* $\mathbb{C}_{2,3}$ *be the bivariate marginal copulas of the copula* $\mathbb{C}$ *corresponding to* F*. Following the principle of inclusion and exclusion (Roberts and Tesman, 2009), it holds that*

$$\begin{aligned} S(t_1, t_2, t_3) = {} & 1 - F_1(t_1) - F_2(t_2) - F_3(t_3) \\ & + F_{1,2}(t_1, t_2) + F_{1,3}(t_1, t_3) + F_{2,3}(t_2, t_3) \\ & - F(t_1, t_2, t_3). \end{aligned}$$

Using Sklar's Theorem (Sklar, 1959), we conclude that

$$\begin{aligned}
&\mathbb{C}^S\{S_1(t_1), S_2(t_2), S_3(t_3)\} \\
&= -2 + S_1(t_1) + S_2(t_2) + S_3(t_3) \\
&\quad + \mathbb{C}_{1,2}\{1 - S_1(t_1), 1 - S_2(t_2)\} + \mathbb{C}_{1,3}\{1 - S_1(t_1), 1 - S_3(t_3)\} + \mathbb{C}_{2,3}\{1 - S_2(t_2), 1 - S_3(t_3)\} \\
&\quad - \mathbb{C}\{1 - S_1(t_1), 1 - S_2(t_2), 1 - S_3(t_3)\}.
\end{aligned}$$

Further,

$$\begin{aligned}
\mathbb{c}^S\{S_1(t_1), S_2(t_2), S_3(t_3)\} &= \frac{\partial^3}{\partial u_1 \partial u_2 \partial u_3} \mathbb{C}^S(u_1, u_2, u_3) \Big|_{\substack{u_1 = S_1(t_1) \\ u_2 = S_2(t_2) \\ u_3 = S_3(t_3)}} \\
&= \mathbb{c}\{1 - S_1(t_1), 1 - S_2(t_2), 1 - S_3(t_3)\} \\
&= \mathbb{c}\{F_1(t_1), F_2(t_2), F_3(t_3)\}.
\end{aligned}$$

4.1.2 Pair-copula constructions in terms of survival components

Now, we address the question how to express the joint density f of event times $(T_1, \ldots, T_d)$ using a pair-copula construction built from bivariate survival copula densities. Recall from Section 2.3 that a d-dimensional R-vine density is constructed from $d(d-1)/2$ unconditional and conditional bivariate copulas. The R-vine structure is defined by a set of linked trees $\mathcal{V}_d = (\mathcal{T}_1, \ldots, \mathcal{T}_d)$ satisfying the three conditions given in Section 2.1 on page 7. Let the corresponding edge set be $E(\mathcal{V}_d) \coloneqq E_1 \cup \cdots \cup E_{d-1}$. Then, the d-dimensional joint density function f of the event times $(T_1, \ldots, T_d)$ can be written as a simplified R-vine density as follows:

$$\begin{aligned}
&f(t_1, \ldots, t_d) \\
&\overset{\text{Sklar (1959)}}{=} \mathbb{c}\{F_1(t_1), \ldots, F_d(t_d)\} \prod_{j=1}^{d} f_j(t_j) \\
&= \prod_{j=1}^{d} f_j(t_j) \prod_{\ell=1}^{d-1} \prod_{e \in E_\ell} \mathbb{c}_{a_e, b_e; D_e}\{F_{a_e|D_e}(t_{a_e}|\boldsymbol{t}_{D_e}), F_{b_e|D_e}(t_{b_e}|\boldsymbol{t}_{D_e})\}. \qquad (4.4)
\end{aligned}$$

If all margins are uniform, we speak of an R-vine copula density. In (4.4),

- $\mathbb{c}_{a_e, b_e; D_e}(\cdot, \cdot)$ denotes the copula density corresponding to the conditional distribution of (T_{a_e}, T_{b_e}) given $\boldsymbol{T}_{D_e} = \boldsymbol{t}_{D_e}$ with $\boldsymbol{T}_{D_e}$ the vector containing all event times with indices in D_e. The corresponding copula will be denoted by $\mathbb{C}_{a_e, b_e; D_e}(\cdot, \cdot)$. Note that the simplifying assumption is applied.
- $F_{a_e|D_e}(\cdot|\boldsymbol{t}_{D_e})$ denotes the conditional distribution of event time T_{a_e} given $\boldsymbol{T}_{D_e} = \boldsymbol{t}_{D_e}$.

In an analogous way, we from now on denote by $S_{a_e|D_e}(\cdot|\boldsymbol{t}_{D_e})$ the conditional survival function of event time T_{a_e} given $\boldsymbol{T}_{D_e} = \boldsymbol{t}_{D_e}$, i.e.

$$S_{a_e|D_e}(t|\boldsymbol{t}_{D_e}) = 1 - F_{a_e|D_e}(t|\boldsymbol{t}_{D_e}).$$

Further, we denote by $\mathbb{C}^{\mathrm{S}}_{a_e,b_e;D_e}(\cdot,\cdot)$ and $\mathbb{c}^{\mathrm{S}}_{a_e,b_e;D_e}(\cdot,\cdot)$ the survival copula and the survival copula density, respectively, corresponding to $\mathbb{C}_{a_e,b_e;D_e}(\cdot,\cdot)$. Then, using (4.3) we are able to rewrite (4.4) in terms of survival copulas:

$$\begin{aligned} & f(t_1,\ldots,t_d) \\ & \overset{(4.3)}{=} \mathbb{c}^{\mathrm{S}}\{S_1(t_1),\ldots,S_d(t_d)\}\prod_{j=1}^{d} f_j(t_j) \\ & = \prod_{j=1}^{d} f_j(t_j)\prod_{\ell=1}^{d-1}\prod_{e\in E_\ell} \mathbb{c}^{\mathrm{S}}_{a_e,b_e;D_e}\{S_{a_e|D_e}(t_{a_e}|\boldsymbol{t}_{D_e}),S_{b_e|D_e}(t_{b_e}|\boldsymbol{t}_{D_e})\}. \end{aligned} \tag{4.5}$$

Recall the important result for pair-copula constructions first given by Joe (1997) that the conditional distribution functions $F_{a_e|D_e}(\cdot|\boldsymbol{t}_{D_e})$, subsequently abbreviated as $F_{a|D}(\cdot|\boldsymbol{t}_D)$, can be evaluated using only the pair-copulas specified in lower tree levels of the underlying R-vine structure. To do so, the corresponding h-functions as defined in Section 2.3 are recursively applied. A similar recursive evaluation is feasible to determine the corresponding conditional survival functions $S_{a|D}(\cdot|\boldsymbol{t}_D)$. Let $a,b\notin D$, $a<b$, and define for $i\in\{a,b\}$ the set $D_{+i}:=D\cup\{i\}$. Recall that

$$F_{a|D_{+b}}\left(t_a|\boldsymbol{t}_{D_{+b}}\right)=h_{a|b;D}\{F_{a|D}(t_a|\boldsymbol{t}_D)\,|F_{b|D}(t_b|\boldsymbol{t}_D)\}=\frac{\partial}{\partial u}\mathbb{C}_{a,b;D}\{F_{a|D}(t_a|\boldsymbol{t}_D),u\}\Big|_{u=F_{b|D}(t_b|\boldsymbol{t}_D)}.$$

To evaluate $S_{a|D_{+b}}\left(t_a|\boldsymbol{t}_{D_{+b}}\right)$, we calculate

$$\begin{aligned} & S_{a|D_{+b}}\left(t_a|\boldsymbol{t}_{D_{+b}}\right) \\ & \quad = 1-F_{a|D_{+b}}\left(t_a|\boldsymbol{t}_{D_{+b}}\right) \\ & \quad = 1-\frac{\partial}{\partial u}\mathbb{C}_{a,b;D}\{F_{a|D}(t_a|\boldsymbol{t}_D),u\}\Big|_{u=F_{b|D}(t_b|\boldsymbol{t}_D)} \\ & \quad \overset{(4.1)}{=} 1-\frac{\partial}{\partial u}\left[1-\{1-F_{a|D}(t_a|\boldsymbol{t}_D)\}-(1-u)+\mathbb{C}^{\mathrm{S}}_{a,b;D}\{1-F_{a|D}(t_a|\boldsymbol{t}_D),1-u\}\right]\Big|_{u=F_{b|D}(t_b|\boldsymbol{t}_D)} \\ & \quad = 1-\left[1+\frac{\partial}{\partial u}\mathbb{C}^{\mathrm{S}}_{a,b;D}\{1-F_{a|D}(t_a|\boldsymbol{t}_D),1-u\}\right]\Big|_{u=F_{b|D}(t_b|\boldsymbol{t}_D)} \\ & \quad \overset{\text{chain rule}}{=} -\frac{\partial}{\partial(1-u)}\mathbb{C}^{\mathrm{S}}_{a,b;D}\{1-F_{a|D}(t_a|\boldsymbol{t}_D),1-u\}\frac{\partial(1-u)}{\partial u}\Big|_{u=F_{b|D}(t_b|\boldsymbol{t}_D)} \\ & \quad = \frac{\partial}{\partial v}\mathbb{C}^{\mathrm{S}}_{a,b;D}\{S_{a|D}(t_a|\boldsymbol{t}_D),v\}\Big|_{v=S_{b|D}(t_b|\boldsymbol{t}_D)}. \end{aligned}$$

In a similar manner, we obtain

$$S_{b|D_{+a}}\left(t_b|\boldsymbol{t}_{D_{+a}}\right)=\frac{\partial}{\partial v}\mathbb{C}^{\mathrm{S}}_{a,b;D}\{v,S_{b|D}(t_b|\boldsymbol{t}_D)\}\Big|_{v=S_{a|D}(t_a|\boldsymbol{t}_D)}.$$

To conclude, the recursive character of the arguments appearing in a pair-copula construction

remains valid when expressing the latter in terms of bivariate (conditional) survival copula densities. The term h-function will subsequently be used in an analogous way for the partial derivatives of the survival pair-copulas with respect to their arguments.

If event time T_b corresponds to a leaf in the first tree of the underlying R-vine structure, it will never occur as a conditioning variable. Then, given the possible recursive evaluation of the conditional survival functions appearing in an R-vine density, there is a closed form expression of $S_{b|D_{+a}}\left(\cdot|\boldsymbol{t}_{D_{+a}}\right)$ only in terms of the survival pair-copulas in lower trees and the survival margins. In particular, we know $S_{b|D_{+a}}\left(\cdot|\boldsymbol{t}_{D_{+a}}\right)$ analytically and can simulate from it. To obtain lower and upper bounds of a prediction interval, the conditional quantile function can be used. The latter is the inverse of the conditional distribution function. Kraus and Czado (2017) show that it also is exclusively based on lower tree pair-copulas and the marginals. The conditional quantile function can be calculated from the conditional survival function as follows:

$$q_\alpha\left(\boldsymbol{t}_{D_{+a}}\right) \coloneqq F^{-1}_{b|D_{+a}}\left(\alpha|\boldsymbol{t}_{D_{+a}}\right) = S^{-1}_{b|D_{+a}}\left(1-\alpha|\boldsymbol{t}_{D_{+a}}\right).$$

We end this section with an important remark on tail-dependence. Note from the definition of rotated bivariate copulas in Section 2.2.3 and (4.3) that the survival pair-copulas $\mathbb{c}^{\mathrm{S}}_{a_e,b_e;D_e}$ correspond to the 180 degree rotations of the corresponding counterparts $\mathbb{c}_{a_e,b_e;D_e}$. Thus, upper tail-dependence and lower tail-dependence correspond to the joint occurrence of very small and very large event times, respectively. The results of this section are summarized in Example 4.2 based on a four-dimensional ordered D-vine density.

Example 4.2. *In case of a four-dimensional D-vine with ordering* $1-2-3-4$*, the pair-copula construction in terms of survival components for the joint density function* f *is given by*

$$\begin{aligned} f\left(t_1,\ldots,t_4\right) = \quad & f_1\left(t_1\right) f_2\left(t_2\right) f_3\left(t_3\right) f_4\left(t_4\right) \\ & \times \mathbb{c}^S_{1,2}\{S_1\left(t_1\right), S_2\left(t_2\right)\}\mathbb{c}^S_{2,3}\{S_2\left(t_2\right), S_3\left(t_3\right)\}\mathbb{c}^S_{3,4}\{S_3\left(t_3\right), S_4\left(t_4\right)\} \\ & \times \mathbb{c}^S_{1,3;2}\{S_{1|2}\left(t_1|t_2\right), S_{3|2}\left(t_3|t_2\right)\}\mathbb{c}^S_{2,4;3}\{S_{2|3}\left(t_2|t_3\right), S_{4|3}\left(t_4|t_3\right)\} \\ & \times \mathbb{c}^S_{1,4;2,3}\{S_{1|2,3}\left(t_1|t_2,t_3\right), S_{4|2,3}\left(t_4|t_2,t_3\right)\}. \end{aligned}$$

For example, the second argument of the pair-copula density $\mathbb{c}^S_{1,4;2,3}$ *in tree* $\mathcal{T}_3$ *of the underlying D-vine tree structure – that is* $S_{4|2,3}\left(t_4|t_2,t_3\right)$ *– can be recursively evaluated using the h-functions corresponding to* $\mathbb{C}^S_{2,3}$ *and* $\mathbb{C}^S_{3,4}$ *specified in tree* $\mathcal{T}_1$ *and* $\mathbb{C}^S_{2,4;3}$ *specified in tree* $\mathcal{T}_2$ *as follows:*

$$\begin{aligned} S_{4|2,3}\left(t_4|t_2,t_3\right) &= h_{4|2;3}\{S_{4|3}\left(t_4|t_3\right)\big|S_{2|3}\left(t_2|t_3\right)\} \\ &= h_{4|2;3}\left[h_{4|3}\{S_4\left(t_4\right)|S_3\left(t_3\right)\}|h_{2|3}\{S_2\left(t_2\right)|S_3\left(t_3\right)\}\right]. \end{aligned} \tag{4.6}$$

Let us consider the subvine only including variables T_2, T_3 *and* T_4*. Since* T_4 *is a leaf, the conditional* α*-quantile* $q_\alpha\left(t_2,t_3\right)$ *can be calculated via inversion of* $S_{4|2,3}\left(\cdot|t_2,t_3\right)$ *as follows:*

$$\begin{aligned} q_\alpha\left(t_2,t_3\right) &= \left(S_{4|2,3}\right)^{-1}\left(1-\alpha|t_2,t_3\right) \\ &\overset{(4.6)}{=} S_4^{-1}\left[h^{-1}_{4|3}\{h^{-1}_{4|2;3}\left(1-\alpha\big|h_{2|3}\left(S_2\left(t_2\right)|S_3\left(t_3\right)\right)\right)\Big|S_3\left(t_3\right)\}\right]. \end{aligned}$$

4.1.3 Modeling of the univariate survival margins

Before copula parameter estimation for right-censored event time data can be discussed, techniques to estimate the survival marginals have to be introduced. In this thesis, survival margins are modeled either parametrically using likelihood optimization or nonparametrically using the popular Kaplan-Meier estimator or the Nelson-Aalen estimator. Details on all estimation techniques can be found in Hougaard (2000).

First, recall that event time data typically are subject to right-censoring, i.e. considering univariate data with sample size n for some observation units the corresponding true realization t_i $(i = 1, \dots, n)$ of the event time T might not be observed, but a lower time c_i stemming from the right-censoring time C might be recorded. Thus, the observed data, which are to be modeled, are given by $y_i = \min(t_i, c_i)$ together with the censoring indicator $\delta_i = \mathbb{1}(t_i \leq c_i)$. A common assumption, which we will adopt throughout, is that censoring times are noninformative and independent of the event times.

If the observed data are supposed to be modeled parametrically, a parametric form with parameters $\boldsymbol{\alpha}$ for the univariate survival function S with corresponding density function f is taken. The loglikelihood function for univariate right-censored data, which is to optimized with respect to $\boldsymbol{\alpha}$, is given by

$$\ell(\boldsymbol{\alpha}; y_1, \dots, y_n, \delta_1, \dots, \delta_n) = \sum_{i=1}^{n} \delta_i \log\{f(y_i)\} + (1 - \delta_i) \log\{S(y_i)\}.$$

Clearly, for each loglikelihood contribution observed true event times and right-censored observations have to be distinguished. In the first case, full information for the corresponding individual is available and – as for complete data – the density function f is evaluated at the observed value. In the latter case, the true event time is arbitrarily larger than the observed value. This is reflected by evaluating the survival function S at the observed value. Examples for common parametric models for univariate time-to-event data are the Weibull or Gamma distribution family.

For nonparametric modeling of the survival function, let $t_{(k)}$ denote a time at which the event of interest (for example death) occurred at least once, and let d_k denote the number of observation units, for which the event of interest occurred at time $t_{(k)}$. Further, denote by n_k the number of observation units at risk at time $t_{(k)}$, i.e. observation units that have not yet experienced the event or been censored at time $t_{(k)}$. The Kaplan-Meier estimate for time t is defined by

$$\hat{S}^{\mathrm{KM}}(t) \coloneqq \prod_{t_{(k)} \leq t} \left(1 - \frac{d_k}{n_k}\right)$$

with $\hat{S}^{\mathrm{KM}}(0) = 1$. While the Kaplan-Meier estimator directly models the survival function, the Nelson-Aalen estimator provides estimates for the so-called cumulative hazard function

$$\Lambda(t) \coloneqq \int_0^t \lambda(s)\, \mathrm{d}s = -\log\{S(t)\},$$

where $\lambda(t)$ is the hazard rate. The latter describes the approximate probability for an individual at time t to instantaneously experience the event of interest conditional on not having experienced it before. The Nelson-Aalen estimate at time t is

$$\hat{\Lambda}(t) \coloneqq \sum_{t_{(k)} \leq t} \frac{d_k}{n_k}$$

and thus, provides an estimate for the survival function at time t through the transformation

$$\hat{S}^{\mathrm{NA}}(t) = \exp\{-\hat{\Lambda}(t)\}.$$

Both estimators result in step functions with jumps only at the observed true event times, i.e. the jump sizes W_i^{KM} and W_i^{NA} for a censored observation y_i with $\delta_i = 0$ equal zero. The jump sizes for a true observed event y_i with $\delta_i = 1$ are determined both by the occurred true events and the censored observations. In particular, note that the Kaplan-Meier estimate only drops to zero, when the last observation is a true event.

For ease of notation, expressions for the d-dimensional survival copula $\mathbb{C}^{\mathrm{S}}$ corresponding to event times $(T_1, \ldots, T_d)$ will in the following be given in terms of the corresponding copula data, i.e. $U_j = S_j(Y_j)$ $(j = 1, \ldots, d)$, where $Y_j = \min(T_j, C_j)$. If the survival margins are unknown, one of the above estimation techniques will be applied to obtain pseudo copula data. Note that the data on the copula level inherit the censoring status of their corresponding values on the original scale. Further, since from now on we exclusively work with survival copulas, we omit the superscript *S*. We also restrict ourselves to the wording *copula* instead of *survival copula*.

4.2 Likelihood estimation of dependence patterns in right-censored event time data

Building upon the provided basics in Section 4.1, in this section the first out of two projects using R-vine copulas in the context of multivariate event time data is presented. The main concepts in this section are based on the work in Barthel (2015) (master's thesis). To the best of our knowledge, R-vine copulas had not been studied for right-censored clustered event times before. The results of the master's thesis were later extended and published in Barthel et al. (2018c). The following content is a slight variation of this publication.

4.2.1 Introduction

In many studies, primary interest lies in the time until a prespecified event occurs. Often, the data appear in clusters of equal size, i.e. the data are balanced. For example, in Laevens et al. (1997) time to mastitis infection in udder quarters of primiparous cows is observed. The cow is the cluster and the infection times of the four udder quarters are the clustered data. For an accurate analysis of clustered data flexible models are needed to describe the underlying dependence pattern. Copulas provide the right tools for this goal. For clusters of size two, a large catalog of bivariate copula families exists. For clusters of size more than two, popular multivariate copulas such as exchangeable (EAC) and nested Archimedean copulas (NAC) (Joe, 1993; Embrechts et al., 2003; Nelsen, 2006; Hofert, 2008) only induce restrictive dependence patterns. For instance, in EAC models all marginal copulas show exactly the same type (and even strength) of tail-dependence. In NAC models, the nesting condition limits all building blocks to stem from the same copula family leading again to the same type (but not strength) of tail-dependence. More flexible models are thus needed to capture complex association patterns present in clustered data. Flexible alternatives for EAC and NAC models include Joe-Hu copulas (Joe and Hu, 1996) and R-vine copulas (Aas et al., 2009; Bedford and Cooke, 2002; Czado, 2010; Kurowicka and Joe, 2011; Kurowicka and Cooke, 2006b). A Joe-Hu copula corresponds to a mixture of positive powers of max-infinitely divisible bivariate copulas. The induced dependence pattern is completely determined by the mixture and by the choice of bivariate copulas. The idea of an R-vine copula is to decompose the joint density of the clustered event times into a cascade of bivariate copula densities using conditioning. So, in both approaches bivariate copulas or bivariate copula densities are the building blocks. Given the variety of well-studied bivariate copulas, it is clear that Joe-Hu copulas and R-vine copulas allow a flexible modeling of the within-cluster association in clustered event time data.

For the above mentioned copula models the focus is usually on complete, i.e. non-censored, data. However, event time data are often subject to right-censoring. This means that for some observations the true event time is not observed but instead a lower (censored) time is registered. For example, in the mastitis study cows may be lost to follow-up (for example due to death) or may experience the event after the end of the study (censored at study end). Since the presence of right-censoring in clustered event time data complicates the statistical analysis substantially, copula based modeling approaches for right-censored clustered data have been less explored and

are restricted to rather simple copula classes such as elliptical or Archimedean copulas. Recently, Geerdens et al. (2016a) studied, for balanced right-censored data, the model flexibility of Joe-Hu copulas (Joe and Hu, 1996) as compared to less elaborate EAC and NAC models. R-vine copulas have not yet been studied for right-censored clustered event times. Therefore, our main objective is to develop a likelihood based estimation approach using the flexible class of R-vine copulas. Using the theorem of Sklar (1959) and following the ideas in Shih and Louis (1995), we proceed in two steps. In step one, the survival margins are modeled. Here, any estimation technique for univariate right-censored event time data can be used, for example maximum likelihood estimation or the nonparametric Kaplan-Meier estimator. Focus, however, lies in detecting the inherent dependence pattern using R-vine copula based likelihood estimation in the second step. Due to right-censoring, numerical integration is needed, making the global likelihood optimization computationally challenging. We introduce a sequential estimation approach to find a fair trade-off between the numerical demand caused by data complexity and the accuracy of the estimates.

In Section 4.2.2, we provide information on the general data setting and introduce the notation used throughout. Following the ideas in Shih and Louis (1995), Section 4.2.3 contains the loglikelihood function for right-censored quadruple event time data. In particular, we provide the loglikelihood expression in terms of R-vine copula components and therewith extend existing R-vine copula concepts to the setting of right-censored clustered time-to-event data. In this section, we also discuss how to deal with numerical aspects of the presented optimization method. A simulation study is performed in Section 4.2.4 to demonstrate the good finite sample performance of our approach.

4.2.2 Data setting and notation

Suppose a study includes n independent individuals. Each of it is to be considered as a cluster of d observation units, which are simultaneously observed for the event of interest. We focus on $d = 3$ and $d = 4$. Let $T_{i,j}$ be the true j-th event time in cluster i ($i = 1, \dots, n$ and $j = 1, \dots, d$). Due to a limited follow-up period, the event times $T_{i,j}$ are subject to right-censoring by $C_{i,j}$, which is the j-th censoring time of cluster i. Thus, for cluster i we observe $Y_{i,j} = \min(T_{i,j}, C_{i,j})$ together with the censoring indicator $\delta_{i,j} = \mathbb{1}(T_{i,j} \leq C_{i,j})$ with $j = 1, \dots, d$. Throughout, we assume that $T_{i,j}$ and $C_{i,j}$ are independent and that censoring is noninformative. Further, we assume that $C_{i,j} = C_i$ holds for all $j = 1, \dots, d$, i.e. all event times are subject to right-censoring by the same censoring time. This setting is called common (univariate) right-censoring. It is illustrated in Figure 4.1 considering four-dimensional data. Here, for cluster k ($i, k \in \{1, \dots, n\}$ and $i \neq k$) only the fourth event time $T_{k,4}$ is observed. All other observations are equal after being censored at the end of the study. From Figure 4.1 it is clear that different joint censoring states need to be distinguished among clusters. Based on the observed data $\boldsymbol{Y}_i = (Y_{i,1}, \dots, Y_{i,4})$ and $\boldsymbol{\delta}_i = (\delta_{i,1}, \dots, \delta_{i,4})$ the latter are defined as follows in the four-dimensional case:

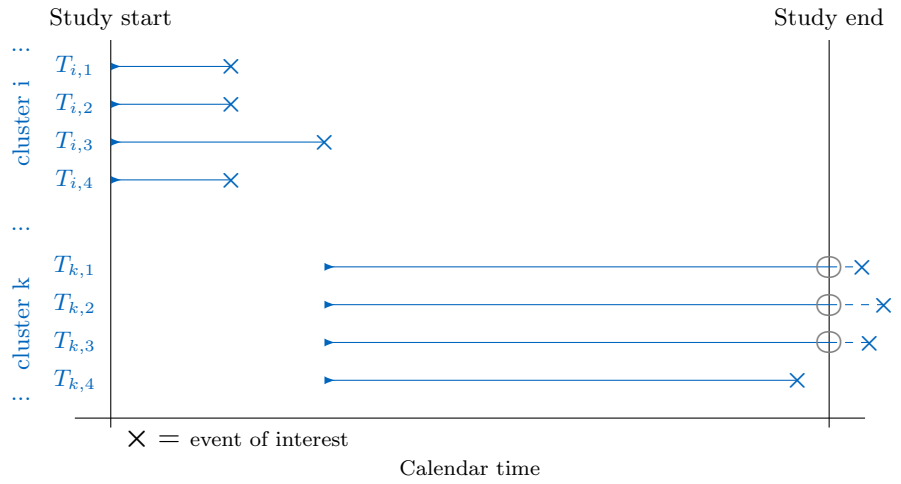


Figure 4.1: Illustration of four-dimensional event time data subject to common right-censoring.

no censoring:	$\Delta_i(1,2,3,4) \coloneqq \delta_{i,1}\delta_{i,2}\delta_{i,3}\delta_{i,4}$
all components censored:	$\Delta_i \coloneqq \prod_{j=1}^{4}(1-\delta_{i,j})$
p-th component not censored:	$\Delta_i(p) \coloneqq \delta_{i,p}\prod_{j=1;j\neq p}^{4}(1-\delta_{i,j})$
p-th, q-th component not censored for $p \neq q$:	$\Delta_i(p,q) \coloneqq \delta_{i,p}\delta_{i,q}\prod_{j=1;j\neq p,q}^{4}(1-\delta_{i,j})$
p-th, q-th, v-th component not censored for $w \neq p,q,v$ and $p \neq q \neq v$:	$\Delta_i(p,q,v) \coloneqq \delta_{i,p}\delta_{i,q}\delta_{i,v}(1-\delta_{i,w})$

4.2.3 Likelihood estimation for four-dimensional event time data

The goal is to develop for d-dimensional event time data as described in the previous section a likelihood estimation strategy. From now on we assume $d = 4$. Let $\mathbb{C}$ with density $\mathbb{c}$ be the survival copula describing the vector (U_1, U_2, U_3, U_4), which corresponds to the vector of observed times (Y_1, Y_2, Y_3, Y_4), i.e. $U_j \coloneqq S_j(Y_j)$ $(j = 1, \ldots, 4)$, where S_j is the survival function for event time T_j. At the moment, we assume S_j $(j = 1, \ldots, 4)$ to be known. Recall that in four dimensions there are only two possible R-vine structures: D-vines and C-vines (see Figure 4.2).

Similar to the univariate case outlined in Section 4.1.3 the joint censoring status of a cluster needs to be taken into account, when constructing an appropriate likelihood expression. Consider the observed data $u_{i,j} = S_j(y_{i,j})$ $(i = 1, \ldots, n$ and $j = 1, \ldots, 4)$ and assume that for example $\boldsymbol{\delta}_i = (1, 0, 0, 1)$. Thus, we have $\Delta_i(1, 4) = 1$ and all other joint censoring indicators equal zero. Then, $u_{i,1}$ and $u_{i,4}$ correspond to true event times. On the other hand, $u_{i,2}$ and $u_{i,3}$ correspond to censoring times, meaning that the copula data linked to the unknown true event times would take values smaller than $u_{i,2}$ and $u_{i,3}$. Thus, the contribution to the loglikelihood is given by

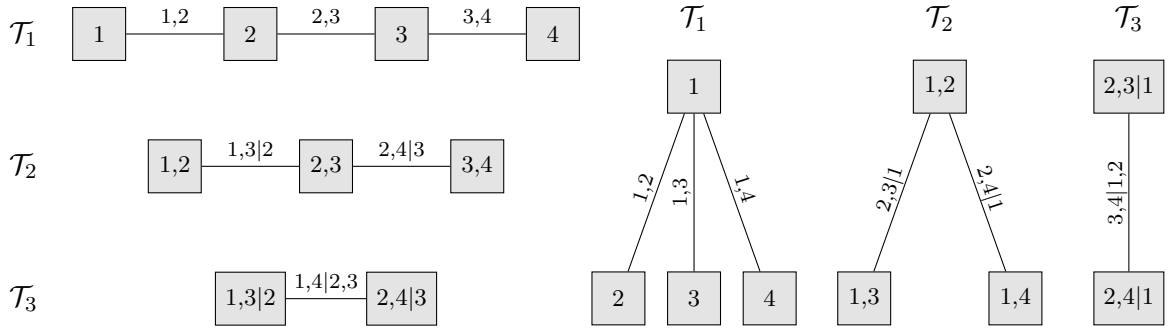


Figure 4.2: Two examples of a four-dimensional R-vine structure: a D-vine on the left and a C-vine on the right.

$$\ell_{i,4}\left(\boldsymbol{\theta}; \boldsymbol{u}_i, \boldsymbol{\delta}_i\right) = \log\left[\frac{\partial^2}{\partial u_{i,1}\partial u_{i,4}}\mathbb{C}\{u_{i,1}, S_2(y_{i,2}), S_3(y_{i,3}), u_{i,4}; \boldsymbol{\theta}\}\right]\Big|_{\substack{u_{i,1}=S_1(y_{i,1})\\ u_{i,4}=S_4(y_{i,4})}},$$

where $\boldsymbol{u}_i = (u_{i,1}, u_{i,2}, u_{i,3}, u_{i,4})$ and with $\boldsymbol{\theta}$ the vector collecting all parameters of the copula $\mathbb{C}$. In general, the contribution of the i-th cluster to the loglikelihood is given by

$$\begin{aligned}
&\ell_{i,4}\left(\boldsymbol{\theta}; \boldsymbol{u}_i, \boldsymbol{\delta}_i\right)\\
&\coloneqq \Delta_i \log\{\mathbb{C}\left(u_{i,1}, u_{i,2}, u_{i,3}, u_{i,4}; \boldsymbol{\theta}\right)\}\\
&+ \sum_{p=1}^{4} \Delta_i\left(p\right) \log\{\frac{\partial}{\partial u_{i,p}}\mathbb{C}\left(u_{i,1}, u_{i,2}, u_{i,3}, u_{i,4}; \boldsymbol{\theta}\right)\}\\
&+ \sum_{p\neq q} \Delta_i\left(p,q\right) \log\{\frac{\partial^2}{\partial u_{i,p}\partial u_{i,q}}\mathbb{C}\left(u_{i,1}, u_{i,2}, u_{i,3}, u_{i,4}; \boldsymbol{\theta}\right)\}\\
&+ \sum_{p\neq q\neq v} \Delta_i\left(p,q,v\right) \log\{\frac{\partial^3}{\partial u_{i,p}\partial u_{i,q}\partial u_{i,v}}\mathbb{C}\left(u_{i,1}, u_{i,2}, u_{i,3}, u_{i,4}; \boldsymbol{\theta}\right)\}\\
&+ \Delta_i\left(1,2,3,4\right) \log\{\mathbb{c}\left(u_{i,1}, u_{i,2}, u_{i,3}\, u_{i,4}; \boldsymbol{\theta}\right)\}.
\end{aligned}$$

The loglikelihood for four-dimensional time-to-event data subject to right-censoring, which is to be maximized with respect to $\boldsymbol{\theta}$, is therefore given by

$$\ell\left(\boldsymbol{\theta}; \boldsymbol{u}_1, \ldots, \boldsymbol{u}_n, \boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_n\right) \coloneqq \sum_{i=1}^{n} \ell_{i,4}\left(\boldsymbol{\theta}; \boldsymbol{u}_i, \boldsymbol{\delta}_i\right). \tag{4.7}$$

Massonnet et al. (2009) and Geerdens et al. (2016a) use this likelihood expression to model dependencies within the mastitis data, which will be discussed in detail in Section 4.3.2. Shih and Louis (1995) and Andersen (2005) consider similar versions for bivariate event time data.

Once we have decided on the R-vine structure to be used, we need the version of the partial derivatives in (4.7) in terms of pair-copula components. For instance, for the D-vine structure considered in Figure 4.2 we have

$$
\begin{aligned}
&\frac{\partial^2 \mathbb{C}\left(u_{i,1}, u_{i,2}, u_{i,3}, u_{i,4}\right)}{\partial u_{i,1} \partial u_{i,4}} \\
&= \int_0^{u_{i,2}} \int_0^{u_{i,3}} \mathbb{c}_{1,2}\left(u_{i,1}, v_{i,2}\right) \mathbb{c}_{2,3}\left(v_{i,2}, v_{i,3}\right) \mathbb{c}_{3,4}\left(v_{i,3}, u_{i,4}\right) \\
&\quad \times \mathbb{c}_{1,3;2}\{\mathbb{C}_{1|2}\left(u_{i,1}|v_{i,2}\right), \mathbb{C}_{3|2}\left(v_{i,3}|v_{i,2}\right)\} \\
&\quad \times \mathbb{c}_{2,4;3}\{\mathbb{C}_{2|3}\left(v_{i,2}|v_{i,3}\right), \mathbb{C}_{4|3}\left(u_{i,4}|v_{i,3}\right)\} \\
&\quad \times \mathbb{c}_{1,4;2,3}\{\mathbb{C}_{1|2,3}\left(u_{i,1}|v_{i,2}, v_{i,3}\right), \mathbb{C}_{4|2,3}\left(u_{i,4}|v_{i,2}, v_{i,3}\right)\} \mathrm{d}v_{i,3} \mathrm{d}v_{i,2} \\
&= \int_0^{u_{i,2}} \int_0^{u_{i,3}} \mathbb{c}_{1,2}\left(u_{i,1}, v_{i,2}\right) \mathbb{c}_{2,3}\left(v_{i,2}, v_{i,3}\right) \mathbb{c}_{3,4}\left(v_{i,3}, u_{i,4}\right) \\
&\quad \times \mathbb{c}_{1,3;2}\{h_{1|2}\left(u_{i,1}|v_{i,2}\right), h_{3|2}\left(v_{i,3}|v_{i,2}\right)\} \\
&\quad \times \mathbb{c}_{2,4;3}\{h_{2|3}\left(v_{i,2}|v_{i,3}\right), h_{4|3}\left(u_{i,4}|v_{i,3}\right)\} \\
&\quad \times \mathbb{c}_{1,4;2,3}\left[h_{1|3;2}\{h_{1|2}\left(u_{i,1}|v_{i,2}\right) \middle| h_{3|2}\left(v_{i,3}|v_{i,2}\right)\},\right. \\
&\qquad\qquad \left. h_{4|2;3}\{h_{4|3}\left(u_{i,4}|v_{i,3}\right) \middle| h_{2|3}\left(v_{i,2}|v_{i,3}\right)\}\right] \mathrm{d}v_{i,3} \mathrm{d}v_{i,2}.
\end{aligned}
$$

The complete collection of D- and C-vine equivalents of the partial derivatives is derived in Barthel (2015, Chapter 3) and is given in Appendix B.1.1.

Practical implementation

We end this section by two remarks concerning the practical implementation of the presented optimization problem.

First, note that in practice, the marginal survival functions, which are assumed to be known in the above discussion, are typically unknown. Clearly, full maximum likelihood optimization of all univariate marginal and copula parameters could be performed. To lower the computational effort and to increase model flexibility, we use the two-stage estimation procedure described in Shih and Louis (1995). A parametric approach can be applied. In stage one, we assume $S_j(\cdot)$ to be known up to some parameter vector $\boldsymbol{\alpha}_j$, i.e. $S_j(\cdot) = S_j(\cdot, \boldsymbol{\alpha}_j)$ $(j = 1, \ldots, 4)$. We obtain the maximum likelihood estimate (MLE) $\hat{\boldsymbol{\alpha}}_j$ of $\boldsymbol{\alpha}_j$ as described in Section 4.1.3 and calculate $\hat{u}_{i,j} = S_j(y_{i,j}, \hat{\boldsymbol{\alpha}}_j)$ $(i = 1, \ldots, n$ and $j = 1, \ldots, 4)$. In stage two, we replace $u_{i,j}$ by the pseudo observation $\hat{u}_{i,j}$ $(i = 1, \ldots, n$ and $j = 1, \ldots, 4)$ and maximize the loglikelihood in (4.7) with respect to $\boldsymbol{\theta}$. Alternatively, a more flexible semiparametric approach can be applied. In stage one, we estimate the marginals nonparametrically as explained in Section 4.1.3. We obtain the Kaplan-Meier estimate (KME) $\hat{S}_j(\cdot)$ of $S_j(\cdot)$ $(j = 1, \ldots, 4)$ and calculate the pseudo observations $\hat{u}_{i,j} = \hat{S}_j(y_{i,j})$. In stage two, we use the latter as substitutes for $u_{i,j}$ and maximize the loglikelihood in (4.7) with respect to $\boldsymbol{\theta}$.

Second, due to right-censoring the use of single and double integrals and hence numerical integration cannot be avoided when evaluating the loglikelihood. Thus, appropriate starting

values are indispensable for a reasonable trade-off between numerical demand and accuracy of the estimates. Due to the rapidly increasing number of parameters for R-vine copulas in higher dimensions this issue also arises for complete data. Herein, the so-called sequential estimation approach of Dißmann et al. (2013) is usually applied. It splits up a d-dimensional estimation problem into $d(d-1)/2$ bivariate ones. First, the parameters of the $d-1$ bivariate copulas in $\mathcal{T}_1$ are estimated. Next, the parameter estimates are used to obtain estimates of the h-functions. These estimates are needed as arguments in the pair-copulas in $\mathcal{T}_2$ when estimating the $d-2$ copula parameters in $\mathcal{T}_2$, etc. Hobæk-Haff et al. (2013) and Stöber and Schepsmeier (2013) provide asymptotic properties for this approach. Since it makes the estimation of high-dimensional R-vine copula models tractable and computationally easy while showing excellent estimation performance, analysis for complete data often solely rely on the sequential estimation approach.

In the setting with right-censored quadruple data, we can mimic this idea and estimate the parameters of the three bivariate copulas in $\mathcal{T}_1$ separately by using the bivariate version of the loglikelihood given in (4.7). However, by construction the arguments in $\mathcal{T}_2$ and $\mathcal{T}_3$ are not directly associated with observed (event or censored) times. As a consequence, estimation via the two-dimensional version of (4.7) is no longer feasible. Instead, after having obtained the parameter estimates for $\mathcal{T}_1$, we substitute them in the loglikelihood (4.7), which we then maximize with respect to the remaining copula parameters in $\mathcal{T}_2$ and $\mathcal{T}_3$. By doing so, we achieve dimension reduction by at least 3 for $d = 4$. We refer to this approach as $\mathcal{T}_1$-sequential estimation. Finally, we use the estimates of the $\mathcal{T}_1$-sequential approach as starting values to solve the computationally heavy optimization problem with respect to all 6 parameters ($d = 4$) of the R-vine copula model simultaneously (step 2 in the two-stage estimation procedure of Shih and Louis (1995)).

For our calculations, we rely on standard optimization methods and the `VineCopula` package in `R` (Schepsmeier et al., 2017), in which the evaluation of h-functions, of the cumulative distribution function and of the density function is implemented for many parametric bivariate copulas.

4.2.4 Simulation study

We investigate the finite sample performance of the loglikelihood approach presented in Section 4.2.3 through an extensive simulation study. To cover a broad range of simulation settings while keeping the numerical effort for a large number of replications reasonable, we restrict ourselves to three dimensions. The goal is to assess the impact of right-censoring on R-vine copula based estimation of the within-cluster association. For this purpose, various degrees of right-censoring, different types of tail-dependence and different strengths of dependence are considered. Our investigations build on the elaborate simulation study in Barthel (2015, Chapter 4) (master's thesis). However, all simulations in Barthel et al. (2018c) and in this thesis are completely rerun considering modified and additional simulation settings, which will be outlined in the following section.

Considered scenarios

To generate multivariate right-censored time-to-event data with a dependence structure specified by an R-vine copula, we simulate in a first step complete copula data using the `R`-package

`VineCopula` (Schepsmeier et al., 2017). We assume the copula $\mathbb{C}$ to be an R-vine copula with density

$$\mathbb{c}\left(u_1, u_2, u_3\right) = \mathbb{c}_{1,2}\left(u_1, u_2\right) \mathbb{c}_{2,3}\left(u_2, u_3\right) \mathbb{c}_{1,3;2}\{\mathbb{C}_{1|2}\left(u_1|u_2\right), \mathbb{C}_{3|2}\left(u_3|u_2\right)\}.$$

Note that in dimension three, all R-vine structures are equivalent up to the labeling of the nodes. Here, the copulas $\mathbb{C}_{1,2}$ and $\mathbb{C}_{2,3}$ are assumed to arise from the same copula family. We investigate both the scenario of lower tail-dependent copulas using the Clayton family and the scenario of upper tail-dependent copulas using the Gumbel family. For ease of comparison, we take Kendall's τ to be the same in both tail-dependence scenarios; we set $\tau_{1,2} = 0.6$ and $\tau_{2,3} = 0.6$ assuming strong dependencies. We assume $\mathbb{C}_{1,3;2}$ to be a Frank copula, which has no tail dependence, with moderate dependence $\tau_{1,3;2} = 0.3$. Two extra simulation settings with $\tau_{1,2} = \tau_{2,3} = \tau_{1,3;2} = 0.1$ (weak dependencies) and $\tau_{1,2} = \tau_{2,3} = \tau_{1,3;2} = 0.3$ (moderate dependencies) are included in Appendix B.2. The three copula families are common choices covering the three standard tail-dependence scenarios for bivariate data. Recall that in an R-vine copula model, these families can be arbitrarily combined allowing for complex dependence structures such as asymmetric tail-dependence behavior. While we focus on Archimedean copulas as building blocks of the considered R-vine copula models, in Barthel (2015) a Gaussian copula is assumed in tree level $\mathcal{T}_2$ showing similar estimation results. The scenarios of weak and moderate dependencies are not considered in Barthel (2015).

In a second step, the inverse probability integral transform is applied to the marginal copula data to obtain the true event times. Note that the proposed modeling strategy handles marginal and dependence modeling separately with no restrictions with regard to the marginal estimation. Thus, the settings for the marginal survival functions mainly serve the purpose to define the transformation from copula data to data on the actual time scale without distorting the dependence structure, which is our focus. Given that the Weibull is a commonly used parametric survival function, we assume this form for the margins of the event times as well as for the censoring mechanism, i.e. $S\left(t\right) = \exp\left(-\left(\frac{t}{\lambda}\right)^{\alpha}\right)$ with shape parameter α and scale parameter λ (in accordance with the parametrization used in `R`). The parameter choices are given in Table 4.1 and are inspired by the marginal estimates of the trivariate tumorigenesis data in Mantel et al. (1977). The latter also motivated the extensive simulation study in Barthel (2015, Section 4.1.1).

Table 4.1: Specification of the Weibull parameters of the survival function for each of the event times T_1, T_2, T_3 and of the two common censoring distributions leading to 25% and 65%, respectively, overall common right-censoring. Further, the individual censoring rates for each of the three margins are shown.

		Event times			Censoring times	
		T_1	T_2	T_3	25%	65%
Weibull parameters	α	3.39	4.20	3.53	6.72	6.72
	λ	3.32	2.21	2.68	3.11	2.17
Marginal censoring		52%	12%	29%	$\times$	
		82%	49%	67%		$\times$

To assess the effect of censoring, we investigate the performance of the estimation procedure for complete data as well as for a moderate overall censoring rate of 25% and for a heavy censoring rate of 65%. Note that the margins are affected to a different extent by the censoring mechanism as caused by distinct survival functions.

Finally, the observed data are obtained by taking the minima of the true event times and the corresponding censoring times. To this data we apply a two-stage approach for known margins as well as for parametrically (MLE) and nonparametrically (KME) estimated margins as described in Section 4.2.3. In case of complete event time data, we use the empirical distribution functions (ECDF) as nonparametric estimates for the marginals. All scenarios are investigated for samples of size 200 and 500. Each sample is replicated 200 times.

Results

We visualize the results of the simulations in Figure 4.3 and Figure 4.4, where the true Kendall's τ values are indicated by a horizontal line. Figure 4.3 shows satisfactory performance of the estimators when common right-censoring is present, even in case of heavy censoring (65%). The two-stage approaches with (non)parametrically estimated margins benefit the most from an increasing sample size. In particular, due to the comparable performance of the parametric and the semiparametric estimation approach, the latter qualifies as an appropriate tool when working with real data. It allows a flexible estimation of the marginals and excludes the risk to misspecify the underlying parametric models. Figure 4.4 shows the censoring effect. Comparing the first and second row illustrates the impact of the marginal censoring rates. Given that event time T_1 is affected most by right-censoring as shown in Table 4.1 we indeed expect that $\tau_{2,3}$ can be estimated in a more accurate way than $\tau_{1,2}$. Also, the method is more sensitive to a higher common right-censoring rate, especially when estimating the parameters of a lower tail-dependent copula, as can be seen by comparing the left-hand side and right-hand side of Figure 4.4. This is due to the lack of information in the data for small copula values, i.e. high event times (see also Figure 4.6). Overall, we can conclude that the presented method is on target for all investigated parameters in the underlying R-vine copula models.

A detailed summary of the simulation results can be found in Table 4.2 and Table 4.3 (Clayton for $\mathcal{T}_1$ and Frank for $\mathcal{T}_2$) and Table 4.4 and Table 4.5 (Gumbel for $\mathcal{T}_1$ and Frank for $\mathcal{T}_2$). Here, θ is the true parameter value, $\bar{\theta}$ is the mean estimate, $\hat{b}(\bar{\theta})$ is the estimated bias, $s^2(\bar{\theta})$ is the estimated squared standard error and $\hat{\text{mse}}(\bar{\theta})$ is the estimated mean squared error of $\bar{\theta}$. The same performance measures are given for the corresponding Kendall's τ values. Table B.2 to Table B.9 in Appendix B.2 show similar results for the two extra simulation settings considering weak and moderate dependencies for all three bivariate copulas.

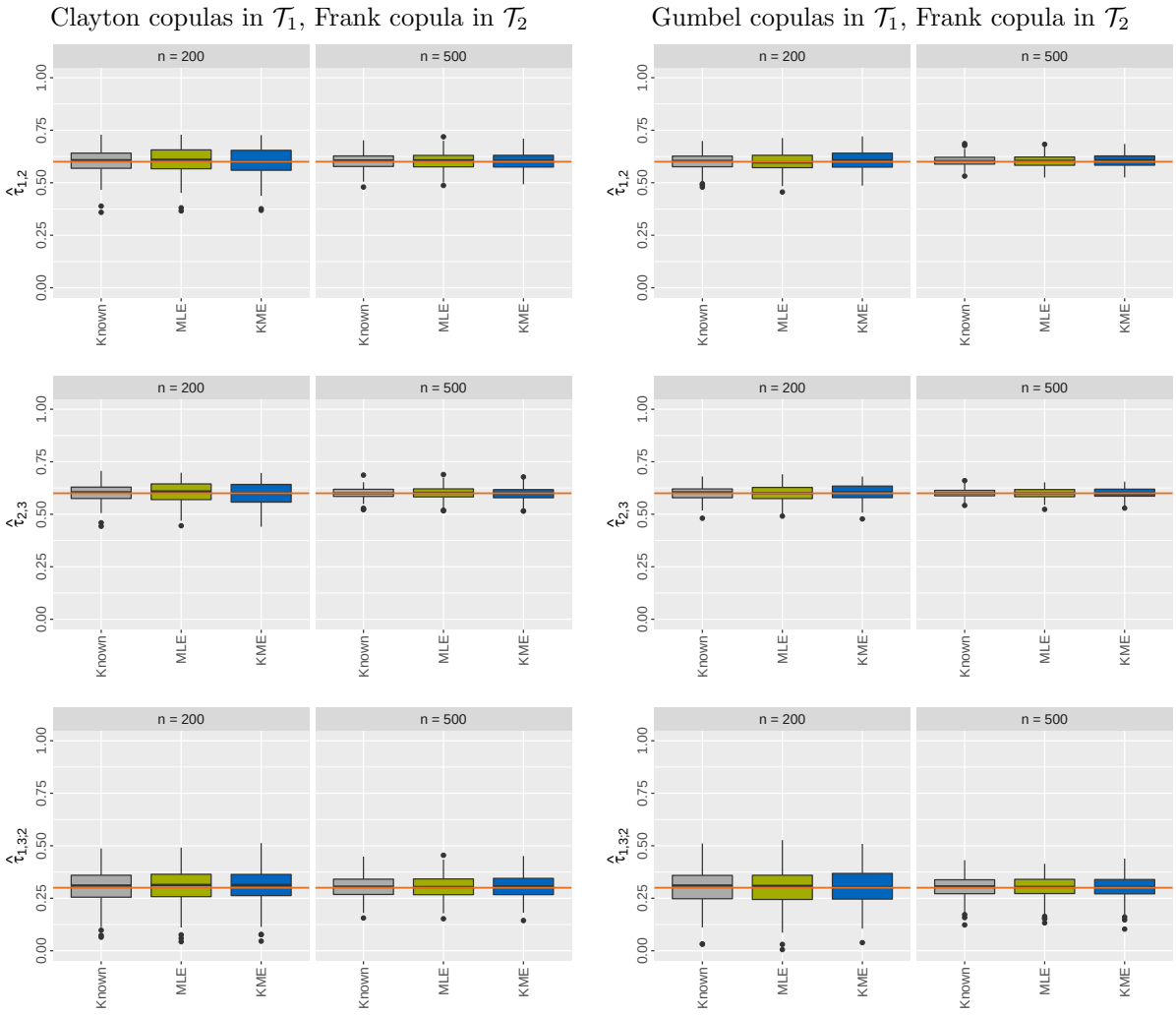


Figure 4.3: Boxplots of the estimated Kendall's τ values for 65% common right-censored event time data with Clayton copulas (left) and Gumbel copulas (right) in $\mathcal{T}_1$, true $\tau_{1,2} = 0.6$, $\tau_{2,3} = 0.6$, $\tau_{1,3;2} = 0.3$ and sample sizes 200 and 500. Known margins, parametrically estimated (MLE) and nonparametrically (KME) estimated margins are considered.

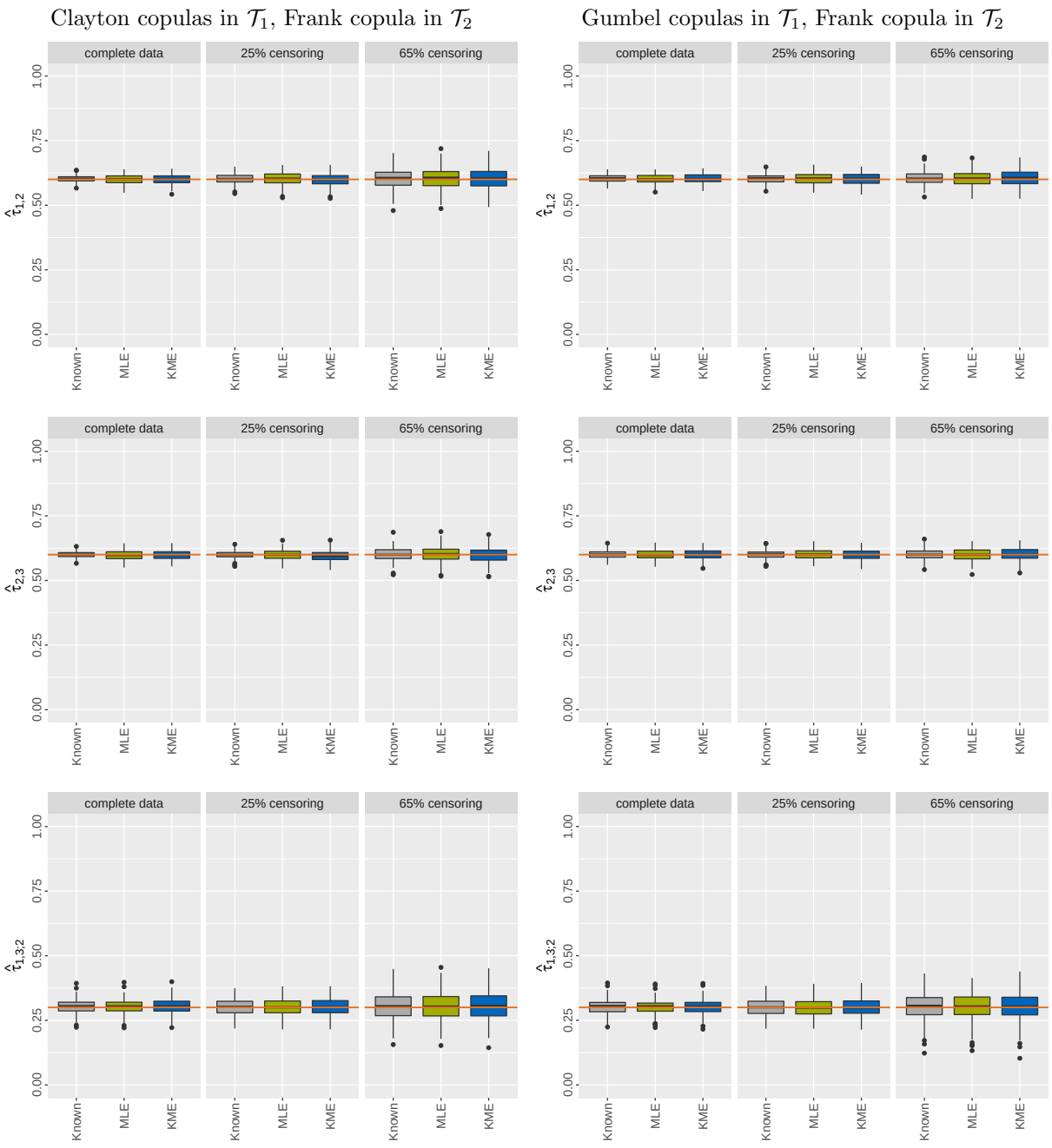


Figure 4.4: Boxplots of the estimated Kendall's τ values for an increasing percentage of common right-censoring with Clayton copulas (left) and Gumbel copulas (right) in $\mathcal{T}_1$, true $\tau_{1,2} = 0.6$, $\tau_{2,3} = 0.6$, $\tau_{1,3;2} = 0.3$ and sample size 500. Known margins, parametrically estimated margins (MLE) and nonparametrically estimated margins (ECDF/KME) are considered.

Table 4.2: Performance measures for the estimation of the copula parameters and Kendall's τ values in case of 65% common right-censored event time data with sample sizes 200 and 500. The copula combination Clayton (C), Clayton (C), Frank (F) with true $\tau_{1,2} = 0.6$, $\tau_{2,3} = 0.6$ and $\tau_{1,3;2} = 0.3$ is investigated. Known margins, parametrically estimated margins (MLE) and nonparametrically estimated margins (KME) are considered.

