Faculty of Engineering and Information Technology
University of Technology, Sydney

Copula-based High Dimensional Dependence Modelling

A thesis submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

by

Wei Wei

November 2014

CERTIFICATE OF AUTHORSHIP/ORIGINALITY

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of Candidate

_____

Acknowledgments

Foremost, I would like to express my sincere gratitude to my supervisors Prof. Longbing Cao for the continuous support of my Ph.D study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study.

I also would like to appreciate my co-supervisor Associate Prof. Jinyan Li for providing me with continuous support throughout my PhD study and research. Without his professional guidance and persistent help, this thesis would not have been possible.

I thank my fellow labmates in Advanced Analystics Institute: Xuhui Fan, Yin Song, Junfu Yin for the stimulating discussions, for the sleepless nights we were working together before deadlines, and for all the fun we have had in the last three years.

I place on record, my sense of gratitude to the team leader, Zhigang Zheng in Australian Taxation Office for his expert and sincere help in the project. I am also grateful to my team members, Allen Lin, Mu Li, Chunming Liu and Wei Cao, for their hard working in project teams.

Last but not the least, I would like to thank my family: my wife and my parents, for their unconditional support, both financially and emotionally throughout the whole PhD studying.

Wei Wei
June 2014 @ UTS

Contents

Certificate . . . . . **i**
Acknowledgment . . . . . **iii**
List of Figures . . . . . **xi**
List of Tables . . . . . **xiii**
List of Publications . . . . . **xv**
Abstract . . . . . **xvii**

Chapter 1 Introduction **1**
1.1 Background 1
1.2 Research Issues 7
1.2.1 RI 1: High Dimensional Data 7
1.2.2 RI 2: Dependence Structure Assumptions 7
1.2.3 RI 3: Dependence Structure Truncation and Optimisation 8
1.2.4 RI 4: High Dimensional Dependence Evaluation 8
1.3 Research Contributions 8
1.4 Thesis Structure 10

Chapter 2 Literature Review and Foundation **13**
2.1 Copula 13
2.1.1 Elliptical Copulas 15
2.1.2 Archimedean Copulas 16
2.1.3 Bivariate Copula Families 17
2.1.4 The Tail Dependence of Copulas 22

2.2 Copula-based Models 23
2.2.1 Time-varying Copula Models 24
2.2.2 Estimate and Inference 28
2.2.3 Goodness-of-Fit Testing and Model Selection 34
2.3 Multivariate Copula Models 36
2.3.1 Multivariate Elliptical Copulas 36
2.3.2 Multivariate Archimedean Copulas 37
2.4 Vine Copula Models 41
2.4.1 Pair Copula Construction 41
2.4.2 Graph Theory 44
2.4.3 Regular Vine 47
2.4.4 Vine Tree Structure Selection Strategies 50
2.4.5 Parameter Estimate 52
2.4.6 Copula Families Selection 53
2.4.7 Vine Optimisation 54
2.4.8 Probability Graphical Models and Copula 54
2.5 Model Comparison Test – Vuong Test 57
2.6 Partial Correlation 59
2.6.1 Partial Correlation Definition 59
2.6.2 Partial Correlation and Conditional Correlation 60
2.7 Time Series Analysis 63
2.7.1 Basic Definition 63
2.7.2 The ARMA Models 65
2.7.3 The GARCH Models 66
2.7.4 The ARMA(1,1)-GARCH(1,1) Model 67
2.7.5 The Ljung Box Test 68
2.8 Applications in Financial Market 69
2.8.1 Risk Managements 69
2.8.2 Derivative Contracts 70
2.8.3 Portfolio Decision Problems 71
2.8.4 Financial Contagion 71

2.8.5 Regular Vine Copula Applications 72
2.8.6 Other Applications 72
2.9 Summary 73

Chapter 3 Modelling the Complex Dependence Structures by Using Truncated Partial Canonical Vine 75
3.1 Introduction 75
3.2 Canonical Vine Autoregressive Model 80
3.3 Partial Canonical Vine Model 82
3.3.1 Canonical Vine 83
3.3.2 Partial Canonical Vine Specification 87
3.3.3 Dependence Structure Construction 88
3.3.4 Copula Families Selection 91
3.3.5 Vine Tree Structure Truncation and Simplification . . . 91
3.3.6 Parameter Estimate 94
3.4 Marginal Models Specification 96
3.4.1 ARMA-GARCH Model 96
3.4.2 Error Distribution 97
3.5 Portfolio of Value at Risk 98
3.6 Dependence Analysis on European Stock Returns 101
3.6.1 Data and Marginal Model Specification 101
3.6.2 The Truncation and Simplification Analysis of Partial Canonical Vine 104
3.6.3 Experiments on Value at Risk 105
3.7 Summary 108

Chapter 4 Modeling Asymmetry and Tail Dependence among Multiple Variables by Using Partial Regular Vine 116
4.1 Introduction 116
4.2 Regular Vine Copula Model with Asymmetric Dependence . . 119
4.2.1 Regular Vine Copula Specification 119
4.2.2 Regular Vine Tree Structure Construction 120

4.2.3 Bivariate Copula Family Selection 123
4.3 Marginal Distribution Specification and Parameter Estimate . 124
4.4 Value at Risk– A widely Used Evaluation in Financial Market 126
4.5 Case Study 128
4.5.1 Non-parametric Dependence Analysis 128
4.5.2 Regular Vine Copula Specification and Tail Dependence Analysis 129
4.5.3 Value at Risk Forecasting 133
4.6 Summary 134

