MIXTURE OF VINES FOR DEPENDENCE MODELING: FINITE MIXTURE AND CD-VINE APPROACHES WITH APPLICATIONS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

ÖMER OZAN EVKAYA

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
STATISTICS

JULY 2018

Approval of the thesis:

MIXTURE OF VINES FOR DEPENDENCE MODELING: FINITE MIXTURE AND CD-VINE APPROACHES WITH APPLICATIONS

submitted by **ÖMER OZAN EVKAYA** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Statistics Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar _____
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Ayşen Dener Akkaya _____
Head of Department, **Statistics**

Assoc. Prof. Dr. Ceylan T. Yozgatlıgil _____
Supervisor, **Statistics Dept., METU**

Prof. Dr. A. Sevtap Selçuk-Kestel _____
Co-supervisor, **Institute of Applied Mathematics, METU**

Examining Committee Members:

Assoc. Prof. Dr. Özlem İlk Dağ
Statistics Dept., METU _____

Assoc. Prof. Dr. Ceylan T. Yozgatlıgil
Statistics Dept., METU _____

Assoc. Prof. Dr. Kasırga Yıldırak
Actuarial Science Dept., Hacettepe University _____

Assoc. Prof. Dr. Ceren Vardar Acar
Statistics Dept., METU _____

Assist. Prof. Dr. Emel Kızılok Kara
Actuarial Science Dept., Kırıkkale University _____

Date: _____

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: ÖMER OZAN EVKAYA

Signature :

ABSTRACT

MIXTURE OF VINES FOR DEPENDENCE MODELING: FINITE MIXTURE AND CD-VINE APPROACHES WITH APPLICATIONS

Evkaya, Ömer Ozan

Ph.D., Department of Statistics

Supervisor : Assoc. Prof. Dr. Ceylan T. Yozgatlıgil

Co-Supervisor : Prof. Dr. A. Sevtap Selçuk-Kestel

July 2018, 165 pages

Recently, there has been an increasing interest on the combination of copulas with a finite mixture model. By a finite mixture, a suitable weighted sum of a parametric densities are tied together in a probabilistic manner. The combination of vine copulas incorporated into a finite mixture model is also beneficial to capture hidden structures in a data set. On the other hand, there are limited number of studies about the mixture of vine copulas. In this dissertation, different mixture of vines are proposed for expressing the complex and hidden dependencies in a multivariate data. Firstly, the mixture of vine copulas with different dependence structures are offered to capture the complex association in higher dimension. For this reason, finite mixture of C-vine is studied with different copula pairs. Thereafter, finite mixture of C- and D-vines have been tested with the same copula family. Lastly, as a novel approach, finite number of C-vines are incorporated into a D-vine copula model to derive the association between several variables. The values of cumulative distribution functions for each component having C-vine structure are combined with D-vine by considering the temporal ordering of the components. The performance of the proposed models are tested using simulated and real data sets, then the corresponding results are interpreted in depth.

Keywords: Vine mixture, Finite mixture model, CD-vine mixture, Optimization

ÖZ

BAĞIMLILIK ANALİZİ İÇİN VİNE COPULA KARIŞIMI: UYGULAMALARLA SONLU KARIŞIM VE CD-VİNE YAKLAŞIMI

Evkaya, Ömer Ozan

Doktora, İstatistik Bölümü

Tez Yöneticisi : Doç. Dr. Ceylan T. Yozgatlıgil

Ortak Tez Yöneticisi : Prof. Dr. A. Sevtap Selçuk-Kestel

Temmuz 2018 , 165 sayfa

Son yıllarda, sonlu karışım modelleri ile copulaların bir araya getirilmesi konusundaki çalışmalara artan bir ilgi vardır. Sonlu karışım modeli yardımıyla, parametrik yoğunluk fonksiyonları uygun ağırlıklar verilerek olasılıksal olarak bir araya getirilmektedir. Verilen veri setinde gizli bağımlılıkların anlaşılmasında, sonlu karışım modelleri ile birleştirilmiş farklı vine copula kombinasyonları da yardımcı olmaktadır. Öte yandan, literatürde sonlu karışım yaklaşımı ile ilgili sınırlı sayıda çalışma bulunmaktadır. Bu doktora tezinde, çok boyutlu verilerde saklı olan bağımlılık yapısı için vine copula karışımları önerilmiştir. İlk olarak, yüksek boyutlu verilerde yer alan saklı kompleks ilişkinin yakalanması için, farklı bağımlılık yapılarına sahip vine copula karışımları sunulmuştur. Buna göre, farklı bağımlılık yapılarını ifade eden copulalar yardımıyla, sonlu sayıda C-vine karışım modelleri önerilmiştir. Ayrıca, sonlu sayıda C- ve D-vine model karışımı aynı copula ailesi kullanılarak test edilmiştir. Son olarak, bir çok değişken arasındaki bağımlılık yapısının incelenmesini sağlayan özgün CD-vine modeli çalışılmıştır. Buna modelde, sonlu sayıda C-vine copula modellerinin birikimli dağılım fonksiyon değerleri, D-vine copula modeli içerisine eklenmiştir. Önerilen modellerin başarısı, benzetim tabanlı ve gerçek veriler kullanılarak test edilmiş, sonuçlar detaylı olarak yorumlanmıştır.

Anahtar Kelimeler: Vine karışımı, Sonlu karışım modeli, CD-vine karışımı, Optimizasyon

To My Growing Family

ACKNOWLEDGEMENTS

First and foremost, I would like to express my very great appreciation to my supervisor Assoc. Prof. Dr. Ceylan T. Yozgatlıgil for his patient guidance, encouragement and valuable advices during the development and preparation of this thesis. Second, I would like to express my sincere gratitude to my co-advisor Prof. Dr. A. Sevtap Selcuk-Kestel for her continuous support of my Ph.D. study and related research. Their guidance and support throughout the study has turned this study to an invaluable learning experience for me.

I would also like to acknowledge my appreciation to my examining committee, Assoc. Prof. Dr. Kasırga Yıldırak, Assoc. Prof. Dr. Özlem İlk Dağ, Assist. Prof. Dr. Emel Kızılok Kara, Assoc. Prof. Dr. Ceren Vardar Acar, Assoc. Prof. Dr. Ümit Aksoy and Prof. Dr. Ömür Uğur for spending their valuable time for reviewing my thesis. Their precious suggestions have improved the quality of my thesis. During the writing process, colleagues from Atılım University, Assoc. Prof. Dr. Mehmet Turan, Dr. Hacer Öz, Res. Assist. Nazmi Oyar had valuable contributions on technical problems. Their experiences have accelerated this writing process certainly.

My sincere thanks go to both Department of Statistics at Middle East Technical University and Departments of Mathematics at Atılım University for allowing me to use computer laboratory during this long simulation process. I would like to thank all academical and technical staff at both universities for their invisible but crucial help. Last but not least, I would like to thank my growing family year by year for supporting me throughout writing this thesis and my life in general.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

In this chapter, first the motivation of the study and related dependence measures are explored by giving short theoretical background. Besides, fundamentals of copula functions are summarized to understand the basic framework. Finally, the main objective of the thesis study is briefly mentioned at the end.

1.1 Background and Motivation

Copulas have became a popular tool in multivariate modeling, and they have been used for many real life problems where the multivariate dependence is of great interest and the usual multivariate normality assumption is violated. They successfully cover both negative and positive dependence, including the case of independence and incorporating the asymmetric dependence structure as well. Interest in copulas has been growing rapidly as a statistical tool to capture the complex dependence structure between several random variables. As copulas have different attractive and distinct properties, the number of copula-based models extensively increased in many research areas including actuarial science, finance, neuroscience and environmental based studies. On the other hand, in contrast to wide range of its applications in bivariate case, the extension of copulas to the higher dimensions has many drawbacks in practice. The reason comes from the fact that standard multivariate copulas like Gaussian (Normal), student-t and other archimedean types lack flexibility to model the dependence in higher dimensions with other limitations such as restriction on parameter space. For this reason, vine copulas have been proposed and considered in

various research fields recently.

Vine copula framework offers a powerful and flexible modeling tool to detect the complex dependence structures by using bivariate copulas. Different vine copulas are available to identify dependence pattern in higher dimensions in terms of different pair copula families. Briefly, vine copulas are graphical type of statistical models designed to overcome the limitations of standard copulas in higher dimensions. They allow to express a multivariate copula by using the suitable bi-variate copulas in a hierarchical manner. For this reason, construction of a vine copula and its inference for multivariate data primarily relies on the pair copula decomposition. Meanwhile, as a result of large number of possible decompositions for conditional density functions inside of the whole multivariate density, there are many ways to generate vines. Among those constructions, two popular types of vine copulas widely used by researchers are called the Canonical(C) and Drawable(D) vines in the literature.

Latterly, there has been an increasing interest on the combination of copulas with a finite mixture model to get more flexible density functions. Within the context of the finite mixture model, a suitable weighted sum of parametric densities are tied together in a probabilistic manner. Such a framework is useful to reveal the complex dependence patterns observed for random variables and more flexible in terms of statistical modeling. The combination of vine copulas incorporated into a finite mixture model is also beneficial for capturing hidden structures in a multivariate data. On the other hand, apart from various studies about mixture of bivariate copula functions, there exist limited number of studies about the mixture of vine copulas in the literature.

In that respect, the study of Kim et al. [22] and the corresponding results is the main motivation and inspiration of this dissertation. As a first step, the possible improvements of the study belonging to [22] are considered. In order to derive various forms of mixture of vines, $2/3$ component, 3 dimensional C-vine mixture models and 2 component, 4 dimensional C-vine mixture with various copula families are considered. Additionally, the same procedure is extended to combine both C-vine and D-vine tree structures under 2 component 4 dimensional case. Finally, as a novel method, instead of using classical finite mixture setup, CD-vine mixture model has

been proposed. Under this framework, finite number of C-vine copula models are incorporated into a D-vine copula model to derive the dependence pattern between several random variables. The multivariate cumulative distribution function (c.d.f) values are obtained for each component of the mixture and then are combined with the help of D-vine copula in terms of temporal dependence among the components. In this new mixture model, the parameter estimation and inference for full multivariate density is required to maximize log-likelihood function in two consecutive steps with required initials. After testing proposed mixture models using simulated data, real life examples were exploited by using two different financial data sets.

1.2 Basics of Dependence Measure

The dependence between random variables is one of the crucial research topic in many fields of mathematics, primarily one of the most widely studied one in statistics. A vast number of studies have been introduced for different dependence concepts in the literature by offering useful applications. From the probabilistic point of view, knowing the distribution function for several random variables is one of the efficient way of understanding the association among multivariate outcomes. For this purpose, copulas or copula functions are very promising to focus on the joint behavior of random variables without putting strict restrictions on their marginal distributions. This method is a very practical tool for identifying multivariate distributions. Even if it has a relatively short history in statistics literature, copulas have been considered in the probability literature for the last 50 years. The interrelation between the dependence properties and measures of association is one of the corner stone for understanding how copulas can be used in this field of study [28]. For that purpose, widely used dependence measures are briefly explained below.

One of the widely accepted measures of dependence is the product moment correlation, simply called as Pearson's linear correlation, denoted by ρ. This linear dependence measure can be easily computed, defined briefly as follows,

Definition 1.2.1. *For two random variables, X and Y with $E[X] < \infty$, $E[Y] < \infty$, $Var[X] < \infty$ and $Var[Y] < \infty$, the Pearson's correlation for population is defined as:*

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}. \tag{1.2.1}$$

where $Cov(X,Y) = E(XY) - E(X)E(Y)$ *calculates the covariance of given random variables,* σ_X *and* σ_Y *are standard deviations of random variables* X *and* Y*, respectively. In this formulation, given fraction in Equation 1.2.1 can be expanded to express the following expression with sample size* n*,*

$$\widehat{\rho}(X,Y) = \frac{\sum_i^n (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_i^n (X_i - \overline{X})^2}\sqrt{\sum_i^n (Y_i - \overline{Y})^2}}. \tag{1.2.2}$$

where $\overline{X}$ *and* $\overline{Y}$ *are sample mean values of* X *and* Y*. These measures are calculated based on the given sample with* n *observations. Three main properties of* $\rho(X,Y)$ *can be summarized as follows:*

- $-1 \leq \rho(X,Y) \leq 1$*, here the lower and upper bounds exhibit perfect negative and positive linear correlation for two random variables, respectively.*
- *If* X *and* Y *are independent, then* $\rho(X,Y) = 0$*. But, the converse of this expression is not true.*
- $\rho(\alpha X + \beta, \gamma Y + \delta) = \text{sgn}(\alpha\gamma)\rho(X,Y)$ *for all* α *and* $\gamma \in \mathbb{R} \setminus 0$ *and* β*,* $\delta \in \mathbb{R}$*, which shows additionally that linear correlation is invariant under strictly increasing linear transformations.*

Although, Pearson correlation is easy to calculate and has many useful properties, it measures only linear correlation between continuous random variables. For this reason, it is not useful to capture non-linear dependence structure. Moreover, there exist some statistical problems with $\rho(X,Y)$, such as not being a robust measure. Thus, other dependence measures are required to avoid these deficiencies, which brings us definition of rank correlation and related dependence measures defined over the concepts of concordance and discordance.

Informally, two random variables are called concordant if large values of one tend to be associated with large values of the other and likewise small values of one with

small values of the other [28]. More precise definitions for the concordance and discordance are glued together as follows:

Definition 1.2.2. *Let (X, Y) be vector of continuous random variables with two observations (x_i, y_i) and (x_j, y_j) for arbitrary values of i and j. Then,*

$$(x_i, y_i)$$

and

$$(x_j, y_j)$$

are concordant if

$$(x_i - x_j)(y_i - y_j) > 0$$

and discordant if

$$(x_i - x_j)(y_i - y_j) < 0$$

Based on these definitions, two widely known correlation measures are defined using different constructions. First one is called Spearman's rho (ρ_S) and is developed to measure the correlation between the ranked data. The second rank based measure is the sample version of the measure of association, namely Kendall's tau (ρ_τ), defined in terms of concordance.

Definition 1.2.3 (Spearman's rho ρ_S)**.** *Let X and Y be two random variables having marginal distribution functions as F_X, F_Y and joint distribution function $F_{X,Y}$. Then, the estimator of ρ_S is defined as*

$$\widehat{\rho_S}(X, Y) = \frac{12}{n(n^2 - 1)} \sum_{i=1}^{n} (rank(X_i) - \frac{n+1}{2})(rank(Y_i) - \frac{n+1}{2}). \qquad (1.2.3)$$

where n is the sample size, $rank(X_i)$ and $rank(Y_i)$ denotes the rank of an observation. Besides, X_i and Y_i gives the position of X_i and Y_i in the ordered random sample. For instance, $rank(X_i)$ is defined as follows, i.e.

$$rank(X_i) = 1 + \sum_{X_j < X_i} j + \frac{1}{2} \sum_{j \neq i, X_j = X_i} j$$

.

and $rank(Y_i)$ has very similar expression. The population version of ρ_S is given by,

$$\rho_S(X, Y) = \rho(F_X(X), F_Y(Y)). \tag{1.2.4}$$

where ρ is the usual linear correlation.

Definition 1.2.4 (Kendall's ρ_τ)**.** *Let $(X_1, Y_1), (X_2, Y_2), \cdots (X_n, Y_n)$ denote a random sample of n observations from a vector (X, Y) of continuous random variables. Let c and d are the number of concordant and discordant pairs within each pair of observations of totally $\binom{n}{2}$ distinct pairs, denoted by (X_i, Y_i) and (X_j, Y_j) in the sample. Then, an estimate of Kendall's rank correlation is defined as,*

$$\widehat{\rho_\tau} = \frac{c - d}{\binom{n}{2}} = \binom{n}{2}^{-1} \sum_{i<j} \text{sgn}(X_i - X_j)(Y_i - Y_j). \tag{1.2.5}$$

The population version of $\rho_\tau(X, Y)$ for independent and identically distributed random vectors with (X_1, Y_1) and (X_2, Y_2) is given by,

$$\rho_\tau(X, Y) = \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) < 0]. \tag{1.2.6}$$

In addition to their important advantages, there is a well-defined relationship between the parameter of the copula functions and above-mentioned rank-based correlation measures. Equivalently, these measures are copula-based and can be expressed in terms of copulas. Although both rank based correlations measure the probability of concordance between random variables with respect to given copula family, values of ρ_S and ρ_τ are still quite different. Detailed explanation and related proofs about both dependence measures can be found in [28] with clear examples. Here, to make a more precise explanation, the relationship between copula functions and both correlation measures is described in the following theorem.

Theorem 1.2.5. *Let X and Y be continuous random variables whose copula function is C. Then the population version of Spearman's rho for X and Y (can be either denoted by $\rho(X, Y)$ or ρ_C) is defined by,*

$$\rho(X,Y) = \rho_C = 12 \underbrace{\iint}_{I^2} C(u,v)\, du\, dv - 3. \tag{1.2.7}$$

In a similar manner, the population version of Kendall's tau for X and Y (can be either denoted by $\tau(X,Y)$ or τ_C) is defined by,

$$\tau(X,Y) = \tau_C = 4 \underbrace{\iint}_{I^2} C(u,v)\, dC(u,v) - 1 = 4E[C(U,V)] - 1. \tag{1.2.8}$$

where $I^2 = I \times I$ is the unit square for $I = [0,1]$.

Above theorem briefly summarizes the role of concordance in copula functions. There are other rank-based correlation measures and all derivations of important results, mainly stated above, are important to understand basis of copula theory. Interested reader can read Chapter 5 of [28] in detail, but, the rest of the findings about correlation measures are not scope of this thesis. For this reason, it is time to go further on the problem of constructing multivariate distributions which are consistent with given marginal distributions and correlations more deeply. For a departure point, the next section introduces the theoretical framework of copula functions.

1.3 Theoretical framework for Copulas

In this part, the generalization of the "non-decreasing" notion for univariate functions, main concept applicable to multivariate functions, is presented. This construction is the key for understanding the basis of copula functions. In this respect, some notations and definitions will be recalled to build the preliminary part of the copula theory step by step. The first focus is the definition of a "2-increasing" function, described as a two dimensional analogy of a one variable non-decreasing function,

Definition 1.3.1. *Let S_1 and S_2 be nonempty subsets of $\overline{R}$ where $\overline{R}$ is the extended real line, i.e. $[-\infty, \infty]$. H be a two-place real function such that $Dom(H) = S_1 \times S_2$ and whose domain is a subset of $\overline{R}^2 = [-\infty, \infty] \times [-\infty, \infty]$, called as extended real*

plane and range is a subset of R. Let $B = [x_1, x_2] \times [y_1, y_2]$ *be a rectangle all of whose vertices are in* $Dom(H)$*. Then, the* H*-volume of* B *is defined as,*

$$V_H(B) = H(x_2, y_2) - H(x_2, y_1) - H(x_1, y_2) + H(x_1, y_1) \tag{1.3.1}$$

Here, a 2*-place real function* H *is called* 2*-increasing (referred as quasi-monotone) if* $V_H(B) \geq 0$ *holds for all rectangles* B *whose vertices are elements of* $Dom(H)$*.*

As a result of Definition 1.3.1, following lemmas can be presented for the continuity of quasi-copulas, indeed copulas.

Lemma 1.3.2. *Let* S_1 *and* S_2 *be nonempty subsets of* $\overline{R}$*, and* H *be a* 2*-increasing function with* $Dom(H) = S_1 \times S_2$*. Let* x_1*,* $x_2 \in S_1$ *with* $x_1 \leq x_2$ *and similarly* y_1*,* $y_2 \in S_2$ *with* $y_1 \leq y_2$*. Then,*

$$H(t, y_2) - H(t, y_1)$$

and

$$H(x_2, t) - H(x_1, t)$$

are non-decreasing functions over S_1 *and* S_2*, respectively. As a direct result of this lemma, a* 2*,increasing function* H *is non-decreasing in each argument results in the following lemma,*

Lemma 1.3.3. *Let* S_1 *and* S_2 *be nonempty subsets of* $\overline{R}$*, and* H *be a* 2*-increasing function with* $Dom(H) = S_1 \times S_2$*. Assume that* a_1 *and* a_2 *be least elements of the subsets* S_1 *and* S_2*, respectively. The function* H *from* $S_1 \times S_2$ *to* R *is grounded if* $H(x, a_2) = H(a_1, y) = 0 \ \forall \ (x, y)$ *in its* $Dom(H)$*. In this case, the function* H *is called grounded* 2*-increasing and it is non-decreasing in each argument.*

After defining the "grounded 2-increasing" function, the last crucial lemma for the construction of copula functions appears as follows,

Lemma 1.3.4. *Let* S_1 *and* S_2 *be nonempty subsets of* $\overline{R}$*, and* H *be a grounded* 2*-increasing function with margins, and* $Dom(H) = S_1 \times S_2$*. Let* $(x_1,\ y_1)$ *and* $(x_2,\ y_2)$ *be any points in its domain. Then,*

$$|H(x_2, y_2) - H(x_1, y_1)| \leq |F(x_2) - F(x_1)| + |G(y_2) - G(y_1)|$$

where F and G are cumulative distribution functions for random variables X and Y, respectively. Above-mentioned definitions and lemmas are required to define copula functions in two steps. Firstly, sub-copulas are defined, which are a certain class of grounded 2-increasing functions with their margins. Then, copula functions as sub-copulas with domain $I^2 = I \times I$ are described.

Definition 1.3.5. *A two-dimensional sub-copula, or simply* 2*-sub-copula is a function* C' *which satisfies the given properties below:*

- $Dom(C') = S_1 \times S_2$ *in which* S_1 *and* S_2 *are subsets of* $I = [0, 1]$.
- $C^{'}$ *is grounded* 2*-increasing function*
- $\forall\ u$ *in* S_1 *and* $\forall\ v$ *in* S_2*, the equalities* $C^{'}(u, 1) = u$ *and* $C^{'}(1, v) = v$ *are satisfied. Here, the range of* $C^{'}$*, (i.e.* $Ran(C^{'})$*) is also a subset of* I *since* $0 \leq C^{'}(u, v) \leq 1$ *holds* $\forall (u, v)$ *in* $Dom(C^{'})$.

Definition 1.3.6. *A two-dimensional copula, shortly copula is a* 2*-sub-copula with domain* I^2 *which satisfies certain properties listed below:*

- *For every* u *and* v *in* I*,*
  $C(u, 0) = C(0, v) = 0$ *and*
  $C(u, 1) = u$ *and* $C(1, v) = v$*;*
- *For every* u_1*,* u_2*,* v_1 *and* v_2 *in* I *with* $u_1 \leq u_2$ *and* $v_1 \leq v_2$*,*
  $C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0$

There exists small differences between sub-copulas and copulas, but this small distinction is still important for Sklar's Theorem, the corner stone of the copula theory. Besides, most of the properties of copulas are indeed the properties of sub-copulas. To illustrate another important inequality, satisfied by both sub-copulas and copulas, given below,

Theorem 1.3.7. *Let* $C^{'}$ *be a subcopula. Then for every* (u, v) *in* $Dom(C^{'})$*,*
$\max(u + v - 1, 0) \leq C^{'} \leq \min(u, v)$ *is satisfied. Since every copula is a sub-copula,*

given inequality holds for copulas for every (u, v) *in* I^2*,*
$W(u, v) \leq C(u, v) \leq M(u, v)$ *where commonly denoted functions* $W(u, v) = \max(u + v - 1, 0)$ *and* $M(u, v) = \min(u, v)$ *are lower and upper bounds for the copula function* $C(u, v)$*.*

This inequality primarily identifies the boundaries of copulas functions, generally referred as "Frechet-Hoeffding bounds". Later on, Sklar's Theorem and other related subjects will be discussed in Chapter 2 by keeping mind this general background. Interested reader can find all details about theoretical framework for the construction of copula functions in Chapter 2 of [28]. For this reason, this section is concluded by the main goal of thesis study, for summarizing the potential of the study.

1.4 Objective of the study

This dissertation aims to make practical and meaningful contribution to the field of mixture of vines. In this study, the mixture of vine copulas are proposed for expressing the different dependencies hidden in a multivariate data set. As a first approach, the finite mixture of C-vine copulas with different dependence structures are proposed to capture the complex hidden association in higher dimension. In that respect, various copula families are considered and in some cases distinct pairs are used to construct each component separately. Furthermore, finite mixture of C- and D-vine models is proposed to define more powerful mixture density function in terms of different tree structures. Afterwords, as a novel approach, finite number of C-vine copula models are incorporated into a D-vine copula model to derive the dependence structure between several random variables. For this mixture model, c.d.f. values obtained from the different components of C-vine copulas are combined with the help of D-vine method by considering the temporal ordering of the components.

The original motivation of CD-vine modeling approach directly relies on one of the agricultural problem, about the dependence among the growing periods of any crop yield. As a summary, for a specific crop, there is a unique requirement for climatic or soil based needs during its growth stages. For an accurate agricultural risk management decisions, the identification and understanding the dependence among

the growth stages for a specific crop yield is very crucial. In that respect, above declared CD-vine model can be a useful tool to detect any hidden dependence both within and between growth periods in terms of pre-selected explanatory variables for a crop yield. With this approach, dependence structure between variables at each time point is explained by C-vine model and dependence among the time points is captured by D-vine model.

In general, the whole research is dedicated to the mixture of vine copulas field and the performance of the proposed models are tested using both simulated and real data sets. Having all experiences on simulated data sets, two different real life examples are studied in the field of finance by giving advantages and drawbacks of the proposed mixture models. For this reason, two different data sets have been investigated including various major stock indices in the world. Finite mixture of C-vine model and the proposed CD-vine framework is implemented with these data sets separately.

The rest of this thesis is organized as follows: Chapter 2 summarizes main properties of copulas and vine copulas with respect to their foundations. Besides, important and widely known copula families are briefly described with their advantages and limitations in bivariate case. Afterwords, the main reason for the rise of vine copulas is explained in detail. The recall for the mixture of vines proposed by [22] with the short literature review is presented in Chapter 3. Furthermore, various finite mixture models are presented and then the novel CD-vine mixture model is proposed. Thereafter, Chapter 4, summarizes the numerical findings of the above mentioned mixture models based on simulated data and real life examples. Finally, in Chapter 5, the main conclusions of this study are highlighted with its advantages and drawbacks by giving possible further research directions and related open problems.

CHAPTER 2

METHODOLOGICAL FRAMEWORK

Under this chapter, the building blocks of copula functions are recalled briefly. There after, widely known and considered bivariate copula families are described with their certain properties. The advantages of those bivariate distributions are discussed with its main drawbacks, which open the next door; the rise of vine copulas. The construction of vine copulas, a kind of graphical multivariate distribution representation, is simply explained with two special cases in the end.

2.1 Copulas and Their Properties

2.1.1 Preliminaries

The history of copula models started with the problem of deriving bivariate distribution function for given marginal distributions, F_1 and F_2 for two random variables X_1 and X_2, respectively [16]. Later on, the notion of copula was first introduced by [35]. The main motivation behind the copula framework is the attempt of answering the questions of how to construct multivariate distributions with different margins and how to separate the dependence structure from their margins [16]. In this sense, the basic definition for copula functions is given as follows:

Definition 2.1.1. *A copula is a multivariate distribution whose marginals are all uniform over* $[0, 1]$. *For a* p*-dimensional vector U on the unit hyper cube, a copula C is defined as,*

$$C(u_1, u_2, ..., u_p) = Pr(U_1 \leq u_1, U_2 \leq u_2, ..., U_p \leq u_p). \tag{2.1.1}$$

This definition is the main result of the Sklar's theorem [35], namely building block of the theory of copulas, given below.

Theorem 2.1.2. (Sklar's Theorem) *Let F be a p-dimensional distribution function with univariate margins F_1, F_2, ..., F_p. Let A_j denote the range of F_j and $A_j = [-\infty, \infty]$ where $j = 1, 2, ..., p$. Then, there exists a copula function C s.t. for all $(x_1, x_2, ..., x_p) \in [-\infty, \infty]$ satisfies the following:*

$$F(x_1, x_2, ..., x_p) = C(F_1(x_1), F_2(x_2), ..., F_p(x_p)). \tag{2.1.2}$$

where the random variables $X_1, ..., X_p$ are assumed to be continuous.

Such a copula C, defined in Equation 2.1.2, is unique whenever all F_1, F_2, ..., F_p are continuous marginal distributions. Conversely, if C is a copula and $F_1, ..., F_p$ are distribution functions, then function F is a multivariate distribution function with margins $F_1, ..., F_p$. In particular, C can be interpreted as the cumulative distribution function of a p dimensional random variable on $[0, 1]^p$ with uniform margins.
Corresponding density of copula function is denoted by c, calculated as $c(x_1 \cdots x_p) = \frac{\partial^p C(x_1 \cdots x_p)}{\partial x_1 \cdots \partial x_p}$

2.1.2 Copula Related Dependence Measures

To understand the relationship among multivariate outcomes, copulas seem to be very promising. More simply, above mentioned joint distribution F given in Equation 2.1.2 contains the dependence structure of the random variables. Such a necessity for copulas comes from numerous fallacies concerning correlation, when especially outside of the elliptical world. In case of linear correlation, there exist certain deficiencies for multivariate models different from normal case. These shortcomings have been solved with an elegant movement into the alternative correlation measures, briefly summarized in Chapter 1.

Thanks to the concepts of concordance and discordance, two widely known rank correlations are defined as Spearman ρ and Kendall's τ. For both rank based correlation measures, the common term in their formulations is naturally the rank of the observation, illustrated below;

$$rank(X_i) = 1 + \sum_{X_j < X_i} j + \frac{1}{2} \sum_{j \neq i, X_j = X_i} j. \tag{2.1.3}$$

where Equation 2.1.3 identifies the position of the i'th observation for a given random variable, X, within in the ordered random sample. These two alternative correlation measures can be recalled one more time in terms of the copula function, as it is mentioned in Chapter 1. Above described nonparametric correlation measures $\widehat{\rho_S}(X, Y)$ and $\widehat{\rho_\tau}(X, Y)$ could be expressed as,

$$\widehat{\rho_S}(X, Y) = 12 \iint_V (C(u, v) - uv)\, du\, dv. \tag{2.1.4}$$

$$\widehat{\rho_\tau}(X, Y) = 4 \iint_V C(u, v) dC(u, v) - 1. \tag{2.1.5}$$

where $V = I^2$ is the unit square in 2-dimensional space and $dC(u, v) = \frac{\partial^2 C(u,v)}{\partial u \partial v} du dv$. Here, both equations 2.1.4 and 2.1.5 allow us to make a conversion from the rank based correlation measure to the parameter of copula functions and vice versa. In this study, as it is suggested in the literature, the main focus is Kendall τ values for measuring dependence among continuous variables in the numerical applications [17]. Furthermore, apart from above measures, another important dependence measures within the context of copulas is briefly explained below.

2.1.2.1 Tail Dependence

It is a well known fact that, linear correlation has vital fallacies for heavy-tailed distributions. Tail dependence phenomenon directly relates to existence of association in extreme values, mainly defined based on tails. Suppose X and Y are two random variables having distributions F_x and F_y, respectively. One can derive two important

asymptotic measures for tail dependence, called the coefficient of upper and lower tail dependence.

The upper version of this coefficient, denoted by λ_u, for X and Y is given as,

$$\lim_{t\to 1^-} P(Y > F_y^{-1}(t)|X > F_x^{-1}(t)) = \lambda_u, \tag{2.1.6}$$

where $\lambda_u \in [0,1]$ and $F^{-1}(t) = \inf(x|F(x \geq t))$ with $t \in (0,1)$. For nonzero values of λ, the pair (X,Y) are called asymptotically dependent in the upper tail. Otherwise, ie. $\lambda = 0$ they are asymptotically independent. Similarly, the coefficient of lower tail dependence, i.e. λ_l is defined below,

$$\lim_{t\to 0^+} P(Y \leq F_y^{-1}(t)|X \leq F_x^{-1}(t)) = \lambda_l, \tag{2.1.7}$$

where $\lambda_l \in [0,1]$ exists. The interpretation of the number λ_l is similar to λ_u. Since the equations 2.1.6 and 2.1.7 include conditional probabilities, there exist a certain connection between those measures and copulas. In brief, λ measure is indeed a function of bivariate copula for X and Y. In case of the continuity of F_x and F_y and existence of given limits, both equations 2.1.6 and 2.1.7 can be rewritten in terms of copula functions, as follows,

$$\begin{aligned}
\lambda_u &= \lim_{t\to 1^-} P(Y > F_y^{-1}(t)|X > F_x^{-1}(t)) \\
&= \lim_{t\to 1^-} \frac{P(Y > F_y^{-1}(t), X > F_x^{-1}(t))}{P(X > F_x^{-1}(t))} \\
&= \lim_{t\to 1^-} \frac{\overline{C}(t,t)}{1-t}.
\end{aligned} \tag{2.1.8}$$

where $\overline{C}(t,t) = 1 - 2t + C(t,t)$ and C is the unique copula function for the pair (X,Y). In a similar manner,

$$\begin{aligned}\lambda_l &= \lim_{t\to 0^+} P(Y \leq F_y^{-1}(t) | X \leq F_x^{-1}(t)) \\ &= \lim_{t\to 0^+} \frac{P(Y \leq F_y^{-1}(t), X \leq F_x^{-1}(t))}{P(X \leq F_x^{-1}(t))} \\ &= \lim_{t\to 0^+} \frac{C(t,t)}{t}.\end{aligned} \tag{2.1.9}$$

Based on equations 2.1.8 and 2.1.9, one can observe that the calculation of λ is easy to handle if the corresponding copula function C has a closed form, to understand the tail dependence as an asymptotic property. This tail dependence measure gives a reasonable information about the joint behavior in extremes asymptotically. On the other hand, for higher dimensions, λ measure for tail dependence is not so straightforward like in bivariate case. When it comes to the mixture of vines, the derivation of these tail dependence measures is not an ordinary calculation certainly.

2.1.3 Parametric Copula Families

This subsection primarily introduces the concept of copulas with their representations mathematically and graphically. The most frequently used copula families can be classified generally as elliptical and archimedean ones. There are more than 100 types of copula families in the literature and many of them are widely used in different research areas. Within this subsection, main copula families are described with their fundamental properties.

For the construction of copula density function, firstly, the marginal distributions should be decided and then these values are transformed into uniform numbers using probability integral transformation. In other respects, the departure point is modeling residuals when the random variables are univariate time series. In this modeling setup, even if these residuals do not ensure the normality conditions, the derivation of multivariate density function is still possible via copulas.

2.1.3.1 Elliptical Copulas

An elliptical copula is the copula function generated from an elliptical distribution by Sklar's theorem. General discussion about elliptical distributions can be found in [15].

Definition 2.1.3. *Let F be the multivariate c.d.f. of an elliptical distribution. Let F_i be CDF of the i'th margin and F_i^{-1} be its inverse function (quantile function) for $i = 1, ..., p$. The elliptical copula determined by F is*

$$C(u_1, ..., u_p) = F[F_1^{-1}(u_1) + ... + F_i^{-1}(u_p)]. \tag{2.1.10}$$

For instance ; normal copulas (derived from bivariate normal with zero means, unit variances and correlation ρ) and student t-copulas (derived from bivariate t-distribution with zero mean, degree of freedom ν and association ρ) are two types of elliptical family. Both of them play a predominant role within the elliptical class and their usage appears in diverse fields of applications.

Definition 2.1.4 (Gaussian (Normal) Copula)**.** *The p-variate Gaussian or Normal copula function could be defined as*

$$C(u_1, \cdots, u_p; \Sigma) = \Phi_{\Sigma}^{p}(\Phi^{-1}(u_1) \cdots \Phi^{-1}(u_p)). \tag{2.1.11}$$

where Σ denotes the correlation matrix (includes dependence structure), Φ_{Σ}^{p} standardized p-variate normal and Φ^{-1} represents the quantile function of a univariate standard normal distribution [20]. In the bivariate case, normal copula family has no tail dependence.

Definition 2.1.5 (Student-t copula)**.** *Assume that $Z \sim N(0, \Sigma)$ and $\nu_S = \frac{\sqrt{\nu}}{\sqrt{S}}$ with S is distributed as $X^2(\nu)$, where ν is the degrees of freedom (d.f.) for given chi-squared variable S. Here, $Y \equiv \nu_S Z$ has a t-distribution with df ν [20]. When the Sklar's Theorem applied with the definition of $\rho_{ij} \equiv \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\,\Sigma_{jj}}}$ for $1 \leq i, j \leq p$, Student-t copula function is derived as,*

$$C(u_1, \cdots, u_p; \nu, \rho) = t_{\nu,\rho}^{p}(t_{\nu}^{-1}(u_1) \cdots t_{\nu}^{-1}(u_p)). \tag{2.1.12}$$

In equation 2.1.12, t_ν^{-1} is the inverse of the classical univariate t-distribution function. Differently from Gaussian copula, bivariate Student t-copula function exhibits both upper and lower tail dependence, which can be calculated by using equations 2.1.8 and 2.1.9, respectively.

2.1.3.2 Archimedean Copulas

Under this subsection, another important class, archimedean copulas are summarized and discussed briefly. With this approach, the study of a multivariate copula could be reduced to a single univariate function [25]. The building block for the construction of such a family, directly based on generator functions ϑ is defined below,

Definition 2.1.6. *An archimedean copula is constructed through a generator ϑ as*

$$C(u_1, ..., u_p) = \vartheta^{[-1]}(\vartheta(u_1) + ... + \vartheta(u_p)). \tag{2.1.13}$$

where $\vartheta^{[-1]}$ is the pseudo-inverse of the generator ϑ, defined by

$$\vartheta^{[-1]}(t) \equiv \begin{cases} \vartheta^{-1}(t) & 0 \leq t \leq \vartheta(0), \\ 0 & \vartheta(0) \leq t \leq \infty. \end{cases} \tag{2.1.14}$$

In bivariate case,

$$C(u_1, u_2) = \vartheta^{[-1]}(\vartheta(u_1) + \vartheta(u_2)). \tag{2.1.15}$$

defines the so-called archimedean bivariate copula function. Here, in Equation 2.1.15, ϑ is called the generator of copula. For instance, Gumbel copula family is generated from $\vartheta = (-ln(t))^\theta$, where $\theta \geq 1$ and it controls the degree of dependence. Similar construction exists for different archimedean type copulas having various tail dependence structures.

Among all archimedean type families, Clayton, Gumbel and Frank are well known copula functions. Different copula families stand for deriving distinct dependence structures. For example, elliptical copulas and Frank copula are preferable to examine the symmetric dependence structures. On the other hand, Clayton and Gumbel are useful to identify the tail dependencies at lower and upper quantiles, respectively.

Table 2.1 summarizes bivariate copulas having certain distributional assumptions. Bivariate standard normal with parameter, ρ, bivariate Student-t distribution with parameter ρ and degrees of freedom, d. Besides, Frank, Clayton, Gumbel and Joe families defined over $[0, 1]$ in Table 2.1 are bivariate copulas with parameter θ. Finally, independence copula family shows no dependence pattern among the variables so that the parameter is directly equals to zero.

Table 2.1: The most widely used copula families and their parameter spaces

Copula Family	$C_\theta(u, v)$ Parameter (θ)
Normal	$\phi_\rho(\phi^{-1}(u), \phi^{-1}(v)) = \int_{-\infty}^{\phi^{-1}(u)} \int_{-\infty}^{\phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} exp(-\frac{s^2-2\rho st+t^2}{2(1-\rho^2)})\, dsdt$ $\rho \in (-1, 1)$
Student-t	$t_{d,\rho}(t_d^{-1}(u), t_d^{-1}(v)) = \int_{-\infty}^{t_d^{-1}(u)} \int_{-\infty}^{t_d^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} (1 + \frac{s^2-2\rho st+t^2}{d(1-\rho^2)})^{-\frac{d+2}{2}}\, dsdt$ $\rho \in (-1, 1); d \in (0, \infty)$
Frank	$(-\frac{1}{\theta}) \ln[1 + \frac{(exp^{-\theta u}-1)(exp^{-\theta v}-1)}{(exp^{-\theta}-1)}]$ $\theta \in (-\infty, \infty) \setminus \{0\}$
Clayton	$(u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}$ $\theta \in (0, \infty)$
Gumbel	$exp[-(u^{-\theta} + v^{-\theta})^{-\frac{1}{\theta}}]$ $\theta \in [1, \infty)$
Joe	$(1 - ((1-u)^\theta + (1-v)^\theta - (1-u)^\theta(1-v)^\theta)^{\frac{1}{\theta}})$ $\theta \in [1, \infty)$
Independence	uv $\theta = 0$

For the dependence structure within the above-mentioned copula families, a set of useful noteworthy pictures are presented. Each copula family has own tail dependence properties and this dependence is controlled by the parameter, generally denoted as θ. To visualize, the following contour plots are beneficial way of understanding each copula function, presented in Figure 2.1.

In Figure 2.1, all the margins are assumed to be normal. On the other hand, generally speaking, margins are converted into uniform numbers before the construction of copula functions. For visualizing various tail dependencies with uniform margins, both surface and contour plots are displayed in Figure 2.2 with the same parameter value. In both Figures 2.1 and 2.2, different tail dependence patterns for each family can be visualized clearly. For instance, Clayton family represents a lower (left) tail dependence but Gumbel and Joe copulas exhibit an upper (right) tail dependence

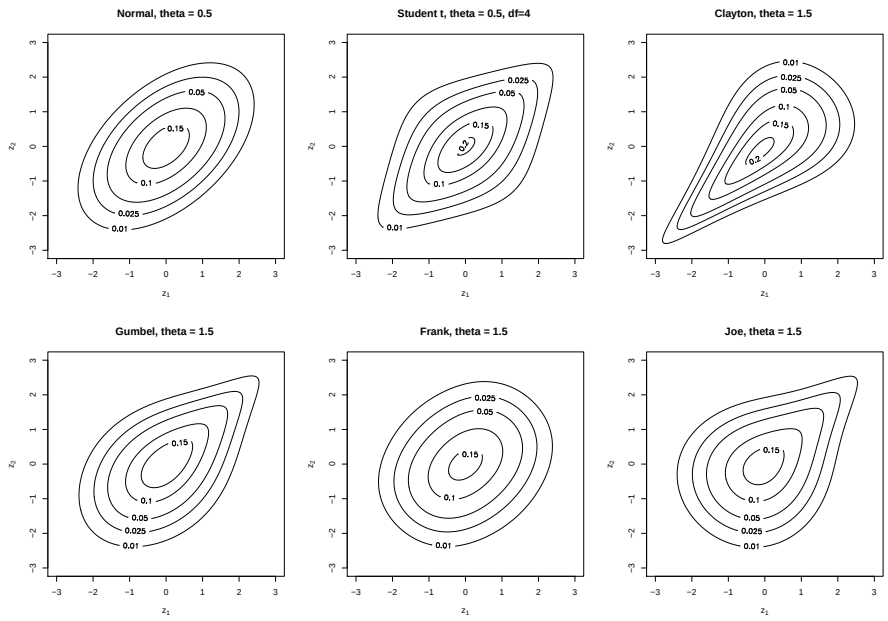


Figure 2.1: Bivariate Copula Families with Normal Margins

structure. Generally, all of those mentioned copula families are asymmetric type of copulas except Normal, Student-t and Frank copulas, presented in Figure 2.1.

As it is mentioned above, the parameter of the related copula family controls the degree of dependence over its natural parameter space. For instance, certain fixed changes on the dependence parameter can be visualized in Figure 2.3 for Clayton copula with its contour plots. Here, whenever the parameter θ increases, the contour lines get more dense in the lower tail to identify the increased dependence.

Along with one parameter copula families, other copula families are constructed based on the rotations or extensions of the existed ones, called as associated copulas. Among those special families, survival (rotated) copulas deserves more attention since they are more convenient to measure the negative dependencies within different tail structures. For example, Clayton, Gumbel, Joe rotated with 90, 180 and 270 degrees are some examples of those new families. To illustrate, various tail dependence structures for rotated versions are graphed in Figure 2.4 for Gumbel family. Similar impacts can be visualized for the rotated versions of other copula functions. All of those copula families appeared in various research fields, and this notion became increasingly popular at the end of nineties. For the interested reader, the monographs of [28] and [20] are referred to give more details about copulas.

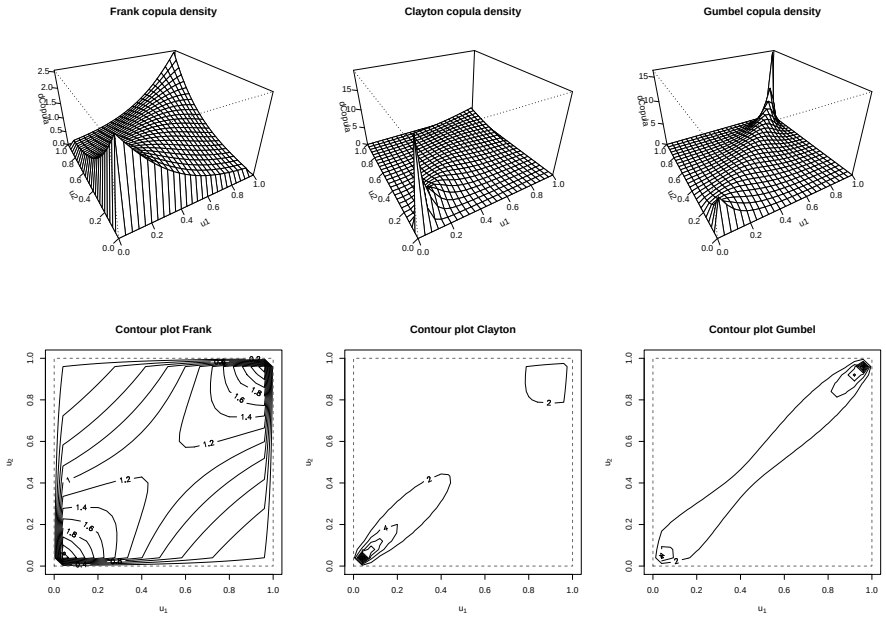


Figure 2.2: Surface and Contour Plots of Frank, Clayton and Gumbel families

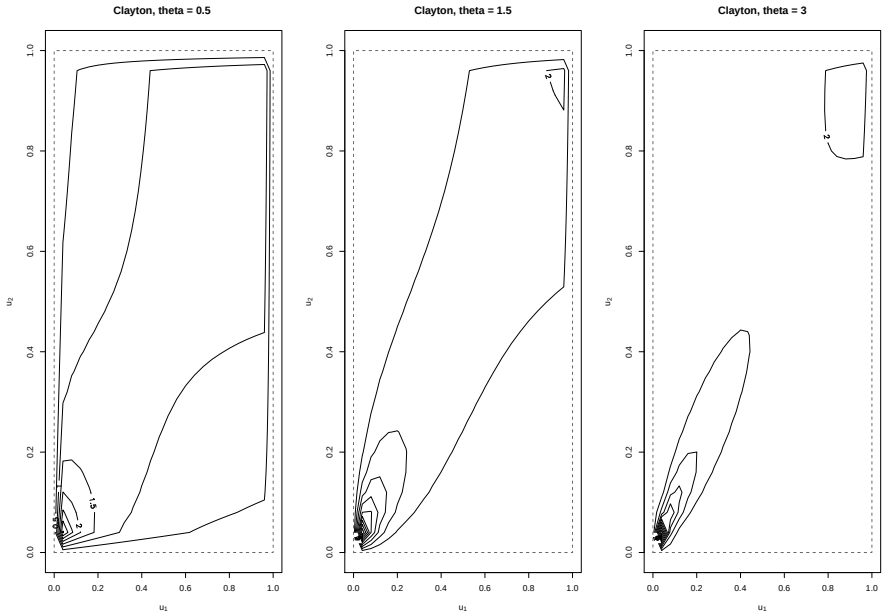


Figure 2.3: Contour Plots of Clayton families with different parameters

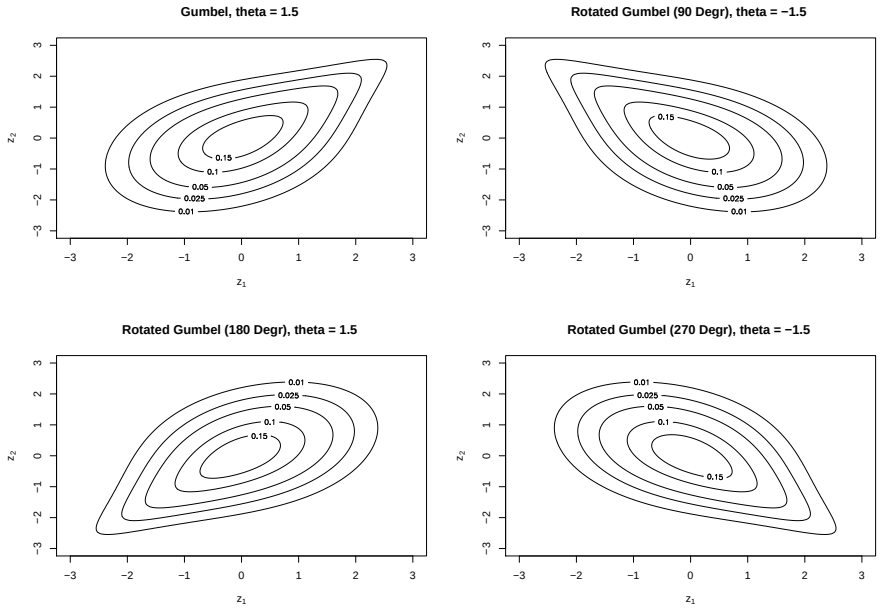


Figure 2.4: Contour Plots of Rotated versions of Gumbel family

2.2 Motivation for Vines

Even if they are very practical to model the dependence structure between variables, the use of copulas is still challenging in higher dimensions, where standard multivariate copulas suffer from rather inflexible structures. Certainly, it is not so realistic to extend elliptical and archimedean copula families directly to higher dimensions. Actually, in some cases it is not possible to do so without losing from their beneficial properties. For this reason, considerable efforts have been undertaken to increase the flexibility of multivariate copula models beyond the scope of elliptical and archimedean copulas.

2.2.1 Preliminaries

Many multivariate data structures exhibit; different marginal distributions having non-symmetric dependencies like heavy tail dependence pattern between some pairs. For such data sets, standard parametric distributions such as the Gaussian or multivariate t distribution are not efficient to identify the correct relationship. Moreover, standard elliptical and archimedean copulas do not allow to use different models

between pairs of variables, even if some of archimedean families can be extended to higher dimensions. For this reason, the analysis of high dimensional data requires more flexible multivariate stochastic models in copula framework. Considerable efforts have been undertaken to increase the flexibility of multivariate copula models and vine copulas (vines) are among the best-received of such efforts.

Vine copulas use unconditional and conditional bivariate copulas to describe a multivariate distribution [6]. A set of linked trees, called as vines, describe a vine copula's factorization of the multivariate copula density function into the density functions of its pair copulas (see [6], [14]). Vine copula models have proven themselves as a very flexible class of multivariate copula models with regard to symmetry and tail dependence for pairs of variables. The full specification requires the choice of vine tree structure, copula families for each pair copula term and their corresponding parameters. Historically, Joe [19] gave a probabilistic construction of multivariate distribution functions based on simple building blocks called pair-copulas. Bedford and Cooke [5], [6] organized these constructions in a graphical way called regular vines and gave expression for the joint density function. Estimation for the Gaussian case was considered in the book by Kurowicka and Cooke [14]. Besides, Aas et al. [1] used the Pair Copula Construction (PCC) to establish flexible multivariate model based on pair-copulas such as bivariate Gaussian, Student-t, Gumbel and Clayton copulas and provided their likelihood expressions. A recent survey by [13] about vine models is proposed and it is available for interested readers.