				Copula parameter					Kendall's τ				
				θ	$\bar{\theta}$	$\hat{b}(\bar{\theta})$	$s^2(\bar{\theta})$	$\hat{\text{mse}}(\bar{\theta})$	τ	$\bar{\tau}$	$\hat{b}(\bar{\tau})$	$s^2(\bar{\tau})$	$\hat{\text{mse}}(\bar{\tau})$
$n = 200$, 65% censoring	Known	C	$\theta_{1,2}$	3.00	3.14	0.1430	0.5762	0.5966	0.60	0.60	0.0025	0.0036	0.0036
		C	$\theta_{2,3}$	3.00	3.06	0.0609	0.2499	0.2536	0.60	0.60	0.0009	0.0016	0.0016
		F	$\theta_{1,3;2}$	2.92	3.03	0.1155	0.8131	0.8264	0.30	0.31	0.0052	0.0062	0.0062
	MLE	C	$\theta_{1,2}$	3.00	3.22	0.2247	0.6998	0.7504	0.60	0.61	0.0073	0.0040	0.0041
		C	$\theta_{2,3}$	3.00	3.12	0.1240	0.3782	0.3936	0.60	0.60	0.0039	0.0024	0.0024
		F	$\theta_{1,3;2}$	2.92	3.04	0.1190	0.8641	0.8783	0.30	0.31	0.0052	0.0066	0.0066
	KME	C	$\theta_{1,2}$	3.00	3.15	0.1542	0.6681	0.6919	0.60	0.60	0.0021	0.0041	0.0041
		C	$\theta_{2,3}$	3.00	3.05	0.0455	0.3857	0.3877	0.60	0.60	-0.0025	0.0025	0.0026
		F	$\theta_{1,3;2}$	2.92	3.07	0.1548	0.8911	0.9151	0.30	0.31	0.0081	0.0066	0.0067
$n = 500$, 65% censoring	Known	C	$\theta_{1,2}$	3.00	3.10	0.0983	0.2246	0.2343	0.60	0.60	0.0044	0.0013	0.0013
		C	$\theta_{2,3}$	3.00	3.03	0.0318	0.1083	0.1093	0.60	0.60	0.0008	0.0007	0.0007
		F	$\theta_{1,3;2}$	2.92	3.00	0.0855	0.3681	0.3754	0.30	0.31	0.0052	0.0027	0.0028
	MLE	C	$\theta_{1,2}$	3.00	3.11	0.1143	0.2703	0.2833	0.60	0.61	0.0051	0.0015	0.0015
		C	$\theta_{2,3}$	3.00	3.05	0.0460	0.1364	0.1386	0.60	0.60	0.0016	0.0008	0.0008
		F	$\theta_{1,3;2}$	2.92	3.00	0.0855	0.3771	0.3844	0.30	0.31	0.0052	0.0028	0.0028
	KME	C	$\theta_{1,2}$	3.00	3.09	0.0887	0.2713	0.2791	0.60	0.60	0.0030	0.0016	0.0016
		C	$\theta_{2,3}$	3.00	3.00	0.0042	0.1323	0.1323	0.60	0.60	-0.0017	0.0008	0.0008
		F	$\theta_{1,3;2}$	2.92	3.02	0.0988	0.3867	0.3964	0.30	0.31	0.0063	0.0029	0.0029

Table 4.3: Performance measures for the estimation of the copula parameters and Kendall's τ values in case of complete and 25% common right-censored event time data with sample size 500. The copula combination Clayton (C), Clayton (C), Frank (F) with true $\tau_{1,2} = 0.6$, $\tau_{2,3} = 0.6$ and $\tau_{1,3;2} = 0.3$ is investigated. Known margins, parametrically estimated margins (MLE) and nonparametrically estimated margins (KME) are considered.

				Copula parameter					Kendall's τ				
				θ	$\bar{\theta}$	$\hat{b}(\bar{\theta})$	$s^2(\bar{\theta})$	$\hat{\text{mse}}(\bar{\theta})$	τ	$\bar{\tau}$	$\hat{b}(\bar{\tau})$	$s^2(\bar{\tau})$	$\hat{\text{mse}}(\bar{\tau})$
$n = 500$, 25% censoring	Known	C	$\theta_{1,2}$	3.00	3.03	0.0282	0.0537	0.0545	0.60	0.60	0.0014	0.0003	0.0003
		C	$\theta_{2,3}$	3.00	3.00	-0.0021	0.0322	0.0322	0.60	0.60	-0.0007	0.0002	0.0002
		F	$\theta_{1,3;2}$	2.92	2.94	0.0275	0.1431	0.1439	0.30	0.30	0.0015	0.0011	0.0011
	MLE	C	$\theta_{1,2}$	3.00	3.03	0.0345	0.0872	0.0884	0.60	0.60	0.0014	0.0006	0.0006
		C	$\theta_{2,3}$	3.00	3.00	-0.0028	0.0599	0.0599	0.60	0.60	-0.0012	0.0004	0.0004
		F	$\theta_{1,3;2}$	2.92	2.95	0.0292	0.1478	0.1487	0.30	0.30	0.0017	0.0011	0.0011
	KME	C	$\theta_{1,2}$	3.00	2.99	-0.0113	0.0904	0.0905	0.60	0.60	-0.0024	0.0006	0.0006
		C	$\theta_{2,3}$	3.00	2.94	-0.0629	0.0621	0.0661	0.60	0.59	-0.0061	0.0004	0.0005
		F	$\theta_{1,3;2}$	2.92	2.96	0.0391	0.1497	0.1512	0.30	0.30	0.0025	0.0011	0.0012
$n = 500$, complete data	Known	C	$\theta_{1,2}$	3.00	3.04	0.0364	0.0235	0.0248	0.60	0.60	0.0025	0.0001	0.0002
		C	$\theta_{2,3}$	3.00	3.00	0.0036	0.0239	0.0239	0.60	0.60	-0.0001	0.0002	0.0002
		F	$\theta_{1,3;2}$	2.92	2.96	0.0457	0.0916	0.0937	0.30	0.30	0.0035	0.0007	0.0007
	MLE	C	$\theta_{1,2}$	3.00	3.01	0.0110	0.0514	0.0515	0.60	0.60	0.0001	0.0003	0.0003
		C	$\theta_{2,3}$	3.00	2.98	-0.0247	0.0517	0.0523	0.60	0.60	-0.0028	0.0003	0.0003
		F	$\theta_{1,3;2}$	2.92	2.96	0.0408	0.0912	0.0929	0.30	0.30	0.0030	0.0007	0.0007
	ECDF	C	$\theta_{1,2}$	3.00	3.02	0.0153	0.0543	0.0545	0.60	0.60	0.0004	0.0003	0.0003
		C	$\theta_{2,3}$	3.00	2.98	-0.0176	0.0551	0.0555	0.60	0.60	-0.0023	0.0004	0.0004
		F	$\theta_{1,3;2}$	2.92	2.97	0.0481	0.0977	0.1000	0.30	0.30	0.0036	0.0007	0.0008

Table 4.4: Performance measures for the estimation of the copula parameters and Kendall's τ values in case of 65% common right-censored event time data with sample sizes 200 and 500. The copula combination Gumbel (G), Gumbel (G), Frank (F) with true $\tau_{1,2} = 0.6$, $\tau_{2,3} = 0.6$ and $\tau_{1,3;2} = 0.3$ is investigated. Known margins, parametrically estimated margins (MLE) and nonparametrically estimated margins (KME) are considered.

				Copula parameter					Kendall's τ				
				θ	$\bar{\theta}$	$\hat{b}(\bar{\theta})$	$s^2(\bar{\theta})$	$\hat{\text{mse}}(\bar{\theta})$	τ	$\bar{\tau}$	$\hat{b}(\bar{\tau})$	$s^2(\bar{\tau})$	$\hat{\text{mse}}(\bar{\tau})$
$n = 200$, 65% censoring	Known	G	$\theta_{1,2}$	2.50	2.53	0.0265	0.0641	0.0648	0.60	0.60	0.0003	0.0016	0.0016
		G	$\theta_{2,3}$	2.50	2.52	0.0201	0.0396	0.0400	0.60	0.60	0.0007	0.0010	0.0010
		F	$\theta_{1,3;2}$	2.92	2.99	0.0705	0.8978	0.9028	0.30	0.30	0.0008	0.0069	0.0069
	MLE	G	$\theta_{1,2}$	2.50	2.52	0.0158	0.0827	0.0830	0.60	0.60	-0.0026	0.0021	0.0021
		G	$\theta_{2,3}$	2.50	2.53	0.0250	0.0570	0.0577	0.60	0.60	0.0004	0.0014	0.0014
		F	$\theta_{1,3;2}$	2.92	3.00	0.0783	0.9986	1.0048	0.30	0.30	0.0009	0.0076	0.0076
	KME	G	$\theta_{1,2}$	2.50	2.58	0.0820	0.1069	0.1136	0.60	0.61	0.0067	0.0023	0.0024
		G	$\theta_{2,3}$	2.50	2.56	0.0558	0.0634	0.0665	0.60	0.60	0.0049	0.0015	0.0015
		F	$\theta_{1,3;2}$	2.92	3.00	0.0805	0.9979	1.0044	0.30	0.30	0.0011	0.0075	0.0075
$n = 500$, 65% censoring	Known	G	$\theta_{1,2}$	2.50	2.54	0.0376	0.0291	0.0305	0.60	0.60	0.0042	0.0007	0.0007
		G	$\theta_{2,3}$	2.50	2.51	0.0106	0.0170	0.0171	0.60	0.60	0.0006	0.0004	0.0004
		F	$\theta_{1,3;2}$	2.92	2.97	0.0494	0.3800	0.3825	0.30	0.30	0.0020	0.0030	0.0030
	MLE	G	$\theta_{1,2}$	2.50	2.53	0.0324	0.0376	0.0386	0.60	0.60	0.0029	0.0009	0.0009
		G	$\theta_{2,3}$	2.50	2.51	0.0107	0.0243	0.0244	0.60	0.60	0.0002	0.0006	0.0006
		F	$\theta_{1,3;2}$	2.92	2.97	0.0507	0.3892	0.3918	0.30	0.30	0.0021	0.0030	0.0030
	KME	G	$\theta_{1,2}$	2.50	2.55	0.0524	0.0445	0.0472	0.60	0.61	0.0056	0.0010	0.0011
		G	$\theta_{2,3}$	2.50	2.52	0.0162	0.0254	0.0257	0.60	0.60	0.0010	0.0006	0.0006
		F	$\theta_{1,3;2}$	2.92	2.98	0.0578	0.4216	0.4249	0.30	0.30	0.0025	0.0033	0.0033

Table 4.5: Performance measures for the estimation of the copula parameters and Kendall's τ values in case of complete and 25% common right-censored event time data with sample size 500. The copula combination Gumbel (G), Gumbel (G), Frank (F) with true $\tau_{1,2} = 0.6$, $\tau_{2,3} = 0.6$ and $\tau_{1,3;2} = 0.3$ is investigated. Known margins, parametrically estimated margins (MLE) and nonparametrically estimated margins (ECDF/KME) are considered.

				Copula parameter					Kendall's τ				
				θ	$\bar{\theta}$	$\hat{b}(\bar{\theta})$	$s^2(\bar{\theta})$	$\hat{\text{mse}}(\bar{\theta})$	τ	$\bar{\tau}$	$\hat{b}(\bar{\tau})$	$s^2(\bar{\tau})$	$\hat{\text{mse}}(\bar{\tau})$
$n = 500$, 25% censoring	Known	G	$\theta_{1,2}$	2.50	2.52	0.0181	0.0129	0.0132	0.60	0.60	0.0021	0.0003	0.0003
		G	$\theta_{2,3}$	2.50	2.51	0.0052	0.0100	0.0101	0.60	0.60	0.0002	0.0003	0.0003
		F	$\theta_{1,3;2}$	2.92	2.93	0.0107	0.1517	0.1518	0.30	0.30	0.0000	0.0012	0.0012
	MLE	G	$\theta_{1,2}$	2.50	2.52	0.0207	0.0198	0.0203	0.60	0.60	0.0021	0.0005	0.0005
		G	$\theta_{2,3}$	2.50	2.51	0.0084	0.0148	0.0148	0.60	0.60	0.0004	0.0004	0.0004
		F	$\theta_{1,3;2}$	2.92	2.92	0.0027	0.1529	0.1529	0.30	0.30	-0.0007	0.0012	0.0012
	KME	G	$\theta_{1,2}$	2.50	2.52	0.0193	0.0207	0.0210	0.60	0.60	0.0018	0.0005	0.0005
		G	$\theta_{2,3}$	2.50	2.50	0.0018	0.0155	0.0155	0.60	0.60	-0.0007	0.0004	0.0004
		F	$\theta_{1,3;2}$	2.92	2.93	0.0106	0.1602	0.1603	0.30	0.30	-0.0001	0.0012	0.0012
$n = 500$, complete data	Known	G	$\theta_{1,2}$	2.50	2.52	0.0212	0.0078	0.0083	0.60	0.60	0.0029	0.0002	0.0002
		G	$\theta_{2,3}$	2.50	2.51	0.0064	0.0086	0.0086	0.60	0.60	0.0005	0.0002	0.0002
		F	$\theta_{1,3;2}$	2.92	2.96	0.0388	0.0997	0.1012	0.30	0.30	0.0028	0.0008	0.0008
	MLE	G	$\theta_{1,2}$	2.50	2.51	0.0136	0.0123	0.0125	0.60	0.60	0.0014	0.0003	0.0003
		G	$\theta_{2,3}$	2.50	2.50	-0.0000	0.0131	0.0131	0.60	0.60	-0.0008	0.0003	0.0003
		F	$\theta_{1,3;2}$	2.92	2.95	0.0276	0.0983	0.0991	0.30	0.30	0.0018	0.0008	0.0008
	ECDF	G	$\theta_{1,2}$	2.50	2.53	0.0255	0.0135	0.0141	0.60	0.60	0.0032	0.0003	0.0003
		G	$\theta_{2,3}$	2.50	2.51	0.0094	0.0144	0.0145	0.60	0.60	0.0006	0.0004	0.0004
		F	$\theta_{1,3;2}$	2.92	2.95	0.0279	0.1040	0.1048	0.30	0.30	0.0018	0.0008	0.0008

4.3 Estimating standard errors in the presence of right-censoring

In Section 4.2, focus was on likelihood based parameter estimation in R-vine copula models for four-dimensional event time data affected by common (univariate) right-censoring. The loglikelihood expression in (4.7) in terms of pair-copula components is already derived in Barthel (2015) (master's thesis) and is part of the publication Barthel et al. (2018c). Simulations in Barthel et al. (2018c) are completely rerun in particular considering additional weak and moderate dependence strengths of the pair-copulas. While for the simulation results (Section 4.2.4) standard errors of the parameter estimates could be empirically obtained from the replications in each simulation scenario, this is not feasible if interest is in real data. Thus, an important extension in Barthel et al. (2018c) as compared to Barthel (2015) (master's thesis) is the estimation of standard errors for the R-vine copula based modeling approach proposed in Section 4.2.

For complete data, Hobæk-Haff et al. (2013), Stöber and Schepsmeier (2013) and Schepsmeier and Stöber (2014) investigate asymptotic theory for R-vine copula based methodology and point out the challenges when interest is in the calculation of standard errors for parameter estimates in R-vine copula models. For right-censored data, the theory developed in these papers needs to be adapted and given the extra data complexity will become even more challenging. As concluded in Hobæk Haff et al. (2010) for complete data, we therefore opt as well for a more tractable alternative for finite samples and develop an appropriate resampling scheme to obtain bootstrap standard errors. Thus, an R-vine copula based parametric bootstrap algorithm is developed in Section 4.3.1. In Section 4.3.2, it will be applied when analyzing the mastitis data in detail.

4.3.1 Parametric bootstrap algorithm

Under common (univariate) right-censoring, as described in Section 4.2.2, standard errors for the estimated parameters of an R-vine copula can be obtained using a parametric bootstrap algorithm (Davison et al., 1997; Massonnet et al., 2009). A similar procedure as the one subsequently proposed is used in Geerdens et al. (2016a) for Joe-Hu copulas (Joe and Hu, 1996) in the context of multivariate right-censored event time data:

Step 1: Fit the R-vine copula model of interest to the copula data $(\hat{u}_{i,j}, \delta_{i,j})$, $i = 1, \ldots, n$ and $j = 1, \ldots, d$, where $\delta_{i,j} = \mathbb{1}(t_{i,j} \leq c_i)$, $\hat{u}_{i,j} = \hat{S}_j(y_{i,j})$ and $y_{i,j} = \min(t_{i,j}, c_i)$ with $\hat{S}_j$ the Kaplan-Meier estimate based on $(y_{i,j}, \delta_{i,j})$. Obtain the vector of copula parameter estimates $\hat{\boldsymbol{\theta}}$, which maximizes the corresponding loglikelihood function.

Step 2: Obtain the Kaplan-Meier estimate $\hat{G}$ of the censoring distribution G based on the observations $(\max(y_{i,1}, \ldots, y_{i,d}), 1 - \delta_{i,1} \cdot \ldots \cdot \delta_{i,d})$, $i = 1, \ldots, n$.

Step 3: Generate B bootstrap samples in the following way: For $b = 1, \ldots, B$, $\quad i = 1, \ldots, n$ and $j = 1, \ldots, d$,

Step 3.1: sample vine copula data $\left(u_{i,1}^{(b)}, \ldots, u_{i,d}^{(b)}\right)$ from the fitted R-vine copula model with parameter vector $\hat{\boldsymbol{\theta}}$.

Step 3.2: Generate event times $\left(t_{i,1}^{(b)}, \ldots, t_{i,d}^{(b)}\right)$ via $t_{i,j}^{(b)} = \hat{S}_j^{-1}\left(u_{i,j}^{(b)}\right)$.

Step 3.3: Generate independent censoring times $c_i^{(b)}$ from $\hat{G}$.

Step 3.4: Obtain observed data by setting $y_{i,j}^{(b)} = \min\left(t_{i,j}^{(b)}, c_i^{(b)}\right)$ and $\delta_{i,j}^{(b)} = \mathbb{1}\left(t_{i,j}^{(b)} \leq c_i^{(b)}\right)$.

Step 3.5: Set $\hat{u}_{i,j}^{(b)} = \hat{S}_j^{(b)}\left(y_{i,j}^{(b)}\right)$ with $\hat{S}_j^{(b)}$ the Kaplan-Meier estimate based on $\left(y_{i,j}^{(b)}, \delta_{i,j}^{(b)}\right)$.

Step 3.6: Given the bootstrap data $\left(\hat{u}_{i,j}^{(b)}, \delta_{i,j}^{(b)}\right)$, fit the R-vine copula model of interest by maximizing the corresponding loglikelihood function to obtain $\hat{\boldsymbol{\theta}}^{(b)}$ for bootstrap sample b.

Step 4: Calculate elementwise the empirical standard deviations of $\hat{\boldsymbol{\theta}}^{(1)}, \ldots, \hat{\boldsymbol{\theta}}^{(B)}$ to obtain bootstrap based standard errors for $\hat{\boldsymbol{\theta}}$.

4.3.2 Data application

In the following, the proposed parametric bootstrap algorithm will be used to obtain standard errors for the R-vine copula parameters estimated for the mastitis data. It will also help to validate nonparametric bootstrapping in the presence of heavy right-censoring as present in the mastitis data. The udder infection data of Laevens et al. (1997) already received considerable attention in a number of papers, for example Duchateau and Janssen (2008), Massonnet et al. (2009) and Geerdens et al. (2016a). The study aims to quantify the impact of mastitis on the milk production and the milk quality. For this, information on the time from parturition to infection is collected for the four udder quarters of a cow. The cow is the cluster and the infection times of the four udder quarters are the clustered data.

For the 407 primiparous cows in the study, the available data consist of the cow identification number, the minimum of the infection time and the censoring time (both in days) for each udder quarter as well as the corresponding censoring indicators. For example, for the first and last cow the data information is given by $\{1, (67, 67, 119, 67), (1, 1, 1, 1)\}$ and $\{407, (279, 279, 279, 263), (0, 0, 0, 1)\}$, respectively, where the ordering in a data quadruple corresponds to left front, right front, left rear and right rear. For the cow with ID 1, the true time

Table 4.6: Censoring patterns of the mastitis data.

#censored observations in a cluster	#cows
0	73
1	49
2	36
3	40
4	209
udder quarter	**percentage of censoring**
front left	64.37%
front right	64.37%
rear left	68.80%
rear right	67.08%

until mastitis infection is recorded for all four udder quarters. For the cow with ID 407, the time until infection is only known for the right rear udder quarter while all other observations are censored. Censoring in the mastitis data occurs at the level of the udder quarters and is common (univariate) in the sense that the same censoring time applies to all udder quarters of an individual cow. Table 4.6 summarizes information on the censoring patterns of the mastitis data. In total, censoring is present in about 66.15% of the observations. Before starting the discussion on model selection, it is important to note that the information loss due to right-censoring complicates accurate model selection and implies the need for careful comparison of possible models. Pairs plots can be used to demonstrate the information loss in a graphical way. Data points corresponding to the two front udder quarters are for example given by $(\hat{u}_i^{\mathrm{FL}}, \hat{u}_i^{\mathrm{FR}})$ with $\hat{u}_i^{\mathrm{FL}} = \hat{S}_{\mathrm{FL}}(y_i^{\mathrm{FL}})$ and $\hat{u}_i^{\mathrm{FR}} = \hat{S}_{\mathrm{FR}}(y_i^{\mathrm{FR}})$, where y_i^{FL} and y_i^{FR} are the observed infection times for the front left and the front right udder quarter of cow i ($i = 1, \ldots, 407$) and $\hat{S}_{\mathrm{FL}}$ and $\hat{S}_{\mathrm{FR}}$ are the corresponding Kaplan-Meier estimates as illustrated in Figure 4.5. Given the heavy censoring, the latter level off away from zero. Thus, scatter plots for the data points on the original time scale would contain only a few points in the upper right corner (of the first quadrant), which in turn leads to an almost empty lower left corner in all pairs plots on the copula scale as shown in Figure 4.6.

In the following, we investigate the dependence structure present in the mastitis data by fitting several R-vine copula models. According to Laevens (personal communication and Laevens et al. (1997)), there is no biological rule that could provide guidance for the dependence modeling. The primary goal is therefore to illustrate how the methodology introduced in Section 4.2 and

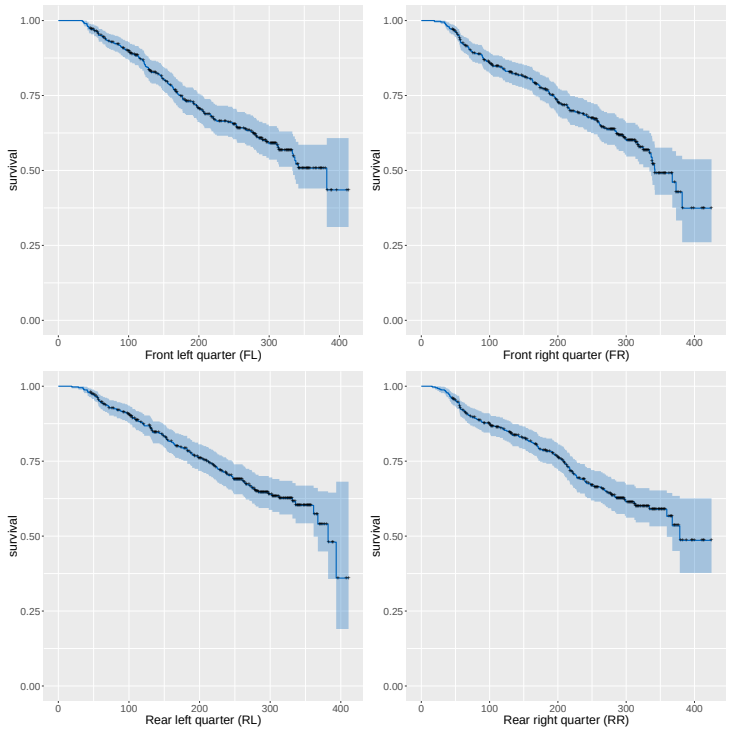


Figure 4.5: Kaplan-Meier estimates of the four udder quarters of the mastitis data illustrating the high censoring rate for all four marginals.

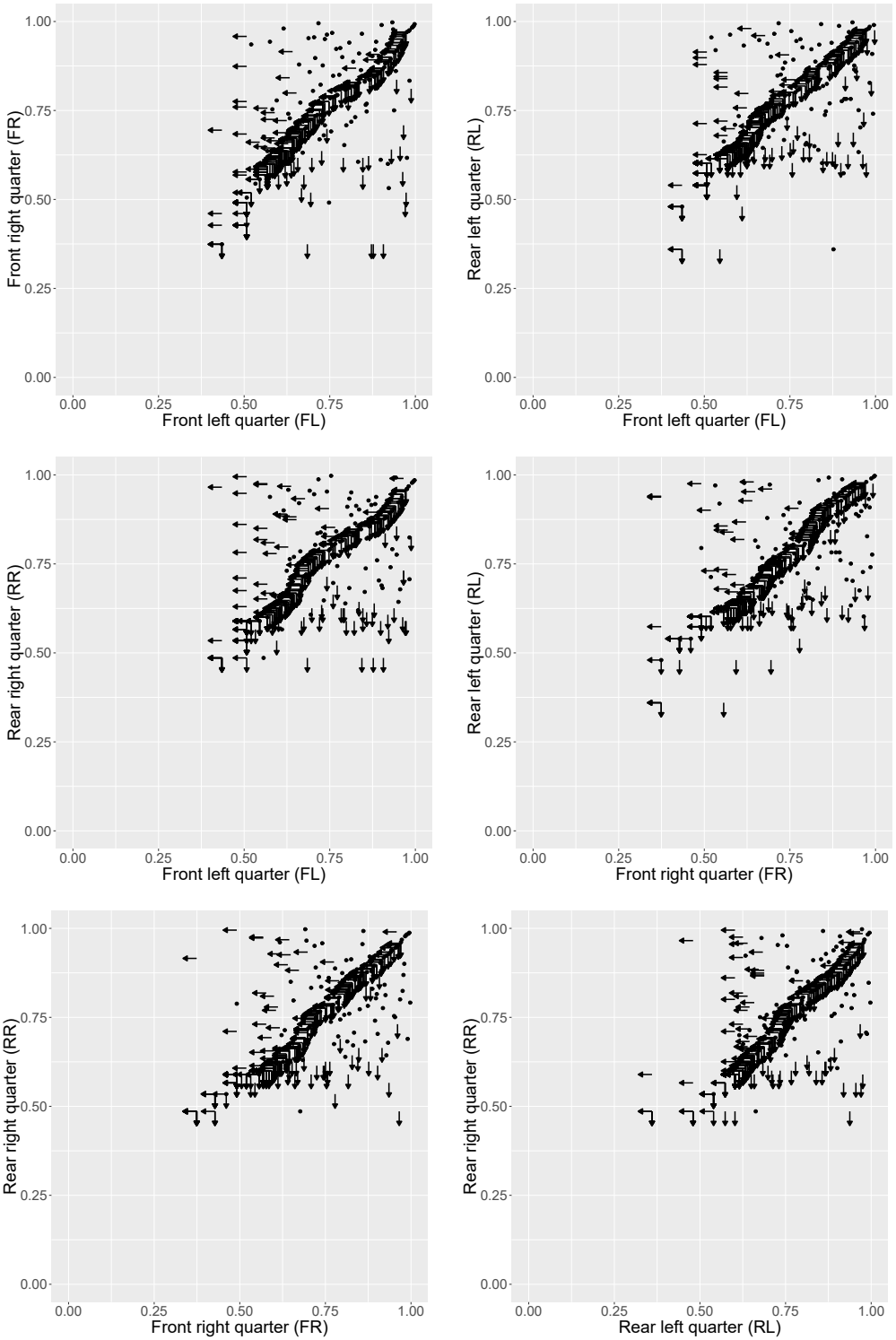


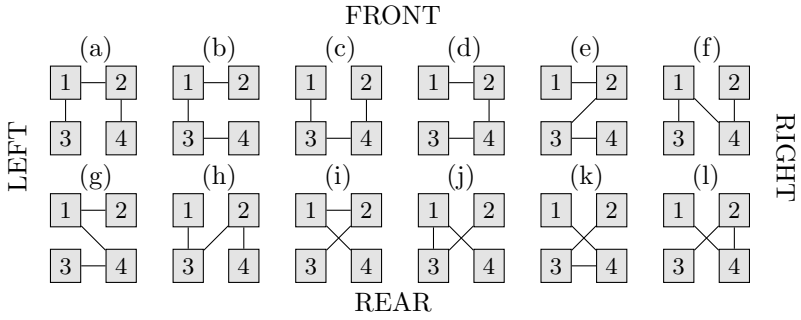
Figure 4.6: Pairs plots of all six udder pairs of the mastitis data based on pseudo observations generated via Kaplan-Meier estimates of the marginals (see Figure 4.5). The effect of right-censoring is reflected by the empty lower left corner in the pairs plots. Observations shown as • are event times for both udder quarters; ← is an event time only for the vertical axis; ↓ is an event time only for the horizontal axis; censored in both components is shown as $\overleftarrow{\downarrow}$.

the parametric bootstrap algorithm proposed in Section 4.3.1 can be applied to real data. In particular, we give insights about the effect of right-censoring in the context of copula estimation. Also, earlier investigations of the mastitis data using EAC models (Massonnet et al., 2009; Geerdens et al., 2016a) assumed equal correlations between all pairs of udder quarters inducing rather restrictive dependence patterns. We will see that these less elaborate models do not sufficiently fit the data.

Given the good performance of the two-stage semiparametric estimation in the simulation study in Section 4.2.4, we flexibly model the marginal survival functions using the Kaplan-Meier estimator and thus do not imply any parametric assumptions for the marginal data. We maximize the loglikelihood (4.7) over all copula parameters using the parameter estimates obtained from the $\mathcal{T}_1$-sequential approach as starting values. We consider R-vine copula models based on one parameter bivariate copulas such that all considered models have the same number of parameters (six). Thus, the AIC and BIC both select the model that gives the highest loglikelihood. We therefore use the loglikelihood for model selection. The use of the loglikelihood value as well as AIC and BIC for model selection in the context of two-stage semiparametric copula estimation for right-censored data has been studied in Chen et al. (2010) and in Geerdens et al. (2016a).

In the following, we assume a D-vine tree structure for the mastitis data. This choice reflects the temporal component of the infection, which may spread from one udder quarter to a neigh-

Table 4.7: D-vine structures considered for the mastitis data and corresponding loglikelihood values obtained via simultaneous estimation of all six parameters ($\mathcal{T}_1$-sequential estimation). Frank copulas are taken in $\mathcal{T}_2$ and $\mathcal{T}_3$.



D-vine		Common family in $\mathcal{T}_1$		
		Clayton	Gumbel	Frank
loglikelihood	(a)	-138.73 (-138.82)	-153.45 (-153.69)	-137.24 (-137.30)
	(b)	-139.19 (-139.25)	-148.91 (-148.99)	-136.05 (-136.10)
	(c)	-127.93 (-127.96)	-142.98 (-143.09)	-124.70 (-124.82)
	(d)	-138.47 (-138.56)	-147.99 (-148.25)	-134.91 (-134.95)
	(e)	-141.45 (-141.56)	-142.29 (-142.44)	-137.62 (-137.65)
	(f)	-129.78 (-129.88)	-145.89 (-145.99)	-130.10 (-130.16)
	(g)	-138.05 (-138.42)	-143.71 (-144.80)	-135.95 (-136.41)
	(h)	-132.53 (-132.60)	-145.33 (-145.40)	-131.17 (-131.28)
	(i)	-140.63 (-140.89)	-145.50 (-145.98)	-139.28 (-139.46)
	(j)	-133.55 (-132.57)	-143.81 (-143.88)	-134.71 (-134.82)
	(k)	-137.42 (-137.63)	-141.61 (-141.81)	-136.92 (-137.17)
	(l)	-134.37 (-134.50)	-145.29 (-145.73)	-134.23 (-134.29)

boring one. All possible 12 D-vines are represented in Table 4.7 by their first tree level, since the latter uniquely determines the whole D-vine structure. For all D-vine copulas the same type of copula is assumed in $\mathcal{T}_1$, however allowing for different parameters. We consider the Clayton, Gumbel or Frank copula, respectively. With this choice we account for possible lower and upper tail-dependence as well as for no tail-dependence inherent in the underlying data. In particular, asymmetric tail-dependence behavior is modeled through combination of the copula families in the considered D-vine copula models. Further, Frank copulas are taken in the two lower tree levels. By doing so, 36 models are investigated in total. Table 4.7 shows the loglikelihood values for the considered models obtained via simultaneous estimation of all six parameters. The loglikelihood values obtained through the $\mathcal{T}_1$-sequential estimation approach are shown in brackets. In general, D-vine structures that capture the dependence along the two flanks perform best, whereas D-vines with two diagonals would generally not be selected. Further, the choice of Frank and Clayton copulas in $\mathcal{T}_1$ is superior to the one of Gumbel copulas. Models with Frank copulas perform slightly better than those with Clayton copulas. Recall that for heavily censored copula data the lower left corner of a pairs plot is empty. However, there might be a considerable amount of observed event times in the upper right corner, where, therefore, most of the information is located (see Figure 4.6). Since Clayton and Frank copulas behave similar in the upper right corner, i.e. for early event times, it is clear that the information loss in case of heavy right-censoring makes it difficult to distinguish between Clayton and Frank copulas. In addition to the presented D-vine copula models, we fitted Gaussian D-vines, i.e. D-vine copulas of which all pair-copulas are Gaussian, to the data. The best performing D-vine structure results in a rather low loglikehood value of -148.92 for both global and $\mathcal{T}_1$-sequential estimation.

To further explore the above findings we consider for the D-vine with structure (c) (the best performing structure in Table 4.7) the 24 additional vine models (besides C-C-C, G-G-G and F-F-F in $\mathcal{T}_1$) having structure (c), where we allow combinations of Clayton, Gumbel and Frank copulas in $\mathcal{T}_1$. The loglikelihood values are listed in Table 4.8. The model with all dependencies captured by Frank copulas remains the best (see Table 4.7), but the loglikelihood values for models which combine Clayton and Frank copulas in $\mathcal{T}_1$ only are slightly smaller. The estimated copula parameters of the four best models are given in Table 4.9 together with their corresponding estimated Kendall's τ values and tail-dependence coefficients.

To obtain standard errors in Table 4.9, 100 replications according to the parametric bootstrapping algorithm in Section 4.3.1 are used, both for global likelihood estimation and for $\mathcal{T}_1$-sequential likelihood estimation. Detailed estimation results for the bootstrap samples given in Appendix B.3 show that using 100 bootstrap replications, the estimates for the standard error of the various parameters are already quite accurate and the empirical means of the bootstrap based parameter estimates are close to the corresponding parameters in the D-vine copula models according to which the bootstrap samples are generated. In case of parametric bootstrapping, we assume the estimated parametric model to be the true one. As a consequence the bootstrap samples and the loglikelihood values obtained after refitting the model in each replication are model dependent. Thus, in order to assess equal performance of the four best models in Table 4.9 nonparametric bootstrap replications are needed. We generate bootstrap

Table 4.8: Loglikelihood values obtained via simultaneous estimation of all six parameters (the $\mathcal{T}_1$-sequential estimation approach). The considered models all have D-vine structure (c) and combinations of Clayton (C), Gumbel (G) and Frank (F) copulas in $\mathcal{T}_1$. Frank copulas are taken in $\mathcal{T}_2$ and $\mathcal{T}_3$.

	Families in $\mathcal{T}_1$: $\text{fam}_{1,3}$–$\text{fam}_{3,4}$–$\text{fam}_{2,4}$		
loglikelihood	C–C–F -127.31 (-127.58)	C–F–C -125.73 (-125.83)	F–C–C -128.67 (-128.78)
	C–F–F -125.39 (-125.49)	F–C–F -127.81 (-127.90)	F–F–C -125.66 (-125.77)
	C–C–G -141.21 (-141.28)	C–G–C -137.64 (-137.66)	G–C–C -145.48 (-145.56)
	C–G–G -138.50 (-138.53)	G–C–G -155.25 (-155.28)	G–G–C -143.59 (-143.61)
	C–G–F -135.04 (-135.08)	C–F–G -137.86 (-138.14)	G–C–F -144.30 (-144.37)
	F–C–G -141.29 (-141.36)	G–F–C -141.28 (-141.47)	F–G–C -136.04 (-136.08)
	G–G–F -136.40 (-136.42)	G–F–G -139.75 (-139.77)	F–G–G -149.23 (-149.27)
	G–F–F -139.38 (-139.56)	F–G–F -132.76 (-132.77)	F–F–G -137.16 (-137.43)

samples by random sampling of 407 cow IDs with replacement. As an alternative the jackknife resampling method could be considered. As already seen in the previous analysis, right-censoring makes estimation less accurate and thus, the censoring percentage among the bootstrap samples has to be carefully monitored. To validate the accuracy of nonparametric bootstrapping in the presence of heavy right-censoring as present for the mastitis data, we perform a small simulation study designed from the D-vine copula fits in Table 4.9. As illustrated in Figure 4.7, in each simulation scenario we assume the corresponding D-vine copula fit to be the true model with parameter vector $\boldsymbol{\theta}$. From this model, we simulate S data sets $D_{\text{sim}s}$ ($s = 1, \ldots, S$) as described in Section 4.3.1 and obtain via refitting based on the corresponding D-vine copula specification estimated parameter vectors $\hat{\boldsymbol{\theta}}_{\text{sim}s}$. Further, for each simulated data set $D_{\text{sim}s}$ a nonparametric bootstrap is performed with bootstrap based samples $D^{(b)}_{\text{sim}s}$ ($b = 1, \ldots, B$), for which refitting based on the corresponding D-vine copula specification results in bootstrap based parameter estimates $\hat{\boldsymbol{\theta}}^{(b)}_{\text{sim}s}$. Proceeding this way for the four D-vine copula models in Table 4.9 using $\mathcal{T}_1$-sequential likelihood estimation and 100 bootstrap replications for each simulated data set, we find that the 90% confidence intervals of the bootstrap based parameter estimates sufficiently cover both the corresponding simulation based parameter estimates and the underlying true parameters. The empirical means of the bootstrap parameter estimates are close to the corresponding simulation based parameter estimates. Results for 16 different data sets are shown in Table B.10 to Table B.13 in Section B.3.

Table 4.9: Estimated copula parameters, Kendall's τ values and tail-dependence coefficients for the four best models fitted to the mastitis data with underlying D-vine structure (c). Results for both the $\mathcal{T}_1$-sequential estimation approach and for joint estimation of all six parameters are shown. Standard errors (in parenthesis) are obtained using the parametric bootstrap algorithm described in 4.3.1.

	$\mathcal{T}_1$-sequential estimation				Global estimation			
	logll	Parameter	Kendall's τ	Lower tail dependence	logll	Parameter	Kendall's τ	Lower tail dependence
F; $\hat{\theta}_{1,3}$	-124.82	6.38 (0.81)	0.53 (0.04)	–	-124.70	6.56 (0.80)	0.54 (0.04)	–
F; $\hat{\theta}_{3,4}$		6.34 (0.79)	0.53 (0.04)	–		6.34 (0.75)	0.53 (0.04)	–
F; $\hat{\theta}_{2,4}$		6.77 (0.80)	0.55 (0.04)	–		6.99 (0.77)	0.56 (0.03)	
F; $\hat{\theta}_{1,4;3}$		1.67 (0.57)	0.18 (0.06)	–		1.68 (0.55)	0.18 (0.06)	
F; $\hat{\theta}_{2,3;4}$		2.81 (0.57)	0.29 (0.05)	–		2.79 (0.55)	0.29 (0.05)	–
F; $\hat{\theta}_{1,2;3,4}$		3.72 (0.63)	0.37 (0.05)	–		3.71 (0.65)	0.37 (0.05)	–
C; $\hat{\theta}_{1,3}$	-125.49	3.60 (0.58)	0.64 (0.04)	0.82 (0.03)	-125.39	3.78 (0.58)	0.65 (0.04)	0.83 (0.02)
F; $\hat{\theta}_{3,4}$		6.34 (0.79)	0.53 (0.04)	–		6.39 (0.75)	0.53 (0.04)	–
F; $\hat{\theta}_{2,4}$		6.77 (0.79)	0.55 (0.04)	–		6.93 (0.74)	0.56 (0.03)	–
F; $\hat{\theta}_{1,4;3}$		1.49 (0.58)	0.16 (0.06)	–		1.51 (0.53)	0.16 (0.05)	–
F; $\hat{\theta}_{2,3;4}$		2.81 (0.53)	0.29 (0.05)	–		2.78 (0.51)	0.29 (0.05)	–
F; $\hat{\theta}_{1,2;3,4}$		3.48 (0.63)	0.35 (0.05)	–		3.48 (0.61)	0.35 (0.05)	–
F; $\hat{\theta}_{1,3}$	-125.77	6.38 (0.81)	0.53 (0.04)	–	-125.66	6.51 (0.79)	0.54 (0.04)	–
F; $\hat{\theta}_{3,4}$		6.34 (0.79)	0.53 (0.04)	–		6.36 (0.72)	0.53 (0.04)	–
C; $\hat{\theta}_{2,4}$		3.90 (0.60)	0.66 (0.03)	0.84 (0.02)		4.10 (0.61)	0.67 (0.03)	0.84 (0.02)
F; $\hat{\theta}_{1,4;3}$		1.54 (0.55)	0.17 (0.06)	–		1.57 (0.55)	0.17 (0.06)	–
F; $\hat{\theta}_{2,3;4}$		2.76 (0.55)	0.29 (0.05)	–		2.79 (0.55)	0.29 (0.05)	–
F; $\hat{\theta}_{1,2;3,4}$		3.86 (0.64)	0.38 (0.05)	–		3.86 (0.65)	0.38 (0.05)	–
C; $\hat{\theta}_{1,3}$	-125.83	3.60 (0.58)	0.64 (0.04)	0.82 (0.03)	-125.73	3.75 (0.61)	0.65 (0.04)	0.83 (0.03)
F; $\hat{\theta}_{3,4}$		6.34 (0.79)	0.53 (0.04)	–		6.40 (0.74)	0.53 (0.04)	–
C; $\hat{\theta}_{2,4}$		3.90 (0.59)	0.66 (0.03)	0.84 (0.02)		4.04 (0.59)	0.67 (0.03)	0.84 (0.02)
F; $\hat{\theta}_{1,4;3}$		1.36 (0.57)	0.15 (0.06)	–		1.39 (0.55)	0.15 (0.06)	–
F; $\hat{\theta}_{2,3;4}$		2.71 (0.53)	0.28 (0.05)	–		2.72 (0.51)	0.28 (0.05)	–
F; $\hat{\theta}_{1,2;3,4}$		3.70 (0.64)	0.37 (0.05)	–		3.71 (0.63)	0.37 (0.05)	–

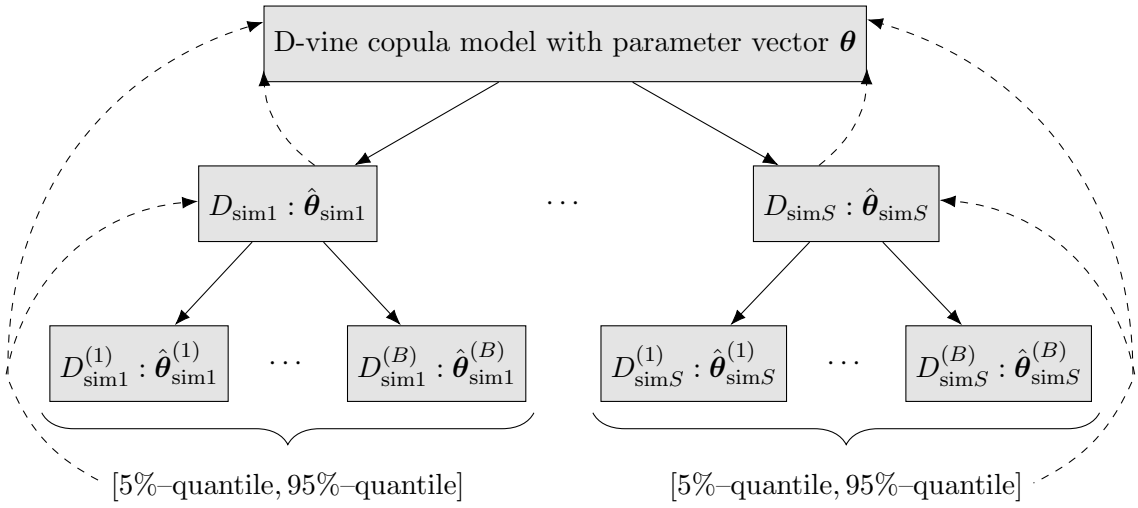


Figure 4.7: Illustration of the simulation setup to validate nonparametric bootstrapping for the mastitis data.

Based on these findings, we are confident to apply a nonparametric bootstrap to the mastitis data. Table 4.10 shows the 90% confidence intervals of the pairwise loglikelihood differences obtained based on 100 nonparametric bootstrap replications for the four D-vine copula models in Table 4.9. While confirming the ranking based on their loglikelihood values, the hypothesis of equal performance cannot be rejected at a confidence level of 10% for any model pair.

Given equal performance of the four D-vine copula models in Table 4.9, it is of particular interest that strong lower tail-dependence and thus a strong association between late event times is detected for Clayton copulas in $\mathcal{T}_1$, while a Frank copula exhibits no tail-behavior. Further, the strength of overall dependence detected for the three udder pairs in $\mathcal{T}_1$ is higher for Clayton copulas as compared to Frank copulas. Figure 4.8 illustrates the normalized contour plots of a bivariate Frank and Clayton copula with Kendall's τ strengths for the front left and rear left udder quarter based on global likelihood estimation, i.e. $\tau_{1,3}^{\text{Frank}} = 0.54$ and $\tau_{1,3}^{\text{Clayton}} = 0.65$. The close match of the contour plots in the upper right quadrant, where most of the data information is located, together with the knowledge about the information loss in the lower left quadrant (recall Figure 4.6), clearly demonstrates the challenges with respect to model selection in the presence of heavy right-censoring.

Table 4.10: 90% confidence intervals for the pairwise loglikelihood differences of the four best D-vine copula fits for the mastitis data based on 100 nonparametric bootstrap samples.

	$\mathcal{T}_1$-sequential estimation	Global estimation
FFF-CFF	[-2.29, 4.96]	[-2.24, 5.10]
FFF-FFC	[-2.76, 5.31]	[-2.64, 5.20]
FFF-CFC	[-3.34, 7.84]	[-3.24, 7.90]
CFF-FFC	[-6.38, 5.24]	[-6.31, 5.21]
CFF-CFC	[-3.47, 4.32]	[-3.54, 4.42]
FFC-CFC	[-2.82, 4.22]	[-2.76, 4.12]

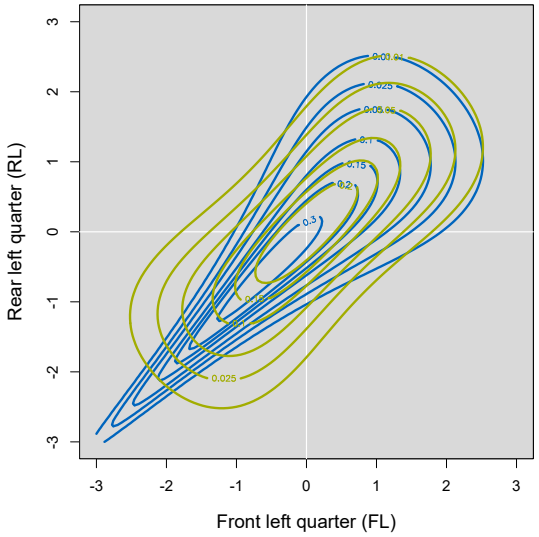


Figure 4.8: Contour plots with standard normalized margins for a Frank (green) and Clayton copula (blue) with Kendall's τ strengths for the front left and rear left udder quarter based on global likelihood estimation, i.e. $\tau_{1,3}^{\text{Frank}} = 0.54$ and $\tau_{1,3}^{\text{Clayton}} = 0.65$.

The results in this section are in line with the findings in Geerdens et al. (2016a), where a Joe-Hu copula that combines a Clayton Laplace transform with bivariate Frank copulas is in the top three of the best models. Both analyses, using an R-vine copula or a Joe-Hu copula, stress the need for flexible copula models for the mastitis data.

We conclude by an important remark on the practical implementation of the optimization procedure. A comparison of the estimation results for both considered estimation methods qualifies the $\mathcal{T}_1$-sequential estimation approach as an important simplification and a valid alternative for the computationally extensive full loglikelihood optimization. Given that heavy censoring goes along with numerical challenges in the full optimization approach, the $\mathcal{T}_1$-sequential approach is the estimation method to apply in practice.

4.4 Modeling recurrent right-censored event time data

In Section 4.2 and Section 4.3, we investigated R-vine copula based likelihood estimation for clustered right-censored event times in a balanced data setting and in the presence of common (univariate) right-censoring. Building upon ideas of this project, extensions to recurrent data were investigated and published in Barthel et al. (2018b). The following section is a slight variation of this manuscript.

4.4.1 Introduction

While in the two previous sections all observation units per cluster simultaneously were under risk of experiencing the event of interest, this is not the case if event times per cluster are recurrent. For example, children with a high risk of developing asthma could be observed. Asthma is a chronic lung disease that inflames and narrows the airways. It causes consecutive episodes of wheezing, chest tightness and shortness of breath, commonly known as asthma attacks.

While the clusters (a child) are independent, the event times (intervals from study entry to event), respectively the gap times (intervals between consecutive events) within a cluster are dependent. Popular survival models that account for within-cluster association are the marginal model (Wei et al., 1989), the (shared) frailty model (Duchateau and Janssen, 2008) and the copula model. However, only the latter allows for direct dependence modeling. The copula model describes the joint survival function of event times or gap times through their survival margins and a copula that fully captures the within-cluster association (Sklar, 1959). Typically, copulas are applied to clusters of equal size (balanced data as in Section 4.2 and Section 4.3), a feature that recurrent event time data often lack. In the above example, one child could have two asthma attacks, while another child experiences three or more asthma attacks. Prenen et al. (2017) and Meyer and Romeo (2015) study copula based inference for unbalanced right-censored clustered data with focus on Archimedean copulas. Unfortunately, the latter only allow for a restrictive dependence structure: all time pairs in a cluster exhibit the same type and strength of association. However, dependence between consecutive events may evolve over time. For example, an asthma attack may weaken the lungs. We therefore advocate D-vine copulas as a flexible alternative to Archimedean copulas (Aas et al., 2009; Czado, 2010). D-vine copulas arise from a line structure, built from freely chosen bivariate (conditional) copulas. As such, they allow for a complex association pattern, while taking the time ordering and the unbalanced nature of recurrent event time data into account. Further, D-vine copulas allow for the two event times, which correspond to leaf variables, to derive an analytical expression of their conditional distribution given all other variables. Thus, predictions for the time until relapse given the individual asthma disease history of a child can be made.

As Meyer and Romeo (2015) but opposed to Prenen et al. (2017), we focus on the analysis of gap times and so an extra challenge arises: not only are the gap times in a cluster associated, but they are also subject to induced dependent right-censoring. Meyer and Romeo (2015) account for this by assuming parametric survival margins in a likelihood based global one-stage estimation strategy. To increase model flexibility we additionally consider nonparametric survival margins

together with global two-stage estimation. To lower the computational burden, we also present alternative sequential estimation techniques.

In summary: for gap time data subject to induced dependent right-censoring we present four novel estimation strategies based on the flexible class of D-vine copulas: one-stage parametric (Section 4.4.4) or two-stage semiparametric (Section 4.4.5) together with global or sequential estimation. Further, we establish guidelines on the best modeling approach for data at hand.

The data setting and notation are given in Section 4.4.2. In Section 4.4.4 to Section 4.4.7, the four estimation strategies are outlined and evaluated under diverse simulation settings. Parametric bootstrap algorithms to obtain standard errors for the parameter estimates of the four estimation strategies are developed in Section 4.4.8. In Section 4.4.9, model selection in the context of unbalanced gap time data subject to induced dependent right-censoring is discussed. The asthma data are analyzed in Section 4.4.10. As an extension to Barthel et al. (2018b) this section further includes a study, where prediction intervals for the times to asthma attacks of each child conditional on the individual disease histories are investigated.

4.4.2 Data setting and notation

Suppose a study includes n independent individuals that are followed-up for a recurrent event. For individual i ($i = 1, \dots, n$) let d_i denote the total number of consecutive events. Thus, individual i corresponds to a cluster of size d_i. Let $T_{i,j}$ be the true j-th event time for cluster i, where $T_{i,j} > 0$ and $T_{i,1} < \ldots < T_{i,d_i}$ ($i = 1, \dots, n$ and $j = 1, \dots, d_i$). Due to a limited study period, the follow-up time of cluster i is subject to right-censoring by C_i. The censoring times are assumed to be noninformative and independent of the event times. The intervals between two subsequent events are referred to as gap times $G_{i,j}$ and are defined by

$$G_{i,1} = T_{i,1} \quad \text{and} \quad G_{i,j} = T_{i,j} - T_{i,j-1} \quad \text{for} \quad i = 1, \dots, n \quad \text{and} \quad j = 2, \dots, d_i.$$

It follows that gap time $G_{i,1}$ is subject to right-censoring by C_i, while subsequent gap times $G_{i,j}$ ($j = 2, \dots, d_i$) are subject to right-censoring by $C_i - T_{i,j-1} = C_i - \sum_{\ell=1}^{j-1} G_{i,\ell}$, which naturally depends on previous gap times. The recurrent nature of the data thus induces dependence between gap times and censoring times: we say that gap times are subject to induced dependent right-censoring. Note that only the last gap time G_{i,d_i} can be right-censored. Hence, for cluster i of size d_i the observed data are given by

$$Y_{i,d_i} = \min\left(G_{i,d_i}, C_i - \sum_{\ell=1}^{d_i-1} G_{i,\ell}\right), \quad \delta_{i,d_i} = \mathbb{1}\left(Y_{i,d_i} = G_{i,d_i}\right) \quad \text{and} \quad Y_{i,j} = G_{i,j}, \quad \delta_{i,j} = 1 \text{ for } j < d_i.$$

Figure 4.9 illustrates the data setting.

Typically, not all individuals experience the same number of events, i.e. the cluster size d_i varies among clusters, resulting in an unbalanced data setting. Let the maximum cluster size be $d = \max\{d_i | i = 1, \dots, n\}$. Denote by n_j ($j = 1, \dots, d$) the number of clusters of size j such that $n = n_1 + n_2 + \ldots + n_{d-1} + n_d$. For ease of notation and for `R`-coding efficiency, we assume that data are ordered by decreasing cluster size. The resulting data format is illustrated in Table 4.11.

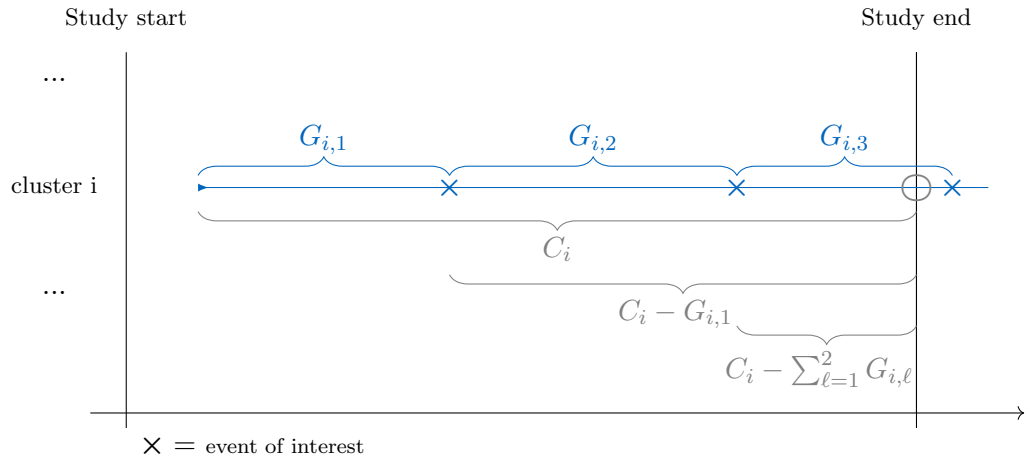


Figure 4.9: Illustration of gap time data subject to induced dependent right-censoring.

4.4.3 D-vine copulas for recurrent data

If interest is in R-vine copulas for dependence modeling of recurrent data, D-vine structures are the natural choice due to the serial variable ordering. For example, Killiches and Czado (2018) model for non-censored unbalanced data repeated measurements using D-vine copulas. In our case of unbalanced gap time data, a natural approach is to choose the joint survival function $S(g_1, \dots, g_d) = \mathbb{P}(G_1 > g_1, \dots, G_d > g_d)$ with survival margins $S_j(g) = \mathbb{P}(G_j > g)$ $(j = 1, \dots, d)$ for the maximum cluster size d and to take the induced d_i-dimensional marginal survival function $S_{1:d_i}(g_1, \dots, g_{d_i})$ for clusters of size $2 \leq d_i < d$. Hence, we consider the vector of gap times $(G_1, \dots, G_d)$. The corresponding D-vine structure with variable order $1{-}2{-}\dots{-}d$ is illustrated in Figure 4.10. In tree $\mathcal{T}_1$, the nodes correspond to the random variables $U_j = S_j(G_j)$ $(j = 1, \dots, d)$, while the edges refer to the bivariate copula density $\mathbb{c}_{k,k+1}(\cdot, \cdot)$ $(k = 1, \dots, d{-}1)$ corresponding to the bivariate distribution of (U_k, U_{k+1}). In tree $\mathcal{T}_\ell$ $(\ell = 2, \dots, d{-}1)$, we define for $k = 1, \dots, d{-}\ell$ the vector $\boldsymbol{u}_{k+1:k+\ell-1} := (u_{k+1}, \dots, u_{k+\ell-1})$ and denote by $\mathbb{c}_{k,k+\ell;k+1:k+\ell-1}(\cdot, \cdot; \boldsymbol{u}_{k+1:k+\ell-1})$ the bivariate conditional copula density linked to the conditional distribution of $(U_k, U_{k+\ell})$ given $\boldsymbol{U}_{k+1:k+\ell-1} = \boldsymbol{u}_{k+1:k+\ell-1}$. As derived in detail in Czado (2010), the copula density $\mathbb{c}_{1:d}$ of $(U_1, \dots, U_d)$ can be expressed as a d-dimensional ordered D-vine copula density as follows:

$$\begin{aligned} &\mathbb{c}_{1:d}(u_1, \dots, u_d) \\ &= \prod_{\ell=1}^{d-1} \prod_{k=1}^{d-\ell} \mathbb{c}_{k,k+\ell;k+1:k+\ell-1}\{\mathbb{C}_{k|k+1:k+\ell-1}(u_k|\boldsymbol{u}_{k+1:k+\ell-1}), \mathbb{C}_{k+\ell|k+1:k+\ell-1}(u_{k+\ell}|\boldsymbol{u}_{k+1:k+\ell-1})\}, \end{aligned} \tag{4.8}$$

where $\mathbb{C}_{k|k+1:k+\ell-1}(\cdot|\boldsymbol{u}_{k+1:k+\ell-1})$ denotes the univariate conditional distribution of U_k given $\boldsymbol{U}_{k+1:k+\ell-1} = \boldsymbol{u}_{k+1:k+\ell-1}$ and $\mathbb{C}_{k+\ell|k+1:k+\ell-1}(\cdot|\boldsymbol{u}_{k+1:k+\ell-1})$ denotes the univariate conditional distribution of $U_{k+\ell}$ given $\boldsymbol{U}_{k+1:k+\ell-1} = \boldsymbol{u}_{k+1:k+\ell-1}$. The copula parameters corresponding to the bivariate copula density $\mathbb{c}_{k,k+\ell;k+1:k+\ell-1}$ are denoted by $\boldsymbol{\theta}_{k,k+\ell;k+1:k+\ell-1}$. The collection of parameters of an ordered d-dimensional D-vine copula is then given by

$$\boldsymbol{\theta}_{1:d} = \{\boldsymbol{\theta}_{k,k+\ell;k+1:k+\ell-1} | k = 1, \dots, d-\ell, \ell = 1, \dots, d-1\}.$$

Table 4.11: Data format considered for the induced dependent right-censored gap time data $(y_{i,1}, y_{i,2}, \ldots, y_{i,d-1}, y_{i,d}, \delta_{i,1}, \delta_{i,2}, \ldots, \delta_{i,d-1}, \delta_{i,d})$, $i = 1, \ldots, n$: ordering by decreasing cluster size.

i	$y_{i,1}$	$y_{i,2}$	$\cdots$	$y_{i,d-1}$	$y_{i,d}$
1	$g_{1,1}$	$g_{1,2}$	$\cdots$	$g_{1,d-1}$	$y_{1,d}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
n_d	$g_{n_d,1}$	$g_{n_d,2}$	$\cdots$	$g_{n_d,d-1}$	$y_{n_d,d}$
$n_d + 1$	$g_{n_d+1,1}$	$g_{n_d+1,2}$	$\cdots$	$y_{n_d+1,d-1}$	
$\vdots$	$\vdots$	$\vdots$		$\vdots$	
$n_d + n_{d-1}$	$g_{n_d+n_{d-1},1}$	$g_{n_d+n_{d-1},2}$	$\cdots$	$y_{n_d+n_{d-1},d-1}$	
$\vdots$	$\vdots$	$\vdots$			
$n_d + \ldots + n_2 + 1$	$y_{n_d+\ldots+n_2+1,1}$				
$\vdots$	$\vdots$				
$n_d + \ldots + n_2 + n_1 = n$	$y_{n,1}$				
i	$\delta_{i,1}$	$\delta_{i,2}$	$\cdots$	$\delta_{i,d-1}$	$\delta_{i,d}$
1	1	1	$\cdots$	1	$\delta_{1,d}$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
n_d	1	1	$\cdots$	1	$\delta_{n_d,d}$
$n_d + 1$	1	1	$\cdots$	$\delta_{n_d+1,d-1}$	
$\vdots$	$\vdots$	$\vdots$		$\vdots$	
$n_d + n_{d-1}$	1	1	$\cdots$	$\delta_{n_d+n_{d-1},d-1}$	
$\vdots$	$\vdots$	$\vdots$			
$n_d + \ldots + n_2 + 1$	$\delta_{n_d+\ldots+n_2+1,1}$				
$\vdots$	$\vdots$				
$n_d + \ldots + n_2 + n_1 = n$	$\delta_{n,1}$				

Thus, in case of only one-parametric pair-copula families the dependence among d variables is described by $d(d-1)/2$ copula parameters. Unless unclear, we do not explicitly include the parameters in the notation of a D-vine copula. As commonly done, we assume in (4.8) that the conditional pair-copulas $\mathbb{c}_{k,k+\ell;k+1:k+\ell-1}$ in trees $\mathcal{T}_\ell$ $(\ell = 2, \ldots, d-1)$ do not depend on the conditioning vector $\boldsymbol{u}_{k+1:k+\ell-1}$. Their arguments $\mathbb{C}_{k|k+1:k+\ell-1}(u_k|\boldsymbol{u}_{k+1:k+\ell-1})$ and $\mathbb{C}_{k+\ell|k+1,k+\ell-1}(u_{k+\ell}|\boldsymbol{u}_{k+1:k+\ell-1})$ indeed do depend on $\boldsymbol{u}_{k+1:k+\ell-1}$. For details on the simplifying assumption, see for example Hobæk Haff et al. (2010) and Stöber et al. (2013).

In the following sections, we develop several procedures to estimate, for gap times subject to induced dependent right-censoring, the parameters of Archimedean and D-vine copulas. We denote $\mathbb{C}_{1:d}$ the copula for the maximum cluster size d and $\mathbb{C}_{1:d_i}$ represents the induced d_i-dimensional marginal copula corresponding to clusters of size $2 \leq d_i < d$. The copula densities are given by $\mathbb{c}_{1:d}$, respectively $\mathbb{c}_{1:d_i}$, with parameter vectors $\boldsymbol{\theta}_{1:d}$, respectively $\boldsymbol{\theta}_{1:d_i}$. Note from Figure 4.10 and (4.8) that lower-dimensional D-vine copula densities $\mathbb{c}_{1:d_i}$ $(d_i < d)$, which correspond to smaller sized clusters, are embedded in the copula density $\mathbb{c}_{1:d}$ and thus explicitly specified through $\mathbb{c}_{1:d}$. This allows easy handling of unbalanced data. Further, recall that for an Archimedean copula $\boldsymbol{\theta}_{1:d}$ is one-dimensional and $\boldsymbol{\theta}_{1:d_i} = \boldsymbol{\theta}_{1:d}$ for all $2 \leq d_i < d$. For a D-vine copula, we have $\boldsymbol{\theta}_{1:d_i} \subset \boldsymbol{\theta}_{1:d}$, where $\boldsymbol{\theta}_{1:d_i}$ contains $d_i(d_i-1)/2$ elements.

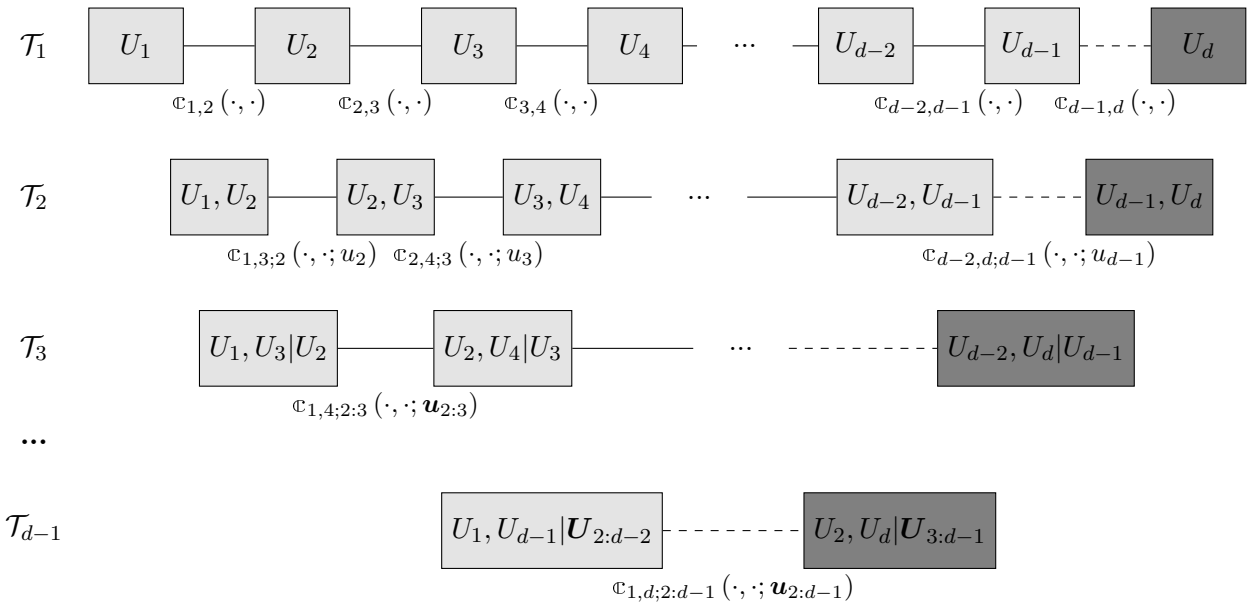


Figure 4.10: D-vine tree structure with order $1-2-\ldots-d$.

Data sampling procedure

Throughout, we support our findings via simulations. To generate unbalanced induced dependent right-censored d-dimensional data, which exhibit a dependence structure specified by a specific copula model $\mathbb{C}_{1:d}$, we use the following procedure. First, we sample data on the copula level from $\mathbb{C}_{1:d}$ using in case of a D-vine copula standard techniques implemented in the `R`-package `VineCopula` (Schepsmeier et al., 2017). Next, we apply – using appropriate assumptions for the survival margins – the inverse probability transform to create gap times $G_{i,j}$ ($i=1,\ldots,n$ and $j=1,\ldots,d$). The corresponding event times $T_{i,j}$ are defined as $T_{i,1}=G_{i,1}$ and $T_{i,j}=\sum_{\ell=1}^{j}G_{i,\ell}$. Based on sampled censoring times, the observed data are obtained as follows: if $T_{i,1}>C_i$ we set $d_i=1$ and retain C_i, if $T_{i,2}>C_i>T_{i,1}$ we set $d_i=2$ and retain $(T_{i,1},C_i)$, etc. In case of the maximum cluster size, i.e. $d_i=d$, we distinguish between $d-1$ events if $T_{i,d}>C_i>T_{i,d-1}$, i.e. we retain $(T_{i,1},\ldots,T_{i,d-1},C_i)$, and d events if $C_i>T_{i,d}$, i.e. we retain $(T_{i,1},\ldots,T_{i,d})$. Finally, the observed gap times for cluster i are given by $Y_{i,j}=T_{i,j}-T_{i,j-1}=G_{i,j}$ for $j=1,\ldots,d_i-1$ and $Y_{i,d_i}=\min(T_{i,d_i}-T_{i,d_i-1},C_i-T_{i,d_i-1})=\min(G_{i,d_i},C_i-\sum_{\ell=1}^{d_i-1}G_{i,\ell})$ together with the right-censoring indicator $\delta_{i,d_i}=\mathbb{1}(Y_{i,d_i}=G_{i,d_i})$. Note that with this procedure the last gap time in a cluster of size $d_i<d$ is always right-censored. Given that many studies have a limited follow-up period, the latter most often holds true in practice, see for example the asthma data.

4.4.4 One-stage parametric copula parameter estimation

In the first modeling strategy, we take a parametric form with parameters $\boldsymbol{\alpha}:=(\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_d)$ for the survival margins S_j ($j=1,\ldots,d$) and parameters $\boldsymbol{\theta}_{1:d}$ for the copula $\mathbb{C}_{1:d}$. As mentioned, gap times are subject to induced dependent censoring. However, given that the latter results from the gap time association, joint estimation of the margins and the dependence structure resolves the issue.

Global likelihood optimization

For cluster i $(i = 1, \ldots, n)$ of size d_i the observed data are defined as

$$(y_{i,1}, \ldots, y_{i,d_i-1}, y_{i,d_i}) = \{g_{i,1}, \ldots, g_{i,d_i-1}, \min\left(g_{i,d_i}, c_i - \sum_{\ell=1}^{d_i-1} g_{i,\ell}\right)\}$$

with censoring indicator $\delta_{i,d_i} = \mathbb{1}\left(y_{i,d_i} = g_{i,d_i}\right)$. The loglikelihood contribution of cluster i is defined following the same arguments as in Section 4.2.3, i.e.

$$\begin{aligned} &\ell_{i,d_i}^{\text{1stage}}(y_{i,1}, \ldots, y_{i,d_i}, \delta_{i,d_i}) \qquad (4.9)\\ &= \delta_{i,d_i} \log\left[\mathbb{c}_{1:d_i}\{S_1\left(y_{i,1}; \boldsymbol{\alpha}_1\right), \ldots, S_{d_i}\left(y_{i,d_i}; \boldsymbol{\alpha}_{d_i}\right); \boldsymbol{\theta}_{1:d_i}\} \cdot f_1\left(y_{i,1}; \boldsymbol{\alpha}_1\right) \cdot \ldots \cdot f_{d_i}\left(y_{i,d_i}; \boldsymbol{\alpha}_{d_i}\right)\right]\\ &\quad + \left(1 - \delta_{i,d_i}\right) \log\left[(-1)^{d_i-1} \frac{\partial^{d_i-1}}{\partial y_{i,1} \cdots \partial y_{i,d_i-1}} \mathbb{C}_{1:d_i}\{S_1\left(y_{i,1}; \boldsymbol{\alpha}_1\right), \ldots, S_{d_i}\left(y_{i,d_i}; \boldsymbol{\alpha}_{d_i}\right); \boldsymbol{\theta}_{1:d_i}\}\right]. \end{aligned}$$

The first term in (4.9) covers the case of y_{i,d_i} being a true gap time, i.e. the last event was observed, the second term in (4.9) corresponds to the case of y_{i,d_i} being a right-censored gap time. For one-stage global parametric estimation, the loglikelihood, which is to be optimized with respect to the marginal parameters $\boldsymbol{\alpha}$ and the copula parameters $\boldsymbol{\theta}_{1:d}$, is then given by

$$\ell^{\text{1stage}}(\boldsymbol{\alpha}, \boldsymbol{\theta}_{1:d}) = \sum_{i=1}^{n} \ell_{i,d_i}^{\text{1stage}}(y_{i,1}, \ldots, y_{i,d_i}, \delta_{i,d_i}). \qquad (4.10)$$

In case of an Archimedean copula all clusters contribute to the estimation of the single parameter $\boldsymbol{\theta}_{1:d}$. For a D-vine, with $\ell = 1, \ldots, d-1, k = \max\left(1, j-\ell+1\right), \ldots, d-\ell$ and $j = 1, \ldots, d$ estimation of the parameters $\boldsymbol{\theta}_{k,k+\ell;k+1:k+\ell-1}$ is based only on clusters i of size $d_i > j$.

Recall that, if $\mathbb{C}_{1:d}$ arises from a d-dimensional ordered D-vine copula, the d_i-dimensional marginal copula densities $\mathbb{c}_{1:d_i}$ $(d_i < d)$ are embedded within $\mathbb{c}_{1:d}$ and are ordered D-vine copula densities themselves. In this case, explicit expressions in terms of pair-copula components are available for the loglikelihood contributions in (4.10). If the last observation is a true gap time, the loglikelihood contribution equals the d_i-dimensional copula density $\mathbb{c}_{1:d_i}$ evaluated at the observed data. According to Joe (1997) the latter can be expressed solely in terms of the pair-copulas in $\mathcal{T}_1$ to $\mathcal{T}_{d_i-1}$. If the last observation is a right-censored gap time, the loglikelihood contribution equals the partial derivative of the copula $\mathbb{C}_{1:d_i}$ with respect to all its arguments except for the last. Given that $\mathbb{C}_{1:d_i}$ arises from a d_i-dimensional ordered D-vine copula, one can show that

$$\frac{\partial^{d_i-1} \mathbb{C}_{1:d_i}(u_{i,1}, \ldots, u_{i,d_i})}{\partial u_{i,1} \cdots \partial u_{i,d_i-1}} = \mathbb{c}_{1:d_i-1}(u_{i,1}, \ldots, u_{i,d_i-1}) \mathbb{C}_{d_i|1:d_i-1}(u_{i,d_i}|\boldsymbol{u}_{i,1:d_i-1}), \qquad (4.11)$$

where $u_{i,j} = S_j\left(y_{i,j}\right)$ $(i = 1, \ldots, n$ and $j = 1, \ldots, d_i)$ is set for ease of notation. The equality in (4.11) is derived in Appendix B.1.2. The first part on the right-hand side $(\mathbb{c}_{1:d_i-1})$ is an $(d_i - 1)$-dimensional ordered D-vine copula density. Thus, according to Joe (1997) an expression only in terms of the pair-copulas in $\mathcal{T}_1$ to $\mathcal{T}_{d_i-2}$ is available. The second part on the right-hand

side ($\mathbb{C}_{d_i|1:d_i-1}$) is a univariate conditional distribution function, which corresponds to the h-function $h_{d_i|1;2:d_i-1}$, i.e. the partial derivative of the pair-copula $\mathbb{C}_{1,d_i;2:d_i-1}$ with respect to the first argument. In a d_i-dimensional ordered D-vine copula, the pair-copula $\mathbb{C}_{1,d_i;2:d_i-1}$ is specified in tree level $\mathcal{T}_{d_i-1}$, and according to Joe (1997) can be recursively evaluated using only pair-copulas in $\mathcal{T}_1$ to $\mathcal{T}_{d_i-2}$. To conclude, for both loglikelihood contributions in (4.10) there exist analytical expressions solely based on the pair-copulas specified by the ordered D-vine copula.

Sequential estimation approach

The global one-stage parametric estimation approach is valid for both Archimedean and D-vine copulas. For data of maximum cluster size d this requires, for D-vine copulas, the joint estimation of $d(d-1)/2$ copula parameters together with the parameters of d survival margins. We aim for a more parsimonious estimation strategy by proceeding sequentially.

We use the fact that for each cluster size $2 \leq d_i < d$ $(i = 1, \ldots, n)$, the copula density $\mathbb{c}_{1:d_i}$ is embedded within the copula density $\mathbb{c}_{1:d}$ and stepwise increase the number of considered gap times from 1 to d. In each step j $(j = 1, \ldots, d)$ estimates obtained from previous steps are fixed such that only the marginal parameters of the j-th gap time and of the pair-copulas incorporating the j-th gap time are to be estimated. In Figure 4.10 estimation proceeds from left to right. Details are given in Algorithm 1 and are additionally illustrated in Figure 4.11. For example, for a model having two-parametric marginal models (like Weibull) the $(d(d-1)/2+2d)$-dimensional optimization is split into d optimizations of dimension $j+1$ $(j = 1, \ldots, d)$.

Algorithm 1 Sequential left-right one-stage parametric estimation.

Input: gap time data $(y_{i,1}, y_{i,2}, \ldots, y_{i,d_i}, \delta_{i,1}, \delta_{i,2}, \ldots, \delta_{i,d_i})$, $i = 1, \ldots, n$, subject to induced dependent right-censoring ordered by decreasing cluster size.

Output: parameter estimates $\hat{\boldsymbol{\alpha}} = (\hat{\boldsymbol{\alpha}}_1, \hat{\boldsymbol{\alpha}}_2, \ldots, \hat{\boldsymbol{\alpha}}_d)$ and $\hat{\boldsymbol{\theta}}_{1:d}$ with $d = \max\{d_i | i = 1, \ldots, n\}$.

1: Set $d = \max\{d_i | i = 1, \ldots, n\}$.
2: Set $N = n_d + \ldots + n_1$.
3: Maximize $\sum_{i=1}^{N} \ell_{i,1}^{\text{1stage}}(y_{i,1}, \delta_{i,1})$ with respect to $\boldsymbol{\alpha}_1$. Denote the maximizer by $\hat{\boldsymbol{\alpha}}_1$.
4: Set $N = n_d + \ldots + n_2$.
5: Fix $\boldsymbol{\alpha}_1$ at $\hat{\boldsymbol{\alpha}}_1$.
6: Maximize $\sum_{i=1}^{N} \ell_{i,2}^{\text{1stage}}(y_{i,1}, y_{i,2}, \delta_{i,2})$ with respect to $\boldsymbol{\alpha}_2$ and $\theta_{1,2}$. Denote the maximizers by $\hat{\boldsymbol{\alpha}}_2$ and $\hat{\theta}_{1,2}$.
7: **for** $j = 3, \ldots, d$ **do**
8: **if** $j < d$ **then** Set $N = n_d + \ldots + n_j$. **end if**
9: **if** $j = d$ **then** Set $N = n_d$. **end if**
10: Fix $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_{j-1}$ at $\hat{\boldsymbol{\alpha}}_1, \ldots, \hat{\boldsymbol{\alpha}}_{j-1}$ and $\boldsymbol{\theta}_{1:j-1}$ at $\hat{\boldsymbol{\theta}}_{1:j-1}$.
11: Maximize $\sum_{i=1}^{N} \ell_{i,j}^{\text{1stage}}(y_{i,1}, \ldots, y_{i,j}, \delta_{i,j})$ with respect to $\boldsymbol{\alpha}_j$ and $\boldsymbol{\theta}_{1:j} \backslash \boldsymbol{\theta}_{1:j-1}$. The estimates obtained in steps 1 to j are $\hat{\boldsymbol{\alpha}}_1, \ldots, \hat{\boldsymbol{\alpha}}_j$, $\hat{\boldsymbol{\theta}}_{1:j}$.
12: **end for**

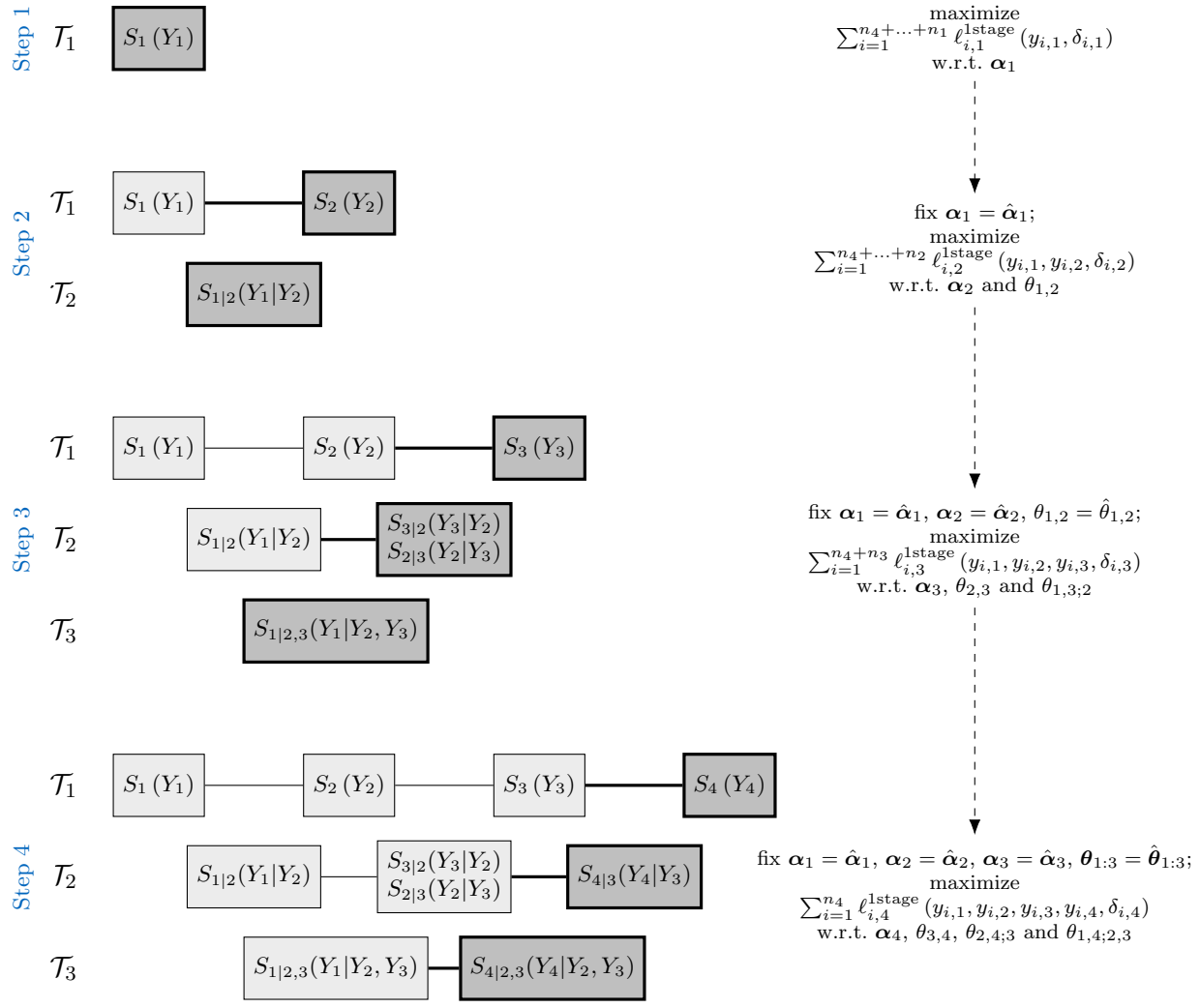


Figure 4.11: Illustration of the sequential left-right one-stage parametric estimation approach assuming four-dimensional data.

Illustrating simulations

To investigate the finite sample performance of the suggested approaches, a wide range of scenarios inspired by the asthma data is considered. We use the procedure outlined in Section 4.4.3 for data sampling. In each scenario, the results are based on 250 data sets. We consider samples of 250, 500 and 1000 clusters, each with a maximum size of 3. The gap times and the censoring times are assumed to follow a Weibull distribution, i.e. $S(g) = \exp(-\lambda g^{\rho})$. The scale ($\lambda$) and shape ($\rho$) parameters are chosen such that the data show about 15% or 30% censoring. A third scenario yields 30% censoring but with censored observations mainly located at late time points (heavy tail - HT). It is assumed that gap 1 differs from gap 2 and gap 3, i.e. the latter are expected to be shorter, reflecting a weakening of the lungs after a first asthma attack. The settings for the gap times and the censoring are summarized in Table 4.12 and further illustrated in Figure 4.12. The dependence between the gap times is modeled using a copula. We look at a

three-dimensional (3d) Archimedean copula, where a single parameter controls the dependence between all gap times. We focus on an intermediate dependence strength as expressed by a Kendall's τ of 0.5 and investigate the scenario of a Clayton copula (upper tail-dependent) and a Gumbel copula (lower tail-dependent). For a three-dimensional (3d) D-vine copula, we take $\tau_{1,2} = \tau_{2,3} = 0.5$ and $\tau_{1,3;2} = 0.25$. We consider scenarios, where both pair-copulas in tree $\mathcal{T}_1$ are Clayton or Gumbel. The pair-copula in tree $\mathcal{T}_2$ is assumed to be Frank. Table 4.13 summarizes all underlying copula models using C for Clayton, G for Gumbel and F for Frank.

Table 4.12: Simulation settings for the marginal survival functions of the three gap times and for the survival function of the censoring times leading to 15%, 30% or 30% HT (heavy tail) censoring.

Weibull parameters	Gap time 1	Gap time 2 - 3	Censoring		
			15%	30%	30% HT
scale λ	0.50	1.00	0.10	0.25	0.10
shape ρ	1.50	1.50	1.50	1.50	3.00

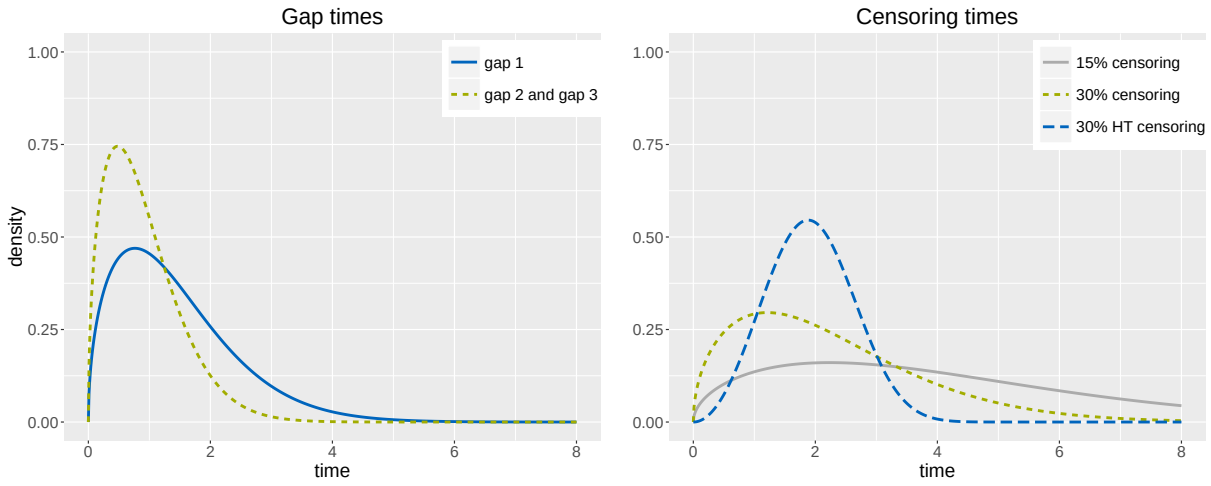


Figure 4.12: Density functions for the Weibull specifications in Table 4.12.

Table 4.13: Simulation settings for Archimedean copulas and D-vine copulas, where $\mathbb{c}$ denotes the (pair-) copula family, τ denotes the Kendall's τ value and θ denotes the parameter corresponding to $\mathbb{c}$ and τ.

	3d Archimedean copula	D-vine copula		
	$\mathbb{c}; \tau; \theta$	$\mathbb{c}_{1,2}; \tau_{1,2}; \theta_{1,2}$	$\mathbb{c}_{2,3}; \tau_{2,3}; \theta_{2,3}$	$\mathbb{c}_{1,3;2}; \tau_{1,3;2}; \theta_{1,3;2}$
Setting 1	C; 0.50; 2.00	C; 0.50; 2.00	C; 0.50; 2.00	F; 0.25; 2.37
Setting 2	G; 0.50; 2.00	G; 0.50; 2.00	G; 0.50; 2.00	F; 0.25; 2.37

The results are obtained under a correct specification of the marginal and copula format and are illustrated in terms of Kendall's τ in Figure 4.13. A detailed summary including copula and marginal parameter estimates is given in Table B.14 to Table B.17 in Appendix B.4. On average and taking the standard deviation into account, estimation is on target. It improves with increasing sample size, but deteriorates with increasing censoring rate and – under fixed censoring rate (30% and 30% HT) – if censored observations are mainly located at late time points. Based on empirical mean and empirical standard deviation results for Clayton based copulas (two top panels of Figure 4.13) are somewhat less accurate than those for Gumbel based copulas (two bottom panels of Figure 4.13). This is due to the lower tail-property of a Clayton copula, which makes the latter more sensitive to right-censoring when modeling a survival function (see Figure 4.6 in Section 4.3.2). For the D-vine copulas, global and sequential optimization perform alike indicating that the latter is a valid alternative for the computationally more demanding global approach.

4.4.5 Two-stage semiparametric copula parameter estimation

In spite of the good performance of the one-stage parametric approaches, model flexibility is increased when using two-stage semiparametric estimation. In stage one, the survival margins (S_j) are estimated nonparametrically $(\hat{S}_j)$. In stage two, the pseudo data $\hat{u}_{i,j} = \hat{S}_j(y_{i,j})$ are used to estimate the copula parameters via likelihood optimization. This approach goes back to Shih and Louis (1995), who consider bivariate survival data subject to independent right-censoring. Extensions to multivariate survival data are for example in Geerdens et al. (2016a) and Barthel et al. (2018c) and were discussed in Section 4.2 and Section 4.3 of this thesis. There, Kaplan-Meier or Nelson-Aalen estimators are applied for nonparametric marginal modeling (see Section 4.1.3). In the presence of induced dependent right-censoring these nonparametric estimators are no longer consistent (Cook and Lawless, 2007) and alternatives are needed.

Univariate marginal modeling

For induced dependent right-censoring de Uña-Álvarez and Meira-Machado (2008) propose a consistent nonparametric estimator for the survival margins. As an estimate for the joint distribution F of the gap time vector $(G_1, \ldots, G_d)$ they define

$$\hat{F}(g_1, \ldots, g_d) = \sum_{i=1}^{n} W_i^{\mathrm{KM}} \mathbb{1}(y_{i,1} \leq g_1, \ldots, y_{i,d} \leq g_d),$$

where W_i^{KM} is the jump of the Kaplan-Meier estimate obtained from observations $(\tilde{y}_i, \delta_{i,d_i})$ with $\tilde{y}_i$ the total follow-up time for cluster i, i.e. $\tilde{y}_i = \min(t_{i,d_i}, c_i) = \sum_{j=1}^{d_i} y_{i,j}$ $(i = 1, \ldots, n)$. An estimate for the j-th marginal survival function is then given by

$$\hat{S}_j^{\mathrm{KM}}(g) = 1 - \sum_{i=1}^{n} W_i^{\mathrm{KM}} \mathbb{1}(y_{i,j} \leq g), \quad j = 1, \ldots, d.$$

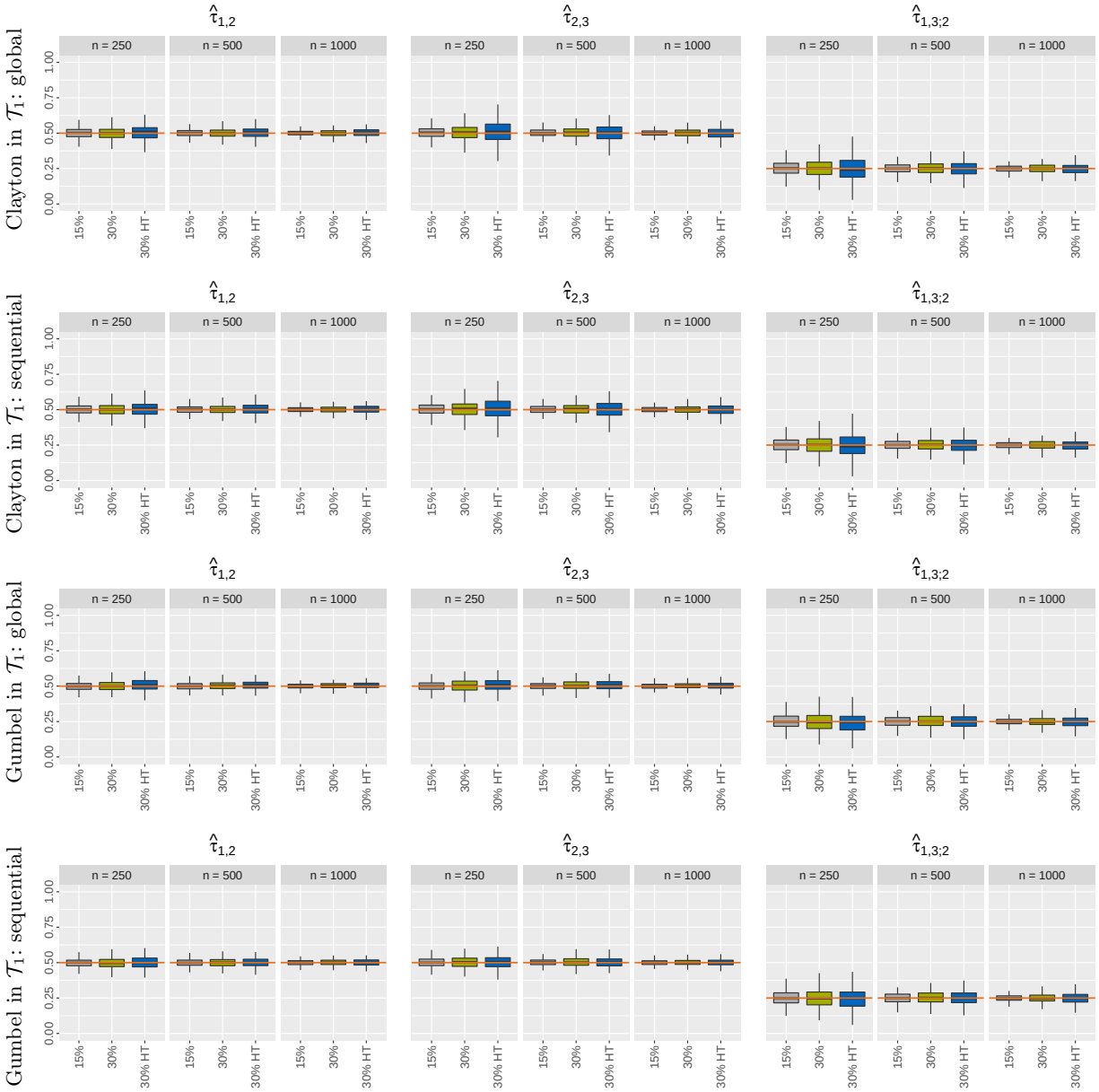


Figure 4.13: Boxplots of Kendall's τ estimates as obtained via one-stage parametric estimation (lines $1 + 3$ = global approach, line $2 + 4$ = sequential approach) for a D-vine copula with Clayton parts (two top panels) and a D-vine copula with Gumbel parts (two bottom panels) based on 250 replications. The true Kendall's τ values are indicated through a horizontal line.

Note that the Kaplan-Meier estimator drops to zero, whenever the largest total time is a true event. After applying the probability integral transform this results in a zero value for the corresponding copula data value. To avoid numerical difficulties in likelihood maximization, we modify the de Uña-Álvarez and Meira-Machado (2008) estimator using the Nelson-Aalen estimator for the cumulative hazard function $\Lambda(\tilde{y})$ of the total times. The corresponding Nelson-Aalen estimate based survival jumps W_i^{NA} $(i = 1, \ldots, n)$ are then obtained via the transformation $\exp(-\Lambda(\tilde{y}))$. The pseudo copula data then are:

$$\hat{u}_{i,j} = \hat{S}_j^{\text{NA}}(y_{i,j}) = 1 \quad - \sum_{\ell \in \{i | 1 \leq i \leq n, d_i \geq j\}} W_\ell^{\text{NA}} \mathbb{1}(y_{\ell,j} \leq y_{i,j}), \quad i = 1, \ldots, n \text{ and } j = 1, \ldots, d_i. \tag{4.12}$$

Global likelihood inference

Based on the pseudo data in (4.12), the copula parameters $\boldsymbol{\theta}_{1:d}$ are estimated via likelihood maximization. As before, right-censoring needs to be taken into account. For cluster i of size d_i $(i = 1, \ldots, n)$ the loglikelihood contribution equals

$$\begin{aligned} &\ell_{i,d_i}^{\text{2stage}}(\hat{u}_{i,1}, \ldots, \hat{u}_{i,d_i}, \delta_{i,d_i}) \qquad (4.13) \\ &= \delta_{i,d_i} \log\left[\mathbb{c}_{1:d_i}\{\hat{u}_{i,1}, \ldots, \hat{u}_{i,d_i}; \boldsymbol{\theta}_{1:d_i}\}\right] + (1 - \delta_{i,d_i}) \log\left[\frac{\partial^{d_i - 1}}{\partial \hat{u}_{i,1} \cdots \partial \hat{u}_{i,d_i - 1}} \mathbb{C}_{1:d_i}\{\hat{u}_{i,1}, \ldots, \hat{u}_{i,d_i}; \boldsymbol{\theta}_{1:d_i}\}\right]. \end{aligned}$$

Hence, the loglikelihood function, which needs to be optimized with respect to $\boldsymbol{\theta}_{1:d}$, is

$$\ell^{\text{2stage}}(\boldsymbol{\theta}_{1:d}) = \sum_{i=1}^{n} \ell_{i,d_i}^{\text{2stage}}(\hat{u}_{i,1}, \ldots, \hat{u}_{i,d_i}, \delta_{i,d_i}). \tag{4.14}$$

Sequential estimation approach

For D-vine copulas, also in case of two-stage estimation, a sequential procedure for likelihood maximization is feasible. It relies on the recursive nature of the arguments of the pair-copulas (Section 4.1.2). In Figure 4.10, estimation proceeds from top to bottom. First, all pair-copula parameters in $\mathcal{T}_1$ are estimated separately. Based on the fitted pair-copulas, the arguments needed in $\mathcal{T}_2$ are calculated by application of the corresponding h-functions. Using the obtained pseudo data all pair-copula parameters in $\mathcal{T}_2$ can be estimated separately, etc. The procedure has been developed for complete data (Aas et al., 2009; Dißmann et al., 2013). In case of right-censoring, an extra challenge arises: from tree $\mathcal{T}_2$ on estimation is no longer based on the observed copula data themselves, but on pseudo data, namely univariate conditional distribution functions, which are evaluated at the observed copula data. For these pseudo observations censoring indicators need to be defined. Recall that for recurrent event time data only the last gap time can be right-censored. Given the construction of an ordered D-vine, the value on the copula scale $\hat{u}_{i,d_i}$ associated with the last gap time of cluster i $(i = 1, \ldots, n)$ corresponds to a leaf node in the d_i-dimensional subvine and thus, only occurs as conditioned variable in the univariate conditional functions. Further, the latter are monotonously increasing in their conditioned argument. Hence, the pseudo observations inherit the censoring status of their observed conditioned

variable. By doing so, the d-dimensional estimation problem is split into $d(d-1)/2$ bivariate ones and the estimation of a high-dimensional D-vine copula becomes tractable and computationally easier. For each pair-copula $\mathbb{c}_{k,k+\ell;k+1:k+\ell-1}\left(\cdot,\cdot;\theta_{k,k+\ell;k+1:k+\ell-1}\right)$ ($\ell = 1,\dots,d-1$ and $k = 1,\dots,k-\ell$) the loglikelihood contribution of cluster i with $d_i \geq k+\ell$ is given by

$$\begin{aligned}
&\ell_{i,k,k+\ell}^{\text{2stage,seq}}\left(\hat{u}_{i,k|k+1:k+\ell-1}, \hat{u}_{i,k+\ell|k+1:k+\ell-1}, \delta_{i,k+\ell}\right) \\
&\quad = \delta_{i,k+\ell} \log\left[\mathbb{c}_{k,k+\ell;k+1:k+\ell-1}\{\hat{u}_{i,k|k+1:k+\ell-1}, \hat{u}_{i,k+\ell|k+1:k+\ell-1}; \theta_{k,k+\ell;k+1:k+\ell-1}\}\right] \\
&\quad + (1-\delta_{i,k+\ell}) \log\left[\frac{\partial}{\partial u}\mathbb{C}_{k,k+\ell;k+1:k+\ell-1}\{u, \hat{u}_{i,k+\ell|k+1:k+\ell-1}; \theta_{k,k+\ell;k+1:k+\ell-1}\}\Big|_{u=\hat{u}_{i,k|k+1:k+\ell-1}}\right],
\end{aligned}$$

where $\hat{u}_{i,k|k+1:k+\ell-1}$ and $\hat{u}_{i,k+\ell|k+1:k+\ell-1}$ are defined as the (pseudo) observations corresponding to $\mathbb{c}_{k,k+\ell;k+1:k+\ell-1}\left(\cdot,\cdot;\theta_{k,k+\ell;k+1:k+\ell-1}\right)$. The corresponding bivariate loglikelihood, which is to be maximized with respect to $\theta_{k,k+\ell;k+1:k+\ell-1}$ is then given by

$$\ell_{k,k+\ell}^{\text{2stage,seq}} = \sum_{i=1}^{N} \ell_{i,k,k+\ell}^{\text{2stage,seq}}\left(\hat{u}_{i,k|k+1:k+\ell-1}, \hat{u}_{i,k+\ell|k+1:k+\ell-1}, \delta_{i,k+\ell}\right),$$

where N is the number of clusters i with $d_i \geq k+\ell$. Details are given in Algorithm 2 and illustrated in Figure 4.14. For complete and balanced data, see for example Hobæk-Haff et al. (2013) and Stöber and Schepsmeier (2013) for asymptotic properties of this approach.

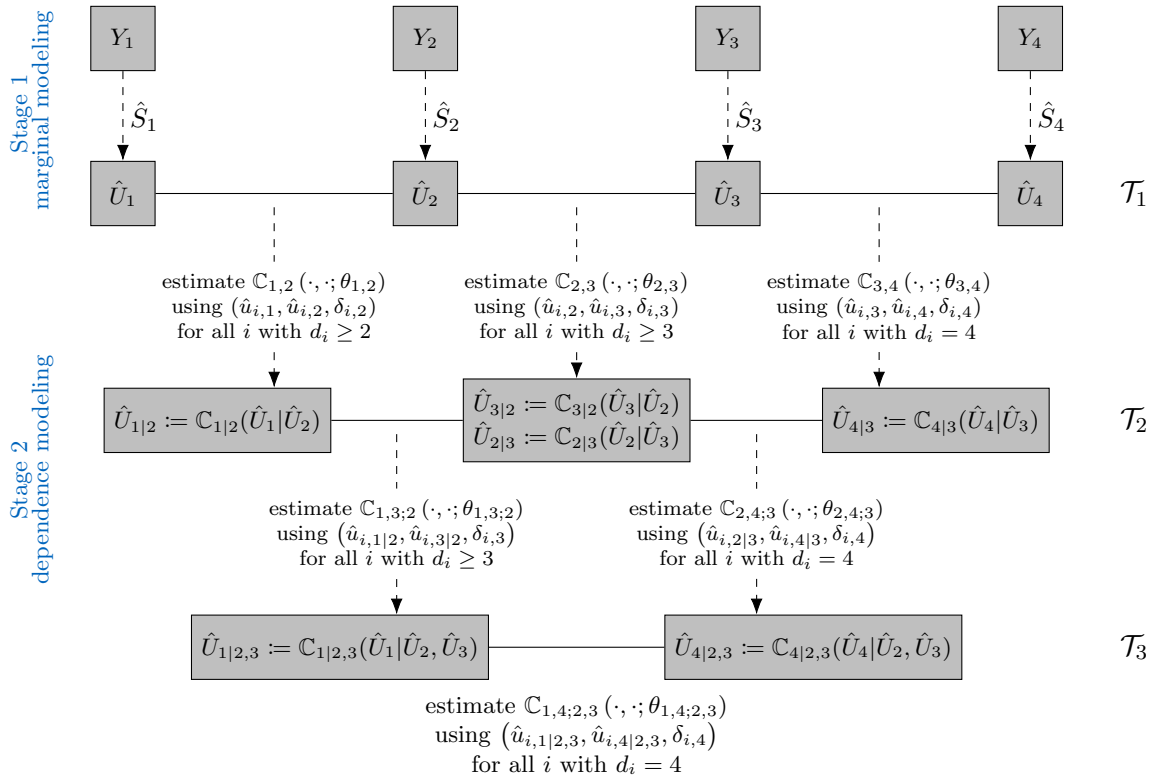


Figure 4.14: Illustration of the sequential top-down two-stage semiparametric estimation approach assuming four-dimensional data.

Algorithm 2 Sequential top-down two-stage semiparametric estimation.

Input: gap time data $(y_{i,1}, y_{i,2}, \ldots, y_{i,d_i}, \delta_{i,1}, \delta_{i,2}, \ldots, \delta_{i,d_i})$, $i = 1, \ldots, n$, subject to induced dependent right-censoring ordered by decreasing cluster size.

Output: parameter estimates $\hat{\boldsymbol{\theta}}_{1:d}$ with $d = \max\{d_i | i = 1, \ldots, n\}$.

Set $d = \max\{d_i | i = 1, \ldots, n\}$.
for $j = 1, \ldots, d$ **do**
  if $j < d$ **then** Set $N = n_d + \ldots + n_j$. **end if**
  if $j = d$ **then** Set $N = n_d$. **end if**
  With $(y_{i,j}, \delta_{i,j})$, $i = 1, \ldots, N$, estimate $\hat{S}_j$ nonparametrically (Section 4.4.5).
  Obtain pseudo copula data $(\hat{u}_{i,j}, \delta_{i,j})$, $i = 1, \ldots, N$, by $\hat{u}_{i,j} = \hat{S}_j(y_{i,j})$.
end for
for $k = 1, \ldots, d-1$ **do**
  if $k < d-1$ **then** Set $N = n_d + \ldots + n_{k+1}$. **end if**
  if $k = d-1$ **then** Set $N = n_d$. **end if**
  Select a copula family for $\mathfrak{c}_{k,k+1}$ and with $(\hat{u}_{i,k}, \hat{u}_{i,k+1}, \delta_{i,k+1})$, $i = 1, \ldots, N$, maximize $\sum_{i=1}^{N} \ell_{i,k,k+1}^{\text{2stage,seq}}(\hat{u}_{i,k}, \hat{u}_{i,k+1}, \delta_{i,k+1})$ with respect to $\theta_{k,k+1}$.
  Using the fitted copula $\mathbb{C}_{k,k+1}(\cdot, \cdot; \hat{\theta}_{k,k+1})$ apply the corresponding conditional cumulative distribution function (CDF) h-functions to calculate $\mathbb{C}_{k|k+1}(\hat{u}_{i,k}|\hat{u}_{i,k+1})$ and $\mathbb{C}_{k+1|k}(\hat{u}_{i,k+1}|\hat{u}_{i,k})$, $i = 1, \ldots, N$.
end for
for $\ell = 2, \ldots, d-1$ **do**
  for $k = 1, \ldots, d-\ell$ **do**
    if $k < d-\ell$ **then** Set $N = n_d + \ldots + n_{k+\ell}$. **end if**
    if $k = d-\ell$ **then** Set $N = n_d$. **end if**
    For $i = 1, \ldots, N$, set $u_i = \mathbb{C}_{k|k+1:k+\ell-1}(\hat{u}_{i,k}|\hat{\boldsymbol{u}}_{i,k+1:k+\ell-1})$
    and $v_i = \mathbb{C}_{k+\ell|k+1:k+\ell-1}(\hat{u}_{i,k+\ell}|\hat{\boldsymbol{u}}_{i,k+1:k+\ell-1})$.
    Set censoring indicator δ_i corresponding to v_i to
    $$\delta_i = \mathbb{1}(d_i > k+\ell) + \mathbb{1}(d_i = k+\ell)\delta_{i,k+\ell}.$$
    Select a copula family for $\mathfrak{c}_{k,k+\ell;k+1:k+\ell-1}$ and with (u_i, v_i, δ_i), $i = 1, \ldots, N$, maximize $\sum_{i=1}^{N} \ell_{i,k,k+\ell}^{\text{2stage,seq}}(u_i, v_i, \delta_i)$ with respect to $\theta_{k,k+\ell;k+1:k+\ell-1}$.
    Using the fitted copula $\mathbb{C}_{k,k+\ell;k+1:k+\ell-1}(\cdot, \cdot; \hat{\theta}_{k,k+\ell;k+1:k+\ell-1})$ apply the corresponding conditional CDF h-functions to calculate
    $\mathbb{C}_{k|k+1:k+\ell}(\hat{u}_{i,k}|\hat{\boldsymbol{u}}_{i,k+1:k+\ell}) = h_{k|k+\ell;k+1:k+\ell-1}(u_i|v_i)$ and
    $\mathbb{C}_{k+\ell|k:k+\ell-1}(\hat{u}_{i,k+\ell}|\hat{\boldsymbol{u}}_{i,k:k+\ell-1}) = h_{k+\ell|k;k+1:k+\ell-1}(v_i|u_i)$, $i = 1, \ldots, N$.
  end for
end for

Illustrating simulations

To investigate the finite sample performance of the suggested two-stage semiparametric approaches, the same simulation settings as for one-stage parametric estimation are used (see Table 4.12 and Table 4.13). The results for Kendall's τ are visualized in Figure 4.15. Detailed results including those for copula parameters are in Table B.18 and Table B.19 in Appendix B.5. Results are calculated under the assumption of a correctly specified copula model. Compared to one-stage parametric estimation, some additional uncertainty is induced by nonparametric marginal estimation. For 15% and 30% censoring, estimation is (on average) accurate. However, for 30% censoring with a heavy tail, estimation is off, i.e. the empirical mean estimates are too high and the empirical standard deviations are larger. Increasing the sample size slightly improves estimation. Clearly, in two-stage estimation not only the amount of censoring but also the censoring position plays a role. In case of many large censored total times, the Nelson-Aalen estimate for the survival function of the total times (usually) levels off away from zero. As such, the estimated survival margins do not drop sufficiently low to zero, which in turn affects the copula data and hence distorts estimation. This issue is not present for one-stage parametric estimation as discussed in Section 4.4.4. Consequently, we recommend to use the latter whenever the tail of the Nelson-Aalen estimate for the survival function of the total times is heavily affected by censoring (leveling off away from zero). The censoring effect is more manifest for a Clayton copula and a D-vine copula with Clayton parts (two top panels of Figure 4.15) as compared to a Gumbel copula and a D-vine copula with Gumbel parts (two bottom panels of Figure 4.15), which again is due to the lower tail-property of Clayton copulas. The simulation results also indicate that, for D-vine copulas, the sequential strategy is a good alternative for the computationally more challenging global approach, when no heavy tail censoring is present.

4.4.6 Guidelines for real life data

Based on the findings of the illustrating simulations with all four estimation strategies Figure 4.16 serves as a guideline to decide for the best estimation approach given real data. It also gives an overview of the four estimation techniques proposed in Section 4.4.4 and Section 4.4.5.

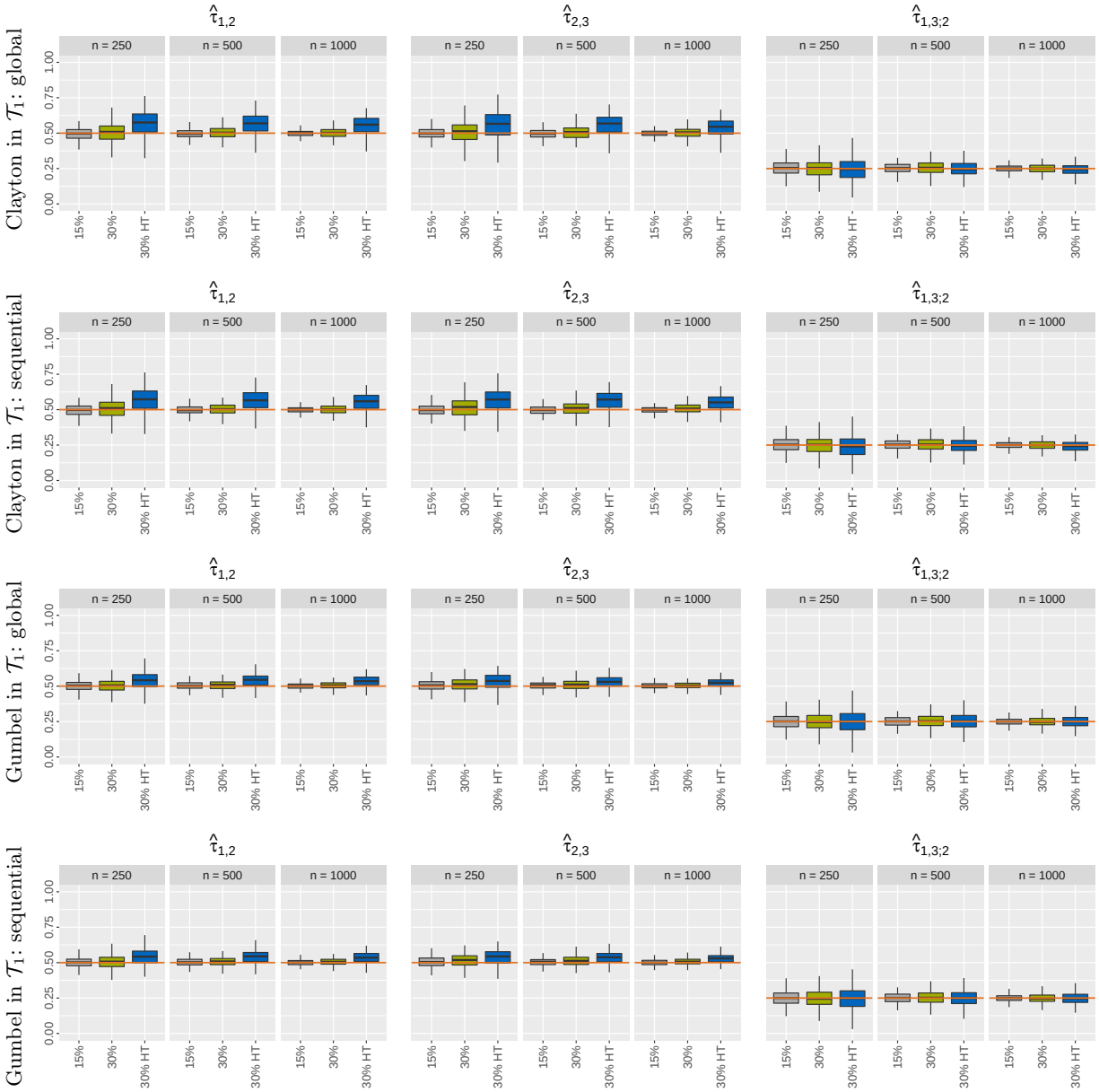


Figure 4.15: Boxplots of Kendall's τ estimates as obtained via two-stage semiparametric estimation (lines 1 + 3 = global approach, line 2 + 4 = sequential approach) for a D-vine copula with Clayton parts (two top panels) and a D-vine copula with Gumbel parts (two bottom panels) based on 250 replications. The true Kendall's τ values are indicated through a horizontal line.