2.2.2 Representation of Vines

A vine copula structure is simply defined as a nested set of trees describing the pairwise copula functions unconditionally at the first tree and conditionally for the rest of connected trees. This construction is presented by Bedford and Cooke [5] as follows

Definition 2.2.1. *Let $V = T_1, ..., T_{p-1}$ denote the regular vine for p variables, where T_i is a connected tree with nodes $N_i = 1, ..., p$ and edges E_i for $i = 2, ..., p-1$. In this tree structure, T_i is a connected tree with nodes $N_i = E_{i-1}$.*
A regular vine with p variables is a vine where two edges in tree i are connected by

an edge in tree $i+1$, only if these edges share a common node. In general, there are totally $p(p-1)/2$ possible edges in a regular vine for p variables [14].

Formally, vine copulas are multivariate copula constructions building on bivariate copulas and this framework is called PCC generally [1]. To clarify, the procedure of PCC for multivariate p.d.f. in p-dimension is defined as follows [5]:

Definition 2.2.2. *Given*

$$f(x_1, \cdots, x_p) = \left(\sum_{t=2}^{p} f(x_t | x_1, \cdots, x_{t-1}) \right) f_1(x_1). \tag{2.2.1}$$

and for distinct values of i, j, $i_1, \cdots, i_k$ with $i < j$ and $i_1 < \cdots < i_k$ define,

$$c_{i,j|i_1,\cdots,i_k} = c_{i,j|i_1,\cdots,i_k}(F(x_i|x_{i_1}, \cdots, x_{i_k}), F(x_j|x_{i_1}, \cdots, x_{i_k})). \tag{2.2.2}$$

where f and c denote probability density function (p.d.f.) of original marginals and copula density function, respectively. Then, one can rewrite the conditional p.d.f, $f(x_t|x_1, \cdots, x_{t-1})$, in terms of conditional copulas as,

$$\begin{aligned} f(x_t|x_1, \cdots, x_{t-1}) &= c_{1,t|2,\cdots,t-1} f(x_t|x_2, \cdots, x_{t-1}) \\ &= \left(\sum_{s=1}^{t-2} c_{s,t|s+1,\cdots,t-1} \right) c_{t-1,t} f_t(x_t). \end{aligned} \tag{2.2.3}$$

Afterwords, by using equation 2.2.3 the joint p.d.f., given in equation 2.2.1 becomes

$$f(x_1, \cdots, x_p) = \left(\sum_{j=1}^{p-1} \sum_{i=1}^{p-j} c_{i,i+j|i+1,\cdots,i+j-1} \right) \left(\sum_{k=1}^{p} f_k(x_k) \right). \tag{2.2.4}$$

for $s = i$, $t = i + j$ given in Equation 2.2.4.

The decomposition presented in Equation 2.2.3 suggests that there is no unique way of deriving the multivariate copula, as different orderings of the variables might lead to a different estimations. For this reason, the full specification of a vine copula requires the choice of vine tree structure, determining copula families for each pair copula and estimating their corresponding parameters. Therefore, regular vine construction

can be made as a sequence of bivariate unconditional and conditional copulas, given below more formally,

Definition 2.2.3. *As T_i, N_i and E_i defined, an p dimensional vine tree structure $V = T_1, \ldots, T_{p-1}$ consists of consecutive $p-1$ linked trees satisfying the following conditions [6]:*

- *Tree T_1 is a tree having nodes 1 to p,*
- *Tree T_j has $p+1-j$ nodes and $p-j$ edges, where $j = 2, \cdots, p-1$*
- *Edges in tree T_j become nodes in tree T_{j+1},*
- *Two nodes in tree T_{j+1} can be joined by an edge only if the corresponding edges in tree T_j share a node, called as Proximity Condition (PC).*

For this tree construction, it has been shown that the edges could be defined by two nodes, called conditioning and conditioned nodes [5], [14].

2.2.3 Regular Vine Tree Structure

In general, vine is kind of a useful graphical tool to label constraints in multivariate distributions. Among those, regular vine, a flexible tool for modeling high-dimensional dependence, is a special case in which all constraints are unconditional or conditional 2-dimensional [20]. By following Definition $2.2.3$, tree structure of a Regular vine (R-vine) on 5 variables can be visualized below, in Figure 2.5.

Based on the R-vine tree structure presented visually, the density of a regular vine distribution is defined by the product of pair copula densities over the $p(p-1)/2$ edges identified by the regular vine tree structure and the product of the marginal densities. To illustrate, the multivariate density function for 5 variables can be written simply as below,

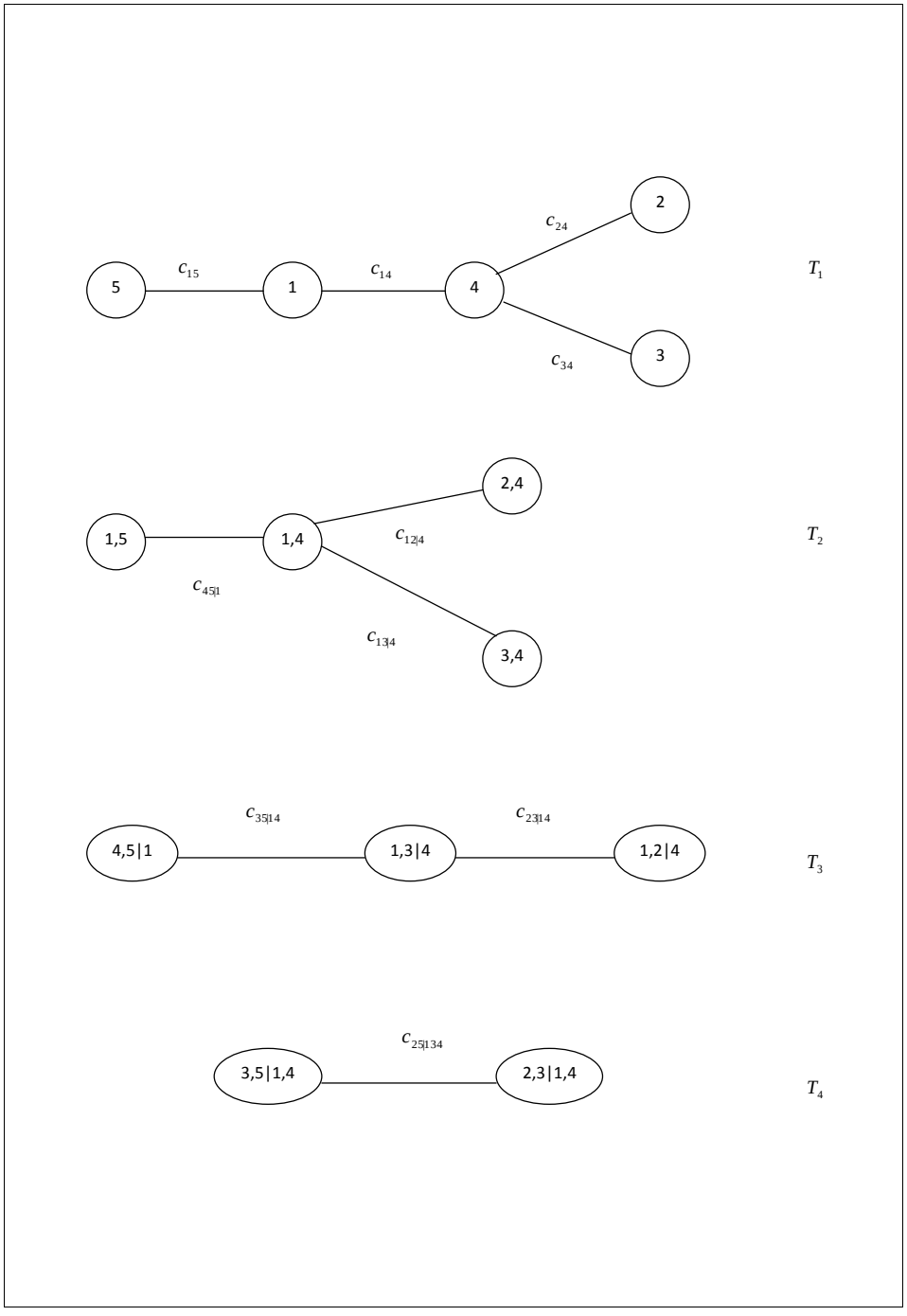


Figure 2.5: Tree Structure of R-Vine for 5 variables

$$
\begin{aligned}
f(x_1, x_2, x_3, x_4, x_5) = f_1(x_1) f_2(x_2) f_3(x_3) f_4(x_4) f_5(x_5) \\
& c_{14}(F_1(x_1), F_4(x_4)) c_{15}(F_1(x_1), F_5(x_5)) c_{24}(F_2(x_2), F_4(x_4)) \\
& c_{34}(F_3(x_3), F_4(x_4)) c_{12|4}(F_1(x_1), F_2(x_2)|F_4(x_4)) \\
& c_{13|4}(F_1(x_1), F_3(x_3)|F_4(x_4)) c_{45|1}(F_4(x_4), F_5(x_5)|F_1(x_1)) \\
& c_{23|14}(F_2(x_2), F_3(x_3)|F_1(x_1), F_4(x_4)) \\
& c_{35|14}(F_3(x_3), F_5(x_5)|F_1(x_1), F_4(x_4)) \\
& c_{25|134}(F_2(x_2), F_5(x_5)|F_1(x_1), F_3(x_3), F_4(x_4)).
\end{aligned}
\tag{2.2.5}
$$

As a technical brief, T_1 has 5 nodes and 4 edges and, nodes $(1,5)$ and $(1,4)$ in tree T_2 are joined by an edge sharing the node 1 coming from T_1 in Figure 2.5. This tree structure is just an example for R-vine, but the main focus of this study includes two special vine trees.

2.2.4 Special Types of Vines

There are infinitely many ways of constructing such a vine tree structure in a multivariate case. Among the others, two of them deserve more attention having special tree structure, introduce (C)anonical and (D)rawable vine copulas.

Definition 2.2.4 (C-vine)**.** *It is a type of regular vine distribution for which each tree has a unique node that is connected to all other nodes of the tree. It uses only star like trees and it is useful for ordering by importance. For example, for* 5 *dimensional multivariate density, assuming that the first variable is selected as the root node of the first tree (*T_1*) and the corresponding C-vine tree structure is represented in Figure 2.6. Besides, the corresponding probability density function can be written as,*

$$
\begin{aligned}
f(x_1, x_2, x_3, x_4, x_5) = {} & f_1(x_1) f_2(x_2) f_3(x_3) f_4(x_4) f_5(x_5) \\
& c_{12}(F_1(x_1), F_2(x_2)) c_{13}(F_1(x_1), F_3(x_3)) c_{14}(F_1(x_1), F_4(x_4)) \\
& c_{15}(F_1(x_1), F_5(x_5)) c_{23|1}(F_2(x_2), F_3(x_3)|F_1(x_1)) \\
& c_{24|1}(F_2(x_2), F_4(x_4)|F_1(x_1)) c_{25|1}(F_2(x_2), F_5(x_5)|F_1(x_1)) \\
& c_{34|12}(F_3(x_3), F_4(x_4)|F_1(x_1), F_2(x_2)) \\
& c_{35|12}(F_3(x_3), F_5(x_5)|F_1(x_1), F_2(x_2)) \\
& c_{45|123}(F_4(x_4), F_5(x_5)|F_1(x_1), F_2(x_2), F_3(x_3)).
\end{aligned}
\tag{2.2.6}
$$

Definition 2.2.5 (D-vine)**.** *D-vine is another special case for regular vine tree structure having no node in any tree is connected to more than two edges. It uses only path like trees and beneficial for temporal ordering of variables. For example, for* 5 *dimensional multivariate density, D-vine model is represented by Figure 2.7 and its density function can be written as,*

$$
\begin{aligned}
f(x_1, x_2, x_3, x_4, x_5) = {} & f_1(x_1) f_2(x_2) f_3(x_3) f_4(x_4) f_5(x_5) \\
& c_{12}(F_1(x_1), F_2(x_2)) c_{23}(F_2(x_2), F_3(x_3)) c_{34}(F_3(x_3), F_4(x_4)) \\
& c_{45}(F_4(x_4), F_5(x_5)) c_{13|2}(F_1(x_1), F_3(x_3)|F_2(x_2)) \\
& c_{24|3}(F_2(x_2), F_4(x_4)|F_3(x_3)) c_{35|4}(F_3(x_3), F_5(x_5)|F_4(x_4)) \\
& c_{14|23}(F_1(x_1), F_4(x_4)|F_2(x_2), F_3(x_3)) \\
& c_{25|34}(F_2(x_2), F_5(x_5)|F_3(x_3), F_4(x_4)) \\
& c_{15|234}(F_1(x_1), F_5(x_5)|F_2(x_2), F_3(x_3), F_4(x_4)).
\end{aligned}
\tag{2.2.7}
$$

For the sake of simplicity, the related dependence parameter is dropped for each unconditional and conditional density functions for equations in 2.2.6 and 2.2.7. The generalizations of both special vines are possible, the corresponding probability density function for C-vine and D-vine for p variables are formulated below, respectively.

$$
f(x_1, \cdots, x_p) = \left[\prod_{k=1}^{p} . f_k(x_k) \right] . \left[\prod_{i=1}^{p-1} \prod_{j=1}^{p-i} c_{i,i+j|1,\cdots,i-1} \right] . \tag{2.2.8}
$$

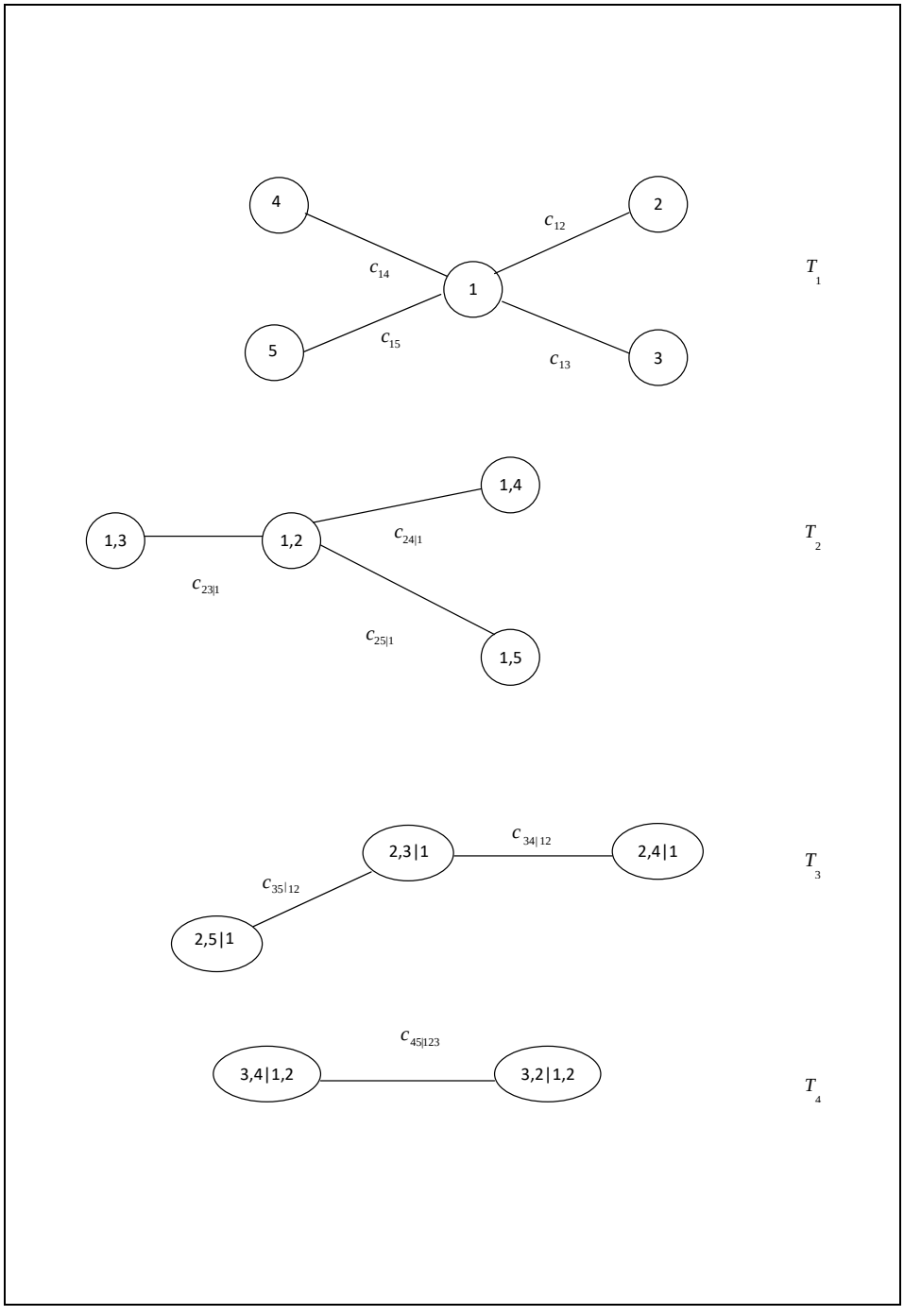


Figure 2.6: Tree Structure of C-Vine for 5 variables

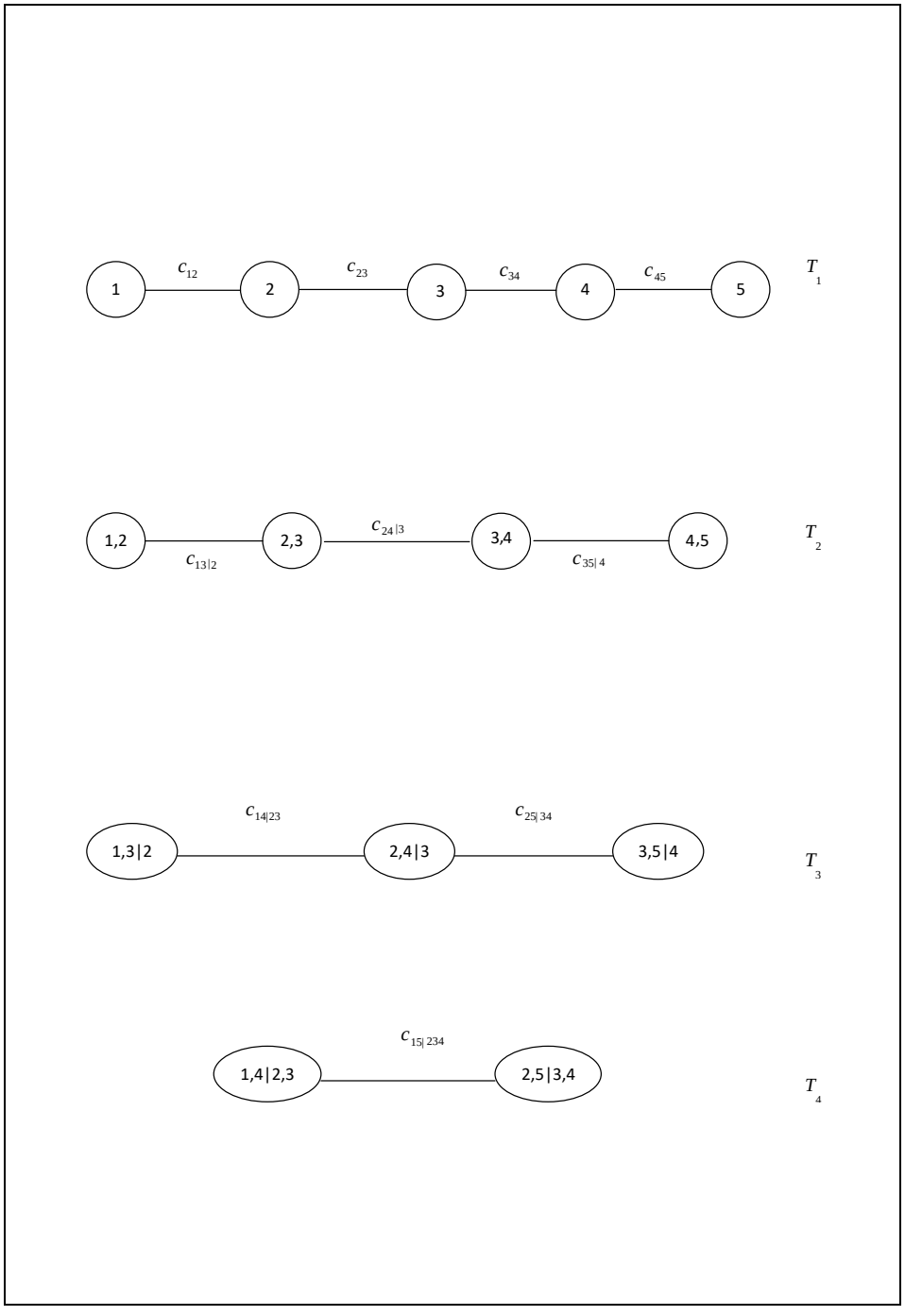


Figure 2.7: Tree Structure of D-Vine for 5 variables

$$f(x_1, \cdots, x_p) = \left[\prod_{k=1}^{p} . f_k(x_k)\right] . \left[\prod_{j=1}^{p-1} \prod_{i=1}^{p-j} c_{i,i+j|i+1,\cdots,i+j-1}\right] . \tag{2.2.9}$$

It is noteworthy to mention that, the likelihood calculation for D-vine is more complicated than that for the C-vine. For the detailed model selection procedure will be established within the context of mixture of vines in Chapter 3 later on, but still the work flow for fitting any suitable vine model is vital to emphasize after defining vine copulas. For this reason, modified sequential selection, discussed in [20], was summarized as a closing part of this section. For any regular vine construction, the following steps could be followed for the model selection:

1. Determine which copula family is suitable in T_1 by checking any tail dependence or asymmetries with the plot of original data.
2. Estimate the related parameters for the selected copula families using the given data.
3. Transform the necessary observations for T_2, by using copula parameters in T_1 and the conditional (h-) functions.
4. Determine now the best suitable family in T_2, like the followed procedure in T_1.
5. Continue this process until describing the properties of last tree, for instance T_5 is the last tree for R-vine in Figure 2.5, and other two specific cases given in Figures 2.6 and 2.7.

Here, the main drawback of this mechanism is not insuring a globally optimal fit. As another approach, information-based model inference or optimal truncation methods are also available in the literature, but those topics are out of scope of this dissertation. Interested reader could follow Chapters 3, 11 and 12 of [20]. In this study, mixture of C-vines has been focused with archimedean type pair copula families. Besides, in 3-dimensional case, C- and D-vine models are identical [1] so that the proposed mixture models cover both special types. In 4-dimensional case, however, C-vine mixture model differs certainly from D-vine structure.

CHAPTER 3

MIXTURE OF VINES

In this chapter, main motivation behind finite mixture models is presented with recent studies in the literature. Afterwords, two different proposed mixture models are discussed within the following three subsequent sections. First of all, finite mixture of C-vines are studied with various scenarios and different parameter estimation methods. Thereafter, same finite mixture model framework is applied to both C- and D-vine to construct another type of multivariate data. Finally, as a novel model, CD-vine approach was proposed under a distinct multivariate setting with complex dependence patterns. Primarily, the parameter estimation part is the focus of this thesis study and numerous optimization methods are considered to find the estimated parameters of mixture models.

3.1 Brief Literature

In the literature, there are limited studies about such type of mixture models in terms of vines. Besides, until this time, only D-vines are considered under finite mixture model to construct a more sophisticated multivariate density functions [22], [33]. Apart from these studies, most of the studies related to mixture model primarily based on the implementation of a limited set of copula families, but applied in various disciplines.

Previously, Vrac et al. [37] studied on partitioning the atmospheric profiles of temperature and humidity to characterize air masses. A model-based clustering approach has been applied to a set of data for capturing air masses and Frank copula family

is considered to identify the dependence between the variables. They developed a model which allows to work on distribution of data, instead of only numerical values to cluster the profiles in terms of mixture density, with the satellite observations. The results of proposed model resulted in relevant clusters for climatology [37]. Within the same year, Cuvelier and Noirhomme-Fraiture [11] proposed a mixture of Clayton copulas for clustering purposes again. They extended the model proposed by [37] from bivariate case to higher dimensions, with n-dimensional Clayton copula. It has been proven that the mixture model with the Clayton copula was successful to identify dependence structure. Besides, clustering based on Clayton family gave better results like other classical methods [11]. Afterwords, Hu [18] employed a mixture of Gaussian, Gumbel and Gumbel Survival to model the dependence of monthly returns between a pair of stock indexes by capturing left or right tail dependence. This method relied on two parameter set; dependence and weight parameters, which control the degree and shape of dependence, respectively. Four stock market indices from different regions were considered in monthly base and empirical findings have powerful insights about risk management and asset pricing [18].

Several years later, Arakelian and Karlis [3] investigated the finite mixture of copulas for clustering purposes with respect to dependence properties of data. Finite mixture of copulas were exploited with its certain properties and illustrated with the help of widely used bivariate copula families, namely Gaussian, Gumbel, Frank, Joe and Clayton copulas. In their study, both static and dynamic weight parameters for the finite mixture model were studied and the potential of their method was illustrated using daily log-returns from major stock markets. From a different perspective, these approaches are extended to higher dimensions with vine copula methodology, the departure point of this thesis [22]. A mixture of D-vine copulas was proposed in order to reveal and fully understand the complex and hidden dependence patterns in a multivariate setting [22]. Both simulated and real life data examples were employed in detail to understand the convenience of the proposed D-vine mixture and daily precipitation data was considered to detect the association among four meteorological stations located in four different municipalities. A similar research has been conducted by [33] for clustering purposes using various set of available data. In that respect, they have been compared the classical clustering methods with D-vine based mixture

model with a limited set of copula families. Thereafter, Kosmidis and Karlis [24] implemented copula framework for clustering purposes using both continuous and discrete multivariate data. Besides, they highlighted outstanding findings about the parameter identification depending on the choice of copulas for the components, including rotated versions for the known families.

In the light of above summarized studies, the primary goal of this thesis is manifold. As a first contribution, the study of [22] is extended by covering C-vines in finite mixture model, instead of looking at only D-vine. Besides, for the construction of mixture models, various copula families are considered for each component. For this study, well known and widely applicable archimedean copulas are focused under proposed models and some of their rotated versions are also covered, like it was discussed by [24]. Secondly, the same finite mixture model is implemented to combine both C- and D-vines. Finally, as a novel contribution of the study, CD-vine model has been proposed, unlike finite mixture model. In this framework, finite C-vine models are incorporated with the help of D-vine copulas. All of above proposed models, main interest of this study is the parameter estimation of the mixture models via different optimization routines. Moreover, model identification problem and improvement of model selection procedure has been discussed, instead of focusing on clustering.

3.2 Proposed Mixture Models

In this dissertation, numerous improvements of the study of [22] are considered. In order to derive distinct finite mixture of vine copulas, two main scenarios are considered. As a first one, 2 component 3 dimensional C-vine mixture models with different copula families are studied. In this modeling setup, mixture of C-vines are equivalent to D-vine mixture because of the dimensionality. In this 3 dimensional case, 3 component C-vine mixtures are also considered with same and different copula families. Secondly, the dimension has been increased and 2 component, 4 dimensional C-vine mixture models are investigated with various copula families. Afterwords, similar approach is studied for the combination of C- and D-vines. In this framework, 2 component 4 dimensional C- and D-vine finite mixture is investigated via predefined multivariate data set. In all cases, the main focus is the Frank pair copula during the

construction.

After studying above finite mixtures, as a novel improvement, CD-vine mixture model have been proposed. Under this setup, finite number of C-vine copula models are incorporated into a D-vine copula model to derive the dependence between several random variables. The multivariate distribution functions are derived empirically for each component of mixture model then tied together by the help of D-vine copula by considering the temporal ordering of the components. In this modeling framework, the parameter estimation and inference for full multivariate density is required to maximize log-likelihood function in two separate steps, the maximization of the log-likelihood function from C-vine first and then from D-vine part. This two step process is similar to the idea of Inference for Margins (IFM) technique, widely considered method for parameter estimations [19].

Within the following subsections, inference for finite mixture of C-vines and the recently established CD-vine mixture model is presented. For this reason, the detailed derivations for each mixture model is elaborated in terms of various copula families. Thereafter, the considered optimization tools are briefly covered at the end, the model selection procedure with modified GOF tests is exhibited finally.

3.2.1 Inference for Mixture of C- and D- Vines

Under finite mixture model, generally, p-dimensional C-vine copula density can be written as below,

$$f(\boldsymbol{x};\boldsymbol{\phi_{CV}}) = \prod_{k=1}^{p} f_k(x_k) \prod_{i=1}^{p-1}\prod_{j=1}^{p-i} c_{i,i+j|1:(i-1)} \\ (F(x_i|x_1,\cdots,x_{i-1}), F(x_{i+j}|x_1,\cdots,,x_{i-1});\boldsymbol{\beta}_{i,(i+j)|1:(i-1)}). \tag{3.2.1}$$

where $f_k(x_k)$ denotes the marginal densities, $c_{i,i+j|1:(j-1)}$ are the bivariate copula density functions with parameter(s) $\boldsymbol{\beta_{i,(i+j)|1:(i-1)}}$, $F(x_k|\cdots)$ denotes the conditional distribution functions and $\boldsymbol{\phi_{CV}}$ is the set of all parameters in p-dimensional C-vine density [6].

There exist one root node in the tree construction of C-vine model which results in following density formulation in 4-dimension given as [1],

$$\begin{aligned} f(x_1, x_2, x_3, x_4; \boldsymbol{\phi_{CV}}) = c_{12}(F(x_1), F(x_2); \beta_{12}) c_{13}(F(x_1), F(x_3); \beta_{13}) \\ c_{14}(F(x_1), F(x_4); \beta_{14}) \\ c_{23|1}(F(x_2|x_1), F(x_3|x_1); \beta_{23|1}) c_{24|1}(F(x_2|x_1), F(x_4|x_1); \beta_{24|1}) \\ c_{34|12}(F(x_3|x_1, x_2), F(x_4|x_1, x_2); \beta_{34|12}) \prod_{k=1}^{4} f_k(x_k). \end{aligned} \tag{3.2.2}$$

Here, the parameter vector for the above written density function in Equation 3.2.2 is $\phi_{CV} = (\beta_{12}, \beta_{13}, \beta_{14}, \beta_{23|1}, \beta_{24|1}, \beta_{34|12})$. Under such multivariate framework, full inference on C-vine copula can be derived using the log-likelihood function presented as [1],

$$\begin{aligned} L(\boldsymbol{\phi}) = \sum_{i=1}^{p-1} \sum_{j=1}^{p-i} \sum_{n=1}^{N} \\ \log c_{i,i+j|1:(i-1)}(F(x_{i,n}|x_{1,n}, ..., x_{i-1,n}), F(x_{i+j,n}|x_{1,n}, ..., x_{i-1,n}); \boldsymbol{\beta}_{i,(i+j)|1:(i-1)}). \end{aligned} \tag{3.2.3}$$

and it requires three consecutive steps [22]:

1. Decide which variable should be used as a root node in the first tree T_1 of a C-vine copula (i.e. joining the variables in which the root node variable is selected based on its significant relations with other variables)
2. Then, specify the family type and parametric shape of each pair-copula function in an assumed C-vine copula model.
3. Estimate all parameters of C-vine by maximizing the log-pseudo likelihood function given in Equation 3.2.3

The finite mixture model is introduced for combining m component C-vine densities to detect complex and hidden dependence structure in a multivariate data. Under this

establishment, EM algorithm is used to estimate the parameters in mixture model. To illustrate, suppose a p-dimensional random vector $\boldsymbol{X}$=($\boldsymbol{X_1}$, $\cdots$, $\boldsymbol{X_p}$) is generated from a mixture of M- component C-vine densities, where its density function is defined as,

$$g(\boldsymbol{x}, \boldsymbol{\theta}) = \sum_{m=1}^{M} \pi_m f(\boldsymbol{x}, \boldsymbol{\phi_m}). \tag{3.2.4}$$

where π_m is the mixing proportion of the m-th component such that $0 < \pi_m < 1$ and $\sum_{m=1}^{M} \pi_m = 1$. Besides, $\boldsymbol{\phi_m}$ is the m-th component-specific parameter vector for the C-vine density described in Equation 3.2.4 [22]. In this notation, θ is the set of all parameters with dimension p and parameter space is denoted by Θ, which includes the full product space (i.e. the simplex of π_m and the cross product space of ϕ_m). Besides, p is the total number of free parameters to be estimated in the mixture model and $p = (M-1) + \sum_{m=1}^{M} dim(\phi_m)$. For the estimation of Equation 3.2.4, both the number of components M and the parameters θ are required to estimate based on EM-algorithm, proposed by [12].

Assume that observations with sample size N randomly drawn from a M component C-vine density given in Equation 3.2.4, denoted as $\boldsymbol{x_k}$=($\boldsymbol{x_{k,1}}$, $\cdots$, $\boldsymbol{x_{k,N}}$) where $k = 1, \cdots, p$. Then, log-likelihood of θ, $g(\boldsymbol{x}, \boldsymbol{\theta})$ is described as,

$$L(\boldsymbol{\theta}) = \log(\prod_{n=1}^{N} g(\boldsymbol{x}_n, \boldsymbol{\theta})) = \log(\prod_{n=1}^{N} \sum_{m=1}^{M} \pi_m f(\boldsymbol{x}_n, \boldsymbol{\phi}_m)). \tag{3.2.5}$$

To show that whether x_n drawn from the m-th component ($z_{nm} = 1$) or not ($z_{nm} = 0$), z_n=($z_{n1}, \cdots, z_{nm}, \cdots, z_{nM}$) latent variables are generated. Here, z_n is independently and identically distributed (i.i.d.) random variable from a multinomial distribution. More formally, z_n is $Mult(M, \pi = (\pi_1, \cdots, \pi_m))$ where the log-likelihood for the complete data set y_n=(x_n,z_n), namely $L(\boldsymbol{\theta})_c$, is given as,

$$
\begin{aligned}
L(\boldsymbol{\theta})_c &= \log \prod_{n=1}^{N} \prod_{m=1}^{M} [\pi_m f(\boldsymbol{x}_n, \boldsymbol{\phi}_m)]^{z_{nm}} \\
&= \sum_{n=1}^{N} \sum_{m=1}^{M} z_{nm} \log(\pi_m) + \sum_{n=1}^{N} \sum_{m=1}^{M} z_{nm} \log(f(\boldsymbol{x}_n, \boldsymbol{\phi}_m)).
\end{aligned}
\tag{3.2.6}
$$

For the inference on $L(\boldsymbol{\theta})_c$, the repeated E-th and M-th steps of the EM algorithm are executed to compute the successive estimates, θ^s, based on initial values θ^0. This repetitive procedure is the key point for parameter estimations and described as follows:

[**E'th-step**]

It calculates the conditional expectation of $L(\theta)_c$ given the observed data and current parameter estimates for θ. This computation is equivalent to finding posterior probability of x_n belongs to the m-th component, given the current values of the parameters which is formulated as [22],

$$
\widehat{z_{nm}}^{(s)} = E[z_{nm}|\boldsymbol{x}, \theta^s] = P[z_{nm} = 1|\boldsymbol{x}, \theta^s] = \frac{\pi_m^{(s)} f(\boldsymbol{x}_n, \boldsymbol{\phi}_m^{(s)})}{\sum_{l=1}^{M} \pi_l^{(s)} f(\boldsymbol{x}_n, \boldsymbol{\phi}_l^{(s)})}. \tag{3.2.7}
$$

[**M'th-step**]

This second step estimates the parameters for each component independently, ($\pi_1^{(s+1)}$, $\cdots$, $\pi_m^{(s+1)}$,$\cdots$, $\pi_M^{(s+1)}$) and ($\phi_1^{(s+1)}, \cdots, \phi_m^{(s+1)}, \cdots, \phi_M^{(s+1)}$) by maximizing the expected complete-data log-likelihood obtained from E'th-step.

For the weights, there exists a closed form solution, denoted by $\pi_m^{(s+1)} = \frac{\sum_{n=1}^{N} \widehat{z_{nm}}^{(s)}}{N}$. Afterwords, the estimation of $\phi_m^{(s+1)}$ in the m-th component C-vine density function is equivalent to obtain the parameter estimates weighted by $\widehat{z_{nm}}^{(s)}$ for the parameters in a C-vine density in Equation 3.2.4. In this case, the second term in Equation

3.2.6 must be maximized to obtain the estimation for ϕ_m. For each component, this maximization can be considered independently.

Under this framework, both steps are continued until the predefined termination criteria is satisfied (i.e. $L(\theta^{s+1}) - L(\theta^s) < 10^{-6}$ or 10^{-3}). As a result of a nice property of EM algorithm, the log-likelihood is not decreased during the iteration. For this reason, the given algorithm is run with suitable starting values drawn from the parameter space and parameter estimates are calculated as a mean value of multiple iterations. For the model comparison part, three well known model selection criteria values are used, as they have been considered in the study of [22].

- Akaike's Information Criterian

$$AIC = -2\log(L(\widehat{\theta})) + 2p. \tag{3.2.8}$$

- Bayesian Information Criterian

$$BIC = -2\log(L(\widehat{\theta})) + p\log(n). \tag{3.2.9}$$

- Consistent AIC (CAIC)

$$CAIC = -2\log(L(\widehat{\theta})) + p(\log(n) + 1). \tag{3.2.10}$$

where $\widehat{\theta}$ is the estimate of p-dimensional θ defined in Equation 3.2.4 and p is the total number of parameters to be estimated in the mixture model. For the model selection part, as a supporting tool, suitable Goodness-of-fit (GOF)tests are considered. For this reason, both Clarke and Vuong GOF tests, proposed by [10] and [38], are modified to compare two different mixture of vines. Additionally, details for this modification is summarized at the end of this chapter. To sum up, the receipt for the full inference on the mixture of C-vines can be summarized as follows:

1. Derive the normalized ranks of p-dimensional observed data
2. Decide the root node of each C-vine density by calculating all pairwise correlations.
3. Consider different copulas for all pairs to specify the family and parametric shape of each pair-copula in an assumed C-vine

4. Given a copula family, investigate a vine mixture with M components and estimate its parameters in each model by employing the EM- algorithm

5. Select the best model by comparing all available models in terms of model selection tools (AIC, BIC, CAIC values and GOF tests).

Above summarized procedure is valid also for D-vine with small differences. In the simulation part, finite mixture of C- and D-vines is proposed and has been tested with simulated data set. For this reason, D-vine modeling steps are summarized below similar to the C-vine model. In general, p-dimensional D-vine copula density can be written as,

$$f(\boldsymbol{x};\boldsymbol{\phi_{DV}}) = \prod_{k=1}^{p} f_k(x_k) \prod_{i=1}^{p-1} \prod_{j=1}^{p-i} c_{j,(j+i)|(j+1):(j+i-1)} \\ (F(x_j|x_{j+1},...,x_{j+i-1}), F(x_{j+1}|x_{j+1},...,x_{j+i-1}); \boldsymbol{\beta}_{j,(j+i)|(j+1):(j+i-1)}). \tag{3.2.11}$$

where $f_k(x_k)$ denotes the marginal densities, $c_{j,(j+i)|(j+1):(j+i-1)}$ are the bivariate copula density functions with parameter(s) $\beta_{j,(j+i)|(j+1):(j+i-1)}$, and ϕ_{DV} is the set of all parameters in D-vine density. For instance, in 4 dimensional case (i.e. $p = 4$), Equation 3.2.11 becomes;

$$\begin{aligned} f(x_1,x_2,x_3,x_4;\boldsymbol{\phi_{DV}}) = {} & c_{12}(F(x_1),F(x_2);\beta_{12}) c_{23}(F(x_2),F(x_3);\beta_{23}) \\ & c_{34}(F(x_3),F(x_4);\beta_{34}) c_{13|2}(F(x_1|x_2),F(x_3|x_2);\beta_{13|2}) \\ & c_{24|3}(F(x_2|x_3),F(x_4|x_3);\beta_{24|3}) \\ & c_{14|23}(F(x_1|x_2,x_3),F(x_4|x_2,x_3);\beta_{14|23}) \prod_{k=1}^{4} f_k(x_k). \end{aligned} \tag{3.2.12}$$

where ϕ_{DV}=($\beta_{12}, \beta_{23}, \beta_{34}, \beta_{13|2}, \beta_{24|3}, \beta_{14|23}$) includes totally 6 copula parameters used in unconditional and conditional pairs. The full inference on D-vine copula is directly based on the identification of the pair copulas and estimation of their parameters, similar to C-vine framework.

Assuming that N observed samples are available like $\mathbf{x}_k = (\mathbf{x}_{k,1}, ..., \mathbf{x}_{k,N})$, where $k = 1, 2, ..., p$ and each random variable X_k follows standard uniform distribution (s.t. $f_k(x_k = 1)$). Commonly, when the distributions of margins are not known, the normalized ranks of the data are used for the inference purposes as they keep the largest amount of information about the joint dependence between variables [29]. In general, log-likelihood function for D-vine density over the parameters and its exemplified version in case of $p = 4$ is given in equations below, respectively.

$$L(\boldsymbol{\phi_{DV}}) = \sum_{i=1}^{p-1} \sum_{j=1}^{p-i} \sum_{n=1}^{N} \log c_{j,(j+i)|(j+1):(j+i-1)} \\ (F(x_{j,n}|x_{j+1,n}, ..., x_{j+i-1,n}), F(x_{j+1,n}|x_{j+1,n}, ..., x_{j+i-1,n}); \boldsymbol{\beta}_{j,(j+i)|(j+1):(j+i-1)}). \tag{3.2.13}$$

where again, ϕ_{DV} is the set of all parameters in D-vine density. For $p = 4$, Equation 3.2.13 becomes as follows, with its corresponding parameters,

$$\begin{aligned} L(\boldsymbol{\phi_{DV}}) = \sum_{n=1}^{N} [&\log c_{12}(x_{1,n}, x_{2,n}; \beta_{12}) + \log c_{23}(x_{2,n}, x_{3,n}; \beta_{23}) + \log c_{34}(x_{3,n}, x_{4,n}; \beta_{34}) \\ &+ \log c_{13|2}(F(x_{1,n}|x_{2,n}), F(x_{3,n}|x_{2,n}); \beta_{13|2}) \\ &+ \log c_{24|3}(F(x_{2,n}|x_{3,n}), F(x_{4,n}|x_{3,n}); \beta_{24|3}) \\ &+ \log c_{14|23}(F(x_{1,n}|x_{2,n}, x_{3,n}), F(x_{4,n}|x_{2,n}, x_{3,n}); \beta_{14|23})]. \end{aligned} \tag{3.2.14}$$

Under this modeling setup, full inference is accomplished by the following consecutive steps [22]:

1. Decide which variables are used at the first tree T_1 of a D-vine copula and order of variables using (tail) dependencies (i.e. joining the variables which have the strongest (tail) dependencies)

2. Specify the family and parametric shape of each pair-copula in an assumed D-vine copula

3. Given a copula family, fit a mixture of vines with M components and estimate its parameters in each model by employing the EM-algorithm

4. Select the best model by comparing all available models in terms of model selection tools (AIC, BIC, CAIC values and GOF tests).

For the components of finite mixture of C-vines, various copula families are implemented like Clayton, Gumbel, Frank families and the possible rotated versions to extend the pool of copula families. Here, this construction is exemplified with Clayton and Survival Clayton families by assuming the case of 2 component, 3 dimensional mixture model. Besides, Gumbel family is illustrated for the mixture model, constructed in 4 dimension. With the same family, CD-vine mixture representation is presented at the end of model construction part.

3.2.1.1 2 Component 3 dimensional C-vine Mixtures

To construct a 2 Component 3 dimensional mixture of C-vines, all pair copulas in the 1'st and 2'nd component are assumed to be Clayton and Survival Clayton family, respectively. For the sake of simplicity, derivation for the 1'st component of mixture model is presented in detail below. Thereafter, the obtained density functions for the 2'nd component are directly summarized.