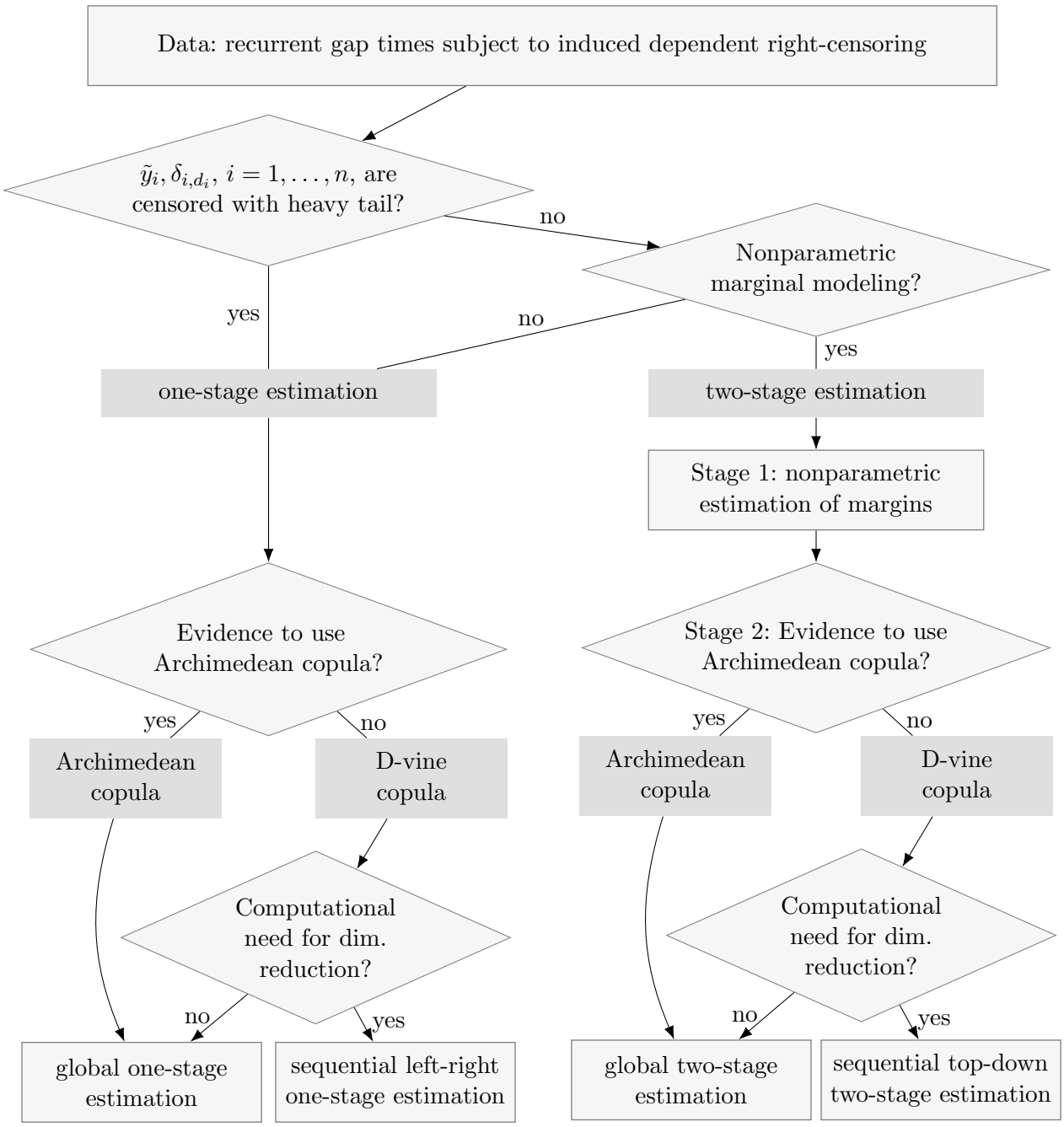


Figure 4.16: Overview and guidelines for usage of the different estimation strategies.

4.4.7 Simulation study

To further explore the finite sample performance of the suggested estimation approaches and to emphasize the flexibility of D-vine copulas over Archimedean copulas with regard to dependence modeling, we additionally investigate settings in which the association varies over time, either in strength or in type. In all scenarios, the results are based on 250 data sets. We consider samples of 250, 500 or 1000 clusters, each with a maximum size of 4, where gap time 4 follows the same distribution as gap times 2 and 3. The censoring times stem from a Weibull survival function with shape and scale parameters chosen such that 15% or 30% of the data are censored. We also consider 30% censoring with large event times being more prone to censoring (HT). Table 4.14 summarizes these settings. The dependence between the four gap times is modeled via a D-vine copula. In trees $\mathcal{T}_2$ and $\mathcal{T}_3$, we consider Frank copulas with $\tau_{1,3;2} = \tau_{2,4;3} = 0.25$ and $\tau_{1,4;2,3} = 0.167$. In $\mathcal{T}_1$, we increase the complexity. First, we fix the dependence strength, but allow the type of association to vary over time: $\mathbb{C}_{1,2}$ is Clayton, $\mathbb{C}_{2,3}$ is Frank, $\mathbb{C}_{3,4}$ is Gumbel. This reflects a slow change from lower to upper tail-dependence. Second, we fix the association type to be Clayton, but allow the strength to increase: $\tau_{1,2} = 0.3$, $\tau_{2,3} = 0.5$, $\tau_{3,4} = 0.7$. See Table 4.15 for a summary using C for Clayton, G for Gumbel and F for Frank.

Table 4.14: Simulation settings for the marginal survival functions of the four gap times and for the survival function of the censoring times leading to 15%, 30% or 30% HT (heavy tail) censoring.

Weibull parameters	Gap time 1	Gap time 2 - 4	Censoring		
			15%	30%	30% HT
scale λ	0.500	1.000	0.085	0.250	0.085
shape ρ	1.500	1.500	1.500	1.500	3.000

Table 4.15: Simulation settings for D-vine copulas.

	D-vine copula (pair-copula families; Kendall's τ; parameter)		
	$\mathbb{C}_{1,2}; \tau_{1,2}; \theta_{1,2}$	$\mathbb{C}_{2,3}; \tau_{2,3}; \theta_{2,3}$	$\mathbb{C}_{3,4}; \tau_{3,4}; \theta_{3,4}$
Setting 1	C; 0.500; 2.00	F; 0.500; 5.76	G; 0.500; 2.00
Setting 2	C; 0.300; 0.86	C; 0.500; 2.00	C; 0.700; 4.67
	$\mathbb{C}_{1,3;2}; \tau_{1,3;2}; \theta_{1,3;2}$	$\mathbb{C}_{2,4;3}; \tau_{2,4;3}; \theta_{2,4;3}$	$\mathbb{C}_{1,4;2,3}; \tau_{1,4;2,3}; \theta_{1,4;2,3}$
Setting 1	F; 0.250; 2.37	F; 0.250; 2.37	F; 0.167; 1.53
Setting 2	F; 0.250; 2.37	F; 0.250; 2.37	F; 0.167; 1.53

The results for Kendall's τ in case of global and sequential one-stage parametric and two-stage semiparametric estimation are illustrated in Figure 4.17 for Setting 1 and in Figure 4.18 for Setting 2. Detailed results including those for the copula and marginal parameters are given in Table B.20 to Table B.24 in Appendix B.6. As before it holds that, under a correct copula format, the one-stage parametric approaches perform well in all censoring scenarios, while the two-stage semiparametric approaches are highly sensitive to the underlying censoring scheme (see

lines 3 and 4 in Figure 4.17 and Figure 4.18). Clearly, except for the two-stage semiparametric approach in case of heavy tail censoring the four proposed estimation strategies allow for accurate estimation of a dependence pattern more complex than that of an Archimedean copula, including varying type and strength of association. Information on runtime for all four proposed estimation strategies is given in Table B.25 in Appendix B.6.

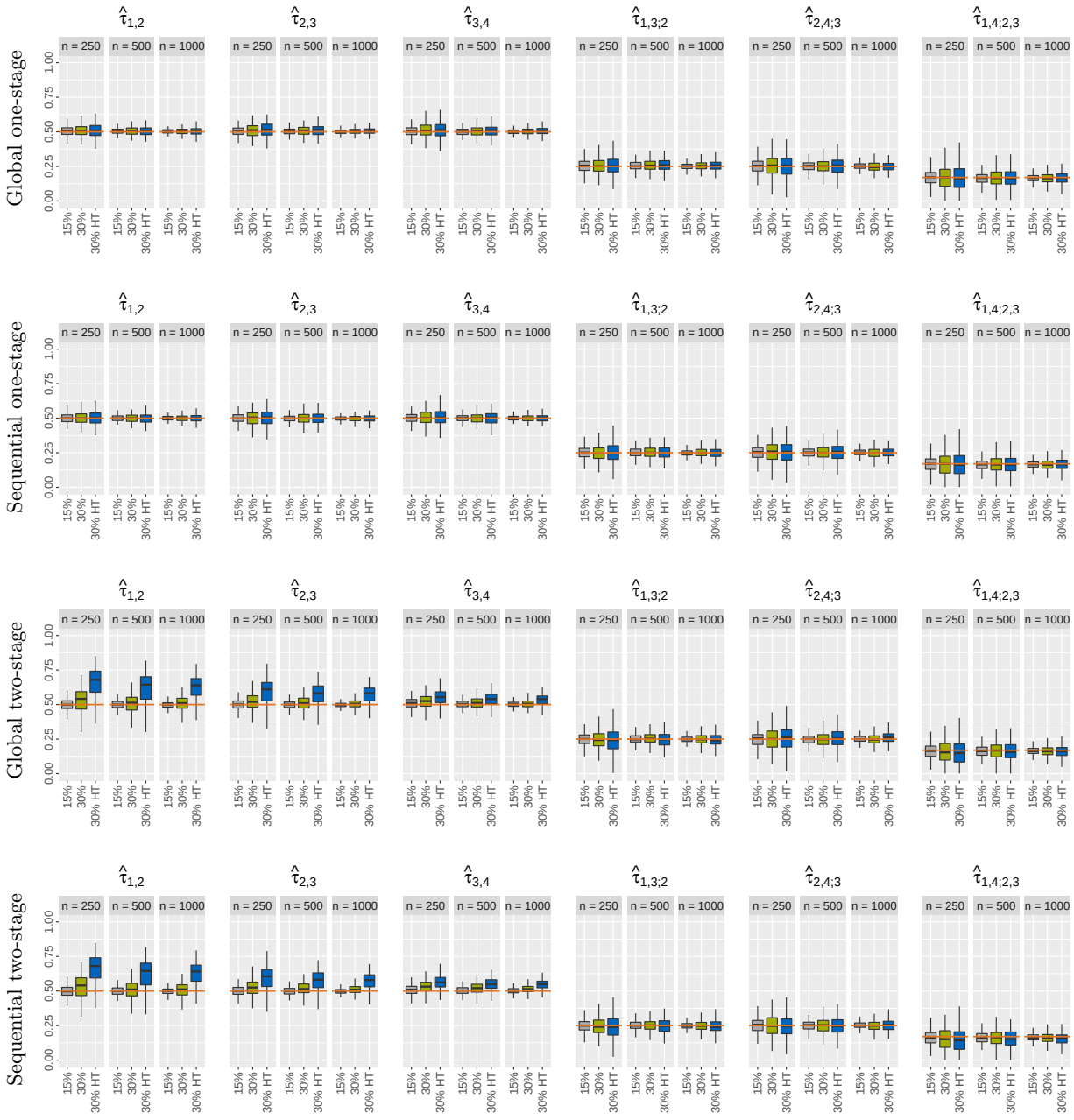


Figure 4.17: Boxplots of Kendall's τ estimates as obtained via one-stage parametric estimation (line 1 = global approach, line 2 = sequential approach) as well as two-stage semiparametric estimation (line 3 = global approach, line 4 = sequential approach) in **Setting 1** (Table 4.15), i.e. for a D-vine copula with changing tail-behavior (lower tail-dependence to upper tail-dependence) based on 250 replications. The true Kendall's τ values are indicated through a horizontal line.

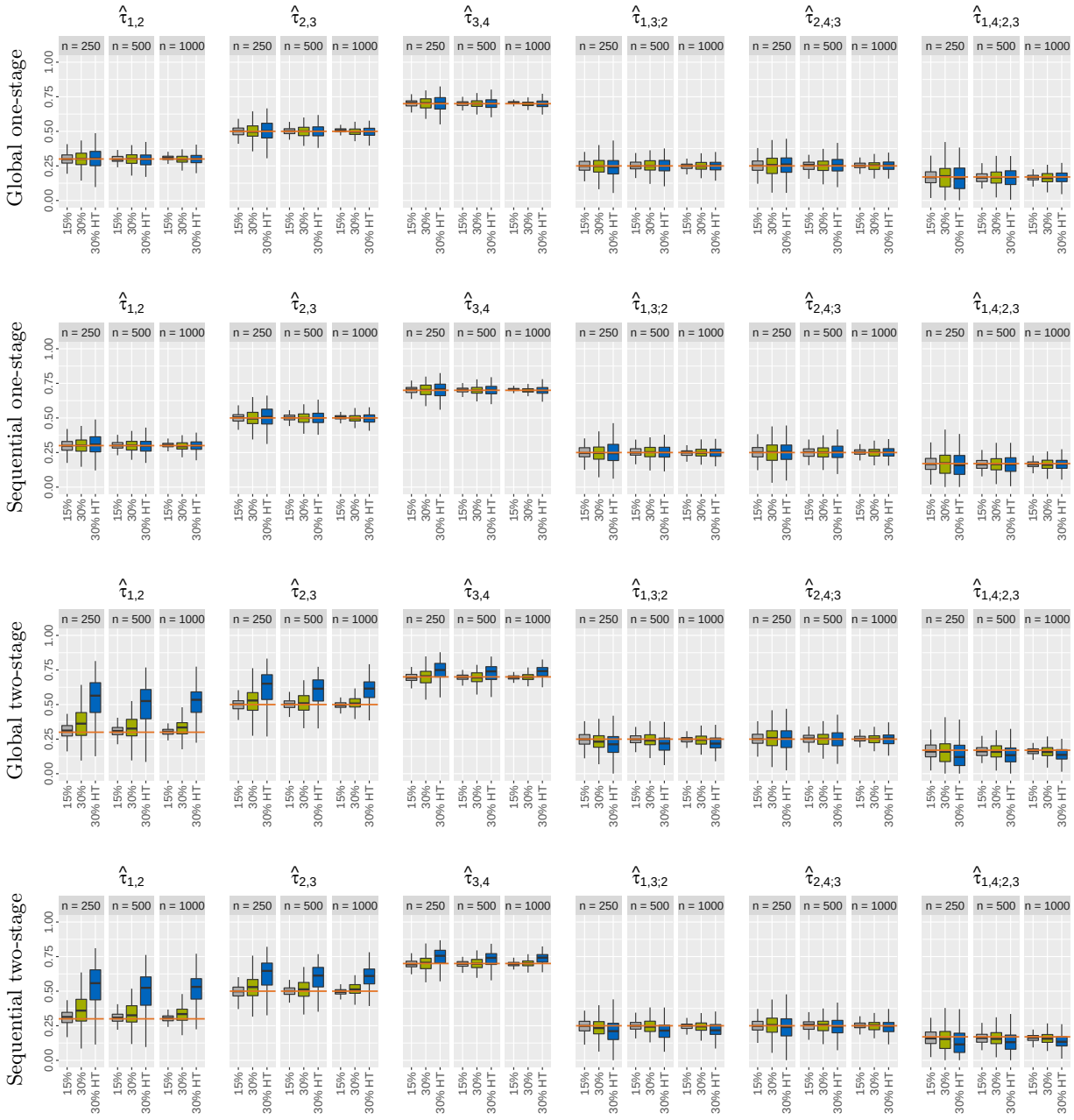


Figure 4.18: Boxplots of Kendall's τ estimates as obtained via one-stage parametric estimation (line 1 = global approach, line 2 = sequential approach) as well as two-stage semi-parametric estimation (line 3 = global approach, line 4 = sequential approach) in **Setting 2** (Table 4.15), i.e. for a D-vine copula with increasing dependence strength (Clayton copula with Kendall's τ values 0.3, 0.5 and 0.7) based on 250 replications. The true Kendall's τ values are indicated through a horizontal line.

4.4.8 Estimation of standard errors

When analyzing real data, standard errors for the copula parameters obtained through application of the four estimation strategies proposed in Section 4.4.4 and Section 4.4.5 need to be estimated. For this purpose, the algorithm for parametric bootstrapping outlined in Section 4.3.1 can be mimicked accounting for the setting of unbalanced, induced dependent right-censored gap time data. Note that we need to distinguish whether the copula parameter estimates were obtained through one-stage parametric or two-stage semiparametric estimation:

Step 1 (one-stage parametric): Under a prespecified parametric format for the marginal survival functions (for example Weibull), follow the global or sequential one-stage parametric estimation approach to fit the D-vine copula model of interest to the data $(y_{i,1}, \ldots, y_{i,d_i}, \delta_{i,1}, \ldots, \delta_{i,d_i})$, $i = 1, \ldots, n$. Obtain the marginal parameter estimates $\hat{\boldsymbol{\alpha}}_j$ $(j = 1, \ldots, d)$ and the D-vine copula parameter estimates $\hat{\boldsymbol{\theta}}_{1:d}$ with $d = \max\{d_i | i = 1, \ldots, n\}$.

Step 1 (two-stage semiparametric): Following the global or sequential two-stage semiparametric estimation approach to fit the D-vine copula model of interest to the data $(y_{i,1}, \ldots, y_{i,d_i}, \delta_{i,1}, \ldots, \delta_{i,d_i})$, $i = 1, \ldots, n$. Obtain the survival marginal estimates $\hat{S}_j^{\mathrm{NA}}$ $(j = 1, \ldots, d)$ and the D-vine copula parameter estimates $\hat{\boldsymbol{\theta}}_{1:d}$ with $d = \max\{d_i | i = 1, \ldots, n\}$.

Step 2: Obtain the Nelson-Aalen estimate $\hat{G}$ of the censoring distribution G based on the observations $(y_{i,1} + y_{i,2} + \ldots + y_{i,d_i}, 1 - \delta_{i,d_i})$, $i = 1, \ldots, n$.

Step 3: Generate B bootstrap samples in the following way: For $b = 1, \ldots, B$ and $i = 1, \ldots, n$,

Step 3.1: sample copula data $\left(u_{i,1}^{(b)}, \ldots, u_{i,d}^{(b)}\right)$ from the fitted D-vine copula model with parameter vector $\hat{\boldsymbol{\theta}}_{1:d}$.

Step 3.2 (one-stage parametric): Generate gap times $\left(g_{i,1}^{(b)}, \ldots, g_{i,d}^{(b)}\right)$ via $g_{i,j}^{(b)} = S_j^{-1}\left(u_{i,j}^{(b)}; \hat{\boldsymbol{\alpha}}_j\right)$ $(j = 1, \ldots, d)$, where $S_j\left(\cdot; \hat{\boldsymbol{\alpha}}_j\right)$ follows the marginal distribution assumed to be known with estimated parameters $\hat{\boldsymbol{\alpha}}_j$.

Step 3.2 (two-stage semiparametric): Generate gap times $\left(g_{i,1}^{(b)}, \ldots, g_{i,d}^{(b)}\right)$ via $g_{i,j}^{(b)} = \left(\hat{S}_j^{\mathrm{NA}}\right)^{-1}\left(u_{i,j}^{(b)}\right)$ $(j = 1, \ldots, d)$, where $\hat{S}_j^{\mathrm{NA}}$ is the nonparametric survival marginal estimate obtained *in Step 1 (two-stage semiparametric).*

Step 3.3: Obtain event times $t_{i,j}^{(b)}$ by setting $t_{i,j}^{(b)} = \sum_{\ell=1}^{j} g_{i,\ell}^{(b)}$.

Step 3.4: Generate independent censoring times $c_i^{(b)}$ from $\hat{G}$.

Step 3.5: Obtain observed data. If $t_{i,1}^{(b)} > c_i^{(b)}$ set $d_i = 1$ and retain $y_{i,1}^{(b)} = c_i^{(b)}$, if $t_{i,2}^{(b)} > c_i^{(b)}$ set $d_i = 2$ and retain $(y_{i,1}^{(b)}, y_{i,2}^{(b)}) = (g_{i,1}^{(b)}, c_i^{(b)} - t_{i,1}^{(b)})$, etc. If $d_i = d$ distinguish between $d-1$ or d events, i.e. $(y_{i,1}^{(b)}, \ldots, y_{i,d-1}^{(b)}, y_{i,d}^{(b)}) = (g_{i,1}^{(b)}, \ldots, g_{i,d-1}^{(b)}, c_i^{(b)} - t_{i,d-1}^{(b)})$ or $(y_{i,1}^{(b)}, \ldots, y_{i,d-1}^{(b)}, y_{i,d}^{(b)}) = (g_{i,1}^{(b)}, \ldots, g_{i,d-1}^{(b)}, g_{i,d}^{(b)})$.
Define $\delta_{i,j}^{(b)} = 1$ for $j < d_i$ and $\delta_{i,d_i}^{(b)} = \mathbb{1}\left(y_{i,d_i}^{(b)} \leq c_i^{(b)} - t_{i,d_i-1}^{(b)}\right)$.

Step 3.6 (one-stage parametric): Under a prespecified parametric format for the marginal survival functions (for example Weibull), follow the global or sequential one-stage parametric estimation approach to fit the D-vine copula model of interest to the bootstrap data $(y_{i,1}^{(b)}, \ldots, y_{i,d_i}^{(b)}, \delta_{i,1}^{(b)}, \ldots, \delta_{i,d_i}^{(b)})$, $i = 1, \ldots, n$. Obtain the marginal parameter estimates $\hat{\boldsymbol{\alpha}}_j^{(b)}$ $(j = 1, \ldots, d)$ and the D-vine copula parameter estimates $\hat{\boldsymbol{\theta}}_{1:d}^{(b)}$ for bootstrap sample b.

Step 3.6 (two-stage semiparametric): Following the global or sequential two-stage semiparametric estimation approach to fit the D-vine copula model of interest to the bootstrap data $(y_{i,1}^{(b)}, \ldots, y_{i,d_i}^{(b)}, \delta_{i,1}^{(b)}, \ldots, \delta_{i,d_i}^{(b)})$, $i = 1, \ldots, n$. Obtain the D-vine copula parameter estimates $\hat{\boldsymbol{\theta}}_{1:d}^{(b)}$ for bootstrap sample b.

Step 4 (one-stage parametric): Calculate elementwise the empirical standard deviations of $\hat{\boldsymbol{\alpha}}_j^{(1)}, \ldots, \hat{\boldsymbol{\alpha}}_j^{(B)}$ $(j = 1, \ldots, d)$ and $\hat{\boldsymbol{\theta}}_{1:d}^{(1)}, \ldots, \hat{\boldsymbol{\theta}}_{1:d}^{(B)}$ to obtain bootstrap based standard errors for $\hat{\boldsymbol{\alpha}}_j$ $(j = 1, \ldots, d)$ and $\hat{\boldsymbol{\theta}}_{1:d}$.

Step 4 (two-stage semiparametric): Calculate elementwise the empirical standard deviations of $\hat{\boldsymbol{\theta}}_{1:d}^{(1)}, \ldots, \hat{\boldsymbol{\theta}}_{1:d}^{(B)}$ to obtain bootstrap based standard errors for $\hat{\boldsymbol{\theta}}_{1:d}$.

4.4.9 Model selection

The simulations in Section 4.4.4, Section 4.4.5 and Section 4.4.7 showed that the four proposed estimation procedures are on target. For two-stage semiparametric estimation we detected high sensitivity in case of heavy tail censoring. Further, Section 4.4.8 provides a bootstrapping algorithm to estimate standard errors based on a D-vine copula model fitted to induced dependent right-censored gap time data. Thus, for analyzing real data the only remaining modeling aspect is to select the best possible copula to describe the data at hand. To do so, we need a suitable model selection tool. Focus is on model accuracy to capture complex dependence patterns such as the ones investigated through the simulations in Section 4.4.7. Against this background, we explore in this section the effect of using an incorrect copula specification and the role of AIC as a valid model selection tool.

For both simulation settings considered in Section 4.4.7, we fit in addition to the correct D-vine copula specification as given in Table 4.15 a four-dimensional Clayton copula (4d) and an incorrect D-vine copula with all pair-copulas being of type Clayton (Clayton vine copula). In case of one-stage parametric estimation the format of the survival margins is correctly specified. In Table 4.16, we list for both simulation settings the preference by AIC, i.e. the proportion of data

Table 4.16: Results on copula selection by AIC under global and sequential one-stage parametric and two-stage semiparametric estimation for four-dimensional data (4d). The D-vine copula model captures in Setting 1: tail-behavior for subsequent gap times changing from lower tail-dependence (Clayton (C)) over no tail-dependence (Frank (F)) to upper tail-dependence (Gumbel (G)) with same overall dependence of $\tau_{1,2} = \tau_{2,3} = \tau_{3,4} = 0.5$; and in Setting 2: for Clayton copulas in $\mathcal{T}_1$ increasing dependence with $\tau_{1,2} = 0.3$, $\tau_{2,3} = 0.5$, $\tau_{3,4} = 0.7$. We consider the fit of a correctly specified D-vine copula, an incorrect Clayton D-vine and a Clayton copula. The AIC preference rate is based on 250 replications and samples of different sizes affected by either 15%, 30% or heavy tail 30% right-censoring.

				D-vine global			D-vine sequential		
				correct	incorrect	4d Clayton	correct	incorrect	4d Clayton
Setting 1 (Table 4.15)	Parametric one-stage	15%	250	1	0	0	1	0	0
			500	1	0	0	1	0	0
			1000	1	0	0	1	0	0
		30%	250	1	0	0	1	0	0
			500	1	0	0	1	0	0
			1000	1	0	0	1	0	0
		30% HT	250	0.996	0.004	0	0.996	0.004	0
			500	1	0	0	1	0	0
			1000	1	0	0	1	0	0
	Semiparametric two-stage	15%	250	1	0	0	1	0	0
			500	1	0	0	1	0	0
			1000	1	0	0	1	0	0
		30%	250	0.956	0.028	0.016	0.964	0.016	0.020
			500	1	0	0	1	0	0
			1000	1	0	0	1	0	0
		30% HT	250	0.844	0.116	0.040	0.836	0.108	0.056
			500	0.928	0.064	0.008	0.928	0.064	0.008
			1000	0.972	0.028	0	0.972	0.028	0
Setting 2 (Table 4.15)	Parametric one-stage	15%	250	0.964	0.036	0	0.964	0.036	0
			500	0.996	0.004	0	0.996	0.004	0
			1000	1	0	0	1	0	0
		30%	250	0.896	0.104	0	0.892	0.108	0
			500	0.960	0.040	0	0.960	0.040	0
			1000	0.992	0.008	0	0.992	0.008	0
		30% HT	250	0.804	0.176	0.020	0.800	0.176	0.024
			500	0.924	0.076	0	0.924	0.076	0
			1000	0.964	0.036	0	0.964	0.036	0
	Semiparametric two-stage	15%	250	0.956	0.044	0	0.960	0.040	0
			500	1	0	0	1	0	0
			1000	1	0	0	1	0	0
		30%	250	0.816	0.172	0.012	0.864	0.124	0.012
			500	0.908	0.092	0	0.936	0.064	0
			1000	0.976	0.024	0	0.992	0.008	0
		30% HT	250	0.424	0.344	0.232	0.472	0.260	0.268
			500	0.552	0.392	0.056	0.632	0.292	0.076
			1000	0.632	0.364	0.004	0.740	0.248	0.012

sets, for which each of the three model specifications performed best based on AIC. It follows that AIC is able to detect the correct copula model for the majority of simulated data sets, indicating that AIC is a valid tool for model selection. As expected, the AIC preference for the correct model increases as sample size grows, but decreases for a higher censoring rate. Also, AIC selects the correct model more often for one-stage parametric estimation as compared to two-stage semiparametric estimation. For the latter, heavy tail censoring again distorts estimation. Finally, for the D-vine copula according to Setting 1 (Clayton, Frank, Gumbel in $\mathcal{T}_1$) the correct vine is selected more often as compared to the D-vine copula according to Setting 2 (only Clayton in $\mathcal{T}_1$). This is to be expected, since the latter resembles a Clayton vine copula and a Clayton copula more closely, making model detection more difficult.

4.4.10 Data application

In this section, we use the proposed methodology to analyze the asthma data, which were already mentioned in previous sections. Data on 232 children are available. The children entered the study at the age of 6 months, at which they were randomized into a placebo group (113 children) or a treatment group (119 children). They were followed up for about 18 months. Due to limited follow-up, the time to the last asthma attack may not be recorded, but a lower right-censoring time may be observed instead. Meyer and Romeo (2015) model the association of gap times via Archimedean copulas and thereby impose the same type and strength of dependence between all gap times. However, an asthma attack further weakens the lungs and thus makes a child more prone to another attack. Therefore, the dependence between gap times is expected to change over time in type and/or strength. The simulations in Section 4.4.7 show that D-vine copulas can capture such features.

Table 4.17 indicates that only few children have more than four asthma attacks, making accurate estimation of the survival margins and of the association from the fifth gap time on rather difficult. Hence, we focus on the first four gap times, i.e. we use the data of attack 1 up to attack 4 even if there is information on subsequent attacks. By doing so, each child experiences at least one asthma attack and 97 children have at least four attacks. For 25 of these children the last asthma attack is right-censored (8 in the treatment group and 17 in the control group). The overall censoring rate is 22.13%.

Dependence modeling

To explore the asthma data and to decide on the estimation strategy, we plot the Nelson-Aalen estimate for the survival function of the total times. We consider the full data sample as well as the data subsamples based on treatment to accommodate a possible effect of the latter on the margins and on the dependence structure. As can be seen in Figure 4.19 each sample is heavily right-censored with accumulation of censored observations at late time points, i.e. the Nelson-Aalen estimates exhibit a heavily right-censored tail, leading to a leveling off at a survival value around 0.6 for the full data set, around 0.7 for the treated group and around 0.5 for the placebo group. Based on the simulation results and the guidelines in Figure 4.16, we opt to apply a one-stage parametric estimation approach to model the dependence structure in the

Table 4.17: Frequency table of the number of children in the asthma data considering the full sample as well as subsamples based on treatment assignment and censoring status of the last asthma attack. The last column (4_{mod}) corresponds to the number of clusters of size 4 after modifying the original data, i.e. even if the original cluster size is larger than 4 only gap times 1 to 4 are considered. For example, for the modified full sample the number of clusters of size 4 with last observation being a true event is 72. This is the sum of all (full sample) entries with cluster size > 4 (both event and censored) and $= 4$ (event). The number of children with a censored fourth attack remains unaffected by the data modification.

	#Children with	Cluster size																
	last attack	2	3	4	5	6	7	8	9	10	11	12	13	14	15	20	39	4_{mod}
Full	event	1	3	1	2	0	0	2	0	1	0	0	1	1	1	1	0	72
	censored	87	44	25	14	10	8	7	8	1	7	2	2	2	0	0	1	25
Treatment	event	0	3	1	1	0	0	1	0	1	0	0	0	0	0	1	0	27
	censored	50	25	8	6	2	2	3	2	1	4	1	0	1	0	0	0	8
Control	event	1	0	0	1	0	0	1	0	0	0	0	1	1	1	0	0	45
	censored	37	19	17	8	8	6	4	6	0	3	1	2	1	0	0	1	17

asthma data. As in Meyer and Romeo (2015), we assume Weibull survival margins, but opposed to them, we allow for a flexible association pattern as modeled by diverse D-vine copulas.

The induced dependent right-censoring present in the asthma data makes model specification challenging. Common data exploration tools cannot be applied. For example, due to the heavy censoring for late gap times, pairs plots on the time scale, respectively on the copula scale, show an empty upper right corner, respectively an empty lower left corner. Thus, visual inspection is obscured (see Figure 4.6 in Section 4.3.2). To unravel the association in the asthma data, we therefore fit a large variety of copula models. We consider the independence copula as well as the four-dimensional (4d) Clayton, Gumbel and Frank copulas, together with several four-dimensional D-vine copulas. For the latter, we consider in tree $\mathcal{T}_1$ all possible permutations of Clayton, Gumbel and/or Frank copulas. In trees $\mathcal{T}_2$ and $\mathcal{T}_3$, all pair-copulas are taken to be Frank. This results in a total of 27 D-vine copulas.

Table 4.18 gives the results of sequential and global one-stage parametric estimation in terms of Kendall's τ for the three best D-vine copulas as selected by AIC, the independence copulas as well as for the Archimedean copulas. Results on marginal estimation are listed in Table B.26 in Appendix B.7. For global loglikelihood maximization the parameter estimates obtained from sequential optimization are used as starting values. For each data sample, all D-vine copulas perform better than the best Archimedean copula based on AIC. While the pair-copula families in $\mathcal{T}_1$ of the best D-vine copula are the same for all samples, the best Archimedean copula varies among the three data sets. Recall that for the latter all dependencies are described by a single parameter. In the asthma data, this dependence is small and close to zero (as confirmed by AIC). D-vine copulas provide a more local view on association and thereby allow varying dependence between gap times. While the Kendall's τ values in trees $\mathcal{T}_2$ and $\mathcal{T}_3$ are quite small, the estimates in tree $\mathcal{T}_1$, i.e. for $\tau_{1,2}$, $\tau_{2,3}$ and $\tau_{3,4}$, increase over time. This finding supports the

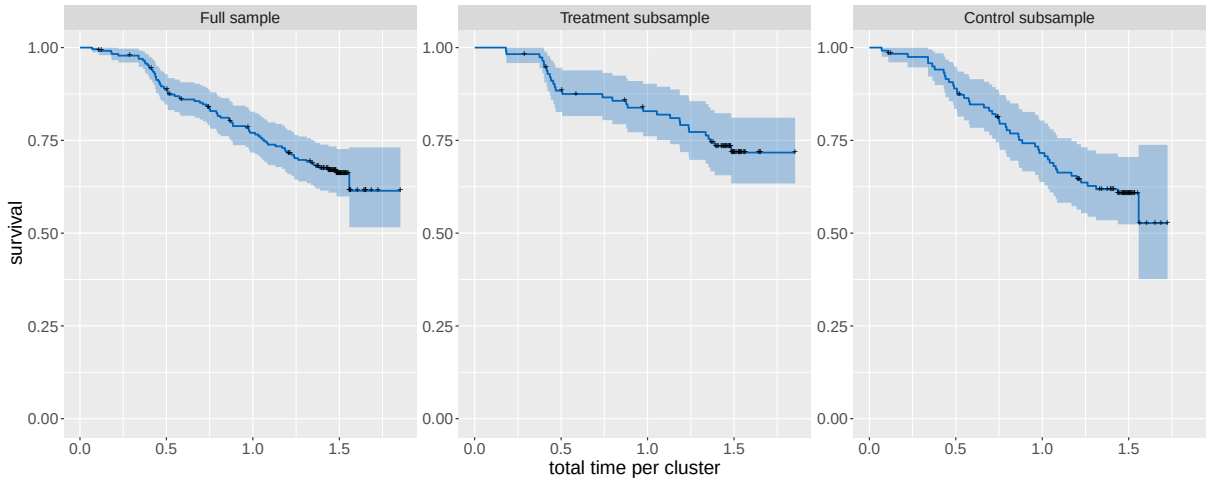


Figure 4.19: Nelson-Aalen estimate of the survival function for the total times (in years).

initial intuition that with each additional asthma attack, children are more prone to a relapse. The fact that a Gumbel copula is chosen for the pair 2-3 suggests that the smaller gap time 2 is, the faster a third asthma attack will follow. The same holds true for pair 3-4. For pair 1-2 there is no clear best copula family, which might be explained by the low Kendall's τ values of on average 0.10. For such a low value the specific features of a copula family such as lower or upper tail-dependence are less pronounced. Interestingly, the estimates for $\tau_{2,3}$ and $\tau_{3,4}$ for the treatment and control group are quite alike, while there is a significant difference for $\tau_{1,2}$. For treated children the occurrences of a first and a second asthma attack are close to independence, while for children in the placebo group the estimate for $\tau_{1,2}$ is about 0.18. This suggests that the medical treatment has a clear influence on the (time to) occurrence of a second asthma attack. However, whenever a treated child has a relapse, subsequent attacks are as likely as for untreated children. In general and most pronounced for the treatment group, Kendall's τ values including the first gap are smaller as compared to those not including the first gap. For D-vine copula models, the sequential and global estimation approach show very similar results with a slight model improvement (based on AIC) in case of global optimization.

The standard errors are based on 1000 bootstrap samples (see Section 4.4.8). In general, standard errors increase for estimates corresponding to late gap times. For them, fewer data are available due to the unbalanced data setting. Note that while in the asthma data there are no clusters of size 1, this may occur in the bootstrap samples (as in many data settings). Table B.27 in Appendix B.7 contains information on average cluster sizes and the average censoring percentage among the bootstrap replications showing that the data generation within the bootstrap succeeds to mimic the features of the asthma data characteristics quite accurately.

Table 4.18: AIC values and Kendall's τ estimates with standard errors based on 1000 bootstrap samples (in parentheses) of copula models fitted to each of the three samples of the asthma data using sequential and global one-stage parametric estimation. In case of Archimedean copulas the Frank (4dF), Gumbel (4dG), Clayton (4dC) and the Independence (4dInd) copula are considered. In case of D-vine copulas only the three best models are shown with Frank being the copula family in trees $\mathcal{T}_2$ and $\mathcal{T}_3$.

			AIC	$\tau_{1,2}/\tau$	$\tau_{2,3}$	$\tau_{3,4}$	$\tau_{1,3;2}$	$\tau_{2,4;3}$	$\tau_{1,4;2,3}$
Sequential one-stage parametric	Full	FGG	210.335	0.121 (0.052)	0.236 (0.059)	0.321 (0.062)	-0.054 (0.064)	0.290 (0.080)	-0.090 (0.079)
		CGG	212.814	0.122 (0.063)	0.244 (0.059)	0.326 (0.061)	-0.055 (0.063)	0.301 (0.080)	-0.089 (0.077)
		GGG	213.341	0.098 (0.047)	0.232 (0.059)	0.317 (0.062)	-0.054 (0.065)	0.290 (0.081)	-0.085 (0.081)
	Treatment	FGG	147.911	0.043 (0.077)	0.233 (0.094)	0.317 (0.103)	0.016 (0.104)	0.420 (0.119)	-0.161 (0.134)
		CGG	147.999	0.051 (0.077)	0.234 (0.094)	0.319 (0.102)	0.015 (0.103)	0.426 (0.117)	-0.160 (0.134)
		GGG	148.479	0.000 (0.035)	0.231 (0.095)	0.316 (0.103)	0.015 (0.107)	0.425 (0.119)	-0.162 (0.138)
	Control	FGG	67.229	0.186 (0.069)	0.240 (0.077)	0.298 (0.082)	-0.114 (0.083)	0.189 (0.106)	-0.034 (0.101)
		FGF	68.784	0.186 (0.069)	0.240 (0.077)	0.285 (0.099)	-0.114 (0.083)	0.165 (0.108)	-0.027 (0.102)
		GGG	69.801	0.165 (0.066)	0.235 (0.077)	0.294 (0.083)	-0.114 (0.085)	0.192 (0.107)	-0.026 (0.104)
Global one-stage parametric	Full	FGG	210.103	0.122 (0.052)	0.258 (0.058)	0.333 (0.061)	-0.050 (0.065)	0.293 (0.081)	-0.090 (0.079)
		CGG	212.582	0.129 (0.064)	0.266 (0.058)	0.338 (0.061)	-0.052 (0.064)	0.304 (0.080)	-0.090 (0.078)
		GGG	213.026	0.100 (0.047)	0.253 (0.059)	0.329 (0.062)	-0.050 (0.066)	0.292 (0.081)	-0.087 (0.081)
		4dF	233.678	0.055 (0.025)					
		4dG	235.377	0.047 (0.030)					
		4dC	236.456	0.064 (0.041)					
		4dInd	237.478						
	Treatment	FGG	147.797	0.046 (0.077)	0.260 (0.092)	0.334 (0.101)	0.017 (0.106)	0.419 (0.120)	-0.162 (0.140)
		CGG	147.876	0.058 (0.080)	0.262 (0.092)	0.336 (0.100)	0.016 (0.104)	0.426 (0.119)	-0.160 (0.137)
		GGG	148.365	0.002 (0.036)	0.258 (0.092)	0.333 (0.101)	0.015 (0.108)	0.424 (0.121)	-0.162 (0.142)
		4dF	154.797	0.036 (0.031)					
		4dG	155.941	0.000 (0.022)					
		4dC	155.496	0.037 (0.051)					
		4dInd	153.941						
	Control	FGG	67.079	0.185 (0.069)	0.259 (0.076)	0.308 (0.082)	-0.111 (0.084)	0.193 (0.107)	-0.034 (0.103)
		FGF	68.722	0.185 (0.069)	0.241 (0.077)	0.286 (0.100)	-0.112 (0.084)	0.167 (0.110)	-0.026 (0.104)
		GGG	69.616	0.166 (0.067)	0.254 (0.077)	0.304 (0.083)	-0.112 (0.086)	0.194 (0.107)	-0.028 (0.106)
		4dF	78.349	0.063 (0.032)					
		4dG	77.859	0.062 (0.037)					
		4dC	79.947	0.073 (0.050)					
		4dInd	80.283						

Conditional prediction

In addition to Barthel et al. (2018b), we further analyze in this thesis quantiles for the time to a relapse conditional on the individual risk profile of a child, which is specified by the previous observed gap times. More precisely, in Section 4.1.2 we outlined that in case of an underlying D-vine copula model there is an analytical expression for the conditional survival function which describes the time to asthma attack $(j+1)$ given the observed disease history $(y_{i,1}, \dots, y_{i,j})$ of child i. Conditional quantiles can thus be obtained through inversion.

For this purpose, using one-stage parametric estimation we fit for each child i in the asthma data a D-vine copula model to the data where the observations corresponding to child i, i.e. $(y_{i,1}, \dots, y_{i,d_i}, \delta_{i,1}, \dots, \delta_{i,d_i})$, are removed. Then, based on the fitted D-vine copula model we calculate the 90% out-of-sample prediction interval, i.e. the conditional 5% and the conditional 95% quantile, for the time until the k-th asthma attack $(k = 2, \dots, d_i)$ given the observed gap times $y_{i,1}, \dots, y_{i,k-1}$. This is done for the complete data set as well as for the subsample based on treatment to which child i belongs.

Table 4.19 provides the percentages of children for who the actual observed gap time lies within the corresponding conditional 90% prediction interval. Results based on both the D-vine copula fitted to the full sample with the considered child's data removed and the D-vine copula fitted to the corresponding subsample with the considered child's data removed are shown. Clearly, the percentage decreases with increasing gap time. For them, less data are available due to the unbalanced data setting. In general, the prediction intervals cover the true observations satisfactorily. Further, Figure 4.20 and Figure 4.21 show for representative children in the treatment group and the control group, respectively, the estimated conditional 90% prediction intervals together with the corresponding true observations. The ID of each child refers to its position in the corresponding subsample. Recall that there are 113 children in the treatment group and 119 in the control group. Results for children with a cluster size of 4 are in the top row of Figure 4.20 and Figure 4.21. For them, there are three prediction intervals corresponding to gap time 2 given gap time 1, gap time 3 given gap times 1 and 2 and gap time 4 given gap times 1 to 3. Results for children with cluster size 3 and cluster size 2 are shown in the middle rows and the bottom rows, respectively. Further, observations corresponding to a true gap time are marked by a cross. Right-censored observations are denoted with a circle. Recall that in the latter case the corresponding true gap time is larger than the observed value.

Table 4.19: Percentage of children for who the true observation of a gap time lies within the corresponding conditional 90% prediction interval. Conditional predictions are based on both the D-vine copula fitted to the full sample with the considered child's data removed and based on the D-vine copula fitted to the subsample (treatment or control group) with the considered child's data removed.

D-vine copula based conditional quantile prediction based on	Treatment group			Control group		
	gap 2	gap 3	gap 4	gap 2	gap 3	gap 4
full sample	100%	94%	91%	99%	98%	85%
subsample	100%	95%	89%	99%	96%	87%

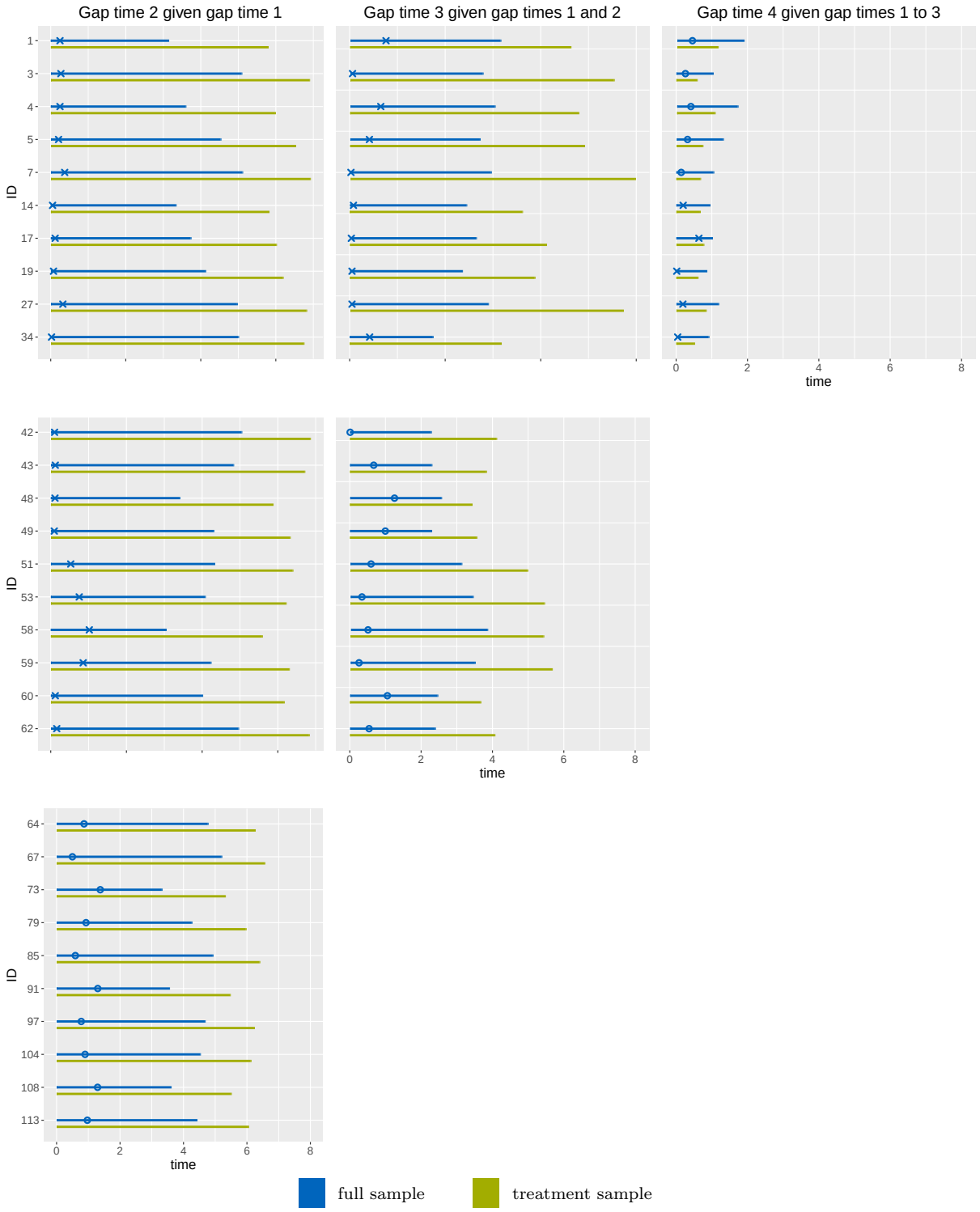


Figure 4.20: Conditional 90% prediction intervals for representative children in the **treatment subsample** of the asthma data. Prediction intervals corresponding to clusters of size 4, size 3 and size 2, respectively, are shown in the top row, the middle row and the bottom row, respectively. Predictions based on the D-vine copula model fitted to the full sample with the considered child's data removed are shown in blue. Predictions based on the D-vine copula model fitted to the treatment subsample with the considered child's data removed are shown in green. Observations corresponding to true gap times are denoted by $\times$. Right-censored observations are denoted by $\circ$.

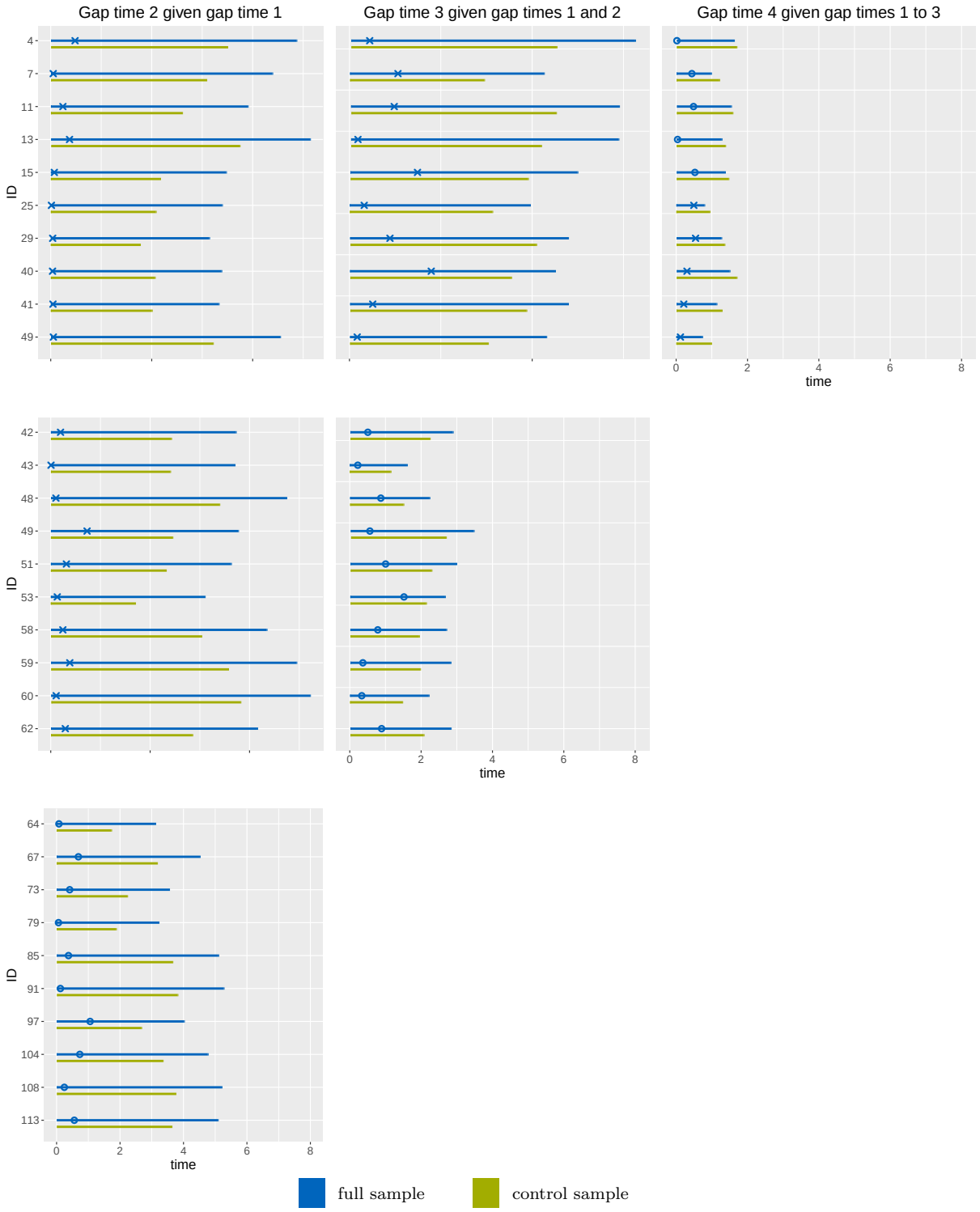


Figure 4.21: Conditional 90% prediction intervals for representative children in the **control subsample** of the asthma data. Prediction intervals corresponding to clusters of size 4, size 3 and size 2, respectively, are shown in the top row, the middle row and the bottom row, respectively. Predictions based on the D-vine copula model fitted to the full sample with the considered child's data removed are shown in blue. Predictions based on the D-vine copula model fitted to the control subsample with the considered child's data removed are shown in green. Observations corresponding to true gap times are denoted by $\times$. Right-censored observations are denoted by $\circ$.

We see that especially the prediction intervals for gap time 2 given gap time 1 are rather large and indicate a heavily right skewed conditional distribution of the gap times. This finding is more evident for results based on the treatment subsample as compared to the full sample, but it is less pronounced based on the control subsample as compared to the full data. This observation might be due to the significantly different association between gap time 1 and gap time 2 for children of the two subsamples (see Table 4.18). For gap time 4 given gap times 1 to 3, the conditional 90% prediction intervals become significantly tighter while still covering the true observations sufficiently well (compare Table 4.19). Note that even if right-censored observations are smaller than the estimated conditional 5% quantile, the corresponding true gap time may still be covered by the prediction interval.

To conclude, the D-vine copula based estimation approaches to model induced dependent right-censored gap time data allow for an analytical closed form expression of the conditional gap time distributions given an individual history of previous gap times. The resulting possibility to obtain conditional quantiles is of particular interest and highly relevant in applications such as the prediction of a relapse in the asthma study.

4.5 Discussion

In this chapter, we investigated R-vine copulas to model dependence patterns in multivariate right-censored event time data. Prior to this work, R-vine theory had only been developed for complete data. First, basic notations and results commonly used in R-vine copula based methodology were reformulated in survival terms in Section 4.1. Then, two data settings, which are typical in time-to-event studies, were discussed in detail.

In Section 4.2 and Section 4.3, balanced data subject to common right-censoring were considered. The developed estimation procedure was conducted in two subsequent steps (two-stage approach). First, the marginal distributions were estimated considering standard parametric and nonparametric estimation techniques for univariate right-censored data (Section 4.1.3). Second, the dependence structure was modeled. The likelihood contributions for right-censored quadruple data in terms of R-vine copula components were provided (Barthel (2015), Appendix B.1.1). Due to the right-censoring single and double integrals showed up in the copula likelihood expression such that numerical integration was needed for its evaluation. Hence, for dependence modeling a sequential estimation approach that facilitates the computational challenges of the likelihood optimization was proposed (Section 4.2.3). For right-censored trivariate data a simulation study gave evidence that the presented estimators are on target (Section 4.2.4). To obtain standard errors for likelihood based parameter estimates an R-vine copula based parametric bootstrapping algorithm for right-censored data was proposed (Section 4.3.1). For the four-dimensional mastitis data, while stressing the general difficulty of model selection in the presence of heavy right-censoring, it was shown how an appropriate R-vine copula model can be selected for data at hand. Both the full and the sequential estimation approach were used (Section 4.3.2). The results qualified the latter as the preferable estimation technique in practice. It provides comparable estimation results while significantly simplifying the numerically challenging optimization

problem. Our findings for the mastitis data were in line with Geerdens et al. (2016a), where the Joe-Hu family was used for flexible dependence modeling in right-censored event time data.

In Section 4.4, dependence between recurrent event times subject to right-censoring was investigated. We addressed several challenges arising when modeling gap time association such as the presence of induced dependent right-censoring and the unbalanced nature of the data. Due to their construction principle, which allows for a temporal ordering of the variables, D-vine copulas were the natural choice for dependence modeling. In total, four estimation strategies were suggested: one-stage parametric estimation (Section 4.4.4) and two-stage semiparametric estimation (Section 4.4.5) combined with global and sequential estimation. Extensive simulations in three and four dimensions underlined the good finite sample performance of all estimation strategies (Section 4.4.7). For the two-stage semiparametric modeling strategies the impact of heavy induced dependent right-censoring was studied. Given its sensitivity with respect to heavy tail censoring, guidelines on the practical use of the four estimation approaches were formulated (Section 4.4.6). Further, methods for standard error estimation (Section 4.4.8) and model selection were provided (Section 4.4.9). The application to data on children suffering from asthma provided new insights on the evolution of the disease (Section 4.4.10). These findings could not be detected by Archimedean copulas, which impose a too restrictive dependence structure to the data. The flexibility of D-vine copula models was further highlighted in the context of conditional prediction of the time until relapse given the individual disease history of children in the asthma data.

Both projects discussed in this chapter stressed the need for more flexible copula models as compared to less elaborated ones such as elliptical or Archimedean copulas in the context of right-censored event time data. They also showed that the data complexity due to right-censoring makes the statistical analysis of multivariate event time data highly challenging with regard to numerical demand and computational manageability.

Chapter 5

Conclusion and outlook

In this thesis, novel methodologies to model and forecast time-series of realized covariance matrices and dependence patterns for multivariate right-censored event time data were proposed. In both research fields, these projects were the first attempts to incorporate regular vines and regular vine copulas into the statistical analysis. While working on the presented results several interesting new aspects and ideas have arisen being potential starting points for future research.

Parsimonious modeling and forecasting of realized volatility time-series in high dimensions

In Chapter 3, one of the striking advantages of the partial correlation vine data transformation approach is that model parsimony is achieved. Not only does careful selection of the underlying regular vine structure result in easy to model time-series data as compared to normally challenging volatility data, but also dependence between the model components, namely realized standard and realized partial correlations as well as realized variances, is less pronounced as between the model components obtained through the Cholesky factorization. To exploit the full potential of this modeling aspect, we advocate for future work to apply the partial correlation vine data transformation approach to high-frequency data in higher dimensions. In this case, multivariate time-series models including factor copulas for dependence modeling might lead to even more model parsimony (Krupskii and Joe, 2013).

R-vine copula based modeling of right-censored time-to-event data including covariates

Given the prevalence of multivariate right-censored data in many biomedical and health care related studies, the availability of reliable and applicable (also for practitioners) statistical models is crucial. In Chapter 4, we showed that despite the increased data complexity due to right-censoring the proposed R-vine copula based models are able to capture the inherent dependence patterns. In particular, less elaborate copula based models were insufficient to detect important data features. Having the basic methodology at hand provides room for further research on the use of R-vine copulas in the presence of censoring.

Often a data set includes one or more covariates. In the context of copula models a covariate can affect the survival margins and/or the dependence structure. If the covariate is at the level of the cluster and only takes a few values, the data set can be split into several subsets and the

proposed copula modeling can be used for each subset separately as done for the asthma data in Section 4.4.10. If the covariate at the level of the cluster is continuous, then one can model the copula parameters as a function of the covariate (for example linear) and further proceed as proposed in Section 4.2. If a covariate is not at the level of the cluster it is not possible to discuss its impact on the association and the covariate can only be included in the margins. For example, a Cox model can be used in the first estimation step. Nonparametric marginal estimation is more involved when covariates are present. An option is to apply the Beran estimator or an extended version of it (Beran, 1981).

R-vine copula based quantile regression for censored response data

A research field, which gained a lot of interest in the recent years, is the prediction of quantiles of a random variable conditioned on the realization of other variables. Possible application areas range from portfolio optimization or risk management in finance to the prediction of water levels in hydrology, the claim sizes for insurances or as already shortly discussed in Section 4.4.10 the time until an asthma attack given the individual health profile of a child. Clearly, the latter is only one example out of many in a clinical context. Here, the response variable might be subject to right-censoring. For complete data, Kraus and Czado (2017) propose a D-vine copula based quantile regression method, which – given the ability for more flexible dependence modeling between the response variable and the covariates – is able to outperform classical methods for quantile prediction. Herrmann (2018) extends these ideas to the more general and even more flexible class of R-vine copulas. The results of this master's thesis will soon be submitted for publication as a scientific research paper.

In case of right-censored response data, the methodology needs additional adaption. This constitutes promising ongoing research. Besides the presence of right-censoring, which as is well known complicates the statistical analysis of data, discrete and categorical explanatory variables are common in regression data. For complete data and using parametric pair-copulas, D-vine and R-vine copula based techniques to handle ordinal categorical variables, which are monotonically related with other explanatory variables, are discussed in Schallhorn et al. (2017) and Chang and Joe (2018), respectively. Nagler (2018) suggests to add random noise to discrete variables and to model the jittered data. However, this approach only works in an unrestricted nonparametric framework. Since the simplifying assumption is a structural assumption, the latter might cause problems when applying simplified nonparametric pair-copula constructions as introduced by Nagler and Czado (2016) in the above context. Geerdens et al. (2016b) propose a nonparametric estimator for a bivariate survival copula for data subject to random right-censoring. For nominal categorical explanatory variables, data splitting as suggested in the previous paragraph and as done for the asthma data in Section 4.4.10 only is an option in case of high sample sizes and few nominal explanatory variables with a small number of categories. Usually, for k categories $k-1$ binary dummy variables are considered in standard regression methods. Here, for each observation unit at most one of the dummy variables will equal 1 while all others are equal to zero. In this case, a possible application of pair-copula constructions in particular under the simplifying assumption has again to be investigated with great care.

Joint modeling of longitudinal and time-to-event data using R-vine copulas

A final project combines the methodology presented in Chapter 4, work on modeling of repeated measurements using D-vine copulas (Killiches and Czado, 2018), ideas from R-vine quantile regression for right-censored data as well as the inclusion of covariates. In many medical studies, along with the possibly right-censored time to a specific event repeated measurements of several biomarkers such as pulse, blood pressure or laboratory tests of tissues are available. While – to the best of our knowledge – existing methodology in the majority of times only captures the influence of a single biomarker measurement history on the event of interest, we propose an R-vine copula model, which simultaneously involves three layers of dependence: the dependence between each of the biomarkers and the event time, the dependence between the different biomarkers, and for each of the biomarkers the temporal dependence within the measurement history.

The variety of possible future projects shows the need for flexible dependence models in many highly relevant and challenging research fields. Thus, R-vine copula models will certainly continue growing in popularity and in their application spectrum.

Bibliography

Aas, K. (2016). Pair-copula constructions for financial applications: A review. *Econometrics*, 4(4):43.

Aas, K., Czado, C., Frigessi, A., and Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, 44(2):182–198.

Andersen, E. W. (2005). Two-stage estimation in copula models used in family studies. *Lifetime Data Analysis*, 11(3):333–350.

Andersen, T. G. and Bollerslev, T. (1997). Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *The Journal of Finance*, 52(3):975–1005.

Andersen, T. G., Bollerslev, T., Christoffersen, P. F., and Diebold, F. X. (2006). Volatility and correlation forecasting. *Handbook of Economic Forecasting*, 1:777–878.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics*, 61(1):43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–625.

Anderson, T. W. (1958). *An introduction to multivariate statistical analysis*. Wiley New York.

Bai, X., Russell, J. R., and Tiao, G. C. (2003). Kurtosis of garch and stochastic volatility models with non-normal innovations. *Journal of Econometrics*, 114(2):349–360.

Baillie, R. T., Bollerslev, T., and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 74(1):3–30.

Barndorff-Nielsen, O. E. and Shephard, N. (2004). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica*, 72(3):885–925.

Barthel, N. (2015). Multivariate Survival Analysis using Vine-Copulas. Master's thesis, Technische Universität München.
https://mediatum.ub.tum.de/doc/1276329/1276329.pdf.

Barthel, N., Czado, C., and Okhrin, Y. (2018a). A partial correlation vine based approach for modeling and forecasting multivariate volatility time-series. *In revision at Computational Statistics & Data Analysis*.

Barthel, N., Geerdens, C., Czado, C., and Janssen, P. (2018b). Dependence modeling for recurrent event times subject to right-censoring with d-vine copulas. *To appear in Biometrics. doi:10.1111/biom.13014*.

Barthel, N., Geerdens, C., Killiches, M., Janssen, P., and Czado, C. (2018c). Vine copula based likelihood estimation of dependence patterns in multivariate event time data. *Computational Statistics & Data Analysis*, 117:109–127.

Bauer, G. H. and Vorkink, K. (2011). Forecasting multivariate realized stock market volatility. *Journal of Econometrics*, 160(1):93–101.

Bedford, T. and Cooke, R. M. (2002). Vines: A new graphical model for dependent random variables. *Annals of Statistics*, pages 1031–1068.

Beran, R. (1981). Nonparametric regression with randomly censored survival data. Technical report, Technical Report, Univ. California, Berkeley.

Bernstein, D. S. (2005). *Matrix mathematics: Theory, facts, and formulas with application to linear systems theory*, volume 41. Princeton university press Princeton.

Brechmann, E. C. and Czado, C. (2015). Coparmultivariate time series modeling using the copula autoregressive model. *Applied Stochastic Models in Business and Industry*, 31(4):495–514.

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

Brechmann, E. C., Heiden, M., and Okhrin, Y. (2018). A multivariate volatility vine copula model. *Econometric Reviews*, 37(4):281–308.

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

Brechmann, E. C. and Joe, H. (2015). Truncation of vine copulas using fit indices. *Journal of Multivariate Analysis*, 138:19–33.

Cappiello, L., Engle, R. F., and Sheppard, K. (2006). Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics*, 4(4):537–572.

Chang, B. and Joe, H. (2018). Prediction based on conditional distributions of vine copulas. *arXiv preprint arXiv:1807.08429*.

Chen, X., Fan, Y., Pouzo, D., and Ying, Z. (2010). Estimation and model selection of semiparametric multivariate survival functions under general censorship. *Journal of Econometrics*, 157(1):129 – 142.

Chiriac, R. and Voev, V. (2011). Modelling and forecasting multivariate realized volatility. *Journal of Applied Econometrics*, 26(6):922–947.

Christensen, K., Kinnebrock, S., and Podolskij, M. (2010). Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. *Journal of Econometrics*, 159(1):116–133.

Cook, R. J. and Lawless, J. (2007). *The statistical analysis of recurrent events*. Springer Science & Business Media.

Cooke, R. M., Joe, H., and Chang, B. (2015). Vine regression.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2):174–196.

Corsi, F., Mittnik, S., Pigorsch, C., and Pigorsch, U. (2008). The volatility of realized volatility. *Econometric Reviews*, 27(1-3):46–78.

Czado, C. (2010). Pair-copula constructions of multivariate copulas. In *Copula Theory and its Applications*, pages 93–109. Springer.

Czado, C. and Min, A. (2011). Bayesian inference for d-vines: estimation and model selection. In *Dependence Modeling: Vine Copula Handbook*, pages 249–264. World Scientific.

Davison, A. C., Hinkley, D. V., et al. (1997). *Bootstrap methods and their application*, volume 1. Cambridge university press.

de Uña-Álvarez, J. and Meira-Machado, L. F. (2008). A simple estimator of the bivariate distribution function for censored gap times. *Statistics & Probability Letters*, 78(15):2440–2445.

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

Doléans-Dade, C. and Meyer, P.-A. (1970). Intégrales stochastiques par rapport aux martingales locales. In *Séminaire de Probabilités IV Université de Strasbourg*, pages 77–107. Springer.

Duchateau, L. and Janssen, P. (2008). *The frailty model*. Springer Science & Business Media.

Elidan, G. (2013). Copulas in machine learning. In *Copulae in mathematical and quantitative finance*, pages 39–60. Springer.

Embrechts, P., Lindskog, F., and McNeil, A. J. (2003). Modelling dependence with copulas and applications to risk management. In *Handbook of Heavy Tailed Distributions in Finance*, pages 329–384. Elsevier.

Engle, R. F. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350.

Erhardt, T. M., Czado, C., and Schepsmeier, U. (2015). Spatial composite likelihood inference using local C-vines. *Journal of Multivariate Analysis*, 138:74–88.

Erhardt, V. and Czado, C. (2012). Modeling dependent yearly claim totals including zero claims in private health insurance. *Scandinavian Actuarial Journal*, 2012(2):106–129.

Fernández, C. and Steel, M. F. (1998). On bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, 93(441):359–371.

Fischer, M., Kraus, D., Pfeuffer, M., and Czado, C. (2017). Stress testing german industry sectors: Results from a vine copula based quantile regression. *Risks*, 5(3):38.

Geerdens, C., Claeskens, G., and Janssen, P. (2016a). Copula based flexible modeling of associations between clustered event times. *Lifetime Data Analysis*, 22(3):363–381.

Geerdens, C., Janssen, P., and Veraverbeke, N. (2016b). Large sample properties of nonparametric copula estimators under bivariate censoring. *Statistics*, 50(5):1036–1055.

Gençay, R., Dacorogna, M., Muller, U. A., Pictet, O., and Olsen, R. (2001). *An introduction to high-frequency finance*. Academic press.

Georges, P., Lamy, A.-G., Nicolas, E., Quibel, G., and Roncalli, T. (2001). Multivariate survival modelling: A unified approach with copulas. *Available at SSRN 1032559*.

Golosnoy, V., Gribisch, B., and Liesenfeld, R. (2012). The conditional autoregressive wishart model for multivariate stock market volatility. *Journal of Econometrics*, 167(1):211–223.

Gouriéroux, C., Jasiak, J., and Sufana, R. (2009). The wishart autoregressive process of multivariate stochastic volatility. *Journal of Econometrics*, 150(2):167–181.

Gruber, L., Czado, C., et al. (2015). Sequential bayesian model selection of regular vine copulas. *Bayesian Analysis*, 10(4):937–963.

Gruber, L. F. and Czado, C. (2018). Bayesian model selection of regular vine copulas. *Bayesian Analysis*, 13(4):1107–1131.

Halbleib, R. and Voev, V. (2014). Forecasting covariance matrices: A mixed approach. *Journal of Financial Econometrics*, 14(2):383–417.

Hansen, P. R., Huang, Z., and Shek, H. H. (2012). Realized GARCH: A joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*, 27(6):877–906.

Hansen, P. R., Lunde, A., and Nason, J. M. (2011). The model confidence set. *Econometrica*, 79(2):453–497.

Hayashi, T., Yoshida, N., et al. (2005). On covariance estimation of non-synchronously observed diffusion processes. *Bernoulli*, 11(2):359–379.

Herrmann, J. (2018). Regular vine copula based quantile regression. Master's thesis, Technische Universität München.

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

Hobæk-Haff, I. et al. (2013). Parameter estimation for pair-copula constructions. *Bernoulli*, 19(2):462–491.

Hofert, M. (2008). Sampling archimedean copulas. *Computational Statistics & Data Analysis*, 52(12):5163–5174.

Hougaard, P. (2000). *Analysis of Multivariate Survival Data*, volume 564. Springer New York.

Jacod, J. (1994). Limit of random measures associated with the increments of a brownian semimartingale. *preprint*, 120:155–162.

Joe, H. (1993). Parametric families of multivariate distributions with given margins. *Journal of Multivariate Analysis*, 46(2):262–282.

Joe, H. (1996). Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters. In Rüschendorf, L., Schweizer, B., and Taylor, M. D., editors, *Distributions with fixed marginals and related topics*, pages 120–141. Institute of Mathematical Statistics, Hayward.

Joe, H. (1997). *Multivariate models and dependence concepts*. Chapman and Hall, London.

Joe, H. (2005). Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of Multivariate Analysis*, 94(2):401–419.

Joe, H. (2014). Dependence Modeling with Copulas, Chapman Hall/CRC. *Published June/July*.

Joe, H., Cooke, R. M., and Kurowicka, D. (2011). Regular vines: generation algorithm and number of equivalence classes. In *Dependence Modeling: Vine Copula Handbook*, pages 219–231. World Scientific.

Joe, H. and Hu, T. (1996). Multivariate distributions from mixtures of max-infinitely divisible distributions. *Journal of Multivariate Analysis*, 57(2):240–265.

Joe, H. and Xu, J. (1996). The estimation method of inference functions for margins for multivariate models. *Journal of Multivariate Analysis*, 166.

Katoh, N., Ibaraki, T., and Mine, H. (1981). An algorithm for finding k minimum spanning trees. *SIAM Journal on Computing*, 10(2):247–255.

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, pages 81–93.

Killiches, M. and Czado, C. (2018). A D-vine copula-based model for repeated measurements extending linear mixed models with homogeneous correlation structure. *Biometrics*, 74(3):997–1005.

Kraus, D. and Czado, C. (2017). D-vine copula based quantile regression. *Computational Statistics & Data Analysis*, 110:1–18.

Krupskii, P. and Joe, H. (2013). Factor copula models for multivariate data. *Journal of Multivariate Analysis*, 120:85–101.

Kurowicka, D. and Cooke, R. (2003). A parameterization of positive definite matrices in terms of partial correlation vines. *Linear Algebra and its Applications*, 372:225–251.

Kurowicka, D. and Cooke, R. (2006a). Completion problem with partial correlation vines. *Linear Algebra and its Applications*, 418(1):188–200.

Kurowicka, D. and Cooke, R. M. (2006b). *Uncertainty Analysis with High Dimensional Dependence Modelling*. John Wiley & Sons.

Kurowicka, D. and Joe, H. (2011). *Dependence modeling: Vine Copula Handbook*. World Scientific.

Laevens, H., Deluyker, H., Schukken, Y. H., De Meulemeester, L., Vandermeersch, R., De Meulenaere, E., and De Kruif, A. (1997). Influence of parity and stage of lactation on the somatic cell count in bacteriologically negative dairy cows. *Journal of Dairy Science*, 80:3219–3226.

Laurent, S., Rombouts, J. V., and Violante, F. (2013). On loss functions and ranking forecasting performances of multivariate volatility models. *Journal of Econometrics*, 173(1):1–10.

Lewandowski, D., Kurowicka, D., and Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, 100(9):1989–2001.

Loaiza Maya, R. A., Gomez-Gonzalez, J. E., and Melo Velandia, L. F. (2015). Latin american exchange rate dependencies: A regular vine copula approach. *Contemporary Economic Policy*, 33(3):535–549.

Mantel, N., Bohidar, N. R., and Ciminera, J. L. (1977). Mantel-Haenszel analyses of litter-matched time-to-response data, with modifications for recovery of interlitter information. *Cancer Research*, 37:3863–3868.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

Massonnet, G., Janssen, P., and Duchateau, L. (2009). Modeling udder data using copula models for quadruples. *Journal of Statistical Planning and Inference*, 139:3865–3877.

Meyer, R. and Romeo, J. S. (2015). Bayesian semiparametric analysis of recurrent failure time data using copulas. *Biometrical Journal*, 57(6):982–1001.

Min, A. and Czado, C. (2010). Bayesian inference for multivariate copulas using pair-copula constructions. *Journal of Financial Econometrics*, 8(4):511–546.

Möller, A., Spazzini, L., Kraus, D., Nagler, T., and Czado, C. (2018). Vine copula based post-processing of ensemble forecasts for temperature. *arXiv preprint arXiv:1811.02255*.

Morales Napoles, O., Cooke, R. M., and Kurowicka, D. (2010). About the number of vines and regular vines on n nodes.

Morgan, J. P. (1996). Riskmetrics technical document.

Nagler, T. (2018). Asymptotic analysis of the jittering kernel density estimator. *Mathematical Methods of Statistics*, 27(1):32–46.

Nagler, T. and Czado, C. (2016). Evading the curse of dimensionality in nonparametric density estimation with simplified vine copulas. *Journal of Multivariate Analysis*, 151:69–89.

Nagler, T., Schellhase, C., and Czado, C. (2017). Nonparametric estimation of simplified vine copula models: comparison of methods. *Dependence Modeling*, 5(1):99–120.

Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer Series in Statistics. Springer-Verlag New York Inc.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1):246–256.

Poignard, B. (2017). *New approaches for high-dimensional multivariate GARCH models*. PhD thesis, PSL Research University.

Prenen, L., Braekers, R., and Duchateau, L. (2017). Extending the Archimedean copula methodology to model multivariate survival data grouped in clusters of variable size. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(2):483–505.

R Core Team (2017). Vienna (Austria): R foundation for statistical computing; 2017. *R: A Language and Environment for Statistical Computing*.

Roberts, F. and Tesman, B. (2009). *Applied Combinatorics*. Chapman and Hall/CRC.

Schallhorn, N., Kraus, D., Nagler, T., and Czado, C. (2017). D-vine quantile regression with discrete variables. *arXiv preprint arXiv:1705.08310*.

Schellhase, C. and Spanhel, F. (2018). Estimating non-simplified vine copulas using penalized splines. *Statistics and Computing*, 28(2):387–409.

Schepsmeier, U. (2016). A goodness-of-fit test for regular vine copula models. *Econometric Reviews*, pages 1–22.

Schepsmeier, U. and Stöber, J. (2014). Derivatives and fisher information of bivariate copulas. *Statistical Papers*, 55(2):525–542.

Schepsmeier, U., Stöber, J., Brechmann, E., and Gräler, B. (2017). VineCopula: Statistical inference of vine copulas. *R Package Version 1.7/r66*.

Shi, P. and Yang, L. (2018). Pair copula constructions for insurance experience rating. *Journal of the American Statistical Association*, 113(521):122–133.

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

Sklar, A. (1959). *Fonctions de répartition à n dimensions et leurs marges*. Publ. Inst. Stat. Université Paris 8.

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

Stöber, J. and Schepsmeier, U. (2013). Estimating standard errors in regular vine copula models. *Computational Statistics*, 28(6):2679–2707.

Voev, V. (2008). Dynamic modelling of large-dimensional covariance matrices. In *High Frequency Financial Econometrics*, pages 293–312. Springer.

Wei, L.-J., Lin, D. Y., and Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American statistical association*, 84(408):1065–1073.

Whittaker, J. (2009). *Graphical models in applied multivariate statistics*. Wiley Publishing.

Zhang, L. (2011). Estimating covariation: Epps effect, microstructure noise. *Journal of Econometrics*, 160(1):33–47.

Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, 100(472):1394–1411.

Appendix A

Supplementary Material to Chapter 3

A.1 Skewed generalized error distribution

The skewed generalized error distribution is specified by the location parameter μ, the scale parameter σ, the shape parameter ν and the skewness parameter ξ. Its density function is given by

$$f\left(\varepsilon|\mu,\sigma,\nu,\xi\right)=\frac{C}{\sigma}\exp\left(-\frac{\left|\varepsilon-\mu+\delta\sigma\right|^{\nu}}{\left[1-\operatorname{sign}\left(\varepsilon-\mu+\delta\sigma\right)\xi\right]^{\nu}\theta^{\nu}\sigma^{\nu}}\right)$$

with

$$\begin{aligned}
C &= \frac{\nu}{2\theta}\Gamma\left(\frac{1}{\nu}\right)^{-1},\\
\theta &= \Gamma\left(\frac{1}{\nu}\right)^{1/2}\Gamma\left(\frac{3}{\nu}\right)^{-1/2}S\left(\xi\right)^{-1},\\
\delta &= 2\xi AS\left(\xi\right)^{-1},\\
S\left(\xi\right) &= \sqrt{1+3\xi^{2}-4A^{2}\xi^{2}},\\
A &= \Gamma\left(\frac{2}{\nu}\right)\Gamma\left(\frac{1}{\nu}\right)^{-1/2}\Gamma\left(\frac{3}{\nu}\right)^{-1/2}.
\end{aligned}$$

For the parameter specification $\nu=2$ and $\xi=0$ the normal distribution is obtained.

A.2 Additional results for the empirical study

Table A.1: RMSE with respect to the complete out-of-sample forecasting horizon (1632 days) for the model components in the Cholesky decomposition based model. The set of superior models according to the MCS approach at a confidence level of 10% is highlighted in gray. The lowest RMSE is highlighted in bold.

	mean	HAR	HN	HSGED	ARFIMA	AN	ASGED
AXP,AXP	0.5215	0.2355	0.2358	0.2359	**0.2334**	0.2336	0.2340
AXP,C	0.6789	0.3730	0.3706	0.3769	0.3698	**0.3684**	0.3750
C,C	0.4579	0.2069	0.2076	0.2094	**0.2063**	0.2070	0.2081
AXP,GE	0.4142	0.2648	0.2646	0.2682	0.2618	**0.2616**	0.2659
C,GE	0.2431	0.1725	0.1726	0.1735	**0.1714**	0.1718	0.1727
GE,GE	0.3676	0.2111	0.2108	0.2112	0.2102	0.2099	**0.2096**
AXP,HD	0.4624	0.2976	0.2998	0.3023	**0.2944**	0.2972	0.2997
C,HD	0.2732	0.2160	0.2176	0.2165	0.2155	0.2163	**0.2153**
GE,HD	0.2352	0.1906	0.1913	0.1925	**0.1900**	0.1904	0.1917
HD,HD	0.3516	0.2165	0.2165	0.2169	0.2158	0.2157	**0.2156**
AXP,IBM	0.3102	0.2190	0.2180	0.2202	0.2174	**0.2166**	0.2189
C,IBM	0.1929	0.1529	0.1539	0.1534	**0.1527**	0.1533	0.1532
GE,IBM	0.1768	0.1401	0.1404	0.1409	**0.1399**	0.1401	0.1404
HD,IBM	0.1321	0.1260	0.1257	0.1254	0.1248	0.1246	**0.1246**
IBM,IBM	0.3390	0.2021	**0.2020**	0.2021	**0.2020**	0.2022	0.2026
AXP,JPM	0.6646	0.3918	0.3895	0.3946	0.3894	**0.3879**	0.3931
C,JPM	0.4075	0.2777	0.2804	0.2803	**0.2767**	0.2789	0.2784
GE,JPM	0.1981	0.1758	0.1757	0.1756	0.1745	**0.1743**	0.1744
HD,JPM	0.1619	0.1552	0.1554	0.1556	**0.1550**	0.1551	0.1556
IBM,JPM	0.1559	0.1498	0.1497	0.1500	**0.1496**	0.1498	0.1500
JPM,JPM	0.4461	0.2114	0.2112	0.2123	**0.2107**	**0.2107**	0.2116

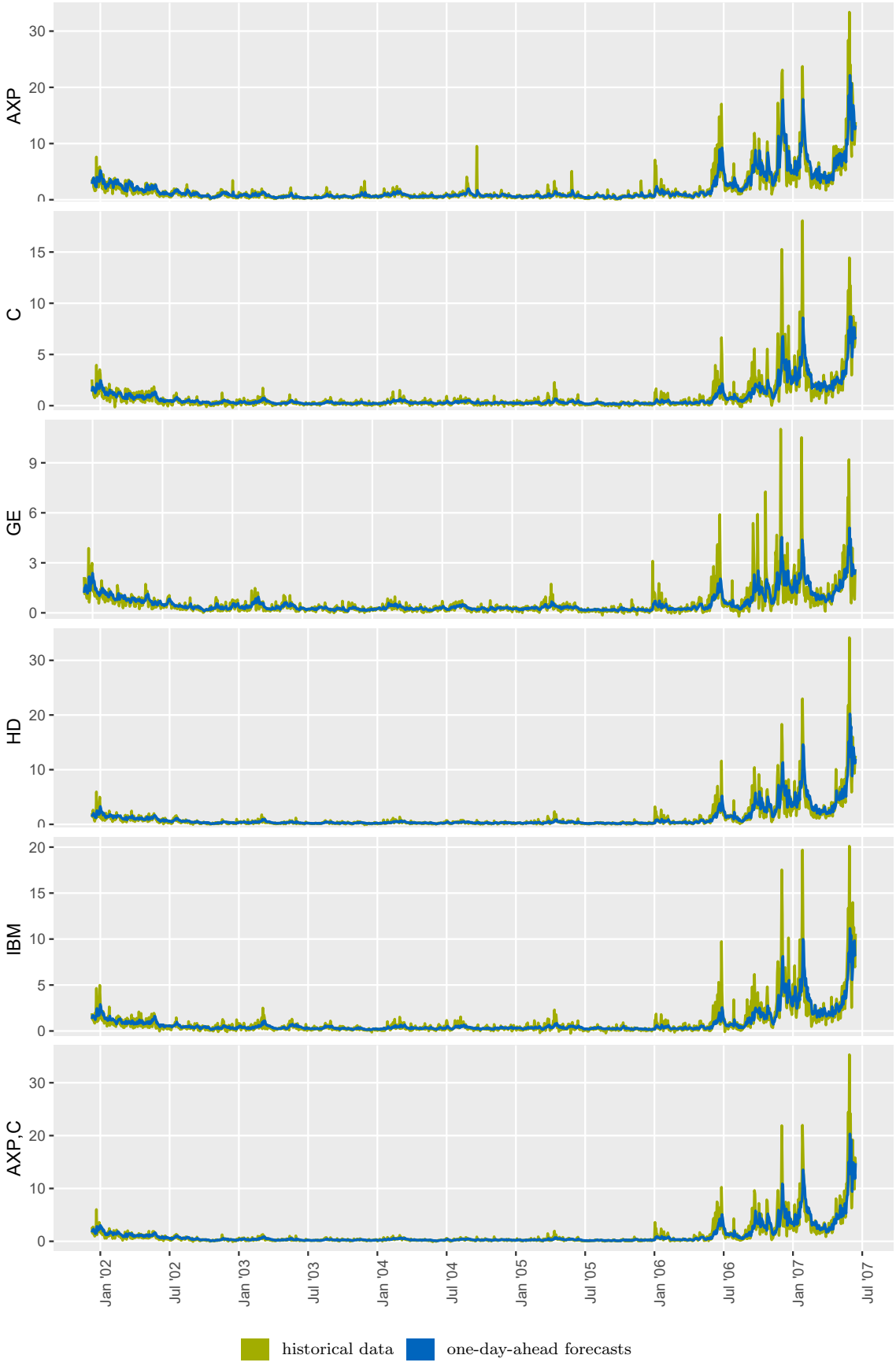


Figure A.1: (part 1/3) Daily realized variance time-series and daily realized covariance time-series together with the time-series of the corresponding daily forecasts based on the partial correlation vine data transformation approach.

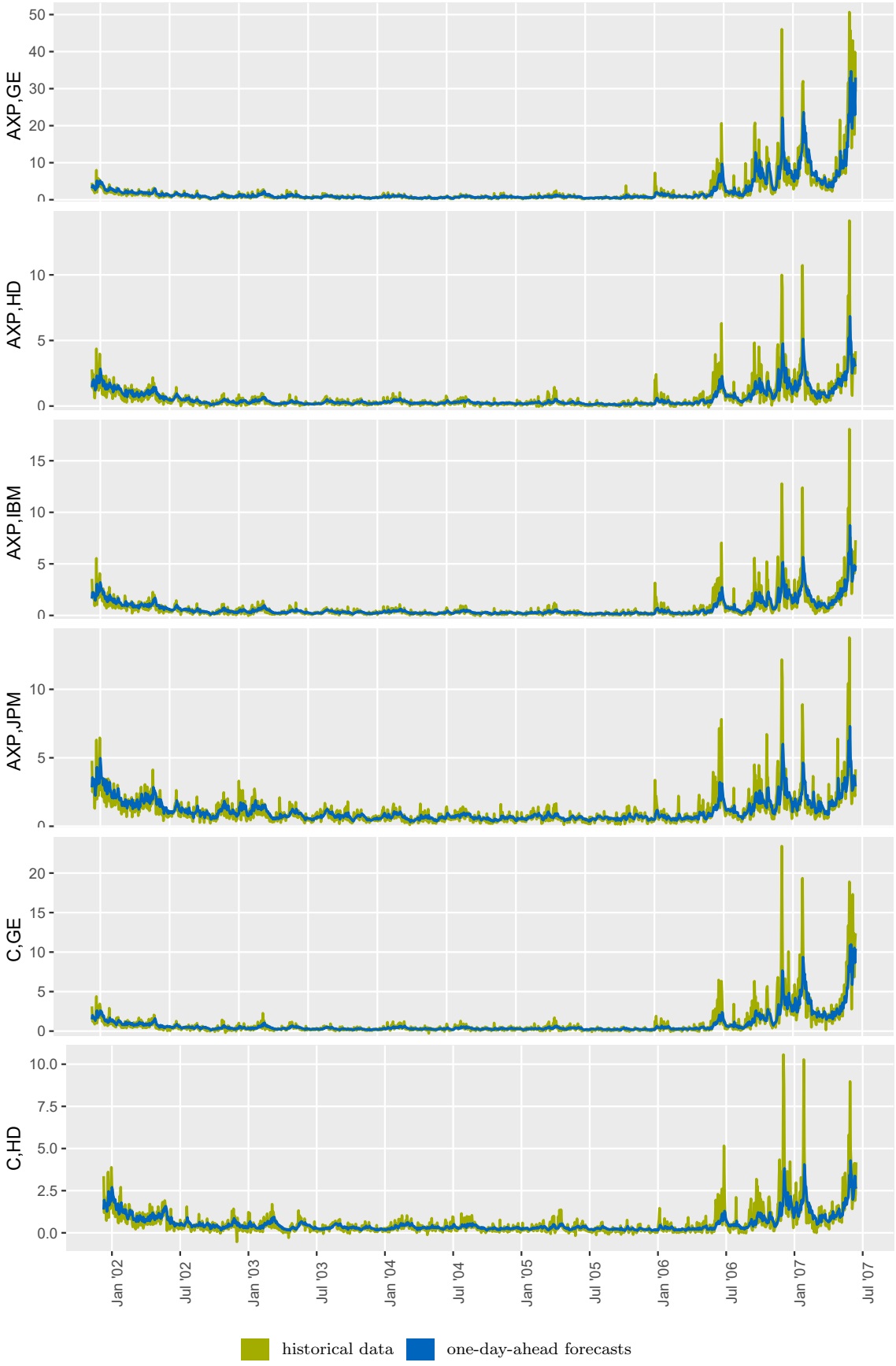


Figure A.1: (part 2/3) Daily realized variance time-series and daily realized covariance time-series together with the time-series of the corresponding daily forecasts based on the partial correlation vine data transformation approach.

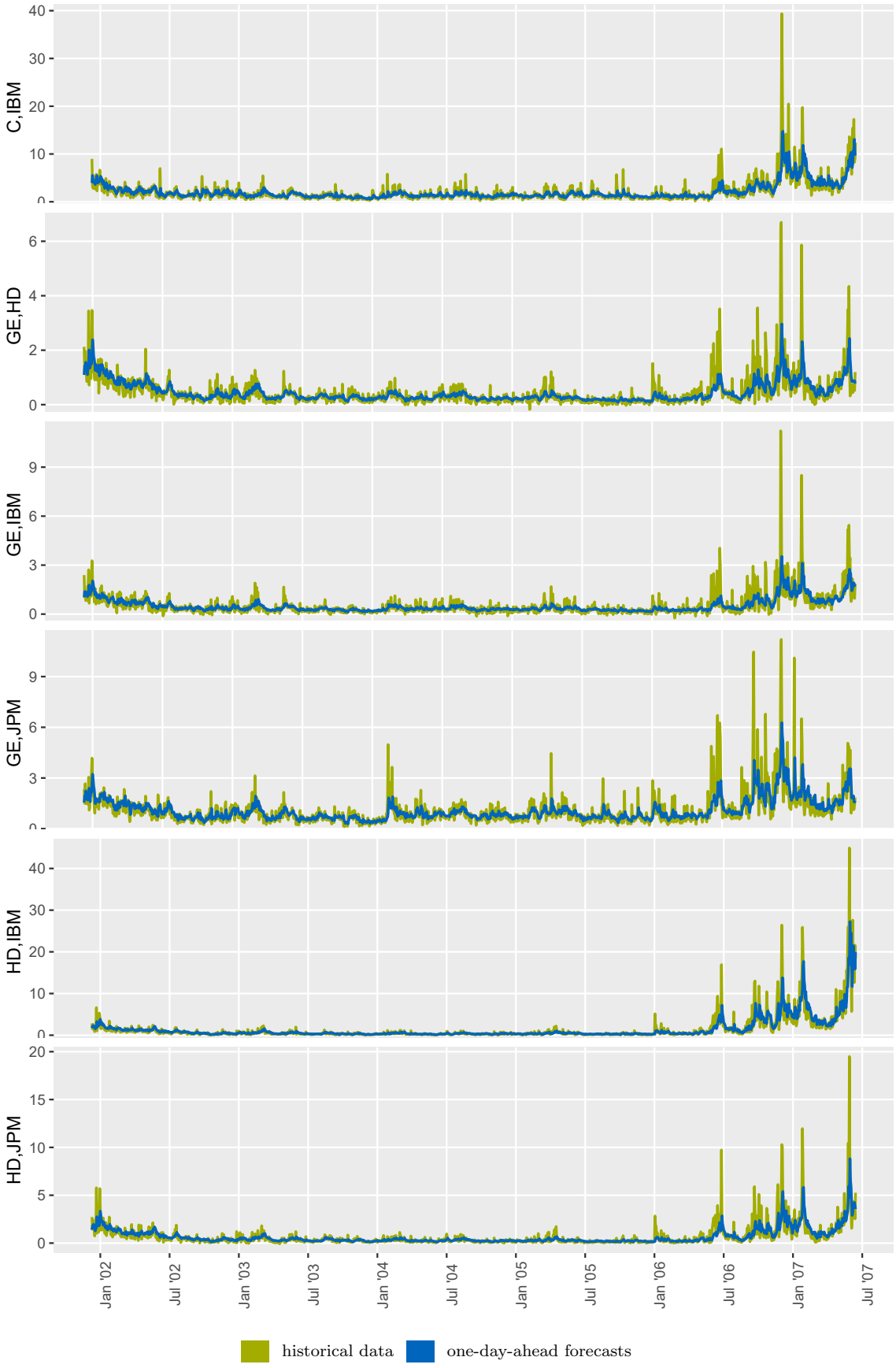


Figure A.1: (part 3/3) Daily realized variance time-series and daily realized covariance time-series together with the time-series of the corresponding daily forecasts based on the partial correlation vine data transformation approach.

Appendix B

Supplementary Material to Chapter 4

B.1 Partial derivatives of R-vine copulas

B.1.1 Partial derivatives in four dimensions

In this section, the partial derivatives for all possible four-dimensional R-vine copulas – 24 in total – are provided. Detailed proofs for the presented expressions are in Barthel (2015, Chapter 3). Recall that in four dimensions there are only D-vine and C-vine structures.

Derivation for an underlying D-vine structure

In four dimensions, there are 12 different D-vine structures. Without loss of generality, we present in Theorem B.1 and Corollary B.2 the partial derivatives of the copula $\mathbb{C}$ assuming the variable order $1-2-3-4$, i.e. the copula density $\mathbb{c}$ can be expressed in terms of pair-copula components as follows

$$\begin{aligned}\mathbb{c}\left(u_1,u_2,u_3,u_4\right) = {} & \mathbb{c}_{1,2}\left(u_1,u_2\right)\mathbb{c}_{2,3}\left(u_2,u_3\right)\mathbb{c}_{3,4}\left(u_3,u_4\right)\\ & \times \mathbb{c}_{1,3;2}\{\mathbb{C}_{1|2}\left(u_1|u_2\right),\mathbb{C}_{3|2}\left(u_3|u_2\right)\}\mathbb{c}_{2,4;3}\{\mathbb{C}_{2|3}\left(u_2|u_3\right),\mathbb{C}_{4|3}\left(u_4|u_3\right)\}\\ & \times \mathbb{c}_{1,4;2,3}\{\mathbb{C}_{1|2,3}\left(u_1|u_2,u_3\right),\mathbb{C}_{4|2,3}\left(u_4|u_2,u_3\right)\}. \end{aligned} \tag{B.1}$$

Note that the partial derivatives for all other variable orders are obtained by index permutation.

Theorem B.1. *For the copula density* (B.1) *the following holds:*

1. $$\begin{aligned}&\mathbb{C}\left(u_1,u_2,u_3,u_4\right)\\ &\quad = \int_0^{u_2}\int_0^{u_3}\mathbb{c}_{2,3}\left(v_2,v_3\right)\mathbb{C}_{1,4;2,3}\{\mathbb{C}_{1|2,3}\left(u_1|v_2,v_3\right),\mathbb{C}_{4|2,3}\left(u_4|v_2,v_3\right)\}\mathrm{d}v_3\mathrm{d}v_2\end{aligned}$$

2.(a) $$\begin{aligned}&\frac{\partial\mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1}\\ &\quad = \int_0^{u_2}\int_0^{u_3}\mathbb{c}_{1,2}\left(u_1,v_2\right)\mathbb{c}_{2,3}\left(v_2,v_3\right)\mathbb{c}_{1,3;2}\{\mathbb{C}_{1|2}\left(u_1|v_2\right),\mathbb{C}_{3|2}\left(v_3|v_2\right)\}\\ &\qquad \times \frac{\partial}{\partial\tilde{u}_1}\mathbb{C}_{1,4;2,3}\{\tilde{u}_1,\mathbb{C}_{4|2,3}\left(u_4|v_2,v_3\right)\}\Big|_{\tilde{u}_1=\mathbb{C}_{1|2,3}(u_1|v_2,v_3)}\mathrm{d}v_3\mathrm{d}v_2\end{aligned}$$

(b) $$\frac{\partial \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_2} = \int_0^{u_3} \mathbb{c}_{2,3}\left(u_2,v_3\right)\mathbb{C}_{1,4;2,3}\{\mathbb{C}_{1|2,3}\left(u_1|u_2,v_3\right),\mathbb{C}_{4|2,3}\left(u_4|u_2,v_3\right)\}\mathrm{d}v_3$$

(c) $$\frac{\partial \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_3} = \int_0^{u_2} \mathbb{c}_{2,3}\left(v_2,u_3\right)\mathbb{C}_{1,4;2,3}\{\mathbb{C}_{1|2,3}\left(u_1|v_2,u_3\right),\mathbb{C}_{4|2,3}\left(u_4|v_2,u_3\right)\}\mathrm{d}v_2$$

(d) $$\begin{aligned}\frac{\partial \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_4} &= \int_0^{u_2}\int_0^{u_3} \mathbb{c}_{2,3}\left(v_2,v_3\right)\mathbb{c}_{3,4}\left(v_3,u_4\right)\mathbb{c}_{2,4;3}\{\mathbb{C}_{2|3}\left(v_2|v_3\right),\mathbb{C}_{4|3}\left(u_4|v_3\right)\}\\ &\quad\times\frac{\partial}{\partial\tilde{u}_4}\mathbb{C}_{1,4;2,3}\{\mathbb{C}_{1|2,3}\left(u_1|v_2,v_3\right),\tilde{u}_4\}\Big|_{\tilde{u}_4=\mathbb{C}_{4|2,3}(u_4|v_2,v_3)}\mathrm{d}v_3\mathrm{d}v_2\end{aligned}$$

3.(a) $$\begin{aligned}\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1\partial u_2} &= \int_0^{u_3} \mathbb{c}_{1,2}\left(u_1,u_2\right)\mathbb{c}_{2,3}\left(u_2,v_3\right)\mathbb{c}_{1,3;2}\{\mathbb{C}_{1|2}\left(u_1|u_2\right),\mathbb{C}_{3|2}\left(v_3|u_2\right)\}\\ &\quad\times\frac{\partial}{\partial\tilde{u}_1}\mathbb{C}_{1,4;2,3}\{\tilde{u}_1,\mathbb{C}_{4|2,3}\left(u_4|u_2,v_3\right)\}\Big|_{\tilde{u}_1=\mathbb{C}_{1|2,3}(u_1|u_2,v_3)}\mathrm{d}v_3\end{aligned}$$

(b) $$\begin{aligned}\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1\partial u_3} &= \int_0^{u_2} \mathbb{c}_{1,2}\left(u_1,v_2\right)\mathbb{c}_{2,3}\left(v_2,u_3\right)\mathbb{c}_{1,3;2}\{\mathbb{C}_{1|2}\left(u_1|v_2\right),\mathbb{C}_{3|2}\left(u_3|v_2\right)\}\\ &\quad\times\frac{\partial}{\partial\tilde{u}_1}\mathbb{C}_{1,4;2,3}\{\tilde{u}_1,\mathbb{C}_{4|2,3}\left(u_4|v_2,u_3\right)\}\Big|_{\tilde{u}_1=\mathbb{C}_{1|2,3}(u_1|v_2,u_3)}\mathrm{d}v_2\end{aligned}$$

(c) $$\begin{aligned}\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1\partial u_4} &= \int_0^{u_2}\int_0^{u_3} \mathbb{c}_{1,2}\left(u_1,v_2\right)\mathbb{c}_{2,3}\left(v_2,v_3\right)\mathbb{c}_{3,4}\left(v_3,u_4\right)\\ &\quad\times\mathbb{c}_{1,3;2}\{\mathbb{C}_{1|2}\left(u_1|v_2\right),\mathbb{C}_{3|2}\left(v_3|v_2\right)\}\mathbb{c}_{2,4;3}\{\mathbb{C}_{2|3}\left(v_2|v_3\right),\mathbb{C}_{4|3}\left(u_4|v_3\right)\}\\ &\quad\times\mathbb{c}_{1,4;2,3}\{\mathbb{C}_{1|2,3}\left(u_1|v_2,v_3\right),\mathbb{C}_{4|2,3}\left(u_4|v_2,v_3\right)\}\mathrm{d}v_3\mathrm{d}v_2\end{aligned}$$

(d) $$\begin{aligned}\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_2\partial u_3} &= \mathbb{c}_{2,3}\left(u_2,u_3\right)\mathbb{C}_{1,4;2,3}\{\mathbb{C}_{1|2,3}\left(u_1|u_2,u_3\right),\mathbb{C}_{4|2,3}\left(u_4|u_2,u_3\right)\}\end{aligned}$$

(e) $$\begin{aligned}\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_2\partial u_4} &= \int_0^{u_3} \mathbb{c}_{2,3}\left(u_2,v_3\right)\mathbb{c}_{3,4}\left(v_3,u_4\right)\mathbb{c}_{2,4;3}\{\mathbb{C}_{2|3}\left(u_2|v_3\right),\mathbb{C}_{4|3}\left(u_4|v_3\right)\}\\ &\quad\times\frac{\partial}{\partial\tilde{u}_4}\mathbb{C}_{1,4;2,3}\{\mathbb{C}_{1|2,3}\left(u_1|u_2,v_3\right),\tilde{u}_4\}\Big|_{\tilde{u}_4=\mathbb{C}_{4|2,3}(u_4|u_2,v_3)}\mathrm{d}v_3\end{aligned}$$

(f) $$\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_3 \partial u_4}$$
$$= \int_0^{u_2} \mathbb{c}_{2,3}\left(v_2,u_3\right) \mathbb{c}_{3,4}\left(u_3,u_4\right) \mathbb{c}_{2,4;3}\{\mathbb{C}_{2|3}\left(v_2|u_3\right), \mathbb{C}_{4|3}\left(u_4|u_3\right)\}$$
$$\times \frac{\partial}{\partial \tilde{u}_4} \mathbb{C}_{1,4;2,3}\{\mathbb{C}_{1|2,3}\left(u_1|v_2,u_3\right), \tilde{u}_4\}\Big|_{\tilde{u}_4=\mathbb{C}_{4|2,3}(u_4|v_2,u_3)} \mathrm{d}v_2$$

4.(a) $$\frac{\partial^3 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1 \partial u_2 \partial u_3}$$
$$= \mathbb{c}_{1,2}\left(u_1,u_2\right) \mathbb{c}_{2,3}\left(u_2,u_3\right) \mathbb{c}_{1,3;2}\{\mathbb{C}_{1|2}\left(u_1|u_2\right), \mathbb{C}_{3|2}\left(u_3|u_2\right)\}$$
$$\times \frac{\partial}{\partial \tilde{u}_1} \mathbb{C}_{1,4;2,3}\{\tilde{u}_1, \mathbb{C}_{4|2,3}\left(u_4|u_2,u_3\right)\}\Big|_{\tilde{u}_1=\mathbb{C}_{1|2,3}(u_1|u_2,u_3)}$$

(b) $$\frac{\partial^3 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1 \partial u_2 \partial u_4}$$
$$= \int_0^{u_3} \mathbb{c}_{1,2}\left(u_1,u_2\right) \mathbb{c}_{2,3}\left(u_2,v_3\right) \mathbb{c}_{3,4}\left(v_3,u_4\right)$$
$$\times \mathbb{c}_{1,3;2}\{\mathbb{C}_{1|2}\left(u_1|u_2\right), \mathbb{C}_{3|2}\left(v_3|u_2\right)\} \mathbb{c}_{2,4;3}\{\mathbb{C}_{2|3}\left(u_2|v_3\right), \mathbb{C}_{4|3}\left(u_4|v_3\right)\}$$
$$\times \mathbb{c}_{1,4;2,3}\{\mathbb{C}_{1|2,3}\left(u_1|u_2,v_3\right), \mathbb{C}_{4|2,3}\left(u_4|u_2,v_3\right)\} \mathrm{d}v_3$$

(c) $$\frac{\partial^3 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1 \partial u_3 \partial u_4}$$
$$= \int_0^{u_2} \mathbb{c}_{1,2}\left(u_1,v_2\right) \mathbb{c}_{2,3}\left(v_2,u_3\right) \mathbb{c}_{3,4}\left(u_3,u_4\right)$$
$$\times \mathbb{c}_{1,3;2}\{\mathbb{C}_{1|2}\left(u_1|v_2\right), \mathbb{C}_{3|2}\left(u_3|v_2\right)\} \mathbb{c}_{2,4;3}\{\mathbb{C}_{2|3}\left(v_2|u_3\right), \mathbb{C}_{4|3}\left(u_4|u_3\right)\}$$
$$\times \mathbb{c}_{1,4;2,3}\{\mathbb{C}_{1|2,3}\left(u_1|v_2,u_3\right), \mathbb{C}_{4|2,3}\left(u_4|v_2,u_3\right)\} \mathrm{d}v_2$$

(d) $$\frac{\partial^3 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_2 \partial u_3 \partial u_4}$$
$$= \mathbb{c}_{2,3}\left(u_2,u_3\right) \mathbb{c}_{3,4}\left(u_3,u_4\right) \mathbb{c}_{2,4;3}\{\mathbb{C}_{2|3}\left(u_2|u_3\right), \mathbb{C}_{4|3}\left(u_4|u_3\right)\}$$
$$\times \frac{\partial}{\partial \tilde{u}_4} \mathbb{C}_{1,4;2,3}\{\mathbb{C}_{1|2,3}\left(u_1|u_2,u_3\right), \tilde{u}_4\}\Big|_{\tilde{u}_4=\mathbb{C}_{4|2,3}(u_4|u_2,u_3)}$$

5. $$\mathbb{c}\left(u_1,u_2,u_3,u_4\right)$$
$$= \mathbb{c}_{1,2}\left(u_1,u_2\right) \mathbb{c}_{2,3}\left(u_2,u_3\right) \mathbb{c}_{3,4}\left(u_3,u_4\right)$$
$$\times \mathbb{c}_{1,3;2}\{\mathbb{C}_{1|2}\left(u_1|u_2\right), \mathbb{C}_{3|2}\left(u_3|u_2\right)\} \mathbb{c}_{2,4;3}\{\mathbb{C}_{2|3}\left(u_2|u_3\right), \mathbb{C}_{4|3}\left(u_4|u_3\right)\}$$
$$\times \mathbb{c}_{1,4;2,3}\{\mathbb{C}_{1|2,3}\left(u_1|u_2,u_3\right), \mathbb{C}_{4|2,3}\left(u_4|u_2,u_3\right)\}$$

Corollary B.2. *In terms of h-functions, for the copula density* (B.1) *the following holds:*

1. $$\begin{aligned}&\mathbb{C}\left(u_1,u_2,u_3,u_4\right)\\&\quad=\int_0^{u_2}\int_0^{u_3}\mathbb{C}_{1,4;2,3}\left[h_{1|3;2}\{h_{1|2}\left(u_1|v_2\right)\big|h_{3|2}\left(v_3|v_2\right)\},h_{4|2;3}\{h_{4|3}\left(u_4|v_3\right)\big|h_{2|3}\left(v_2|v_3\right)\}\right]\\&\qquad\times\mathbb{c}_{2,3}\left(v_2,v_3\right)\mathrm{d}v_3\mathrm{d}v_2\end{aligned}$$

2.(a) $$\begin{aligned}&\frac{\partial\mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1}\\&\quad=\int_0^{u_2}\int_0^{u_3}\mathbb{c}_{1,2}\left(u_1,v_2\right)\mathbb{c}_{2,3}\left(v_2,v_3\right)\mathbb{c}_{1,3;2}\{h_{1|2}\left(u_1|v_2\right),h_{3|2}\left(v_3|v_2\right)\}\\&\qquad\times h_{4|1;2,3}\left[h_{4|2;3}\{h_{4|3}\left(u_4|v_3\right)\big|h_{2|3}\left(v_2|v_3\right)\}\middle|h_{1|3;2}\{h_{1|2}\left(u_1|v_2\right)\big|h_{3|2}\left(v_3|v_2\right)\}\right]\mathrm{d}v_3\mathrm{d}v_2\end{aligned}$$

(b) $$\begin{aligned}&\frac{\partial\mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_2}\\&\quad=\int_0^{u_3}\mathbb{C}_{1,4;2,3}\left[h_{1|3;2}\{h_{1|2}\left(u_1|u_2\right)\big|h_{3|2}\left(v_3|u_2\right)\},h_{4|2;3}\{h_{4|3}\left(u_4|v_3\right)\big|h_{2|3}\left(u_2|v_3\right)\}\right]\\&\qquad\times\mathbb{c}_{2,3}\left(u_2,v_3\right)\mathrm{d}v_3\end{aligned}$$

(c) $$\begin{aligned}&\frac{\partial\mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_3}\\&\quad=\int_0^{u_2}\mathbb{C}_{1,4;2,3}\left[h_{1|3;2}\{h_{1|2}\left(u_1|v_2\right)\big|h_{3|2}\left(u_3|v_2\right)\},h_{4|2;3}\{h_{4|3}\left(u_4|u_3\right)\big|h_{2|3}\left(v_2|u_3\right)\}\right]\\&\qquad\times\mathbb{c}_{2,3}\left(v_2,u_3\right)\mathrm{d}v_2\end{aligned}$$

(d) $$\begin{aligned}&\frac{\partial\mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_4}\\&\quad=\int_0^{u_2}\int_0^{u_3}\mathbb{c}_{2,3}\left(v_2,v_3\right)\mathbb{c}_{3,4}\left(v_3,u_4\right)\mathbb{c}_{2,4;3}\{h_{2|3}\left(v_2|v_3\right),h_{4|3}\left(u_4|v_3\right)\}\\&\qquad\times h_{1|4;2,3}\left[h_{1|3;2}\{h_{1|2}\left(u_1|v_2\right)\big|h_{3|2}\left(v_3|v_2\right)\}\middle|h_{4|2;3}\{h_{4|3}\left(u_4|v_3\right)\big|h_{2|3}\left(v_2|v_3\right)\}\right]\mathrm{d}v_3\mathrm{d}v_2\end{aligned}$$

3.(a) $$\begin{aligned}&\frac{\partial^2\mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1\partial u_2}\\&\quad=\int_0^{u_3}\mathbb{c}_{1,2}\left(u_1,u_2\right)\mathbb{c}_{2,3}\left(u_2,v_3\right)\mathbb{c}_{1,3;2}\{h_{1|2}\left(u_1|u_2\right),h_{3|2}\left(v_3|u_2\right)\}\\&\qquad\times h_{4|1;2,3}\left[h_{4|2;3}\{h_{4|3}\left(u_4|v_3\right)\big|h_{2|3}\left(u_2|v_3\right)\}\middle|h_{1|3;2}\{h_{1|2}\left(u_1|u_2\right)\big|h_{3|2}\left(v_3|u_2\right)\}\right]\mathrm{d}v_3\end{aligned}$$

(b) $$\begin{aligned}&\frac{\partial^2\mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1\partial u_3}\\&\quad=\int_0^{u_2}\mathbb{c}_{1,2}\left(u_1,v_2\right)\mathbb{c}_{2,3}\left(v_2,u_3\right)\mathbb{c}_{1,3;2}\{h_{1|2}\left(u_1|v_2\right),h_{3|2}\left(u_3|v_2\right)\}\\&\qquad\times h_{4|1;2,3}\left[h_{4|2;3}\{h_{4|3}\left(u_4|u_3\right)\big|h_{2|3}\left(v_2|u_3\right)\}\middle|h_{1|3;2}\{h_{1|2}\left(u_1|v_2\right)\big|h_{3|2}\left(u_3|v_2\right)\}\right]\mathrm{d}v_2\end{aligned}$$

(c)
$$\begin{aligned}\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1 \partial u_4} &= \int_0^{u_2}\int_0^{u_3} \mathbb{c}_{1,2}\left(u_1,v_2\right)\mathbb{c}_{2,3}\left(v_2,v_3\right)\mathbb{c}_{3,4}\left(v_3,u_4\right)\\ &\quad\times \mathbb{c}_{1,3;2}\{h_{1|2}\left(u_1|v_2\right),h_{3|2}\left(v_3|v_2\right)\}\mathbb{c}_{2,4;3}\{h_{2|3}\left(v_2|v_3\right),h_{4|3}\left(u_4|v_3\right)\}\\ &\quad\times \mathbb{c}_{1,4;2,3}\left[h_{1|3;2}\{h_{1|2}\left(u_1|v_2\right)\left|h_{3|2}\left(v_3|v_2\right)\right.\},h_{4|2;3}\{h_{4|3}\left(u_4|v_3\right)\left|h_{2|3}\left(v_2|v_3\right)\right.\}\right]\mathrm{d}v_3\mathrm{d}v_2\end{aligned}$$

(d)
$$\begin{aligned}\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_2 \partial u_3} &= \mathbb{c}_{2,3}\left(u_2,u_3\right)\mathbb{C}_{1,4;2,3}\left[h_{1|3;2}\{h_{1|2}\left(u_1|u_2\right)\left|h_{3|2}\left(u_3|u_2\right)\right.\},h_{4|2;3}\{h_{4|3}\left(u_4|u_3\right)\left|h_{2|3}\left(u_2|u_3\right)\right.\}\right]\end{aligned}$$

(e)
$$\begin{aligned}\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_2 \partial u_4} &= \int_0^{u_3} \mathbb{c}_{2,3}\left(u_2,v_3\right)\mathbb{c}_{3,4}\left(v_3,u_4\right)\mathbb{c}_{2,4;3}\{h_{2|3}\left(u_2|v_3\right),h_{4|3}\left(u_4|v_3\right)\}\\ &\quad\times h_{1|4;2,3}\left[h_{1|3;2}\{h_{1|2}\left(u_1|u_2\right)\left|h_{3|2}\left(v_3|u_2\right)\right.\}\middle|h_{4|2;3}\{h_{4|3}\left(u_4|v_3\right)\left|h_{2|3}\left(u_2|v_3\right)\right.\}\right]\mathrm{d}v_3\end{aligned}$$

(f)
$$\begin{aligned}\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_3 \partial u_4} &= \int_0^{u_2} \mathbb{c}_{2,3}\left(v_2,u_3\right)\mathbb{c}_{3,4}\left(u_3,u_4\right)\mathbb{c}_{2,4;3}\{h_{2|3}\left(v_2|u_3\right),h_{4|3}\left(u_4|u_3\right)\}\\ &\quad\times h_{1|4;2,3}\left[h_{1|3;2}\{h_{1|2}\left(u_1|v_2\right)\left|h_{3|2}\left(u_3|v_2\right)\right.\}\middle|h_{4|2;3}\{h_{4|3}\left(u_4|u_3\right)\left|h_{2|3}\left(v_2|u_3\right)\right.\}\right]\mathrm{d}v_2\end{aligned}$$

4.(a)
$$\begin{aligned}\frac{\partial^3 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1 \partial u_2 \partial u_3} &= \mathbb{c}_{1,2}\left(u_1,u_2\right)\mathbb{c}_{2,3}\left(u_2,u_3\right)\mathbb{c}_{1,3;2}\{h_{1|2}\left(u_1|u_2\right),h_{3|2}\left(u_3|u_2\right)\}\\ &\quad\times h_{4|1;2,3}\left[h_{4|2;3}\{h_{4|3}\left(u_4|u_3\right)\left|h_{2|3}\left(u_2|u_3\right)\right.\}\middle|h_{1|3;2}\{h_{1|2}\left(u_1|u_2\right)\left|h_{3|2}\left(u_3|u_2\right)\right.\}\right]\end{aligned}$$

(b)
$$\begin{aligned}\frac{\partial^3 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1 \partial u_2 \partial u_4} &= \int_0^{u_3} \mathbb{c}_{1,4;2,3}\left[h_{1|3;2}\{h_{1|2}\left(u_1|u_2\right)\left|h_{3|2}\left(v_3|u_2\right)\right.\},h_{4|2;3}\{h_{4|3}\left(u_4|v_3\right)\left|h_{2|3}\left(u_2|v_3\right)\right.\}\right]\\ &\quad\times \mathbb{c}_{1,2}\left(u_1,u_2\right)\mathbb{c}_{2,3}\left(u_2,v_3\right)\mathbb{c}_{3,4}\left(v_3,u_4\right)\mathbb{c}_{1,3;2}\{h_{1|2}\left(u_1|u_2\right),h_{3|2}\left(v_3|u_2\right)\}\\ &\quad\times \mathbb{c}_{2,4;3}\{h_{2|3}\left(u_2|v_3\right),h_{4|3}\left(u_4|v_3\right)\}\mathrm{d}v_3\end{aligned}$$

(c)
$$\begin{aligned}\frac{\partial^3 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1 \partial u_3 \partial u_4} &= \int_0^{u_2} \mathbb{c}_{1,4;2,3}\left[h_{1|3;2}\{h_{1|2}\left(u_1|v_2\right)\left|h_{3|2}\left(u_3|v_2\right)\right.\},h_{4|2;3}\{h_{4|3}\left(u_4|u_3\right)\left|h_{2|3}\left(v_2|u_3\right)\right.\}\right]\\ &\quad\times \mathbb{c}_{1,2}\left(u_1,v_2\right)\mathbb{c}_{2,3}\left(v_2,u_3\right)\mathbb{c}_{3,4}\left(u_3,u_4\right)\mathbb{c}_{1,3;2}\{h_{1|2}\left(u_1|v_2\right),h_{3|2}\left(u_3|v_2\right)\}\\ &\quad\times \mathbb{c}_{2,4;3}\{h_{2|3}\left(v_2|u_3\right),h_{4|3}\left(u_4|u_3\right)\}\mathrm{d}v_2\end{aligned}$$

(d) $$\frac{\partial^3 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_2 \partial u_3 \partial u_4}$$
$$= \mathbb{c}_{2,3}\left(u_2,u_3\right)\mathbb{c}_{3,4}\left(u_3,u_4\right)\mathbb{c}_{2,4;3}\{h_{2|3}\left(u_2|u_3\right),h_{4|3}\left(u_4|u_3\right)\}$$
$$\times h_{1|4;2,3}\left[h_{1|3;2}\{h_{1|2}\left(u_1|u_2\right)\left|h_{3|2}\left(u_3|u_2\right)\right.\}\middle|h_{4|2;3}\{h_{4|3}\left(u_4|u_3\right)\left|h_{2|3}\left(u_2|u_3\right)\right.\}\right]$$

5. $$\mathbb{c}\left(u_1,u_2,u_3,u_4\right)$$
$$= \mathbb{c}_{1,2}\left(u_1,u_2\right)\mathbb{c}_{2,3}\left(u_2,u_3\right)\mathbb{c}_{3,4}\left(u_3,u_4\right)$$
$$\times \mathbb{c}_{1,3;2}\{h_{1|2}\left(u_1|u_2\right),h_{3|2}\left(u_3|u_2\right)\}\mathbb{c}_{2,4;3}\{h_{2|3}\left(u_2|u_3\right),h_{4|3}\left(u_4|u_3\right)\}$$
$$\times \mathbb{c}_{1,4;2,3}\left[h_{1|3;2}\{h_{1|2}\left(u_1|u_2\right)\left|h_{3|2}\left(u_3|u_2\right)\right.\},h_{4|2;3}\{h_{4|3}\left(u_4|u_3\right)\left|h_{2|3}\left(u_2|u_3\right)\right.\}\right]$$

Derivation for an underlying C-vine structure

In total, there are 12 different C-vine structures. Without loss of generality, we present in Theorem B.3 and Corollary B.4 the partial derivatives of the copula $\mathbb{C}$ assuming that the copula density $\mathbb{c}$ can be expressed in terms of pair-copula components as follows

$$\begin{aligned}\mathbb{c}\left(u_1,u_2,u_3,u_4\right) = {} & \mathbb{c}_{1,2}\left(u_1,u_2\right)\mathbb{c}_{1,3}\left(u_1,u_3\right)\mathbb{c}_{1,4}\left(u_1,u_4\right)\\ & \times \mathbb{c}_{2,3;1}\{\mathbb{C}_{2|1}\left(u_2|u_1\right),\mathbb{C}_{3|1}\left(u_3|u_1\right)\}\mathbb{c}_{2,4;1}\{\mathbb{C}_{2|1}\left(u_2|u_1\right),\mathbb{C}_{4|1}\left(u_4|u_1\right)\}\\ & \times \mathbb{c}_{3,4;1,2}\{\mathbb{C}_{3|1,2}\left(u_3|u_1,u_2\right),\mathbb{C}_{4|1,2}\left(u_4|u_1,u_2\right)\}. \end{aligned} \tag{B.2}$$

Note that the partial derivatives for all other variable orders are obtained by index permutation.