For the first component, all pairs are modeled by Clayton copula and its copula function for pairs (x_1, x_2) with parameter β_{12} is defined as,

$$C(x_1, x_2; \beta_{12}) = (x_1^{-\beta_{12}} + x_2^{-\beta_{12}} - 1)^{\frac{-1}{\beta_{12}}} \exp(-[(-\log x_1)^{\beta_{12}} + (-\log x_2)^{\beta_{12}}]^{\frac{1}{\beta_{12}}}). \tag{3.2.15}$$

and its probability density function obtained using the second order derivative of $C(x_1, x_2; \beta_{12})$ as follows (after cancellations and possible rearrangements);

$$
\begin{aligned}
c(x_1, x_2; \beta_{12}) &= \frac{\partial^2 C(x_1, x_2; \beta_{12})}{\partial x_1 \partial x_2} \\
&= \frac{\partial[\frac{-1}{\beta_{12}}(x_1^{-\beta_{12}} + x_2^{-\beta_{12}} - 1)^{\frac{-1}{\beta_{12}}-1}(-\beta_{12})x_2^{-\beta_{12}-1}]}{\partial x_1} \\
&= \frac{\partial[(x_1^{-\beta_{12}} + x_2^{-\beta_{12}} - 1)^{\frac{-1}{\beta_{12}}-1}x_2^{-\beta_{12}-1}]}{\partial x_1} \\
&= (-1-\beta_{12})\frac{(x_1^{-\beta_{12}} + x_2^{-\beta_{12}} - 1)^{\frac{-1}{\beta_{12}}-2}(-\beta_{12})x_1^{-\beta_{12}-1}x_2^{-\beta_{12}-1}}{\beta_{12}} \\
&= (\beta_{12}+1)(x_1x_2)^{-\beta_{12}-1}(x_1^{-\beta_{12}} + x_2^{-\beta_{12}} - 1)^{\frac{-1}{\beta_{12}}-2}.
\end{aligned}
$$

Similarly, $c(x_1, x_3; \beta_{13})$ can be written as,

$$
\begin{aligned}
c(x_1, x_3; \beta_{13}) &= \frac{\partial^2 C(x_1, x_3; \beta_{13})}{\partial x_1 \partial x_3} \\
&= \frac{\partial[\frac{-1}{\beta_{13}}(x_1^{-\beta_{13}} + x_3^{-\beta_{13}} - 1)^{\frac{-1}{\beta_{13}}-1}(-\beta_{13})x_3^{-\beta_{13}-1}]}{\partial x_1} \\
&= \frac{\partial[(x_1^{-\beta_{13}} + x_3^{-\beta_{13}} - 1)^{\frac{-1}{\beta_{13}}-1}x_3^{-\beta_{13}-1}]}{\partial x_1} \\
&= (-1-\beta_{13})\frac{(x_1^{-\beta_{13}} + x_3^{-\beta_{13}} - 1)^{\frac{-1}{\beta_{13}}-2}(-\beta_{13})x_1^{-\beta_{13}-1}x_3^{-\beta_{13}-1}}{\beta_{13}} \\
&= (\beta_{13}+1)(x_1x_3)^{-\beta_{13}-1}(x_1^{-\beta_{13}} + x_3^{-\beta_{13}} - 1)^{\frac{-1}{\beta_{13}}-2}. \qquad (3.2.16)
\end{aligned}
$$

For the calculation of conditional density of $c_{23|1}$, it is necessary to use h-functions and derive $F(x_2|x_1)$ and $F(x_3|x_1)$. Their computations are briefly presented below in detail.

$$
\begin{aligned}
x_{21} = F(x_2|x_1) &= \frac{\partial C(x_1, x_2; \beta_{12})}{\partial x_1} \\
&= x_1^{-\beta_{12}-1}(x_1^{-\beta_{12}} + x_2^{-\beta_{12}} - 1)^{\frac{-1}{\beta_{12}}-1}. \qquad (3.2.17)
\end{aligned}
$$

$$
\begin{aligned}
x_{31} = F(x_3|x_1) &= \frac{\partial C(x_1, x_3; \beta_{13})}{\partial x_1} \\
&= x_1^{-\beta_{13}-1}(x_1^{-\beta_{13}} + x_3^{-\beta_{13}} - 1)^{\frac{-1}{\beta_{13}}-1}. \qquad (3.2.18)
\end{aligned}
$$

Then, obtained values of x_{21} and x_{31} are used as inputs of the conditional density function $c_{23|1}$ with Clayton formula given above. Thereafter, conditional density $c_{23|1}$ is derived as,

$$\begin{aligned} c_{23|1} &= c_{23|1}(x_{21}, x_{31}; \beta_{23|1}) \\ &= (\beta_{23|1}+1)(x_{21}x_{31})^{-\beta_{23|1}-1}((x_{21})^{-\beta_{23|1}} + (x_{31})^{-\beta_{23|1}} - 1)^{\frac{-1}{\beta_{23|1}}-2}. \end{aligned} \tag{3.2.19}$$

when x_{21} and x_{31} are replaced in Equation 3.2.19 to write $c_{23|1}$ as a function of x_1, x_2 and x_3 with required parameters, namely β_{12}, β_{13} and $\beta_{23|1}$, obtained as follows,

$$\begin{aligned} c_{23|1} &= c_{23|1}(x_{21}, x_{31}; \beta_{23|1}) \\ &= (\beta_{23|1}+1)(x_1^{-\beta_{12}-1}(x_1^{-\beta_{12}} + x_2^{-\beta_{12}} - 1)^{\frac{-1}{\beta_{12}}-1} \\ &\quad x_1^{-\beta_{13}-1}(x_1^{-\beta_{13}} + x_3^{-\beta_{13}} - 1)^{\frac{-1}{\beta_{13}}-1})^{-\beta_{23|1}-1} \\ &\quad ((x_1^{-\beta_{12}-1}(x_1^{-\beta_{12}} + x_2^{-\beta_{12}} - 1)^{\frac{-1}{\beta_{12}}-1})^{-\beta_{23|1}} \\ &\quad +(x_1^{-\beta_{13}-1}(x_1^{-\beta_{13}} + x_3^{-\beta_{13}} - 1)^{\frac{-1}{\beta_{13}}-1})^{-\beta_{23|1}} - 1)^{\frac{-1}{\beta_{23|1}}-2} \\ &= (\beta_{23|1}+1)(x_1^{-\beta_{12}-\beta_{13}-2})^{-\beta_{23|1}-1} \\ &\quad ((x_1^{-\beta_{12}} + x_2^{-\beta_{12}} - 1)^{\frac{-1}{\beta_{12}}-1}(x_1^{-\beta_{13}} + x_3^{-\beta_{13}} - 1)^{\frac{-1}{\beta_{13}}-1})^{\frac{-1}{\beta_{23|1}}-1} \\ &\quad ((C_{12}(x_1, x_2; \beta_{12}))^{-\beta_{23|1}} + (C_{13}(x_1, x_3; \beta_{13}))^{-\beta_{23|1}} - 1)^{\frac{-1}{\beta_{23|1}}-2}. \end{aligned} \tag{3.2.20}$$

where both

$$C_{12} = (x_1^{-\beta_{12}-1}(x_1^{-\beta_{12}} + x_2^{-\beta_{12}} - 1)^{\frac{-1}{\beta_{12}}-1}). \tag{3.2.21}$$

$$C_{13} = (x_1^{-\beta_{13}-1}(x_1^{-\beta_{13}} + x_3^{-\beta_{13}} - 1)^{\frac{-1}{\beta_{13}}-1}). \tag{3.2.22}$$

are functions of pairs (x_1, x_2) and (x_1, x_3) with the corresponding parameters β_{12} and β_{13}, respectively. Then, 3-dimensional C-vine density function, which describes the dependence pattern in 1'st component with Clayton pairs, given below,

$$\log f_1 = [\log c_{12} + \log c_{13} + \log c_{23|1}]$$
$$= \log(\beta_{12}+1) + (-\beta_{12}-1)\log(x_1x_2) + (\frac{-1}{\beta_{12}}-2)\log(x_1^{-\beta_{12}} + x_2^{-\beta_{12}} - 1)$$
$$+ \log(\beta_{13}+1) + (-\beta_{13}-1)\log(x_1x_3) + (\frac{-1}{\beta_{13}}-2)\log(x_1^{-\beta_{13}} + x_3^{-\beta_{13}} - 1)$$
$$+ \log(\beta_{23|1}+1) + (-\beta_{23|1}-1)(-\beta_{12}-\beta_{13}-2)\log(x_1) +$$
$$(-\beta_{23|1}-1)((\frac{-1}{\beta_{12}}-1)\log(x_1^{-\beta_{12}} + x_2^{-\beta_{12}} - 1) + (\frac{-1}{\beta_{13}}-1)\log(x_1^{-\beta_{13}} + x_3^{-\beta_{13}} - 1))$$
$$+(\frac{-1}{\beta_{23|1}}-2)\log((x_{21})^{-\beta_{23|1}} + (x_{31})^{-\beta_{23|1}} - 1). \tag{3.2.23}$$

In a similar manner, all unconditional and conditional densities of 2'nd component based on Survival Clayton family can be described. Two main identities for the construction of survival copulas in bivariate case (rotated with $180°$) are presented below for both copula distribution and density functions, abbreviated as $C^{180}(x_1, x_2)$ and $c^{180}(x_1, x_2)$, respectively. The meaning of such a rotation can be visualized in Figure 2.4 in Chapter 2.

$$C^{180}(x_1, x_2) = x_1 + x_2 - 1 + C(1 - x_1, 1 - x_2). \tag{3.2.24}$$

$$c^{180}(x_1, x_2) = c(1 - x_1, 1 - x_2). \tag{3.2.25}$$

where $C(1 - x_1, 1 - x_2)$ and $c(1 - x_1, 1 - x_2)$ are transformed versions of original functions in bivariate case. In this study, one of the rotated copula families is Survival Clayton (SClay). Here, the log-likelihood function construction for the 2'nd component, having all pair copulas are SClay are described below. For 3 dimensional case, c_{12}^{SClay} and c_{13}^{SClay} are defined by using above identities given in Equations 3.2.24 and 3.2.25.

$$c_{12}^{SClay}(x_1, x_2; \beta_{12}) = c_{12}^{SClay} = (1+\beta_{12})((x_1-1)(x_2-1))^{-\beta_{12}-1}$$
$$((1-x_1)^{-\beta_{12}} + (1-x_2)^{-\beta_{12}} - 1)^{\frac{-1}{\beta_{12}}-2}. \tag{3.2.26}$$

$$c_{13}^{SClay}(x_1, x_3; \beta_{13}) = c_{13}^{SClay} = (1+\beta_{13})((x_1-1)(x_3-1))^{-\beta_{13}-1}$$
$$((1-x_1)^{-\beta_{13}} + (1-x_3)^{-\beta_{13}} - 1)^{\frac{-1}{\beta_{13}}-2}. \tag{3.2.27}$$

For the conditional density function of SClay, again $x_{21}^{SClay} = F(x_2|x_1)$ and $x_{31}^{SClay} = F(x_3|x_1)$ values are derived for the calculation of $c_{23|1}^{SClay}$, their computations are presented below in detail,

$$x_{21}^{SClay} = F(x_2|x_1) = \frac{\partial C^{180}(x_1, x_2)}{\partial x_1}$$
$$= 1 - (1-x_1)^{-\beta_{12}-1}((1-x_1)^{-\beta_{12}} + (1-x_2)^{-\beta_{12}} - 1)^{\frac{-1}{\beta_{12}}-1}. \quad (3.2.28)$$

$$x_{31}^{SClay} = F(x_3|x_1) = \frac{\partial C^{180}(x_1, x_3)}{\partial x_1}$$
$$= 1 - (1-x_1)^{-\beta_{13}-1}((1-x_1)^{-\beta_{13}} + (1-x_3)^{-\beta_{13}} - 1)^{\frac{-1}{\beta_{13}}-1}. \quad (3.2.29)$$

Thereafter, conditional density function, $c_{23|1}^{SClay}$, has been derived by following the similar construction using $x_{21}^{SClay} = x_{21}$ and $x_{31}^{SClay} = x_{31}$ are input values.

$$c_{23|1}^{SClay}(x_{21}, x_{31}; \beta_{23|1}) = c_{23|1}^{SClay} = (1+\beta_{23|1})$$
$$((x_{21}-1)(x_{31}-1))^{-\beta_{23|1}-1}$$
$$((1-x_{21})^{-\beta_{23|1}} + (1-x_{31})^{-\beta_{23|1}} - 1)^{\frac{-1}{\beta_{23|1}}-2}. \quad (3.2.30)$$

After deriving all those density functions, log-likelihood function from C-vine having SClay pair copulas can be defined as,

$$\begin{aligned}
\log f_1 &= [\log c_{12}^{SClay} + \log c_{13}^{SClay} + \log c_{23|1}^{SClay}] \\
&= \log(\beta_{12}+1) + (-\beta_{12}-1)\log((x_1-1)(x_2-1)) \\
&+ (\frac{-1}{\beta_{12}} - 2)\log((1-x_1)^{-\beta_{12}} + (1-x_2)^{-\beta_{12}} - 1) \\
&+ \log(\beta_{13}+1) + (-\beta_{13}-1)\log((x_1-1)(x_3-1)) \\
&+ (\frac{-1}{\beta_{13}} - 2)\log((1-x_1)^{-\beta_{13}} + (1-x_3)^{-\beta_{13}} - 1) \\
&+ \log(\beta_{23|1}+1) + (-\beta_{12}-1)(log((1-x_1)^{-\beta_{12}-1}((1-x_1)^{-\beta_{12}} \\
&+ (1-x_2)^{-\beta_{12}})^{\frac{-1}{\beta_{12}}-1}) + log((1-x_1)^{-\beta_{13}-1}((1-x_1)^{-\beta_{13}} + (1-x_3)^{-\beta_{13}})^{\frac{-1}{\beta_{13}}-1})) \\
&+ (\frac{-1}{\beta_{23|1}} - 2)\log(((1-x_1)^{-\beta_{12}-1}((1-x_1)^{-\beta_{12}} + (1-x_2)^{-\beta_{12}} - 1)^{\frac{-1}{\beta_{12}}-1})^{-\beta_{23|1}} \\
&+ ((1-x_1)^{-\beta_{13}-1}((1-x_1)^{-\beta_{13}} + (1-x_3)^{-\beta_{13}})^{\frac{-1}{\beta_{13}}-1})^{-\beta_{23|1}} - 1).
\end{aligned} \quad (3.2.31)$$

Under the finite mixture model, the density function for 2'nd component is derived by using Equation 3.2.31 and maximized by EM algorithm similar to the 1'st component.

These log-likelihood constructions for each components can be derived by using all other copula families in a similar manner easily.

3.2.1.2 2 Component 4 dimensional C-vine Mixtures

When the dimension is increased from 3 to 4, totally 3 unconditional copulas, 2 conditional copulas with only one conditioning variable and 1 conditional copula with two conditioning variables are required to write the multivariate density function. For the sake of simplicity, 4 dimensional C-vine density is discussed for only Gumbel family for the 1'st component of the mixture model.

In addition to the calculation of conditional copulas with only one conditioning variable, additional 1 conditional copula with two conditioning variable is required in 4 dimensional case. To illustrate, by recalling the 3 dimensional case,

$$\phi_{CV_1} = (\beta^1_{12}, \beta^1_{13}, \beta^1_{14}, \beta^1_{23|1}, \beta^1_{24|1}, \beta^1_{34|12})$$

includes totally 6 pair copula parameters in 4 dimensional C-vine construction. The written superscript value denotes that parameters belong to the 1'st component. Assuming that, all pairs are Gumbel, the joint density calculation has required to obtain 6 pair copula parameters. Let $f_1(x_{1:4}; \phi_{CV_1})$ be the joint density modeled by C-vine with Gumbel pairs, one can write the following 4 dimensional density function as,

$$\log f_1 = \log c_{12} + \log c_{13} + \log c_{14} + \log c_{23|1} + \log c_{24|1} + \log c_{34|12}. \qquad (3.2.32)$$

with the corresponding parameter space ϕ_{CV_1}. Similarly, all those bivariate copula density function in terms of Gumbel family can be calculated as second order partial derivative of $C(x_i, x_j; \theta)$ in general. Here, calculations of all c_{12}, c_{13} and c_{14} are presented below,

$$c_{12} = \frac{C(x_1, x_2; \beta_{12})}{x_1 x_2} [\log x_1 \log x_2]^{\beta^1_{12}-1}$$
$$G_{12}(x_1, x_2; \beta^1_{12})[G_{12}(x_1, x_2; \beta^1_{12}) + \frac{\beta^1_{12} - 1}{(-\log x_1)^{\beta^1_{12}} + (-\log x_2)^{\beta^1_{12}}}]. \qquad (3.2.33)$$

$$c_{13} = \frac{C(x_1, x_3; \beta_{13})}{x_1 x_3} [\log x_1 \log x_3]^{\beta^1_{13}-1}$$
$$G_{13}(x_1, x_3; \beta^1_{13})[G_{13}(x_1, x_3; \beta^1_{13}) + \frac{\beta^1_{13} - 1}{(-\log x_1)^{\beta^1_{13}} + (-\log x_3)^{\beta^1_{13}}}]. \quad (3.2.34)$$

$$c_{14} = \frac{C(x_1, x_4; \beta_{14})}{x_1 x_4} [\log x_1 \log x_4]^{\beta^1_{14}-1}$$
$$G_{14}(x_1, x_4; \beta^1_{14})[G_{14}(x_1, x_4; \beta^1_{14}) + \frac{\beta^1_{14} - 1}{(-\log x_1)^{\beta^1_{14}} + (-\log x_4)^{\beta^1_{14}}}]. \quad (3.2.35)$$

where

$$G_{12}(x_1, x_2; \beta^1_{12}) = [(-\log x_1)^{\beta^1_{12}} + (-\log x_2)^{\beta^1_{12}}]^{\frac{1}{\beta^1_{12}}-1}. \quad (3.2.36)$$

$$G_{13}(x_1, x_3; \beta^1_{13}) = [(-\log x_1)^{\beta^1_{13}} + (-\log x_3)^{\beta^1_{13}}]^{\frac{1}{\beta^1_{13}}-1}. \quad (3.2.37)$$

$$G_{14}(x_1, x_4; \beta^1_{14}) = [(-\log x_1)^{\beta^1_{14}} + (-\log x_4)^{\beta^1_{14}}]^{\frac{1}{\beta^1_{14}}-1}. \quad (3.2.38)$$

To obtain the unconditional density functions, described expressions in Equations 3.2.36, 3.2.37 and 3.2.38 are used in Equations 3.2.33, 3.2.34 and 3.2.35, respectively and then their logarithms are evaluated to find the first three terms in Equation 3.2.32.

Afterwords, for the next tree, $\log c_{23|1}$ and $\log c_{24|1}$ are required. $\log c_{24|1}$ can be obtained similar to the $\log c_{23|1}$ by changing the role of x_3 by x_4 in the conditional density calculation. In this construction, $c_{23|1}$ copula density incorporates x_{21} and x_{31} with parameters β^1_{12}, β^1_{13} and $\beta^1_{23|1}$. Similarly, $c_{24|1}$ copula function incorporates the input values defined as x_{21} and x_{41} and includes the parameters β^1_{12}, β^1_{14} and $\beta^1_{24|1}$.

For the computation of $c_{23|1}$ and $c_{24|1}$, conditional CDF's are required to compute with the help of h functions [1]. For this purpose, $F(x_2|x_1)$, $F(x_3|x_1)$ and $F(x_4|x_1)$ are calculated as inputs of the conditional bivariate densities of $c_{23|1}$ and $c_{24|1}$ below,

$$x_{21} = F(x_2|x_1) = h_{21}(x_2|x_1) = \frac{\partial C(x_1, x_2; \beta_{12})}{\partial x_1}$$
$$= C(x_1, x_2; \beta^1_{12}) \frac{(-\log x_1)^{\beta^1_{12}-1}}{x_1} G_{12}(x_1, x_2; \beta^1_{12})$$
$$= exp(-[(-\log x_1)^{\beta^1_{12}} + (-\log x_2)^{\beta^1_{12}}]^{\frac{1}{\beta^1_{12}}}) \frac{(-\log x_1)^{\beta^1_{12}-1}}{x_1} G_{12}(x_1, x_2; \beta^1_{12}). \tag{3.2.39}$$

$$x_{31} = F(x_3|x_1) = h_{31}(x_3|x_1) = \frac{\partial C(x_1, x_3; \beta_{13})}{\partial x_1}$$
$$= C(x_1, x_3; \beta^1_{13}) \frac{(-\log x_1)^{\beta^1_{13}-1}}{x_1} G_{13}(x_1, x_3; \beta^1_{13})$$
$$= exp(-[(-\log x_1)^{\beta^1_{13}} + (-\log x_3)^{\beta^1_{12}}]^{\frac{1}{\beta^1_{13}}}) \frac{(-\log x_1)^{\beta^1_{13}-1}}{x_1} G_{13}(x_1, x_3; \beta^1_{13}). \tag{3.2.40}$$

$$x_{41} = F(x_4|x_1) = h_{41}(x_4|x_1) = \frac{\partial C(x_1, x_4; \beta_{14})}{\partial x_1}$$
$$= C(x_1, x_4; \beta^1_{14}) \frac{(-\log x_1)^{\beta^1_{14}-1}}{x_1} G_{14}(x_1, x_4; \beta^1_{14})$$
$$= exp(-[(-\log x_1)^{\beta^1_{14}} + (-\log x_4)^{\beta^1_{14}}]^{\frac{1}{\beta^1_{14}}}) \frac{(-\log x_1)^{\beta^1_{14}-1}}{x_1} G_{14}(x_1, x_4; \beta^1_{14}). \tag{3.2.41}$$

Above described input values given in Equations 3.2.39, 3.2.40 and 3.2.41 are replaced for finding both $\log c_{23|1}$ and $\log c_{24|1}$ function, formulated below,

$$\log c_{23|1} = (-((-\log(x_{21}))^{\beta^1_{23|1}} + (-\log(x_{31}))^{\beta^1_{23|1}})^{\beta^1_{23|1}})$$
$$- \log(x_{21}x_{31}) + (-2 + \frac{2}{\beta^1_{23|1}}) \log(((-\log(x_{21}))^{\beta^1_{23|1}} + (-\log(x_{31}))^{\beta^1_{23|1}}))$$
$$+(\beta^1_{23|1} - 1) \log(\log(x_{21}) \log(x_{31}))$$
$$+ \log(1 + (\beta^1_{23|1} - 1)(((-\log(x_{21}))^{\beta^1_{23|1}} + (-\log(x_{31}))^{\beta^1_{23|1}}))^{\frac{-1}{\beta^1_{23|1}}}) \tag{3.2.42}$$

$$\log c_{24|1} = (-((-\log(x_{21}))^{\beta^1_{24|1}} + (-\log(x_{41}))^{\beta^1_{24|1}})^{\beta^1_{24|1}})$$
$$-\log(x_{21}x_{41}) + (-2 + \frac{2}{\beta^1_{24|1}})\log(((-\log(x_{21}))^{\beta^1_{24|1}} + (-\log(x_{41}))^{\beta^1_{24|1}}))$$
$$+(\beta^1_{24|1} - 1)\log(\log(x_{21})\log(x_{41}))$$
$$+\log(1 + (\beta^1_{24|1} - 1)(((-\log(x_{21}))^{\beta^1_{24|1}} + (-\log(x_{41}))^{\beta^1_{24|1}})^{\frac{-1}{\beta^1_{24|1}}})$$
(3.2.43)

where x_{21}, x_{31} and x_{41} are defined in Equations 3.2.39, 3.2.40 and 3.2.41 previously. In the above setup, the values of functions G_{12}, G_{13} and G_{14} are required either to calculate both $\log c_{23|1}$ and $\log c_{24|1}$. Finally, the most challenging part of the density Equation of 3.2.32 is the last term. Here, $\log c_{34|12}$ requires two conditioning variable and h-functions must be used in a different manner. As an input of $c_{34|12}$ function, $F(x_3|x_1, x_2)$ and $F(x_4|x_1, x_2)$ must be derived using h-functions recursively. Based on the general formula given below,

$$F(x_j|x_1, x_2, \cdots, x_{j-1}) = \frac{\partial C_{j,j-1|1,\cdots,j-2}(F(x_j|x_1, \cdots, x_{j-2}), F(x_{j-1}|x_1, \cdots, x_{j-2}))}{\partial F(x_{j-1}|x_1, \cdots, x_{j-2})}.$$
(3.2.44)

For $j = 3$, one can derive the following conditional CDF $F(x_3|x_1, x_2)$ with two conditioning variables x_1, x_2.

$$F(x_3|x_1, x_2) = \frac{\partial C_{2,3|1}(F(x_3|x_1), F(x_2|x_1))}{\partial F(x_2|x_1)} = \frac{\partial C_{2,3|1}}{\partial x_{21}}.$$
(3.2.45)

where $x_{21} = F(x_2|x_1)$ from the definition of h-functions. So, one need to compute the following partial derivative for $F(x_3|x_1, x_2)$ with Gumbel family, represented below,

$$F(x_3|x_1, x_2) = \frac{\partial(exp(-[(-\log x_{21})^{\beta_{23|1}} + (-\log x_{31})^{\beta_{23|1}}]^{\frac{1}{\beta_{23|1}}}))}{\partial x_{21}}$$
$$= (exp(-[(-\log x_{21})^{\beta_{23|1}} + (-\log x_{31})^{\beta_{23|1}}]^{\frac{1}{\beta_{23|1}}}))$$
$$(\frac{-1}{\beta_{23|1}})[(-\log x_{21})^{\beta_{23|1}} + (-\log x_{31})^{\beta_{23|1}}]^{\frac{1}{\beta_{23|1}}-1}(\beta_{23|1})(-\log x_{21})^{\beta_{23|1}-1}(\frac{-1}{x_{21}}).$$
(3.2.46)

Similarly, $F(x_4|x_1, x_2)$ can be obtained using the similar conditional density function. The only difference is the replacement of x_{31} and x_{41} and its derivation given as,

$$\begin{aligned} F(x_4|x_1,x_2) &= \frac{\partial(exp(-[(-\log x_{21})^{\beta_{24|1}} + (-\log x_{41})^{\beta_{24|1}}]^{\frac{1}{\beta_{24|1}}}))}{\partial x_{21}} \\ &= (exp(-[(-\log x_{21})^{\beta_{24|1}} + (-\log x_{41})^{\beta_{24|1}}]^{\frac{1}{\beta_{24|1}}})) \\ (\frac{-1}{\beta_{24|1}})[(-\log x_{21})^{\beta_{24|1}} + (-\log x_{41})^{\beta_{24|1}}]^{\frac{1}{\beta_{24|1}}-1}(\beta_{24|1})(-\log x_{21})^{\beta_{24|1}-1}(\frac{-1}{x_{21}}). \end{aligned} \tag{3.2.47}$$

Finally, $c_{34|12}$ density function is derived by using both $F(x_3|x_1, x_2) = x_{3|12}$ and $F(x_4|x_1, x_2) = x_{4|12}$ as input values. Based on this computation, Equation 3.2.32 is completely derived and ready to use for generating 4-dimensional C-vine with all pairs are Gumbel family.

$$\begin{aligned} c_{34|12} = \exp(-[(-\log x_{3|12})^{\beta^1_{34|12}} + (-\log x_{4|12})^{\beta^1_{34|12}}]^{\frac{1}{\beta^1_{34|12}}}) \\ \frac{(-\log x_{3|12})^{\beta^1_{34|12}-1}}{x_{3|12}} \frac{(-\log x_{4|12})^{\beta^1_{34|12}-1}}{x_{4|12}} \\ G_{34|12}(x_{3|12}, x_{4|12}; \beta^1_{34|12}) \\ [G_{34|12}(x_{3|12}, x_{4|12}; \beta^1_{34|12}) + \frac{\beta^1_{34|12} - 1}{(-\log x_{3|12})^{\beta^1_{34|12}} + (-\log x_{4|12})^{\beta^1_{34|12}}}]. \end{aligned} \tag{3.2.48}$$

where

$$G_{34|12}(x_{3|12}, x_{4|12}; \beta^1_{34|12}) = [(-\log x_{3|12})^{\beta^1_{34|12}} + (-\log x_{4|12})^{\beta^1_{34|12}}]^{\frac{1}{\beta^1_{34|12}}-1}. \tag{3.2.49}$$

Now, take the logarithm of the function $c_{34|12}$ presented in Equation 3.2.48 to complete the all terms of the log-likelihood function of C-vine density having Gumbel families for each pair in 4-dimension. For the rest, logarithm of all unconditional c_{12}, c_{13}, c_{14} and conditional $c_{23|1}$, $c_{24|1}$ and $c_{34|12}$ density functions are replaced in Equation (3.2.32).

In this section, only the detailed construction for Clayton- SurClay in 3-dimensional case and Gumbel family in 4-dimensional case is presented. Certainly, similar calculations are available for other considered archimedean families. However, it has been

thought that above derivations are enough to give a general idea for a mixture model construction.

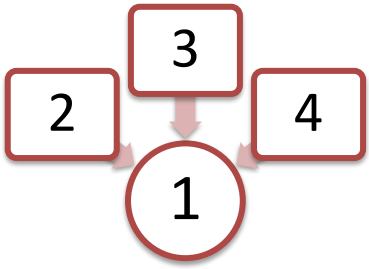
3.2.2 CD-Vine Approach

As a new mixture model, finite C-vine copulas are combined with a D-vine framework. The main motivation behind such a mixture model construction is based on the fact that the dependence pattern of multivariate data can change over time. Equivalently, C-vine mixture model can include different copulas over various time intervals and each C-vine copula model can be attached to each other by using D-vine methodology. Even if the finite mixture model based mixture of C-vine copula density is established with its all parameters in multivariate data (case $p = 4$), the main purpose of the study is to combine these densities using D-vine copula method instead of using finite mixture framework. The main difference for the construction of CD-vine that it requires the parameter knowledge belonging to C-vine part before modeling D-vine for the dependence among the components. To illustrate, the following mixture model is established with required properties. Assume that 3 component C-vine copula densities are generated using same or different pair copulas where $p = 4$ (to differ from D-vine structure, $p \geq 4$ must be satisfied). To illustrate, Figure 3.1 is presented based on 3 component C-vine density with different pair copula families for Components 1, 2 and 3, respectively. Under this setup, $f_1(x_{1:4}; \phi_{CV_1})$, $f_2(x_{1:4}; \phi_{CV_2})$ and $f_3(x_{1:4}; \phi_{CV_3})$ denote the multivariate C-vine density for each component with their corresponding parameter space. For instance; ϕ_{CV_1} includes totally 6 pair copula parameters given in the form of $\phi_{CV_1} = (\beta^1_{12}, \beta^1_{13}, \beta^1_{14}, \beta^1_{23|1}, \beta^1_{24|1}, \beta^1_{34|12})$. In total, for 3 component C-vine density modeling with different pair copulas in each component has required to obtain 18 pair copula parameters.

In this modeling framework, CV_1, CV_2 and CV_3 are multivariate random variables to denote the observed values obtained from each component with their corresponding density functions $f_1(x_{1:4}; \phi_{CV_1})$, $f_2(x_{1:4}; \phi_{CV_2})$ and $f_3(x_{1:4}; \phi_{CV_3})$ respectively. To visualize, the following Figure 3.2 presents the mixture of CD-vine based on observed values coming from C-vines with the help of D-vine copula approach.

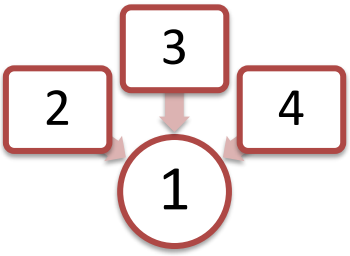
Based on Figure 3.2, DV_{12} and DV_{23} at T_1 are unconditional pair copulas to combine

Component 1



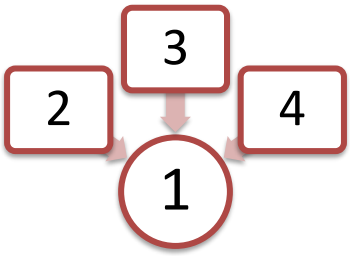
The parameter 1 is the root node in C-vine model, with suitable copula families

Component 2



The parameter 1 is the root node in C-vine model, with suitable copula families

Component 3



The parameter 1 is the root node in C-vine model, with suitable copula families

Figure 3.1: 3-Component C-vine density with different pair copulas

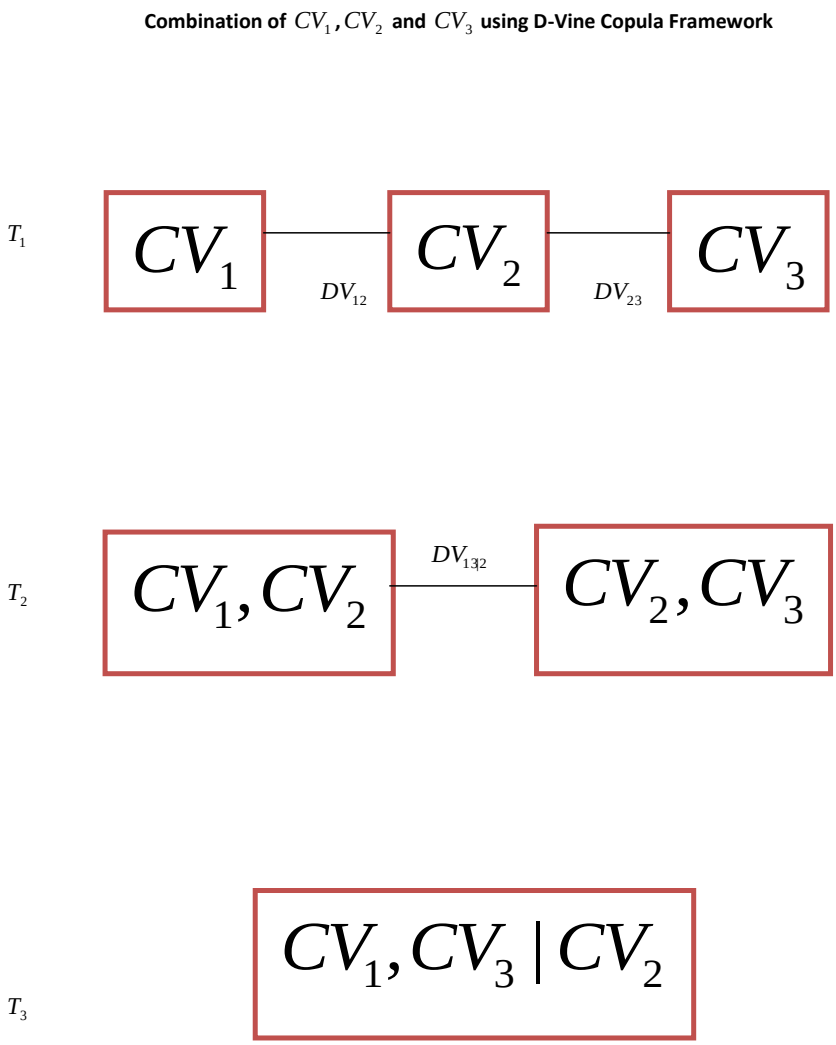


Figure 3.2: CD-Vine Mixture Model Tree Structure

the pairs (CV_1, CV_2) and (CV_2, CV_3) with copula parameters $\theta_{DV_{12}}$ and $\theta_{DV_{23}}$ respectively. For the next tree, T_2, conditional bivariate copula, denoted by $DV_{13|2}$, comes into play to derive the density of the pairs (CV_1, CV_3) given CV_2 based on D-vine copula tree structure. In this mixture, totally there are

$$M \times (p(p-1)/2) + M(M-1)/2 = 18 + 3 = 21$$

parameters required to estimate for expressing the full inference on CD-vine mixture model. One can write the mixture model density for this example $p = 4$ and $M = 3$ as follows,

$$\begin{aligned} g(\boldsymbol{x}; \boldsymbol{\Gamma}) = DV_{12}(F(cv_1), F(cv_2); \theta_{DV_{12}}) DV_{23}(F(cv_2), F(cv_3); \theta_{DV_{23}}) \\ DV_{13|2}(F(cv_1|cv_2), F(cv_3|cv_2); \theta_{DV_{13|2}}) \\ f_1(x_1, x_2, x_3, x_4; \phi_{CV_1}) f_2(x_1, x_2, x_3, x_4; \phi_{CV_2}) \\ f_3(x_1, x_2, x_3, x_4; \phi_{CV_3}). \end{aligned}$$

where cv_1, cv_2 and cv_3 are observations specifically extracted from the multivariate C-vine densities defined via empirical multivariate cumulative distribution functions. $\boldsymbol{\Gamma}$ denotes the all necessary copula parameters to be estimated for maximizing the log-likelihood of function $g(\boldsymbol{x}; \boldsymbol{\Gamma})$ based on given sample, where

$$\boldsymbol{\Gamma} = (\phi_{CV_1}, \phi_{CV_2}, \phi_{CV_3}, \theta_{DV_{12}}, \theta_{DV_{23}}, \theta_{DV_{13|2}}).$$

This CD-vine mixture setup is explained below with Gumbel family in detail.

Consider a simple case including 3 component C-vine copula densities generated by Gumbel copula pairs where $p = 4$. Under this assumption, let $f_1(x_{1:4}; \phi_{CV_1})$, $f_2(x_{1:4}; \phi_{CV_2})$ and $f_3(x_{1:4}; \phi_{CV_3})$ denote the multivariate C-vine densities for each component with their corresponding parameter space ϕ_{CV_1}, ϕ_{CV_2} and ϕ_{CV_3}, which can be expressed with the following general notation.

$$\phi_{CV_i} = (\beta^i_{12}, \beta^i_{13}, \beta^i_{14}, \beta^i_{23|1}, \beta^i_{24|1}, \beta^i_{34|12}). \tag{3.2.50}$$

for $i = 1, 2, 3$ where i is the component number. Gumbel family is useful to identify the upper tail dependencies, with a positive association. Generally, the bivariate copula function $C(x_1, x_2; \theta)$ is given below with the corresponding parameter space,

$\theta \in [1, \infty)$. For the full inference of CD-vine mixture under the proposed setup, Gumbel copula and its probability density function recalled one more time as follows,

$$C(x_1, x_2; \theta) = exp(-[(-\log x_1)^\theta + (-\log x_2)^\theta]^{\frac{1}{\theta}}). \tag{3.2.51}$$

$$c(x_1, x_2; \theta) = \frac{\partial C(x_1, x_2; \theta)}{\partial x_1 \partial x_2}. \tag{3.2.52}$$

After the construction of the pair copulas via Gumbel family for each component, the dependence between the components is modeled by using temporal ordering of the variables. Thus, CD-vine mixture model based on Gumbel copula pairs for each tree are generated. For the full inference on CD-vine mixture model, the following log-likelihood function can be written based on N observations randomly drawn from CD-vine mixture model,

$$\begin{aligned} L(\mathbf{\Gamma}) &= \log(\prod_{n=1}^{N} g(\boldsymbol{x}; \mathbf{\Gamma})) \\ &= \sum_{n=1}^{N} [(\log DV_{12}(F(cv_1), F(cv_2); \theta_{DV_{12}}) + \log DV_{23}(F(cv_2), F(cv_3); \theta_{DV_{23}}) \\ &\quad + \log DV_{13|2}(F(cv_1|cv_2), F(cv_3|cv_2); \theta_{DV_{13|2}})) \\ &\quad + (\log f_1(x_{1:4}; \phi_{CV_1}) + \log f_2(x_{1:4}; \phi_{CV_2}) + \log f_3(x_{1:4}; \phi_{CV_3}))]. \end{aligned} \tag{3.2.53}$$

where the sum of $\log f_i(x_{1:4}; \phi_{CV_i})$ for $i = 1, 2, 3$ as last three red terms represent the contribution to the log-likelihood function from each C-vine component, and the first three blue terms denote the contribution to the function $L(\mathbf{\Gamma})$ from mixture model by considering the interactions of each component with D-vine copula. Since the copula parameters of blue colored terms directly dependent on the parameters required in red colored terms, two step maximization is beneficial to estimate all parameters stated in the related parameter space at the beginning. For this reason, it is important to first maximize the red colored terms of the log-likelihood function to derive $\phi_{CV_1}, \phi_{CV_2}, \phi_{CV_3}$. Afterwords, one can estimate $\theta_{DV_{12}}, \theta_{DV_{23}}, \theta_{DV_{13|2}}$ by maximizing the blue colored terms of $L(\mathbf{\Gamma})$ in Equation 3.2.53.

To illustrate, the first term of red colored terms of the log-likelihood function is summarized below for the first component, as a product of conditional and unconditional Gumbel densities,

$$\begin{aligned}\log f_1(x_{1:4};\phi_{CV_1}) = \log[&c_{12}(x_1,x_2;\beta^1_{12})c_{13}(x_1,x_3;\beta^1_{13})c_{14}(x_1,x_4;\beta^1_{14})\\ &c_{23|1}(F(x_2|x_1),F(x_3|x_1);\beta^1_{23|1})c_{24|1}(F(x_2|x_1),F(x_4|x_1);\beta^1_{24|1})\\ &c_{34|12}(F(x_3|x_1,x_2),F(x_4|x_1,x_2);\beta^1_{34|12})].\end{aligned} \tag{3.2.54}$$

Given $\log f_1(x_{1:4};\phi_{CV_1})$ function in Equation 3.2.54 is simplified and corresponding dependence parameters are suppressed for the rest of CD-vine mixture model, presented below,

$$\log f_1 = [\log c_{12} + \log c_{13} + \log c_{14} + \log c_{23|1} + \log c_{24|1} + \log c_{34|12}]. \tag{3.2.55}$$

In Equation 3.2.55, all log-transformed versions of Gumbel copulas are denoted by $\log c_{12}$, $\log c_{13}$ $\log c_{14}$, $\log c_{23|1}$, $\log c_{24|1}$ and $c_{34|12}$ with their corresponding parameters. These functions are derived previously in Equations 3.2.33, 3.2.34, 3.2.35, 3.2.42, 3.2.43 and 3.2.48. For this reason, the detailed construction of is reviewed only for $\log c_{12}$ below,

$$\begin{aligned}c_{12} &= \frac{\partial^2 C(x_1,x_2;\beta_{12})}{\partial x_1 \partial x_2}\\ &= \frac{\partial}{\partial x_1}[C(x_1,x_2)(\frac{1}{x_2}(-\log x_2)^{\beta^1_{12}-1})[(-\log x_1)^{\beta^1_{12}} + (-\log x_2)^{\beta^1_{12}}]^{\frac{1}{\beta^1_{12}}-1}]\\ &= \frac{1}{x_2}(-\log x_2)^{\beta^1_{12}-1}[\frac{\partial C_{12}}{\partial x_1}G_{12}(x_1,x_2;\beta^1_{12}) + C_{12}\frac{\partial G_{12}(x_1,x_2;\beta^1_{12})}{\partial x_1}]\\ &= \frac{C(x_1,x_2;\beta_{12})}{x_1x_2}[\log x_1 \log x_2]^{\beta^1_{12}-1}G_{12}(x_1,x_2;\beta^1_{12})[G_{12}(x_1,x_2;\beta^1_{12})\\ &+ \frac{\beta^1_{12}-1}{[(-\log x_1)^{\beta^1_{12}} + (-\log x_2)^{\beta^1_{12}}]^{\beta^1_{12}}}]\\ &= \frac{C(x_1,x_2;\beta_{12})}{x_1x_2}[\log x_1 \log x_2]^{\beta^1_{12}-1}G_{12}(x_1,x_2;\beta^1_{12})\\ &[G_{12}(x_1,x_2;\beta^1_{12}) + \frac{\beta^1_{12}-1}{(-\log x_1)^{\beta^1_{12}} + (-\log x_2)^{\beta^1_{12}}}].\end{aligned} \tag{3.2.56}$$

where $G_{12}(x_1, x_2; \beta_{12}^1) = [(-\log x_1)^{\beta_{12}^1} + (-\log x_2)^{\beta_{12}^1}]^{\frac{1}{\beta_{12}^1}-1}$. After taking the natural logarithm of c_{12} density function and arranging all the terms, the following formula has been derived for $\log c_{12}$,

$$\log c_{12} = \log(C(x_1, x_2; \beta_{12})) - \log(x_1 x_2) + (\beta_{12}^1 - 1)\log(\log(x_1)\log(x_2)) + \log(G_{12}) + \log(G_{12} + \frac{\beta_{12}^1 - 1}{[(-\log x_1)^{\beta_{12}^1} + (-\log x_2)^{\beta_{12}^1}]}). \tag{3.2.57}$$

In a similar manner, one can write the formula for $\log c_{13}$ and $\log c_{14}$ using G_{13} and G_{14}, given in Equations 3.2.37 and 3.2.38, as follows,

$$\log c_{13} = \log(C(x_1, x_3; \beta_{13})) - \log(x_1 x_3) + (\beta_{13}^1 - 1)\log(\log(x_1)\log(x_3)) + \log(G_{13}) + \log(G_{13} + \frac{\beta_{13}^1 - 1}{[(-\log x_1)^{\beta_{13}^1} + (-\log x_3)^{\beta_{13}^1}]}). \tag{3.2.58}$$

$$\log c_{14} = \log(C(x_1, x_4; \beta_{14})) - \log(x_1 x_4) + (\beta_{14}^1 - 1)\log(\log(x_1)\log(x_4)) + \log(G_{14}) + \log(G_{14} + \frac{\beta_{14}^1 - 1}{[(-\log x_1)^{\beta_{14}^1} + (-\log x_4)^{\beta_{14}^1}]}). \tag{3.2.59}$$

For the computations of $\log c_{23|1}$ and $\log c_{24|1}$, Equations (3.2.42) and (3.2.43) are already available. Besides, the construction of $\log c_{34|12}$ is obvious after deriving Equation (3.2.48). To sum up, $\log f_1(x_{1:4}; \phi_{CV_1})$ can written as a sum of all above mentioned unconditional and conditional Gumbel copula densities for the first component. Here, new analogue functions, $\widetilde{G}_{12}$, $\widetilde{G}_{13}$, $\widetilde{G}_{14}$, $\widetilde{G}_{23|1}$, $\widetilde{G}_{24|1}$, $\widetilde{G}_{34|12}$, $\widetilde{x}_{21}$, $\widetilde{x}_{31}$, $\widetilde{x}_{41}$, $\widetilde{x}_{3|12}$, $\widetilde{x}_{4|12}$ are generated as they given below, for writing $\log f_1(x_{1:4}; \phi_{CV_1})$ function in a more rigorous way.

$$\widetilde{G}_{12} = \widetilde{G}_{12}(x_1, x_2; \beta_{12}^1) = [(-\log x_1)^{\beta_{12}^1} + (-\log x_2)^{\beta_{12}^1}]. \tag{3.2.60}$$

$$\widetilde{G}_{13} = \widetilde{G}_{13}(x_1, x_3; \beta^1_{13}) = [(-\log x_1)^{\beta^1_{13}} + (-\log x_3)^{\beta^1_{13}}]. \tag{3.2.61}$$

$$\widetilde{G}_{14} = \widetilde{G}_{14}(x_1, x_4; \beta^1_{14}) = [(-\log x_1)^{\beta^1_{14}} + (-\log x_4)^{\beta^1_{14}}]. \tag{3.2.62}$$

$$\widetilde{x}_{21} = \exp\left(-\widetilde{G}_{12}\right)^{\frac{1}{\beta^1_{12}}} \left(\widetilde{G}_{12}\right)^{\frac{1}{\beta_{12}}-1} \frac{(-\log(x_1))^{\beta_{12}-1}}{x_1}. \tag{3.2.63}$$

$$\widetilde{x}_{31} = \exp\left(-\widetilde{G}_{13}\right)^{\frac{1}{\beta^1_{13}}} \left(\widetilde{G}_{13}\right)^{\frac{1}{\beta_{13}}-1} \frac{(-\log(x_1))^{\beta_{13}-1}}{x_1}. \tag{3.2.64}$$

$$\widetilde{x}_{41} = \exp\left(-\widetilde{G}_{14}\right)^{\frac{1}{\beta^1_{14}}} \left(\widetilde{G}_{14}\right)^{\frac{1}{\beta_{14}}-1} \frac{(-\log(x_1))^{\beta_{14}-1}}{x_1}. \tag{3.2.65}$$

where $\widetilde{G}_{12}$, $\widetilde{G}_{13}$ and $\widetilde{G}_{14}$ are defined above as in Equations (3.2.60), (3.2.61) and (3.2.62), respectively.

$$\widetilde{G}_{23|1} = \widetilde{G}_{23|1}(\widetilde{x}_{21}, \widetilde{x}_{31}; \beta^1_{23|1}) = [(-\log \widetilde{x}_{21})^{\beta^1_{23|1}} + (-\log \widetilde{x}_{31})^{\beta^1_{23|1}}]. \tag{3.2.66}$$

$$\widetilde{G}_{24|1} = \widetilde{G}_{24|1}(\widetilde{x}_{21}, \widetilde{x}_{41}; \beta^1_{24|1}) = [(-\log \widetilde{x}_{21})^{\beta^1_{24|1}} + (-\log \widetilde{x}_{41})^{\beta^1_{24|1}}]. \tag{3.2.67}$$

where $\widetilde{x}_{21}$, $\widetilde{x}_{31}$ and $\widetilde{x}_{41}$ are defined above as in Equation 3.2.63, 3.2.64 and 3.2.65, respectively.

$$\widetilde{x}_{3|12} = \exp\left(-\widetilde{G}_{23|1}\right)^{\frac{1}{\beta^1_{23|1}}} \left(\widetilde{G}_{23|1}\right)^{\frac{1}{\beta_{23|1}}-1} \frac{(-\log(\widetilde{x}_{21}))^{\beta_{23|1}-1}}{\widetilde{x}_{21}}. \tag{3.2.68}$$

$$\widetilde{x}_{4|12} = \exp\left(-\widetilde{G}_{24|1}\right)^{\frac{1}{\beta^1_{24|1}}} \left(\widetilde{G}_{24|1}\right)^{\frac{1}{\beta_{24|1}}-1} \frac{(-\log(\widetilde{x}_{21}))^{\beta_{24|1}-1}}{\widetilde{x}_{21}}. \tag{3.2.69}$$

where $\widetilde{G}_{23|1}$ and $\widetilde{G}_{24|1}$ are presented in Equations 3.2.66 and 3.2.67, respectively. Finally,

$$\widetilde{G}_{34|12} = \widetilde{G}_{34|12}(\widetilde{x}_{3|12}, \widetilde{x}_{4|12}; \beta^1_{34|12}) = [(-\log \widetilde{x}_{3|12})^{\beta^1_{34|12}} + (-\log \widetilde{x}_{4|12})^{\beta^1_{34|12}}]. \tag{3.2.70}$$

in which, both $\widetilde{x}_{3|12}$ is described in Equation (3.2.68) and similarly, $\widetilde{x}_{4|12}$ is defined in Equation (3.2.69). After combining above displayed necessary functions given in Equations 3.2.60-3.2.70, $\log f_1(x_{1:4}; \phi_{CV_1})$ function can be written as,

$$\begin{aligned}
\log f_1(x_{1:4}; \phi_{CV_1}) = (-\widetilde{G}_{12})^{\frac{1}{\beta^1_{12}}} - \log(x_1 x_2) + (-2 + \frac{2}{\beta^1_{12}}) \log(\widetilde{G}_{12}) \\
+(\beta^1_{12} - 1) \log(\log(x_1) \log(x_2)) + \log(1 + (\beta^1_{12} - 1)(\widetilde{G}_{12})^{\frac{-1}{\beta^1_{12}}}) \\
(-\widetilde{G}_{13})^{\frac{1}{\beta^1_{13}}} - \log(x_1 x_3) + (-2 + \frac{2}{\beta^1_{13}}) \log(\widetilde{G}_{13}) \\
+(\beta^1_{13} - 1) \log(\log(x_1) \log(x_3)) + \log(1 + (\beta^1_{13} - 1)(\widetilde{G}_{13})^{\frac{-1}{\beta^1_{13}}}) \\
(-\widetilde{G}_{14})^{\frac{1}{\beta^1_{14}}} - \log(x_1 x_4) + (-2 + \frac{2}{\beta^1_{14}}) \log(\widetilde{G}_{14}) \\
+(\beta^1_{14} - 1) \log(\log(x_1) \log(x_4)) + \log(1 + (\beta^1_{14} - 1)(\widetilde{G}_{14})^{\frac{-1}{\beta^1_{14}}}) \\
(-\widetilde{G}_{23|1})^{\frac{1}{\beta^1_{23|1}}} - \log(\widetilde{x}_{21} \widetilde{x}_{31}) + (-2 + \frac{2}{\beta^1_{23|1}}) \log(\widetilde{G}_{23|1}) \\
+(\beta^1_{23|1} - 1) \log(\log(\widetilde{x}_{21}) \log(\widetilde{x}_{31})) + \log(1 + (\beta^1_{23|1} - 1)(\widetilde{G}_{23|1})^{\frac{-1}{\beta^1_{23|1}}}) \\
(-\widetilde{G}_{24|1})^{\frac{1}{\beta^1_{24|1}}} - \log(\widetilde{x}_{21} \widetilde{x}_{41}) + (-2 + \frac{2}{\beta^1_{24|1}}) \log(\widetilde{G}_{24|1}) \\
+(\beta^1_{24|1} - 1) \log(\log(x'_{21}) \log(\widetilde{x}_{41})) + \log(1 + (\beta^1_{24|1} - 1)(\widetilde{G}_{24|1})^{\frac{-1}{\beta^1_{24|1}}}) \\
(-\widetilde{G}_{34|12})^{\frac{1}{\beta^1_{34|12}}} - \log(\widetilde{x}_{3|12} \widetilde{x}_{4|12}) + (-2 + \frac{2}{\beta^1_{34|12}}) \log(\widetilde{G}_{34|12}) \\
+(\beta^1_{34|12} - 1) \log(\log(\widetilde{x}_{3|12}) \log(\widetilde{x}_{4|12})) + \log(1 + (\beta^1_{34|12} - 1)(\widetilde{G}_{34|12})^{\frac{-1}{\beta^1_{34|12}}}).
\end{aligned} \tag{3.2.71}$$

As a second part of the mixture model, it is possible to transform the obtained values cv_1 ,cv_2, cv_3 to pseudo-observations one more time to construct the D-vine part of the CD-vine mixture model. At this step, estimated parameters ϕ_{CV_i} for $i = 1, 2, 3$ are considered for the computation of $\theta_{DV_{12}}$, $\theta_{DV_{23}}$, $\theta_{DV_{13|2}}$ of again Gumbel pairs. Under such a mixed modeling, pseudo-observations defined over cv_1 and cv_2 tied with

Gumbel family. However, cv_2 is combined with cv_3 as unconditional Gumbel copulas at tree T_1 (so it is required to write c_{23}). As a final step, (cv_1, cv_2) and (cv_2, cv_3) are combined together and obtain the conditional Gumbel pair copula $c_{13|2}$ for tree T_2. To maximize the second part of the CD-vine mixture model, following functions are defined to conclude the full inference on parameter estimation,

$$\begin{aligned}\log DV_{12} &= \log C(x_1, x_2; \theta_{DV_{12}}) + [(-\log x_1) + (-\log x_2)] \\ &+ (\theta_{DV_{12}} - 1)(\log(-\log x_1) + \log(-\log x_2)) \\ &+ (\frac{1}{\theta_{DV_{12}}} - 1)\log((-\log x_1)^{\theta_{DV_{12}}} + (-\log x_2)^{\theta_{DV_{12}}}) \\ &+ \log[G_{12_{DV}}(x_1, x_2; \theta_{DV_{12}}) + \frac{\theta_{DV_{12}} - 1}{[(-\log x_1)^{\theta_{DV_{12}}} + (-\log x_2)^{\theta_{DV_{12}}}]}].\end{aligned} \tag{3.2.72}$$

$$\begin{aligned}\log DV_{23} &= \log C(x_2, x_3; \theta_{DV_{23}}) + [(-\log x_2) + (-\log x_3)] \\ &+ (\theta_{DV_{23}} - 1)(\log(-\log x_2) + \log(-\log x_3)) \\ &+ (\frac{1}{\theta_{DV_{23}}} - 1)\log((-\log x_2)^{\theta_{DV_{23}}} + (-\log x_3)^{\theta_{DV_{23}}}) \\ &+ \log[G_{23_{DV}}(x_2, x_3; \theta_{DV_{23}}) + \frac{\theta_{DV_{23}} - 1}{[(-\log x_2)^{\theta_{DV_{23}}} + (-\log x_3)^{\theta_{DV_{23}}}]}].\end{aligned} \tag{3.2.73}$$

where

$$G_{12_{DV}}(x_1, x_2; \theta_{DV_{12}}) = [(-\log x_1)^{\theta_{DV_{12}}} + (-\log x_2)^{\theta_{DV_{12}}}]^{\frac{1}{\theta_{DV_{12}}}-1}. \tag{3.2.74}$$

$$G_{23_{DV}}(x_2, x_3; \theta_{DV_{23}}) = [(-\log x_2)^{\theta_{DV_{23}}} + (-\log x_3)^{\theta_{DV_{23}}}]^{\frac{1}{\theta_{DV_{23}}}-1}. \tag{3.2.75}$$

For the derivation of $DV_{13|2}$, one can rearrange the related terms as it is given for $c_{23|1}$. By using the same idea, $DV_{13|2}$ is constructed by using x_{12} and x_{32} given below,

$$x_{12} = C(x_1, x_2; \theta_{DV_{12}})\frac{(-\log x_2)^{\theta_{DV_{12}}-1}}{x_2}G_{12_{DV}}(x_1, x_2; \theta_{DV_{12}}). \tag{3.2.76}$$

$$x_{32} = C(x_2, x_3; \theta_{DV_{23}})\frac{(-\log x_2)^{\theta_{DV_{12}}-1}}{x_2}G_{23_{DV}}(x_2, x_3; \theta_{DV_{23}}). \tag{3.2.77}$$

where x_1, x_2 and x_3 are equivalently used for second step pseudo-observations defined over cv_1, cv_2 and cv_3; $G_{12_{DV}}(x_1, x_2; \theta_{DV_{12}})$ is the analogue of function $G_{12}(x_1, x_2; \beta_{12}^1)$ defined previously and $G_{23_{DV}}(x_2, x_3; \theta_{DV_{23}})$ is a new function defined in Equation

3.2.75. Using these functions, the closed form of $DV_{13|2}$ can be written as,

$$\begin{aligned} DV_{13|2} = exp(-[(-\log x_{12})^{\theta_{DV_{13|2}}} + (-\log x_{32})^{\theta_{DV_{13|2}}}]^{\frac{1}{\theta_{DV_{13|2}}}}) \\ \frac{(-\log x_{12})^{\theta_{DV_{13|2}}-1}}{x_{12}} \frac{(-\log x_{32})^{\theta_{DV_{13|2}}-1}}{x_{32}} \\ DV^{12;32}(x_{12}, x_{32}; \theta_{DV_{13|2}})[DV^{12;32}(x_{12}, x_{32}; \theta_{DV_{13|2}}) \\ + \frac{(\theta_{DV_{13|2}} - 1)}{[(-\log x_{12})^{\theta_{DV_{13|2}}} + (-\log x_{32})^{\theta_{DV_{13|2}}}]}]. \end{aligned} \tag{3.2.78}$$

where $DV^{12;32}(x_{12}, x_{32}; \theta_{DV_{13|2}}) = [(-\log x_{12})^{\theta_{DV_{13|2}}} + (-\log x_{32})^{\theta_{DV_{13|2}}}]^{\frac{1}{\theta_{DV_{13|2}}}-1}$. More clearly, $\log DV_{13|2}$ can be written as follows,

$$\begin{aligned} \log DV_{13|2} = (-[E_{12}^{DV}(x_1, x_2; \theta_{DV_{12}})]^{\theta_{DV_{13|2}}} + [E_{23}^{DV}(x_2, x_3; \theta_{DV_{23}})]^{\theta_{DV_{13|2}}})^{\frac{1}{\theta_{DV_{13|2}}}} \\ + E_{12}^{DV}(x_1, x_2; \theta_{DV_{12}}) + E_{23}^{DV}(x_2, x_3; \theta_{DV_{23}}) \\ + (\theta_{DV_{13|2}} - 1)[\log[E_{12}^{DV}(x_1, x_2; \theta_{DV_{12}})] + \log[E_{23}^{DV}(x_2, x_3; \theta_{DV_{23}})]] \\ + (\frac{1}{\theta_{DV_{13|2}}} - 1) \log[[E_{12}^{DV}(x_1, x_2; \theta_{DV_{12}})]^{\theta_{DV_{13|2}}} + [E_{23}^{DV}(x_2, x_3; \theta_{DV_{23}})]^{\theta_{DV_{13|2}}}] \\ + \log[[[E_{12}^{DV}(x_1, x_2; \theta_{DV_{12}})]^{\theta_{DV_{13|2}}} + [E_{23}^{DV}(x_2, x_3; \theta_{DV_{23}})]^{\theta_{DV_{13|2}}}]^{\frac{1}{\theta_{DV_{13|2}}}-1} \\ + \frac{\theta_{DV_{13|2}} - 1}{[E_{12}^{DV}(x_1, x_2; \theta_{DV_{12}})]^{\theta_{DV_{13|2}}} + [E_{23}^{DV}(x_2, x_3; \theta_{DV_{23}})]^{\theta_{DV_{13|2}}}}]. \end{aligned} \tag{3.2.79}$$

where

$$\begin{aligned} E_{12}^{DV}(x_1, x_2; \theta_{DV_{12}}) = -\log C_{12} - \log(\frac{1}{x_2}) - (\theta_{DV_{12}} - 1)\log(-\log x_2) \\ - \log G_{12_{DV}}(x_1, x_2; \theta_{DV_{12}}). \end{aligned} \tag{3.2.80}$$

$$\begin{aligned} E_{23}^{DV}(x_2, x_3; \theta_{DV_{23}}) = -\log C_{23} - \log(\frac{1}{x_2}) - (\theta_{DV_{23}} - 1)\log(-\log x_2) \\ - \log G_{23_{DV}}(x_2, x_3; \theta_{DV_{23}}). \end{aligned} \tag{3.2.81}$$

For the maximization of the blue colored terms of the Equation (3.2.53), the sum of Equations (3.2.72), (3.2.73) and (3.2.79) are required. This last optimization problem over the parameters of D-vine part almost complete the inference of parameters for the proposed CD-vine mixture model. This whole sum described above can be summarized

as,

$$
\begin{aligned}
&(\log DV_{12}(F(cv_1), F(cv_2); \theta_{DV_{12}}) + \log DV_{23}(F(cv_2), F(cv_3); \theta_{DV_{23}}) \\
&+\log DV_{13|2}(F(cv_1|cv_2), F(cv_3|cv_2); \theta_{DV_{13|2}})) \\
&= (\theta_{DV_{12}} - 1)\log(-\log x_1) + (-\log x_1) + (\theta_{DV_{23}} - 1)\log(-\log x_3) + (-\log x_3) \\
&+ \log[G_{12_{DV}}(x_1, x_2; \theta_{DV_{12}}) + \frac{\theta_{DV_{12}} - 1}{[(-\log x_1)^{\theta_{DV_{12}}} + (-\log x_2)^{\theta_{DV_{12}}}]}] \\
&+ \log[G_{23_{DV}}(x_2, x_3; \theta_{DV_{23}}) + \frac{\theta_{DV_{23}} - 1}{[(-\log x_2)^{\theta_{DV_{23}}} + (-\log x_3)^{\theta_{DV_{23}}}]}] \\
&(-[E_{12}^{DV}(x_1, x_2; \theta_{DV_{12}})]^{\theta_{DV_{13|2}}} + [E_{23}^{DV}(x_2, x_3; \theta_{DV_{23}})]^{\theta_{DV_{13|2}}})^{\frac{1}{\theta_{DV_{13|2}}}} \\
&+(\theta_{DV_{13|2}} - 1)[\log[E_{12}^{DV}(x_1, x_2; \theta_{DV_{12}})] + \log[E_{23}^{DV}(x_2, x_3; \theta_{DV_{23}})]] \\
&+(\frac{1}{\theta_{DV_{13|2}}} - 1)\log[[E_{12}^{DV}(x_1, x_2; \theta_{DV_{12}})]^{\theta_{DV_{13|2}}} + [E_{23}^{DV}(x_2, x_3; \theta_{DV_{23}})]^{\theta_{DV_{13|2}}}] \\
&+ \log[[[E_{12}^{DV}(x_1, x_2; \theta_{DV_{12}})]^{\theta_{DV_{13|2}}} + [E_{23}^{DV}(x_2, x_3; \theta_{DV_{23}})]^{\theta_{DV_{13|2}}}]^{\frac{1}{\theta_{DV_{13|2}}}-1} \\
&+ \frac{\theta_{DV_{13|2}} - 1}{[E_{12}^{DV}(x_1, x_2; \theta_{DV_{12}})]^{\theta_{DV_{13|2}}} + [E_{23}^{DV}(x_2, x_3; \theta_{DV_{23}})]^{\theta_{DV_{13|2}}}}].
\end{aligned}
\tag{3.2.82}
$$

Finally, above mentioned two-stage maximization procedure for CD-vine model can be summarized with the following steps to estimate all parameters,

1. Construct three C-vine models in 4-dimensional case having sample size N using Gumbel family for each pair, but for instance different parameter values.

2. For the parameter estimation, first put some initials for the related parameters of each C-vine component to maximize the red part of the likelihood function given in 3.2.53.

3. After deriving the estimates for each component, generate C-vine copula data in 3 dimension with estimated parameters and evaluate (EMCDF) of it to derive $N \times 3$ matrix based on approximated parameters

4. Define pseudo-observations over the each column of $N \times 3$ matrix and use them to maximize the blue part of the likelihood function with some predetermined initials for parameters of D-vine copula in 3 dimension.

5. Compare the original and estimated parameters based on different iterations, investigate the best model from driven data and any model identification problem in CD-vine mixture model.

For all above mentioned procedures for both finite mixture of C-vines and CD-vine mixture model is summarized at the end of the study. For both proposed methods, their pseudo codes are presented in Appendix C separately.

3.3 About Parameter Estimation

For all above described mixture models, the most fragile part is the estimation of parameters. The performance and success of the proposed models is highly related to the implementation of suitable optimization algorithm for the maximization of the log-likelihood. With this objective, various derivative free optimization routines are implemented for finding the parameter estimates. For this reason, those different tools are briefly stated in terms of their main features before going further on simulated examples.

3.3.1 Derivative Free Optimization Tools

In this subsection, the considered derivative free algorithms are briefly explained. The reason for such methods is directly depends on the complex nonlinear functions resulted from mixture of vines. Furthermore, it is not so practical to derive analytical derivatives for all mixture densities based on various dependence parameters.

3.3.1.1 Hooke-Jeeves Algorithm

Hooke-Jeeves (hjk) is a kind of direct search method that use only function values. This algorithm creates a set of search directions to identify the search span completely [23]. An iterative process has been used to combine exploratory moves and heuristic pattern moves. More clearly, the detailed steps for this approach is presented below in Algorithm 1. In this following algorithm, x_b is the base point, $f_b = f(x_b)$ value of

the base function, x_C is called stencil center and s_f is the status flag for displaying the failure of exploratory move, namely $ExploratoryMove(x_b, x_C, f, h, s_f)$. Then, in the second part ($PatternMove(x, f, h)$), x is the initial iterate for the best point and f is the objective function and this part completes the exploratory move. In this algorithm, given $PatternMove(x, f, h)$ is the repeated process until the satisfaction of stopping criteria.

Above described hjk algorithm shares the property of its suitable extension to bound constraint problems naturally. In this calculation, both consecutive steps, Exploratory and Pattern moves restrict parameter space to a feasible set. Simply, hjk method with a box constraint is abbreviated as hjkb for the rest of the study. For further properties of hjk algorithm, the well known book by [21] is available for the interested reader.

3.3.1.2 Differential Evolution Algorithm

Differential Evolution (DEoptim) algorithm is a kind of heuristic method to solve optimization problems globally in an effective way [27]. Equivalently, DEoptim implements the idea of differential evolution to optimize a real valued function by using a suitable parameter vector. This is a technique that each generation transforms a set of parameters into another one to find the best set of members more likely to minimize the given objective function [27]. From a given parameter vector, the next one is generated by the scaled difference of two randomly selected parameter vectors within this algorithm.

There are certain crucial parameters for this algorithm, denoted simply by NP, CR and F to express the number of parameter vectors in the population, crossover probability and a positive scale factor, respectively. Here, at the initial generation DEoptim guesses the optimal value from NP and each generation creates a new population from the current members via differential mutation method. The algorithm stops after some set number of generations, or after the objective function value has been reduced below some threshold.

Algorithm 1 hjk algorithm

Require: $ExploratoryMove(x_b, x_C, f, h, s_f)$

$f_b \leftarrow f(x_b)$, $d = 0$ where $x_C \leftarrow x_b + d$ with $x_{cb} = x_b$,

$s_f = 0, f_{cb} \leftarrow f(x_{cb}), x_t = x_C$

for $j = 1$ to N **do**

$p \leftarrow x_t + hv_j$; h is the pattern size and v_j is the j'th column of a direction matrix.

if $f(p) \geq f_b$ **then**

$p \leftarrow x_t - hv_j$

else

$x_t \leftarrow x_{cb} \leftarrow p$

$f_b = f(x_{cb})$

end if

end for

if $x_{cb} \neq x_b$ **then**

$s_f = 1$

$x_b \leftarrow x_{cb}$

end if

Require: $PatternMove(x, f, h)$ $x_b = x, x_C = x, s_f = 1$, x is the initial for best point

Call $ExploratoryMove(x_b, x_C, f, h, s_f)$

while $s_f = 1$ **do**

i) $d \leftarrow x - x_b; x_b = x; x_C = x + d$

ii) Call $ExploratoryMove(x_b, x_C, f, h, s_f)$

if $s_f = 0, x_C = x$ **then**

Call $ExploratoryMove(x_b, x_C, f, h, s_f)$

end if

end while

Require: $hjk(x, f, h_k)$

for $k = 1$ to $\cdots$ **do**

$PatternMove(x, f, h)$

end for

To sum up, DEoptim algorithm directly depends on repeated calculations to bring the population to a global minimum of an objective function. Briefly, DEoptim algorithm is explained below in Algorithm 2. The described consecutive steps for DEoptim algorithm, initialization, mutation, recombination and selection continue until the stopping criterion is reached. For further reading about DEoptim, [31] is suggested for the interested readers.

Algorithm 2 DEoptim algorithm

$x_{i,g}$: current population members, where i and g denote the vectors that make up the population and indexes generation where $i = 1, 2, \cdots, N$.

Define upper and lower bounds for each parameter and consider randomly selected initials uniformly over $[x_j^L, x_j^U]$.

$x_j^L \leq x_{j,i,1} \leq x_j^U$

Then, a trial mutant parameter vector, from three randomly selected members of the population

$v_{i,g+1} \leftarrow x_{r1,g} + F(x_{r2,g} - x_{r3,g})$

In general $F < 1$, it is called mutation factor and $v_{i,g+1}$ is called the donor vector.

Consider a uniform number $rand_{j,i}$, I_{rand} is a random integer from $[1, 2, \cdots, NP]$ where NP is the number of parameters. $v_{i,g+1} \neq x_{i,g}$ is satisfied by I_{rand}.

if $rand_{j,i} \leq CR$ or $j = I_{rand}$ **then**

$u_{j,i,g+1} \leftarrow v_{j,i,g+1}$

else if $rand_{j,i} > CR$ and $j \neq I_{rand}$ **then**

$u_{j,i,g+1} \leftarrow x_{j,i,g}$

end if

where $i = 1, 2, \cdots, N$; $j = 1, 2, \cdots, NP$.

Then, target vector $x_{i,g}$ is compared with trial vector $v_{i,g+1}$, one with the lowest objective function value is attained to the next generation

if $f(u_{i,g+1}) \leq f(x_{i,g})$ **then**

$x_{i,g+1} \leftarrow u_{i,g+1}$

else if $f(u_{i,g+1}) > f(x_{i,g})$ **then**

$x_{i,g+1} \leftarrow x_{i,g}$

end if

3.3.1.3 Spectral Projected Gradient Algorithm

For an unconstrained optimization problem, spectral projected gradient (spg) algorithm is a version of Quasi-Newton secant methods, originally comes from the recursive relationship, defined as,

$$x^{k+1} = x^k - \alpha_k B_k^{-1} \nabla f(x^k). \tag{3.3.1}$$

where $B_{k+1}s_k = y_k$ and $s_{k-1} = x^k - x^{k-1}$, $y_{k-1} = \nabla f(x^k) - \nabla f(x^{k-1})$ are crucial variables for the initial step length in Barzilai-Borwein (BB) iteration algorithm [4]. Namely, in this BB case the unconstrained minimization is defined as $x^{k+1} = x^k + \alpha_k d_k$ where $d_k = -\lambda_k \nabla f(x^k)$ [9]. In this framework, spg method is a spectral method to solve convex constrained optimization problems having the form $x^{k+1} = x^k + \alpha_k d_k$ where d_k is the search direction defined as follows,

$$d^k = P_\Omega(x^k - \lambda_k \nabla f(x^k)) - x^k. \tag{3.3.2}$$

In this Equation 3.3.2, P_Ω is the Euclidean projection onto set Ω and λ_k is the related spectral step length choice described as,

$$\lambda_k = \frac{s_{k-1}^T s_{k-1}}{s_{k-1}^T y_{k-1}}. \tag{3.3.3}$$

After giving brief aspects of the above-mentioned optimization problem with respect to spectral step length, spg algorithm is summarized as follows in Algorithm 3. For further details about the convergence properties and some extensions of spg, [8] is suggested for the interested readers.

Algorithm 3 spg algorithm

Assume $\gamma \in (0,1)$ is the decrease parameter, $M \geq 1$ with safeguarding parameters $0 < \sigma_1 < \sigma_2 < 1$ for quadratic interpolation and $0 < \lambda_{\min} < \lambda_{\max} < \infty$ for the spectral step length. x^0 and $\lambda_0 \in (\lambda_{min}, \lambda_{\max})$ are given,

if $x^0 \notin \Omega$ **then**

Redefine $x_0 = P_\Omega(x_0)$ and set $k \leftarrow 0$

end if

if $||P_\Omega(x^k - \lambda_k \nabla f(x^k)) - x^k||_\infty \leq \epsilon$ **then**

x_k is an approximate stationary point

end if

$d^k \leftarrow P_\Omega(x^k - \lambda_k \nabla f(x^k)) - x^k$ and calculate α_k with parameters γ, M, σ_1 and σ_2.

$x^{k+1} \leftarrow x^k + \alpha_k d_k$

$s_k \leftarrow x^{k+1} - x^k$ and $y_k \leftarrow \nabla f(x^{k+1}) - \nabla f(x^k)$

if $s_k^T y_k \leq 0$ **then**

$\lambda_{k+1} \leftarrow \lambda_{\max}$

else

$\lambda_{k+1} \leftarrow \max(\lambda_{\min}, \min(\frac{s_k^T s_k}{s_k^T y_k}, \lambda_{\max}))$

end if

Set $k \leftarrow k+1$ and turn back.

α_k is called the non-monotone line search and it is calculated based on

$f_{max} \leftarrow \max(f_{x_{k-j}|0 \leq j \leq \min(k, M-1)})$ where M is the integer parameter to impose a functional decrease every M iterations, set $\alpha \leftarrow 1$.

if $f_{x^k + _k} \leq f_{max} + \gamma \alpha \nabla f(x^k)^T d_k$ **then**

$\alpha_k \leftarrow \alpha$

end if

3.3.1.4 Memetic Algorithms with Local Search Chains (Rmalschains) Algorithm

Another global optimization tool used in this study is simply called MA-LS Chains, abbreviated as Rmalschains within Chapter 4 whenever it is used. Briefly, it is a hybridization between Evolutionary Algorithm (EA) and Local Search (LS) by taking advantages of both techniques at the same time [7]. In continuous optimization problems, MA-LS-Chains has proven in terms of their efficiency in literature.

Rmalschains uses steady-state genetic algorithm (SSGA) as EA and allows to improve the same solution many times while creating an LS chain. After finding the initial population, the repeated mechanism is maintained in the population. Whenever SSGA captures a new best solution, then this solution is improved by LS chain. Certainly, the details of Rmalschains algorithm requires various methods like Covariance Matrix Adaptation Evolution Strategy (CMA-ES), which works best. For the sake of simplicity, only general scheme of Rmalschains is described below in Algorithm 4 and for further details [26] is referred.

Algorithm 4 Rmalschains algorithm

Generate $Initial_{population}$

while $\neg termination - condition$ **do**

Perform $SSGA$ with n_{frec} evulations where $SSGA$ is steady-state genetic algorithm and n_{frec} is certain amount of evaluations. Build set S_{LS}, refined by LS chains where S_{LS} is the set with the individuals of the population that have never been improved.

Take best individual c_{LS}

if $c_{LS} \in LS$ **then**

Initialize LS operator with LS state stored with c_{LS}.

else

Initialize LS operator with default LS.

end if

Apply LS algorithm to c_{LS} with I_{str} and give c^r_{LS}.

$c_{LS} \leftarrow c^r_{LS}$

Store final LS state with c^r_{LS}

end while

3.4 About Model Selection

Apart from naive information criteria values, some practical Goodness of Fit (GOF) tests are proposed for the comparison of vine mixtures. For that reason, Clarke and Vuong tests, available for vine copula models, are extended for mixture models. First of all, with the aim of comparing two mixture models to identify whether they are equivalent or significantly different from each other, Clarke test for mixture model is implemented. Thereafter, a new version of Vuong test is studied to conclude that which mixture model is preferable. With the upcoming two subsections, both mixture versions of Clarke and Vuong tests are briefly described.

3.4.1 Clarke test for Mixture of Vines (ClarkeMixV)

As an extension of the test proposed by [10], the new test allows to compare two vine mixture model. Besides, this test enables us to derive information about the mixture model generated via simulated data and constructed based on the estimated parameters of the best model.

Definition 3.4.1. *Let $g_1(\boldsymbol{x};\boldsymbol{\Gamma_1})$ and $g_2(\boldsymbol{x};\boldsymbol{\Gamma_2})$ be two competing vine mixtures with their corresponding estimated parameters $\widehat{\Gamma_1}$ and $\widehat{\Gamma_2}$, respectively. To test the statistical indistinguishability of the two models,*

- $H_0 : P(m_i > 0) = 0.5$
- $H_1 : P(m_i > 0) \neq 0.5$ *for* $i = 1, \cdots, N$

where $m_i = log(\frac{g_1(\boldsymbol{x};\widehat{\Gamma_1})}{g_2(\boldsymbol{x};\widehat{\Gamma_2})})$ for sample size N. Under statistical equivalence of the two models, the following statistics is calculated

$$C_{stats} = \sum_{i=1}^{N} \mathbf{1}_{(0,\infty)}(m_i). \tag{3.4.1}$$

where $\mathbf{1}$ denotes the indicator function, and C_{stats} is distributed Binomial with parameters N and $p = 0.5$ in Equation 3.4.1. Based on the obtained critical values, mixture

model g_1 is statistically equivalent to model g_2 if C_{stats} is not significantly different from $E(C_{stats}) = N \cdot p = \frac{N}{2}$.

The result of new Clarke test enables to make inference on indiscernibility of both base and the fitted models via simulated data.

3.4.2 Vuong test for Mixture of Vines (VuongMixV)

Classical Vuong test can be performed to select the superior model between two vine copula densities [38]. Under this modeling setup, it is redefined for comparing two mixture models in terms of the likelihood-ratio.

Definition 3.4.2. *Let $g_1(\boldsymbol{x}; \mathbf{\Gamma_1})$ and $g_2(\boldsymbol{x}; \mathbf{\Gamma_2})$ be two competing vine mixtures with their corresponding estimated parameters $\widehat{\Gamma_1}$ and $\widehat{\Gamma_2}$, respectively. Besides, $m_i = log(\frac{g_1(\boldsymbol{x};\widehat{\Gamma_1})}{g_2(\boldsymbol{x};\widehat{\Gamma_2})})$ is given for N observations. The standardized sum ν is calculated as follows,*

$$V_{stats} = \nu = \frac{\frac{1}{N} \cdot \sum_{i=1}^{N} m_i}{\sqrt{\sum_{i=1}^{N} (m_i - \overline{m})^2}}. \tag{3.4.2}$$

and it is proven that ν is asymptotically standard normal. Based on the hypothesis test given below,

- $H_0 : E(m_i) = 0$ *for all* $i = 1, \cdots, N$
- $H_1 : E(m_i) \neq 0$ *at least one of* $i = 1, \cdots, N$

one of the mixture models is preferred at a certain level, α, based on the following comparisons.

- *if* $\nu > \phi^{-1}(1 - \frac{\alpha}{2})$*, the mixture model g_1 is preferable to model g_2*
- *if* $\nu < -\phi^{-1}(1 - \frac{\alpha}{2})$*, the mixture model g_2 is preferable to model g_1*
- *if* $|\nu| \leq \phi^{-1}(1 - \frac{\alpha}{2})$ *is satisfied, there is no decision about the selection among both models.*

where ϕ^{-1} is the inverse of the standard normal distribution function in the comparison part above.

For the model comparison part, Vuong test enables to decide which mixture model is the best one among the mentioned mixture models by using simulated data set. Both GOF tests are core tests for the model selection part by supporting or disproving the result of classical information criteria values.

CHAPTER 4

NUMERICAL RESULTS

In this section, the above mentioned multivariate frameworks are studied with a wide range of simulated data sets. For that purpose, numerical results are presented in detail first. Besides, after observing the properties of different mixture models, two different real life applications have been investigated and the recent findings of the study was highlighted by discussing both benefits and limitations of the study.

4.1 Simulation Study

As a simulated data, various mixture models have been considered and the number of observations has been changed from $N = 50$ (small sample) to $N = 1000$ (large sample) to observe the effect of sample size on the parameter estimation part. The estimated parameters for each pair of both components are obtained using mean values of 1000 different replications. Under this data generating process, distinct mixture models have been compared based on the classical information criteria values. Besides, the modified GOF tests, explained previously in the last section of Chapter 3, were implemented as a supplementary tool for model comparison.

For the parameter estimation, it is required to incorporate a suitable optimization routine with meaningful initial values. In that respect, different optimization tools were considered to estimate the parameters of mixture models. Especially, for such mixture of vines, derivative free optimization algorithms have certain advantages with respect to avoided mass calculations on the derivatives of log-likelihood function. Generally, most of the log-likelihood maximization process for various scenarios

were considered with fixed initial parameter guesses from the corresponding parameter space.

With a purpose of understanding the nature of mixture models, various scenarios have been investigated with different tail properties. In the first subsection, 2 component 3 dimensional C-vine mixtures have been considered with the same and different pair copula construction scheme. Thereafter, within the same framework, 3 component 3 dimensional C-vine mixtures are studied. In addition to these simulated data sets, 2 component 4 dimensional C-vine mixtures were considered. Furthermore, finite mixture of C- and D-vine is elaborated with a same copula family for each component. Finally, as a novel contribution, CD-vine mixture approach has been considered. For each scenario mentioned briefly above, model comparison results with the estimated parameter values and certain statistics have been summarized.

4.1.1 Finite Mixture of C-vines

In this part, both 3 and 4 dimensional C-vine mixtures were studied. Here, 2 and 3 component mixtures have been investigated for 3 dimensional case, meanwhile, 2 component mixture case was considered for 4 dimensional multivariate data. The results of the considered mixture scenarios were summarized in detail.

4.1.2 2 Component 3 dimensional C-vine Mixtures

Under this subsection, two different C-vine mixture models have been studied with various dependence properties. Firstly, Frank-Frank case was considered with strong dependence assumption. Thereafter, Clayton and Joe families were selected for the pair copula construction of the 1^{st} and 2^{nd} components, respectively. Within the second scenario, different tail properties are discussed based on the compared models.

4.1.2.1 Frank-Frank case

Suppose the density for each of the components of 3 dimensional C-vine mixture model are relied on Frank and Frank copula families, respectively. Besides, both

C-vine model is constructed with the higher θ values for each pair copula density, which expresses Strong/Strong (SS) dependence case. In this mixture model, only Frank pairs are considered for the densities belonging to 1^{st} and 2^{nd} component, with parameter values $(\beta^1_{12} = 8, \beta^1_{13} = 7, \beta^1_{23|1} = 6)$ and $(\beta^2_{12} = 9, \beta^2_{13} = 6, \beta^2_{23|1} = 5)$.

For the comparison, the focus of the simulation is a mixture of C-vines with assumed properties to check whether the proposed model is able to capture different patterns of dependence structures or not. To illustrate, different mixtures were investigated to compare with the base model and to see the parameter estimation performance over various sample size. Here, all parameter estimations are obtained via derivative free optimization routine, called as $hjkb$ method, with some predefined constraints (Lower and Upper bounds are set to 1 and 9 for the parameters, respectively). For the model selection, Table 4.1 was generated based on the model selection criteria values for various mixtures, where SClay represents for Survival Clayton family for the rest of the study.

Table 4.1: 2 Component 3 dimensional C-vine mixtures with different pairs, strong dependence (Base is Frank-Frank)

Mixt.	**Size**	**Model Selection**		
C1-C2	N	AIC	BIC	CAIC
Frank-Frank	50	**-67.38451**	**-54.00035**	**-47.00035**
	100	**-148.1205**	**-129.8843**	**-122.8843**
	250	**-383.487**	**-358.8367**	**-351.8367**
	500	**-791.2834**	**-761.7811**	**-754.7811**
	1000	**-1591.905**	**-1557.55**	**-1550.55**
Clayton-Clayton	50	-58.77387	-45.38971	-38.38971
	100	-117.8449	-99.60869	-92.60869
	250	-302.5222	-277.872	-270.872
	500	-613.8753	-584.3731	-577.3731
	1000	-1235.783	-1201.429	-1194.429
Clayton-SClay	50	-60.33558	-46.95141	-39.95141
	100	-126.2253	-107.9891	-100.9891
	250	-321.5344	-296.8841	-289.8841
	500	-642.4542	-612.952	-605.952
	1000	-1274.26	-1239.905	-1232.905

Based on the comparison given in Table 4.1, Frank-Frank case for the pairs of the first and second component seems to be the best model as it is expected since the simulated data is coming from the 2 component C-vine mixture with Frank-Frank pairs. Besides, there exists an consistency among different information criteria values when the base model and the fitted one has the same pair copula family. Additionally, corresponding parameter estimates and their useful statistics are given for each mixture model.

In Table 4.2 and 4.3, the most plausible parameters are obtained in the mixture model with Frank-Frank pair copulas. For the whole parameter space, the best approximations belong to the Frank-Frank case. Furthermore, except the parameters β^2_{12} and $\beta^2_{23|1}$, all estimations have very small bias values. Both weight and density parameters have been approximated reasonably for all sample size values. This result briefly exhibits that, the mixture model has captured the dependence structure hidden in the simulated data correctly, based on the 2 component 3-dimensional C-vine mixture.

Table 4.2: Parameter Estimations for 1'st component (Base is Frank-Frank)

		Parameters of the Mixture Model	
		β^1_{12}-β^1_{13}-$\beta^1_{23\|1}$ (8-7-6) ($\pi_1 = 0.5$)	
Mixt.	**Size**	Est.	St.dv.
C1	N		(Bias)
Frank-Frank	50	**7.67-6.11-6.15** 0.51	1.94-2.51-2.76 (-0.33)-(-0.89)-(0.15)
	100	**7.76-6.13-5.93** 0.5	1.96-2.36-2.69 (-0.24)-(-0.87)-(-0.07)
	250	**8-6.42-5.8** 0.48	1.71-2.06-2.53 (0)-(-0.58)-(-0.2)
	500	**7.96-6.68-6.02** 0.49	1.77-1.81-2.33 (-0.04)-(-0.32)-(0.02)
	1000	**8.17-6.39-6.13** 0.49	1.33-1.79-2.19 (0.17)-(-0.61)-(0.13)
Clayton-Clayton	50	2.49-2.28-2.24 0.78	2.37-2.32-2.54 (-5.51)-(-4.72)-(-3.76)
	100	8.34-7.89-8.05 0.08	1.04-1.22-1.78 (0.34)-(0.89)-(2.05)
	250	8.68-8.18-7.7 0.08	0.62-0.9-1.85 (0.68)-(1.18)-(1.7)
	500	8.82-8.39-7.57 0.07	0.39-0.66-1.66 (0.82)-(1.39)-(1.57)
	1000	8.94-8.47-7.83 0.07	0.19-0.54-1.39 (0.94)-(1.47)-(1.83)
Clayton-SClay	50	3.74-3.45-5.63 0.32	2-1.91-3.14 (-4.26)-(-3.55)-(-0.37)
	100	3.12-2.84-4.31 0.36	1.22-1.15-2.83 (-4.88)-(-4.16)-(-1.69)
	250	2.89-2.6-3.4 0.38	0.7-0.6-2.09 (-5.11)-(-4.4)-(-2.6)
	500	2.82-2.53-2.81 0.4	0.45-0.41-1.4 (-5.18)-(-4.47)-(-3.19)
	1000	2.77-2.5-2.46 0.4	0.31-0.29-0.79 (-5.23)-(-4.5)-(-3.54)

Table 4.3: Parameter Estimations for 2'nd component (Base is Frank-Frank)

		Parameters of the Mixture Model	
		β^2_{12}-β^2_{13}-$\beta^2_{23\|1}$ (9-6-5) ($\pi_2 = 0.5$)	
Mixt.	**Size**	Est.	St.dv.
C2	N		(Bias)
Frank-Frank	50	**7.72-6.26-5.66**	2.03-2.42-2.79
		0.49	(-1.28)-(0.26)-(0.66)
	100	**7.78-6.19-5.79**	1.98-2.25-2.62
		0.5	(-1.22)-(0.19)-(0.79)
	250	**7.98-6.13-5.8**	1.67-1.94-2.45
		0.52	(-1.02)-(0.13)-(0.8)
	500	**7.99-5.89-5.92**	1.58-1.85-2.36
		0.51	(-1.01)-(-0.11)-(0.92)
	1000	**7.94-6.18-6**	1.67-1.61-2.28
		0.51	(-1.06)-(0.18)-(1)
Clayton-Clayton	50	7.06-6.88-7.31	2.62-2.75-2.94
		0.22	(-1.94)-(0.88)-(2.31)
	100	1.61-1.11-1.05	0.34-0.18-0.13
		0.92	(-7.39)-(-4.89)-(-3.95)
	250	1.52-1.03-1.01	0.2-0.07-0.04
		0.92	(-7.48)-(-4.97)-(-3.99)
	500	1.48-1.01-1	0.13-0.03-0.01
		0.93	(-7.52)-(-4.99)-(-4)
	1000	1.47-1-1	0.09-0.01-0
		0.93	(-7.53)-(-5)-(-4)
Clayton-SClay	50	2.86-1.75-1.37	1.37-0.95-0.76
		0.68	(-6.14)-(-4.25)-(-3.63)
	100	2.76-1.65-1.24	0.9-0.59-0.38
		0.64	(-6.24)-(-4.35)-(-3.76)
	250	2.69-1.59-1.14	0.58-0.37-0.21
		0.62	(-6.31)-(-4.41)-(-3.86)
	500	2.68-1.6-1.1	0.4-0.26-0.13
		0.6	(-6.32)-(-4.4)-(-3.9)
	1000	2.67-1.57-1.06	0.3-0.18-0.09
		0.6	(-6.33)-(-4.43)-(-3.94)

4.1.2.2 Clayton-Joe case

Within the second scenario, tail dependence structure for each component has been changed. To illustrate, suppose the densities of $1'st$ and the $2'nd$ component of 3 dimensional C-vine mixture are modeled by Clayton and Joe pair copula families, respectively. Assume that both C-vine component is constructed with the higher θ values ($\beta^1_{12} = 6, \beta^1_{13} = 7, \beta^1_{23|1} = 8$) and ($\beta^2_{12} = 5, \beta^2_{13} = 6, \beta^2_{23|1} = 9$) for each pair copula density at the beginning, which expresses again Strong/Strong (SS) dependence. In this mixture model, Clayton pairs exhibit strong left tail dependence and relatively weak right tail dependence. On the other hand, Joe family stands to express strong right tail dependence and relatively weak left tail dependence.

Table 4.4: 2 Component 3 dimensional C-vine mixtures with different pairs, strong dependence (Base is Clayton-Joe)

Mixt.	**Size**	**Model Selection**		
C1-C2	N	AIC	BIC	CAIC
Clayton-Joe	50	-180.0698	-166.6857	-159.6857
	100	-420.7195	-402.4834	-395.4834
	250	-1240.829	-1216.179	-1209.179
	500	-2694.689	-2665.187	-2658.187
	1000	**-5704.382**	**-5670.028**	**-5663.028**
Clayton-Clayton	50	-88.33761	-67.30536	-56.30536
	100	-205.0571	-186.8209	-179.8209
	250	-495.2029	-470.5527	-463.5527
	500	-998.8634	-969.3611	-962.3611
	1000	-1993.513	-1959.159	-1952.159
Clayton-Gumbel	50	**-201.7108**	**-188.3266**	**-181.3266**
	100	**-469.9157**	**-451.6796**	**-444.6796**
	250	**-1304.756**	**-1280.105**	**-1273.105**
	500	**-2726.712**	**-2697.21**	**-2690.21**
	1000	-5492.561	-5458.206	-5451.206
Clayton-SClay	50	-175.4084	-162.0242	-155.0242
	100	-422.963	-404.7268	-397.7268
	250	-1251.251	-1226.601	-1219.601
	500	-2726.414	-2696.912	-2689.912
	1000	-5653.508	-5619.154	-5612.154

Based on all values AIC, BIC and CAIC, there exists a model identification problem for the mixture models Clayton-Joe, Clayton-Gumbel and Clayton-SClay. The weight parameters seem to be unbalanced for each component, even if the starting values are equal ($\pi_1 = \pi_2 = 0.5$). When the sample size was increased, the fitted mixture model Clayton-Joe is the best one as it is expected at the beginning (for $N = 1000$). Besides, even if the model identification problem occurs based on model information criteria values, the parameters of the mixture Clayton-Joe are the most plausible ones, has been investigated with corresponding statistics in Tables 4.5 and 4.6. To sum up, the main reason for this model identification problem was similar dependence structure for Joe, Gumbel and SClay copula families that has been used for the construction of the 2'nd component. More rigorously, all pair copulas represent right tail dependence so that such mixtures have similar dependence patterns.

Parameter estimations for each mixture model was presented in Tables 4.5 and 4.6. In those mixture models, the best parameter estimations belong to the 2'nd component of the mixture model of Clayton-Joe, but the weight parameters are unbalanced for all of them. As a result of the instability for the mixing proportions, the parameter estimations for copulas of the 2'nd component (Joe copula pairs) are more plausible than the 1'st component. Certainly, the increase on the sample size has a positive effect on the reduction for difference between the true parameter and estimated values for the best model.

To visualize the impact of various weights in the mixture model, detailed graphs are generated to express the dependence structure of the mixture model based on Clayton-Joe copula pairs for the 1'st and 2'nd component. In Figures A.1-A.5 numerous fixed weights are considered over the range $[0, 1]$ with step size $h = 0.25$ for each component. For the visualization of different mixtures of Clayton-Joe case, varying weight parameters was investigated with respect to strong dependence scenario for each pair copula.

Table 4.5: Parameter Estimations for 1'st component (Base is Clayton-Joe)

		Parameters of the Mixture Model	
		β^1_{12}-β^1_{13}-$\beta^1_{23\|1}$ (6-7-8) ($\pi_1 = 0.5$)	
Mixt.	**Size**	Est.	St.dv.
C1	N		(Bias)
Clayton-Joe	50	**5.77-5.92-8.74** **0.19**	2.54-2.61-1 (-0.23)-(-1.08)-(0.74)
	100	**6.33-6.48-8.89** **0.13**	2.45-2.51-0.67 (0.33)-(-0.52)-(0.89)
	250	**7.55-7.63-8.98** **0.08**	2.01-2.02-0.28 (1.55)-(0.63)-(0.98)
	500	**8.29-8.36-8.99** **0.05**	1.26-1.28-0.1 2.29-1.36-0.99
	1000	**8.73-8.65-9** **0.03**	0.53-0.56-0.02 (2.73)-(1.65)-(1)
Clayton-Clayton	50	8.6-8.88-8.98 0.31	0.75-0.46-0.19 (2.6)-(1.88)-(0.98)
	100	8.38-8.74-8.81 0.34	1.19-1.04-0.92 (2.38)-(1.74)-(0.81)
	250	8.65-9-9 0.32	0.42-0.03-0 (2.65)-(2)-(1)
	500	8.67-9-9 0.32	0.35-0-0 (2.67)-(2)-(1)
	1000	8.73-9-9 0.32	0.27-0.01-0.01 (2.73-2-1)
Clayton-Gumbel	50	6.16-6.21-8.93 0.1	2.83-2.85-0.51 (0.16)-(-0.79)-(0.93)
	100	6.51-6.58-8.96 0.05	2.79-2.82-0.44 (0.51)-(-0.42)-(0.96)
	250	7.33-7.4-8.99 0.02	2.56-2.57-0.16 (1.33)-(0.4)-(0.99)
	500	8.45-8.46-9 0.01	1.27-1.27-0 (2.45)-(1.46)-(1)
	1000	8.45-8.46-9 0.01	1.27-1.27-0 (2.45)-(1.46)-(1)
Clayton-SClay	50	5.57-5.73-8.77 0.21	2.51-2.59-0.92 (-0.43)-(-1.27)-(0.77)
	100	6.23-6.33-8.88 0.13	2.53-2.57-0.58 (0.23)-(-0.67)-(0.88)
	250	7.69-7.74-8.98 0.07	1.9-1.92-0.25 (1.69)-(0.74)-(0.98)
	500	8.52-8.5-9 0.04	1.07-1.08-0.09 (2.52)-(1.5)-(1)
	1000	8.69-8.77-9 0.04	0.53-0.51-0 (2.69)-(1.77)-(1)

Table 4.6: Parameter Estimations for 2'nd component (Base is Clayton-Joe)

		Parameters of the Mixture Model	
		β^2_{12}-β^2_{13}-$\beta^2_{23\|1}$ (5-6-9) ($\pi_2 = 0.5$)	
Mixt.	**Size**	Est.	St.dv.
C2	N		(Bias)
Clayton-Joe	50	**4.89-5.66-6.45** **0.81**	1.11-1.28-1.48 (-0.11)-(-0.34)-(-2.55)
	100	**4.73-5.55-6.93** **0.87**	0.72-0.84-1.09 (-0.27)-(-0.45)-(-2.07)
	250	**4.75-5.66-7.7** **0.92**	0.42-0.48-0.66 (-0.25)-(-0.34)-(-1.3)
	500	**4.83-5.76-8.09** **0.95**	0.33-0.39-0.5 (-0.17)-(-0.24)-(-0.91)
	1000	**4.91-5.89-8.28** **0.97**	0.24-0.29-0.53 (-0.09)-(-0.11)-(-0.72)
Clayton-Clayton	50	1.45-1.74-3.15 0.69	0.54-0.67-1.74 (-3.55)-(-4.26)-(-5.85)
	100	1.6-1.86-3.34 0.66	1.61-1.59-1.78 (-3.4)-(-4.14)-(-5.66)
	250	1.12-1.36-3.09 0.68	0.18-0.25-0.88 (-3.88)-(-4.64)-(-5.91)
	500	1.05-1.28-3.17 0.68	0.1-0.15-0.65 (-3.95)-(-4.72)-(-5.83)
	1000	1.02-1.25-3.12 0.68	0.07-0.1-0.45 (-3.98)-(-4.75)-(-5.88)
Clayton-Gumbel	50	3.42-3.89-4.02 0.9	0.57-0.67-0.82 (-1.58)-(-2.11)-(-4.98)
	100	3.43-3.91-4.17 0.95	0.36-0.43-0.54 (-1.57)-(-2.09)-(-4.83)
	250	3.5-3.99-4.37 0.98	0.23-0.27-0.34 (-1.5)-(-2.01)-(-4.63)
	500	3.52-4.03-4.48 0.99	0.16-0.18-0.26 (-1.48)-(-1.97)-(-4.52)
	1000	3.52-4.03-4.48 0.99	0.16-0.18-0.26 (-1.48)-(-1.97)-(-4.52)
Clayton-SClay	50	4.17-4.95-5.72 0.79	1.28-1.41-1.66 (-0.83)-(-1.05)-(-3.28)
	100	3.94-4.8-6.19 0.87	0.71-0.85-1.05 (-1.06)-(-1.2)-(-2.81)
	250	4.01-4.93-6.93 0.93	0.42-0.51-0.68 (-0.99)-(-1.07)-(-2.07)
	500	4.08-5.03-7.31 0.96	0.26-0.32-0.53 (-0.92)-(-0.97)-(-1.69)
	1000	4.13-5.11-7.73 0.96	0.22-0.26-0.35 (-0.87)-(-0.89)-(-1.27)

For the model comparison, instead of only looking at information criteria values, modified version of Clarke and Vuong GOF tests are applied. The details of ClarkeMixV and VuongMixV are briefly explained at the end of Chapter 3. In Table 4.7, Clayton-Joe-($M1$), Clayton-Clayton($M2$), Clayton-Gumbel($M3$) and Clayton-SClay($M4$) are compared in terms of ClarkeMixV and VuongMixV GOF tests. In this comparison, $M1$, as a base model, is compared with the other mixtures only.

Table 4.7: GOF Matrix for Model Selection of Clayton-Joe($M1$) with $\alpha = 0.05$

Model Comparison Values	**GOF Matrix**
$N = 50\ E[Cstats] = 25$	
$\phi^{-1}(1-\frac{\alpha}{2}) = 1.96$	$\begin{bmatrix} 0 & \mathbf{112.3945} & \mathbf{46.7813} & \mathbf{65.0243} \\ \mathbf{0.0278} & 0 & \cdots & \cdots \\ \mathbf{-0.0067} & \cdots & 0 & \cdots \\ \mathbf{0.0111} & \cdots & \cdots & 0 \end{bmatrix}$
$N = 250\ E[Cstats] = 125$	
$\phi^{-1}(1-\frac{\alpha}{2}) = 1.96$	$\begin{bmatrix} 0 & \mathbf{456.588} & \mathbf{199.6862} & \mathbf{202.9848} \\ \mathbf{0.0151} & 0 & \cdots & \cdots \\ \mathbf{-0.0088} & \cdots & 0 & \cdots \\ \mathbf{-0.0083} & \cdots & \cdots & 0 \end{bmatrix}$
$N = 1000\ E[Cstats] = 500$	
$\phi^{-1}(1-\frac{\alpha}{2}) = 1.96$	$\begin{bmatrix} 0 & \mathbf{2357.97} & \mathbf{1054.076} & \mathbf{1023.538} \\ \mathbf{0.0022} & 0 & \cdots & \cdots \\ \mathbf{0.0019} & \cdots & 0 & \cdots \\ \mathbf{0.0024} & \cdots & \cdots & 0 \end{bmatrix}$

In the presented GOF matrices in Table 4.7, the first row from left to right includes the Clarke test statistic ($Cstats_{ClarkeMixV}$) for the comparison of $M1/M2$, $M1/M3$ and $M1/M4$. Based on different sample size, when the obtained test statistic is different from $E[Cstats]$, the compared models are significantly different from each other. Meanwhile, the first column of GOF matrix from top to bottom exhibits the Vuong test statistic ($\nu_{VuongMixV}$) value for the same comparisons, namely $M1/M2$, $M1/M3$ and $M1/M4$. Here, these statistics are compared with $\phi^{-1}(1-\frac{\alpha}{2})$ value to decide which model is more preferable. However, based on the Vuong test statistics, $\nu_{M1/M2}$, $\nu_{M1/M3}$ and $\nu_{M1/M4}$ compared with $\phi^{-1}(1-\frac{0.05}{2}) = 1.96$, there is no decision about the selection among the models $M1$ and others at $\%95$ significance level. Equivalently, there is no significant result for supporting or disproving the empirical findings presented in Table 4.4.

4.1.3 3 Component 3 dimensional C-vine Mixtures

Under this subsection, two different C-vine mixture models have been studied with various dependence structure. Firstly, SClay-SClay-SClay case was considered with strong dependence (SS) assumption. Thereafter, Clayton, Frank and SClay families were selected for the pair copula construction of the 1^{st}, 2^{nd} and 3^{rd} components, respectively. Within these scenarios, the main differences of both cases have been highlighted after the model comparison part.

4.1.3.1 SClay-SClay-SClay case

For 3 component C-vine mixture, the first multivariate data is generated based on Survival-Clayton (SClay) copula pairs with equal weights. Within this synthetic data, all copula pairs are modeled with strong dependence assumption. Here, totally 9 copula density parameters are selected as $(\beta^1_{12} = 4, \beta^1_{13} = 5, \beta^1_{23|1} = 6)$, $(\beta^2_{12} = 5, \beta^2_{13} = 6, \beta^2_{23|1} = 8)$ and $(\beta^3_{12} = 4, \beta^3_{13} = 6, \beta^3_{23|1} = 9)$ for each component. Model comparison results for the mixture models with equal weight parameters are presented below in Table 4.8.

Table 4.8 expresses the best model with respect to model information criteria values. Apparently, Clay-Clay-Clay is the poorest one, but there is a model confusion between the base model SClay-SClay-SClay and Clay-SClay-SClay mixture model. In this comparison, model information criteria values have tendency to select less complex mixture model in terms of weight parameters. Similar to 2 component mixture case, the problem comes from the unbalanced weight parameters, tabulated in Table 4.9 in detail. Besides, since Kendall's τ values of a copula family and its survival version is same, there exists a model identification problem based on information criterion values. In Table 4.9, the worst parameter estimations belong to the Clay-Clay-Clay mixture and especially estimated parameters for copula pairs are not plausible enough for the conditional part. However, the increase on the number of components worsens the parameter estimations for each component, apart from the 1'st one. Besides, the results for the parameter estimations highlight, SClay-SClay-SClay results in the best estimations as it is expected.