Theorem B.3. *For the copula density* (B.2) *the following holds:*

1. $$\mathbb{C}\left(u_1,u_2,u_3,u_4\right)$$
$$= \int_0^{u_1}\int_0^{u_2} \mathbb{c}_{1,2}\left(v_1,v_2\right)\mathbb{C}_{3,4;1,2}\{\mathbb{C}_{3|1,2}\left(u_3|v_1,v_2\right),\mathbb{C}_{4|1,2}\left(u_4|v_1,v_2\right)\}\mathrm{d}v_2\mathrm{d}v_1$$

2.(a) $$\frac{\partial \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1}$$
$$= \int_0^{u_2} \mathbb{c}_{1,2}\left(u_1,v_2\right)\mathbb{C}_{3,4;1,2}\{\mathbb{C}_{3|1,2}\left(u_3|u_1,v_2\right),\mathbb{C}_{4|1,2}\left(u_4|u_1,v_2\right)\}\mathrm{d}v_2$$

(b) $$\frac{\partial \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_2}$$
$$= \int_0^{u_1} \mathbb{c}_{1,2}\left(v_1,u_2\right)\mathbb{C}_{3,4;1,2}\{\mathbb{C}_{3|1,2}\left(u_3|v_1,u_2\right),\mathbb{C}_{4|1,2}\left(u_4|v_1,u_2\right)\}\mathrm{d}v_1$$

(c) $$\frac{\partial \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_3}$$
$$= \int_0^{u_1}\int_0^{u_2} \mathbb{c}_{1,2}\left(v_1,v_2\right)\mathbb{c}_{1,3}\left(v_1,u_3\right)\mathbb{c}_{2,3;1}\{\mathbb{C}_{2|1}\left(v_2|v_1\right),\mathbb{C}_{3|1}\left(u_3|v_1\right)\}$$
$$\times \frac{\partial}{\partial \tilde{u}_3}\mathbb{C}_{3,4;1,2}\{\tilde{u}_3,\mathbb{C}_{4|12}\left(u_4|v_1,v_2\right)\}\Big|_{\tilde{u}_3=\mathbb{C}_{3|1,2}(u_3|v_1,v_2)}\mathrm{d}v_2\mathrm{d}v_1$$

(d) $$\frac{\partial \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_4} = \int_0^{u_1}\int_0^{u_2} \mathbb{c}_{1,2}\left(v_1,v_2\right)\mathbb{c}_{1,4}\left(v_1,u_4\right)\mathbb{c}_{2,4;1}\{\mathbb{C}_{2|1}\left(v_2|v_1\right),\mathbb{C}_{4|1}\left(u_4|v_1\right)\} \times \frac{\partial}{\partial \tilde{u}_4}\mathbb{C}_{3,4;1,2}\{\mathbb{C}_{3|1,2}\left(u_3|v_1,v_2\right),\tilde{u}_4\}\Big|_{\tilde{u}_4=\mathbb{C}_{4|1,2}(u_4|v_1,v_2)} \mathrm{d}v_2\mathrm{d}v_1$$

3.(a) $$\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1\partial u_2} = \mathbb{c}_{1,2}\left(u_1,u_2\right)\mathbb{C}_{3,4;1,2}\{\mathbb{C}_{3|1,2}\left(u_3|u_1,u_2\right),\mathbb{C}_{4|1,2}\left(u_4|u_1,u_2\right)\}$$

(b) $$\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1\partial u_3} = \int_0^{u_2} \mathbb{c}_{1,2}\left(u_1,v_2\right)\mathbb{c}_{1,3}\left(u_1,u_3\right)\mathbb{c}_{2,3;1}\{\mathbb{C}_{2|1}\left(v_2|u_1\right),\mathbb{C}_{3|1}\left(u_3|u_1\right)\} \times \frac{\partial}{\partial \tilde{u}_3}\mathbb{C}_{3,4;1,2}\{\tilde{u}_3,\mathbb{C}_{4|1,2}\left(u_4|u_1,v_2\right)\}\Big|_{\tilde{u}_3=\mathbb{C}_{3|1,2}(u_3|u_1,v_2)} \mathrm{d}v_2$$

(c) $$\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1\partial u_4} = \int_0^{u_2} \mathbb{c}_{1,2}\left(u_1,v_2\right)\mathbb{c}_{1,4}\left(u_1,u_4\right)\mathbb{c}_{2,4;1}\{\mathbb{C}_{2|1}\left(v_2|u_1\right),\mathbb{C}_{4|1}\left(u_4|u_1\right)\} \times \frac{\partial}{\partial \tilde{u}_4}\mathbb{C}_{3,4;1,2}\{\mathbb{C}_{3|1,2}\left(u_3|u_1,v_2\right),\tilde{u}_4\}\Big|_{\tilde{u}_4=\mathbb{C}_{4|1,2}(u_4|u_1,v_2)} \mathrm{d}v_2$$

(d) $$\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_2\partial u_3} = \int_0^{u_1} \mathbb{c}_{1,2}\left(v_1,u_2\right)\mathbb{c}_{1,3}\left(v_1,u_3\right)\mathbb{c}_{2,3;1}\{\mathbb{C}_{2|1}\left(u_2|v_1\right),\mathbb{C}_{3|1}\left(u_3|v_1\right)\} \times \frac{\partial}{\partial \tilde{u}_3}\mathbb{C}_{3,4;1,2}\{\tilde{u}_3,\mathbb{C}_{4|1,2}\left(u_4|v_1,u_2\right)\}\Big|_{\tilde{u}_3=\mathbb{C}_{3|1,2}(u_3|v_1,u_2)} \mathrm{d}v_1$$

(e) $$\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_2\partial u_4} = \int_0^{u_1} \mathbb{c}_{1,2}\left(v_1,u_2\right)\mathbb{c}_{1,4}\left(v_1,u_4\right)\mathbb{c}_{2,4;1}\{\mathbb{C}_{2|1}\left(u_2|v_1\right),\mathbb{C}_{4|1}\left(u_4|v_1\right)\} \times \frac{\partial}{\partial \tilde{u}_4}\mathbb{C}_{3,4;1,2}\{\mathbb{C}_{3|1,2}\left(u_3|v_1,u_2\right\},\tilde{u}_4)\Big|_{\tilde{u}_4=\mathbb{C}_{4|1,2}(u_4|v_1,u_2)} \mathrm{d}v_1$$

(f) $$\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_3\partial u_4} = \int_0^{u_1}\int_0^{u_2} \mathbb{c}_{1,2}\left(v_1,v_2\right)\mathbb{c}_{1,3}\left(v_1,u_3\right)\mathbb{c}_{1,4}\left(v_1,u_4\right) \times \mathbb{c}_{2,3;1}\{\mathbb{C}_{2|1}\left(v_2|v_1\right),\mathbb{C}_{3|1}\left(u_3|v_1\right)\}\mathbb{c}_{2,4;1}\{\mathbb{C}_{2|1}\left(v_2|v_1\right),\mathbb{C}_{4|1}\left(u_4|v_1\right)\} \times \mathbb{c}_{3,4;1,2}\{\mathbb{C}_{3|1,2}\left(u_3|v_1,v_2\right),\mathbb{C}_{4|1,2}\left(u_4|v_1,v_2\right)\}\mathrm{d}v_2\mathrm{d}v_1$$

4.(a) $$\frac{\partial^3 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1 \partial u_2 \partial u_3}$$
$$= \mathbb{c}_{1,2}\left(u_1,u_2\right)\mathbb{c}_{1,3}\left(u_1,u_3\right)\mathbb{c}_{2,3;1}\{\mathbb{C}_{2|1}\left(u_2|u_1\right),\mathbb{C}_{3|1}\left(u_3|u_1\right)\}$$
$$\times \frac{\partial}{\partial \tilde{u}_3}\mathbb{C}_{3,4;1,2}\{\tilde{u}_3,\mathbb{C}_{4|1,2}\left(u_4|u_1,u_2\right)\}\Big|_{\tilde{u}_3=\mathbb{C}_{3|1,2}(u_3|u_1,u_2)}$$

(b) $$\frac{\partial^3 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1 \partial u_2 \partial u_4}$$
$$= \mathbb{c}_{1,2}\left(u_1,u_2\right)\mathbb{c}_{1,4}\left(u_1,u_4\right)\mathbb{c}_{2,4;1}\{\mathbb{C}_{2|1}\left(u_2|u_1\right),\mathbb{C}_{4|1}\left(u_4|u_1\right)\}$$
$$\times \frac{\partial}{\partial \tilde{u}_4}\mathbb{C}_{3,4;1,2}\{\mathbb{C}_{3|1,2}\left(u_3|u_1,u_2\right),\tilde{u}_4\}\Big|_{\tilde{u}_4=\mathbb{C}_{4|1,2}(u_4|u_1,u_2)}$$

(c) $$\frac{\partial^3 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1 \partial u_3 \partial u_4}$$
$$= \int_0^{u_2} \mathbb{c}_{1,2}\left(u_1,v_2\right)\mathbb{c}_{1,3}\left(u_1,u_3\right)\mathbb{c}_{1,4}\left(u_1,u_4\right)$$
$$\times \mathbb{c}_{2,3;1}\{\mathbb{C}_{2|1}\left(v_2|u_1\right),\mathbb{C}_{3|1}\left(u_3|u_1\right)\}\mathbb{c}_{2,4;1}\{\mathbb{C}_{2|1}\left(v_2|u_1\right),\mathbb{C}_{4|1}\left(u_4|u_1\right)\}$$
$$\times \mathbb{c}_{3,4;1,2}\{\mathbb{C}_{3|1,2}\left(u_3|u_1,v_2\right),\mathbb{C}_{4|1,2}\left(u_4|u_1,v_2\right)\}\mathrm{d}v_2$$

(d) $$\frac{\partial^3 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_2 \partial u_3 \partial u_4}$$
$$= \int_0^{u_1} \mathbb{c}_{1,2}\left(v_1,u_2\right)\mathbb{c}_{1,3}\left(v_1,u_3\right)\mathbb{c}_{1,4}\left(v_1,u_4\right)$$
$$\times \mathbb{c}_{2,3;1}\{\mathbb{C}_{2|1}\left(u_2|v_1\right),\mathbb{C}_{3|1}\left(u_3|v_1\right)\}\mathbb{c}_{2,4;1}\{\mathbb{C}_{2|1}\left(u_2|v_1\right),\mathbb{C}_{4|1}\left(u_4|v_1\right)\}$$
$$\times \mathbb{c}_{3,4;1,2}\{\mathbb{C}_{3|1,2}\left(u_3|v_1,u_2\right),\mathbb{C}_{4|1,2}\left(u_4|v_1,u_2\right)\}\mathrm{d}v_1$$

5. $$\mathbb{c}\left(u_1,u_2,u_3,u_4\right)$$
$$= \mathbb{c}_{1,2}\left(u_1,u_2\right)\mathbb{c}_{1,3}\left(u_1,u_3\right)\mathbb{c}_{1,4}\left(u_1,u_4\right)$$
$$\times \mathbb{c}_{2,3;1}\{\mathbb{C}_{2|1}\left(u_2|u_1\right),\mathbb{C}_{3|1}\left(u_3|u_1\right)\}\mathbb{c}_{2,4;1}\{\mathbb{C}_{2|1}\left(u_2|u_1\right),\mathbb{C}_{4|1}\left(u_4|u_1\right)\}$$
$$\times \mathbb{c}_{3,4;1,2}\{\mathbb{C}_{3|1,2}\left(u_3|u_1,u_2\right),\mathbb{C}_{4|1,2}\left(u_4|u_1,u_2\right)\}$$

Corollary B.4. *In terms of h-functions, for the copula density* (B.2) *the following holds:*

1. $$\mathbb{C}\left(u_1,u_2,u_3,u_4\right)$$
$$= \int_0^{u_1}\int_0^{u_2} \mathbb{C}_{3,4;1,2}\left[h_{3|2;1}\{h_{3|1}\left(u_3|v_1\right)\big|h_{2|1}\left(v_2|v_1\right)\},h_{4|2;1}\{h_{4|1}\left(u_4|v_1\right)\big|h_{2|1}\left(v_2|v_1\right)\}\right]$$
$$\times \mathbb{c}_{1,2}\left(v_1,v_2\right)\mathrm{d}v_2\mathrm{d}v_1$$

2.(a) $$\frac{\partial \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1}$$
$$= \int_0^{u_2} \mathbb{C}_{3,4;1,2}\left[h_{3|2;1}\{h_{3|1}\left(u_3|u_1\right)\big|h_{2|1}\left(v_2|u_1\right)\},h_{4|2;1}\{h_{4|1}\left(u_4|u_1\right)\big|h_{2|1}\left(v_2|u_1\right)\}\right]$$
$$\times \mathbb{c}_{1,2}\left(u_1,v_2\right)\mathrm{d}v_2$$

(b) $$\frac{\partial \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_2} = \int_0^{u_1} \mathbb{C}_{3,4;1,2}\left[h_{3|2;1}\{h_{3|1}\left(u_3|v_1\right)\left|h_{2|1}\left(u_2|v_1\right)\right.\}, h_{4|2;1}\{h_{4|1}\left(u_4|v_1\right)\left|h_{2|1}\left(u_2|v_1\right)\right.\}\right] \times \mathbb{c}_{1,2}\left(v_1,u_2\right)\mathrm{d}v_1$$

(c) $$\frac{\partial \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_3} = \int_0^{u_1}\int_0^{u_2} \mathbb{c}_{1,2}\left(v_1,v_2\right)\mathbb{c}_{1,3}\left(v_1,u_3\right)\mathbb{c}_{2,3;1}\{h_{2|1}\left(v_2|v_1\right),h_{3|1}\left(u_3|v_1\right)\} \times h_{4|3;1,2}\left[h_{4|2;1}\{h_{4|1}\left(u_4|v_1\right)\left|h_{2|1}\left(v_2|v_1\right)\right.\}\middle|h_{3|2;1}\{h_{3|1}\left(u_3|v_1\right)\left|h_{2|1}\left(v_2|v_1\right)\right.\}\right]\mathrm{d}v_2\mathrm{d}v_1$$

(d) $$\frac{\partial \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_4} = \int_0^{u_1}\int_0^{u_2} \mathbb{c}_{1,2}\left(v_1,v_2\right)\mathbb{c}_{1,4}\left(v_1,u_4\right)\mathbb{c}_{2,4;1}\left(h_{2|1}\left(v_2|v_1\right),h_{4|1}\left(u_4|v_1\right)\right) \times h_{3|4;1,2}\left[h_{3|2;1}\{h_{3|1}\left(u_3|v_1\right)\left|h_{2|1}\left(v_2|v_1\right)\right.\}\middle|h_{4|2;1}\{h_{4|1}\left(u_4|v_1\right)\left|h_{2|1}\left(v_2|v_1\right)\right.\}\right]\mathrm{d}v_2\mathrm{d}v_1$$

3.(a) $$\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1\partial u_2} = \mathbb{c}_{1,2}\left(u_1,u_2\right)\mathbb{C}_{3,4;1,2}\left[h_{3|2;1}\{h_{3|1}\left(u_3|u_1\right)\left|h_{2|1}\left(u_2|u_1\right)\right.\}, h_{4|2;1}\{h_{4|1}\left(u_4|u_1\right)\left|h_{2|1}\left(u_2|u_1\right)\right.\}\right]$$

(b) $$\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1\partial u_3} = \int_0^{u_2} \mathbb{c}_{1,2}\left(u_1,v_2\right)\mathbb{c}_{1,3}\left(u_1,u_3\right)\mathbb{c}_{2,3;1}\{h_{2|1}\left(v_2|u_1\right),h_{3|1}\left(u_3|u_1\right)\} \times h_{4|3;1,2}\left[h_{4|2;1}\{h_{4|1}\left(u_4|u_1\right)\left|h_{2|1}\left(v_2|u_1\right)\right.\}\middle|h_{3|2;1}\{h_{3|1}\left(u_3|u_1\right)\left|h_{2|1}\left(v_2|u_1\right)\right.\}\right]\mathrm{d}v_2$$

(c) $$\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1\partial u_4} = \int_0^{u_2} \mathbb{c}_{1,2}\left(u_1,v_2\right)\mathbb{c}_{1,4}\left(u_1,u_4\right)\mathbb{c}_{2,4;1}\{h_{2|1}\left(v_2|u_1\right),h_{4|1}\left(u_4|u_1\right)\} \times h_{3|4;1,2}\left[h_{3|2;1}\{h_{3|1}\left(u_3|u_1\right)\left|h_{2|1}\left(v_2|u_1\right)\right.\}\middle|h_{4|2;1}\{h_{4|1}\left(u_4|u_1\right)\left|h_{2|1}\left(v_2|u_1\right)\right.\}\right]\mathrm{d}v_2$$

(d) $$\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_2\partial u_3} = \int_0^{u_1} \mathbb{c}_{1,2}\left(v_1,u_2\right)\mathbb{c}_{1,3}\left(v_1,u_3\right)\mathbb{c}_{2,3;1}\{h_{2|1}\left(u_2|v_1\right),h_{3|1}\left(u_3|v_1\right)\} \times h_{4|3;1,2}\left[h_{4|2;1}\{h_{4|1}\left(u_4|v_1\right)\left|h_{2|1}\left(u_2|v_1\right)\right.\}\middle|h_{3|2;1}\{h_{3|1}\left(u_3|v_1\right)\left|h_{2|1}\left(u_2|v_1\right)\right.\}\right]\mathrm{d}v_1$$

(e) $$\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_2\partial u_4} = \int_0^{u_1} \mathbb{c}_{1,2}\left(v_1,u_2\right)\mathbb{c}_{1,4}\left(v_1,u_4\right)\mathbb{c}_{2,4;1}\{h_{2|1}\left(u_2|v_1\right),h_{4|1}\left(u_4|v_1\right)\} \times h_{3|4;1,2}\left[h_{3|2;1}\{h_{3|1}\left(u_3|v_1\right)\left|h_{2|1}\left(u_2|v_1\right)\right.\}\middle|h_{4|2;1}\{h_{4|1}\left(u_4|v_1\right)\left|h_{2|1}\left(u_2|v_1\right)\right.\}\right]\mathrm{d}v_1$$

(f) $$\frac{\partial^2 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_3 \partial u_4}$$
$$= \int_0^{u_1}\int_0^{u_2} \mathbb{c}_{3,4;1,2}\left[h_{3|2;1}\{h_{3|1}\left(u_3|v_1\right)\big|h_{2|1}\left(v_2|v_1\right)\}, h_{4|2;1}\{h_{4|1}\left(u_4|v_1\right)\big|h_{2|1}\left(v_2|v_1\right)\}\right]$$
$$\mathbb{c}_{1,2}\left(v_1,v_2\right)\mathbb{c}_{1,3}\left(v_1,u_3\right)\mathbb{c}_{1,4}\left(v_1,u_4\right)\mathbb{c}_{2,3;1}\{h_{2|1}\left(v_2|v_1\right), h_{3|1}\left(u_3|v_1\right)\}$$
$$\times\, \mathbb{c}_{2,4;1}\{h_{2|1}\left(v_2|v_1\right), h_{4|1}\left(u_4|v_1\right)\}\mathrm{d}v_2\mathrm{d}v_1$$

4.(a) $$\frac{\partial^3 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1 \partial u_2 \partial u_3}$$
$$= \mathbb{c}_{1,2}\left(u_1,u_2\right)\mathbb{c}_{1,3}\left(u_1,u_3\right)\mathbb{c}_{2,3;1}\{h_{2|1}\left(u_2|u_1\right), h_{3|1}\left(u_3|u_1\right)\}$$
$$\times\, h_{4|3;1,2}\left[h_{4|2;1}\{h_{4|1}\left(u_4|u_1\right)\big|h_{2|1}\left(u_2|u_1\right)\}\middle|h_{3|2;1}\{h_{3|1}\left(u_3|u_1\right)\big|h_{2|1}\left(u_2|u_1\right)\}\right]$$

(b) $$\frac{\partial^3 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1 \partial u_2 \partial u_4}$$
$$= \mathbb{c}_{1,2}\left(u_1,u_2\right)\mathbb{c}_{1,4}\left(u_1,u_4\right)\mathbb{c}_{2,4;1}\{h_{2|1}\left(u_2|u_1\right), h_{4|1}\left(u_4|u_1\right)\}$$
$$\times\, h_{3|4;1,2}\left[h_{3|2;1}\{h_{3|1}\left(u_3|u_1\right)\big|h_{2|1}\left(u_2|u_1\right)\}\middle|h_{4|2;1}\{h_{4|1}\left(u_4|u_1\right)\big|h_{2|1}\left(u_2|u_1\right)\}\right]$$

(c) $$\frac{\partial^3 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_1 \partial u_3 \partial u_4}$$
$$= \int_0^{u_2} \mathbb{c}_{3,4;1,2}\left[h_{3|2;1}\{h_{3|1}\left(u_3|u_1\right)\big|h_{2|1}\left(v_2|u_1\right)\}, h_{4|2;1}\{h_{4|1}\left(u_4|u_1\right)\big|h_{2|1}\left(v_2|u_1\right)\}\right]$$
$$\mathbb{c}_{1,2}\left(u_1,v_2\right)\mathbb{c}_{1,3}\left(u_1,u_3\right)\mathbb{c}_{1,4}\left(u_1,u_4\right)\mathbb{c}_{2,3;1}\{h_{2|1}\left(v_2|u_1\right), h_{3|1}\left(u_3|u_1\right)\}$$
$$\times\, \mathbb{c}_{2,4;1}\{h_{2|1}\left(v_2|u_1\right), h_{4|1}\left(u_4|u_1\right)\}dv_2$$

(d) $$\frac{\partial^3 \mathbb{C}\left(u_1,u_2,u_3,u_4\right)}{\partial u_2 \partial u_3 \partial u_4}$$
$$= \int_0^{u_1} \mathbb{c}_{3,4;1,2}\left[h_{3|2;1}\{h_{3|1}\left(u_3|v_1\right)\big|h_{2|1}\left(u_2|v_1\right)\}, h_{4|2;1}\{h_{4|1}\left(u_4|v_1\right)\big|h_{2|1}\left(u_2|v_1\right)\}\right]$$
$$\mathbb{c}_{1,2}\left(v_1,u_2\right)\mathbb{c}_{1,3}\left(v_1,u_3\right)\mathbb{c}_{1,4}\left(v_1,u_4\right)\mathbb{c}_{2,3;1}\{h_{2|1}\left(u_2|v_1\right), h_{3|1}\left(u_3|v_1\right)\}$$
$$\times\, \mathbb{c}_{2,4;1}\{h_{2|1}\left(u_2|v_1\right), h_{4|1}\left(u_4|v_1\right)\}\mathrm{d}v_1$$

5. $$\mathbb{c}\left(u_1,u_2,u_3,u_4\right)$$
$$= \mathbb{c}_{1,2}\left(u_1,u_2\right)\mathbb{c}_{1,3}\left(u_1,u_3\right)\mathbb{c}_{1,4}\left(u_1,u_4\right)$$
$$\times\, \mathbb{c}_{2,3;1}\{h_{2|1}\left(u_2|u_1\right), h_{3|1}\left(u_3|u_1\right)\}\mathbb{c}_{2,4;1}\{h_{2|1}\left(u_2|u_1\right), h_{4|1}\left(u_4|u_1\right)\}$$
$$\times\, \mathbb{c}_{3,4;1,2}\left[h_{3|2;1}\{h_{3|1}\left(u_3|u_1\right)\big|h_{2|1}\left(u_2|u_1\right)\}, h_{4|2;1}\{h_{4|1}\left(u_4|u_1\right)\big|h_{2|1}\left(u_2|u_1\right)\}\right]$$

B.1.2 Conditional distribution of the leaf variable in a D-vine copula

Theorem B.5. *If the d-dimensional copula $\mathbb{C}_{1:d}$ arises from an ordered simplified D-vine copula the following holds:*

$$\frac{\partial^{d-1}\mathbb{C}_{1:d}(u_1,\dots,u_d)}{\partial u_1\cdots\partial u_{d-1}} = \mathbb{c}_{1:d-1}(u_1,\dots,u_{d-1})\mathbb{C}_{d|1:d-1}(u_d|\boldsymbol{u}_{1:d-1}).$$

Proof. To proof the equality, we first note that

$$\frac{\partial^{d-1}\mathbb{C}_{1:d}\left(u_1,\dots,u_d\right)}{\partial u_1\cdots\partial u_{d-1}} = \int_0^{u_d} \mathbb{c}_{1:d}\left(u_1,\dots,u_{d-1},v\right)\mathrm{d}v. \tag{B.3}$$

Then, using

$$\begin{aligned}\mathbb{C}_{d|1:d-1}(u_d|\boldsymbol{u}_{1:d-1}) &= \int_0^{u_d} \mathbb{c}_{d|1:d-1}\left(v|\boldsymbol{u}_{1:d-1}\right)\mathrm{d}v \\ &= \int_0^{u_d} \frac{\mathbb{c}_{1:d}\left(u_1,\dots,u_{d-1},v\right)}{\mathbb{c}_{1:d-1}\left(u_1,\dots,u_{d-1}\right)}\mathrm{d}v\end{aligned}$$

we obtain

$$\mathbb{c}_{1:d-1}\left(u_1,\dots,u_{d-1}\right)\mathbb{C}_{d|1:d-1}(u_d|\boldsymbol{u}_{1:d-1}) = \int_0^{u_d} \mathbb{c}_{1:d}\left(u_1,\dots,u_{d-1},v\right)\mathrm{d}v,$$

which combined with (B.3) concludes the proof.

□

B.2 Additional simulation results for Section 4.2.4

Table B.1: Overview of considered simulation settings with references to corresponding results.

Kendall's τ values	Copula family in $\mathcal{T}_1$	Censoring	Sample size	Table and page
$\tau_{1,2} = \tau_{2,3} = \tau_{1,3;2} = 0.3$	Clayton	65%	200 500	Table B.2 (page 145)
		25% complete data	500 500	Table B.3 (page 146)
	Gumbel	65%	200 500	Table B.4 (page 147)
		25% complete data	500 500	Table B.5 (page 148)
$\tau_{1,2} = \tau_{2,3} = \tau_{1,3;2} = 0.1$	Clayton	65%	200 500	Table B.6 (page 149)
		25% complete data	500 500	Table B.7 (page 150)
	Gumbel	65%	200 500	Table B.8 (page 151)
		25% complete data	500 500	Table B.9 (page 152)

Table B.2: Performance measures for the estimation of the copula parameters and Kendall's τ values in case of 65% common right-censored event time data with sample sizes 200 and 500. The copula combination Clayton (C), Clayton (C), Frank (F) with true $\tau_{1,2} = \tau_{2,3} = \tau_{1,3;2} = 0.3$ is investigated. Known margins, parametrically estimated margins (MLE) and nonparametrically estimated margins (KME) are considered.

				Copula parameter					Kendall's τ				
				θ	$\bar{\theta}$	$\hat{b}(\bar{\theta})$	$s^2(\bar{\theta})$	$\hat{\text{mse}}(\bar{\theta})$	τ	$\bar{\tau}$	$\hat{b}(\bar{\tau})$	$s^2(\bar{\tau})$	$\hat{\text{mse}}(\bar{\tau})$
$n = 200$, 65% censoring	Known	C	$\theta_{1,2}$	0.86	0.90	0.0410	0.1461	0.1478	0.30	0.30	-0.0015	0.0078	0.0078
		C	$\theta_{2,3}$	0.86	0.87	0.0170	0.0768	0.0771	0.30	0.30	-0.0023	0.0045	0.0046
		F	$\theta_{1,3;2}$	2.92	3.08	0.1652	0.8452	0.8725	0.30	0.31	0.0092	0.0063	0.0064
	MLE	C	$\theta_{1,2}$	0.86	0.91	0.0495	0.1525	0.1550	0.30	0.30	0.0002	0.0080	0.0080
		C	$\theta_{2,3}$	0.86	0.88	0.0274	0.0858	0.0865	0.30	0.30	-0.0004	0.0050	0.0050
		F	$\theta_{1,3;2}$	2.92	3.09	0.1771	0.8877	0.9190	0.30	0.31	0.0100	0.0065	0.0066
	KME	C	$\theta_{1,2}$	0.86	0.91	0.0506	0.1576	0.1602	0.30	0.30	0.0001	0.0082	0.0082
		C	$\theta_{2,3}$	0.86	0.88	0.0201	0.0856	0.0860	0.30	0.30	-0.0023	0.0051	0.0051
		F	$\theta_{1,3;2}$	2.92	3.13	0.2138	0.9328	0.9786	0.30	0.31	0.0129	0.0068	0.0069
$n = 500$, 65% censoring	Known	C	$\theta_{1,2}$	0.86	0.86	0.0070	0.0580	0.0581	0.30	0.30	-0.0032	0.0034	0.0034
		C	$\theta_{2,3}$	0.86	0.86	0.0058	0.0288	0.0288	0.30	0.30	-0.0010	0.0017	0.0017
		F	$\theta_{1,3;2}$	2.92	3.02	0.1038	0.3549	0.3657	0.30	0.31	0.0069	0.0027	0.0027
	MLE	C	$\theta_{1,2}$	0.86	0.87	0.0087	0.0612	0.0612	0.30	0.30	-0.0030	0.0036	0.0036
		C	$\theta_{2,3}$	0.86	0.86	0.0074	0.0308	0.0309	0.30	0.30	-0.0008	0.0018	0.0018
		F	$\theta_{1,3;2}$	2.92	3.02	0.1065	0.3635	0.3748	0.30	0.31	0.0071	0.0027	0.0028
	KME	C	$\theta_{1,2}$	0.86	0.86	0.0065	0.0614	0.0615	0.30	0.30	-0.0035	0.0036	0.0036
		C	$\theta_{2,3}$	0.86	0.86	0.0018	0.0304	0.0304	0.30	0.30	-0.0021	0.0018	0.0018
		F	$\theta_{1,3;2}$	2.92	3.03	0.1154	0.3665	0.3798	0.30	0.31	0.0078	0.0027	0.0028

Table B.3: Performance measures for the estimation of the copula parameters and Kendall's τ values in case of complete and 25% common right-censored event time data with sample size 500. The copula combination Clayton (C), Clayton (C), Frank (F) with true $\tau_{1,2} = \tau_{2,3} = \tau_{1,3;2} = 0.3$ is investigated. Known margins, parametrically estimated margins (MLE) and nonparametrically estimated margins (ECDF/KME) are considered.

				Copula parameter					Kendall's τ				
				θ	$\bar{\theta}$	$\hat{b}(\bar{\theta})$	$s^2(\bar{\theta})$	$\hat{\text{mse}}(\bar{\theta})$	τ	$\bar{\tau}$	$\hat{b}(\bar{\tau})$	$s^2(\bar{\tau})$	$\hat{\text{mse}}(\bar{\tau})$
$n = 500$, 25% censoring	Known	C	$\theta_{1,2}$	0.86	0.87	0.0170	0.0145	0.0148	0.30	0.30	0.0029	0.0008	0.0009
		C	$\theta_{2,3}$	0.86	0.86	0.0014	0.0102	0.0102	0.30	0.30	-0.0005	0.0006	0.0006
		F	$\theta_{1,3;2}$	2.92	2.96	0.0425	0.1144	0.1162	0.30	0.30	0.0030	0.0009	0.0009
	MLE	C	$\theta_{1,2}$	0.86	0.87	0.0175	0.0185	0.0188	0.30	0.30	0.0027	0.0011	0.0011
		C	$\theta_{2,3}$	0.86	0.86	-0.0009	0.0136	0.0136	0.30	0.30	-0.0014	0.0008	0.0008
		F	$\theta_{1,3;2}$	2.92	2.96	0.0401	0.1181	0.1197	0.30	0.30	0.0028	0.0009	0.0009
	KME	C	$\theta_{1,2}$	0.86	0.86	0.0037	0.0180	0.0180	0.30	0.30	-0.0006	0.0011	0.0011
		C	$\theta_{2,3}$	0.86	0.84	-0.0141	0.0140	0.0142	0.30	0.30	-0.0047	0.0009	0.0009
		F	$\theta_{1,3;2}$	2.92	2.96	0.0454	0.1220	0.1241	0.30	0.30	0.0032	0.0009	0.0009
$n = 500$, complete data	Known	C	$\theta_{1,2}$	0.86	0.87	0.0136	0.0067	0.0069	0.30	0.30	0.0027	0.0004	0.0004
		C	$\theta_{2,3}$	0.86	0.86	0.0059	0.0071	0.0071	0.30	0.30	0.0008	0.0004	0.0004
		F	$\theta_{1,3;2}$	2.92	2.97	0.0536	0.0847	0.0875	0.30	0.30	0.0042	0.0007	0.0007
	MLE	C	$\theta_{1,2}$	0.86	0.86	0.0052	0.0108	0.0108	0.30	0.30	0.0004	0.0007	0.0007
		C	$\theta_{2,3}$	0.86	0.85	-0.0025	0.0108	0.0108	0.30	0.30	-0.0015	0.0007	0.0007
		F	$\theta_{1,3;2}$	2.92	2.96	0.0397	0.0821	0.0837	0.30	0.30	0.0030	0.0006	0.0006
	ECDF	C	$\theta_{1,2}$	0.86	0.88	0.0222	0.0112	0.0117	0.30	0.30	0.0045	0.0007	0.0007
		C	$\theta_{2,3}$	0.86	0.87	0.0138	0.0115	0.0117	0.30	0.30	0.0024	0.0007	0.0007
		F	$\theta_{1,3;2}$	2.92	2.96	0.0423	0.0858	0.0876	0.30	0.30	0.0032	0.0007	0.0007

Table B.4: Performance measures for the estimation of the copula parameters and Kendall's τ values in case of 65% common right-censored event time data with sample sizes 200 and 500. The copula combination Gumbel (G), Gumbel (G), Frank (F) with true $\tau_{1,2} = \tau_{2,3} = \tau_{1,3;2} = 0.3$ is investigated. Known margins, parametrically estimated margins (MLE) and nonparametrically estimated margins (KME) are considered.

				Copula parameter					Kendall's τ				
				θ	$\bar{\theta}$	$\hat{b}(\bar{\theta})$	$s^2(\bar{\theta})$	$\hat{\text{mse}}(\bar{\theta})$	τ	$\bar{\tau}$	$\hat{b}(\bar{\tau})$	$s^2(\bar{\tau})$	$\hat{\text{mse}}(\bar{\tau})$
$n = 200$, 65% censoring	Known	G	$\theta_{1,2}$	1.43	1.44	0.0086	0.0121	0.0122	0.30	0.30	0.0001	0.0029	0.0029
		G	$\theta_{2,3}$	1.43	1.44	0.0084	0.0082	0.0082	0.30	0.30	0.0013	0.0020	0.0020
		F	$\theta_{1,3;2}$	2.92	3.02	0.0992	0.8832	0.8930	0.30	0.30	0.0034	0.0067	0.0067
	MLE	G	$\theta_{1,2}$	1.43	1.44	0.0148	0.0138	0.0140	0.30	0.30	0.0026	0.0032	0.0032
		G	$\theta_{2,3}$	1.43	1.44	0.0131	0.0106	0.0108	0.30	0.30	0.0028	0.0025	0.0025
		F	$\theta_{1,3;2}$	2.92	3.04	0.1179	0.9273	0.9411	0.30	0.30	0.0047	0.0070	0.0070
	KME	G	$\theta_{1,2}$	1.43	1.47	0.0464	0.0163	0.0184	0.30	0.32	0.0170	0.0035	0.0038
		G	$\theta_{2,3}$	1.43	1.47	0.0365	0.0112	0.0125	0.30	0.31	0.0139	0.0025	0.0027
		F	$\theta_{1,3;2}$	2.92	3.07	0.1511	0.9618	0.9846	0.30	0.31	0.0074	0.0072	0.0072
$n = 500$, 65% censoring	Known	G	$\theta_{1,2}$	1.43	1.44	0.0080	0.0053	0.0053	0.30	0.30	0.0022	0.0012	0.0012
		G	$\theta_{2,3}$	1.43	1.43	0.0030	0.0032	0.0033	0.30	0.30	0.0003	0.0008	0.0008
		F	$\theta_{1,3;2}$	2.92	2.96	0.0454	0.3610	0.3630	0.30	0.30	0.0018	0.0027	0.0027
	MLE	G	$\theta_{1,2}$	1.43	1.44	0.0081	0.0064	0.0064	0.30	0.30	0.0018	0.0015	0.0015
		G	$\theta_{2,3}$	1.43	1.43	0.0038	0.0040	0.0040	0.30	0.30	0.0005	0.0010	0.0010
		F	$\theta_{1,3;2}$	2.92	2.96	0.0446	0.3685	0.3705	0.30	0.30	0.0017	0.0028	0.0028
	KME	G	$\theta_{1,2}$	1.43	1.45	0.0210	0.0074	0.0078	0.30	0.31	0.0077	0.0017	0.0017
		G	$\theta_{2,3}$	1.43	1.44	0.0127	0.0042	0.0044	0.30	0.30	0.0048	0.0010	0.0010
		F	$\theta_{1,3;2}$	2.92	2.98	0.0620	0.3768	0.3806	0.30	0.30	0.0031	0.0028	0.0028

Table B.5: Performance measures for the estimation of the copula parameters and Kendall's τ values in case of complete and 25% common right-censored event time data with sample size 500. The copula combination Gumbel (G), Gumbel (G), Frank (F) with true $\tau_{1,2} = \tau_{2,3} = \tau_{1,3;2} = 0.3$ is investigated. Known margins, parametrically estimated margins (MLE) and nonparametrically estimated margins (ECDF/KME) are considered.

				Copula parameter					Kendall's τ				
				θ	$\bar{\theta}$	$\hat{b}(\bar{\theta})$	$s^2(\bar{\theta})$	$\hat{\text{mse}}(\bar{\theta})$	τ	$\bar{\tau}$	$\hat{b}(\bar{\tau})$	$s^2(\bar{\tau})$	$\hat{\text{mse}}(\bar{\tau})$
$n = 500$, 25% censoring	Known	G	$\theta_{1,2}$	1.43	1.44	0.0071	0.0029	0.0029	0.30	0.30	0.0025	0.0007	0.0007
		G	$\theta_{2,3}$	1.43	1.43	0.0022	0.0025	0.0025	0.30	0.30	0.0002	0.0006	0.0006
		F	$\theta_{1,3;2}$	2.92	2.95	0.0312	0.1125	0.1135	0.30	0.30	0.0020	0.0009	0.0009
	MLE	G	$\theta_{1,2}$	1.43	1.44	0.0090	0.0037	0.0037	0.30	0.30	0.0031	0.0009	0.0009
		G	$\theta_{2,3}$	1.43	1.43	0.0041	0.0030	0.0030	0.30	0.30	0.0010	0.0007	0.0007
		F	$\theta_{1,3;2}$	2.92	2.94	0.0245	0.1139	0.1145	0.30	0.30	0.0014	0.0009	0.0009
	KME	G	$\theta_{1,2}$	1.43	1.44	0.0136	0.0039	0.0041	0.30	0.31	0.0053	0.0009	0.0009
		G	$\theta_{2,3}$	1.43	1.44	0.0077	0.0031	0.0031	0.30	0.30	0.0027	0.0007	0.0007
		F	$\theta_{1,3;2}$	2.92	2.95	0.0295	0.1168	0.1177	0.30	0.30	0.0019	0.0009	0.0009
$n = 500$, complete data	Known	G	$\theta_{1,2}$	1.43	1.44	0.0069	0.0023	0.0023	0.30	0.30	0.0026	0.0005	0.0005
		G	$\theta_{2,3}$	1.43	1.43	0.0028	0.0023	0.0023	0.30	0.30	0.0006	0.0006	0.0006
		F	$\theta_{1,3;2}$	2.92	2.96	0.0474	0.0895	0.0917	0.30	0.30	0.0036	0.0007	0.0007
	MLE	G	$\theta_{1,2}$	1.43	1.43	0.0041	0.0027	0.0027	0.30	0.30	0.0011	0.0006	0.0006
		G	$\theta_{2,3}$	1.43	1.43	0.0004	0.0028	0.0028	0.30	0.30	-0.0008	0.0007	0.0007
		F	$\theta_{1,3;2}$	2.92	2.95	0.0305	0.0873	0.0882	0.30	0.30	0.0021	0.0007	0.0007
	ECDF	G	$\theta_{1,2}$	1.43	1.44	0.0134	0.0031	0.0033	0.30	0.31	0.0055	0.0007	0.0007
		G	$\theta_{2,3}$	1.43	1.44	0.0093	0.0030	0.0031	0.30	0.30	0.0035	0.0007	0.0007
		F	$\theta_{1,3;2}$	2.92	2.95	0.0337	0.0892	0.0904	0.30	0.30	0.0024	0.0007	0.0007

Table B.6: Performance measures for the estimation of the copula parameters and Kendall's τ values in case of 65% common right-censored event time data with sample sizes 200 and 500. The copula combination Clayton (C), Clayton (C), Frank (F) with true $\tau_{1,2} = \tau_{2,3} = \tau_{1,3;2} = 0.1$ is investigated. Known margins, parametrically estimated margins (MLE) and nonparametrically estimated margins (KME) are considered.

				Copula parameter					Kendall's τ				
				θ	$\bar{\theta}$	$\hat{b}(\bar{\theta})$	$s^2(\bar{\theta})$	$\hat{\text{mse}}(\bar{\theta})$	τ	$\bar{\tau}$	$\hat{b}(\bar{\tau})$	$s^2(\bar{\tau})$	$\hat{\text{mse}}(\bar{\tau})$
$n = 200$, 65% censoring	Known	C	$\theta_{1,2}$	0.22	0.27	0.0487	0.0550	0.0573	0.10	0.11	0.0106	0.0074	0.0075
		C	$\theta_{2,3}$	0.22	0.26	0.0341	0.0353	0.0365	0.10	0.11	0.0077	0.0051	0.0052
		F	$\theta_{1,3;2}$	0.91	1.02	0.1120	0.6702	0.6827	0.10	0.11	0.0101	0.0075	0.0076
	MLE	C	$\theta_{1,2}$	0.22	0.27	0.0511	0.0565	0.0592	0.10	0.11	0.0113	0.0075	0.0076
		C	$\theta_{2,3}$	0.22	0.26	0.0363	0.0364	0.0377	0.10	0.11	0.0084	0.0052	0.0053
		F	$\theta_{1,3;2}$	0.91	1.02	0.1108	0.7013	0.7136	0.10	0.11	0.0099	0.0078	0.0079
	KME	C	$\theta_{1,2}$	0.22	0.28	0.0564	0.0575	0.0606	0.10	0.11	0.0132	0.0076	0.0077
		C	$\theta_{2,3}$	0.22	0.26	0.0386	0.0361	0.0376	0.10	0.11	0.0094	0.0052	0.0053
		F	$\theta_{1,3;2}$	0.91	1.02	0.1103	0.7159	0.7281	0.10	0.11	0.0098	0.0080	0.0081
$n = 500$, 65% censoring	Known	C	$\theta_{1,2}$	0.22	0.23	0.0099	0.0237	0.0238	0.10	0.10	-0.0002	0.0037	0.0037
		C	$\theta_{2,3}$	0.22	0.23	0.0101	0.0137	0.0138	0.10	0.10	0.0017	0.0021	0.0021
		F	$\theta_{1,3;2}$	0.91	0.96	0.0507	0.2776	0.2801	0.10	0.10	0.0047	0.0032	0.0032
	MLE	C	$\theta_{1,2}$	0.22	0.23	0.0107	0.0241	0.0242	0.10	0.10	0.0001	0.0037	0.0037
		C	$\theta_{2,3}$	0.22	0.23	0.0094	0.0134	0.0135	0.10	0.10	0.0014	0.0021	0.0021
		F	$\theta_{1,3;2}$	0.91	0.96	0.0514	0.2776	0.2802	0.10	0.10	0.0048	0.0032	0.0032
	KME	C	$\theta_{1,2}$	0.22	0.24	0.0134	0.0234	0.0236	0.10	0.10	0.0013	0.0036	0.0036
		C	$\theta_{2,3}$	0.22	0.23	0.0087	0.0135	0.0136	0.10	0.10	0.0011	0.0021	0.0021
		F	$\theta_{1,3;2}$	0.91	0.96	0.0552	0.2797	0.2828	0.10	0.11	0.0052	0.0032	0.0033

Table B.7: Performance measures for the estimation of the copula parameters and Kendall's τ values in case of complete and 25% common right-censored event time data with sample size 500. The copula combination Clayton (C), Clayton (C), Frank (F) with true $\tau_{1,2} = \tau_{2,3} = \tau_{1,3;2} = 0.1$ is investigated. Known margins, parametrically estimated margins (MLE) and nonparametrically estimated margins (ECDF/KME) are considered.

				Copula parameter					Kendall's τ				
				θ	$\bar{\theta}$	$\hat{b}(\bar{\theta})$	$s^2(\bar{\theta})$	$\hat{\text{mse}}(\bar{\theta})$	τ	$\bar{\tau}$	$\hat{b}(\bar{\tau})$	$s^2(\bar{\tau})$	$\hat{\text{mse}}(\bar{\tau})$
$n = 500$, 25% censoring	Known	C	$\theta_{1,2}$	0.22	0.23	0.0031	0.0078	0.0078	0.10	0.10	-0.0001	0.0013	0.0013
		C	$\theta_{2,3}$	0.22	0.22	0.0024	0.0054	0.0054	0.10	0.10	-0.0000	0.0009	0.0009
		F	$\theta_{1,3;2}$	0.91	0.95	0.0395	0.0908	0.0924	0.10	0.10	0.0040	0.0011	0.0011
	MLE	C	$\theta_{1,2}$	0.22	0.22	-0.0002	0.0080	0.0080	0.10	0.10	-0.0015	0.0013	0.0013
		C	$\theta_{2,3}$	0.22	0.22	-0.0009	0.0054	0.0054	0.10	0.10	-0.0013	0.0009	0.0009
		F	$\theta_{1,3;2}$	0.91	0.95	0.0379	0.0887	0.0902	0.10	0.10	0.0038	0.0010	0.0010
	KME	C	$\theta_{1,2}$	0.22	0.22	0.0017	0.0077	0.0077	0.10	0.10	-0.0007	0.0012	0.0012
		C	$\theta_{2,3}$	0.22	0.22	0.0009	0.0057	0.0057	0.10	0.10	-0.0007	0.0009	0.0009
		F	$\theta_{1,3;2}$	0.91	0.94	0.0355	0.0894	0.0907	0.10	0.10	0.0036	0.0010	0.0011
$n = 500$, complete data	Known	C	$\theta_{1,2}$	0.22	0.22	0.0024	0.0036	0.0036	0.10	0.10	0.0003	0.0006	0.0006
		C	$\theta_{2,3}$	0.22	0.23	0.0031	0.0036	0.0036	0.10	0.10	0.0006	0.0006	0.0006
		F	$\theta_{1,3;2}$	0.91	0.94	0.0285	0.0675	0.0683	0.10	0.10	0.0029	0.0008	0.0008
	MLE	C	$\theta_{1,2}$	0.22	0.22	-0.0005	0.0040	0.0040	0.10	0.10	-0.0009	0.0007	0.0007
		C	$\theta_{2,3}$	0.22	0.22	0.0012	0.0038	0.0038	0.10	0.10	-0.0002	0.0006	0.0006
		F	$\theta_{1,3;2}$	0.91	0.93	0.0197	0.0657	0.0660	0.10	0.10	0.0019	0.0008	0.0008
	ECDF	C	$\theta_{1,2}$	0.22	0.23	0.0100	0.0043	0.0044	0.10	0.10	0.0033	0.0007	0.0007
		C	$\theta_{2,3}$	0.22	0.23	0.0112	0.0042	0.0043	0.10	0.10	0.0038	0.0007	0.0007
		F	$\theta_{1,3;2}$	0.91	0.93	0.0214	0.0670	0.0675	0.10	0.10	0.0021	0.0008	0.0008

Table B.8: Performance measures for the estimation of the copula parameters and Kendall's τ values in case of 65% common right-censored event time data with sample sizes 200 and 500. The copula combination Gumbel (G), Gumbel (G), Frank (F) with true $\tau_{1,2} = \tau_{2,3} = \tau_{1,3;2} = 0.1$ is investigated. Known margins, parametrically estimated margins (MLE) and nonparametrically estimated margins (KME) are considered.

				Copula parameter					Kendall's τ				
				θ	$\bar{\theta}$	$\hat{b}(\bar{\theta})$	$s^2(\bar{\theta})$	$\hat{\text{mse}}(\bar{\theta})$	τ	$\bar{\tau}$	$\hat{b}(\bar{\tau})$	$s^2(\bar{\tau})$	$\hat{\text{mse}}(\bar{\tau})$
$n = 200$, 65% censoring	Known	G	$\theta_{1,2}$	1.11	1.12	0.0087	0.0049	0.0050	0.10	0.10	0.0036	0.0030	0.0030
		G	$\theta_{2,3}$	1.11	1.12	0.0068	0.0036	0.0036	0.10	0.10	0.0029	0.0023	0.0023
		F	$\theta_{1,3;2}$	0.91	0.97	0.0641	0.6779	0.6821	0.10	0.10	0.0050	0.0076	0.0076
	MLE	G	$\theta_{1,2}$	1.11	1.12	0.0091	0.0048	0.0049	0.10	0.10	0.0039	0.0030	0.0030
		G	$\theta_{2,3}$	1.11	1.12	0.0070	0.0037	0.0038	0.10	0.10	0.0030	0.0024	0.0024
		F	$\theta_{1,3;2}$	0.91	0.97	0.0649	0.6984	0.7026	0.10	0.11	0.0050	0.0078	0.0079
	KME	G	$\theta_{1,2}$	1.11	1.13	0.0231	0.0058	0.0063	0.10	0.11	0.0145	0.0034	0.0036
		G	$\theta_{2,3}$	1.11	1.13	0.0179	0.0041	0.0044	0.10	0.11	0.0114	0.0025	0.0027
		F	$\theta_{1,3;2}$	0.91	0.98	0.0717	0.6913	0.6964	0.10	0.11	0.0058	0.0078	0.0078
$n = 500$, 65% censoring	Known	G	$\theta_{1,2}$	1.11	1.11	0.0032	0.0017	0.0018	0.10	0.10	0.0013	0.0011	0.0011
		G	$\theta_{2,3}$	1.11	1.11	0.0004	0.0013	0.0013	0.10	0.10	-0.0006	0.0008	0.0008
		F	$\theta_{1,3;2}$	0.91	0.92	0.0156	0.2675	0.2678	0.10	0.10	0.0009	0.0031	0.0031
	MLE	G	$\theta_{1,2}$	1.11	1.11	0.0030	0.0018	0.0018	0.10	0.10	0.0012	0.0012	0.0012
		G	$\theta_{2,3}$	1.11	1.11	0.0007	0.0013	0.0013	0.10	0.10	-0.0004	0.0009	0.0009
		F	$\theta_{1,3;2}$	0.91	0.92	0.0147	0.2674	0.2676	0.10	0.10	0.0008	0.0031	0.0031
	KME	G	$\theta_{1,2}$	1.11	1.12	0.0097	0.0020	0.0021	0.10	0.11	0.0063	0.0013	0.0013
		G	$\theta_{2,3}$	1.11	1.12	0.0051	0.0014	0.0014	0.10	0.10	0.0031	0.0009	0.0009
		F	$\theta_{1,3;2}$	0.91	0.93	0.0233	0.2635	0.2641	0.10	0.10	0.0018	0.0030	0.0030

Table B.9: Performance measures for the estimation of the copula parameters and Kendall's τ values in case of complete and 25% common right-censored event time data with sample size 500. The copula combination Gumbel (G), Gumbel (G), Frank (F) with true $\tau_{1,2} = \tau_{2,3} = \tau_{1,3;2} = 0.1$ is investigated. Known margins, parametrically estimated margins (MLE) and nonparametrically estimated margins (ECDF/KME) are considered.