Table 4.8: 3 Component 3 dimensional C-vine mixtures with different pairs, strong dependence (Base is SClay-SClay-SClay)

Mixt.	**Size**	**Model Selection**		
C1-C2-C3	N	AIC	BIC	CAIC
SClay-SClay-SClay	50	-87.75049	-66.71824	-55.71824
	100	-202.8378	-174.181	-163.181
	250	-543.4935	-504.7574	-493.7574
	500	-1103.723	-1057.362	-1046.362
	1000	-2163.763	-2109.778	-2098.778
Clay-Clay-Clay	50	-74.49535	-53.4631	-42.4631
	100	-128.6865	-100.0297	-89.02966
	250	-332.4111	-293.675	-282.675
	500	-761.267	-714.9063	-703.9063
	1000	-1662.728	-1608.743	-1597.743
Clay-SClay-SClay	50	**-95.98086**	**-74.94861**	**-63.94861**
	100	**-219.144**	**-190.4871**	**-179.4871**
	250	**-643.1008**	**-604.3647**	**-593.3647**
	500	**-1374.556**	**-1328.195**	**-1317.195**
	1000	**-3613.678**	**-3559.693**	**-3548.693**

Table 4.9: Parameter Estimations based on different C-vine mixtures for (Base is SClay-SClay-SClay)

		Parameters of the Mixture Model					
		β^1_{12}-β^1_{13}-$\beta^1_{23\|1}$ (4-5-6) ($\pi_1 = 1/3$)		β^2_{12}-β^2_{13}-$\beta^2_{23\|1}$ (5-6-8) ($\pi_2 = 1/3$)		β^3_{12}-β^3_{13}-$\beta^3_{23\|1}$ (4-6-9) ($\pi_3 = 1/3$)	
Mixt.	**Size**	Est.	St.dv.	Est.	St.dv.	Est.	St.dv.
C1-C2-C3	N		(Bias)		(Bias)		(Bias)
SClay-SClay-SClay	50	**4.06-5.27-6.36** **0.34**	2.3-2.54-3.08 (0.06)-(0.27)-(0.36)	**4.32-5.38-6.75** **0.33**	2.39-2.61-2.91 (-0.68)-(-0.62)-(-1.25)	**4.38-5.38-6.54** **0.32**	2.53-2.58-3.02 (0.38)-(-0.62)-(-2.46)
	100	**3.95-4.97-6.43** **0.34**	2.16-2.35-2.89 (-0.05)-(-0.03)-(0.43)	**3.9-4.72-6.39** **0.33**	2.16-2.27-2.9 (-1.1)-(-1.28)-(-1.61)	**3.87-4.91-6.12** **0.33**	2.12-2.38-3.01 (-0.13)-(-1.09)-(-2.88)
	250	**3.42-4.44-6.08** **0.34**	1.69-2.04-2.74 (-0.58)-(-0.56)-(0.08)	**3.72-4.66-6.47** **0.33**	1.86-2.16-2.75 (-1.28)-(-1.34)-(-1.53)	**3.83-5-6.24** **0.33**	1.95-2.26-2.73 (-0.17)-(-1)-(-2.76)
	500	**3.33-4.39-5.77** **0.32**	1.62-2.05-2.63 (-0.67)-(-0.61)-(-0.23)	**3.45-4.44-6.46** **0.34**	1.54-2.04-2.58 (-1.55)-(-1.56)-(-1.54)	**3.67-4.93-6.38** **0.34**	1.7-2.12-2.64 (-0.33)-(-1.07)-(-2.62)
	1000	**3.17-4.13-5.68** **0.34**	1.46-1.85-2.44 (-0.83)-(-0.87)-(-0.32)	**3.45-4.63-6.43** **0.34**	1.56-2.09-2.36 (-1.55)-(-1.37)-(-1.57)	**3.82-4.88-6.42** **0.32**	1.84-2.23-2.46 (-0.18)-(-1.12)-(-2.58)
Clay-Clay-Clay	50	6.57-7-7.34 0.31	3.19-3.21-2.88 (2.57)-(2)-(1.34)	8.42-8.8-8.98 0.16	1.03-0.8-0.18 (3.42)-(2.8)-(0.98)	3.12-3.57-4.46 0.53	3.21-3.11-3.01 (-0.88)-(-2.43)-(-4.54)
	100	1.17-1.4-2.19 0.71	0.26-0.34-0.93 (-2.83)-(-3.6)-(-3.81)	8.58-8.92-8.99 0.18	0.59-0.28-0.14 (3.58)-(2.92)-(0.99)	8.59-8.92-8.99 0.1	0.59-0.3-0.19 (4.59)-(2.92)-(-0.01)
	250	1.06-1.29-2.13 0.7	0.11-0.18-0.63 (-2.94)-(-3.71)-(-3.87)	8.64-8.99-9 0.18	0.45-0.08-0 (3.64)-(2.99)-(1)	8.64-8.99-9 0.11	0.45-0.08-0 (4.64)-(2.99)-(0)
	500	1.02-1.26-2.17 0.71	0.05-0.12-0.44 (-2.98)-(-3.74)-(-3.83)	8.66-9-9 0.11	0.4-0.04-0 (3.66)-(3)-(1)	8.65-9-9 0.18	0.4-0-0 (4.65)-(3)-(0)
	1000	1-1.25-2.16 0.72	0.02-0.08-0.29 (-3)-(-3.75)-(-3.84)	8.66-8.99-9 0	0.34-0.07-0 (3.66)-(2.99)-(1)	8.55-8.97-9 0.28	0.46-0.23-0 (4.55)-(2.97)-(0)
Clay-SClay-SClay	50	5.37-5.48-8.67 0.19	2.42-2.46-1.18 (1.37)-(0.48)-(2.67)	4.13-5.41-6.44 0.41	2.01-2.33-2.69 (-0.87)-(-0.59)-(-1.56)	4.28-5.6-6.27 0.4	2.19-2.39-2.7 (0.28)-(-0.4)-(-2.73)
	100	5.17-5.22-8.69 0.16	2.35-2.36-1.16 (1.17)-(0.22)-(2.69)	3.85-5.2-6.32 0.42	1.62-2.09-2.43 (-1.15)-(-0.8)-(-1.68)	3.88-5.34-6.39 0.41	1.68-2.14-2.45 (-0.12)-(-0.66)-(-2.61)
	250	5.37-5.46-8.84 0.13	2.14-2.19-0.74 (1.37)-(0.46)-(2.84)	3.68-4.95-6.58 0.39	1.43-1.89-2.06 (-1.32)-(-1.05)-(-1.42)	3.7-5.18-6.66 0.47	1.35-1.84-1.89 (-0.3)-(-0.82)-(-2.34)
	500	5.89-5.97-8.94 0.11	2.08-2.11-0.41 (1.89)-(0.97)-(2.94)	3.77-5.12-6.6 0.45	1.2-1.73-1.76 (-1.23)-(-0.88)-(-1.4)	3.52-4.86-6.58 0.44	1.27-1.74-1.87 (-0.48)-(-1.14)-(-2.42)
	1000	6.87-7.02-8.99 0.08	1.94-2.02-0.11 (2.87)-(2.02)-(2.99)	3.74-4.81-6.96 0.6	1.28-1.54-1.15 (-1.26)-(-1.19)-(-1.04)	3.77-5.21-6.23 0.32	1.25-1.88-1.79 (-0.23)-(-0.79)-(-2.77)

4.1.3.2 Clayton-Frank-SClay case

As a second finite mixture based on 3-dimensional Cvines, the multivariate data is constructed based on distinct tail dependencies. The 1'st, 2'nd and 3'rd component are modeled via Clayton, Frank and SClay copula pairs, respectively. For this mixture, various tail dependencies has been considered with the help of Clayton, Frank and SClay for each component of the mixture model. Within this scheme, two different cases were considered for the multivariate dependence structure. In the first one, strong dependence was assumed (SS case) for each parameter of the whole mixture model. As a second simulation study, the conditional density takes a smaller value for the corresponding parameter, ie. each component has both strong and weak dependence structure (SW case). For both mixtures, the model comparison results and the parameter estimations are presented below.

For SS case, the selected parameters are $(\beta^1_{12} = 6, \beta^1_{13} = 7, \beta^1_{23|1} = 8)$, $(\beta^2_{12} = 5, \beta^2_{13} = 6, \beta^2_{23|1} = 9)$ and $(\beta^2_{12} = 4, \beta^2_{13} = 5, \beta^2_{23|1} = 7)$ for Clayton, Frank and SClay copula pairs in the mixture construction. The obtained mixture model has the following properties, presented in Figure 4.1 in detail. For SW case, the selected parameters are $(\beta^1_{12} = 6, \beta^1_{13} = 7, \beta^1_{23|1} = 2.8)$, $(\beta^2_{12} = 5, \beta^2_{13} = 6, \beta^2_{23|1} = 2.9)$ and $(\beta^2_{12} = 4, \beta^2_{13} = 5, \beta^2_{23|1} = 2.7)$ for Clayton, Frank and SClay copula pairs, respectively. In this model, the conditional copula density exhibits weaker dependence for each component compared to unconditional densities. The simulated data has been investigated under the equal weight scheme and parameters are estimated using $hjkb$ optimization algorithm.

In Table 4.10, the mixture model captured the dependence pattern correctly for $N = 50$. On the other hand, the increase in the sample size resulted in a tendency of selecting less complex models by information criteria values. Besides, the mixing proportion for 2'nd component is larger than the weights of other components so that the included tail dependencies can not be identified, presented in Table 4.13. For this reason, the best parameter estimates belong to the Frank pairs used in the construction of 2'nd component of the mixture model. Furthermore, this result addresses a possible sample size issue for such mixture models.

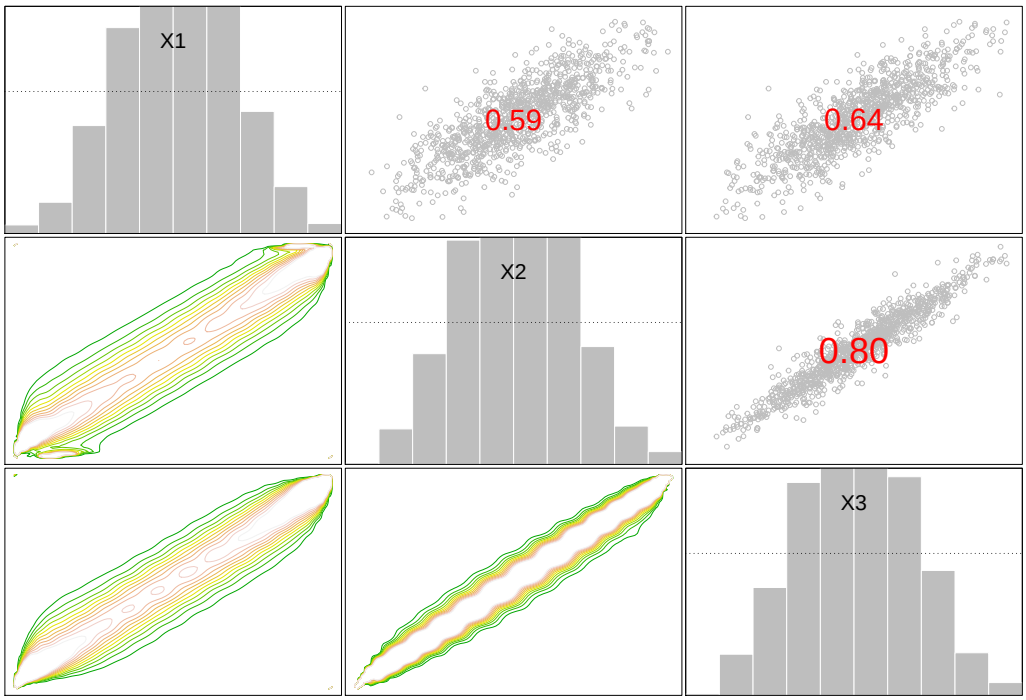


Figure 4.1: Dependence Structure for Clay-Frank-SClay mixture model with equal weights

Table 4.10: 3 Component 3 dimensional C-vine mixtures with different pairs, strong-strong (SS) dependence (Base is Clay-Frank-SClay)

Mixt.	**Size**	**Model Selection**		
C1-C2-C3	N	AIC	BIC	CAIC
Clay-Frank-SClay	50	**-62.51385**	**-41.48159**	**-30.48159**
	100	-129.3742	-100.7173	-89.71729
	250	-365.1885	-326.4525	-315.4525
	500	-809.7136	-763.3529	-752.3529
	1000	-1600.604	-1546.618	-1535.618
SClay-Frank-SClay	50	-61.40312	-40.37086	-29.37086
	100	**-156.3246**	**-127.6677**	**-116.6677**
	250	**-479.842**	**-441.1059**	**-430.1059**
	500	**-1033.387**	**-987.0263**	**-976.0263**
	1000	**-2202.955**	**-2148.97**	**-2137.97**
SClay-SClay-SClay	50	-58.50578	-37.47353	-26.47353
	100	-135.4923	-106.8354	-95.8354
	250	-344.9862	-306.2502	-295.2502
	500	-684.5649	-638.2042	-627.2042
	1000	-1276.612	-1222.626	-1211.626

Table 4.11: Parameter Estimations based on different C-vine mixtures (Base is Clay-Frank-SClay (SS))

		Parameters of the Mixture Model					
		β^1_{12}-β^1_{13}-$\beta^1_{23\|1}$ (6-7-8) ($\pi_1 = 1/3$)		β^2_{12}-β^2_{13}-$\beta^2_{23\|1}$ (5-6-9) ($\pi_2 = 1/3$)		β^3_{12}-β^3_{13}-$\beta^3_{23\|1}$ (4-5-7) ($\pi_3 = 1/3$)	
Mixt.	**Size**	Est.	St.dv.	Est.	St.dv.	Est.	St.dv.
C1-C2-C3	N		(Bias)		(Bias)		(Bias)
Clay-Frank-SClay	50	**4.47-4.48-8.91**	2.68-2.7-0.62	**5.56-6.85-7.8**	1.87-1.9-1.46	**3.87-4.25-6.9**	2.59-2.68-2.8
		0.15	(-1.53)-(-2.52)-(0.91)	0.61	(0.56)-(0.85)-(-1.2)	0.23	(-0.13)-(-0.75)-(-0.1)
	100	**3.55-3.56-8.93**	2.27-2.28-0.51	**5.18-6.5-8.07**	1.2-1.29-1.02	**3.17-3.84-7**	2.47-2.74-2.83
		0.14	(-2.45)-(-3.44)-(0.93)	0.74	(0.18)-(0.5)-(-0.93)	0.12	(-0.83)-(-1.16)-(0)
	250	**2.97-2.97-8.99**	1.81-1.79-0.14	**5.07-6.34-8.31**	0.75-0.84-0.68	**2.99-3.59-6.2**	2.31-2.69-2.99
		0.13	(-3.03)-(-4.03)-(0.99)	0.78	(0.07)-(0.34)-(-0.69)	0.09	(-1.01)-(-1.41)-(-0.8)
	500	**2.89-2.9-9**	1.6-1.64-0.09	**4.94-6.18-8.47**	0.46-0.49-0.51	**5.15-3.01-5.54**	3.28-2.36-3.11
		0.12	(-3.11)-(-4.1)-(1)	0.83	(-0.06)-(0.18)-(-0.53)	0.05	(1.15)-(-1.99)-(-1.46)
	1000	**2.36-2.36-9**	0.61-0.6-0	**4.9-6.12-8.49**	0.31-0.37-0.39	**4.11-3.27-5.68**	2.88-2.52-2.72
		0.13	(-3.64)-(-4.64)-(1)	0.82	(-0.1)-(0.12)-(-0.51)	0.04	(0.11)-(-1.73)-(-1.32)
SClay-Frank-SClay	50	3.29-4.43-8.11	2.61-2.89-2.08	5.6-6.6-8.01	1.62-1.64-1.29	3.65-3.96-7.77	2.81-2.8-2.46
		0.11	(-2.71)-(-2.57)-(0.11)	0.76	(0.6)-(0.6-(-0.99)	0.12	(-0.35)-(-1.04)-(0.77)
	100	3.36-3.57-7.54	2.68-2.56-2.6	5.19-6.27-8.38	1.05-1.09-0.83	3.41-3.75-7.78	2.78-2.7-2.44
		0.08	(-2.64)-(-3.43)-(-0.46)	0.85	(0.19)-(0.27)-(-0.62)	0.07	(-0.59)-(-1.25)-(0.78)
	250	3.06-4.4-7.47	2.43-2.91-2.62	5.08-6.13-8.68	0.54-0.58-0.46	3.9-3.68-7.3	3-2.59-2.76
		0.03	(-2.94)-(-2.6)-(-0.53)	0.92	(0.08)-(0.13)-(-0.32)	0.04	(-0.1)-(-1.32)-(0.3)
	500	4.01-4.88-5.72	3.05-2.89-3.06	5.08-6.07-8.79	0.33-0.38-0.32	4.84-3.88-6.37	3.07-2.73-3.01
		0.02	(-1.99)-(-2.12)-(-2.28)	0.95	(0.08)-(0.07)-(-0.21)	0.02	(0.84)-(-1.12)-(-0.63)
	1000	3.71-3.5-8.3	1.35-1.51-1.97	5.07-6.42-8.85	0.12-0.2-0.09	5.12-8.53-8.61	1.05-1.3-1.16
		0.01	(-2.29)-(-3.5)-(0.3)	0.97	(0.07)-(0.42)-(-0.15)	0.01	(1.12)-(3.53)-(1.61)
SClay-SClay-SClay	50	3.5-3.8-5.81	2.62-2.66-3.31	3.37-3.95-5.48	2.6-2.82-3.4	3.09-4.12-5.5	2.31-2.78-3.29
		0.32	(-2.5)-(-3.2)-(-2.19)	0.36	(-1.63)-(-2.05)-(-3.52)	0.33	(-0.91)-(-0.88)-(-1.5)
	100	3.19-3.63-5.65	2.34-2.5-3.21	2.82-3.56-4.87	2.26-2.66-3.26	2.81-3.57-5.17	2.06-2.59-3.21
		0.33	(-2.81)-(-3.37)-(-2.35)	0.34	(-2.18)-(-2.44)-(-4.13)	0.34	(-1.19)-(-1.43)-(-1.83)
	250	2.7-3.6-4.6	2.02-2.57-3.06	2.37-2.93-4.57	1.86-2.28-3.11	2.77-3.75-4.91	1.87-2.59-2.97
		0.31	(-3.3)-(-3.4)-(-3.4)	0.34	(-2.63)-(-3.07)-(-4.43)	0.35	(-1.23)-(-1.25)-(-2.09)
	500	2.85-3.61-4.97	2.04-2.46-2.93	2.25-2.73-3.42	2.03-2.38-2.68	3.02-3.8-5.11	2.12-2.64-2.87
		0.33	(-3.15)-(-3.39)-(-3.03)	0.32	(-2.75)-(-3.27)-(-5.58)	0.35	(-0.98)-(-1.2)-(-1.89)
	1000	2.85-3.98-4.95	1.77-2.6-2.76	3.33-4.14-5.81	2.18-2.66-2.46	1.41-1.77-2.11	1.12-1.59-1.86
		0.33	(-3.15)-(-3.02)-(-3.05)	0.36	(-1.67)-(-1.86)-(-3.19)	0.31	(-2.59)-(-3.23)-(-4.89)

Table 4.12: 2 Component 3 dimensional C-vine mixtures with different pairs, strong-weak (SW) dependence (Base is Clay-Frank-SurClay)

Mixt.	**Size**	**Model Selection**		
C1-C2-C3	N	AIC	BIC	CAIC
Clay-Frank-SClay	50	-38.00444	-16.97218	-5.972182
	100	-90.06918	-61.41231	-50.41231
	250	-259.976	-221.24	-210.24
	500	-549.7543	-503.3936	-492.3936
	1000	-1182.312	-1128.326	-1117.326
Clay-Clay-Clay	50	**-48.51457**	**-27.48232**	**-16.48232**
	100	**-117.967**	**-89.31018**	**-78.31018**
	250	**-337.8324**	**-299.0963**	**-288.0963**
	500	**-708.4535**	**-662.0928**	**-651.0928**
	1000	**-1469.647**	**-1415.661**	**-1404.661**
SClay-SClay-SClay	50	-34.45827	-13.42602	-2.426019
	100	-83.83367	-55.17679	-44.17679
	250	-170.2798	-131.5438	-120.5438
	500	-335.8053	-289.4446	-278.4446
	1000	-660.6078	-606.6225	-595.6225

Table 4.13: Parameter Estimations based on different C-vine mixtures (Base is Clay-Frank-SClay (SW))

		Parameters of the Mixture Model					
		β^1_{12}-β^1_{13}-$\beta^1_{23\|1}$ (6-7-2.8) ($\pi_1 = 1/3$)		β^2_{12}-β^2_{13}-$\beta^2_{23\|1}$ (5-6-2.9) ($\pi_2 = 1/3$)		β^3_{12}-β^3_{13}-$\beta^3_{23\|1}$ (4-5-2.7) ($\pi_3 = 1/3$)	
Mixt.	**Size**	Est.	St.dv.	Est.	St.dv.	Est.	St.dv.
C1-C2-C3	N		(Bias)		(Bias)		(Bias)
Clay-Frank-SClay	50	**5.27-5.17-8.79**	2.76-2.7-0.94	**5.5-6.41-2.95**	1.75-1.67-1.95	**3.38-4.4-6.45**	2.62-2.74-3.33
		0.07	(-0.73)-(-1.83)-(5.99)	0.76	(0.5)-(0.41)-(0.05)	0.17	(-0.62)-(-0.6)-(3.75)
	100	**4.29-4.32-8.83**	2.53-2.56-0.85	**5.21-6.2-2.58**	1.21-1.21-1.04	**3.5-3.97-6.57**	2.76-2.82-3.25
		0.06	(-1.71)-(-2.68)-(6.03)	0.84	(0.21)-(0.2)-(-0.32)	0.1	(-0.5)-(-1.03)-(3.87)
	250	**3.69-3.71-8.91**	2.39-2.45-0.59	**5.04-6.15-2.58**	0.63-0.76-0.47	**3.84-3.4-5.99**	2.96-2.75-3.42
		0.06	(-2.31)-(-3.29)-(6.11)	0.89	(0.04)-(0.15)-(-0.32)	0.06	(-0.16)-(-1.6)-(3.29)
	500	**4.07-3.67-8.98**	2.62-2.02-0.27	**4.87-5.96-2.62**	0.36-0.55-0.31	**5.28-4.28-4.67**	3.31-3.06-3.64
		0.05	(-1.93)-(-3.33)-(6.18)	0.92	(-0.13)-(-0.04)-(-0.28)	0.03	(1.28)-(-0.72)-(1.97)
	1000	**2.83-2.88-9**	0.99-1.05-0	**4.88-5.97-2.69**	0.23-0.3-0.18	**4.88-5.48-6.15**	3.4-2.75-3.1
		0.05	(-3.17)-(-4.12)-(6.2)	0.94	(-0.12)-(-0.03)-(-0.21)	0.01	(0.88)-(0.48)-(3.45)
Clay-Clay-Clay	50	8.32-8.6-8.26	1.04-0.82-1.64	8.32-8.6-8.28	1.03-0.81-1.6	1.35-1.62-1.26	0.41-0.55-0.46
		0.09	(2.32)-(1.6)-(5.46)	0.13	(3.32)-(2.6)-(5.38)	0.77	(-2.65)-(-3.38)-(-1.44)
	100	8.48-8.84-8.54	0.75-0.42-1.21	8.47-8.85-8.43	0.75-0.42-1.42	1.3-1.56-1.18	0.31-0.4-0.3
		0.19	(2.48)-(1.84)-(5.74)	0.06	(3.47)-(2.85)-(5.53)	0.75	(-2.7)-(-3.44)-(-1.52)
	250	8.63-8.93-8.88	0.51-0.21-0.57	8.62-8.92-8.89	0.52-0.24-0.52	1.24-1.51-1.12	0.21-0.26-0.2
		0.25	(2.63)-(1.93)-(6.08)	0.02	(3.62)-(2.92)-(5.99)	0.73	(-2.76)-(-3.49)-(-1.58)
	500	8.69-8.97-8.95	0.41-0.12-0.29	8.71-8.96-8.96	0.4-0.16-0.23	1.23-1.51-1.07	0.15-0.19-0.12
		0.26	(2.69)-(1.97)-(6.15)	0.01	(3.71)-(2.96)-(6.06)	0.73	(-2.77)-(-3.49)-(-1.63)
	1000	8.8-8.99-8.99	0.28-0.08-0.17	8.8-8.99-8.98	0.3-0.07-0.19	1.22-1.51-1.06	0.11-0.13-0.11
		0.27	(2.8)-(1.99)-(6.19)	0	(3.8)-(2.99)-(6.08)	0.73	(-2.78)-(-3.49)-(-1.64)
SClay-SClay-SClay	50	3.4-4.45-4.69	2.76 2.99 3.65	3.02-3.37-3.55	2.77-2.92-3.48	4.12-4.62-5.13	2.85-2.93-3.65
		0.28	(-2.6)-(-2.55)-(1.89)	0.47	(-1.98)-(-2.63)-(0.65)	0.25	(0.12)-(-0.38)-(2.43)
	100	3.49-4.4-4.11	2.74-2.9-3.5	2.71-3.14-2.87	2.5-2.75-3.06	4.14-4.23-4.12	2.92-2.86-3.47
		0.26	(-2.51)-(-2.6)-(1.31)	0.49	(-2.29)-(-2.86)-(-0.03)	0.25	(0.14)-(-0.77)-(1.42)
	250	2.68-5.66-2.6	2.53-2.8-2.81	1.53-1.78-1.4	1.52-1.99-1.56	5.68-4.56-4.3	2.77-2.61-3.33
		0.19	(-3.32)-(-1.34)-(-0.2)	0.65	(-3.47)-(-4.22)-(-1.5)	0.15	(1.68)-(-0.44)-(1.6)
	500	1.14-6.24-1.29	0.55-2.54-1.23	1.46-1.66-1.26	1.4-1.88-1.22	6.27-5.18-2.91	2-2.28-2.65
		0.19	(-4.86)-(-0.76)-(-1.51)	0.66	(-3.54)-(-4.34)-(-1.64)	0.15	(2.27)-(0.18)-(0.21)
	1000	6.56-6.03-2.03	1.84-1.78-1.29	1.17-1.28-1.06	0.84-1.11-0.47	1.05-7.04-1	0.2-2.25-0.03
		0.13	(0.56)-(-0.97)-(-0.77)	0.72	(-3.83)-(-4.72)-(-1.84)	0.16	(-2.95)-(2.04)-(-1.7)

Similar model identification problem has been identified in the comparison given in Table 4.12. Model information criteria values have tendency to select less complex models with same copula pairs. In Table 4.13, the unbalanced weight problem emerges for the mixture model so that the best parameter estimates belong to the 2'nd component in the base model. Here, Clay-Frank-SClay mixture with SW dependence pattern has larger weights for 2'nd component in any sample size, which results in the fact that plausible parameter estimates exist only for Frank pairs. Certainly, there is a positive relation between the accuracy of the parameters and the sample size. On the other hand, other two mixture models, given in Table 4.13, have very poor parameter estimates.

4.1.4 2 Component 4 dimensional C-vine Mixtures

Under this framework, the dimension for each vine model has been increased so that C-vine and D-vine models are differed. For that purpose, 2 Component, 4 dimensional C-Vine mixtures are constructed with same and distinct copula families. Furthermore, Rmalschains, DEoptim and spg functions were implemented for the parameter estimations in this part. Similar to $hjkb$ function, all considered optimization algorithms do not require the gradient information for the obtained log-likelihood function for each component and allow to search global optimum value. For different mixture models, above mentioned optimization tools are considered separately. In general, the computational time cost is higher for all algorithms when the dimension is increased, because of their nature for searching the best candidate for the parameters. This common limitation reduces both the sample size and the total number of iterations for the simulation under this section. As a first mixture model, Joe-Joe case has been considered for each component under 2 Component 4 dimensional case. Here, the considered parameters are $(\beta^1_{12} = 8, \beta^1_{13} = 7, \beta^1_{14} = 6, \beta^1_{23|1} = 9, \beta^1_{24|1} = 8, \beta^1_{34|12} = 7)$ and $(\beta^2_{12} = 9, \beta^2_{13} = 6, \beta^2_{14} = 5, \beta^2_{23|1} = 9, \beta^2_{24|1} = 8, \beta^2_{34|12} = 7)$ for the 1'st and 2'nd component, respectively. For the conditional part, the same parameters are considered within the construction of mixture models. In the first comparison, the mixture models are investigated only for small sample size data, i.e. $N = 50, 100, 250$ with lesser iterations $max1 = 100$.

Table 4.14: 2 Component 4 dimensional C-vine mixtures, (SS) dependence (Base is Joe-Joe)

Mixt.	**Size**	**Model Selection**		
C1-C2	N	AIC	BIC	CAIC
Joe-Joe	50	-282.0111	-257.1548	-244.1548
	100	-649.3973	-615.5301	-602.5301
	250	-1905.531	-1859.752	-1846.752
Clayton-Clayton	50	112.3849	137.2412	150.2412
	100	-149.0358	-115.1686	-102.1686
	250	-390.0704	-344.2914	-331.2914
Gumbel-Gumbel	50	**-479.515**	**-454.6587**	**-441.6587**
	100	**-1011.443**	**-977.5762**	**-964.5762**
	250	**-2478.381**	**-2432.602**	**-2419.602**

In Table 4.14, there exists Joe and Gumbel family confusion in the mixture model because of same dependence structure. At the same time, Clayton-Clayton mixture is the poorest model when it is compared to others. However, not surprisingly, the parameter estimates given for Joe-Joe case is more plausible within a limited sample size. Even if AIC, BIC and CAIC have tendency to select Gumbel-Gumbel mixture case, the number of components is identified incorrectly.

Table 4.15: 1'st component parameter estimations for C-vine mixtures (Base is Joe-Joe)

		Parameters of the Mixture Model	
		β^1_{12}-β^1_{13}-β^1_{14}-$\beta^1_{23\|1}$-$\beta^1_{24\|1}$-$\beta^1_{34\|12}$ (8-7-6-9-8-7) ($\pi_1 = 0.5$)	
Comp.	**Size**	Est.	St.dv.
C1	N		(Bias)
Joe-Joe	50	**6.61-5.26-5.19-5.4-5.24-5.26** 0.49	1.55-2.38-2.35-1.71-1.94-2.16 (-1.39)-(-1.74)-(-0.81)-(-3.6)-(-2.76)-(-1.74)
	100	**6.89-5.04-4.93-5.45-5.18-4.64** 0.52	1.38-1.97-1.87-1.75-1.64-1.96 (-1.11)-(-1.96)-(-1.07)-(-3.55)-(-2.82)-(-2.36)
	250	**6.78-4.66-4.44-5.84-5.48-3.9** 0.53	1.42-2.11-1.9-1.48-1.61-1.68 (-1.22)-(-2.34)-(-1.56)-(-3.16)-(-2.52)-(-3.1)
Clayton-Clayton	50	3.99-3.73-3.51-3.73-3.68-4.07 0.63	1.59-0.91-0.75-0.9-1.01-1.38 (-4.01)-(-3.27)-(-2.49)-(-5.27)-(-4.32)-(-2.93)
	100	4.03-3.37-3.1-3.46-3.24-3.44 0.76	1.29-0.21-0.12-0.73-0.14-0.33 (-3.97)-(-3.63)-(-2.9)-(-5.54)-(-4.76)-(-3.56)
	250	3.33-3.68-3.21-3.55-3.25-3.48 0.48	0.3-0.45-0.26-0.7-0.35-0.81 (-4.67)-(-3.32)-(-2.79)-(-5.45)-(-4.75)-(-3.52)
Gumbel-Gumbel	50	4.88-4.37-3.91-3.5-3.25-3.05 1	0.99-0.79-0.74-0.92-0.81-0.82 (-3.12)-(-2.63)-(-2.09)-(-5.5)-(-4.75)-(-3.95)
	100	4.66-4.09-3.61-4.35-3.82-3.02 0.97	0.99-0.85-0.74-0.74-0.62-0.35 (-3.34)-(-2.91)-(-2.39)-(-4.65)-(-4.18)-(-3.98)
	250	4.85-4.37-3.92-4.01-3.71-3.05 0.93	0.86-0.76-0.67-1.11-0.91-0.62 (-3.15)-(-2.63)-(-2.08)-(-4.99)-(-4.29)-(-3.95)

In both Tables 4.15-4.16, the estimated parameters of various models have been presented for small sample size. Here, the best estimates belong to Joe-Joe case, but the accuracy is not plausible enough because of the considered optimization tool. Most probably, low accuracy is related to the nature of Rmalschains algorithm since it requires large number of objective function evaluations. Furthermore, model information criteria values have tendency to select 1 component classical vine model with Gumbel family.

In the second mixture model, Frank copula pairs are considered for each component with the same parameters. Here, the main difference is the symmetric dependence behavior in Frank copula, instead of only looking at upper tail dependence pattern

Table 4.16: 2'nd component parameter estimations for C-vine mixtures (Base is Joe-Joe)

		Parameters of the Mixture Model	
		β^2_{12}-β^2_{13}-β^2_{14}-$\beta^2_{23\|1}$-$\beta^2_{24\|1}$-$\beta^2_{34\|12}$ (9-6-5-9-8-7) ($\pi_2 = 0.5$)	
Comp.	**Size**	Est.	St.dv.
C2	N		(Bias)
Joe-Joe	50	**6.62-4.83-4.66-5.56-4.93-4.73** 0.51	1.47-2.13-2.09-1.78-1.62-2.14 (-2.38)-(-1.17)-(-0.34)-(-3.44)-(-3.07)-(-2.27)
	100	**6.73-4.37-4.18-5.73-5.03-4.3** 0.48	1.39-1.98-1.93-1.9-1.62-2.09 (-2.27)-(-1.63)-(-0.82)-(-3.27)-(-2.97)-(-2.7)
	250	**6.74-4.63-4.44-5.63-5.4-4.18** 0.47	1.56-2-1.97-1.47-1.59-1.92 (-2.26)-(-1.37)-(-0.56)-(-3.37)-(-2.6)-(-2.82)
Clayton-Clayton	50	6.54-4.42-3.8-4.67-4.04-4.35 0.37	1.74-1.36-1.06-1.45-1.43-1.6 (-2.46)-(-1.58)-(-1.2)-(-4.33)-(-3.96)-(-2.65)
	100	6.62-4.6-3.55-4.82-3.89-4.08 0.24	1.66-1.21-0.77-1.57-0.97-1.15 (-2.38)-(-1.4)-(-1.45)-(-4.18)-(-4.11)-(-2.92)
	250	6.44-3.2-3.26-3.15-3.29-3.25 0.52	0.94-0.14-0.19-0.23-0.28-0.82 (-2.56)-(-2.8)-(-1.74)-(-5.85)-(-4.71)-(-3.75)
Gumbel-Gumbel	50	4.38-6.52-6.27-5.59-5.99-6.26 0	2.69-1.91-2.08-2.32-2.22-2.47 (-4.62)-(0.52)-(1.27)-(-3.41)-(-2.01)-(-0.74)
	100	3.74-7.29-7.16-4.32-4.11-5.15 0.03	2.54-1.46-1.64-1.82-2.23-2.37 (-5.26)-(1.29)-(2.16)-(-4.68)-(-3.89)-(-1.85)
	250	4.9-6.79-6.61-4.37-4.32-4.2 0.07	3.08-1.59-1.94-2.24-2.18-2.25 (-4.1)-(0.79)-(1.61)-(-4.63)-(-3.68)-(-2.8)

comes from Joe family. Furthermore, for the comparison of various 2 component 4 dimensional mixtures, the results of spg function is summarized below. Similar to the above case, the model comparison is studied only for the parameter estimation belonging to small sample size data, i.e. $N = 50, 100$ with lesser iterations $max1 = 100$.

Table 4.17: 2 Component 4 dimensional C-vine mixtures, (SS) dependence (Base is Frank-Frank)

Mixt.	**Size**	**Model Selection**		
C1-C2	N	AIC	BIC	CAIC
Frank-Frank	50	**-145.2674**	**-120.4111**	**-107.4111**
	100	**-329.6863**	**-295.819**	**-282.819**
Joe-Joe	50	-139.1808	-114.3245	-101.3245
	100	-327.9397	-294.0725	-281.0725
Clayton-Clayton	50	-27.39868	-2.542382	10.45762
	100	-52.82008	-18.95287	-5.952865

In model comparison part, given in Table 4.17, model information criteria values attains minimum for Frank-Frank mixture model. Besides, the parameter estimates for the model of Frank-Frank mixture is more promising when it is compared to Joe-Joe mixture case. For a small sample size data, the base model can be identified by using classical information criteria values when the considered copula families are same for each component.

Table 4.18: 1'st component parameter estimations for C-vine mixtures (Base is Frank-Frank)

		Parameters of the Mixture Model	
		β^1_{12}-β^1_{13}-β^1_{14}-$\beta^1_{23\|1}$-$\beta^1_{24\|1}$-$\beta^1_{34\|12}$ (8-7-6-9-8-7) ($\pi_1 = 0.5$)	
Comp.	**Size**	Est.	St.dv.
C1	N		(Bias)
Frank-Frank	50	**7.62-6.58-5.89-7.69-6.91-7.94** 0.45	2.05-2.13-2.36-2.02-2.39-2.06 (-0.38)-(-0.42)-(-0.11)-(-1.31)-(-1.09)-(0.94)
	100	**7.9-6.28-5.57-7.65-7.3-7.26** 0.46	1.62-2.11-2.22-2.07-1.91-2.02 (-0.1)-(-0.72)-(-0.43)-(-1.35)-(-0.7)-(0.26)
Joe-Joe	50	5.16-4.22-4.18-4.03-3.78-4.88 0.39	2.88-2.93-2.98-2.58-2.55-2.84 (-2.84)-(-2.78)-(-1.82)-(-4.97)-(-4.22)-(-2.12)
	100	3.1-2.97-2.95-3.09-2.85-4.39 0.77	0.99-2.24-2.19-2.04-1.88-2.06 (-4.9)-(-4.03)-(-3.05)-(-5.91)-(-5.15)-(-2.61)
Clayton-Clayton	50	7.02-8.12-7.06-8.99-6.53-2.46 0	0.46-0.71-0.62-0.05-0.58-1.55 (-0.98)-(1.12)-(1.06)-(-0.01)-(-1.47)-(-4.54)
	100	2.09-2.02-2-2.01-2-2 1	0.24-0.08-0-0.04-0-0 (-5.91)-(-4.98)-(-4)-(-6.99)-(-6)-(-5)

For the 1'st component, the parameter estimates for Frank-Frank case is very close to the true values in Table 4.18. Furthermore, weight estimate and each bias value for the estimated parameters is plausible enough for Frank-Frank case. Interpretation of the results of Table 4.19 is very similar to the discussed results above.

For the implemented optimization tools, there is no certain consensus about the selection of best algorithm. For this reason, different functions are considered in different mixture cases. Having the model identification experience in 3 dimensional mixture models in the previous subsection, each component is modeled with the same copula pair. Although, other combinations might be generated for numerous C-vine mixtures. For studying the estimation accuracy, different comparisons are investigated in terms of Frank-Frank case with different sample size and iterations. For instance, the result of spg algorithm for the parameter estimations of Frank-Frank C-vine mixture model is presented below for different N and $max1$ values (Tables 4.20 and 4.21).

Table 4.19: 2'nd component parameter estimations for C-vine mixtures (Base is Frank-Frank)

		Parameters of the Mixture Model	
		β^2_{12}-β^2_{13}-β^2_{14}-$\beta^2_{23\|1}$-$\beta^2_{24\|1}$-$\beta^2_{34\|12}$ (9-6-5-9-8-7) ($\pi_2 = 0.5$)	
Comp.	**Size**	Est.	St.dv.
C2	N		(Bias)
Frank-Frank	50	**8.12-6.15-5.21-7.84-7.3-7.01** 0.55	1.43-1.82-2.03-1.84-2.07-2.33 (-0.88)-(0.15)-(0.21)-(-1.16)-(-0.7)-(0.01)
	100	**8.18-5.94-5.15-8.13-7.51-7.24** 0.54	1.34-1.87-1.84-1.68-1.87-2.22 (-0.82)-(-0.06)-(0.15)-(-0.87)-(-0.49)-(0.24)
Joe-Joe	50	3.32-3.32-3.46-3.29-3.01-4.72 0.61	1.55-2.38-2.59-2.04-2-2.52 (-5.68)-(-2.68)-(-1.54)-(-5.71)-(-4.99)-(-2.28)
	100	2.59-7.11-7.09-4.3-4.56-7.97 0.23	0.99-2.65-2.65-2.93-2.99-2.16 (-6.41)-(1.11)-(2.09)-(-4.7)-(-3.44)-(0.97)
Clayton-Clayton	50	2.42-2.07-2-2.07-2-2 1	0.6-0.18-0.01-0.25-0-0 (-6.58)-(-3.93)-(-3)-(-6.93)-(-6)-(-5)
	100	9-7.54-4.56-4.77-3.11-5.83 0	0-0-0-0-0-0 (0)-(1.54)-(-0.44)-(-4.23)-(-4.89)-(-1.17)

In both Tables 4.20-4.21, there is certain positive impact of both sample size and total number of iterations on the accuracy of parameter estimates. In terms of spg function, the parameter estimates of Frank-Frank C-vine mixture model are reasonable, but, some improvements must be required in any way. Another similar comparison is presented for the results of DEoptim function with a fixed number of iterations, namely $max1 = 100$, below in Table 4.22.

Naturally, there is no best derivative free optimization tool for all problems so that other methods are also applicable for the estimation. Within the subsequent sections, for finite mixture of C- and D-vines and CD-vine mixture methodology, DEoptim function has been used. The main reason for this selection is primarily based on the parameter estimation accuracy presented in Table 4.22 and its algorithm characteristic in terms of finding global optimum values.

Table 4.20: 1'st component parameter estimations for C-vine mixtures with Frank-Frank case

		Parameters of the Mixture Model	
		β^1_{12}-β^1_{13}-β^1_{14}-$\beta^1_{23\|1}$-$\beta^1_{24\|1}$-$\beta^1_{34\|12}$ (8-7-6-9-8-7) ($\pi_1 = 0.5$)	
Comp.	**Size**	Est.	St.dv.
C1	N		(Bias)
$Frank_{250}$	50	**8.12-6.39-5.41-8.06-7.58-7.29** 0.45	2.24-2.61-2.78-2.47-2.64-3.23 (0.12)-(-0.61)-(-0.59)-(-0.94)-(-0.42)-(0.29)
	100	**8.01-6.08-5.09-8.21-7.68-7.42** 0.45	2.11-2.56-2.54-2.09-2.45-2.77 (0.01)-(-0.92)-(-0.91)-(-0.79)-(-0.32)-(0.42)
	250	**8.16-6.69-5.93-8.19-7.52-7.5** 0.51	1.7-2.21-2.06-2.24-2.05-2.3 (0.16)-(-0.31)-(-0.07)-(-0.81)-(-0.48)-(0.5)
	500	**8.63-6.65-5.56-8.32-7.39-7.13** 0.49	1.63-1.9-1.99-1.85-1.99-2.47 (0.63)-(-0.35)-(-0.44)-(-0.68)-(-0.61)-(0.13)
	1000	**8.23-6.4-5.56-8.31-7.48-7.66** 0.50	1.61-1.73-1.7-1.75-1.87-2.14 (0.23)-(-0.6)-(-0.44)-(-0.69)-(-0.52)-(0.66)
$Frank_{500}$	50	**8.27-6.55-5.69-8.07-7.38-7.66** 0.52	1.96-2.43-2.54-2.31-2.63-2.85 (0.27)-(-0.45)-(-0.31)-(-0.93)-(-0.62)-(0.66)
	100	**8.1-6.51-5.66-8.17-7.7-7.6** 0.52	1.94-2.24-2.26-2.26-2.2-2.54 (0.1)-(-0.49)-(-0.34)-(-0.83)-(-0.3)-(0.6)
	250	**8.44-6.68-5.7-8.33-7.64-7.16** 0.49	1.59-1.98-2.05-1.97-2.16-2.46 (0.44)-(-0.32)-(-0.3)-(-0.67)-(-0.36)-(0.16)
	500	**8.3-6.59-5.42-8.35-7.39-7.48** 0.5	1.57-1.83-1.94-1.88-2.02-2.25 (0.3)-(-0.41)-(-0.58)-(-0.65)-(-0.61)-(0.48)
	1000	**8.36-6.29-5.31-8.21-7.5-7.55** 0.48	1.44-1.74-1.63-1.93-1.86-2.17 (0.36)-(-0.71)-(-0.69)-(-0.79)-(-0.5)-(0.55)
$Frank_{1000}$	50	**7.99-6.3-5.5-8.2-7.46-7.98** 0.50	2.18-2.53-2.54-2.35-2.54-2.73 (-0.01)-(-0.7)-(-0.5)-(-0.8)-(-0.54)-(0.98)
	100	**8.29-6.51-5.67-8.2-7.58-7.64** 0.48	1.89-2.35-2.34-2.14-2.32-2.66 (0.29)-(-0.49)-(-0.33)-(-0.8)-(-0.42)-(0.64)
	250	**8.41-6.61-5.65-8.3-7.51-7.22** 0.50	1.62-2.08-2.03-1.9-2.13-2.45 (0.41)-(-0.39)-(-0.35)-(-0.7)-(-0.49)-(0.22)
	500	**8.37-6.48-5.47-8.32-7.55-7.45** 0.48	1.56-1.91-1.85-1.92-1.97-2.27 (0.37)-(-0.52)-(-0.53)-(-0.68)-(-0.45)-(0.45)
	1000	**8.41-6.44-5.43-8.23-7.46-7.42** 0.51	1.4-1.69-1.6-1.76-1.79-2.13 (0.41)-(-0.56)-(-0.57)-(-0.77)-(-0.54)-(0.42)

Table 4.21: 2'nd component parameter estimations for C-vine mixtures with Frank-Frank case

		Parameters of the Mixture Model	
		β^2_{12}-β^2_{13}-β^2_{14}-$\beta^2_{23\|1}$-$\beta^2_{24\|1}$-$\beta^2_{34\|12}$ (9-6-5-9-8-7) ($\pi_2 = 0.5$)	
Comp.	**Size**	Est.	St.dv.
C2	N		(Bias)
$Frank_{250}$	50	**8.21-6.34-5.53-8.43-7.75-7.72** 0.55	1.93-2.21-2.45-2.02-2.34-2.85 (-0.79)-(0.34)-(0.53)-(-0.57)-(-0.25)-(0.72)
	100	**8.36-6.35-5.47-8.09-7.28-7.87** 0.55	1.92-2.13-2.13-2.12-2.33-2.34 (-0.64)-(0.35)-(0.47)-(-0.91)-(-0.72)-(0.87)
	250	**8.34-6.29-5.26-8.22-7.48-7.52** 0.49	1.94-2.11-1.99-2-2.19-2.33 (-0.66)-(0.29)-(0.26)-(-0.78)-(-0.52)-(0.52)
	500	**8.03-6.14-5.41-8.41-7.61-7.68** 0.51	1.76-1.84-1.74-1.75-1.89-2.18 (-0.97)-(0.14)-(0.41)-(-0.59)-(-0.39)-(0.68)
	1000	**8.72-6.81-5.68-8.18-7.36-6.85** 0.50	1.43-1.72-1.67-1.95-2.05-2.06 (-0.28)-(0.81)-(0.68)-(-0.82)-(-0.64)-(-0.15)
$Frank_{500}$	50	**8.04-6.41-5.3-8.15-7.77-7.77** 0.48	2.19-2.51-2.56-2.43-2.34-2.92 (-0.96)-(0.41)-(0.3)-(-0.85)-(-0.23)-(0.77)
	100	**8.29-6.2-5.45-8.08-7.38-7.6** 0.48	1.97-2.29-2.3-2.28-2.42-2.59 (-0.71)-(0.2)-(0.45)-(-0.92)-(-0.62)-(0.6)
	250	**8.25-6.4-5.24-8.34-7.47-7.81** 0.51	1.87-2-1.95-1.85-2.14-2.14 (-0.75)-(0.4)-(0.24)-(-0.66)-(-0.53)-(0.81)
	500	**8.41-6.52-5.53-8.32-7.66-7.41** 0.5	1.67-1.87-1.87-1.92-1.89-2.21 (-0.59)-(0.52)-(0.53)-(-0.68)-(-0.34)-(0.41)
	1000	**8.54-6.73-5.65-8.25-7.38-7.09** 0.52	1.54-1.65-1.52-1.91-1.85-2.17 (-0.46)-(0.73)-(0.65)-(-0.75)-(-0.62)-(0.09)
$Frank_{1000}$	50	**8.21-6.56-5.53-8.15-7.82-7.53** 0.50	2.24-2.45-2.5-2.36-2.37-2.87 (-0.79)-(0.56)-(0.53)-(-0.85)-(-0.18)-(0.53)
	100	**8.31-6.4-5.39-8.33-7.47-7.42** 0.52	1.88-2.19-2.22-2.1-2.24-2.71 (-0.69)-(0.4)-(0.39)-(-0.67)-(-0.53)-(0.42)
	250	**8.27-6.33-5.31-8.32-7.5-7.73** 0.50	1.87-2.03-2.01-1.94-2.14-2.32 (-0.73)-(0.33)-(0.31)-(-0.68)-(-0.5)-(0.73)
	500	**8.33-6.53-5.56-8.28-7.45-7.42** 0.52	1.75-1.84-1.81-1.85-1.91-2.17 (-0.67)-(0.53)-(0.56)-(-0.72)-(-0.55)-(0.42)
	1000	**8.5-6.68-5.66-8.25-7.46-7.17** 0.49	1.57-1.77-1.65-1.86-1.78-2.24 (-0.5)-(0.68)-(0.66)-(-0.75)-(-0.54)-(0.17)

Table 4.22: Parameter estimations of the 2 component 4 dimensional Frank-Frank mixture for $max1 = 100$

		Parameters of the Mixture Model	
		β^1_{12}-β^1_{13}-β^1_{14}-$\beta^1_{23\|1}$-$\beta^1_{24\|1}$-$\beta^1_{34\|12}$ (8-7-6-9-8-7) ($\pi_1 = 0.5$)	
Comp.	**Size**	Est.	St.dv.
C1	N		(Bias)
Frank	50	**8.28-6-5.38-7.63-6.99-7.84** 0.5	2.12-2.29-2.68-2.62-2.78-2.82 (0.28)-(-1)-(-0.62)-(-1.37)-(-1.01)-(0.84)
	100	**8.38-6.9-6.17-7.69-7.25-7.69** 0.5	1.64-2.27-2.19-2.43-2.5-2.59 (0.38)-(-0.1)-(0.17)-(-1.31)-(-0.75)-(0.69)
	250	**8.11-6.42-5.31-8.42-7.69-7.34** 0.55	1.71-2.19-2.02-1.71-1.85-2.14 (0.11)-(-0.58)-(-0.69)-(-0.58)-(-0.31)-(0.34)
		β^2_{12}-β^2_{13}-β^2_{14}-$\beta^2_{23\|1}$-$\beta^2_{24\|1}$-$\beta^2_{34\|12}$ (9-6-5-9-8-7) ($\pi_1 = 0.5$)	
Comp.	**Size**	Est.	St.dv.
C2	N		(Bias)
Frank	50	**8.19-6.81-6.15-8.24-7.31-7.66** 0.5	2.17-2.35-2.61-2.23-2.59-2.92 (-0.81)-(0.81)-(1.15)-(-0.76)-(-0.69)-(0.66)
	100	**8.15-6.18-5.32-8.43-7.7-7.31** 0.5	2.13-2.26-2.3-2.03-2.42-2.85 (-0.85)-(0.18)-(0.32)-(-0.57)-(-0.3)-(0.31)
	250	**8.14-6.22-5.27-8.41-7.73-7.77** 0.45	1.89-2.38-2.16-1.89-1.94-2.38 (-0.86)-(0.22)-(0.27)-(-0.59)-(-0.27)-(0.77)

4.1.5 Finite Mixture of C- and D-vines

Above described finite mixture models can be easily adjusted to cover both C- and D-vine structure within different components. Instead of taking weighted sum of two different C-vine density functions, both C- and D- vine densities can be combined under the finite mixture framework for the 1'st and 2'nd component, respectively. The main motivation for such a mixture model is having no certain prior information about the tree structure for a vine model. Furthermore, D-vine is more flexible than C-vine because of its ability to select pairs freely and this property adds more power for such a mixture model to consider complex probability density functions. To illustrate, only one simulated data was investigated under this subsection, similar to above discussed simulation results.

For a finite mixture of C- and D-vine models, 2 Component, 4 dimensional case was considered with different pair copula functions. More rigorously, Clayton and Frank copula families were implemented for the construction of finite mixture of C- and D-vine models having the same parameter values. Under this mixture model, when the weight parameter, i.e. π_1 ranges from 0 to 1, the mixture model has been changed

from 1 component C-vine to 1 component D-vine model with the corresponding copulas and their parameters. Whenever $\pi_i \in [0,1]$ for $i = 1,2$, the multivariate hidden dependence structure varies from C-vine to D- vine. To illustrate, following figures (Figure 4.2-4.4) were generated to understand the dependence structure among the variables of two distinct components.

In this mixture model, the generated complex multivariate data includes both positive and negative dependence structures with respect to Clayton and Frank copula pairs for 1'st and 2'nd components, respectively. The selected parameters are $(\beta^1_{12} = 8, \beta^1_{13} = 7, \beta^1_{14} = 6, \beta^1_{23|1} = 1.7, \beta^1_{24|1} = 2.5, \beta^1_{34|12} = 1.4)$ and $(\beta^2_{12} = 9, \beta^2_{13} = 6, \beta^2_{14} = 5, \beta^2_{23|1} = -1.8, \beta^2_{24|1} = -1.6, \beta^2_{34|12} = -2.3)$ with various weight parameters. Figures 4.2-4.4 summarize the change on the dependence among the variables under numerous C- and D- vine finite mixture models based on different weight parameters.

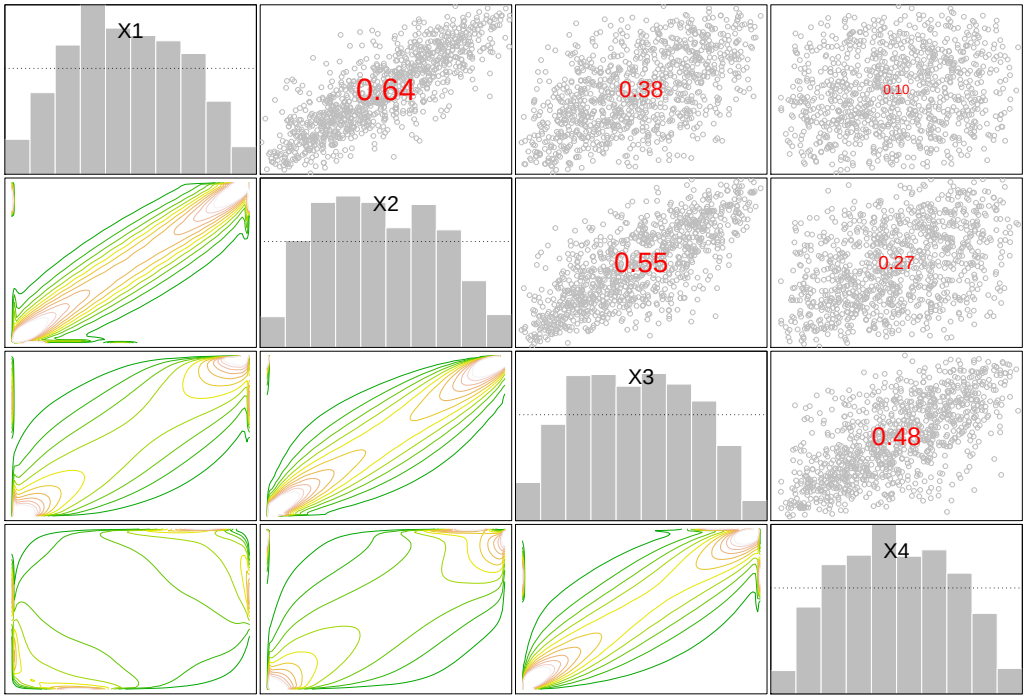


Figure 4.2: Dependence Structure of 2-Component C- and D-vine mixture with $\pi_1 = 0.25$ and $\pi_2 = 0.75$ for the 1'st and 2'nd components

In Figure 4.2, the dominant component was the 2'nd one with both strong and weak dependence parameters. Within the mixture model, the most and least significant

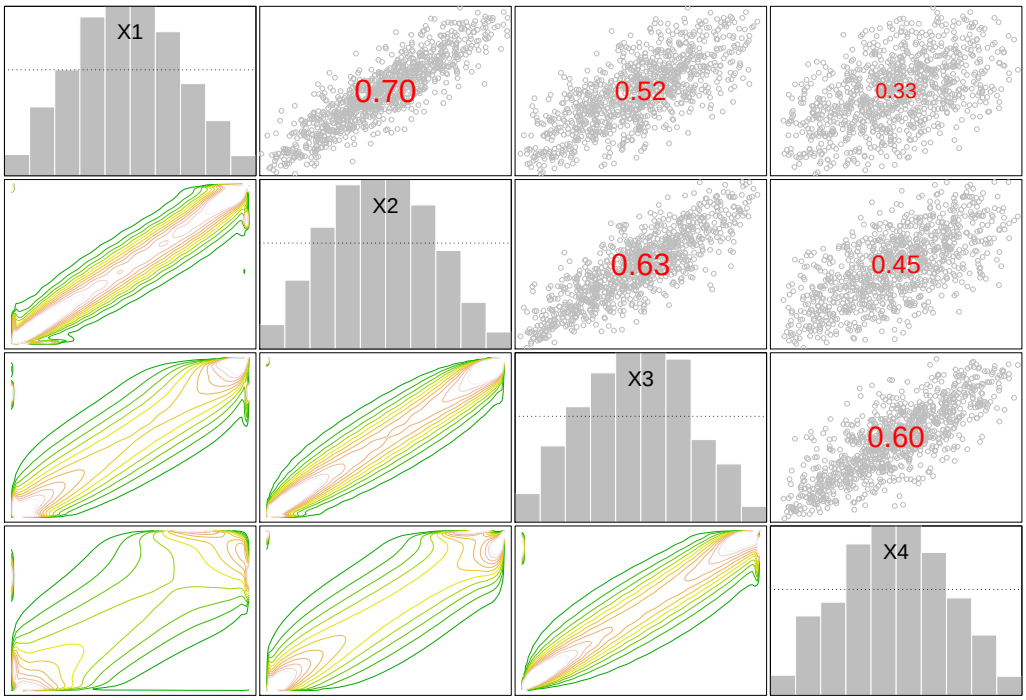


Figure 4.3: Dependence Structure of 2-Component C- and D-vine mixture with $\pi_1 = 0.5$ and $\pi_2 = 0.5$ for the 1'st and 2'nd components

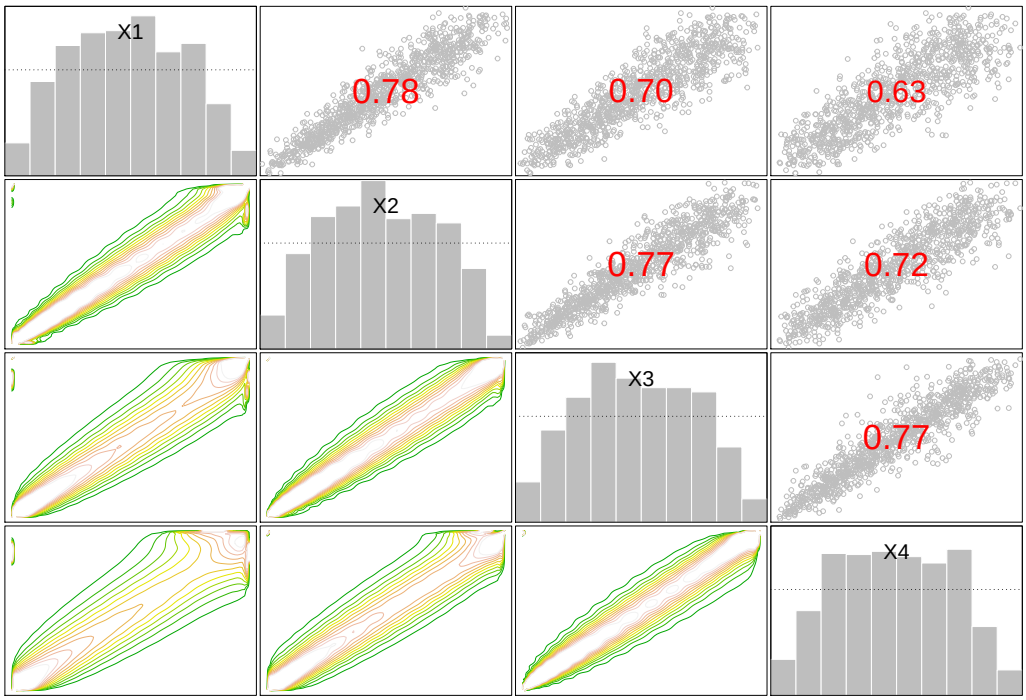


Figure 4.4: Dependence Structure of 2-Component C- and D-vine mixture with $\pi_1 = 0.75$ and $\pi_2 = 0.25$ for the 1'st and 2'nd components

dependence structure exist between the pairs (X_1, X_2) and (X_1, X_4), respectively. Even if the conditional densities were constructed using negative parameters in 2'nd component, the mixture model do not exhibit any negative dependence pattern. The similar dependence structure exist for the case visualized in Figure 4.3 and 4.4, as the weight of the 1'st component increases, the dependence among the variables belonging finite mixture of to C- and D-vine has been increased either.

In contrast to above visualized Clayton-Frank SW case, for the comparison of various mixture models, same copula families (Frank pairs) are considered with parameters $(\beta^1_{12} = 8, \beta^1_{13} = 7, \beta^1_{14} = 6, \beta^1_{23|1} = 9, \beta^1_{24|1} = 8, \beta^1_{34|12} = 7)$ and $(\beta^2_{12} = 9, \beta^2_{23} = 6, \beta^2_{34} = 5, \beta^2_{13|2} = 9, \beta^2_{24|3} = 8, \beta^2_{14|23} = 7)$. The model comparison part and the estimated parameters have been presented below in detail.

For model comparison, differently from the model comparison in finite mixture of C-vines, the performance of the proposed model has been considered under misspecified data. For instance, in the above mentioned scenario with Frank pairs, the simulated data has been generated using three different cases, like the 1'st component is generated based on C-vine structure and 2'nd component is constructed also using C-vine model. Here, there exists three multivariate data for the base model, for FrankCV-FrankDV case, where notations CV and DV exhibits C- and D-vine finite mixture model with Frank pairs. The other two data sets generated from FrankCV-FrankCV and FrankDV–FrankDV mixture models with equal weights. In this case, data set comes from CV-DV, CV-CV and DV-DV combination for 2 component mixture model with Frank pairs.

For the parameter estimation part, DEoptim function has been implemented for this mixture model. Besides, the number of iterations and the observations has been reduced as a result of the computational cost of DEoptim function, but the obtained parameter estimations for such a mixture are very promising indeed. The parameter estimations are presented separately for different number of iterations for $max1 = 100$ and $max1 = 250$.

In above Tables 4.25 and 4.26, parameter estimations of the same model fitted to three different data set. In the base model, FrankCV-FrankDV case is tested in terms of data set comes from CV-DV, CV-CV and DV-DV mixtures. Here, model identification

Table 4.23: 2 component 4-dimensional C- and D-vine mixtures with $max1 = 100$ (Base is FrankCV-FrankDV)

Mult. Data	**Size**	**Model Selection**		
C1-C2	N	AIC	BIC	CAIC
FrankCV-FrankDV	50	-181.4804	-156.6241	-143.6241
	100	-443.2836	-409.4164	-396.4164
	250	-1248.77	-1202.991	-1189.991
FrankCV-FrankCV	50	**-196.2398**	**-171.3835**	**-158.3835**
	100	**-457.866**	**-423.9988**	**-410.9988**
	250	**-1258.743**	**-1212.964**	**-1199.964**
FrankDV-FrankDV	50	-192.4237	-167.5674	-154.5674
	100	-449.6968	-415.8295	-402.8295
	250	-1237.198	-1191.419	-1178.419

Table 4.24: 2 component 4-dimensional C- and D-vine mixtures with $max1 = 250$ (Base is FrankCV-FrankDV)

Mult. Data	**Size**	**Model Selection**		
C1-C2	N	AIC	BIC	CAIC
FrankCV-FrankDV	50	-182.1744	-157.3181	-144.3181
	100	-439.8372	-405.97	-392.97
	250	-1251.043	-1205.264	-1192.264
FrankCV-FrankCV	50	**-202.0385**	**-177.1822**	**-164.1822**
	100	**-461.5943**	**-427.7271**	**-414.7271**
	250	**-1266.251**	**-1220.472**	**-1207.472**
FrankDV-FrankDV	50	-198.3203	-173.464	-160.464
	100	-448.103	-414.2358	-401.2358
	250	-1230.419	-1184.64	-1171.64

Table 4.25: Parameter Estimations for C- and D-vine mixtures with $max1 = 100$ (Base is FrankCV-FrankDV)

		Parameters of the Mixture Model	
		β^1_{12}-β^1_{13}-β^1_{14}-$\beta^1_{23\|1}$-$\beta^1_{24\|1}$-$\beta^1_{34\|12}$ (8-7-6-9-8-7) ($\pi_1 = 0.5$)	
Mult. Data	**Size**	Est.	St.dv.
C1	N		(Bias)
FrankCV-DV	50	**6.05-8.44-8.05-6.92-7.72-8.02** 0.1	3.28-2.24-2.6-3.74-3.32-3.38 (-1.95)-(1.44)-(2.05)-(-2.08)-(-0.28)-(1.02)
	100	**5.91-7.48-8.23-7.11-7.39-7.4** 0.05	3.32-3.28-2.58-3.8-3.58-3.9 (-2.09)-(0.48)-(2.23)-(-1.89)-(-0.61)-(0.4)
	250	**8.1-7.02-6.04-8.95-7.97-6.94** 0.97	0.57-0.52-0.5-0.74-0.53-0.82 (0.1)-(0.02)-(0.04)-(-0.05)-(-0.03)-(-0.06)
FrankCV-CV	50	8.32-6.98-6.1-8.44-7.6-7.71 0.89	1.34-1.32-1.18-1.31-1.28-1.6 (0.32)-(-0.02)-(0.1)-(-0.56)-(-0.4)-(0.71)
	100	8.35-7.05-6.05-8.83-7.7-7.3 0.94	1.07-0.94-0.96-0.97-1-1.09 (0.35)-(0.05)-(0.05)-(-0.17)-(-0.3)-(0.3)
	250	8.1-6.99-5.98-9.05-7.97-7.15 0.97	0.71-0.63-0.61-0.65-0.7-0.63 (0.1)-(-0.01)-(-0.02)-(0.05)-(-0.03)-(0.15)
FrankDV-DV	50	5.87-7.95-7.42-6.87-7.41-6.73 0.07	3.28-2.88-3.23-3.94-3.6-4.17 (-2.13)-(0.95)-(1.42)-(-2.13)-(-0.59)-(-0.27)
	100	5.47-8.1-7.79-5.19-8.59-7.47 0.03	3.39-2.82-3.1-3.87-2.9-3.84 (-2.53)-(1.1)-(1.79)-(-3.81)-(0.59)-(0.47)
	250	4.17-7.24-8.27-7.65-8-8.42 0.01	3.28-3.1-2.84-3.51-3.2-3.31 (-3.83)-(0.24)-(2.27)-(-1.35)-(0)-(1.42)
		β^2_{12}-β^2_{13}-β^2_{14}-$\beta^2_{23\|1}$-$\beta^2_{24\|1}$-$\beta^2_{34\|12}$ (9-6-5-9-8-7) ($\pi_2 = 0.5$)	
Mult. Data	**Size**	Est.	St.dv.
C2	N		(Bias)
FrankCV-DV	50	**8.71-5.99-5.13-7.94-7.9-5.44** 0.9	1.3-1.33-1.43-1.65-1.8-1.82 (-0.29)-(-0.01)-(0.13)-(-1.06)-(-0.1)-(-1.56)
	100	**8.98-6.04-4.99-8.46-8.14-5.97** 0.95	0.89-0.83-0.84-1.27-0.98-1.38 (-0.02)-(0.04)-(-0.01)-(-0.54)-(0.14)-(-1.03)
	250	**5.6-7.53-8.61-7.17-6.67-8.86** 0.03	3.72-2.83-2.46-3.68-3.8-2.9 (-3.4)-(1.53)-(3.61)-(-1.83)-(-1.33)-(1.86)
FrankCV-CV	50	5.68-8.58-8.3-7.16-6-8.46 0.11	3.49-2.21-2.62-3.65-3.92-3.16 (-3.32)-(2.58)-(3.3)-(-1.84)-(-2)-(1.46)
	100	6.01-7.84-8.25-6.69-7.79-8.99 0.06	3.53-2.9-2.65-3.63-3.22-2.71 (-2.99)-(1.84)-(3.25)-(-2.31)-(-0.21)-(1.99)
	250	4.98-7.4-8.58-6.12-6.78-8.73 0.03	3.5-2.77-2.27-4.04-3.64-3 (-4.02)-(1.4)-(3.58)-(-2.88)-(-1.22)-(1.73)
FrankDV-DV	50	9.03-5.89-4.87-8.29-8.1-5.83 0.93	1-1.02-1.23-1.47-1.36-1.69 (0.03)-(-0.11)-(-0.13)-(-0.71)-(0.1)-(-1.17)
	100	8.9-5.95-4.91-8.63-7.98-6.11 0.97	0.8-0.59-0.82-0.95-0.9-1.05 (-0.1)-(-0.05)-(-0.09)-(-0.37)-(-0.02)-(-0.89)
	250	9-6.01-5.02-8.94-8.05-6.56 0.99	0.55-0.45-0.45-0.53-0.56-0.63 (0)-(0.01)-(0.02)-(-0.06)-(0.05)-(-0.44)

Table 4.26: Parameter Estimations for C- and D-vine mixtures with $max1 = 250$ (Base is FrankCV-FrankDV)

		Parameters of the Mixture Model	
		β^1_{12}-β^1_{13}-β^1_{14}-$\beta^1_{23\|1}$-$\beta^1_{24\|1}$-$\beta^1_{34\|12}$ (8-7-6-9-8-7) ($\pi_1 = 0.5$)	
Mult. Data	**Size**	Est.	St.dv.
C1	N		(Bias)
FrankCV-DV	50	**6.74-8.25-7.73-6.74-7.57-7.55**	3.16-2.63-2.58-3.88-3.27-3.66
		0.16	(-1.26)-(1.25)-(1.73)-(-2.26)-(-0.43)-(0.55)
	100	**5.07-7.63-8.03-6.97-7.67-8.27**	3.35-3.09-2.76-3.8-3.43-3.41
		0.04	(-2.93)-(0.63)-(2.03)-(-2.03)-(-0.33)-(1.27)
	250	**8.08-7.03-5.98-8.87-7.93-6.99**	0.65-0.56-0.51-0.77-0.69-0.76
		0.97	(0.08)-(0.03)-(-0.02)-(-0.13)-(-0.07)-(-0.01)
FrankCV-CV	50	8.17-6.94-5.96-8.73-7.83-7.47	1.41-1.34-1.3-1.36-1.45-1.59
		0.91	(0.17)-(-0.06)-(-0.04)-(-0.27)-(-0.17)-(0.47)
	100	8.03-6.91-5.91-8.92-7.89-7.14	1.02-0.94-0.89-0.96-0.96-1.1
		0.95	(0.03)-(-0.09)-(-0.09)-(-0.08)-(-0.11)-(0.14)
	250	8.12-7.01-6.02-9.05-8-7.06	0.66-0.58-0.57-0.67-0.66-0.62
		0.97	(0.12)-(0.01)-(0.02)-(0.05)-(0)-(0.06)
FrankDV-DV	50	5.54-8.04-7.68-6.56-7.54-7.68	3.22-2.87-3-3.93-3.6-3.8
		0.06	(-2.46)-(1.04)-(1.68)-(-2.44)-(-0.46)-(0.68)
	100	4.97-7.63-7.93-6.8-7.65-8.39	3.35-3.17-2.78-3.8-3.56-3.32
		0.03	(-3.03)-(0.63)-(1.93)-(-2.2)-(-0.35)-(1.39)
	250	5.13-7.19-7.89-8.26-7.33-8.14	3.22-3.1-2.9-2.89-3.69-3.38
		0.01	(-2.87)-(0.19)-(1.89)-(-0.74)-(-0.67)-(1.14)
		β^2_{12}-β^2_{13}-β^2_{14}-$\beta^2_{23\|1}$-$\beta^2_{24\|1}$-$\beta^2_{34\|12}$ (9-6-5-9-8-7) ($\pi_2 = 0.5$)	
Mult. Data	**Size**	Est.	St.dv.
C2	N		(Bias)
FrankCV-DV	50	**8.51-6.1-5.16-8.06-7.87-5.78**	1.57-1.5-1.68-1.91-1.76-2.11
		0.84	(-0.49)-(0.1)-(0.16)-(-0.94)-(-0.13)-(-1.22)
	100	**8.94-5.91-4.91-8.59-7.95-6.04**	0.85-0.73-0.81-1.05-1.07-1.07
		0.96	(-0.06)-(-0.09)-(-0.09)-(-0.41)-(-0.05)-(-0.96)
	250	**5.7-7.22-8.16-6.35-7.43-9.27**	3.44-3.08-2.6-3.81-3.55-2.3
		0.03	(-3.3)-(1.22)-(3.16)-(-2.65)-(-0.57)-(2.27)
FrankCV-CV	50	5.84-7.74-8.59-6.79-6.8-8.48	3.36-2.94-2.36-3.93-3.75-3.18
		0.09	(-3.16)-(1.74)-(3.59)-(-2.21)-(-1.2)-(1.48)
	100	6.02-7.18-8.35-6.93-7.01-9.01	3.4-3.06-2.59-3.7-3.69-2.65
		0.05	(-2.98)-(1.18)-(3.35)-(-2.07)-(-0.99)-(2.01)
	250	4.97-7.39-8.58-6.6-7.07-8.74	3.45-2.79-2.29-3.92-3.61-2.95
		0.03	(-4.03)-(1.39)-(3.58)-(-2.4)-(-0.93)-(1.74)
FrankDV-DV	50	8.77-5.82-5.01-8.64-8.06-5.78	1.18-1.03-1.07-1.21-1.37-1.39
		0.94	(-0.23)-(-0.18)-(0.01)-(-0.36)-(0.06)-(-1.22)
	100	8.94-5.89-4.91-8.81-8.12-6.12	0.84-0.68-0.79-0.97-1.05-1.06
		0.97	(-0.06)-(-0.11)-(-0.09)-(-0.19)-(0.12)-(-0.88)
	250	8.9-5.95-4.97-8.83-7.99-6.56	0.56-0.41-0.48-0.64-0.57-0.59
		0.99	(-0.1)-(-0.05)-(-0.03)-(-0.17)-(-0.01)-(-0.44)

problem emerges one more time for the finite mixture of C- and D-vine copula with Frank pairs. In terms of small data set, the best estimates of the parameters belonging to the DV-part, although this result became reverse when $N = 250$. Overall, FrankCV–FrankDV model fails to identify the simulated data from CV-DV mixture and parameter estimates are only significant for one of the components based on the weight parameter. Additionally, when the sample size increased, there is a weight change from D-vine to C-vine in terms of components.

4.1.6 CD-vine Mixture Model

As a novel contribution of this study, the mixture of C-vines with D-vine methodology has been discussed. Here, for the pre-analysis part, various C-vine densities have been investigated in 4-dimension with distinct pair copula families like Clayton, Frank or Gumbel for all components when $M = 3$. For instance, in 3 component 4 dimensional CD-vine mixture model, as it is explained in Chapter 3, each component is modeled via C-vine and then the dependence among the components has been captured with the help of D-vine. In this methodology, the dependence between the values of Empirical Multivariate Cumulative Distribution function (EMCDF) belonging to each component, i.e. $F(cv_1)$, $F(cv_2)$ and $F(cv_3)$ was the crucial and facilitator inter-step for CD-vine mixture modeling.

As an optimization routine, previously considered DEoptim function was implemented under two step maximization process for CD-vine mixture model. In this proposed model, similar to the widely evaluated Inference For Margins (IFM) method, first the parameters of density for each component was estimated independently. Afterwords, these obtained parameters are used for the construction of $F(cv_1)$, $F(cv_2)$ and $F(cv_3)$ values, the EMCDF data having dependence information about each component. Finally, new 3-dimensional EMCDF data was tied together with the help of D-vine methodology.

4.1.6.1 Dependence only within components

Suppose each component has Clayton pairs but there is no temporal association between the variables under each component and dependence among the calculated values of $F(cv_1)$, $F(cv_2)$ and $F(cv_3)$ for D-vine specification. In this framework, the most suitable copula family for D-vine modeling part has been detected as the independence copula. To illustrate, Figures 4.5 and 4.6 can be considered, to show the existing dependence occurs only within the components under strong dependence scenario. Before observing independence among the values of $F(cv_1)$, $F(cv_2)$ and $F(cv_3)$ given in Figure 4.6, the dependence within the the components originated from the selected C-vine densities based on Clayton pairs is visualized in Figure 4.5.

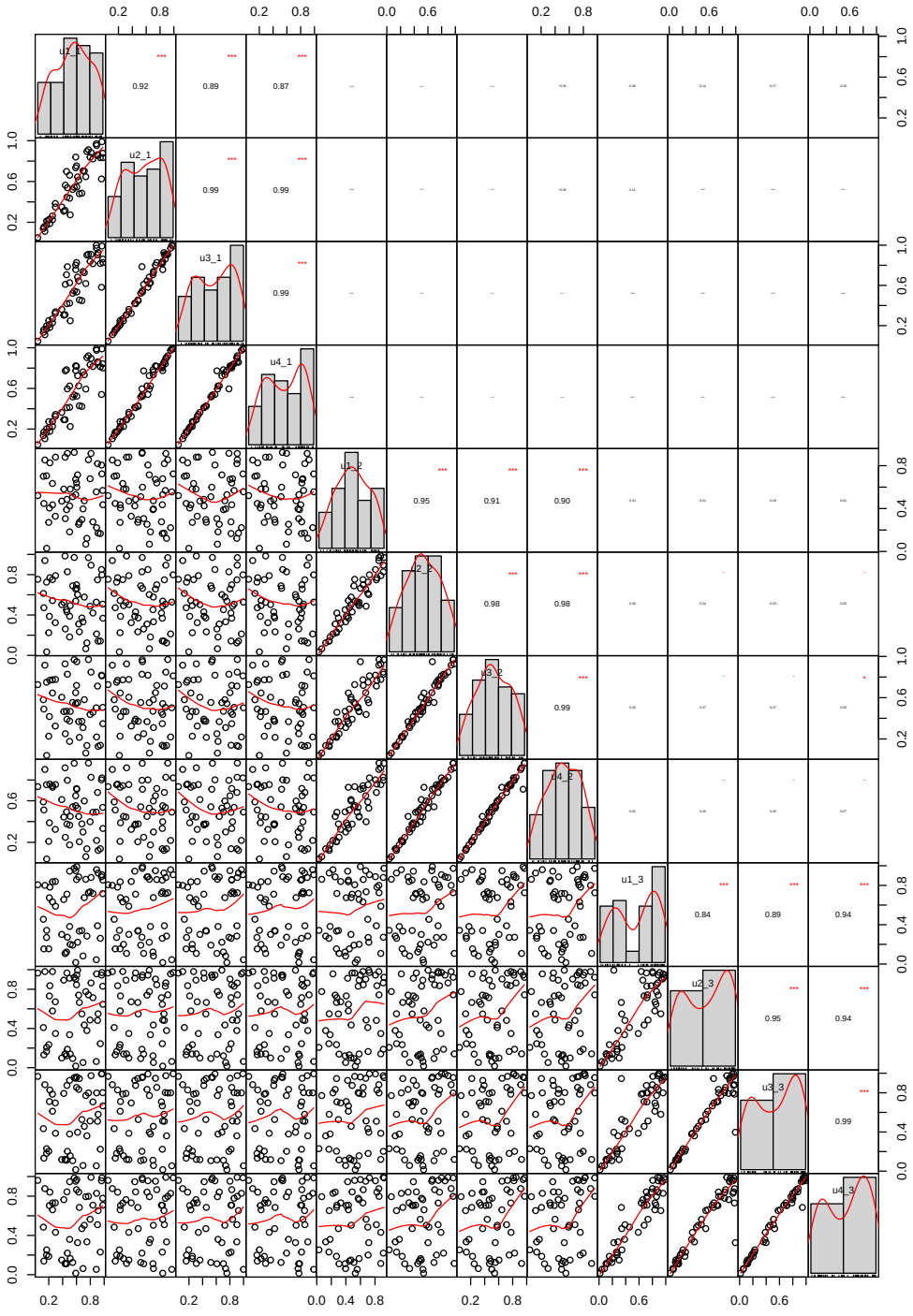


Figure 4.5: Dependence Structure among all variables, each component was generated by Clayton pairs and there is no association among components

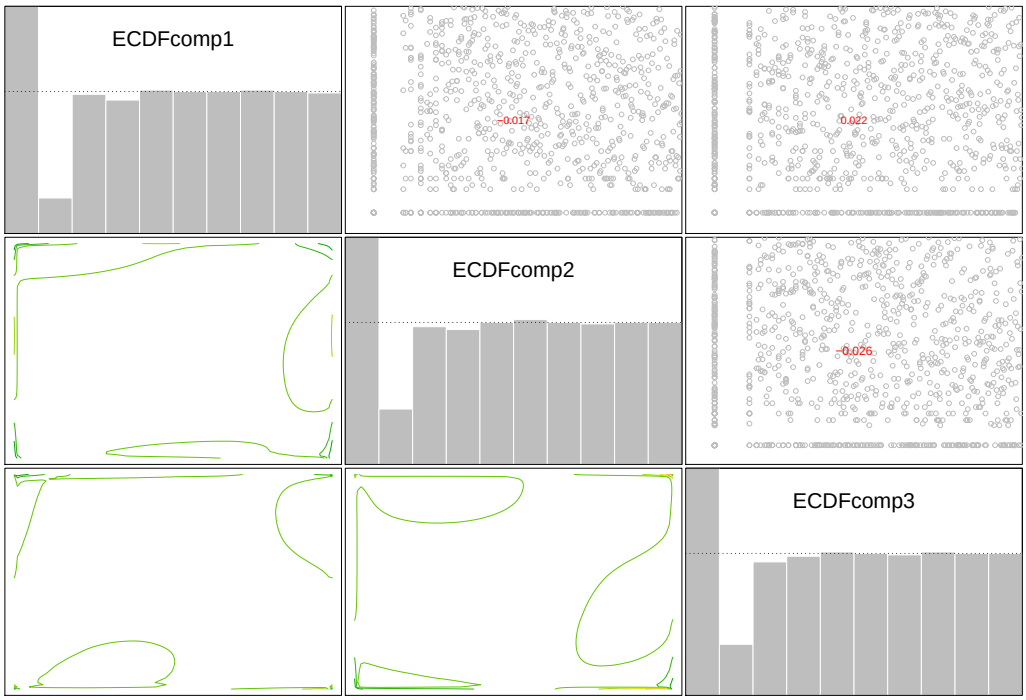


Figure 4.6: Dependence Structure of $F(cv_1)$, $F(cv_2)$ and $F(cv_3)$ from a 3-Component C-vine mixture with Clayton pairs

In Figure 4.6, such a CD-vine construction requires independence copula family for D-vine part in the second step. For this reason, CD-vine mixture model, which includes Clayton pairs in C-vine and independence copula pairs in D-vine construction has been studied. In this modeling setup, the parameters of each component are selected as $(\beta^1_{12} = 8, \beta^1_{13} = 7, \beta^1_{14} = 6, \beta^1_{23|1} = 9, \beta^1_{24|1} = 8, \beta^1_{34|12} = 7)$, $(\beta^2_{12} = 9, \beta^2_{13} = 6, \beta^2_{14} = 5, \beta^2_{23|1} = 9, \beta^2_{24|1} = 8, \beta^2_{34|12} = 7)$ and $(\beta^3_{12} = 4, \beta^3_{13} = 5, \beta^3_{14} = 7, \beta^3_{23|1} = 9, \beta^3_{24|1} = 8, \beta^3_{34|12} = 7)$. Naturally, the parameter set for the D-vine part in this model is denoted by $(\beta^{DV}_{12}, \beta^{DV}_{23}, \beta^{DV}_{13|2})$. Simulation results for the case of Clayton-Independence pairs belonging to C- and D-vine part are summarized. For different number of iterations (i.e. $max1 = 100$ and $max1 = 250$), similar results (Tables 4.27 and 4.28) are presented below.

Table 4.27: Parameter Estimations for CD-vine mixture with Clayton-Independence copula pairs for $max1 = 100$

		Parameters of the Mixture Model	
		$\beta^1_{12}, \beta^1_{13}, \beta^1_{14}, \beta^1_{23\|1}, \beta^1_{24\|1}, \beta^1_{34\|12}$	
C-vine part	**Size**	(8-7-6-9-8-7)	St.dv.
CV1	N	Est.	(Bias)
Clayton	50	**8.17-7.13-6.12-8.91-7.95-7.64**	0.79-0.67-0.59-0.94-0.91-1.32 (0.17)-(0.13)-(0.12)-(-0.09)-(-0.05)-(0.64)
	100	**8.05-7.04-6.03-9.12-8.1-7.03**	0.58-0.52-0.46-0.68-0.61-0.78 (0.05)-(0.04)-(0.03)-(0.12)-(0.1)-(0.03)
	250	**8.06-7.05-6.04-8.96-7.98-7.13**	0.34-0.3-0.26-0.54-0.5-0.62 (0.06)-(0.05)-(0.04)-(-0.04)-(-0.02)-(0.13)
		$\beta^2_{12}, \beta^2_{13}, \beta^2_{14}, \beta^2_{23\|1}, \beta^2_{24\|1}, \beta^2_{34\|12}$	
		(9-6-5-9-8-7)	St.dv.
CV2	N	Est.	(Bias)
Clayton	50	**9.1-6.06-5.06-9.03-8.02-7.4**	0.7-0.51-0.45-0.92-0.78-1.26 (0.1)-(0.06)-(0.06)-(0.03)-(0.02)-(0.4)
	100	**9.07-6.04-5.04-8.94-7.97-7.2**	0.59-0.41-0.36-0.72-0.66-0.84 (0.07)-(0.04)-(0.04)-(-0.06)-(-0.03)-(0.2)
	250	**9.07-6.05-5.05-8.92-7.92-7.04**	0.37-0.26-0.22-0.52-0.46-0.51 (0.07)-(0.05)-(0.05)-(-0.08)-(-0.08)-(0.04)
		$\beta^3_{12}, \beta^3_{13}, \beta^3_{14}, \beta^3_{23\|1}, \beta^3_{24\|1}, \beta^3_{34\|12}$	
		(4-5-7-9-8-7)	St.dv.
CV3	N	Est.	(Bias)
Clayton	50	**4.05-5.05-7.06-9.05-8.03-7.35**	0.39-0.44-0.58-1.03-0.99-1.18 (0.05)-(0.05)-(0.06)-(0.05)-(0.03)-(0.35)
	100	**4.01-5-7-9.13-8.14-7.15**	0.25-0.29-0.39-0.66-0.65-0.83 (0.01)-(0)-(0)-(0.13)-(0.14)-(0.15)
	250	**4.03-5.04-7.05-9.11-8.07-7.04**	0.17-0.2-0.26-0.49-0.45-0.57 (0.03)-(0.04)-(0.05)-(0.11)-(0.07)-(0.04)
		$\beta^{DV}_{12}, \beta^{DV}_{23}, \beta^{DV}_{13\|2}$	
D-vine part	**Size**	(0 / 0 / 0)	St.dv.
DV	N	Est.	(Bias)
Indep	50	**-1.396413e-07 / -2.385398e-07 / -1.157744e-07**	1.133487e-07 / 1.85544e-07 / 1.934761e-07 (-1.396413e-07)-(-2.385398e-07)-(-1.157744e-07)
	100	**-0.0006554143 / -0.0007067731 / -0.0006302542**	0.000579263 / 0.0006077416 / 0.0005635711 (-0.0006554143)-(-0.0007067731)-(-0.0006302542)
	250	**-0.0007759255 / -0.0006201724 / -0.000564488**	0.0007099019 / 0.0005251071 / 0.0005843439 (-0.0007759255)-(-0.0006201724)-(-0.000564488)

In Table 4.27, obtained parameter estimates are very promising for CD-vine mixture with Clayton-Independence case. Besides, there is positive impact of the increase on the sample size over the accuracy, when the standard deviation and bias values are investigated. In this type CD-vine mixture case, it is known that the D-vine part requires independence copulas in advance, that makes the above derived values are reasonable based on the simulated data. Similar parameter estimations are presented for $max1 = 250$ below, in Table 4.28.

Table 4.28: Parameter Estimations for CD-vine mixture with Clayton-Independence copula pairs for $max1 = 250$

		Parameters of the Mixture Model	
		$\beta^1_{12}, \beta^1_{13}, \beta^1_{14}, \beta^1_{23\|1}, \beta^1_{24\|1}, \beta^1_{34\|12}$	
C-vine part	**Size**	(8-7-6-9-8-7)	St.dv.
CV1	N	Est.	(Bias)
Clayton	50	**8.19-7.19-6.16-8.99-7.98-7.25**	0.73-0.64-0.5-0.95-0.86-1.14 (0.19)-(0.19)-(0.16)-(-0.01)-(-0.02)-(0.25)
	100	**8.09-7.07-6.06-8.94-7.93-7.26**	0.61-0.53-0.47-0.81-0.74-0.89 (0.09)-(0.07)-(0.06)-(-0.06)-(-0.07)-(0.26)
	250	**8.06-7.05-6.04-9.01-8.01-7.01**	0.38-0.33-0.29-0.56-0.5-0.55 (0.06)-(0.05)-(0.04)-(0.01)-(0.01)-(0.01)
		$\beta^2_{12}, \beta^2_{13}, \beta^2_{14}, \beta^2_{23\|1}, \beta^2_{24\|1}, \beta^2_{34\|12}$	
		(9-6-5-9-8-7)	St.dv.
CV2	N	Est.	(Bias)
Clayton	50	**9.1-6.08-5.07-9.03-8.03-7.26**	0.7-0.48-0.42-0.93-0.86-1.21 (0.1)-(0.08)-(0.07)-(0.03)-(0.03)-(0.26)
	100	**9.06-6.04-5.03-9.03-8.02-7.23**	0.59-0.41-0.36-0.72-0.67-0.9 (0.06)-(0.04)-(0.03)-(0.03)-(0.02)-(0.23)
	250	**9.04-6.03-5.03-9.03-8.02-7.07**	0.39-0.27-0.23-0.52-0.47-0.6 (0.04)-(0.03)-(0.03)-(0.03)-(0.02)-(0.07)
		$\beta^3_{12}, \beta^3_{13}, \beta^3_{14}, \beta^3_{23\|1}, \beta^3_{24\|1}, \beta^3_{34\|12}$	
		(4-5-7-9-8-7)	St.dv.
CV3	N	Est.	(Bias)
Clayton	50	**4.04-5.05-7.08-9.02-8.02-7.53**	0.37-0.43-0.57-0.91-0.87-1.23 (0.04)-(0.05)-(0.08)-(0.02)-(0.02)-(0.53)
	100	**4.03-5.04-7.05-9.03-8.03-7.26**	0.26-0.29-0.39-0.74-0.65-0.89 (0.03)-(0.04)-(0.05)-(0.03)-(0.03)-(0.26)
	250	**4.02-5.02-7.03-9.02-8.03-7.08**	0.16-0.18-0.24-0.52-0.47-0.51 (0.02)-(0.02)-(0.03)-(0.02)-(0.03)-(0.08)
		$\beta^{DV}_{12}, \beta^{DV}_{23}, \beta^{DV}_{13\|2}$	
D-vine part	**Size**	(0/0/0)	St.dv.
DV	N	Est.	(Bias)
Indep	50	**-0.0006874875 / -0.0006659864 / -0.0006871078**	0.000694688 / 0.0007766992 / 0.0006787283 (-0.0006874875)-(-0.0006659864)-(-0.0006871078)
	100	**-0.0006154219 / -0.0006480791 / -0.0006335947**	0.0005524461 / 0.0006746858 / 0.0006316712 (-0.0006154219)-(-0.0006480791)-(-0.0006335947)
	250	**-0.0007068852 / -0.0006224004 / -0.0007062854**	0.000673962 / 0.0006086603 / 0.0006704548 (-0.0007068852)-(-0.0006224004)-(-0.0007062854)

In this two-step maximization with the case of Clayton-Independence CD-vine mixture model, the parameter estimations for both components are promising enough. When there is no dependence among the components via $F(cv_1)$, $F(cv_2)$ and $F(cv_3)$ values, D-vine part requires only Independence copula pairs and it is straightforward to

capture high accuracy for the parameter estimations of such a CD-vine mixture model. On the other hand, upcoming scenario, presented below, is not so easy to accomplish the model identification under the proposed CD-vine mixture.

4.1.6.2 Dependence among the components

Within the context of CD-vine mixture, another dependence pattern occurs whenever EMCDF values of each component were calculated at the same values, which means realization of P($u_1 \leq a, u_2 \leq b, u_3 \leq c, u_4 \leq d$) where a, b, c, d are arbitrary real values belonging to the domain of multivariate density function. For this scenario, Frank pairs with strong dependence case were constructed for each component. Afterwords, $F(cv_1)$, $F(cv_2)$ and $F(cv_3)$ values are calculated for the same random matrix including $a, b, c, d \in [0, 1]$. In this scenario, there exist a dependence among the components via EMCDF values, namely among $F(cv_1)$, $F(cv_2)$ and $F(cv_3)$ so that independence copula is not an appropriate choice anymore.

For the second simulated data, same 4 dimensional 3 component C-vine models are generated based on the selected parameters, used previously. For the sake of simplicity, some archimedean copula pairs are considered for the second step of CD-vine mixture model. On the other hand, this dependence pattern within the components via $F(cv_1)$, $F(cv_2)$ and $F(cv_3)$ is not predictable at the beginning. For this reason, the best D-vine model should be selected after finishing the first step. However, for the simulation part, pre-selected pairs are forced to fit for the obtained data set to investigate the performance of the CD-vine model when copula families are not specified correctly. Parameter estimation results are available for Clayton-Frank CD-vine mixture model for small sample size whenever the true structure includes other copula pairs rather than Frank pairs in D-vine part.

In C-vine part, the parameter estimation accuracy is very high for Clayton pairs for each component. Thereafter, the dependence among the components can be modeled via D-vine framework. Even if the forced parameters of D-vine part are $\beta_{12}^{DV} = 5, \beta_{23}^{DV} = 7, \beta_{13|2}^{DV} = 9$, certainly second part fails in terms of accuracy of the parameter estimates since the original dependence structure is different. Parameter estimates for the Frank pairs in D-vine part hits the initially determined upper constraints

Table 4.29: Parameter Estimations for CD-vine mixture with Clayton-Frank copula pairs for $max1 = 100$

		Parameters of the Mixture Model	
		$\beta^1_{12}, \beta^1_{13}, \beta^1_{14}, \beta^1_{23\|1}, \beta^1_{24\|1}, \beta^1_{34\|12}$	
C-vine part	**Size**	(8-7-6-9-8-7)	St.dv.
CV1	N	Est.	(Bias)
Clayton	50	**8.08-7.08-6.08-9.06-8.05-7.3**	0.73-0.66-0.59-0.9-0.82-1.12 (0.08)-(0.08)-(0.08)-(0.06)-(0.05)-(0.3)
	100	**8.07-7.07-6.07-9.19-8.11-7.1**	0.55-0.47-0.41-0.76-0.65-0.87 (0.07)-(0.07)-(0.07)-(0.19)-(0.11)-(0.1)
	250	**8.05-7.04-6.03-9.11-8.11-6.99**	0.34-0.31-0.28-0.56-0.51-0.58 (0.05)-(0.04)-(0.03)-(0.11)-(0.11)-(-0.01)
		$\beta^2_{12}, \beta^2_{13}, \beta^2_{14}, \beta^2_{23\|1}, \beta^2_{24\|1}, \beta^2_{34\|12}$	
		(9-6-5-9-8-7)	St.dv.
CV2	N	Est.	(Bias)
Clayton	50	**9.15-6.11-5.1-9.03-7.97-7.44**	0.68-0.46-0.41-0.88-0.8-1.2 (0.15)-(0.11)-(0.1)-(0.03)-(-0.03)-(0.44)
	100	**9.02-6.04-5.03-9.06-8.06-6.89**	0.57-0.4-0.34-0.71-0.61-0.79 (0.02)-(0.04)-(0.03)-(0.06)-(0.06)-(-0.11)
	250	**9.03-6.02-5.02-9.01-8-7.08**	0.37-0.25-0.22-0.54-0.48-0.59 (0.03)-(0.02)-(0.02)-(0.01)-(0)-(0.08)
		$\beta^3_{12}, \beta^3_{13}, \beta^3_{14}, \beta^3_{23\|1}, \beta^3_{24\|1}, \beta^3_{34\|12}$	
		(4-5-7-9-8-7)	St.dv.
CV3	N	Est.	(Bias)
Clayton	50	**4.07-5.1-7.13-9.12-8.07-7.25**	0.34-0.42-0.58-0.87-0.84-1.37 (0.07)-(0.1)-(0.13)-(0.12)-(0.07)-(0.25)
	100	**3.98-4.99-6.99-9.01-8.04-7.21**	0.22-0.25-0.33-0.68-0.61-0.82 (-0.02)-(-0.01)-(-0.01)-(0.01)-(0.04)-(0.21)
	250	**4-5.01-7.01-9.01-8.03-7.09**	0.15-0.18-0.24-0.49-0.42-0.52 (0)-(0.01)-(0.01)-(0.01)-(0.03)-(0.09)
		$\beta^{DV}_{12}, \beta^{DV}_{23}, \beta^{DV}_{13\|2}$	
D-vine part	**Size**	(5 / 7 / 9)	St.dv.
DV	N	Est.	(Bias)
Frank	50	**9 / 9 / 9**	8.440675e-07 / 8.169323e-07 / 9.463762e-07 (-6.705599e-08)-(6.348697e-07)-(-9.163034e-07)
	100	**9 / 9 / 9**	8.930808e-07 / 9.90206e-07 / 9.533107e-07 (-8.499318e-07)-(9.165051e-07)-(-1.317546e-07)
	250	**9 / 9 / 9**	8.689216e-07 / 8.037634e-07 / 8.283851e-07 (-6.877701e-07)-(1.241574e-06)-(-8.252853e-07)

in all cases. For this reason, Frank pairs are not suitable for modeling the dependence between the components. Generally, this problem occurs in any kind of family restriction, since it is not suitable to identify the pair copula function for D-vine part at the beginning. Nevertheless, for a real life application, this two step maximization process can be cultivated by capturing the dependence among the components before evaluating D-vine model.

4.2 Real Life Application

In this subsection, the above proposed mixture models were tested using two different financial data sets. In the first application, finite mixture of C-vines has been tested with 2 components for log-return time series data. Secondly, the proposed CD-vine mixture model was executed for a similar log-return data set. The details of the selected financial data sets and the corresponding numerical results are summarized within the following subsections.

4.2.1 Finite Mixture of Log-returns

As a first application, Istanbul stock exchange data set is considered with other notable indices like S&P 500, and DAX. Data set is available from June 5, 2009 to February 22, 2011, including 536 rows in daily base and retrieved from the UC Irvine Machine Learning Repository [2]. Under the proposed model, the stock exchange data will be modeled by finite mixture of C-vines and for the model comparison purposes, the suitable selections are pooled together for $M = 1$ and $M = 2$.

Firstly, the considered data set with its brief explanations is presented below in Table 4.30. For the first application, the pre-analysis part has been started with all given indices. Before time series modeling, the summary statistics for each time series are given below in Table 4.31 with their plots in Figure 4.7.

Table 4.30: Log-return series definition

Stock Index	Definition
ISE_{TL}	İstanbul stock exchange, TL based BIST100
ISE_{USD}	İstanbul stock exchange, USD based BIST100
SP	Standard & Poor 500 Index
DAX	Germany Stock market return index
$FTSE$	UK Stock market return index
$NIKKEI$	Japan Stock market return index
$BOVESPA$	Brazil Stock market return index
EU	MSCI European index
EM	MSCI emerging markets index

In Table 4.30, well known international major log-return indices are considered with

ISE stock index. For a brief and more clear definition, EU is designed to represents the performance of large and mid-cap equities across 15 developed countries in Europe based on Global Investable Market Index (GIMI) methodology. Meanwhile, EM relies on the same methodology but it exhibits a similar performance of securities in 24 Emerging Markets including Turkey, this is the reason for the selection of index EM for the rest of the study. Further information about these stock indices can be found in [36].

Table 4.31: Summary statistics for Log-return series

Stock Index	Min.	Max.	Mean	Variance	Skewness	Kurtosis
ISE_{TL}	-0.062208	0.068952	0.001629	0.0002645265	-0.09471169	4.391144
ISE_{USD}	-0.084716	0.100621	0.001552	0.0004461357	-0.07994364	4.795950
SP	-0.0542620	0.0683664	0.0006433	0.0001986115	-0.09029299	6.068512
DAX	-0.0523312	0.0589505	0.0007208	0.0002119113	-0.10698160	5.104065
$FTSE$	-0.0548160	0.0503227	0.0005103	0.0001601665	-0.16194822	5.254229
$NIKKEI$	-0.0504476	0.0612293	0.0003077	0.0002205262	0.03233025	4.297928
$BOVESPA$	-0.0538495	0.0637915	0.0009353	0.0002480795	0.02934667	4.853999
EU	-0.0488168	0.0670425	0.0004706	0.0001687406	-0.03082931	5.607954
EM	-0.0385645	0.0478045	0.0009359	0.0001102742	0.08273992	5.510568

In Table 4.31, the important statistical properties for each time series are summarized. Based on the obtained values, the measures of dispersion for the log-returns are very similar, with different means. Furthermore, skewness values exhibit different symmetric properties. For instance, right tail structure occurs for only the series $NIKKEI$, $BOVESPA$ and EM, the remaining log-returns display negative skewness values, i.e. left tail property. Based on the kurtosis values, the highest and smallest peakedness exist for the series SP and $NIKKEI$, respectively. For each univariate series, large values of kurtosis display a significant departure from normality and express fatter tails for the series.

As a first insight for the data set, the dependence among each univariate time series are presented in Figure 4.8 with the corresponding Kendall's τ values within the upper diagonal windows and bivariate scatter plots in the lower diagonal graphs. Furthermore, the most correlated log-return set includes ISE_{USD} with DAX, $FTSE$ and EM. These series are considered for further analysis to fit a suitable mixture model. Besides, the presented Kendall's τ values exhibit a plausible strong positive dependence for ISE_{USD} and it could be considered as a root node for C-vine model

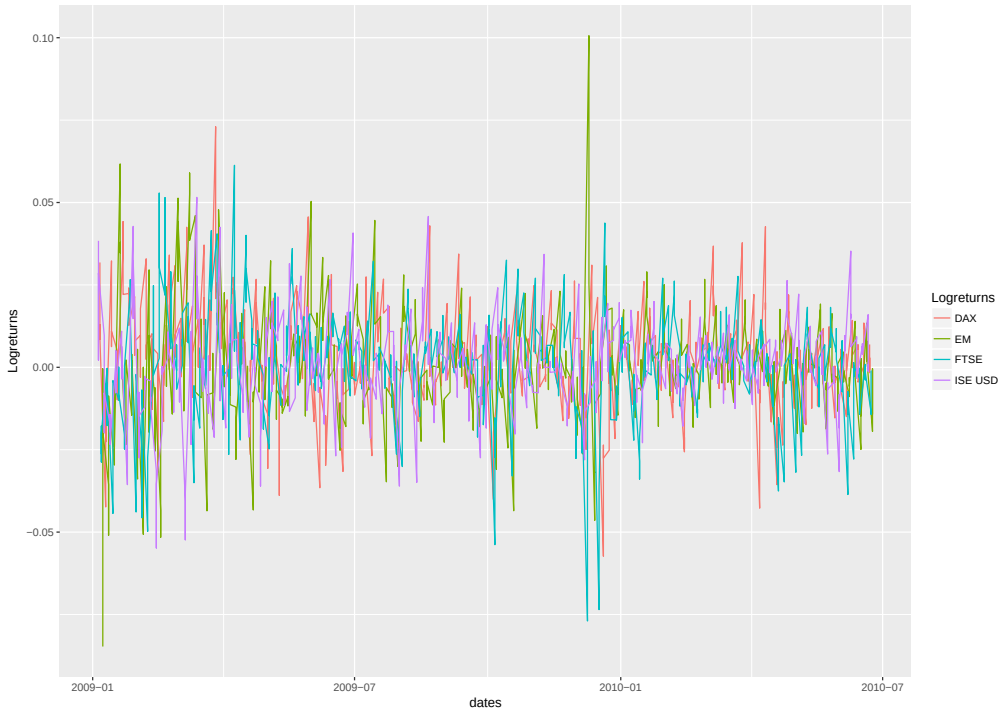


Figure 4.7: Univariate log-returns between 05-Jan-09 and 22-Febr-11

within a multivariate data set, i.e. $(ISE_{USD}, DAX, FTSE, EM)$. Here, the original univariate log-return series with their ACF and PACF plots are presented for ISE_{USD}, DAX, $FTSE$ and EM in the following Figures 4.9, 4.10, 4.11 and 4.12, respectively.