				Copula parameter					Kendall's τ				
				θ	$\bar{\theta}$	$\hat{b}(\bar{\theta})$	$s^2(\bar{\theta})$	$\hat{\text{mse}}(\bar{\theta})$	τ	$\bar{\tau}$	$\hat{b}(\bar{\tau})$	$s^2(\bar{\tau})$	$\hat{\text{mse}}(\bar{\tau})$
$n = 500$, 25% censoring	Known	G	$\theta_{1,2}$	1.11	1.11	0.0031	0.0013	0.0013	0.10	0.10	0.0015	0.0009	0.0009
		G	$\theta_{2,3}$	1.11	1.11	0.0008	0.0011	0.0011	0.10	0.10	-0.0002	0.0007	0.0007
		F	$\theta_{1,3;2}$	0.91	0.94	0.0364	0.0898	0.0911	0.10	0.10	0.0037	0.0010	0.0011
	MLE	G	$\theta_{1,2}$	1.11	1.11	0.0033	0.0014	0.0014	0.10	0.10	0.0016	0.0009	0.0009
		G	$\theta_{2,3}$	1.11	1.11	0.0015	0.0012	0.0012	0.10	0.10	0.0003	0.0008	0.0008
		F	$\theta_{1,3;2}$	0.91	0.94	0.0335	0.0889	0.0900	0.10	0.10	0.0034	0.0010	0.0010
	KME	G	$\theta_{1,2}$	1.11	1.12	0.0072	0.0015	0.0015	0.10	0.10	0.0048	0.0009	0.0010
		G	$\theta_{2,3}$	1.11	1.12	0.0044	0.0012	0.0012	0.10	0.10	0.0027	0.0007	0.0008
		F	$\theta_{1,3;2}$	0.91	0.94	0.0344	0.0885	0.0897	0.10	0.10	0.0035	0.0010	0.0010
$n = 500$, complete data	Known	G	$\theta_{1,2}$	1.11	1.11	0.0025	0.0011	0.0011	0.10	0.10	0.0013	0.0007	0.0007
		G	$\theta_{2,3}$	1.11	1.11	0.0008	0.0010	0.0010	0.10	0.10	-0.0001	0.0007	0.0007
		F	$\theta_{1,3;2}$	0.91	0.93	0.0253	0.0679	0.0685	0.10	0.10	0.0025	0.0008	0.0008
	MLE	G	$\theta_{1,2}$	1.11	1.11	0.0015	0.0011	0.0011	0.10	0.10	0.0004	0.0007	0.0007
		G	$\theta_{2,3}$	1.11	1.11	0.0002	0.0011	0.0011	0.10	0.10	-0.0006	0.0007	0.0007
		F	$\theta_{1,3;2}$	0.91	0.92	0.0168	0.0664	0.0667	0.10	0.10	0.0016	0.0008	0.0008
	ECDF	G	$\theta_{1,2}$	1.11	1.12	0.0062	0.0012	0.0012	0.10	0.10	0.0041	0.0008	0.0008
		G	$\theta_{2,3}$	1.11	1.12	0.0045	0.0011	0.0012	0.10	0.10	0.0028	0.0007	0.0007
		F	$\theta_{1,3;2}$	0.91	0.93	0.0187	0.0673	0.0676	0.10	0.10	0.0018	0.0008	0.0008

B.3 Additional bootstrapping results for Section 4.3.2

In this chapter, detailed results of the parametric and nonparametric copula bootstrap as applied to the four R-vine copula models that best describe the mastitis data (see Table 4.9 in Section 4.3.2) are shown. All results are based on 100 replications. Each table contains in lines 1-4 the results for the copula parameters, in lines 5-8 the results for the Kendall's τ values and in lines 9-12 the results for the lower tail-dependence coefficients λ^{L} (LTD). Since Frank (F) copulas do not exhibit any tail-dependence, the latter are only reported for Clayton (C) copulas. First, the underlying model, i.e. the model estimated for the mastitis data, is given. Second, we provide estimation results, when refitting 100 simulated (based on the parametric model) data sets before censoring is induced, i.e. in case of complete data. This serves as a benchmark to assess the impact of information loss due to right-censoring. Second, we give estimation results in case of 66% censoring (as in the mastitis data) using the parametric bootstrap (PB) introduced in Section 4.3.1. Fourth, results based on a nonparametric bootstrap (NPB) are shown. We always list the mean estimate together with the corresponding standard error (in parenthesis).

1st best model

- global likelihood estimation

Parameter	model	$F;\ \hat{\theta}_{1,3}: 6.56$	$F;\ \hat{\theta}_{3,4}: 6.34$	$F;\ \hat{\theta}_{2,4}: 6.99$	$F;\ \hat{\theta}_{1,4;3}: 1.68$	$F;\ \hat{\theta}_{2,3;4}: 2.79$	$F;\ \hat{\theta}_{1,2;3,4}: 3.71$
	complete data	6.641 (0.406)	6.324 (0.377)	7.045 (0.444)	1.712 (0.295)	2.794 (0.304)	3.715 (0.353)
	66% cens. PB	6.727 (0.797)	6.376 (0.746)	7.102 (0.770)	1.792 (0.549)	2.886 (0.553)	3.740 (0.650)
	66% cens. NPB	6.526 (0.756)	6.493 (0.865)	7.128 (0.728)	1.812 (0.688)	2.887 (0.628)	3.593 (0.741)
Kendall's τ	model	$F;\ \hat{\tau}_{1,3}: 0.54$	$F;\ \hat{\tau}_{3,4}: 0.53$	$F;\ \hat{\tau}_{2,4}: 0.56$	$F;\ \hat{\tau}_{1,4;3}: 0.18$	$F;\ \hat{\tau}_{2,3;4}: 0.29$	$F;\ \hat{\tau}_{1,2;3,4}: 0.37$
	complete data	0.545 (0.019)	0.530 (0.019)	0.563 (0.020)	0.184 (0.030)	0.288 (0.027)	0.366 (0.027)
	66% cens. PB	0.547 (0.036)	0.531 (0.036)	0.565 (0.034)	0.192 (0.055)	0.295 (0.049)	0.366 (0.050)
	66% cens. NPB	0.538 (0.035)	0.536 (0.040)	0.566 (0.031)	0.193 (0.069)	0.295 (0.055)	0.353 (0.058)
LTD	model						
	complete data						
	66% cens. PB						
	66% cens. NPB						

- $\mathcal{T}_1$-sequential likelihood estimation

Parameter	model	$F;\ \hat{\theta}_{1,3}: 6.38$	$F;\ \hat{\theta}_{3,4}: 6.34$	$F;\ \hat{\theta}_{2,4}: 6.77$	$F;\ \hat{\theta}_{1,4;3}: 1.67$	$F;\ \hat{\theta}_{2,3;4}: 2.81$	$F;\ \hat{\theta}_{1,2;3,4}: 3.72$
	complete data	6.641 (0.406)	6.324 (0.377)	7.045 (0.444)	1.712 (0.295)	2.794 (0.304)	3.715 (0.353)
	66% cens. PB	6.561 (0.808)	6.367 (0.786)	6.860 (0.802)	1.771 (0.565)	2.917 (0.566)	3.733 (0.629)
	66% cens. NPB	6.377 (0.768)	6.407 (0.906)	6.833 (0.784)	1.767 (0.677)	2.876 (0.616)	3.602 (0.725)
Kendall's τ	model	$F;\ \hat{\tau}_{1,3}: 0.53$	$F;\ \hat{\tau}_{3,4}: 0.53$	$F;\ \hat{\tau}_{2,4}: 0.55$	$F;\ \hat{\tau}_{1,4;3}: 0.18$	$F;\ \hat{\tau}_{2,3;4}: 0.29$	$F;\ \hat{\tau}_{1,2;3,4}: 0.37$
	complete data	0.537 (0.019)	0.530 (0.019)	0.553 (0.020)	0.183 (0.030)	0.291 (0.027)	0.366 (0.027)
	66% cens. PB	0.540 (0.038)	0.530 (0.038)	0.553 (0.036)	0.189 (0.057)	0.298 (0.050)	0.365 (0.049)
	66% cens. NPB	0.531 (0.037)	0.531 (0.043)	0.552 (0.035)	0.188 (0.069)	0.294 (0.054)	0.354 (0.057)
LTD	model						
	complete data						
	66% cens. PB						
	66% cens. NPB						

2nd best model

- global likelihood estimation

Parameter	model	C; $\hat{\theta}_{1,3}$: 3.78	F; $\hat{\theta}_{3,4}$: 6.39	F; $\hat{\theta}_{2,4}$: 6.93	F; $\hat{\theta}_{1,4;3}$: 1.51	F; $\hat{\theta}_{2,3;4}$: 2.78	F; $\hat{\theta}_{1,2;3,4}$: 3.48
	complete data	3.810 (0.214)	6.364 (0.384)	6.972 (0.447)	1.538 (0.286)	2.768 (0.298)	3.516 (0.335)
	66% cens. PB	3.824 (0.580)	6.420 (0.754)	7.052 (0.742)	1.624 (0.530)	2.858 (0.510)	3.491 (0.614)
	66% cens. NPB	3.675 (0.591)	6.502 (0.859)	7.108 (0.707)	1.619 (0.689)	2.835 (0.645)	3.349 (0.7)
Kendall's τ	model	C; $\hat{\tau}_{1,3}$: 0.65	F; $\hat{\tau}_{3,4}$: 0.53	F; $\hat{\tau}_{2,4}$: 0.56	F; $\hat{\tau}_{1,4;3}$: 0.16	F; $\hat{\tau}_{2,3;4}$: 0.29	F; $\hat{\tau}_{1,2;3,4}$: 0.35
	complete data	0.655 (0.013)	0.532 (0.019)	0.560 (0.020)	0.167 (0.030)	0.286 (0.027)	0.350 (0.027)
	66% cens. PB	0.653 (0.035)	0.533 (0.036)	0.562 (0.033)	0.175 (0.054)	0.293 (0.046)	0.346 (0.049)
	66% cens. NPB	0.644 (0.036)	0.536 (0.040)	0.565 (0.030)	0.173 (0.070)	0.290 (0.057)	0.334 (0.057)
LTD	model	C; $\hat{\lambda}^{\mathrm{L}}_{1,3}$: 0.83					
	complete data	0.833 (0.008)					
	66% cens. PB	0.831 (0.024)					
	66% cens. NPB	0.825 (0.025)					

- $\mathcal{T}_1$-sequential likelihood estimation

Parameter	model	C; $\hat{\theta}_{1,3}$: 3.60	F; $\hat{\theta}_{3,4}$: 6.34	F; $\hat{\theta}_{2,4}$: 6.77	F; $\hat{\theta}_{1,4;3}$: 1.49	F; $\hat{\theta}_{2,3;4}$: 2.81	F; $\hat{\theta}_{1,2;3,4}$: 3.48
	complete data	3.629 (0.206)	6.315 (0.382)	6.806 (0.441)	1.508 (0.287)	2.797 (0.297)	3.523 (0.335)
	66% cens. PB	3.653 (0.578)	6.379 (0.785)	6.863 (0.786)	1.613 (0.576)	2.884 (0.534)	3.493 (0.633)
	66% cens. NPB	3.566 (0.599)	6.407 (0.906)	6.833 (0.784)	1.587 (0.678)	2.826 (0.643)	3.361 (0.684)
Kendall's τ	model	C; $\hat{\tau}_{1,3}$: 0.64	F; $\hat{\tau}_{3,4}$: 0.53	F; $\hat{\tau}_{2,4}$: 0.55	F; $\hat{\tau}_{1,4;3}$: 0.16	F; $\hat{\tau}_{2,3;4}$: 0.29	F; $\hat{\tau}_{1,2;3,4}$: 0.35
	complete data	0.644 (0.013)	0.530 (0.019)	0.553 (0.020)	0.164 (0.030)	0.289 (0.027)	0.350 (0.027)
	66% cens. PB	0.643 (0.037)	0.531 (0.038)	0.554 (0.035)	0.173 (0.059)	0.295 (0.048)	0.346 (0.051)
	66% cens. NPB	0.637 (0.039)	0.531 (0.043)	0.552 (0.035)	0.170 (0.070)	0.289 (0.057)	0.335 (0.056)
LTD	model	C; $\hat{\lambda}^{\mathrm{L}}_{1,3}$: 0.82					
	complete data	0.826 (0.009)					
	66% cens. PB	0.824 (0.026)					
	66% cens. NPB	0.819 (0.028)					

3rd best model

- global likelihood estimation

Parameter	model	F; $\hat{\theta}_{1,3}$: 6.51	F; $\hat{\theta}_{3,4}$: 6.36	C; $\hat{\theta}_{2,4}$: 4.10	F; $\hat{\theta}_{1,4;3}$: 1.57	F; $\hat{\theta}_{2,3;4}$: 2.79	F; $\hat{\theta}_{1,2;3,4}$: 3.86
	complete data	6.592 (0.407)	6.343 (0.376)	4.138 (0.237)	1.606 (0.288)	2.798 (0.302)	3.869 (0.357)
	66% cens. PB	6.676 (0.792)	6.406 (0.719)	4.134 (0.612)	1.668 (0.549)	2.889 (0.550)	3.887 (0.650)
	66% cens. NPB	6.539 (0.738)	6.471 (0.835)	4.139 (0.593)	1.674 (0.682)	2.836 (0.633)	3.704 (0.752)
Kendall's τ	model	F; $\hat{\tau}_{1,3}$: 0.54	F; $\hat{\tau}_{3,4}$: 0.53	C; $\hat{\tau}_{2,4}$: 0.67	F; $\hat{\tau}_{1,4;3}$: 0.17	F; $\hat{\tau}_{2,3;4}$: 0.29	F; $\hat{\tau}_{1,2;3,4}$: 0.38
	complete data	0.543 (0.019)	0.531 (0.019)	0.674 (0.013)	0.174 (0.030)	0.289 (0.027)	0.377 (0.027)
	66% cens. PB	0.545 (0.036)	0.533 (0.035)	0.671 (0.033)	0.179 (0.056)	0.296 (0.049)	0.377 (0.050)
	66% cens. NPB	0.539 (0.034)	0.535 (0.039)	0.671 (0.031)	0.179 (0.070)	0.29 (0.056)	0.362 (0.058)
LTD	model			C; $\hat{\lambda}^{\mathrm{L}}_{2,4}$: 0.84			
	complete data			0.845 (0.008)			
	66% cens. PB			0.843 (0.022)			
	66% cens. NPB			0.843 (0.020)			

- $\mathcal{T}_1$-sequential likelihood estimation

Parameter	model	F; $\hat{\theta}_{1,3} : 6.38$	F; $\hat{\theta}_{3,4} : 6.34$	C; $\hat{\theta}_{2,4} : 3.90$	F; $\hat{\theta}_{1,4;3} : 1.54$	F; $\hat{\theta}_{2,3;4} : 2.76$	F; $\hat{\theta}_{1,2;3,4} : 3.86$
	complete data	6.460 (0.403)	6.315 (0.375)	3.937 (0.228)	1.568 (0.288)	2.769 (0.301)	3.868 (0.358)
	66% cens. PB	6.562 (0.808)	6.373 (0.791)	3.907 (0.598)	1.634 (0.550)	2.860 (0.553)	3.862 (0.635)
	66% cens. NPB	6.377 (0.768)	6.407 (0.906)	3.946 (0.617)	1.649 (0.675)	2.833 (0.626)	3.721 (0.739)
Kendall's τ	model	F; $\hat{\tau}_{1,3} : 0.53$	F; $\hat{\tau}_{3,4} : 0.53$	C; $\hat{\tau}_{2,4} : 0.66$	F; $\hat{\tau}_{1,4;3} : 0.17$	F; $\hat{\tau}_{2,3;4} : 0.29$	F; $\hat{\tau}_{1,2;3,4} : 0.38$
	complete data	0.537 (0.019)	0.530 (0.019)	0.663 (0.013)	0.170 (0.030)	0.286 (0.027)	0.377 (0.027)
	66% cens. PB	0.540 (0.038)	0.530 (0.039)	0.658 (0.034)	0.175 (0.056)	0.293 (0.049)	0.375 (0.048)
	66% cens. NPB	0.531 (0.037)	0.531 (0.043)	0.660 (0.034)	0.176 (0.069)	0.290 (0.055)	0.363 (0.057)
LTD	model			C; $\hat{\lambda}^{\mathrm{L}}_{2,4} : 0.84$			
	complete data			0.838 (0.009)			
	66% cens. PB			0.834 (0.023)			
	66% cens. NPB			0.836 (0.023)			

4th best model

- global likelihood estimation

Parameter	model	C; $\hat{\theta}_{1,3} : 3.75$	F; $\hat{\theta}_{3,4} : 6.40$	C; $\hat{\theta}_{2,4} : 4.04$	F; $\hat{\theta}_{1,4;3} : 1.39$	F; $\hat{\theta}_{2,3;4} : 2.72$	F; $\hat{\theta}_{1,2;3,4} : 3.71$
	complete data	3.788 (0.212)	6.377 (0.380)	4.070 (0.235)	1.414 (0.274)	2.706 (0.292)	3.744 (0.339)
	66% cens. PB	3.782 (0.612)	6.437 (0.742)	4.090 (0.590)	1.491 (0.552)	2.780 (0.509)	3.747 (0.625)
	66% cens. NPB	3.664 (0.586)	6.501 (0.852)	4.093 (0.572)	1.490 (0.670)	2.758 (0.641)	3.550 (0.721)
Kendall's τ	model	C; $\hat{\tau}_{1,3} : 0.65$	F; $\hat{\tau}_{3,4} : 0.53$	C; $\hat{\tau}_{2,4} : 0.67$	F; $\hat{\tau}_{1,4;3} : 0.15$	F; $\hat{\tau}_{2,3;4} : 0.28$	F; $\hat{\tau}_{1,2;3,4} : 0.37$
	complete data	0.654 (0.013)	0.533 (0.019)	0.670 (0.013)	0.154 (0.029)	0.281 (0.026)	0.368 (0.026)
	66% cens. PB	0.650 (0.037)	0.534 (0.035)	0.669 (0.032)	0.161 (0.057)	0.286 (0.046)	0.366 (0.048)
	66% cens. NPB	0.643 (0.036)	0.536 (0.039)	0.669 (0.030)	0.160 (0.069)	0.283 (0.057)	0.350 (0.057)
LTD	model	C; $\hat{\lambda}^{\mathrm{L}}_{1,3} : 0.83$		C; $\hat{\lambda}^{\mathrm{L}}_{2,4} : 0.84$			
	complete data	0.832 (0.008)		0.843 (0.008)			
	66% cens. PB	0.829 (0.025)		0.841 (0.022)			
	66% cens. NPB	0.824 (0.024)		0.842 (0.020)			

- $\mathcal{T}_1$-sequential likelihood estimation

Parameter	model	C; $\hat{\theta}_{1,3} : 3.60$	F; $\hat{\theta}_{3,4} : 6.34$	C; $\hat{\theta}_{2,4} : 3.90$	F; $\hat{\theta}_{1,4;3} : 1.36$	F; $\hat{\theta}_{2,3;4} : 2.71$	F; $\hat{\theta}_{1,2;3,4} : 3.70$
	complete data	3.630 (0.206)	6.314 (0.378)	3.926 (0.229)	1.389 (0.275)	2.690 (0.291)	3.741 (0.339)
	66% cens. PB	3.653 (0.578)	6.381 (0.791)	3.914 (0.588)	1.464 (0.565)	2.770 (0.530)	3.703 (0.644)
	66% cens. NPB	3.566 (0.599)	6.407 (0.906)	3.946 (0.617)	1.463 (0.672)	2.742 (0.643)	3.560 (0.712)
Kendall's τ	model	C; $\hat{\tau}_{1,3} : 0.64$	F; $\hat{\tau}_{3,4} : 0.53$	C; $\hat{\tau}_{2,4} : 0.66$	F; $\hat{\tau}_{1,4;3} : 0.15$	F; $\hat{\tau}_{2,3;4} : 0.28$	F; $\hat{\tau}_{1,2;3,4} : 0.37$
	complete data	0.644 (0.013)	0.530 (0.019)	0.662 (0.013)	0.151 (0.029)	0.279 (0.026)	0.368 (0.026)
	66% cens. PB	0.643 (0.037)	0.531 (0.038)	0.659 (0.034)	0.158 (0.058)	0.285 (0.048)	0.363 (0.050)
	66% cens. NPB	0.637 (0.039)	0.531 (0.043)	0.66 (0.034)	0.157 (0.070)	0.282 (0.058)	0.351 (0.057)
LTD	model	C; $\hat{\lambda}^{\mathrm{L}}_{1,3} : 0.82$		C; $\hat{\lambda}^{\mathrm{L}}_{2,4} : 0.84$			
	complete data	0.826 (0.009)		0.838 (0.009)			
	66% cens. PB	0.824 (0.026)		0.835 (0.023)			
	66% cens. NPB	0.819 (0.028)		0.836 (0.023)			

Table B.10: Results for model FFF in Table 4.9: In each row, the 90% confidence intervals of the differences between the nonparametric bootstrap based parameter estimates $(\hat{\boldsymbol{\theta}}_{\text{sim}}^{(1)}, \ldots, \hat{\boldsymbol{\theta}}_{\text{sim}}^{(100)})$ with the true parameters underlying the D-vine copula model ($\boldsymbol{\theta}$) and with the simulation based parameter estimates ($\hat{\boldsymbol{\theta}}_{\text{sim}}$) are shown in the first and the second column, respectively. The third column contains the parameters estimated for the simulated data set compared to the mean of the bootstrap based parameter estimates in the fourth column. Four different seeds for data simulation are considered.

			[5%-quantile, 95%-quantile] of			
		$\boldsymbol{\theta}$	$\left(\hat{\boldsymbol{\theta}}_{\text{sim}}^{(1)}, \ldots, \hat{\boldsymbol{\theta}}_{\text{sim}}^{(100)}\right) - \boldsymbol{\theta}$	$\left(\hat{\boldsymbol{\theta}}_{\text{sim}}^{(1)}, \ldots, \hat{\boldsymbol{\theta}}_{\text{sim}}^{(100)}\right) - \hat{\boldsymbol{\theta}}_{\text{sim}}$	$\hat{\boldsymbol{\theta}}_{\text{sim}}$	mean $\left(\hat{\boldsymbol{\theta}}_{\text{sim}}^{(1)}, \ldots, \hat{\boldsymbol{\theta}}_{\text{sim}}^{(100)}\right)$
Simulation 1	$\theta_{1,3}$	6.38	[-1.835, 0.366]	[-0.944, 1.257]	5.489	5.613
	$\theta_{3,4}$:	6.34	[-1.703, 0.397]	[-0.960, 1.140]	5.594	5.582
	$\theta_{2,4}$:	6.77	[-0.774, 1.640]	[-1.015, 1.398]	7.007	7.062
	$\theta_{1,4;3}$:	1.67	[-0.679, 1.197]	[-0.868, 1.008]	1.858	1.874
	$\theta_{2,3;4}$:	2.81	[-0.652, 0.820]	[-0.615, 0.857]	2.773	2.867
	$\theta_{1,2;3,4}$:	3.72	[-1.825, 0.125]	[-0.935, 1.015]	2.828	2.828
Simulation 2	$\theta_{1,3}$:	6.38	[-1.188, 0.517]	[-0.670, 1.036]	5.862	5.962
	$\theta_{3,4}$:	6.34	[-2.160, -0.237]	[-0.814, 1.109]	4.991	5.056
	$\theta_{2,4}$:	6.77	[-1.283, 1.278]	[-0.963, 1.598]	6.446	6.587
	$\theta_{1,4;3}$:	1.67	[-0.705, 1.116]	[-0.944, 0.877]	1.907	1.873
	$\theta_{2,3;4}$:	2.81	[-1.057, 0.653]	[-0.722, 0.988]	2.474	2.616
	$\theta_{1,2;3,4}$:	3.72	[-1.464, 0.213]	[-0.802, 0.875]	3.055	3.084
Simulation 3	$\theta_{1,3}$:	6.38	[-0.952, 1.382]	[-1.090, 1.245]	6.518	6.554
	$\theta_{3,4}$:	6.34	[-1.608, 0.011]	[-0.679, 0.941]	5.407	5.434
	$\theta_{2,4}$:	6.77	[-1.012, 0.996]	[-1.008, 1.000]	6.762	6.778
	$\theta_{1,4;3}$:	1.67	[-0.102, 1.617]	[-0.784, 0.935]	2.351	2.387
	$\theta_{2,3;4}$:	2.81	[0.167, 2.145]	[-0.954, 1.024]	3.930	3.937
	$\theta_{1,2;3,4}$:	3.72	[-0.903, 0.698]	[-0.815, 0.786]	3.629	3.607
Simulation 4	$\theta_{1,3}$:	6.38	[-1.553, 0.798]	[-1.126, 1.225]	5.954	6.075
	$\theta_{3,4}$:	6.34	[-2.122, -0.221]	[-0.724, 1.177]	4.939	5.103
	$\theta_{2,4}$:	6.77	[-1.520, 0.609]	[-0.767, 1.363]	6.012	6.251
	$\theta_{1,4;3}$:	1.67	[0.197, 1.943]	[-0.963, 0.783]	2.828	2.728
	$\theta_{2,3;4}$:	2.81	[0.223, 2.229]	[-0.807, 1.198]	3.840	3.937
	$\theta_{1,2;3,4}$:	3.72	[-0.683, 1.082]	[-1.009, 0.756]	4.042	3.958

Table B.11: Results for model CFF in Table 4.9: In each row, the 90% confidence intervals of the differences between the nonparametric bootstrap based parameter estimates $(\hat{\boldsymbol{\theta}}_{\text{sim}}^{(1)}, \ldots, \hat{\boldsymbol{\theta}}_{\text{sim}}^{(100)})$ with the true parameters underlying the D-vine copula model ($\boldsymbol{\theta}$) and with the simulation based parameter estimates ($\hat{\boldsymbol{\theta}}_{\text{sim}}$) are shown in the first and the second column, respectively. The third column contains the parameters estimated for the simulated data set compared to the mean of the bootstrap based parameter estimates in the fourth column. Four different seeds for data simulation are considered.

			[5%-quantile, 95%-quantile] of			
		$\boldsymbol{\theta}$	$\left(\hat{\boldsymbol{\theta}}_{\text{sim}}^{(1)}, \ldots, \hat{\boldsymbol{\theta}}_{\text{sim}}^{(100)}\right) - \boldsymbol{\theta}$	$\left(\hat{\boldsymbol{\theta}}_{\text{sim}}^{(1)}, \ldots, \hat{\boldsymbol{\theta}}_{\text{sim}}^{(100)}\right) - \hat{\boldsymbol{\theta}}_{\text{sim}}$	$\hat{\boldsymbol{\theta}}_{\text{sim}}$	mean $\left(\hat{\boldsymbol{\theta}}_{\text{sim}}^{(1)}, \ldots, \hat{\boldsymbol{\theta}}_{\text{sim}}^{(100)}\right)$
Simulation 1	$\theta_{1,3}$:	3.60	[-0.935, 0.952]	[-0.983, 0.905]	3.645	3.598
	$\theta_{3,4}$:	6.34	[-1.433, 1.257]	[-1.281, 1.408]	6.185	6.303
	$\theta_{2,4}$:	6.77	[0.335, 2.869]	[-0.981, 1.553]	8.081	8.272
	$\theta_{1,4;3}$:	1.49	[-0.193, 1.315]	[-0.738, 0.770]	2.030	2.024
	$\theta_{2,3;4}$:	2.81	[-1.189, 0.503]	[-0.917, 0.775]	2.540	2.479
	$\theta_{1,2;3,4}$:	3.48	[0.053, 2.065]	[-0.778, 1.234]	4.316	4.463
Simulation 2	$\theta_{1,3}$:	3.60	[-1.166, 0.310]	[-0.665, 0.811]	3.096	3.116
	$\theta_{3,4}$:	6.34	[-1.901, -0.128]	[-0.735, 1.039]	5.170	5.251
	$\theta_{2,4}$:	6.77	[-1.428, 1.083]	[-1.025, 1.486]	6.362	6.500
	$\theta_{1,4;3}$:	1.49	[-0.699, 1.336]	[-0.982, 1.053]	1.769	1.737
	$\theta_{2,3;4}$:	2.81	[-1.130, 0.687]	[-0.833, 0.984]	2.515	2.552
	$\theta_{1,2;3,4}$:	3.48	[-1.725, -0.015]	[-0.889, 0.820]	2.649	2.680
Simulation 3	$\theta_{1,3}$:	3.60	[-1.365, 0.318]	[-0.719, 0.964]	2.951	3.055
	$\theta_{3,4}$:	6.34	[-2.632, -0.481]	[-0.945, 1.206]	4.650	4.738
	$\theta_{2,4}$:	6.77	[-2.205, 0.153]	[-1.080, 1.278]	5.640	5.721
	$\theta_{1,4;3}$:	1.49	[0.138, 1.569]	[-0.602, 0.830]	2.225	2.289
	$\theta_{2,3;4}$:	2.81	[-0.069, 2.062]	[-0.904, 1.227]	3.647	3.718
	$\theta_{1,2;3,4}$:	3.48	[-0.316, 1.615]	[-0.818, 1.113]	3.986	4.071
Simulation 4	$\theta_{1,3}$:	3.60	[-0.446, 1.899]	[-0.992, 1.353]	4.143	4.198
	$\theta_{3,4}$:	6.34	[-1.105, 1.649]	[-1.453, 1.301]	6.684	6.748
	$\theta_{2,4}$:	6.77	[-0.813, 1.763]	[-1.111, 1.466]	7.063	7.087
	$\theta_{1,4;3}$:	1.49	[0.090, 1.671]	[-0.734, 0.848]	2.309	2.348
	$\theta_{2,3;4}$:	2.81	[-1.429, 0.164]	[-0.657, 0.936]	2.040	2.185
	$\theta_{1,2;3,4}$:	3.48	[-0.567, 1.690]	[-0.941, 1.316]	3.858	4.004

Table B.12: Results for model FFC in Table 4.9: In each row, the 90% confidence intervals of the differences between the nonparametric bootstrap based parameter estimates $(\hat{\boldsymbol{\theta}}_{\text{sim}}^{(1)}, \ldots, \hat{\boldsymbol{\theta}}_{\text{sim}}^{(100)})$ with the true parameters underlying the D-vine copula model ($\boldsymbol{\theta}$) and with the simulation based parameter estimates ($\hat{\boldsymbol{\theta}}_{\text{sim}}$) are shown in the first and the second column, respectively. The third column contains the parameters estimated for the simulated data set compared to the mean of the bootstrap based parameter estimates in the fourth column. Four different seeds for data simulation are considered.

			[5%-quantile, 95%-quantile] of			
		$\boldsymbol{\theta}$	$\left(\hat{\boldsymbol{\theta}}_{\text{sim}}^{(1)}, \ldots, \hat{\boldsymbol{\theta}}_{\text{sim}}^{(100)}\right) - \boldsymbol{\theta}$	$\left(\hat{\boldsymbol{\theta}}_{\text{sim}}^{(1)}, \ldots, \hat{\boldsymbol{\theta}}_{\text{sim}}^{(100)}\right) - \hat{\boldsymbol{\theta}}_{\text{sim}}$	$\hat{\boldsymbol{\theta}}_{\text{sim}}$	mean $\left(\hat{\boldsymbol{\theta}}_{\text{sim}}^{(1)}, \ldots, \hat{\boldsymbol{\theta}}_{\text{sim}}^{(100)}\right)$
Simulation 1	$\theta_{1,3}$:	6.38	[-1.835, 0.366]	[-0.944, 1.257]	5.489	5.613
	$\theta_{3,4}$:	6.34	[-1.751, 0.271]	[-0.963, 1.059]	5.549	5.528
	$\theta_{2,4}$:	3.90	[-0.756, 1.037]	[-0.812, 0.981]	3.952	3.962
	$\theta_{1,4;3}$:	1.54	[-0.661, 1.063]	[-0.797, 0.927]	1.672	1.675
	$\theta_{2,3;4}$:	2.76	[-0.653, 0.841]	[-0.715, 0.779]	2.824	2.909
	$\theta_{1,2;3,4}$:	3.86	[-1.963, 0.100]	[-1.099, 0.964]	2.999	3.009
Simulation 2	$\theta_{1,3}$:	6.38	[-0.952, 1.382]	[-1.090, 1.245]	6.518	6.554
	$\theta_{3,4}$:	6.34	[-1.603, -0.027]	[-0.674, 0.902]	5.408	5.430
	$\theta_{2,4}$:	3.90	[-0.915, 0.613]	[-0.740, 0.788]	3.721	3.710
	$\theta_{1,4;3}$:	1.54	[-0.038, 1.616]	[-0.714, 0.940]	2.212	2.250
	$\theta_{2,3;4}$:	2.76	[0.226, 2.054]	[-0.867, 0.961]	3.856	3.844
	$\theta_{1,2;3,4}$:	3.86	[-0.998, 0.758]	[-0.936, 0.820]	3.801	3.781
Simulation 3	$\theta_{1,3}$:	6.38	[-0.495, 2.690]	[-1.399, 1.787]	7.284	7.420
	$\theta_{3,4}$:	6.34	[-1.290, 1.208]	[-1.000, 1.498]	6.047	6.155
	$\theta_{2,4}$:	3.90	[-0.638, 1.075]	[-0.795, 0.919]	4.053	4.064
	$\theta_{1,4;3}$:	1.54	[-1.526, -0.076]	[-0.736, 0.715]	0.746	0.689
	$\theta_{2,3;4}$:	2.76	[0.345, 2.104]	[-0.723, 1.036]	3.831	3.928
	$\theta_{1,2;3,4}$:	3.86	[-1.399, 0.716]	[-0.881, 1.235]	3.344	3.470
Simulation 4	$\theta_{1,3}$:	6.38	[-1.628, 0.815]	[-1.201, 1.242]	5.954	5.954
	$\theta_{3,4}$:	6.34	[-2.175, -0.157]	[-0.785, 1.234]	4.946	5.042
	$\theta_{2,4}$:	3.90	[-1.226, 0.120]	[-0.659, 0.688]	3.329	3.343
	$\theta_{1,4;3}$:	1.54	[0.314, 1.948]	[-0.838, 0.795]	2.689	2.669
	$\theta_{2,3;4}$:	2.76	[0.404, 2.202]	[-0.718, 1.080]	3.885	3.986
	$\theta_{1,2;3,4}$:	3.86	[-0.443, 1.465]	[-0.778, 1.129]	4.198	4.226

Table B.13: Results for model CFC in Table 4.9: In each row, the 90% confidence intervals of the differences between the nonparametric bootstrap based parameter estimates $(\hat{\boldsymbol{\theta}}_{\text{sim}}^{(1)}, \ldots, \hat{\boldsymbol{\theta}}_{\text{sim}}^{(100)})$ with the true parameters underlying the D-vine copula model ($\boldsymbol{\theta}$) and with the simulation based parameter estimates ($\hat{\boldsymbol{\theta}}_{\text{sim}}$) are shown in the first and the second column, respectively. The third column contains the parameters estimated for the simulated data set compared to the mean of the bootstrap based parameter estimates in the fourth column. Four different seeds for data simulation are considered.

			[5%-quantile, 95%-quantile] of			
		$\boldsymbol{\theta}$	$\left(\hat{\boldsymbol{\theta}}_{\text{sim}}^{(1)}, \ldots, \hat{\boldsymbol{\theta}}_{\text{sim}}^{(100)}\right) - \boldsymbol{\theta}$	$\left(\hat{\boldsymbol{\theta}}_{\text{sim}}^{(1)}, \ldots, \hat{\boldsymbol{\theta}}_{\text{sim}}^{(100)}\right) - \hat{\boldsymbol{\theta}}_{\text{sim}}$	$\hat{\boldsymbol{\theta}}_{\text{sim}}$	mean $\left(\hat{\boldsymbol{\theta}}_{\text{sim}}^{(1)}, \ldots, \hat{\boldsymbol{\theta}}_{\text{sim}}^{(100)}\right)$
Simulation 1	$\theta_{1,3}$:	3.60	[-0.757, 1.064]	[-1.134, 0.687]	3.974	3.795
	$\theta_{3,4}$:	6.34	[-1.077, 1.645]	[-1.097, 1.625]	6.357	6.544
	$\theta_{2,4}$:	3.90	[-1.079, 0.904]	[-0.977, 1.005]	3.795	3.807
	$\theta_{1,4;3}$:	1.36	[-1.250, 0.390]	[-0.568, 1.072]	0.682	0.807
	$\theta_{2,3;4}$:	2.71	[-0.377, 1.652]	[-1.087, 0.943]	3.416	3.431
	$\theta_{1,2;3,4}$:	3.70	[-1.212, 0.621]	[-0.949, 0.884]	3.439	3.428
Simulation 2	$\theta_{1,3}$:	3.60	[-1.500, 0.155]	[-0.748, 0.906]	2.846	2.913
	$\theta_{3,4}$:	6.34	[-1.392, 0.592]	[-0.779, 1.205]	5.723	5.748
	$\theta_{2,4}$:	3.90	[-0.823, 1.112]	[-0.957, 0.978]	4.030	4.060
	$\theta_{1,4;3}$:	1.36	[-0.668, 1.241]	[-0.934, 0.975]	1.631	1.626
	$\theta_{2,3;4}$:	2.71	[-1.015, 0.473]	[-0.682, 0.806]	2.374	2.482
	$\theta_{1,2;3,4}$:	3.70	[-2.054, 0.061]	[-1.110, 1.005]	2.758	2.716
Simulation 3	$\theta_{1,3}$:	3.60	[-0.987, 0.578]	[-0.734, 0.831]	3.344	3.345
	$\theta_{3,4}$:	6.34	[-1.547, 0.128]	[-0.705, 0.970]	5.494	5.508
	$\theta_{2,4}$:	3.90	[-0.819, 0.771]	[-0.781, 0.809]	3.859	3.841
	$\theta_{1,4;3}$:	1.36	[0.164, 1.656]	[-0.575, 0.917]	2.103	2.153
	$\theta_{2,3;4}$:	2.71	[0.157, 2.001]	[-0.919, 0.925]	3.783	3.773
	$\theta_{1,2;3,4}$:	3.70	[-1.079, 0.720]	[-0.894, 0.905]	3.517	3.508
Simulation 4	$\theta_{1,3}$:	3.60	[-0.446, 1.899]	[-0.992, 1.353]	4.143	4.198
	$\theta_{3,4}$:	6.34	[-1.041, 1.709]	[-1.436, 1.315]	6.731	6.794
	$\theta_{2,4}$:	3.90	[-0.575, 1.235]	[-0.796, 1.015]	4.117	4.110
	$\theta_{1,4;3}$:	1.36	[0.108, 1.662]	[-0.685, 0.869]	2.157	2.225
	$\theta_{2,3;4}$:	2.71	[-1.431, 0.183]	[-0.738, 0.875]	2.014	2.038
	$\theta_{1,2;3,4}$:	3.70	[-0.851, 1.388]	[-0.974, 1.265]	3.824	3.926

B.4 Additional simulation results for Section 4.4.4

Table B.14: Simulation results using one-stage parametric estimation for three-dimensional data. A Clayton (3dC) copula (top panel right) and a Gumbel (3dG) copula (bottom panel right) each with Kendall's $\tau = 0.5$ are considered. A D-vine copula including Clayton copulas (top panel left), respectively Gumbel copulas (bottom panel left), with $\tau_{1,2} = \tau_{2,3} = 0.5$ in $\mathcal{T}_1$ and a Frank (F) copula with $\tau_{1,3;2} = 0.25$ in $\mathcal{T}_2$ is considered. For the D-vine copulas global and sequential likelihood estimation is reported. The **empirical mean (empirical standard deviation)** for the **Kendall's τ estimates** are presented based on 250 replications and samples of size 250 and 500 affected by either 15%, 30% or heavy tail 30% right-censoring.

			D-vine copula model			Archimedean copula
			C; $\tau_{1,2}$: 0.50	C; $\tau_{2,3}$: 0.50	F; $\tau_{1,3;2}$: 0.25	3dC; τ : 0.50
Global	15%	250	0.502 (0.035)	0.505 (0.041)	0.250 (0.052)	0.503 (0.033)
		500	0.501 (0.027)	0.503 (0.029)	0.251 (0.036)	0.501 (0.024)
		1000	0.500 (0.019)	0.500 (0.020)	0.250 (0.024)	
	30%	250	0.501 (0.049)	0.504 (0.058)	0.250 (0.068)	0.505 (0.041)
		500	0.501 (0.033)	0.505 (0.041)	0.251 (0.046)	0.501 (0.029)
		1000	0.501 (0.024)	0.501 (0.029)	0.250 (0.034)	
	30% HT	250	0.503 (0.059)	0.502 (0.083)	0.247 (0.080)	0.504 (0.051)
		500	0.503 (0.041)	0.501 (0.057)	0.249 (0.053)	0.500 (0.038)
		1000	0.502 (0.028)	0.499 (0.038)	0.248 (0.037)	
Sequential	15%	250	0.501 (0.036)	0.505 (0.042)	0.250 (0.052)	
		500	0.501 (0.028)	0.503 (0.030)	0.251 (0.036)	
		1000	0.500 (0.019)	0.500 (0.020)	0.250 (0.024)	
	30%	250	0.500 (0.049)	0.504 (0.059)	0.250 (0.068)	
		500	0.501 (0.033)	0.505 (0.041)	0.251 (0.046)	
		1000	0.501 (0.025)	0.500 (0.029)	0.250 (0.034)	
	30% HT	250	0.503 (0.060)	0.502 (0.081)	0.247 (0.080)	
		500	0.503 (0.041)	0.501 (0.057)	0.249 (0.053)	
		1000	0.502 (0.029)	0.499 (0.038)	0.248 (0.037)	
			G; $\tau_{1,2}$: 0.50	G; $\tau_{2,3}$: 0.50	F; $\tau_{1,3;2}$: 0.25	3dG; τ : 0.50
Global	15%	250	0.498 (0.033)	0.501 (0.036)	0.251 (0.050)	0.500 (0.030)
		500	0.499 (0.027)	0.501 (0.027)	0.250 (0.035)	0.501 (0.021)
		1000	0.500 (0.018)	0.500 (0.019)	0.250 (0.024)	
	30%	250	0.501 (0.039)	0.503 (0.044)	0.249 (0.066)	0.504 (0.035)
		500	0.502 (0.030)	0.504 (0.033)	0.251 (0.046)	0.502 (0.025)
		1000	0.502 (0.021)	0.502 (0.023)	0.249 (0.036)	
	30% HT	250	0.507 (0.042)	0.507 (0.046)	0.245 (0.077)	0.504 (0.040)
		500	0.506 (0.031)	0.506 (0.035)	0.248 (0.048)	0.503 (0.029)
		1000	0.505 (0.022)	0.504 (0.025)	0.248 (0.038)	
Sequential	15%	250	0.498 (0.034)	0.501 (0.035)	0.250 (0.050)	
		500	0.499 (0.027)	0.501 (0.027)	0.250 (0.035)	
		1000	0.499 (0.019)	0.500 (0.019)	0.250 (0.024)	
	30%	250	0.501 (0.039)	0.503 (0.044)	0.249 (0.066)	
		500	0.500 (0.030)	0.503 (0.033)	0.252 (0.046)	
		1000	0.500 (0.022)	0.501 (0.023)	0.249 (0.036)	
	30% HT	250	0.499 (0.045)	0.502 (0.047)	0.247 (0.078)	
		500	0.501 (0.033)	0.502 (0.036)	0.249 (0.049)	
		1000	0.500 (0.023)	0.501 (0.025)	0.249 (0.038)	

Table B.15: Simulation results using one-stage parametric estimation for three-dimensional data. A Clayton (3dC) copula (top panel right) and a Gumbel (3dG) copula (bottom panel right) each with Kendall's $\tau = 0.5$ are considered. A D-vine copula including Clayton copulas (top panel left), respectively Gumbel copulas (bottom panel left), with $\tau_{1,2} = \tau_{2,3} = 0.5$ in $\mathcal{T}_1$ and a Frank (F) copula with $\tau_{1,3;2} = 0.25$ in $\mathcal{T}_2$ is considered. For the D-vine copulas global and sequential likelihood estimation is reported. The **empirical mean (empirical standard deviation)** for the **copula parameter estimates** are presented based on 250 replications and samples of size 250 and 500 affected by either 15%, 30% or heavy tail 30% right-censoring.

			D-vine copula model			Archimedean copula
			C; $\theta_{1,2} : 2.00$	C; $\theta_{2,3} : 2.00$	F; $\theta_{1,3;2} : 2.37$	3dC; $\theta : 2.00$
Global	15%	250	2.036 (0.292)	2.071 (0.353)	2.396 (0.554)	2.039 (0.266)
		500	2.019 (0.218)	2.039 (0.242)	2.389 (0.375)	2.020 (0.198)
		1000	2.009 (0.149)	2.010 (0.162)	2.376 (0.257)	2.016 (0.119)
	30%	250	2.044 (0.412)	2.091 (0.503)	2.407 (0.722)	2.069 (0.345)
		500	2.030 (0.269)	2.068 (0.341)	2.398 (0.489)	2.018 (0.233)
		1000	2.016 (0.197)	2.019 (0.233)	2.384 (0.361)	2.023 (0.171)
	30% HT	250	2.082 (0.494)	2.121 (0.670)	2.388 (0.870)	2.078 (0.424)
		500	2.054 (0.339)	2.061 (0.471)	2.380 (0.563)	2.019 (0.307)
		1000	2.030 (0.226)	2.012 (0.302)	2.364 (0.389)	2.012 (0.203)
Sequential	15%	250	2.033 (0.293)	2.067 (0.359)	2.391 (0.551)	
		500	2.020 (0.224)	2.039 (0.247)	2.387 (0.375)	
		1000	2.008 (0.151)	2.009 (0.164)	2.375 (0.257)	
	30%	250	2.042 (0.415)	2.088 (0.506)	2.401 (0.721)	
		500	2.028 (0.273)	2.066 (0.343)	2.395 (0.488)	
		1000	2.014 (0.198)	2.017 (0.233)	2.383 (0.361)	
	30% HT	250	2.084 (0.499)	2.122 (0.665)	2.387 (0.867)	
		500	2.055 (0.340)	2.060 (0.472)	2.377 (0.563)	
		1000	2.030 (0.228)	2.011 (0.301)	2.363 (0.390)	
			G; $\theta_{1,2} : 2.00$	G; $\theta_{2,3} : 2.00$	F; $\theta_{1,3;2} : 2.37$	3dG; $\theta : 2.00$
Global	15%	250	2.001 (0.131)	2.014 (0.145)	2.397 (0.536)	2.006 (0.119)
		500	2.003 (0.108)	2.009 (0.109)	2.381 (0.374)	2.005 (0.084)
		1000	2.001 (0.072)	2.003 (0.077)	2.371 (0.256)	2.008 (0.055)
	30%	250	2.015 (0.159)	2.030 (0.179)	2.387 (0.705)	2.024 (0.143)
		500	2.016 (0.120)	2.024 (0.134)	2.401 (0.485)	2.013 (0.102)
		1000	2.012 (0.083)	2.011 (0.091)	2.368 (0.378)	2.007 (0.068)
	30% HT	250	2.041 (0.173)	2.047 (0.191)	2.357 (0.840)	2.029 (0.165)
		500	2.033 (0.128)	2.035 (0.149)	2.369 (0.509)	2.019 (0.117)
		1000	2.022 (0.090)	2.019 (0.100)	2.357 (0.397)	2.010 (0.082)
Sequential	15%	250	2.001 (0.133)	2.016 (0.145)	2.393 (0.535)	
		500	2.003 (0.110)	2.012 (0.109)	2.381 (0.373)	
		1000	2.001 (0.074)	2.004 (0.077)	2.371 (0.256)	
	30%	250	2.015 (0.159)	2.030 (0.179)	2.387 (0.705)	
		500	2.008 (0.122)	2.019 (0.135)	2.406 (0.484)	
		1000	2.004 (0.085)	2.006 (0.092)	2.372 (0.379)	
	30% HT	250	2.012 (0.181)	2.027 (0.192)	2.383 (0.851)	
		500	2.011 (0.132)	2.020 (0.150)	2.382 (0.513)	
		1000	2.006 (0.094)	2.008 (0.101)	2.368 (0.399)	

Table B.16: Simulation results using one-stage parametric estimation for three-dimensional data. A **Clayton** copula (top panel) with Kendall's $\tau = 0.5$ and a D-vine copula including Clayton copulas (bottom panel) with $\tau_{1,2} = \tau_{2,3} = 0.5$ in $\mathcal{T}_1$ and a Frank copula with $\tau_{1,3;2} = 0.25$ in $\mathcal{T}_2$ is considered. For the D-vine copula global and sequential likelihood estimation is reported. The **empirical mean (empirical standard deviation)** for the **marginal parameter estimates** are presented based on 250 replications and samples of size 250 and 500 affected by either 15%, 30% or heavy tail 30% right-censoring.

			3d Clayton copula					
			$\lambda_1 : 0.50$	$\rho_1 : 1.50$	$\lambda_2 : 1.00$	$\rho_2 : 1.50$	$\lambda_3 : 1.00$	$\rho_3 : 1.50$
Global	15%	250	0.495 (0.042)	1.510 (0.072)	0.994 (0.077)	1.512 (0.083)	0.997 (0.079)	1.517 (0.088)
		500	0.500 (0.032)	1.504 (0.052)	1.001 (0.055)	1.505 (0.055)	1.001 (0.055)	1.501 (0.058)
		1000	0.499 (0.021)	1.502 (0.039)	0.999 (0.038)	1.502 (0.039)	0.999 (0.041)	1.502 (0.044)
	30%	250	0.493 (0.041)	1.512 (0.083)	0.995 (0.094)	1.518 (0.114)	1.001 (0.101)	1.525 (0.119)
		500	0.498 (0.033)	1.509 (0.064)	1.006 (0.066)	1.507 (0.069)	1.006 (0.072)	1.498 (0.076)
		1000	0.498 (0.021)	1.504 (0.048)	0.997 (0.047)	1.505 (0.049)	1.000 (0.054)	1.503 (0.059)
	30% HT	250	0.494 (0.041)	1.516 (0.092)	0.995 (0.119)	1.516 (0.130)	1.014 (0.154)	1.526 (0.146)
		500	0.499 (0.033)	1.508 (0.070)	1.010 (0.086)	1.510 (0.079)	1.013 (0.109)	1.503 (0.103)
		1000	0.499 (0.021)	1.504 (0.048)	0.999 (0.054)	1.502 (0.060)	1.001 (0.076)	1.500 (0.069)
			Clayton based D-vine model					
			$\lambda_1 : 0.50$	$\rho_1 : 1.50$	$\lambda_2 : 1.00$	$\rho_2 : 1.50$	$\lambda_3 : 1.00$	$\rho_3 : 1.50$
Global	15%	250	0.496 (0.044)	1.516 (0.076)	0.998 (0.081)	1.512 (0.079)	1.004 (0.082)	1.512 (0.087)
		500	0.498 (0.029)	1.508 (0.049)	0.999 (0.055)	1.503 (0.058)	1.000 (0.056)	1.505 (0.062)
		1000	0.493 (0.017)	1.507 (0.040)	0.977 (0.016)	1.502 (0.040)	0.992 (0.033)	1.507 (0.041)
	30%	250	0.497 (0.048)	1.518 (0.084)	1.007 (0.102)	1.521 (0.102)	1.017 (0.132)	1.519 (0.123)
		500	0.498 (0.030)	1.507 (0.059)	0.999 (0.065)	1.502 (0.069)	1.001 (0.080)	1.509 (0.084)
		1000	0.495 (0.020)	1.506 (0.044)	0.972 (0.023)	1.494 (0.047)	0.991 (0.055)	1.501 (0.054)
	30% HT	250	0.497 (0.046)	1.517 (0.091)	1.001 (0.122)	1.509 (0.128)	1.035 (0.218)	1.519 (0.159)
		500	0.498 (0.030)	1.505 (0.063)	0.994 (0.091)	1.496 (0.085)	1.018 (0.139)	1.510 (0.106)
		1000	0.500 (0.020)	1.503 (0.048)	0.997 (0.060)	1.499 (0.059)	1.010 (0.082)	1.502 (0.070)
Sequential	15%	250	0.497 (0.046)	1.517 (0.085)	0.999 (0.082)	1.513 (0.080)	1.005 (0.083)	1.512 (0.088)
		500	0.498 (0.030)	1.508 (0.054)	0.999 (0.055)	1.503 (0.061)	1.000 (0.057)	1.505 (0.062)
		1000	0.500 (0.021)	1.503 (0.041)	0.999 (0.037)	1.501 (0.042)	1.005 (0.038)	1.503 (0.042)
	30%	250	0.497 (0.048)	1.518 (0.090)	1.008 (0.103)	1.522 (0.102)	1.017 (0.133)	1.519 (0.124)
		500	0.498 (0.031)	1.508 (0.061)	0.999 (0.066)	1.502 (0.070)	1.001 (0.080)	1.510 (0.085)
		1000	0.500 (0.022)	1.504 (0.045)	0.997 (0.048)	1.500 (0.050)	1.005 (0.058)	1.501 (0.055)
	30% HT	250	0.497 (0.047)	1.513 (0.091)	1.002 (0.124)	1.508 (0.128)	1.033 (0.211)	1.517 (0.157)
		500	0.498 (0.030)	1.504 (0.064)	0.994 (0.092)	1.496 (0.084)	1.019 (0.139)	1.510 (0.107)
		1000	0.500 (0.020)	1.503 (0.049)	0.997 (0.060)	1.499 (0.060)	1.011 (0.082)	1.502 (0.070)

Table B.17: Simulation results using one-stage parametric estimation for three-dimensional data. A **Gumbel** copula (top panel) with Kendall's $\tau = 0.5$ and a D-vine copula including Gumbel copulas (bottom panel) with $\tau_{1,2} = \tau_{2,3} = 0.5$ in $\mathcal{T}_1$ and a Frank copula with $\tau_{1,3;2} = 0.25$ in $\mathcal{T}_2$ is considered. For the D-vine copula global and sequential likelihood estimation is reported. The **empirical mean (empirical standard deviation)** for the **marginal parameter estimates** are presented based on 250 replications and samples of size 250 and 500 affected by either 15%, 30% or heavy tail 30% right-censoring.

			3d Gumbel copula					
			$\lambda_1 : 0.50$	$\rho_1 : 1.50$	$\lambda_2 : 1.00$	$\rho_2 : 1.50$	$\lambda_3 : 1.00$	$\rho_3 : 1.50$
Global	15%	250	0.500 (0.043)	1.508 (0.084)	1.004 (0.084)	1.513 (0.079)	1.004 (0.080)	1.501 (0.085)
		500	0.498 (0.031)	1.505 (0.054)	1.002 (0.057)	1.503 (0.055)	0.989 (0.059)	1.496 (0.062)
		1000	0.500 (0.020)	1.498 (0.035)	0.997 (0.038)	1.497 (0.038)	1.000 (0.044)	1.496 (0.042)
	30%	250	0.501 (0.043)	1.506 (0.090)	1.002 (0.097)	1.509 (0.099)	1.003 (0.112)	1.498 (0.103)
		500	0.500 (0.032)	1.505 (0.060)	1.006 (0.068)	1.505 (0.070)	0.991 (0.084)	1.497 (0.076)
		1000	0.500 (0.022)	1.500 (0.039)	1.000 (0.049)	1.500 (0.049)	1.000 (0.055)	1.498 (0.053)
	30% HT	250	0.501 (0.043)	1.507 (0.095)	0.997 (0.111)	1.507 (0.116)	1.016 (0.161)	1.503 (0.132)
		500	0.500 (0.030)	1.503 (0.067)	1.006 (0.082)	1.502 (0.082)	0.995 (0.124)	1.495 (0.098)
		1000	0.501 (0.021)	1.498 (0.042)	1.003 (0.055)	1.498 (0.061)	1.010 (0.086)	1.500 (0.070)
			Gumbel based D-vine model					
			$\lambda_1 : 0.50$	$\rho_1 : 1.50$	$\lambda_2 : 1.00$	$\rho_2 : 1.50$	$\lambda_3 : 1.00$	$\rho_3 : 1.50$
Global	15%	250	0.485 (0.038)	1.525 (0.081)	0.961 (0.038)	1.517 (0.085)	0.987 (0.078)	1.526 (0.094)
		500	0.489 (0.025)	1.515 (0.053)	0.969 (0.027)	1.507 (0.058)	0.985 (0.049)	1.514 (0.061)
		1000	0.493 (0.017)	1.507 (0.040)	0.977 (0.016)	1.502 (0.040)	0.992 (0.033)	1.507 (0.041)
	30%	250	0.486 (0.041)	1.521 (0.089)	0.954 (0.048)	1.510 (0.097)	0.991 (0.133)	1.522 (0.130)
		500	0.491 (0.028)	1.511 (0.060)	0.963 (0.035)	1.496 (0.069)	0.984 (0.085)	1.508 (0.082)
		1000	0.495 (0.020)	1.506 (0.044)	0.972 (0.023)	1.494 (0.047)	0.991 (0.055)	1.501 (0.054)
	30% HT	250	0.490 (0.042)	1.510 (0.091)	0.948 (0.056)	1.488 (0.103)	1.014 (0.197)	1.517 (0.152)
		500	0.493 (0.028)	1.503 (0.063)	0.955 (0.048)	1.481 (0.071)	0.993 (0.128)	1.501 (0.098)
		1000	0.496 (0.019)	1.501 (0.048)	0.968 (0.030)	1.486 (0.054)	0.990 (0.083)	1.496 (0.066)
Sequential	15%	250	0.497 (0.046)	1.517 (0.085)	0.999 (0.079)	1.516 (0.086)	1.009 (0.091)	1.518 (0.095)
		500	0.489 (0.025)	1.515 (0.053)	0.969 (0.027)	1.507 (0.058)	0.985 (0.049)	1.514 (0.061)
		1000	0.500 (0.021)	1.503 (0.041)	0.999 (0.037)	1.501 (0.042)	1.005 (0.038)	1.503 (0.042)
	30%	250	0.498 (0.047)	1.518 (0.090)	1.007 (0.102)	1.521 (0.099)	1.015 (0.142)	1.518 (0.123)
		500	0.499 (0.031)	1.507 (0.061)	0.999 (0.070)	1.503 (0.071)	1.002 (0.093)	1.505 (0.080)
		1000	0.500 (0.022)	1.504 (0.045)	0.997 (0.048)	1.500 (0.050)	1.005 (0.058)	1.501 (0.055)
	30% HT	250	0.497 (0.047)	1.515 (0.093)	1.006 (0.117)	1.513 (0.112)	1.046 (0.211)	1.528 (0.153)
		500	0.498 (0.030)	1.504 (0.064)	0.996 (0.088)	1.499 (0.080)	1.016 (0.137)	1.509 (0.100)
		1000	0.500 (0.020)	1.503 (0.049)	0.998 (0.060)	1.499 (0.060)	1.007 (0.087)	1.501 (0.067)

B.5 Additional simulation results for Section 4.4.5

Table B.18: Simulation results using two-stage semiparametric estimation for three-dimensional data. A Clayton (3dC) copula (top panel right) and a Gumbel (3dG) copula (bottom panel right) with Kendall's $\tau = 0.5$ is considered. A D-vine copula including Clayton copulas (top panel left), respectively Gumbel copulas (bottom panel left), with $\tau_{1,2} = \tau_{2,3} = 0.5$ in $\mathcal{T}_1$ and a Frank (F) copula with $\tau_{1,3;2} = 0.25$ in $\mathcal{T}_2$ is considered. For the D-vine copulas global and sequential likelihood estimation is reported. The **empirical mean (empirical standard deviation)** for the **Kendall's τ estimates** are presented based on 250 replications and samples of size 250, 500 and 1000 affected by either 15%, 30% or heavy tail 30% right-censoring.