To determine the best and plausible time series model for each series, the help of some available Cran packages like forecast and so on are exploited for the pre-analysis part. As a well known fact that, to fit a copula model, residuals of each time series must be derived for further elaborations. The following Table 4.32 summarize the best plausible models for each time series.

Table 4.32: ARIMA Model summary for each Logreturns

Log-returns	**Model Fit**	**log-likelihood**	**LB [M-Li]**
ISE_{USD}	ARIMA(4,0,3)	1451.894	0.6822 [0*]
DAX	ARIMA(7,0,7)	1539.295	0.9333 [0*]
$FTSE$	ARIMA(3,0,3)	1514.276	0.658 [0*]
EM	ARIMA(2,0,2)	1465.914	0.4913 [0*]

In above Table 4.32, naive $ARIMA(p, d, q)$ models are determined for each series with Ljung-Box (LB) and McLeod-Li (M-Li) test results to derive the model residuals.

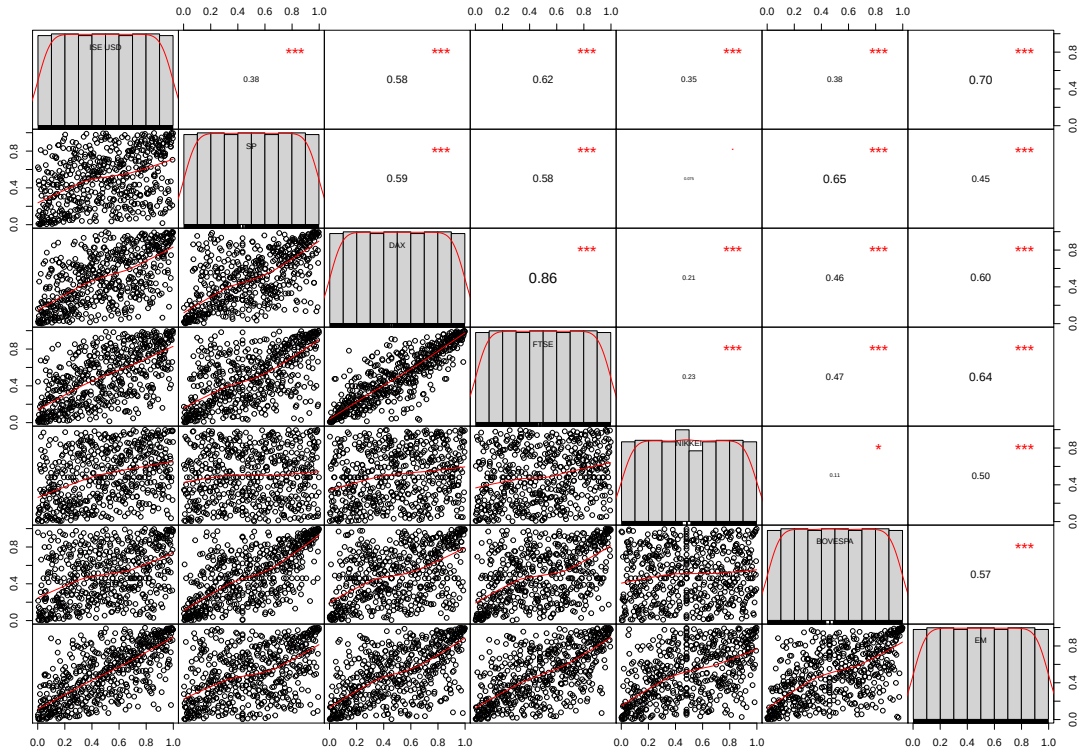


Figure 4.8: Dependence Structure among the log-returns with Kendall's τ values

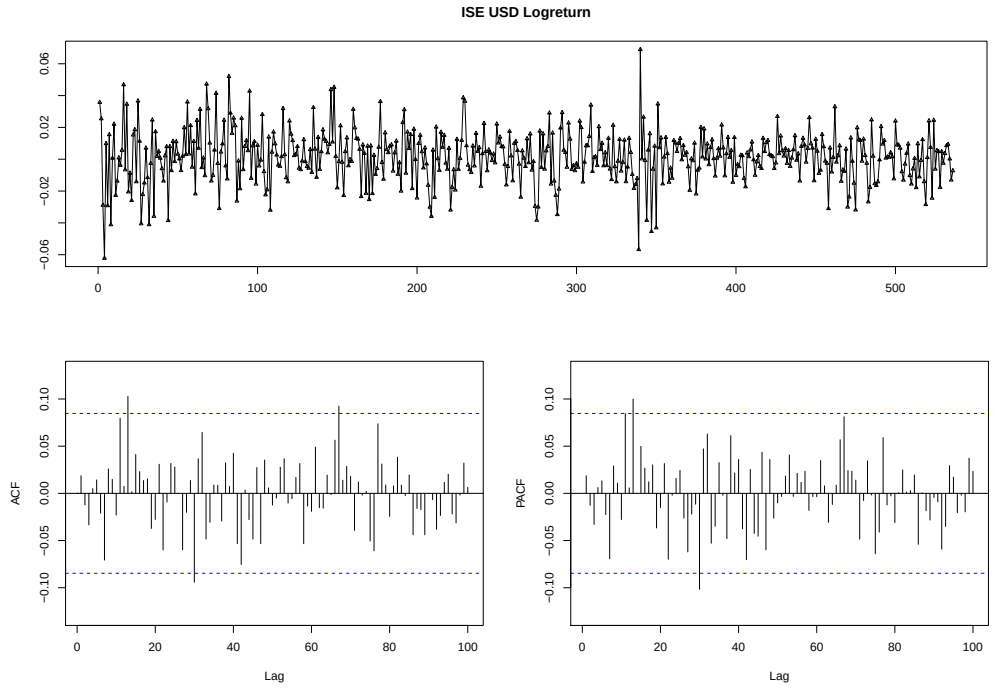


Figure 4.9: Time series of ISE_{USD} between 05-Jan-09/22-Febr-11 and its Acf-Pacf Plots

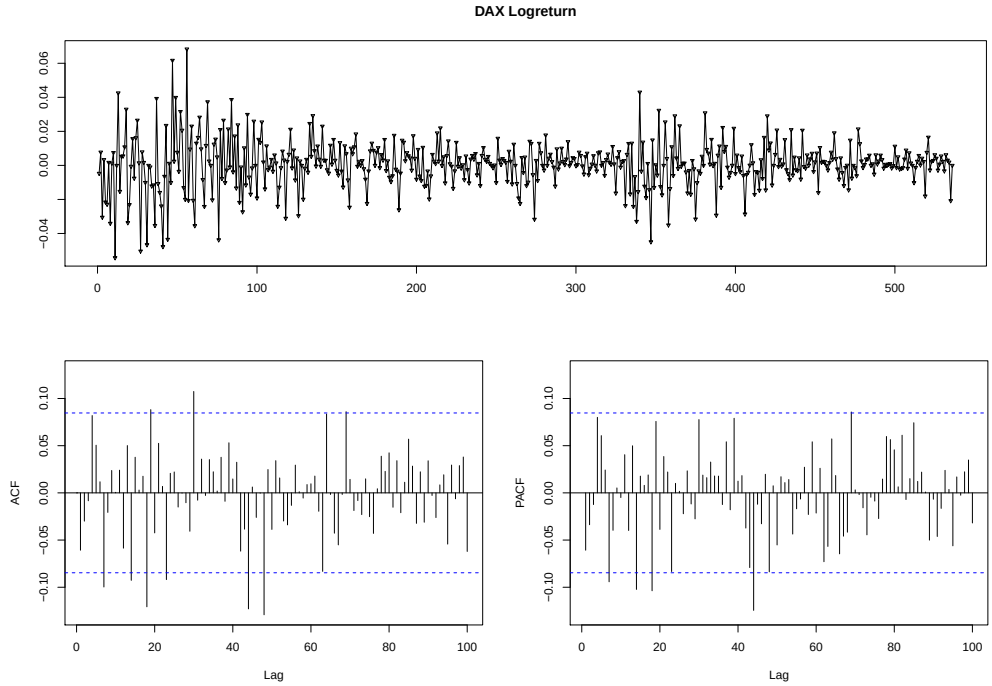


Figure 4.10: Time series of DAX between 05-Jan-09/22-Febr-11 and its Acf-Pacf Plots

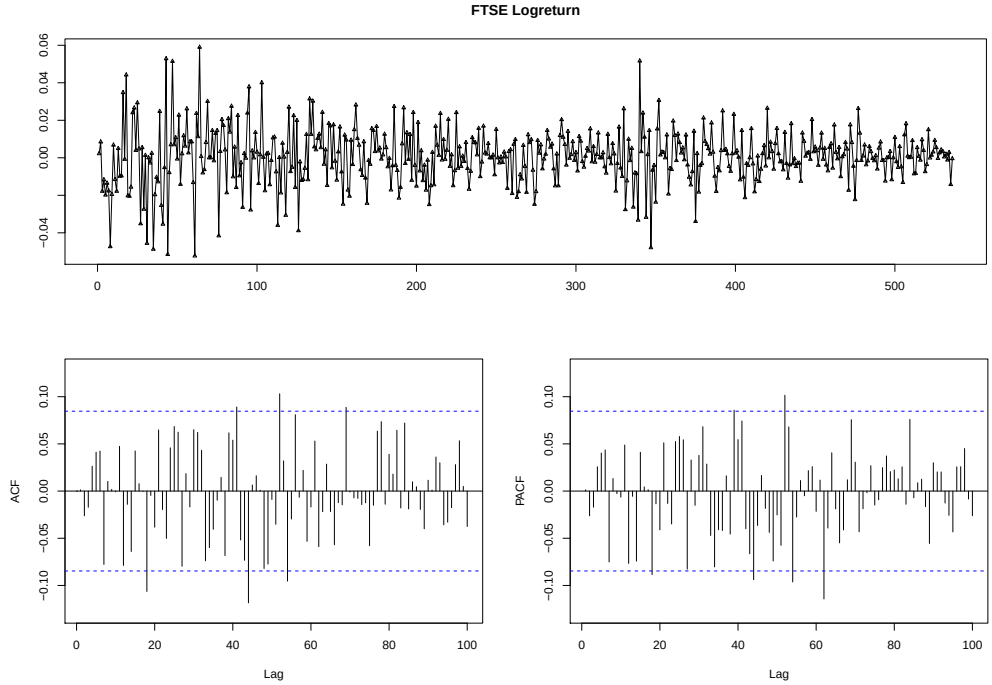


Figure 4.11: Time series of $FTSE$ between 05-Jan-09/22-Febr-11 and its Acf-Pacf Plots

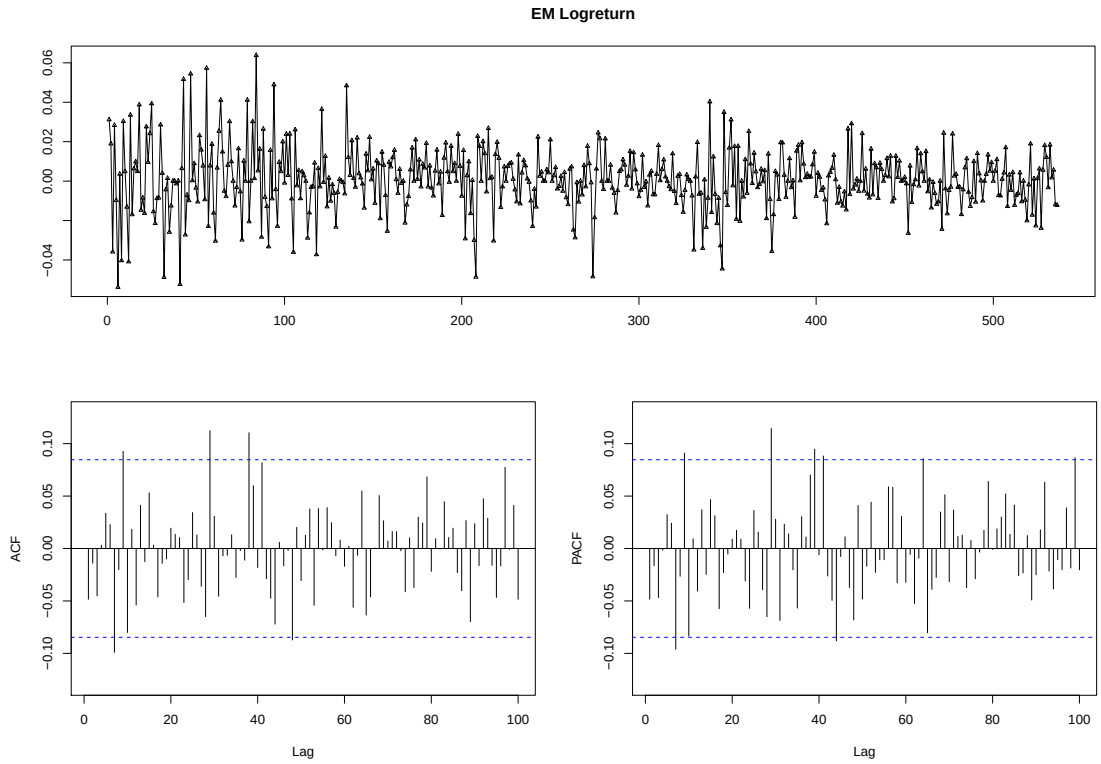


Figure 4.12: Time series of EM between 05-Jan-09/22-Febr-11 and its Acf-Pacf Plots

On the other hand, classical $ARIMA(p, d, q)$ models capture only best linear forecast without considering nonlinearity. For that reason, any possible model volatility must be detected and included in the model, with the help of GARCH models. This stepwise methodology results in checking residuals in detail and finding a suitable GARCH model, tabulated in Table 4.33. The subscripts in the Model fits identify the considered conditional distribution in GARCH modeling setup.

At this step, various GARCH(1,1) models with desired conditional distributions, such as normal, student-t and their skewed versions, are investigated. Furthermore, the Exponential and GJR type GARCH models (eGARCH and gjrGARCH) are also considered for the comparison with the classical version to determine the best model. Especially, the suitability of gjrGARCH models was previously mentioned in the study of [3] because of its power to capture positive and negative shocks on the conditional variance. The model having the largest likelihood and smaller information criteria values was decided for each univariate time series. Certainly, other diagnostic tests are considered to identify any evidence about the serial correlation, stability of the parameters and so on.

Table 4.33: ARCH/GARCH Model summary for each Log-returns

Log-returns	**Model Fit**	**log-likelihood**	**LB**	**WN**
ISE_{USD}	$gjrGARCH(1,1)_{sstd}$	1482.08	0.6822	0.1983
DAX	$GARCH(1,1)_{std}$	1631.04	0.9333	0.6727
$FTSE$	$GARCH(1,1)_{sstd}$	1571.069	0.658	0.8366
EM	$GARCH(1,1)_{std}$	1512.232	0.4913	0.3680

After the comparison, it has been observed that for each case, considered models work well enough to continue with the final model residuals. Furthermore, WN abbreviates the result of white noise test for the obtained residuals and the p-values of WN for each series are large enough. For a visual diagnostic for $GARCH(1,1)$ models of each series are given in Figures 4.13-4.16 below.

After completing time series modeling part, the obtained residuals are stored as inputs for the corresponding vine models. In this study, classical vine copulas ($M = 1$ case for the mixture model) and the mixture of vines ($M = 2$ case only) are fitted to the residual data. In this exploration, the existed dependence among the residuals

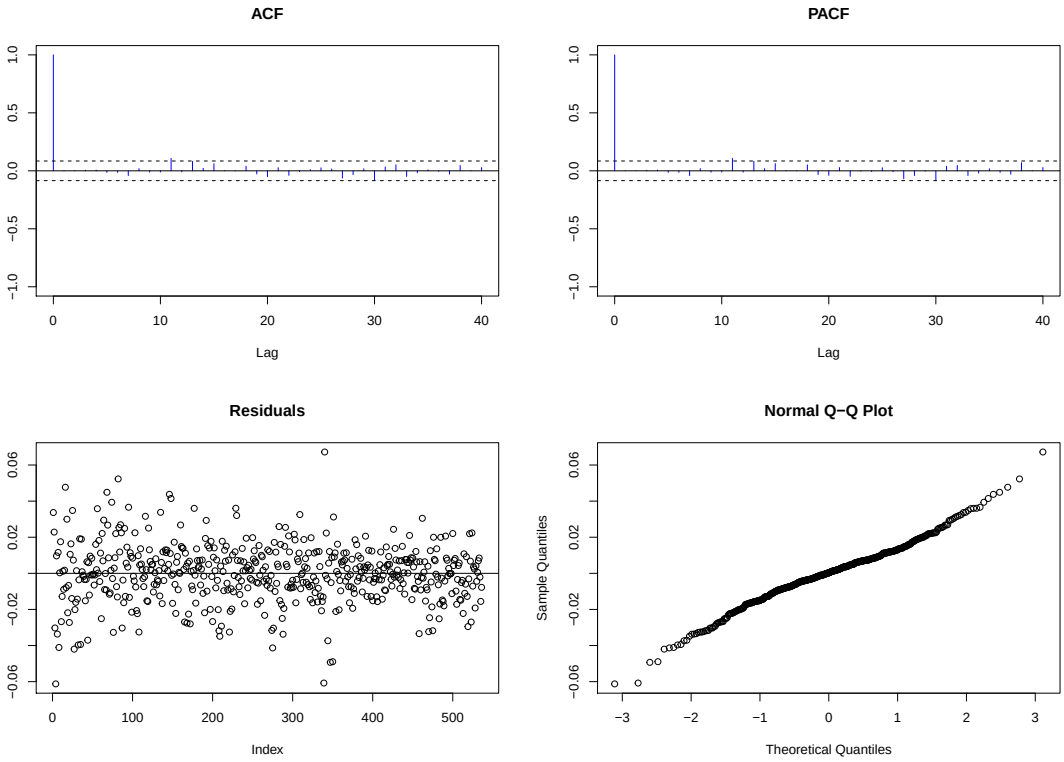


Figure 4.13: GARCH Model summary for the series ISE_{USD}

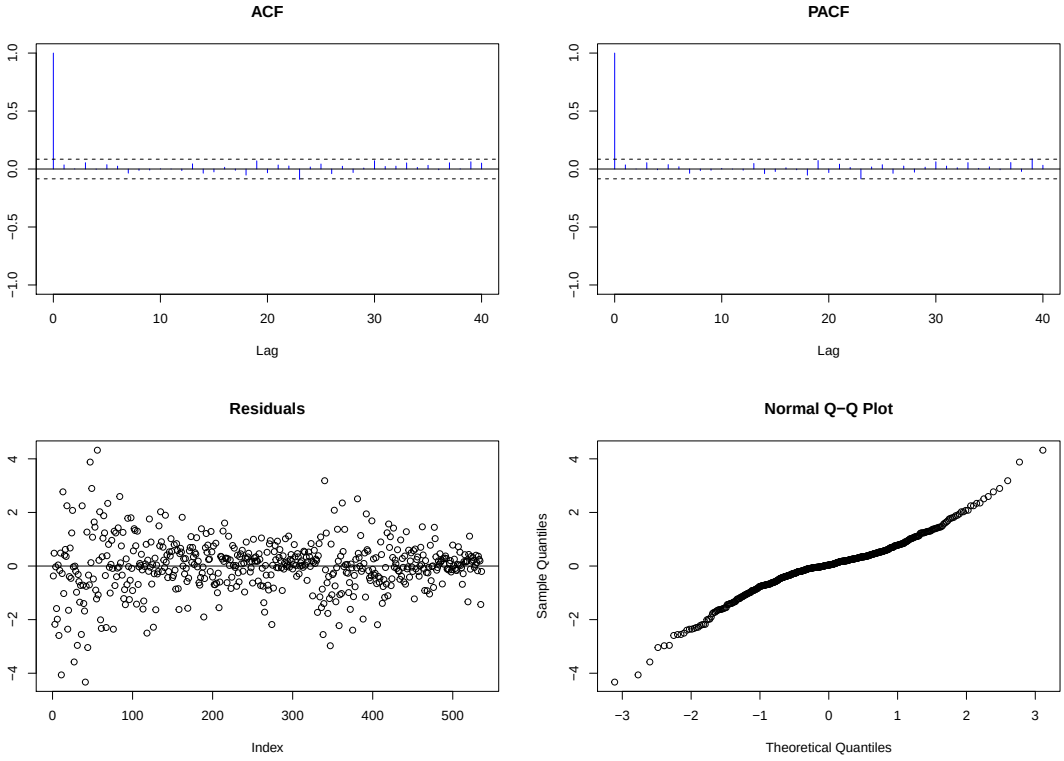


Figure 4.14: GARCH Model summary for the series DAX

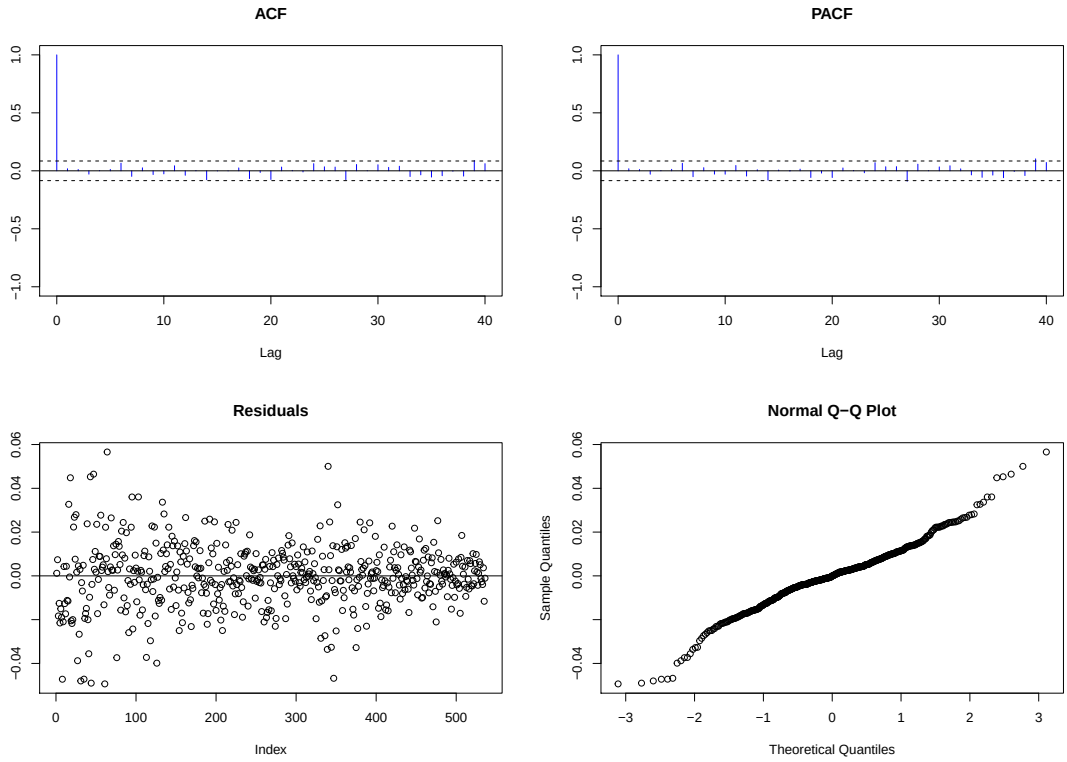


Figure 4.15: GARCH Model summary for the series $FTSE$

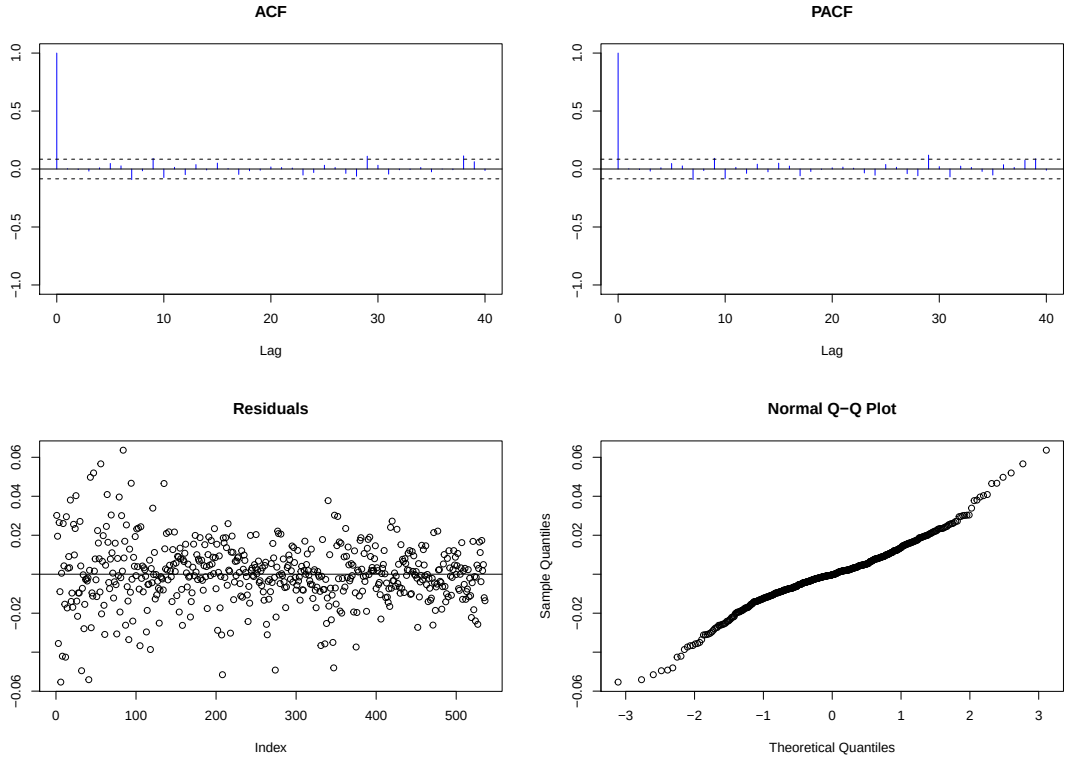


Figure 4.16: GARCH Model summary for the series EM

of the time series models can be visualized in Figure 4.17 below. Besides, when the residuals are converted into pseudo-observations (copula data), this association further illustrated with Figure 4.18 including Kendall's τ values.

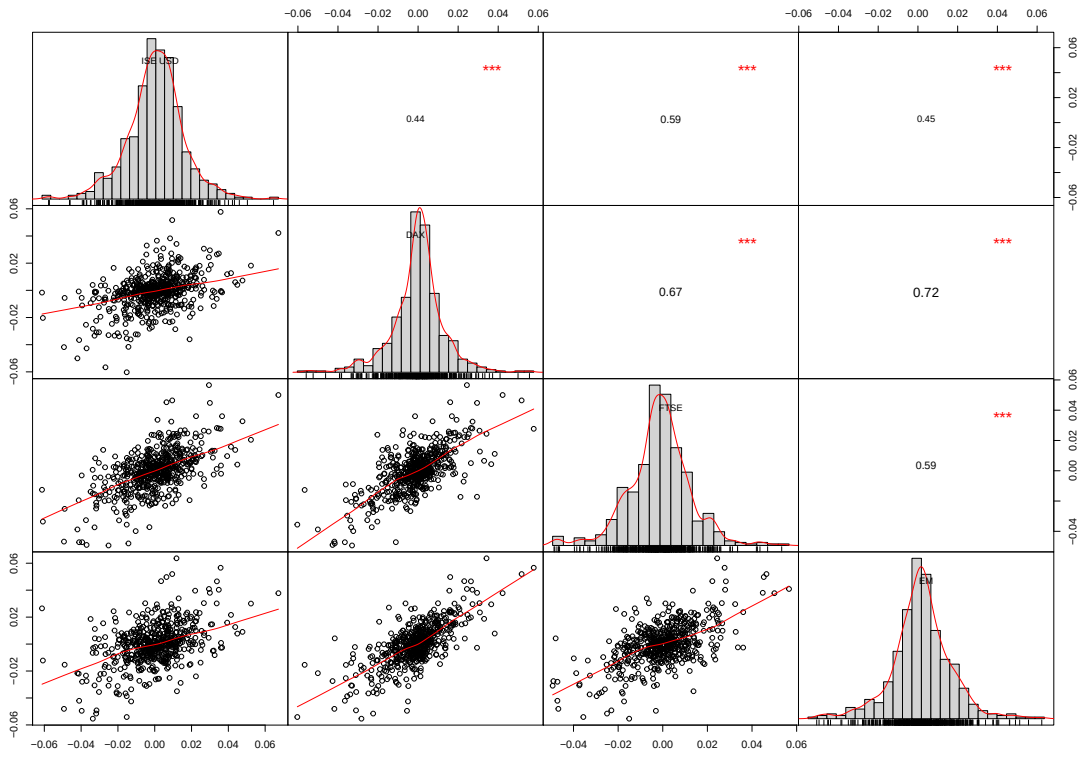


Figure 4.17: Dependence between the obtained residuals for each log-return series

For the obtained residuals, different vine models are investigated from $M = 1$ to $M = 2$ with only Frank-Frank scenario for the finite mixture model because of the symmetric dependence structure, visualized in Figure 4.18. For the general vine models, Normal and Frank copula families are considered in the model comparison part. Other one-tailed type copula families like Clayton or Gumbel is not suitable to model this dependence pattern since they are mainly based on positive correlation. In this framework, main interest is to identify the dependence structure between four log-return values with a root node variable, ISE_{USD}.

In this Table 4.34, Frank-Frank mixture model is compared with the one-component classical C-vine models. For the mixture ones, the initial guesses for the parameters are retrieved from the estimated parameter for CVFrank model. More clearly, initial parameters are $(\beta^1_{12} = 3, \beta^1_{13} = 4, \beta^1_{14} = 3, \beta^1_{23|1} = 4, \beta^1_{24|1} = 5, \beta^1_{34|12} = 2)$ and

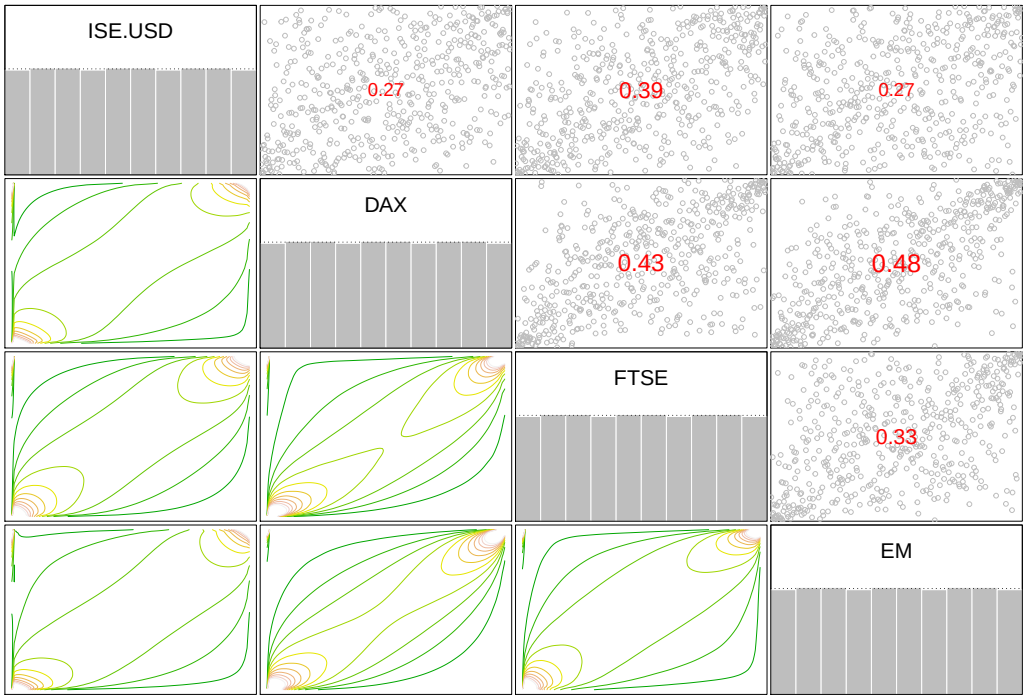


Figure 4.18: Dependence between the corresponding pseudo-observations belonging to the residuals

Table 4.34: Comparison of C-vine models including Frank-Frank mixture scenario

Vine Model	**Loglikelihood**
CVNormal	446.0360
CVFrank	409.4262
$Frank - Frank$	416.2312

$(\beta^2_{12} = 2, \beta^2_{13} = 3, \beta^2_{14} = 2, \beta^2_{23|1} = 3, \beta^2_{24|1} = 4, \beta^2_{34|12} = 1)$ for 1'st and 2'nd component, respectively. Based on the log-likelihood values, the obtained mixture is superior than CVFrank but it is as not plausible as CVNormal model. In this modeling setup, as a best model, CVNormal shows that there exists a weaker dependence in the center and stonger dependence at tails when it is compared to both CVFrank and Frank-Frank mixture case.

4.2.2 CD-vine mixture of Log-returns

For the second application, Daily Closing Prices of major indices given in Table 4.35 between years 1991-1998 are considered. The data set is retrieved from the repository of CRAN data sets package and does not include weekends and holidays, similar to previous data set [30]. For the implementation of 3 component 4-dimensional CD-vine mixture model, last three years, i.e. 1995-1997 were focused from the complete data set. The brief definitions for the indices are tabulated in Table 4.35 with their time series plots between 1995-1997 in Figure 4.19.

Table 4.35: Log-return series definition

Stock Index	Definition
DAX	Germany Stock market return index
$FTSE$	UK Stock market return index
SMI	Swiss Market Index
CAC	France CAC 40 Stock Market Index

As a first insight, the summary statistics of the whole three years and individually each separate year was given below in Table 4.36. Furthermore,for each year, the dependence between the selected log-return values is visualized in Figures 4.20, 4.21 and 4.22, that motivates the background of CD-vine mixture model implementation.

In this CD-vine modeling setup, the dependence structure for each year, equivalently each component requires different C-vines. The overall procedure for the time series analysis and CD-vine mixture fitting part was summarized in Appendix part with Figures B.1 and B.2, respectively. Thereafter, each univariate time series have been modeled similar to the first application above. After checking stationarity using available tests, $ARIMA(p, d, q)$ models have been tested with suitable parameters.

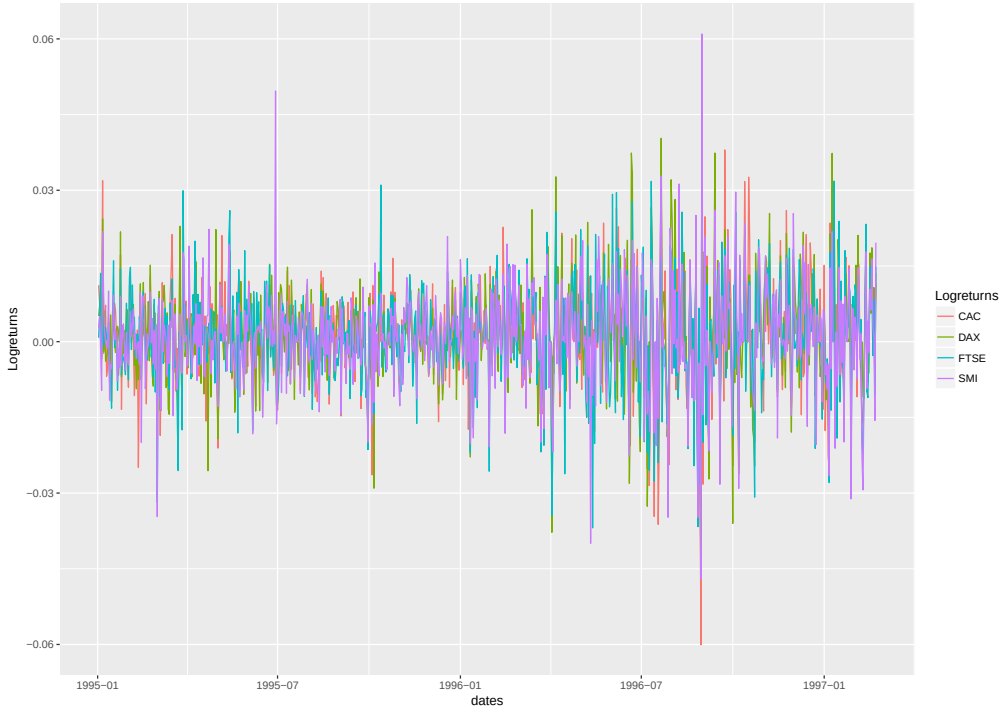


Figure 4.19: Univariate log-returns between 1995-01-02 and 1997-12-31

Table 4.36: Summary statistics for Log-return series, for years 1995-1997 and each three year separately (CV1 for 1995, CV2 for 1996 and CV3 for 1997)

Stock Index	**Min.**	**Max.**	**Mean**	**Variance**	**Skewness**	**Kurtosis**
DAX	-0.060068	0.043207	0.001328	1.204847e-04	-0.2800984	2.44286
SMI	-0.046951	0.049680	0.001335	9.530134e-05	-0.2045915	2.6556
CAC	-0.0436528	0.0609773	0.0010594	1.229722e-04	-0.04038411	2.013503
$FTSE$	-0.0310272	0.0312507	0.0007407	6.00198e-05	-0.1434858	1.355695
DAX_{CV1}	-0.0318230	0.0243396	0.0007990	5.347527e-05	-0.2151643	1.684495
SMI_{CV1}	-0.025545	0.049680	0.001062	5.302119e-05	0.8042204	7.620212
CAC_{CV1}	-0.0346849	0.0319273	0.0005016	8.394722e-05	0.03933226	1.221863
$FTSE_{CV1}$	-0.0144040	0.0217781	0.0004479	3.301096e-05	0.07156261	0.3041095
DAX_{CV2}	-0.037787	0.032663	0.001528	8.204311e-05	-0.317721	1.670497
SMI_{CV2}	-0.034389	0.031049	0.001596	7.136043e-05	-0.5082251	2.291705
CAC_{CV2}	-0.0399492	0.0295422	0.0012602	1.065269e-04	-0.4598755	1.401845
$FTSE_{CV2}$	-0.0220118	0.0265099	0.0009132	4.26914e-05	-0.2069626	1.015416
DAX_{CV3}	-0.060068	0.043207	0.001657	2.264325e-04	-0.2886622	0.792924
SMI_{CV3}	-0.0469512	0.0370533	0.0013475	1.621124e-04	-0.2886056	0.7564879
CAC_{CV3}	-0.0436528	0.0609773	0.0014163	1.789076e-04	0.08112503	1.648378
$FTSE_{CV3}$	-0.0310272	0.0312507	0.0008611	1.046882e-04	-0.1752442	0.3311307

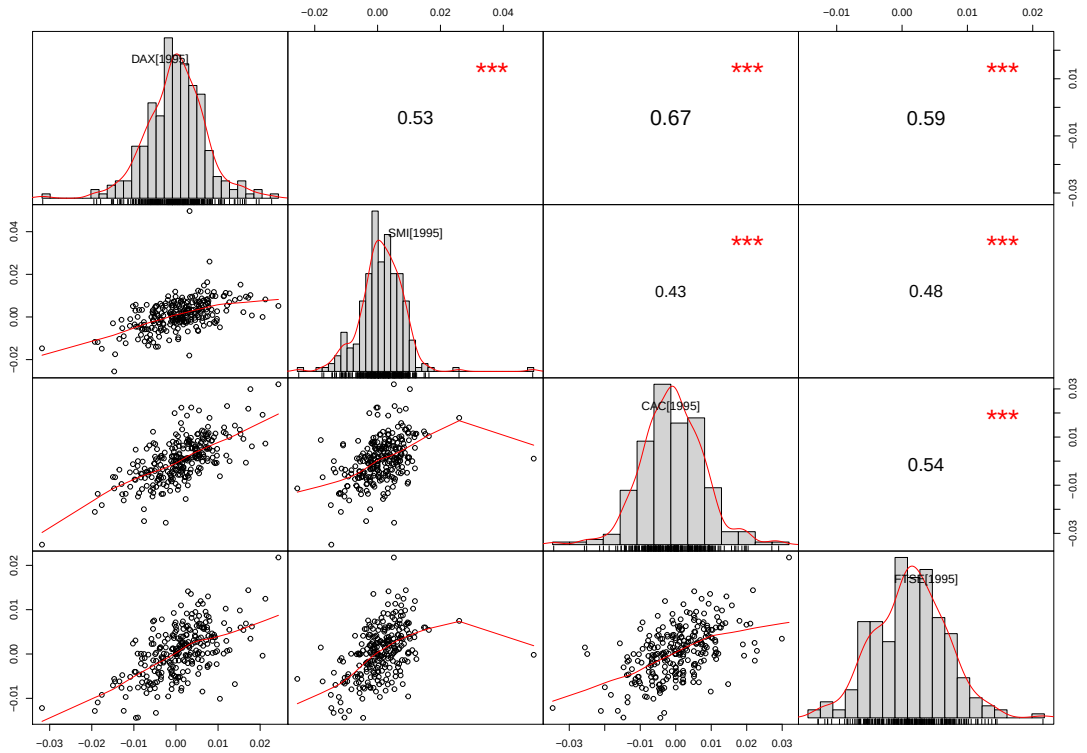


Figure 4.20: Dependence Structure among the log-returns with Kendall's τ values for year 1995

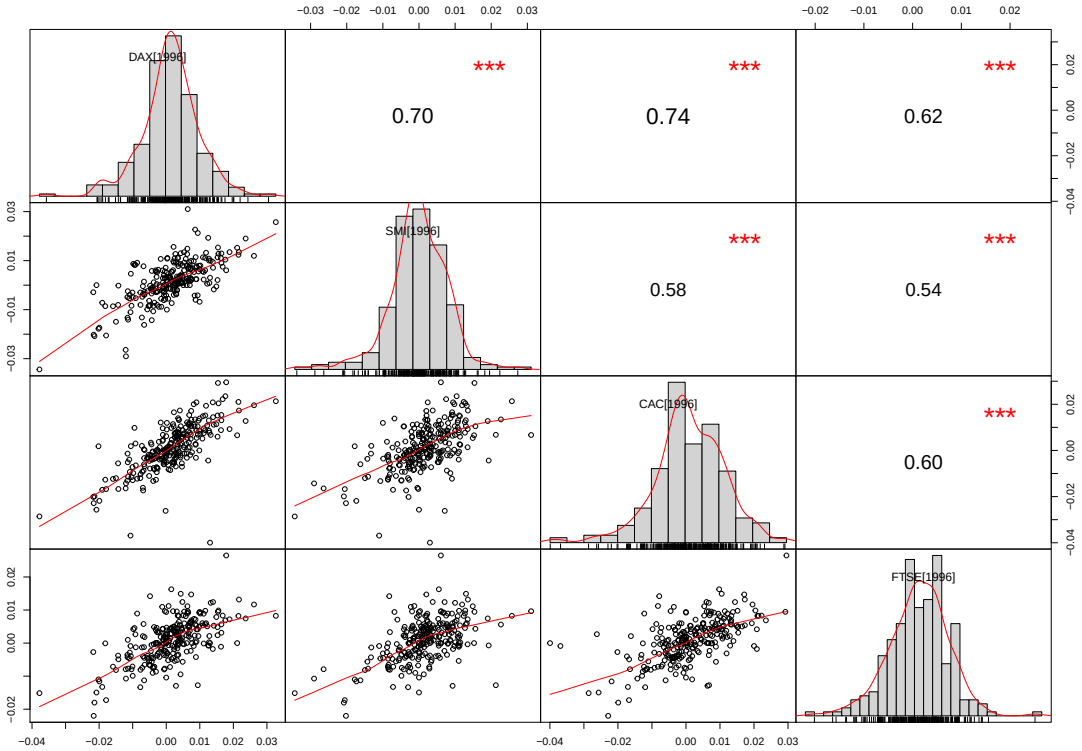


Figure 4.21: Dependence Structure among the log-returns with Kendall's τ values for year 1996

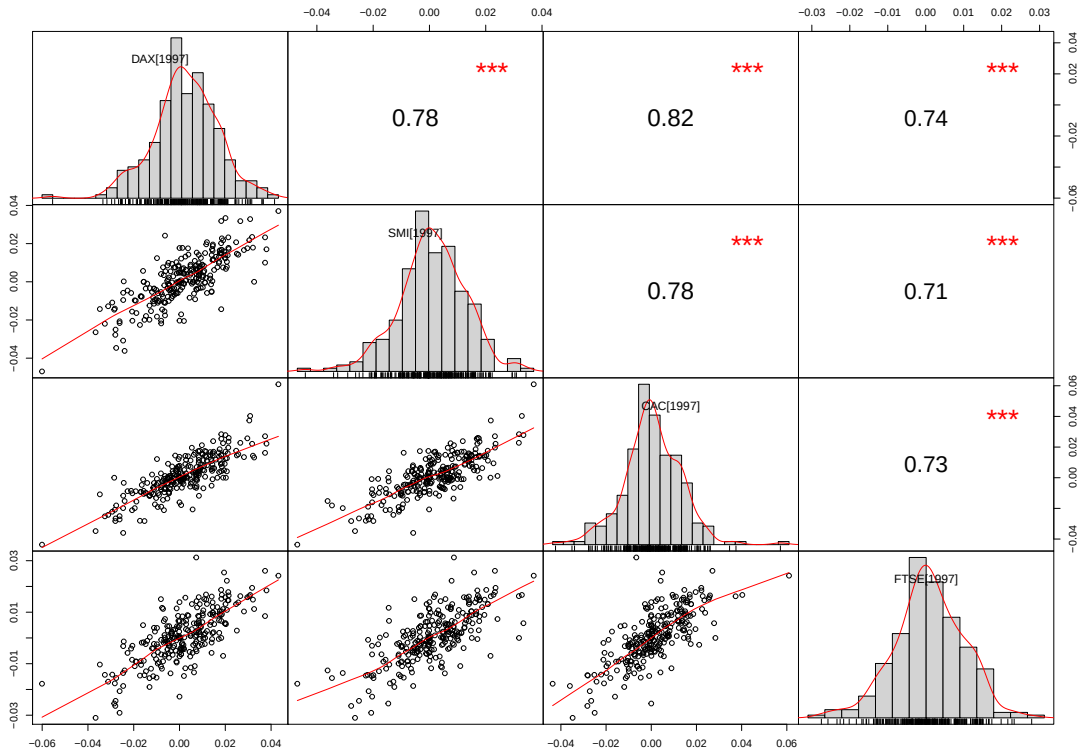


Figure 4.22: Dependence Structure among the log-returns with Kendall's τ values for year 1997

Afterwords, if there exist any necessity, various $GARCH(1,1)$ models are considered with different distributions. For the sake of simplicity, only standard $GARCH(1,1)$ model has been tested over the residuals of the $ARIMA(p,d,q)$ models, if it is required. The model summary for each series belonging to each component is summarized in Table 4.37.

Table 4.37: ARIMA Model summary for each Logreturns for each component

Log-returns	Model Fit	log-likelihood	LB [M-Li]
DAX_{CV1}	ARIMA(1,0,1) $\rightarrow GARCH(1,1)_{std}$	922.4105	0.485 [0.6581]
SMI_{CV1}	ARIMA(1,0,1) $\rightarrow GARCH(1,1)_{std}$	937.9006	0.6262 [0.9997]
CAC_{CV1}	ARIMA(1,0,0) $\rightarrow GARCH(1,1)_{sstd}$	861.8276	0.8368 [0.8005]
$FTSE_{CV1}$	ARIMA(2,0,2) $\rightarrow GARCH(1,1)_{std}$	980.7856	0.3024 [0.4483]
DAX_{CV2}	ARIMA(4,0,0) $\rightarrow GARCH(1,1)_{sstd}$	875.944	0.6678 [0^*]
SMI_{CV2}	ARIMA(2,1,2) $\rightarrow GARCH(1,1)_{std}$	890.4312	0.4692 [0.3076]
CAC_{CV2}	ARIMA(2,0,0) $\rightarrow GARCH(1,1)_{sstd}$	835.8359	0.2972 [0.0107^*]
$FTSE_{CV2}$	ARIMA(0,0,0) $\rightarrow GARCH(1,1)_{sstd}$	948.8898	0.5206 [0.9545]
DAX_{CV3}	ARIMA(9,0,1) $\rightarrow GARCH(1,1)_{snorm}$	738.8651	0.2092 [0.029^*]
SMI_{CV3}	ARIMA(2,0,0) $\rightarrow GARCH(1,1)_{snorm}$	779.1635	0.135 [0^*]
CAC_{CV3}	ARIMA(5,0,0) $\rightarrow GARCH(1,1)_{sstd}$	766.9586	0.5108 [0^*]
$FTSE_{CV3}$	ARIMA(0,0,1) $\rightarrow GARCH(1,1)_{std}$	832.1742	0.3364 [0.7369]

In a similar manner, for each univariate series, the classical Box-Jenkins models have been considered with (LB:Ljung-Box, M-Li:McLeod-Li) test results. In some cases, there exist an arch effect, indicated by the value of McLeod-Li test result, so then various $GARCH(1,1)$ models have been compared to determine any further accuracy in the model. In Table 4.37, these procedure was given together for a related log-returns. Generally, the result of $GARCH(1,1)$ model was quite well for various series except DAX_{CV2}, CAC_{CV2}, DAX_{CV3}, SMI_{CV3} and CAC_{CV3}. On the other hand, modeling residuals with a $GARCH(1,1)$ model is still reasonable when the p-values of McLeod-Li test for the original $ARIMA(p,d,q)$ and $GARCH(1,1)$ models are compared. Finally, residuals, obtained from above two step modeling part, are captured for the CD-vine mixture model. Now, the dependence among log-returns in terms of their transformed residuals within each year is identified in terms of Kendall τ values, graphically visualized in Figures 4.23, 4.24 and 4.25.

For each year, the dependence structure seems to be two-tailed and there is no certain accumulated points at the lower and upper tails. For this reason, only Frank copula pairs are considered in the first step of the CD-vine model. For the C-vine construction,

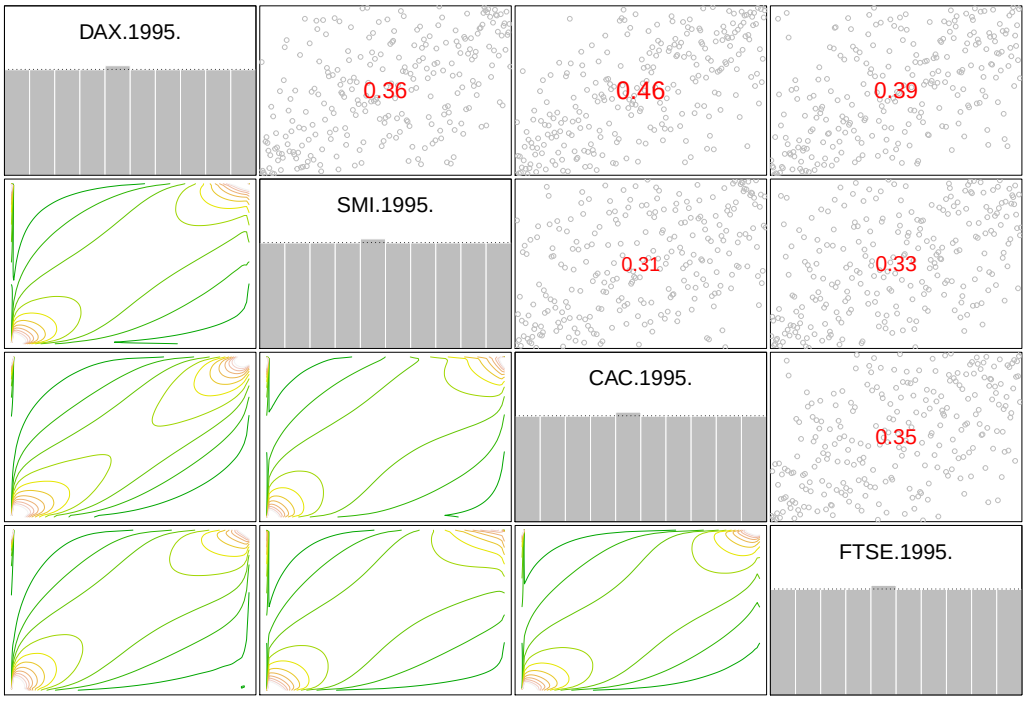


Figure 4.23: Dependence structure among the transformed residuals in 1995

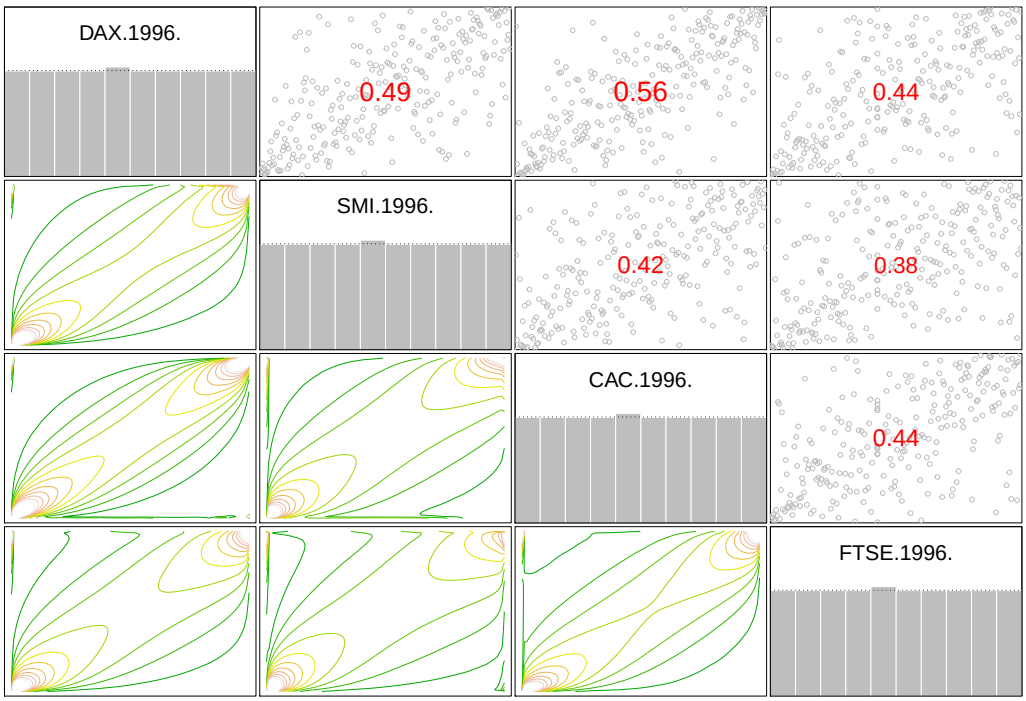


Figure 4.24: Dependence structure among the transformed residuals in 1996

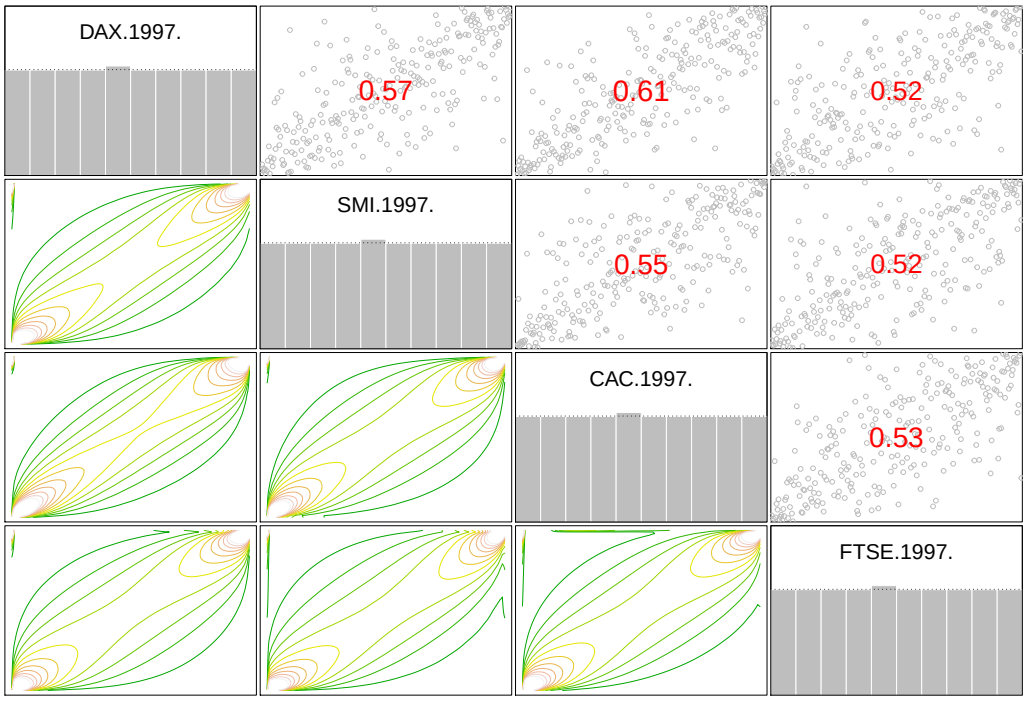


Figure 4.25: Dependence structure among the transformed residuals in 1997

based on the above Figures 4.23-4.25, DAX log-return seems to be the best choice for being a root node in terms of Kendall τ values. For the next step of CD-vine mixture model, the following flowchart visualizes the upcoming calculations in Figure 4.26.

For the sake of simplicity, same copula function has been considered for all components. Thereafter, the calculated values of $F(cv_1)$, $F(cv_2)$ and $F(cv_3)$ based on the obtained Frank models with parameters are incorporated into D-vine model. For D-vine model, other available copulas are considered to model each pair with a different family. The details of the fitted model with their loglikelihood (Log-lh) values, dependence parameters and corresponding Kendall τ values are tabulated below. Furthermore, tree structure of $F(cv_1)$, $F(cv_2)$ and $F(cv_3)$ based on the selected copula family can be visualized in Figure 4.27.

Based on the results in Table 4.38, for each year, the considered stock values have symmetric dependence pattern based on Frank copula family. Especially, for unconditional densities, the dependence parameter values are very high. For instance, ordered numbers $7.28 - 8.20 - 6.22$ for Model $CV3$ express that DAX_{CV3} log-return is

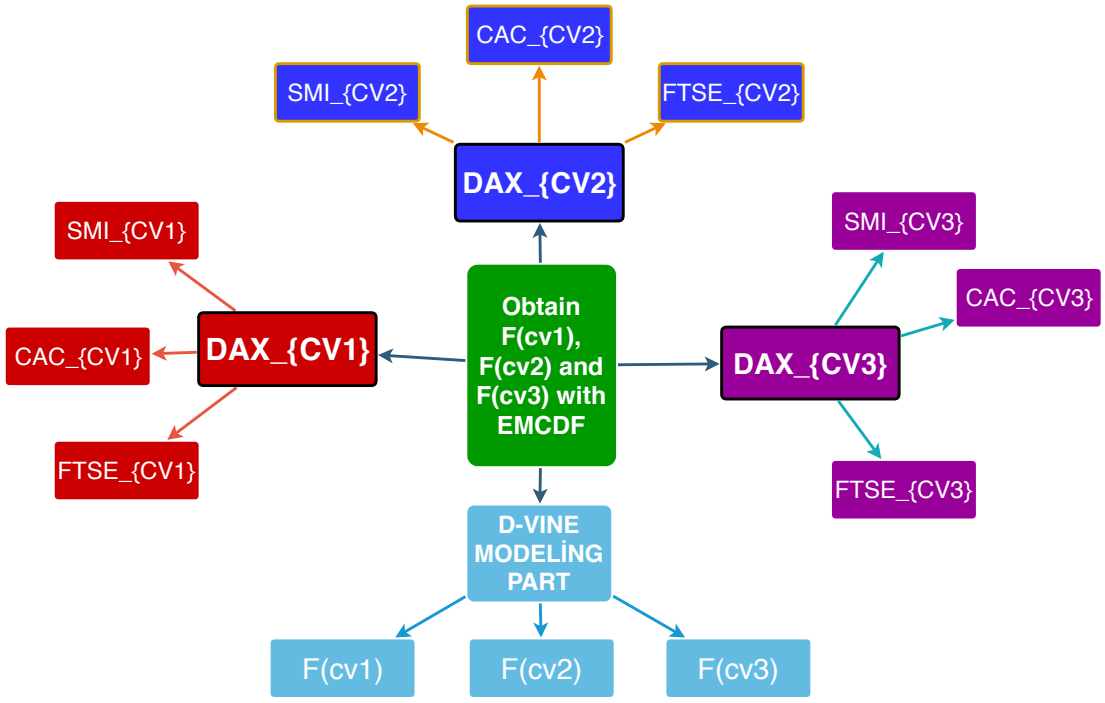


Figure 4.26: CD-vine mixture model Diagram for the residual data

Table 4.38: Comparison of C-vine models for each component

Model for $CV1$	**Log-lh**	**Parameters [Kendall** τ**]**
Frank	176.1582	3.76 [0.37]-5.12 [0.46]-3.99 [0.39]-1.02 [0.11]-1.66 [0.18]-1.23 [0.13]
Model for $CV2$		
Frank	268.9389	5.65 [0.50]-7.07 [0.57]-4.78 [0.44]-1.11 [0.12]-1.50 [0.16]-1.83 [0.20]
Model for $CV3$		
Frank	386.019	7.28 [0.57]-8.20 [0.61]-6.22 [0.53]-2.65 [0.28]-2.56 [0.27]-1.72 [0.19]
Model for DV		
Gumbel-Gumbel-Survival Gumbel	1201.872	13.54 [0.93]-16.88 [0.94]-1.20 [0.17]

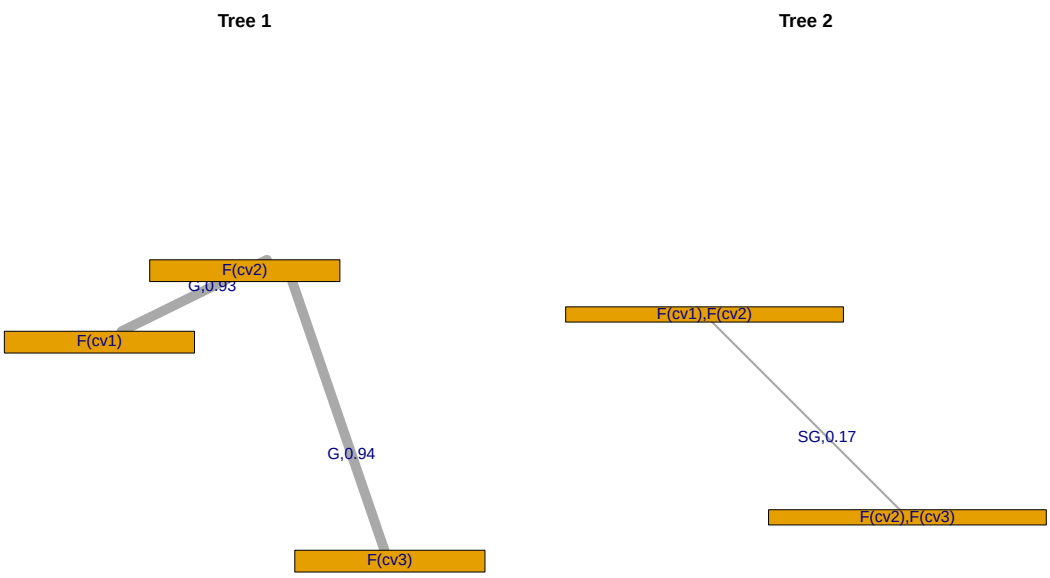


Figure 4.27: D-vine model for $F(cv_1)$, $F(cv_2)$ and $F(cv_3)$

positively dependent with other stock indices, namely SMI_{CV3}, CAC_{CV3} and $FTSE_{CV3}$, in 1997. This dependence pattern is almost similar for the previous years and conditional density functions have smaller parameter values compared to the parameters of unconditional ones.

In D-vine part, the dependence among the years is described with one-tailed copula families. In Figure 4.27, G and SG stands for Gumbel and Survival Gumbel family with their empirical Kendall τ values. For instance, the dependence among years 1995 and 1996 with the values of $F(cv_1)$, $F(cv_2)$ and $F(cv_3)$ is modeled by Gumbel family, which displays the upper tail dependence for these years. On the other hand, the dependence among years 1995 and 1997, conditioned on 1996 exhibits lower tail dependence ($\tau = 0.17$) based on Survival Gumbel family.

CHAPTER 5

CONCLUSION AND OUTLOOK

In this last chapter, the whole study is briefly summarized and discussed in detail from diverse perspectives. Firstly, the main results are summarized and the limitations of the study have been highlighted. Thereafter, the rise of open problems from the proposed mixture models over vine copulas is described. Finally, the schedule for future studies is explained with possible solutions to the open problems of the study.

5.1 Main Results

In this dissertation, the various mixture of vines are proposed and their performances are tested by using both simulated and real life data. Within the finite mixture context, firstly the results of [22] are improved for C-vine model with numerous tail dependence structures. For that purpose, as a departure point, 2 and 3 component, 3 dimensional C-vine mixtures are discussed including same and distinct copula pairs. Previously, it has been proven that whenever the same copula family is selected to model the pairs belonging to the components, parameter estimate accuracy is plausible enough based on EM algorithm and the classical optimization tools. However, different pair copulas require much more attention in terms of model comparison and parameter estimates. Additionally, the same mixture approach is extended to the 4-dimensional space under C-vine framework and the parameter estimation part is completed by using elegant derivative free optimization algorithms. Afterwords, C- and D-vine models are considered two components and they are combined again using finite mixture model. In these mixture models, only Frank family is considered with various

dependence patterns. Finally, CD-vine mixture, another fruitful contribution of the study is discussed with specific scenarios both using simulated and real life data set. In this modeling setup, CD-vine mixture model requires two step maximization process for the parameter estimation part. Eventually, in the application part, the main focus is the dependence among the several major stock indices within the mixture of vines. For that purpose, for considering two different applications, two financial data sets have been investigated.

5.2 Discussions for findings

Apparently, this is the first study in the literature for making contribution on the selection of different copula pairs for the construction of C-vine models. Furthermore, within the same methodology, both C- and D-vine models are combined in terms of finite mixture framework and this idea can be generalized for an arbitrary R-vine model directly. As a novel contribution, the dependence among the components, which are modeled by C-vine copula, has been considered in terms of D-vine copula for the first time two specific scenarios. In this subsection, all the experienced results are summarized and the recent findings of this thesis are examined step by step.

In the case of 3-dimensional mixture data over 2 or 3 components, numerous conclusions can be drawn having both advantages and drawbacks. Firstly, whenever the choice of copula family is different for each component, mixture model setup has failed in terms of parameter estimation under considered optimization tools. Furthermore, classical information criteria values selected to use in this study, are not sufficient enough to decide which model is the best one. Furthermore, as a second conclusion DEoptim function results are very promising for Frank-Frank mixture case in 4-dimension rather than other considered functions. In another proposed finite mixture, when the structure of two components are different, C- and D-vine for instance, the model identification problem has emerged in terms of the multivariate data. In this C- and D- vine mixture, the fitted model capture the dependence structure of only one of the components.

CD-vine approach has been studied under different dependence assumptions belonging to variables and components. The result of independence case is very satisfactory in terms of accuracy of the parameter estimates. For this case, Clayton-Independence CD-vine mixture was tested and the parameter estimations are close enough to the true parameters decided in the simulation part. On the other hand, whenever the dependence exists among the components via multivariate distribution functions, it is not possible to detect which copula pairs are suitable in D-vine part.

Additionally, the proposed ClarkeMixV and VuongMixV GOF tests have been tested only for model comparison of 2 component 3 dimensional mixture with Clayton-Joe pairs. In these test results, ClarkeMixV has proved that the compared models are statistically different from each other successfully. On the other hand, VuongMixV test statistics can not distinguish two models in terms of which one is more preferable than the others.

For the first application, finite C-vine mixture model was considered with other copula pairs in 4-dimensional case. The model comparison part exhibited that the mixture model is better than the classical 1 component C-vine case with Frank copula pairs. However, the classical C-vine with Normal pairs seems to be better than Frank based models because of its tail dependence pattern. This selected model exhibits weaker dependence in the center and stronger dependence at tails when compared with mixture of Frank copulas. This modeling setup is very useful to understand the co-movements of the considered stocks in the market.

In the second application, CD-vine mixture model has been implemented to another set of log-returns to identify their dependence structure by considering temporal dependence among different years. Firstly, data set is partitioned into 3 subsets with equal size for years 1995, 1996 and 1997. For each year, Frank copula pairs are considered for the construction of C-vine part with DAX log-return as a root node as a result of pre-analysis. Thereafter, obtained F_{CV_1}, F_{CV_2} and F_{CV_3} are modeled with D-vine including various copula families to identify their correct dependence structure. In this application, for each year there exists a significant dependence among the selected log-returns. Furthermore, one tailed copulas in D-vine part have certain clues about the stock values.

5.3 Future Research Directions

Certainly, the obtained results for all simulated data set and the application part has not been completed. There are lots of new research questions, that come from the above discussed vine mixture models. Especially, such mixtures can be generalized for various copula families instead of restricting our interest for only some archimedean copulas. Besides, the parameter estimation part deserves more attention in terms of the selection of initial guesses during the maximization process, the choice of the optimization tools including their suitable parameters and the implementation of the gradient information for the objective function to enhance the derived parameters. Apart from above mentioned improvements, the likelihood construction for the mixture models might be required other tools like the variants of EM algorithm at the beginning. The corresponding open problems about these subtitles are highlighted with the future research schedules below.

5.3.1 For mixture models

In this study, there are various copula families implemented for the mixture models with especially strong dependence assumption. Clearly, the number of combinations and the selection of copula families can be improved to work on numerous mixture models. Besides, instead of only considering some archimedean families, it is possible to extend the same mixture model procedure for elliptical families and so on. Furthermore, the parameters of selected copula families might take negative values to exhibit negative dependence under Frank or rotated versions of other archimedean copulas.

Another further study can be developed in terms of the pair copula families within the considered vine tree structure. In this study, each component has been modeled by using same copula families, but, this mixture methodology can be applicable in case of selecting distinct copula pairs to model the dependence among the variables. For instance, the dependence for the pairs (u_1, u_2), (u_1, u_3) can be modeled via Clayton and Gumbel family while the conditional density for the pair $(u_2, u_3|u_1)$ can be constructed with a Frank copula for a C-vine density of the component. Certainly, this kind of selection increase the flexibility of the proposed mixture models, but the

parameter estimation part requires more elegant tools to get rid of unexpected failures.

C- and D-vine finite mixture model comparison has been tested only in terms of Frank copula pairs that there should be further improvements of such mixtures. Besides, this mixture model can be developed for arbitrary R-vine copula as well, of course with a high demand for caution on the parameter estimation. In the proposed CD-vine mixture model, other copula families and the selection of different families for each bivariate copulas within the vine density is also possible, one of the planned further research for the authors.

Apart from above mentioned structural improvements, another modeling setup can be considered in terms of mixing proportions. In all studied scenarios, weight parameters are considered as a fixed value, not changing with time. On the other hand, these mixing proportions can be examined on some covariate information and this is another way of considering time varying copula under finite mixture model. This model extension idea was presented in empirical findings of the study of [3] in bivariate case, but not for mixture of vines. This issued problem is another further plan of the authors to investigate time varying copula methodology for mixture of vines, especially for C-vines.

5.3.2 About Parameter Estimations

For the maximization process of log-likelihood function, various derivative free optimization tools have been investigated. However, there are still many alternatives to increase the accuracy of the estimated parameters for the proposed mixtures with the help of more distinguished optimization algorithm. This concern emerges in all of the scenarios especially when the dimension has been increased and the number of components and the complexity is boosted. Another recently issued open problem for this study is finding more accurate parameter estimation under vine mixture models.

For the parameter accuracy, the gradient information for the objective functions can be implemented. Especially, in 3-dimensional mixture models, these analytical derivatives might have beneficial to increase the accuracy with some extra computational effort. However, meanwhile, those gradient functions have certain computational drawbacks

under such available optimization routines. For that reason, both analytical and numerical gradient information should be elaborated carefully before considering any optimization tool. For higher dimensions, suitably implemented derivative free algorithms can still have various advantages.

Within all the applied algorithms, the initial guesses for the parameter estimation part has been considered as a fixed number. Generally, the sensitivity of the estimation part is directly related to the choice of initials and number of initial guesses. For that purpose, above studied optimization algorithms can be developed with random initials and multiple starting points to enhance the final parameter estimation. Besides, there are several methods to optimize an objective function in a global sense with the set of initial values, such as Particle Swarm Optimization (PSO) and Artificial Bee Colony (ABC). Such approaches will be studied in the near future to increase to accuracy of the parameter estimates especially when there is no available analytic gradient function.

Another contribution will be the hybridization of two or more available optimization methods for the log-likelihood function of mixture models. Similar to the idea of hybrid forecasting tools, parameter estimates of various implemented methods can be combined with certain weights. For instance, the combination of the parameter estimates belonging to two optimization methods, which have various features, can enhance the estimation accuracy.

Overall, the asymptotic behavior of the parameters is another open problem for this study. Furthermore, from the statistical point of view, consistency and robustness of the estimations deserve more attention for further analysis. In that respect, the model diagnostic part will be considered urgently with all other improvements.

5.3.3 Log-Likelihood Construction

The EM algorithm approach for the likelihood of mixture function is borrowed from the existing literature. Besides, this framework is one of the best approaches in such problems whenever the data set is not directly observed. On the other hand, there are various variants of the EM algorithm, which can be implemented within the problem

of this thesis study. Another open problem for the mixture model methodology emerges from the selection of the method for the construction of log-likelihood function of such mixture densities.

More clearly, other deterministic and stochastic variants of classical EM algorithm could be considered to speed up the algorithm and gain increase on the rate of convergence. In the literature, there are various kinds of versions for EM algorithm, such as Classification EM (CEM), Accelerated EM (AEM), Expectation Conditional Maximization (ECM) and so on to cope with the convergence speed problem of the original one. Furthermore, stochastic variants are more related to other limitations of EM approach, like getting trapped in local maximum value [32]. Those stochastic methods like Stochastic EM (SEM), Data Augmentation algorithm (DA) and Monte Carlo EM (MCEM) allow us to overcome the problems of untruly finding only local maximum for the log-likelihood function in this mixture model framework.

5.3.4 Model Selection

In model comparison and selection context, classical information criteria values and new GOF tests are implemented in this study. However, the success of classical model information criteria values and the modified GOF tests is not so promising. Especially, when the copula pairs are differed while modeling the components of mixture model, based on the values of AIC, BIC and CAIC, there exists a model identification problem. Furthermore, the modified Vuong test, VuongMix, can not decide which model is more preferable. Indeed, model information criteria values have tendency to select less complex model whenever the multivariate data exhibits complex dependence pattern. The only plausible model selection result is reached for the mixture of C-vines with Frank-Frank pairs in both 3 and 4-dimensional cases. Besides, some copula pairs are confused whenever they have similar or same dependence patterns.

In this sense, the model identification based on a suitable selection tool is required for mixture of vines. The above-mentioned problem has been stated by [22] in a similar manner. One of the main drawback of this study is that the lack of suitable of GOF tests for comparing such mixture of vines. Additionally, even if the classical

AIC, BIC and CAIC values are enough to identify finite mixture model involved on same copula families, this study highlighted that they are not enough for mixtures with different copula pairs. Most probably, the main reason for this identification problem is the selection of copula families and considered dependence parameters. As [3] stated in their empirical findings, in some cases EM algorithm can not identify the correct mixture model as a result of dependence structure of the selected copula functions.

To sum up, the performance of such measures are not very promising and also no measure is examined in terms of the distance between the estimated and the correct models. For this reason, another beneficial measure for the comparison might be Kullback-Leibler (KL) divergence, previously proposed by [39] and considered again within the context of mixture of D-vines by [33]. Furthermore, the most plausible method to compare such mixture models based on the corresponding pair copulas will be the construction of a suitable GOF test. In this respect, the idea studied by [34] can be borrowed as a departure point for building a model selection tool for mixtures. These issued problems are at the top of the list of further plans.

5.3.5 Further Benefits

Mixture of copulas, indeed mixture of vines as a contribution of this study, has various potentials to capture more realistic data generating processes. Especially, these mixture models can serve as a new tool for understanding distinct dependence patterns in terms of both degree and the structure of it. In that respect, such mixture models have great advantages to examine extreme co-movements in international markets. Furthermore, as it is suggested by [18], mixture of vines will be very beneficial for empirical modelings in finance and economics, including problems about risk management and asset pricing models.

Even if the main focus of the thesis study is the data generating processes and parameter estimations under various optimization tools, this work has another potential in terms of model-based clustering. Finite mixture models based on copulas, eventually vines for higher dimension, have very fruitful opportunities to investigate the number of clusters depends on distinct dependence structure. For further improvements of the

study, another fertile research field is clustering and its reflections on financial problems, as it is clearly pointed out in the empirical findings of the study of [3].

Additionally, apart from financial problems, mixture of vines can be useful to examine the dependence patterns emerged in environmental science. More clearly, dependence among the weather stations and its impacts on crop modeling have certain benefits to derive carefully designed risk management strategies in agriculture. In the application part of [22], the dependence between weather stations in terms of precipitation has been discussed with clustering only. This methodology can be extended directly with the empirical findings of this study to understand and classify droughts, the connection between dry periods and crop yield in any location. Such a real life problem should be investigated by using mixture models, as it is strongly suggested by the authors of this study.

REFERENCES

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

[2] O. Akbilgic, H. Bozdogan, and M. E. Balaban. A novel hybrid rbf neural networks model as a forecaster statistics and computing. *PhD Thesis Istanbul University*, 2013.

[3] V. Arakelian and D. Karlis. Clustering dependencies via mixtures of copulas. *Communications in Statistics - Simulation and Computation*, 43(7):1644–1661, 2013.

[4] J. Barzilai and J. Borwein. Two point step size gradient methods. *IMA Journal of Numerical Analysis*, 8(1):141–148, 1988.

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

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

[7] C. Bergmeir, D. Molina, and J. Benítez. Memetic algorithms with local search chains in r: The rmalschains package. *Journal of Statistical Software, Articles*, 75(4):1–33, 2016.

[8] E. G. Birgin, J. M. Martinez, and M. Raydan. Inexact spectral projected gradient methods on convex sets. *IMA Journal of Numerical Analysis*, 23(4):539–559, 2003.

[9] E. G. Birgin, J. M. Martinez, and M. Raydan. Spectral projected gradient methods: Review and perspectives. *Journal of Statistical Software*, 60(3):1–21, 2014.

[10] K. A. Clarke. A simple distribution-free test for nonnested model selection. *Political Analysis*, 15(3):347–363, 2007.

[11] E. Cuvelier and F. M. Noirhomme. Clayton copula and mixture decomposition. *Applied Stochastic Models and Data Analysis*, page 699–708, 2005.

[12] L. N. Dempster, A.P. and D. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society Series B*, 39(1):1–38, 1977.

[13] J. DisMann, E. C. Brechmann, C. Czado, and D. Kurowicka. Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics and Data Analysis*, 59:52–69, 2013.

[14] K. Dorota and C. Roger. *Uncertainty Analysis with High Dimensional Dependence Modelling*. John Wiley and Sons, Ltd, 2006.

[15] K.-T. Fang, H.-B. Fang, and D. V. Rosen. A family of bivariate distributions with non-elliptical contours. *Comm. Statistics Theory Methods*, 29:1885–1898, 2000.

[16] M. Frechet. Sur les tableaux de corrélation dont les marges sont données. *Ann Univ Lyon Sect*, A-9:53–77, 1951.

[17] C. Genest and A.-C. Favre. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, 12(4):347–368, 2007.

[18] L. Hu. Dependence patterns across financial markets: a mixed copula approach. *Applied Financial Economics*, 16(10):717–729, 2006.

[19] H. Joe. Families of m-variate distributions with given margins and m(m-1)/2 bivariate dependence parameters. *IMS Lecture Notes-Monograph Series. Hayward, CA*, pages 120–141, 1996.

[20] H. Joe. *Dependence Modeling with Copulas*. Chapman and Hall/CRC, 2014.

[21] C. T. Kelley. *Iterative Methods for Optimization*. SIAM In Applied Mathematics, 1999.

[22] D. Kim, J.-M. Kim, S.-M. Liao, and Y.-S. Jung. Mixture of d-vine copulas for modeling dependence. *Computational Statistics and Data Analysis*, 64(C):1–19, 2013.

[23] G. Kirgat and A. Surde. Review of hooke and jeeves direct search solution method analysis applicable to mechanical design engineering. *International Journal of Innovations in Engineering Research and Technology*, 1(2):1–14, 2014.

[24] I. Kosmidis and D. Karlis. Model-based clustering using copulas with applications. *Statistics and Computing*, 26(5):1079–1099, 2016.

[25] R. D. Matteis. Fitting copulas to data. *Diploma thesis Institute of Mathematics of the University of Zurich*, 2001.

[26] D. Molina, M. Lozano, C. García-Martínez, and F. Herrera. Memetic algorithms for continuous optimisation based on local search chains. *Evolutionary Computation*, 18(1):27–63, 2010.

[27] K. Mullen, D. Ardia, D. Gil, D. Windover, and J. Cline. DEoptim: An R package for global optimization by differential evolution. *Journal of Statistical Software*, 40(6):1–26, 2011.

[28] R. B. Nelsen. *An Introduction to Copulas (Springer Series in Statistics)*. Springer-Verlag, 2006.

[29] D. Oakes. A model for association in bivariate survival data. 44(3):414–422, 1982.

[30] O. Okhrin, S. Trimborn, S. Zhang, and Q. M. Zhou. gofcopula: Goodness-of-fit tests for copulae r package version 0.2-3. 2016.

[31] K. Price, R. Storn, and J. Lampinen. *Differential Evolution – A Practical Approach to Global Optimization*. Springer-Verlag, 2006.

[32] A. Roche. Em algorithm and variants: An informal tutorial. 2012.

[33] A. Roy and S. K. Parui. Pair-copula based mixture models and their application in clustering. *Pattern Recognition*, 47(4):1689–1697, 2014.

[34] U. Schepsmeier. Efficient information based goodness-of-fit tests for vine copula models with fixed margins: A comprehensive review. *Journal of Multivariate Analysis*, 138:34–52, 2015.

[35] M. Sklar. *Fonctions de Répartition À N Dimensions Et Leurs Marges*. Université Paris 8, 1959.

[36] T. M. I. Strategy. Msci emerging markets index. June 18, 2018.

[37] M. Vrac, A. Chédin, and E. Diday. Clustering a global field of atmospheric profiles by mixture decomposition of copulas. *Journal of Atmospheric and Oceanic Technology*, 22(10):1445–1459, 2005.

[38] Q. H. Vuong. Ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57(2):307–333, 1989.

[39] Q. Wang, S. R. Kulkarni, and S. Verdu. Divergence estimation for multidimensional densities via k-nearest-neighbor distances. *IEEE Transactions on Information Theory*, 55(5):2392–2405, 2009.

APPENDIX A

DEPENDENCE PATTERN FOR C-VINE MIXTURE WITH VARIOUS WEIGHTS

In the presented figures below, Figures A.1-A.5 exhibits directly the 1 component Clayton-Joe mixture model, whenever $\pi_2 = 1$ or $\pi_1 = 1$, respectively. Besides, for all mixture models, because of selected parameters the strong dependence structure has been considered with weight parameters π_1 and π_2 for the 1'st and 2'nd component. Besides, given pictures include Kendall τ values in all figures.

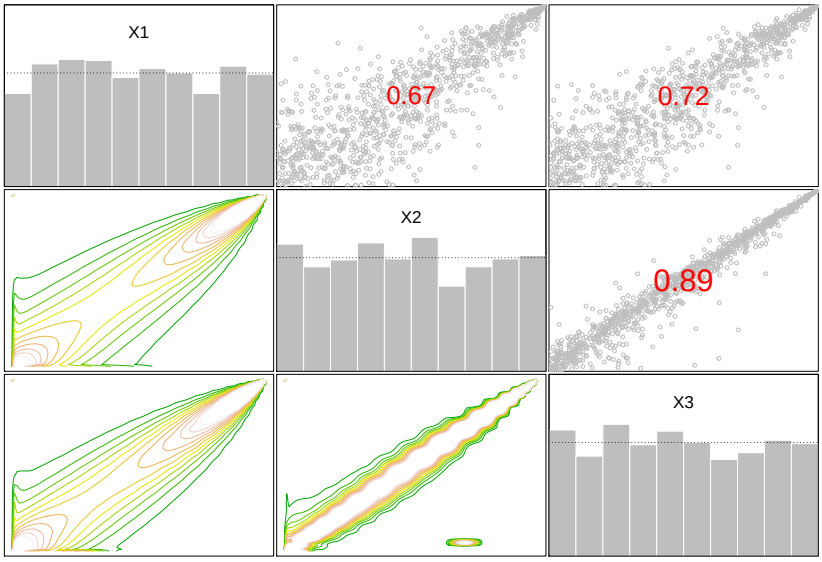


Figure A.1: Dependence Structure of 2-Component Clayton-Joe mixture with $\pi_1 = 0$ and $\pi_2 = 1$.

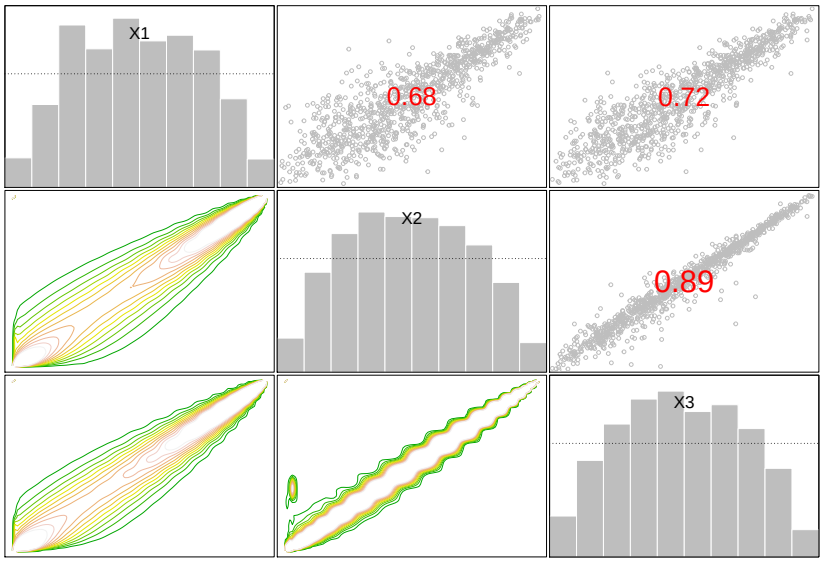


Figure A.2: Dependence Structure of 2-Component Clayton-Joe mixture with $\pi_1 = 0.25$ and $\pi_2 = 0.75$.

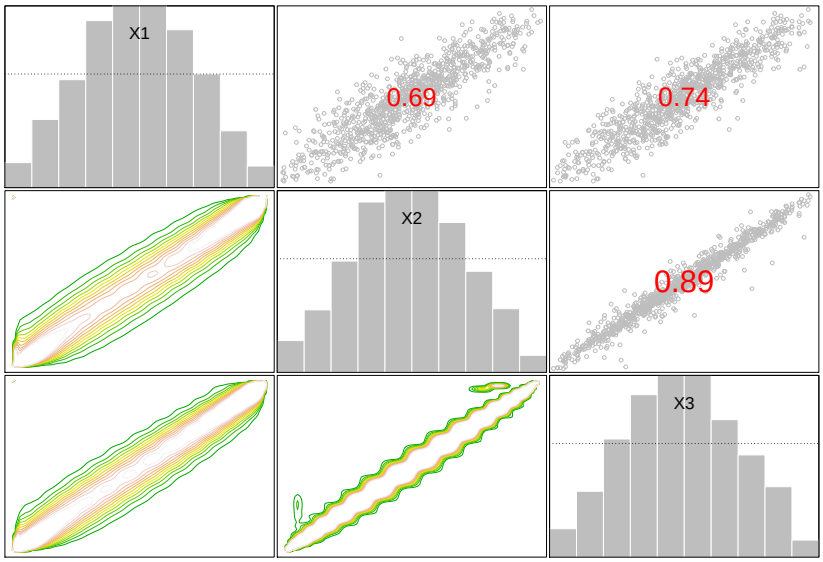


Figure A.3: Dependence Structure of 2-Component Clayton-Joe mixture with $\pi_1 = 0.5$ and $\pi_2 = 0.5$.

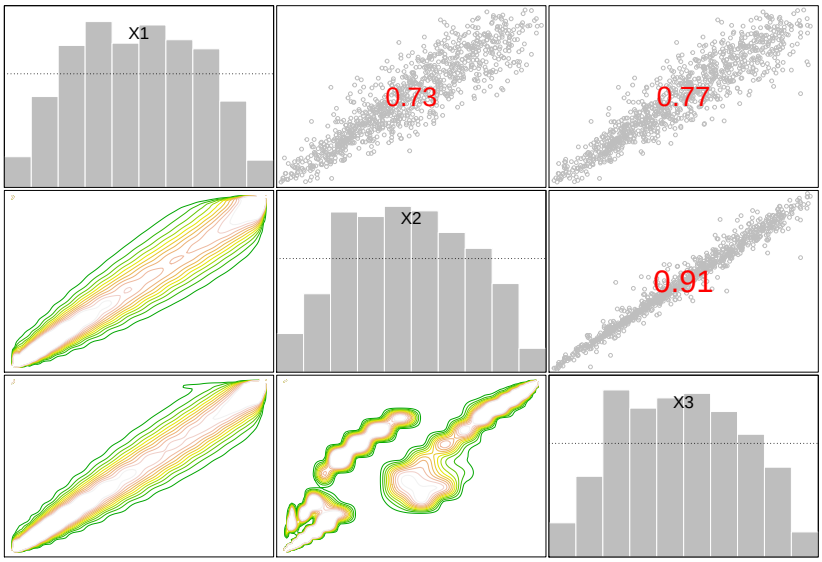


Figure A.4: Dependence Structure of 2-Component Clayton-Joe mixture with $\pi_1 = 0.75$ and $\pi_2 = 0.25$.

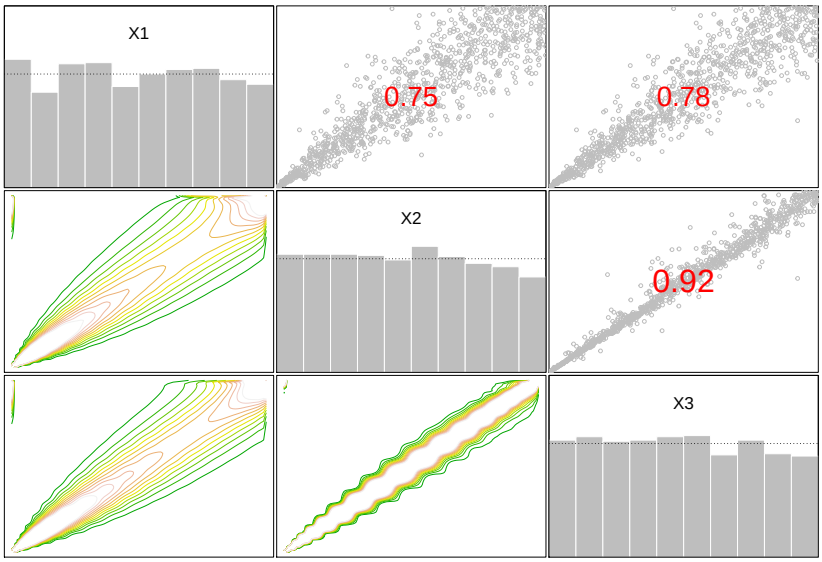


Figure A.5: Dependence Structure of 2-Component Clayton-Joe mixture with $\pi_1 = 1$ and $\pi_2 = 0$.

APPENDIX B

CD-VINE MIXTURE MODEL FLOW CHART

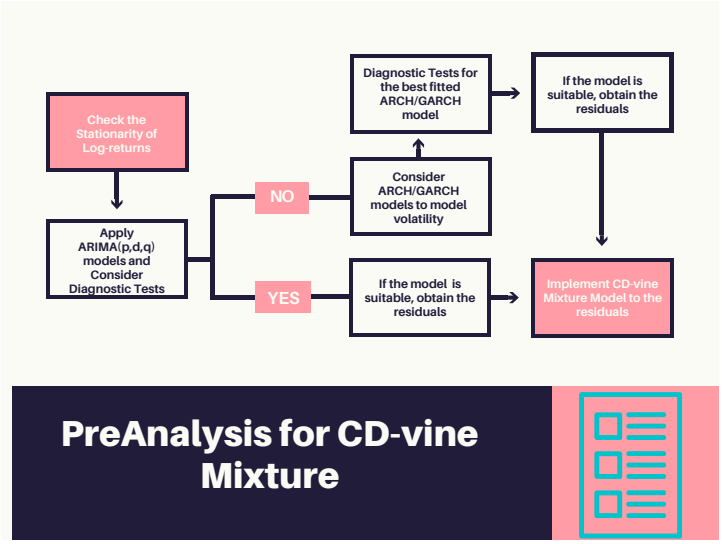


Figure B.1: Flow Chart of pre-analysis for CD-vine mixture

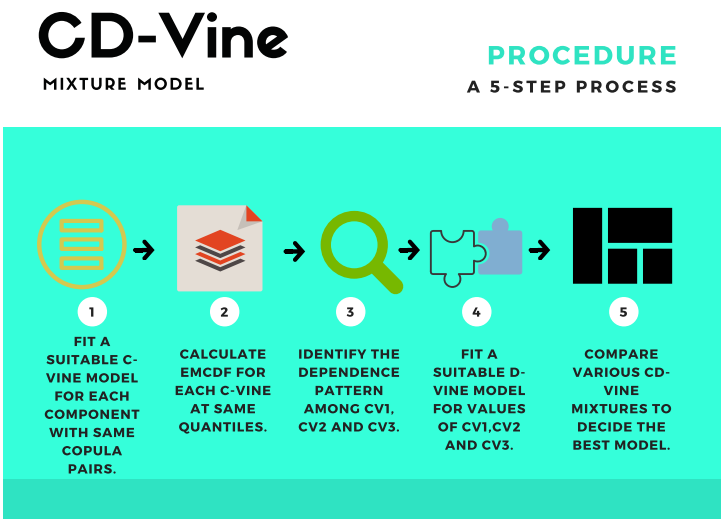


Figure B.2: Flow Chart of CD-vine mixture model procedure

APPENDIX C

ALGORITHMS FOR MIXTURE MODELS

Algorithm 5 2 component 4-dimensional Mixture of C-vines

Require: Sample two C-vine models $CV1$ and $CV2$

for $k = 1$ to $max1$ **do**

$g_{mix} \leftarrow \pi_1 f(\boldsymbol{x}; \boldsymbol{\phi_{CV1}}) + \pi_2 f(\boldsymbol{x}; \boldsymbol{\phi_{CV2}})$,

for $t = 1$ to $LargeNumber$ **do**

for $i = 1$ to N **do**

$fCV_p[i,1] \leftarrow c12_p[i,1] \times c13_p[i,1] \times c14_p[i,1] \times c23\_1_p[i,1] \times c24\_1_p[i,1] \times c34\_12_p[i,1]$, where fCV_p is the density function.

Calculate $\widehat{z_{nm}}$

end for

for $p = 1, 2$. Calculate Log-likelihood $logL0_{CV}[t]$

if $|logL0_{CV}[t] - logL0_{CV}[t-1]| \leq \epsilon$ **then**

Break

end if

where $\epsilon = 10^{-6}$

Update weights using $\widehat{z_{nm}}$ the parameters of each component via optimization $PhiCV1_{new}[k,c] = op.par[c]$ and $PhiCV2_{new}[k,c] = op.par[c]$ for $c = 1, \cdots dd$, where dd is number of parameters.

end for

Store the updated weights for the next iteration $Pi_{last}[k,] = Pi_{new}$

Calculate $AIC[k,]$, $BIC[k,]$ and $CAIC[k,]$ for iteration k

end for

Require:

Get the final estimates by mean($PhiCV1_{new}, PhiCV2_{new}$).

Algorithm 6 CD-vine Mixture Model (1'st Step)

Require: Sample $CV1$, $CV2$ and $CV3$

for $k = 1$ to $max1$ **do**

for $t = 1$ to $LargeNumber$ **do**

for $i = 1$ to N **do**

$fCV_p[i,1] \leftarrow c12_p[i,1] \times c13_p[i,1] \times c14_p[i,1] \times c23\_1_p[i,1] \times c24\_1_p[i,1] \times c34\_12_p[i,1]$

end for

for $p = 1, 2, 3$

Calculate $logL0_{CV}[t]$

if $|logL0_{CV}[t] - logL0_{CV}[t-1]| \leq \epsilon$ **then**

Break

end ifwhere $\epsilon = 10^{-3}$

$PhiCV1_{new}[k,c] = op.par[c]$

$PhiCV2_{new}[k,c] = op.par[c]$

$PhiCV3_{new}[k,c] = op.par[c]$

end for

for $c = 1, \cdots dd$ where dd is number of parameters for the component

Calculate $AIC[k,]$, $BIC[k,]$ and $CAIC[k,]$ for iteration k

end for

Require:

mean($PhiCV1_{new}, PhiCV2_{new}, PhiCV3_{new}$)

Algorithm 7 CD-vine Mixture Model (2'nd Step)

for $kk = 1$ to $max1$ **do**

Calculate $F(cv1)$, $F(cv2)$ and $F(cv3)$ based on estimated parameters

for $tt = 1$ to $LargeNumber$ **do**

for $j = 1$ to N **do**

Calculate density for D-vine

$f_{DV}[j,1] = c12_{DV}[j,1] \times c23_{DV}[j,1] \times c13\_2_{DV}[j,1]$

end for

Calculate Loglikelihood $logL0_{DV}[tt]$

if $|logL0_{DV}[tt] - logL0_{DV}[tt-1]| \leq \epsilon$ **then**

Break

end ifwhere $\epsilon = 10^{-3}$

Second Step Maximization via optimization and update

$PhiDV_{new}[kk, c] = op.par[c]$ for $c = 1, \cdots 3$

end for

Calculate $AIC[kk,]$, $BIC[kk,]$ and $CAIC[kk,]$ for iteration kk

end for

Require:

mean($PhiDV_{new}$)

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Evkaya, Ömer Ozan
Nationality: Turkish
Date and Place of Birth: 03.07.1987, Avanos/Nevşehir
Marital Status: Single
Phone: +90 555 872 1751

EDUCATION

Degree	Institution	Year of Graduation
M.Sc.	Actuarial Science, Inst. of App. Math., METU	2012
B.S.	Mathematics, METU	2009
High School	Hacıbektaş Çok Programlı Lisesi	2004

PROFESSIONAL EXPERIENCE

Year	Place	Enrollment
2009-2018	Mathematics Department, Atılım University	Research Assistant

PUBLICATIONS

1. Evkaya, O., Yildirak, S. K. & Selcuk-Kestel, A. S. (2017). Assessment of Index-based Drought Insurance. *Ekonomik Yaklaşım*, 28(105).

2. G. Erkan, O. Evkaya and S. Turkan, The Determination of the Affecting Factors

of the Number of Babies Born Alive in Multiple Pregnancies with Poisson Models. *Turkiye Klinikleri Biyoistatistik Dergisi*.

International Conference Publications

1. Evkaya, O., Yozgatlıgil, C. & Selcuk-Kestel, A. S. (2017). Finite Mixture of C-vines for Complex Dependence. *Proceedings of the 20th European Young Statisticians Meeting*, 14-18 August, 2017, Uppsala, Sweeden, 125–130.

Presentations

1. **EuroMath-EuroScience 2018, Krakow, Poland**, "Mathematics in Theatre: Beckett's Quad Play and Consuming Space: Joint Workshop with Gizem Gürer for young researchers", 7-11 March 2018.
2. **International Statistics Congress (ISC2017), Ankara, TURKEY**, "Drought Forecasting with Time Series and Machine Learning Approaches", 6-8 December 2017.
3. **Advanced Modelling in Mathematical Finance. A Workshop in Honour of Azize Hayfavi's 75'th Birthday, Ankara, TURKEY**, "Copula(s), advantages and limitations with its applications", 23 October 2017.
4. **20th European Young Statisticians Meeting, Uppsala, SWEDEN**, "Finite Mixture of C-vines for Complex Dependence", 14-18 August 2017.
5. **Copula Theory and its Applications. To commemorate the 75'th birthday of Professor Roger B. Nelsen, Almeria, SPAIN**, "Multivariate Analysis of Drought Characteristics Using Vine Copula Approach", 3-5 July 2017.
6. **3rd International Researchers, Statisticians and Young Statisticians Congress (IRSYC 2017), Konya, TURKEY**, "The Investigation of the Number of Babies Born Alive in Multiple Pregnancies Using Poisson Models", 24-26 May 2017.
7. **Recent Developments in Dependence Modeling with Applications in Finance and Insurance (Fourth Edition), Aegina, GREECE**, "Mixture of Vine Copulas for Complex and Hidden Dependence", 22-23 May 2017

8. **X'th International Statistics Days Conf.(ISDC 2016), Giresun, TURKEY**, "Bivariate Drought Analysis using Copulas", 7-9 October 2016.

9. **IX'th International Statistics Congress, Antalya, TURKEY**, "Electricity Consumption and Economic Growth in Turkey: Is Copula framework possible ?", 28 October-1 November 2015.

10. **Summer School on Copulas for Hydrology and Environmental Sciences, Pau, FRANCE**, "Copula Based Drought Indices (joint work with Tobias, E. and Ondrej, L.)", 6-10 July 2015.

11. **X'th International Conference on Applied and Computational Mathematics (ICACM), Ankara, TURKEY**, "Basis Risk Comparison of Bayesian based Wheat Yield Estimation against drought", 3-6 October, 2012.