			D-vine copula model			Archimedean copula
			C; $\tau_{1,2} : 0.50$	C; $\tau_{2,3} : 0.50$	F; $\tau_{1,3;2} : 0.25$	3dC; $\tau : 0.50$
Global	15%	250	0.495 (0.042)	0.497 (0.046)	0.253 (0.053)	0.496 (0.040)
		500	0.496 (0.034)	0.498 (0.035)	0.253 (0.037)	0.497 (0.028)
		1000	0.498 (0.023)	0.498 (0.023)	0.251 (0.025)	0.499 (0.019)
	30%	250	0.504 (0.071)	0.505 (0.074)	0.249 (0.070)	0.504 (0.061)
		500	0.501 (0.044)	0.504 (0.049)	0.253 (0.047)	0.498 (0.042)
		1000	0.503 (0.035)	0.502 (0.037)	0.251 (0.035)	0.498 (0.031)
	30% HT	250	0.558 (0.110)	0.556 (0.101)	0.245 (0.088)	0.546 (0.094)
		500	0.558 (0.089)	0.549 (0.092)	0.246 (0.061)	0.543 (0.081)
		1000	0.551 (0.069)	0.536 (0.072)	0.242 (0.042)	0.538 (0.066)
Sequential	15%	250	0.495 (0.042)	0.497 (0.045)	0.252 (0.053)	
		500	0.497 (0.034)	0.498 (0.035)	0.253 (0.037)	
		1000	0.498 (0.023)	0.498 (0.023)	0.251 (0.025)	
	30%	250	0.504 (0.070)	0.511 (0.071)	0.246 (0.069)	
		500	0.502 (0.043)	0.510 (0.048)	0.251 (0.046)	
		1000	0.503 (0.035)	0.507 (0.035)	0.250 (0.035)	
	30% HT	250	0.558 (0.107)	0.564 (0.089)	0.242 (0.086)	
		500	0.557 (0.087)	0.556 (0.083)	0.243 (0.060)	
		1000	0.550 (0.067)	0.544 (0.064)	0.240 (0.041)	
			G; $\tau_{1,2} : 0.50$	G; $\tau_{2,3} : 0.50$	F; $\tau_{1,3;2} : 0.25$	3dG; $\tau : 0.50$
Global	15%	250	0.501 (0.038)	0.506 (0.039)	0.251 (0.051)	0.504 (0.033)
		500	0.503 (0.028)	0.505 (0.028)	0.251 (0.036)	0.503 (0.022)
		1000	0.501 (0.019)	0.502 (0.020)	0.250 (0.025)	0.504 (0.015)
	30%	250	0.503 (0.049)	0.511 (0.049)	0.249 (0.067)	0.511 (0.042)
		500	0.507 (0.033)	0.510 (0.037)	0.254 (0.047)	0.508 (0.028)
		1000	0.505 (0.024)	0.505 (0.025)	0.249 (0.037)	0.506 (0.020)
	30% HT	250	0.535 (0.065)	0.531 (0.058)	0.251 (0.087)	0.534 (0.063)
		500	0.536 (0.053)	0.527 (0.052)	0.253 (0.061)	0.531 (0.048)
		1000	0.533 (0.039)	0.523 (0.035)	0.249 (0.044)	0.526 (0.039)
Sequential	15%	250	0.501 (0.038)	0.505 (0.039)	0.250 (0.051)	
		500	0.503 (0.028)	0.504 (0.028)	0.251 (0.036)	
		1000	0.501 (0.020)	0.501 (0.021)	0.250 (0.025)	
	30%	250	0.504 (0.049)	0.515 (0.050)	0.247 (0.067)	
		500	0.507 (0.033)	0.513 (0.037)	0.253 (0.046)	
		1000	0.505 (0.024)	0.507 (0.025)	0.249 (0.036)	
	30% HT	250	0.535 (0.065)	0.531 (0.058)	0.251 (0.087)	
		500	0.537 (0.053)	0.533 (0.050)	0.251 (0.059)	
		1000	0.534 (0.039)	0.528 (0.033)	0.247 (0.043)	

Table B.19: Simulation results using two-stage semiparametric estimation for three-dimensional data. A Clayton (3dC) copula (top panel right) and a Gumbel (3dG) copula (bottom panel right) with Kendall's $\tau = 0.5$ is considered. A D-vine copula including Clayton copulas (top panel left), respectively Gumbel copulas (bottom panel left), with $\tau_{1,2} = \tau_{2,3} = 0.5$ in $\mathcal{T}_1$ and a Frank (F) copula with $\tau_{1,3;2} = 0.25$ in $\mathcal{T}_2$ is considered. For the D-vine copulas global and sequential likelihood estimation is reported. The **empirical mean (empirical standard deviation)** for the **copula parameter estimates** are presented based on 250 replications and samples of size 250, 500 and 1000 affected by either 15%, 30% or heavy tail 30% right-censoring.

			D-vine copula model			Archimedean copula
			C; $\theta_{1,2} : 2.00$	C; $\theta_{2,3} : 2.00$	F; $\theta_{1,3;2} : 2.37$	3dC; $\theta : 2.00$
Global	15%	250	1.985 (0.329)	2.011 (0.370)	2.422 (0.567)	1.994 (0.315)
		500	1.990 (0.274)	2.003 (0.288)	2.416 (0.388)	1.984 (0.225)
		1000	1.991 (0.180)	1.994 (0.188)	2.384 (0.265)	1.993 (0.148)
	30%	250	2.114 (0.601)	2.131 (0.617)	2.392 (0.747)	2.090 (0.511)
		500	2.041 (0.359)	2.075 (0.414)	2.420 (0.494)	2.012 (0.338)
		1000	2.041 (0.288)	2.035 (0.299)	2.393 (0.375)	1.996 (0.245)
	30% HT	250	2.785 (1.094)	2.736 (1.062)	2.377 (0.959)	2.590 (0.923)
		500	2.701 (0.881)	2.599 (0.836)	2.352 (0.653)	2.505 (0.752)
		1000	2.554 (0.662)	2.402 (0.625)	2.302 (0.439)	2.409 (0.601)
Sequential	15%	250	1.990 (0.327)	2.009 (0.368)	2.414 (0.564)	
		500	1.995 (0.272)	2.003 (0.284)	2.411 (0.387)	
		1000	1.993 (0.178)	1.992 (0.186)	2.382 (0.265)	
	30%	250	2.116 (0.596)	2.176 (0.607)	2.365 (0.737)	
		500	2.046 (0.356)	2.123 (0.415)	2.400 (0.487)	
		1000	2.041 (0.285)	2.079 (0.289)	2.379 (0.372)	
	30% HT	250	2.769 (1.072)	2.773 (0.956)	2.336 (0.939)	
		500	2.682 (0.861)	2.641 (0.767)	2.326 (0.640)	
		1000	2.541 (0.646)	2.463 (0.568)	2.279 (0.428)	
			G; $\theta_{1,2} : 2.00$	G; $\theta_{2,3} : 2.00$	F; $\theta_{1,3;2} : 2.37$	3dG; $\theta : 2.00$
Global	15%	250	2.014 (0.152)	2.036 (0.162)	2.397 (0.552)	2.026 (0.135)
		500	2.018 (0.114)	2.028 (0.114)	2.389 (0.381)	2.018 (0.092)
		1000	2.009 (0.079)	2.012 (0.084)	2.371 (0.263)	2.016 (0.060)
	30%	250	2.034 (0.205)	2.067 (0.206)	2.395 (0.731)	2.061 (0.177)
		500	2.035 (0.135)	2.051 (0.156)	2.429 (0.496)	2.037 (0.116)
		1000	2.025 (0.098)	2.024 (0.103)	2.375 (0.388)	2.029 (0.080)
	30% HT	250	2.191 (0.299)	2.164 (0.261)	2.442 (0.963)	2.184 (0.281)
		500	2.183 (0.232)	2.139 (0.224)	2.437 (0.679)	2.154 (0.206)
		1000	2.157 (0.177)	2.105 (0.152)	2.372 (0.468)	2.122 (0.167)
Sequential	15%	250	2.014 (0.152)	2.033 (0.162)	2.392 (0.548)	
		500	2.018 (0.115)	2.022 (0.116)	2.389 (0.381)	
		1000	2.008 (0.079)	2.007 (0.085)	2.373 (0.264)	
	30%	250	2.034 (0.205)	2.084 (0.211)	2.373 (0.718)	
		500	2.036 (0.134)	2.065 (0.159)	2.416 (0.491)	
		1000	2.024 (0.098)	2.035 (0.106)	2.366 (0.385)	
	30% HT	250	2.191 (0.299)	2.164 (0.261)	2.442 (0.963)	
		500	2.186 (0.233)	2.165 (0.221)	2.409 (0.653)	
		1000	2.158 (0.177)	2.131 (0.149)	2.350 (0.457)	

B.6 Additional simulation results for Section 4.4.7

Table B.20: Simulation results using global and sequential one-stage parametric and two-stage semiparametric estimation for four-dimensional data. The D-vine copula model (**Setting 1 in Table 4.15**) captures tail-behavior for subsequent gap times changing from lower tail-dependence (Clayton (C)) over no tail-dependence (Frank (F)) to upper tail-dependence (Gumbel (G)) with same overall dependence of Kendall's $\tau_{1,2} = \tau_{2,3} = \tau_{3,4} = 0.5$. The **empirical mean (empirical standard deviation)** of the **Kendall's τ estimates** are presented based on 250 replications and samples of different sizes affected by either 15%, 30% or heavy tail 30% right-censoring.

				D-vine copula model					
				C; $\tau_{1,2}$: 0.500	F; $\tau_{2,3}$: 0.500	G; $\tau_{3,4}$: 0.500	F; $\tau_{1,3;2}$: 0.250	F; $\tau_{2,4;3}$: 0.250	F; $\tau_{1,4;2,3}$: 0.167
One-stage parametric	global	15%	250	0.505 (0.033)	0.504 (0.035)	0.501 (0.036)	0.253 (0.050)	0.252 (0.055)	0.169 (0.061)
			500	0.505 (0.021)	0.503 (0.024)	0.500 (0.027)	0.254 (0.034)	0.250 (0.037)	0.165 (0.040)
			1000	0.503 (0.015)	0.500 (0.017)	0.500 (0.020)	0.250 (0.023)	0.252 (0.025)	0.166 (0.028)
		30%	250	0.508 (0.043)	0.510 (0.046)	0.509 (0.052)	0.255 (0.062)	0.253 (0.079)	0.167 (0.089)
			500	0.504 (0.030)	0.507 (0.034)	0.503 (0.037)	0.257 (0.044)	0.250 (0.052)	0.166 (0.057)
			1000	0.503 (0.023)	0.503 (0.022)	0.502 (0.026)	0.254 (0.033)	0.247 (0.036)	0.164 (0.040)
		30% HT	250	0.507 (0.054)	0.513 (0.051)	0.512 (0.057)	0.257 (0.070)	0.249 (0.087)	0.165 (0.096)
			500	0.504 (0.036)	0.510 (0.037)	0.506 (0.043)	0.259 (0.045)	0.250 (0.061)	0.165 (0.063)
			1000	0.503 (0.028)	0.505 (0.024)	0.506 (0.027)	0.256 (0.035)	0.250 (0.038)	0.166 (0.045)
	sequential	15%	250	0.499 (0.037)	0.499 (0.036)	0.500 (0.037)	0.251 (0.049)	0.252 (0.055)	0.168 (0.060)
			500	0.500 (0.023)	0.499 (0.026)	0.500 (0.027)	0.252 (0.035)	0.250 (0.037)	0.165 (0.040)
			1000	0.500 (0.016)	0.497 (0.017)	0.500 (0.020)	0.248 (0.023)	0.252 (0.025)	0.165 (0.027)
		30%	250	0.500 (0.046)	0.500 (0.052)	0.504 (0.053)	0.249 (0.061)	0.255 (0.077)	0.164 (0.087)
			500	0.499 (0.032)	0.499 (0.037)	0.500 (0.038)	0.253 (0.046)	0.250 (0.052)	0.165 (0.056)
			1000	0.499 (0.024)	0.497 (0.024)	0.500 (0.026)	0.250 (0.035)	0.248 (0.036)	0.164 (0.040)
		30% HT	250	0.500 (0.057)	0.501 (0.059)	0.506 (0.058)	0.252 (0.071)	0.248 (0.086)	0.164 (0.095)
			500	0.498 (0.038)	0.498 (0.043)	0.500 (0.045)	0.251 (0.048)	0.252 (0.062)	0.164 (0.063)
			1000	0.499 (0.029)	0.496 (0.028)	0.500 (0.028)	0.250 (0.038)	0.252 (0.038)	0.166 (0.045)
Two-stage semiparametric	global	15%	250	0.498 (0.046)	0.500 (0.038)	0.507 (0.040)	0.250 (0.050)	0.250 (0.056)	0.165 (0.060)
			500	0.499 (0.032)	0.500 (0.028)	0.505 (0.030)	0.251 (0.035)	0.248 (0.037)	0.164 (0.040)
			1000	0.499 (0.024)	0.498 (0.019)	0.502 (0.021)	0.248 (0.024)	0.251 (0.026)	0.164 (0.028)
		30%	250	0.530 (0.092)	0.519 (0.064)	0.522 (0.059)	0.249 (0.066)	0.251 (0.080)	0.157 (0.086)
			500	0.507 (0.070)	0.508 (0.050)	0.512 (0.043)	0.255 (0.047)	0.246 (0.054)	0.162 (0.058)
			1000	0.509 (0.051)	0.506 (0.034)	0.507 (0.032)	0.250 (0.038)	0.247 (0.037)	0.162 (0.041)
		30% HT	250	0.637 (0.150)	0.583 (0.114)	0.548 (0.067)	0.240 (0.087)	0.255 (0.090)	0.151 (0.091)
			500	0.616 (0.114)	0.572 (0.084)	0.534 (0.058)	0.246 (0.055)	0.257 (0.068)	0.160 (0.071)
			1000	0.614 (0.115)	0.564 (0.079)	0.534 (0.042)	0.246 (0.053)	0.260 (0.049)	0.161 (0.046)
	sequential	15%	250	0.497 (0.047)	0.500 (0.039)	0.507 (0.041)	0.249 (0.049)	0.251 (0.057)	0.163 (0.059)
			500	0.499 (0.033)	0.500 (0.028)	0.504 (0.030)	0.251 (0.035)	0.249 (0.037)	0.163 (0.040)
			1000	0.499 (0.024)	0.498 (0.019)	0.501 (0.022)	0.248 (0.024)	0.252 (0.026)	0.164 (0.028)
		30%	250	0.529 (0.092)	0.523 (0.062)	0.530 (0.057)	0.247 (0.065)	0.248 (0.079)	0.152 (0.083)
			500	0.506 (0.071)	0.515 (0.047)	0.520 (0.041)	0.253 (0.047)	0.249 (0.051)	0.159 (0.056)
			1000	0.509 (0.052)	0.511 (0.033)	0.515 (0.031)	0.248 (0.038)	0.247 (0.036)	0.159 (0.041)
		30% HT	250	0.638 (0.147)	0.583 (0.108)	0.559 (0.060)	0.239 (0.086)	0.242 (0.089)	0.143 (0.085)
			500	0.616 (0.113)	0.573 (0.077)	0.546 (0.051)	0.245 (0.054)	0.247 (0.066)	0.154 (0.068)
			1000	0.614 (0.114)	0.565 (0.074)	0.546 (0.036)	0.245 (0.053)	0.250 (0.047)	0.155 (0.044)

Table B.21: Simulation results using global and sequential one-stage parametric and two-stage semiparametric estimation for four-dimensional data. The D-vine copula model (**Setting 2 in Table 4.15**) captures for Clayton (C) copulas in $\mathcal{T}_1$ increasing dependence with $\tau_{1,2} = 0.3$, $\tau_{2,3} = 0.5$, $\tau_{3,4} = 0.7$. The **empirical mean (empirical standard deviation)** of the **Kendall's τ estimates** are presented based on 250 replications and samples of different sizes affected by either 15%, 30% or heavy tail 30% right-censoring.

				D-vine copula model				
			C; $\tau_{1,2}$: 0.300	C; $\tau_{2,3}$: 0.500	C; $\tau_{3,4}$: 0.700	F; $\tau_{1,3;2}$: 0.250	F; $\tau_{2,4;3}$: 0.250	F; $\tau_{1,4;2,3}$: 0.167
One-stage parametric	global	15% 250	0.299 (0.043)	0.500 (0.039)	0.701 (0.028)	0.249 (0.049)	0.253 (0.053)	0.170 (0.062)
		15% 500	0.300 (0.028)	0.500 (0.027)	0.700 (0.020)	0.251 (0.036)	0.251 (0.037)	0.165 (0.041)
		15% 1000	0.308 (0.018)	0.507 (0.015)	0.707 (0.011)	0.248 (0.024)	0.253 (0.025)	0.167 (0.026)
		30% 250	0.300 (0.063)	0.499 (0.059)	0.702 (0.044)	0.248 (0.067)	0.253 (0.079)	0.169 (0.094)
		30% 500	0.299 (0.044)	0.499 (0.043)	0.700 (0.031)	0.251 (0.051)	0.251 (0.053)	0.167 (0.056)
		30% 1000	0.298 (0.031)	0.497 (0.029)	0.699 (0.021)	0.250 (0.035)	0.250 (0.034)	0.165 (0.040)
		30% HT 250	0.300 (0.080)	0.499 (0.079)	0.699 (0.060)	0.244 (0.079)	0.254 (0.088)	0.163 (0.102)
		30% HT 500	0.295 (0.052)	0.497 (0.053)	0.698 (0.042)	0.251 (0.054)	0.253 (0.062)	0.166 (0.069)
		30% HT 1000	0.298 (0.041)	0.495 (0.034)	0.698 (0.030)	0.250 (0.042)	0.252 (0.039)	0.166 (0.048)
	sequential	15% 250	0.300 (0.045)	0.501 (0.039)	0.701 (0.028)	0.249 (0.048)	0.252 (0.054)	0.168 (0.061)
		15% 500	0.302 (0.030)	0.503 (0.026)	0.701 (0.020)	0.252 (0.035)	0.249 (0.038)	0.164 (0.041)
		15% 1000	0.304 (0.020)	0.505 (0.016)	0.706 (0.012)	0.248 (0.024)	0.253 (0.024)	0.167 (0.027)
		30% 250	0.300 (0.064)	0.499 (0.060)	0.702 (0.044)	0.248 (0.066)	0.252 (0.079)	0.167 (0.092)
		30% 500	0.298 (0.045)	0.498 (0.044)	0.699 (0.031)	0.253 (0.050)	0.250 (0.052)	0.166 (0.056)
		30% 1000	0.298 (0.032)	0.497 (0.030)	0.699 (0.021)	0.249 (0.036)	0.250 (0.034)	0.165 (0.040)
		30% HT 250	0.305 (0.082)	0.504 (0.080)	0.702 (0.060)	0.250 (0.081)	0.251 (0.087)	0.162 (0.098)
		30% HT 500	0.296 (0.054)	0.498 (0.054)	0.698 (0.042)	0.251 (0.053)	0.254 (0.061)	0.164 (0.068)
		30% HT 1000	0.299 (0.041)	0.495 (0.035)	0.698 (0.031)	0.250 (0.042)	0.251 (0.039)	0.165 (0.047)
Two-stage semiparametric	global	15% 250	0.309 (0.054)	0.499 (0.046)	0.693 (0.035)	0.247 (0.051)	0.253 (0.055)	0.164 (0.060)
		15% 500	0.308 (0.039)	0.500 (0.034)	0.696 (0.024)	0.249 (0.037)	0.252 (0.038)	0.162 (0.041)
		15% 1000	0.303 (0.027)	0.496 (0.025)	0.696 (0.017)	0.246 (0.024)	0.253 (0.025)	0.164 (0.027)
		30% 250	0.362 (0.119)	0.523 (0.089)	0.697 (0.061)	0.232 (0.070)	0.254 (0.078)	0.157 (0.093)
		30% 500	0.330 (0.082)	0.509 (0.067)	0.695 (0.044)	0.244 (0.054)	0.249 (0.055)	0.162 (0.056)
		30% 1000	0.329 (0.061)	0.508 (0.050)	0.697 (0.031)	0.244 (0.039)	0.251 (0.035)	0.159 (0.041)
		30% HT 250	0.519 (0.189)	0.614 (0.142)	0.736 (0.085)	0.211 (0.087)	0.250 (0.090)	0.136 (0.100)
		30% HT 500	0.490 (0.159)	0.594 (0.117)	0.721 (0.075)	0.214 (0.067)	0.253 (0.071)	0.137 (0.077)
		30% HT 1000	0.496 (0.140)	0.596 (0.101)	0.730 (0.058)	0.221 (0.061)	0.251 (0.050)	0.141 (0.055)
	sequential	15% 250	0.311 (0.055)	0.498 (0.047)	0.693 (0.035)	0.248 (0.050)	0.251 (0.055)	0.162 (0.059)
		15% 500	0.309 (0.040)	0.499 (0.034)	0.696 (0.024)	0.250 (0.037)	0.251 (0.039)	0.161 (0.040)
		15% 1000	0.304 (0.028)	0.495 (0.025)	0.696 (0.017)	0.247 (0.025)	0.252 (0.026)	0.164 (0.027)
		30% 250	0.364 (0.116)	0.526 (0.085)	0.700 (0.058)	0.233 (0.069)	0.251 (0.077)	0.152 (0.090)
		30% 500	0.331 (0.082)	0.513 (0.063)	0.698 (0.042)	0.244 (0.054)	0.249 (0.055)	0.160 (0.055)
		30% 1000	0.330 (0.060)	0.512 (0.047)	0.700 (0.029)	0.244 (0.040)	0.250 (0.036)	0.158 (0.041)
		30% HT 250	0.300 (0.080)	0.499 (0.079)	0.699 (0.060)	0.244 (0.079)	0.254 (0.088)	0.163 (0.102)
		30% HT 500	0.489 (0.155)	0.596 (0.107)	0.726 (0.064)	0.212 (0.067)	0.246 (0.072)	0.133 (0.075)
		30% HT 1000	0.495 (0.134)	0.595 (0.095)	0.733 (0.050)	0.220 (0.061)	0.242 (0.050)	0.138 (0.053)

Table B.22: Simulation results using global and sequential one-stage parametric and two-stage semiparametric estimation for four-dimensional data. The D-vine copula model (**Setting 1 in Table 4.15**) captures tail-behavior for subsequent gap times changing from lower tail-dependence (Clayton (C)) over no tail-dependence (Frank (F)) to upper tail-dependence (Gumbel (G)) with same overall dependence of Kendall's $\tau_{1,2} = \tau_{2,3} = \tau_{3,4} = 0.5$. The **empirical mean (empirical standard deviation)** of the **copula parameter estimates** are presented based on 250 replications and samples of different sizes affected by either 15%, 30% or heavy tail 30% right-censoring.

				D-vine copula model				
			C; $\theta_{1,2}$: 2.00	F; $\theta_{2,3}$: 5.74	G; $\theta_{3,4}$: 2.00	F; $\theta_{1,3;2}$: 2.37	F; $\theta_{2,4;3}$: 2.37	F; $\theta_{1,4;2,3}$: 1.54
One-stage parametric	global	15%	250 2.062 (0.273)	5.837 (0.631)	2.013 (0.143)	2.425 (0.531)	2.414 (0.588)	1.574 (0.594)
			500 2.049 (0.173)	5.805 (0.455)	2.005 (0.109)	2.417 (0.362)	2.377 (0.386)	1.530 (0.387)
			1000 2.032 (0.121)	5.747 (0.307)	2.004 (0.079)	2.372 (0.245)	2.395 (0.265)	1.528 (0.266)
		30%	250 2.096 (0.363)	5.983 (0.871)	2.061 (0.224)	2.456 (0.675)	2.447 (0.850)	1.571 (0.871)
			500 2.049 (0.247)	5.902 (0.637)	2.021 (0.154)	2.465 (0.476)	2.386 (0.549)	1.539 (0.552)
			1000 2.033 (0.187)	5.806 (0.411)	2.013 (0.108)	2.418 (0.356)	2.352 (0.377)	1.517 (0.384)
		30% HT	250 2.106 (0.455)	6.065 (0.981)	2.078 (0.247)	2.484 (0.766)	2.414 (0.938)	1.559 (0.953)
			500 2.055 (0.295)	5.969 (0.692)	2.039 (0.175)	2.479 (0.485)	2.395 (0.643)	1.536 (0.615)
			1000 2.039 (0.229)	5.843 (0.437)	2.030 (0.113)	2.448 (0.379)	2.378 (0.404)	1.538 (0.437)
	sequential	15%	250 2.016 (0.295)	5.756 (0.655)	2.011 (0.145)	2.398 (0.521)	2.411 (0.589)	1.563 (0.588)
			500 2.012 (0.184)	5.745 (0.476)	2.006 (0.109)	2.398 (0.366)	2.381 (0.388)	1.524 (0.384)
			1000 2.005 (0.130)	5.699 (0.313)	2.004 (0.080)	2.355 (0.247)	2.400 (0.266)	1.525 (0.266)
		30%	250 2.038 (0.378)	5.812 (0.954)	2.038 (0.221)	2.384 (0.655)	2.462 (0.829)	1.544 (0.854)
			500 2.007 (0.254)	5.767 (0.686)	2.011 (0.157)	2.420 (0.489)	2.394 (0.550)	1.535 (0.547)
			1000 2.000 (0.189)	5.704 (0.431)	2.004 (0.108)	2.380 (0.370)	2.364 (0.379)	1.517 (0.383)
		30% HT	250 2.051 (0.467)	5.866 (1.081)	2.054 (0.250)	2.427 (0.772)	2.396 (0.917)	1.545 (0.941)
			500 2.010 (0.311)	5.757 (0.777)	2.016 (0.182)	2.398 (0.509)	2.416 (0.652)	1.523 (0.610)
			1000 2.003 (0.232)	5.677 (0.495)	2.006 (0.114)	2.379 (0.399)	2.400 (0.408)	1.536 (0.436)
Two-stage semiparametric	global	15%	250 2.015 (0.369)	5.776 (0.689)	2.042 (0.163)	2.388 (0.537)	2.395 (0.602)	1.529 (0.584)
			500 2.009 (0.258)	5.757 (0.513)	2.027 (0.124)	2.392 (0.373)	2.358 (0.385)	1.512 (0.384)
			1000 1.999 (0.192)	5.708 (0.353)	2.012 (0.087)	2.353 (0.251)	2.386 (0.273)	1.511 (0.268)
		30%	250 2.417 (0.872)	6.236 (1.312)	2.122 (0.260)	2.393 (0.711)	2.422 (0.856)	1.475 (0.840)
			500 2.138 (0.593)	5.971 (0.956)	2.065 (0.184)	2.446 (0.511)	2.352 (0.564)	1.508 (0.563)
			1000 2.122 (0.434)	5.874 (0.634)	2.036 (0.136)	2.379 (0.405)	2.351 (0.387)	1.493 (0.399)
		30% HT	250 4.303 (2.138)	8.131 (2.795)	2.261 (0.328)	2.322 (0.929)	2.479 (0.981)	1.420 (0.889)
			500 3.643 (1.576)	7.545 (1.964)	2.178 (0.249)	2.350 (0.578)	2.482 (0.748)	1.495 (0.697)
			1000 3.549 (1.347)	7.303 (1.678)	2.164 (0.189)	2.347 (0.575)	2.496 (0.519)	1.487 (0.451)
	sequential	15%	250 2.012 (0.374)	5.776 (0.701)	2.041 (0.167)	2.379 (0.530)	2.401 (0.608)	1.513 (0.577)
			500 2.006 (0.262)	5.753 (0.515)	2.022 (0.125)	2.390 (0.374)	2.373 (0.390)	1.507 (0.382)
			1000 1.999 (0.196)	5.703 (0.354)	2.008 (0.089)	2.351 (0.252)	2.393 (0.277)	1.509 (0.268)
		30%	250 2.413 (0.871)	6.313 (1.292)	2.158 (0.258)	2.367 (0.701)	2.394 (0.839)	1.423 (0.802)
			500 2.131 (0.598)	6.083 (0.921)	2.099 (0.181)	2.415 (0.504)	2.374 (0.541)	1.469 (0.543)
			1000 2.120 (0.440)	5.970 (0.619)	2.071 (0.137)	2.360 (0.405)	2.353 (0.384)	1.467 (0.390)
		30% HT	250 4.300 (2.125)	8.059 (2.647)	2.312 (0.311)	2.301 (0.922)	2.338 (0.943)	1.334 (0.826)
			500 3.643 (1.561)	7.523 (1.816)	2.228 (0.232)	2.334 (0.572)	2.370 (0.707)	1.433 (0.660)
			1000 3.550 (1.339)	7.276 (1.558)	2.217 (0.172)	2.340 (0.574)	2.385 (0.496)	1.430 (0.426)

Table B.23: Simulation results using global and sequential one-stage parametric and two-stage semiparametric estimation for four-dimensional data. The D-vine copula model (**Setting 2 in Table 4.15**) captures for Clayton (C) copulas in $\mathcal{T}_1$ increasing dependence with $\tau_{1,2} = 0.3$, $\tau_{2,3} = 0.5$, $\tau_{3,4} = 0.7$. The **empirical mean (empirical standard deviation)** of the **copula parameter estimates** are presented based on 250 replications and samples of different sizes affected by either 15%, 30% or heavy tail 30% right-censoring.

				D-vine copula model				
			C; $\theta_{1,2}$: 0.86	C; $\theta_{2,3}$: 2.00	C; $\theta_{3,4}$: 4.67	F; $\theta_{1,3;2}$: 2.37	F; $\theta_{2,4;3}$: 2.37	F; $\theta_{1,4;2,3}$: 1.54
One-stage parametric	global	15%	250	0.866 (0.178)	2.024 (0.315)	4.756 (0.616)	2.381 (0.523)	2.420 (0.573)	1.579 (0.607)
			500	0.864 (0.117)	2.014 (0.216)	4.702 (0.440)	2.389 (0.378)	2.394 (0.388)	1.523 (0.393)
			1000	0.893 (0.076)	2.063 (0.128)	4.836 (0.256)	2.355 (0.248)	2.408 (0.260)	1.540 (0.257)
		30%	250	0.878 (0.256)	2.046 (0.490)	4.861 (1.046)	2.377 (0.717)	2.453 (0.860)	1.600 (0.932)
			500	0.864 (0.178)	2.023 (0.351)	4.731 (0.694)	2.405 (0.539)	2.398 (0.561)	1.548 (0.552)
			1000	0.854 (0.128)	1.988 (0.230)	4.678 (0.476)	2.375 (0.370)	2.379 (0.363)	1.525 (0.392)
		30% HT	250	0.897 (0.339)	2.090 (0.639)	4.912 (1.381)	2.358 (0.861)	2.475 (0.959)	1.549 (1.033)
			500	0.852 (0.211)	2.023 (0.422)	4.754 (0.932)	2.407 (0.577)	2.427 (0.664)	1.544 (0.671)
			1000	0.860 (0.168)	1.977 (0.267)	4.691 (0.671)	2.385 (0.441)	2.401 (0.415)	1.540 (0.463)
	sequential	15%	250	0.868 (0.184)	2.030 (0.314)	4.754 (0.623)	2.375 (0.503)	2.416 (0.573)	1.567 (0.598)
			500	0.870 (0.123)	2.032 (0.211)	4.716 (0.447)	2.400 (0.371)	2.376 (0.396)	1.513 (0.393)
			1000	0.875 (0.081)	2.043 (0.132)	4.808 (0.268)	2.351 (0.247)	2.405 (0.258)	1.539 (0.257)
		30%	250	0.880 (0.264)	2.049 (0.501)	4.859 (1.042)	2.381 (0.711)	2.441 (0.853)	1.579 (0.916)
			500	0.862 (0.183)	2.018 (0.358)	4.722 (0.710)	2.423 (0.527)	2.389 (0.553)	1.538 (0.548)
			1000	0.854 (0.132)	1.989 (0.235)	4.675 (0.477)	2.372 (0.377)	2.377 (0.363)	1.521 (0.390)
		30%HT	250	0.920 (0.351)	2.134 (0.659)	4.978 (1.386)	2.415 (0.879)	2.439 (0.952)	1.527 (0.979)
			500	0.858 (0.222)	2.026 (0.429)	4.752 (0.922)	2.402 (0.569)	2.435 (0.654)	1.531 (0.666)
			1000	0.864 (0.170)	1.978 (0.274)	4.680 (0.682)	2.382 (0.449)	2.396 (0.416)	1.528 (0.456)
Two-stage semiparametric	global	15%	250	0.914 (0.227)	2.025 (0.372)	4.596 (0.739)	2.353 (0.539)	2.427 (0.591)	1.527 (0.588)
			500	0.902 (0.167)	2.021 (0.273)	4.616 (0.514)	2.369 (0.397)	2.406 (0.400)	1.500 (0.393)
			1000	0.875 (0.112)	1.982 (0.200)	4.608 (0.374)	2.339 (0.256)	2.403 (0.266)	1.515 (0.262)
		30%	250	1.255 (0.674)	2.351 (0.882)	4.868 (1.418)	2.215 (0.737)	2.461 (0.842)	1.476 (0.926)
			500	1.032 (0.381)	2.152 (0.581)	4.702 (0.971)	2.332 (0.575)	2.378 (0.580)	1.505 (0.548)
			1000	1.006 (0.278)	2.107 (0.408)	4.659 (0.673)	2.316 (0.413)	2.389 (0.371)	1.472 (0.400)
		30% HT	250	2.761 (1.686)	3.818 (1.901)	6.297 (2.415)	2.017 (0.894)	2.426 (0.979)	1.283 (0.981)
			500	2.282 (1.245)	3.307 (1.407)	5.622 (1.812)	2.030 (0.690)	2.442 (0.773)	1.275 (0.760)
			1000	2.238 (1.036)	3.211 (1.133)	5.693 (1.405)	2.096 (0.637)	2.397 (0.532)	1.299 (0.530)
	sequential	15%	250	0.919 (0.231)	2.017 (0.375)	4.608 (0.743)	2.366 (0.530)	2.402 (0.592)	1.506 (0.575)
			500	0.904 (0.170)	2.013 (0.269)	4.609 (0.515)	2.380 (0.395)	2.397 (0.407)	1.491 (0.390)
			1000	0.877 (0.116)	1.971 (0.199)	4.602 (0.370)	2.343 (0.260)	2.395 (0.272)	1.511 (0.262)
		30%	250	1.259 (0.660)	2.365 (0.850)	4.923 (1.393)	2.220 (0.727)	2.426 (0.825)	1.427 (0.882)
			500	1.037 (0.378)	2.181 (0.556)	4.748 (0.946)	2.330 (0.573)	2.379 (0.575)	1.480 (0.532)
			1000	1.011 (0.276)	2.133 (0.394)	4.726 (0.645)	2.315 (0.418)	2.381 (0.378)	1.455 (0.392)
		30% HT	250	0.897 (0.339)	2.090 (0.639)	4.912 (1.381)	2.358 (0.861)	2.475 (0.959)	1.549 (1.033)
			500	2.256 (1.205)	3.266 (1.284)	5.670 (1.621)	2.010 (0.688)	2.362 (0.780)	1.236 (0.735)
			1000	2.211 (0.994)	3.165 (1.051)	5.722 (1.260)	2.088 (0.638)	2.307 (0.527)	1.270 (0.507)

Table B.24: Simulation results using global and sequential one-stage parametric estimation. In the top panels, the underlying D-vine copula model (**Setting 1 in Table 4.15**) captures tail-behavior for subsequent gap times changing from lower tail-dependence (Clayton (C)) over no tail-dependence (Frank (F)) to upper tail-dependence (Gumbel (G)) with same overall dependence of Kendall's $\tau_{1,2} = \tau_{2,3} = \tau_{3,4} = 0.5$. In the bottom panels, the underlying D-vine copula model (**Setting 2 in Table 4.15**) captures for Clayton (C) copulas in $\mathcal{T}_1$ increasing dependence with $\tau_{1,2} = 0.3$, $\tau_{2,3} = 0.5$, $\tau_{3,4} = 0.7$. The **empirical mean (empirical standard deviation)** of the **marginal parameter estimates** are presented based on 250 replications and samples of different sizes affected by either 15%, 30% or heavy tail 30% right-censoring.

				$\lambda_1 : 0.50$	$\rho_1 : 1.50$	$\lambda_2 : 1.00$	$\rho_2 : 1.50$	$\lambda_3 : 1.00$	$\rho_3 : 1.50$	$\lambda_4 : 1.00$	$\rho_4 : 1.50$
Setting 1 (Table 4.15)	global	15%	250	0.491 (0.036)	1.507 (0.074)	0.980 (0.061)	1.517 (0.068)	0.965 (0.034)	1.513 (0.087)	0.984 (0.074)	1.530 (0.101)
			500	0.492 (0.026)	1.508 (0.053)	0.986 (0.045)	1.513 (0.052)	0.972 (0.023)	1.509 (0.063)	0.986 (0.058)	1.518 (0.066)
			1000	0.495 (0.018)	1.503 (0.036)	0.989 (0.028)	1.508 (0.039)	0.978 (0.014)	1.501 (0.043)	0.988 (0.041)	1.509 (0.047)
		30%	250	0.493 (0.040)	1.508 (0.093)	0.981 (0.079)	1.513 (0.093)	0.951 (0.053)	1.499 (0.113)	1.004 (0.158)	1.528 (0.143)
			500	0.494 (0.029)	1.505 (0.064)	0.989 (0.061)	1.515 (0.070)	0.964 (0.037)	1.500 (0.085)	0.987 (0.111)	1.512 (0.101)
			1000	0.497 (0.019)	1.501 (0.042)	0.993 (0.042)	1.509 (0.048)	0.970 (0.026)	1.495 (0.054)	0.997 (0.075)	1.510 (0.070)
		30% HT	250	0.494 (0.040)	1.507 (0.095)	0.984 (0.106)	1.509 (0.112)	0.939 (0.075)	1.485 (0.122)	1.025 (0.259)	1.526 (0.188)
			500	0.495 (0.029)	1.505 (0.064)	0.988 (0.074)	1.510 (0.082)	0.949 (0.054)	1.483 (0.087)	0.986 (0.165)	1.500 (0.122)
			1000	0.498 (0.019)	1.501 (0.044)	0.994 (0.052)	1.508 (0.056)	0.960 (0.041)	1.484 (0.055)	0.986 (0.116)	1.498 (0.078)
	sequential	15%	250	0.500 (0.042)	1.506 (0.082)	1.001 (0.073)	1.510 (0.073)	1.005 (0.074)	1.518 (0.091)	1.005 (0.083)	1.525 (0.105)
			500	0.501 (0.031)	1.504 (0.058)	1.004 (0.053)	1.507 (0.056)	1.006 (0.055)	1.511 (0.065)	1.004 (0.065)	1.511 (0.068)
			1000	0.502 (0.021)	1.500 (0.039)	1.003 (0.035)	1.504 (0.042)	1.004 (0.038)	1.503 (0.045)	1.002 (0.045)	1.504 (0.048)
		30%	250	0.501 (0.044)	1.508 (0.095)	1.003 (0.091)	1.509 (0.093)	1.016 (0.119)	1.522 (0.121)	1.031 (0.169)	1.533 (0.142)
			500	0.501 (0.033)	1.502 (0.067)	1.007 (0.068)	1.510 (0.072)	1.013 (0.088)	1.516 (0.093)	1.010 (0.120)	1.515 (0.102)
			1000	0.502 (0.021)	1.500 (0.045)	1.006 (0.048)	1.506 (0.050)	1.006 (0.062)	1.507 (0.059)	1.015 (0.081)	1.513 (0.070)
		30% HT	250	0.502 (0.042)	1.502 (0.093)	1.007 (0.113)	1.508 (0.111)	1.019 (0.163)	1.515 (0.140)	1.046 (0.252)	1.527 (0.180)
			500	0.501 (0.031)	1.502 (0.067)	1.008 (0.081)	1.509 (0.084)	1.020 (0.128)	1.514 (0.108)	1.023 (0.179)	1.514 (0.125)
			1000	0.502 (0.021)	1.499 (0.046)	1.009 (0.059)	1.510 (0.058)	1.013 (0.097)	1.510 (0.069)	1.017 (0.129)	1.513 (0.082)
Setting 2 (Table 4.15)	global	15%	250	0.500 (0.042)	1.507 (0.080)	1.000 (0.077)	1.511 (0.074)	1.003 (0.072)	1.521 (0.078)	1.521 (0.078)	1.524 (0.084)
			500	0.501 (0.030)	1.505 (0.055)	1.004 (0.058)	1.509 (0.053)	1.005 (0.056)	1.512 (0.053)	1.512 (0.053)	1.511 (0.058)
			1000	0.493 (0.018)	1.499 (0.036)	0.972 (0.018)	1.507 (0.040)	0.971 (0.016)	1.504 (0.036)	0.970 (0.017)	1.505 (0.040)
		30%	250	0.501 (0.043)	1.508 (0.095)	1.005 (0.110)	1.512 (0.100)	1.017 (0.122)	1.534 (0.111)	1.534 (0.111)	1.531 (0.131)
			500	0.501 (0.033)	1.503 (0.065)	1.006 (0.079)	1.512 (0.070)	1.013 (0.082)	1.514 (0.077)	1.514 (0.077)	1.517 (0.088)
			1000	0.502 (0.021)	1.500 (0.043)	1.007 (0.053)	1.508 (0.048)	1.007 (0.061)	1.510 (0.055)	1.011 (0.066)	1.511 (0.064)
		30% HT	250	0.501 (0.042)	1.505 (0.096)	1.008 (0.139)	1.509 (0.113)	1.027 (0.172)	1.527 (0.143)	1.527 (0.143)	1.559 (0.186)
			500	0.501 (0.031)	1.502 (0.066)	1.011 (0.091)	1.512 (0.080)	1.016 (0.122)	1.514 (0.096)	1.514 (0.096)	1.521 (0.121)
			1000	0.502 (0.021)	1.499 (0.045)	1.008 (0.068)	1.510 (0.055)	1.011 (0.089)	1.510 (0.063)	1.019 (0.119)	1.517 (0.087)
	sequential	15%	250	0.500 (0.042)	1.506 (0.082)	1.001 (0.077)	1.508 (0.077)	1.003 (0.072)	1.515 (0.084)	1.515 (0.084)	1.519 (0.088)
			500	0.499 (0.029)	1.503 (0.058)	0.999 (0.056)	1.506 (0.057)	0.998 (0.053)	1.511 (0.059)	1.511 (0.059)	1.509 (0.063)
			1000	0.502 (0.021)	1.500 (0.039)	0.977 (0.016)	1.512 (0.043)	0.974 (0.016)	1.508 (0.042)	0.972 (0.016)	1.509 (0.043)
		30%	250	0.501 (0.044)	1.507 (0.095)	1.004 (0.110)	1.511 (0.101)	1.015 (0.121)	1.527 (0.116)	1.527 (0.116)	1.527 (0.134)
			500	0.502 (0.033)	1.500 (0.066)	1.008 (0.079)	1.510 (0.073)	1.016 (0.084)	1.514 (0.082)	1.514 (0.082)	1.513 (0.088)
			1000	0.502 (0.021)	1.500 (0.045)	1.007 (0.054)	1.507 (0.051)	1.008 (0.062)	1.509 (0.059)	1.011 (0.066)	1.510 (0.065)
		30% HT	250	0.502 (0.043)	1.501 (0.094)	1.005 (0.138)	1.502 (0.112)	1.009 (0.169)	1.509 (0.134)	1.509 (0.134)	1.531 (0.185)
			500	0.501 (0.031)	1.502 (0.067)	1.010 (0.092)	1.507 (0.079)	1.018 (0.123)	1.515 (0.096)	1.515 (0.096)	1.517 (0.119)
			1000	0.502 (0.021)	1.498 (0.047)	1.008 (0.067)	1.508 (0.057)	1.012 (0.090)	1.509 (0.065)	1.019 (0.119)	1.513 (0.084)

Runtime comparison

Table B.25: Average computation time in seconds for the simulation settings considered in the extensive simulation study of Section 4.4.7 based on 250 replications. Results for one-stage parametric and two-stage semiparametric together with sequential and global proceeding are compared. Calculations were run on a Linux cluster with Intel Xeon E5-2690 v3 CPU 64 GB RAM.

			One-stage parametric		Two-stage semiparametric	
			sequential	global	sequential	global
Setting 1 (Table 4.15)	15%	250	129.65	1077.56	5.61	51.95
		500	159.34	1328.51	11.64	101.72
		1000	338.29	2752.166	23.31	198.96
	30%	250	58.19	513.40	3.73	39.62
		500	109.37	952.54	7.30	77.97
		1000	218.12	1945.72	14.38	149.63
	30% HT	250	60.44	545.05	3.75	41.06
		500	109.96	996.14	7.52	78.49
		1000	217.78	2072.00	14.41	152.21
Setting 2 (Table 4.15)	15%	250	101.20	427.07	6.40	53.33
		500	193.52	747.64	12.97	101.00
		1000	373.06	1303.48	25.82	203.71
	30%	250	60.98	281.48	3.95	35.48
		500	109.94	451.16	7.73	69.79
		1000	205.77	775.33	15.03	137.75
	30% HT	250	57.05	301.71	3.74	32.08
		500	108.21	506.98	7.16	60.23
		1000	194.64	791.06	14.18	113.58

B.7 Additional results for the asthma data

Table B.26: Marginal parameter estimates with standard errors based on 1000 bootstrap samples (in parentheses) of copula models fitted to each of the three samples of the asthma data using sequential and global one-stage parametric estimation. In case of Archimedean copulas the Frank (4dF), Gumbel (4dG), Clayton (4dC) and the Independence (4dInd) copula are considered. In case of D-vine copulas only the three best models are shown with Frank being the pair-copula family in trees $\mathcal{T}_2$ and $\mathcal{T}_3$.

			λ_1	ρ_1	λ_2	ρ_2	λ_3	ρ_3	λ_4	ρ_4
Sequential estimation	Full	FGG	1.900 (0.135)	1.005 (0.058)	1.285 (0.120)	0.612 (0.043)	1.365 (0.183)	0.698 (0.058)	1.664 (0.360)	0.726 (0.068)
		CGG	1.900 (0.135)	1.005 (0.058)	1.234 (0.121)	0.600 (0.043)	1.347 (0.182)	0.696 (0.059)	1.611 (0.353)	0.719 (0.068)
		GGG	1.900 (0.135)	1.005 (0.058)	1.273 (0.115)	0.626 (0.044)	1.364 (0.185)	0.704 (0.059)	1.655 (0.365)	0.732 (0.069)
	Treatment	FGG	1.759 (0.179)	1.174 (0.099)	1.043 (0.148)	0.595 (0.065)	1.062 (0.242)	0.619 (0.086)	1.554 (0.657)	0.711 (0.112)
		CGG	1.759 (0.179)	1.174 (0.099)	1.027 (0.149)	0.591 (0.065)	1.060 (0.240)	0.620 (0.086)	1.518 (0.641)	0.706 (0.112)
		GGG	1.759 (0.176)	1.174 (0.099)	1.050 (0.144)	0.599 (0.066)	1.065 (0.245)	0.621 (0.086)	1.547 (0.653)	0.707 (0.112)
	Control	FGG	2.057 (0.209)	0.898 (0.074)	1.596 (0.205)	0.639 (0.059)	1.602 (0.256)	0.756 (0.079)	1.834 (0.485)	0.745 (0.092)
		FGF	2.057 (0.209)	0.900 (0.074)	1.596 (0.205)	0.639 (0.059)	1.602 (0.256)	0.756 (0.079)	1.889 (0.547)	0.763 (0.096)
		GGG	2.057 (0.209)	0.900 (0.074)	1.563 (0.197)	0.653 (0.060)	1.601 (0.260)	0.763 (0.080)	1.811 (0.507)	0.750 (0.093)
Global estimation	Full	FGG	1.891 (0.135)	1.008 (0.058)	1.247 (0.116)	0.602 (0.041)	1.330 (0.180)	0.684 (0.057)	1.662 (0.354)	0.716 (0.067)
		CGG	1.899 (0.135)	1.005 (0.058)	1.200 (0.118)	0.590 (0.041)	1.307 (0.179)	0.682 (0.057)	1.568 (0.345)	0.708 (0.066)
		GGG	1.900 (0.135)	0.989 (0.056)	1.241 (0.112)	0.613 (0.042)	1.325 (0.181)	0.689 (0.057)	1.617 (0.361)	0.720 (0.068)
		4dF	1.869 (0.135)	1.008 (0.056)	1.303 (0.110)	0.620 (0.044)	1.543 (0.170)	0.698 (0.055)	2.660 (0.420)	0.761 (0.067)
		4dG	1.879 (0.161)	0.995 (0.063)	1.302 (0.131)	0.616 (0.048)	1.585 (0.234)	0.696 (0.063)	2.827 (0.592)	0.766 (0.084)
		4dC	1.890 (0.136)	1.006 (0.058)	1.273 (0.115)	0.612 (0.044)	1.478 (0.189)	0.680 (0.055)	2.525 (0.434)	0.731 (0.063)
		4dInd	1.900 (0.137)	1.005 (0.060)	1.293 (0.110)	0.620 (0.044)	1.657 (0.176)	0.690 (0.056)	3.068 (0.435)	0.742 (0.066)
	Treatment	FGG	1.757 (0.179)	1.175 (0.100)	1.018 (0.146)	0.590 (0.063)	1.025 (0.240)	0.608 (0.085)	1.503 (0.640)	0.700 (0.110)
		CGG	1.760 (0.179)	1.173 (0.099)	1.001 (0.146)	0.586 (0.062)	1.021 (0.238)	0.609 (0.084)	1.463 (0.613)	0.695 (0.109)
		GGG	1.763 (0.180)	1.172 (0.099)	1.026 (0.142)	0.595 (0.064)	1.028 (0.239)	0.610 (0.084)	1.497 (0.635)	0.696 (0.109)
		4dF	1.730 (0.177)	1.175 (0.099)	1.055 (0.136)	0.596 (0.068)	1.312 (0.241)	0.609 (0.083)	3.584 (1.116)	0.776 (0.116)
		4dG	1.759 (0.196)	1.174 (0.101)	1.050 (0.150)	0.599 (0.071)	1.379 (0.264)	0.601 (0.090)	3.989 (1.423)	0.753 (0.129)
		4dC	1.753 (0.182)	1.173 (0.095)	1.038 (0.135)	0.594 (0.070)	1.290 (0.247)	0.597 (0.086)	3.541 (1.118)	0.750 (0.111)
		4dInd	1.759 (0.182)	1.174 (0.095)	1.051 (0.136)	0.559 (0.072)	1.378 (0.243)	0.601 (0.084)	3.989 (1.118)	0.753 (0.109)
	Control	FGG	2.039 (0.209)	0.906 (0.075)	1.548 (0.200)	0.626 (0.057)	1.566 (0.252)	0.740 (0.076)	1.803 (0.493)	0.736 (0.091)
		FGF	2.039 (0.209)	0.906 (0.075)	1.561 (0.202)	0.632 (0.057)	1.601 (0.257)	0.754 (0.078)	1.881 (0.559)	0.762 (0.096)
		GGG	2.038 (0.209)	0.881 (0.071)	1.526 (0.192)	0.637 (0.058)	1.563 (0.256)	0.736 (0.077)	1.787 (0.509)	0.740 (0.092)
		4dF	2.011 (0.203)	0.902 (0.072)	1.604 (0.191)	0.647 (0.061)	1.751 (0.281)	0.767 (0.080)	2.405 (0.500)	0.763 (0.092)
		4dG	2.028 (0.233)	0.891 (0.075)	1.600 (0.214)	0.643 (0.065)	1.786 (0.316)	0.761 (0.087)	2.525 (0.631)	0.773 (0.102)
		4dC	2.047 (0.212)	0.900 (0.071)	1.565 (0.189)	0.639 (0.060)	1.655 (0.257)	0.744 (0.081)	2.252 (0.497)	0.739 (0.086)
		4dInd	2.057 (0.215)	0.898 (0.071)	1.583 (0.176)	0.648 (0.062)	1.883 (0.269)	0.759 (0.080)	2.766 (0.523)	0.756 (0.086)

Table B.27: Average cluster sizes and average censoring percentage among the 1000 bootstrap replications used for standard error calculation of the copula and marginal parameter estimates in the asthma data. Results for all three subsamples are shown. In case of Archimedean copulas the Frank (4dF), Gumbel (4dG), Clayton (4dC) and the Independence (4dInd) copula are considered. In case of D-vine copulas only the three best models are shown with Frank being the pair-copula family in $\mathcal{T}_2$ and $\mathcal{T}_3$. One stage-parametric estimation is considered. For D-vine copulas both sequential (top panels) and global (bottom panels) estimation are performed.

			#size 1	#size 2	#size 3	#size 4 (event)	#size 4 (censored)	%censoring
Sequential estimation	Full	FGG	17.85	63.74	46.36	23.68	80.38	21.67
		CGG	17.84	63.49	46.51	23.88	80.28	21.67
		GGG	17.85	64.65	47.37	24.02	78.12	22.08
	Treatment	FGG	8.96	39.86	24.32	8.01	31.85	25.34
		CGG	8.96	39.73	24.39	8.06	31.86	25.32
		GGG	8.96	39.77	24.76	8.11	31.40	25.50
	Control	FGG	8.49	24.47	22.61	14.81	48.62	18.61
		FGF	8.49	24.47	22.61	14.37	49.06	18.50
		GGG	8.49	25.01	23.26	15.15	47.10	19.10
Global estimation	Full	FGG	17.99	65.68	45.82	22.91	79.61	21.90
		CGG	17.85	65.24	46.06	23.12	79.72	21.85
		GGG	18.10	66.08	46.76	23.25	77.80	22.22
		4dF	18.98	62.26	45.53	21.78	83.45	21.22
		4dG	18.59	62.60	45.05	21.32	84.45	21.04
		4dC	18.17	62.83	44.30	20.57	86.14	20.71
		4dInd	17.82	63.26	44.06	20.36	86.51	20.69
	Treatment	FGG	8.97	40.65	23.98	7.79	31.61	25.51
		CGG	8.96	40.55	23.96	7.87	31.68	25.47
		GGG	8.94	40.57	24.42	7.92	31.15	25.67
		4dF	9.31	39.51	22.99	6.65	34.55	24.42
		4dG	9.01	39.88	23.09	6.20	34.82	24.34
		4dC	9.07	39.54	23.21	6.11	35.07	24.23
		4dInd	8.94	39.70	23.38	6.18	34.81	24.32
	Control	FGG	8.62	25.45	22.51	14.35	48.07	18.87
		FGF	8.62	25.21	22.37	14.19	48.61	18.70
		GGG	8.76	25.65	23.01	14.62	46.96	19.23
		4dF	9.13	23.87	22.35	14.04	49.62	18.37
		4dG	8.80	23.79	21.88	13.98	50.55	18.05
		4dC	8.49	23.77	21.93	13.66	51.15	17.85
		4dInd	8.43	24.00	21.37	13.67	51.53	17.73