Chapter 5 High-dimensional Dependence Modelling through Truncated Partial Regular Vine Copula 145
5.1 Introduction 145
5.2 Truncated Partial Regular Vine 150
5.2.1 Partial Regular Vine Specification 150
5.2.2 Partial Regular Vine Tree Structure Construction . . . 151
5.2.3 Partial Regular Vine Tree Structure Truncation 155
5.2.4 Bivariate Copula Selection 156
5.2.5 Marginal Distribution Specification and Parameter Estimation 158
5.3 Evaluation Methods 159
5.3.1 In-Sample Evaluation Methods 160
5.3.2 Out-of-Sample Evaluation Method 160
5.4 Experimental Results 162
5.4.1 Data and Marginal Distributions Specification 162
5.4.2 Partial Regular Vine Truncation Analysis 164
5.4.3 In-Sample Performance Analysis 166
5.4.4 Out-of-Sample Performance Analysis 167
5.5 Summary 169

Chapter 6 Optimal Allocation of High Dimensional Assets through Truncated Partial Canonical Vines 174

6.1 Introduction 174
6.2 The Portfolio Optimisation 177
6.2.1 CRRA Optimisation Function 177
6.2.2 Truncated Partial Canonical Vine 178
6.2.3 Marginal Distribution and Estimate 180
6.3 Evaluation Methods 181
6.4 Case Study 182
6.4.1 Data and Model Specification 182
6.4.2 Experiment Results and Analysis 184
6.5 Summary 186

Chapter 7 Conclusions and Future Work 193
7.1 Conclusions 193
7.2 Future Work 195

Appendix A Dependence Measurements 197
A.1 Pearson Correlation 197
A.2 Kendall's Tau 198

Appendix B The Determinant of Partial Correlation-based Canonical Vine and Regular Vine 200

Appendix C Model Selection Criteria 203

Appendix D List of Symbols 204

Bibliography 206

List of Figures

1.1 The scatter plot of copulas 3
1.2 The profile of work in this thesis 12

2.1 Fully nested Archimedean construction 38
2.2 Partially nested Archimedean construction 39
2.3 Hierarchically nested Archimedean construction 40
2.4 An example graph with 5 nodes and 5 edges 46
2.5 An example graph with a cycle 47
2.6 An example tree with 5 nodes and 4 edges 47

3.1 An example of dependence structure 76
3.2 CAVA dependence structure 81
3.3 PCV model work flow 82
3.4 PCV dependence structure 84
3.5 The canonical vine dependence structure with 7 variables . . . 86
3.6 The results of truncated and simplified partial canonical vine against ρ_{sign} 114
3.7 The VaR forecasting with portfolio returns 115

4.1 The partial regular vine tree structure 131
4.2 The tail dependence of $v4$ and $v6$ 137
4.3 The tail dependence of $v4$ and $v5$ 138
4.4 The tail dependence of $v4$ and $v8$ 139
4.5 The tail dependence of $v4$ and $v7$ 140

4.6 The tail dependence of $v2$ and $v7$ 141
4.7 The tail dependence of $v2$ and $v3$ 142
4.8 The tail dependence of $v1$ and $v2$ 143
4.9 The VaR forecasting of portfolio returns 144

5.1 Performance analysis of truncated partial regular vines 172
5.2 The VaR forecast of portfolio returns by using $TPRV_{0.1}$ 173

6.1 The truncated and simplified vine for *S&P500* 190
6.2 The truncated and simplified vine for *Stoxx50E* 191
6.3 The portfolio value results analysis 192

List of Tables

2.1 The Tail Dependence of the Bivariate Copula Family 24

3.1 Indices and Stocks 102
3.2 Descriptive Statistics for the Indices 103
3.3 Results of Ljung Box Test 104
3.4 Determinants and Numbers of Pair Copulas 106
3.5 Truncated and Simplified Partial Canonical Vine Analysis . . 109
3.6 Number of Exceedances of VaR Forecast 112
3.7 The Results of VaR Backtesting 113

4.1 The Tail Dependence of Copula Family 123
4.2 Non-parametric Tail Dependence Analysis 129
4.3 Results of BP and LP Tests 130
4.4 Tail Dependence Analysis by Using Various Copula 135
4.5 The Backtesting Results of Value at Risk Forecasting 136

5.1 The Tail Dependence of Bivariate Copulas 157
5.2 Ljung-Box Test Results 163
5.3 The Total Number of Edges in Truncated Partial Regular Vine against Truncation Value 165
5.4 The Comparison between $TPRV_{0.1}$ and PRV 166
5.5 Results of Bivariate Copula Selection and Model Selection . . 167
5.6 Vuong Test Results 168
5.7 Backtesting Results of Value at Risk for Currencies 170

5.8 Backtesting Results of Value at Risk for Currencies 171

6.1 The Partial Correlation Analysis 183
6.2 The Comparison between Truncated, Simplified and Non-truncated Vines 184
6.3 The Results of the Opportunity Costs, Final Amounts and Utilities for $S\&P500$ 188
6.4 The Results of the Opportunity Costs, Final Amounts and Utilities for $Stoxx50E$ 189

List of Publications

Papers Published

- **Wei Wei**, Junfu Yin, Jinyan Li, Longbing Cao (2014), Modelling Asymmetry and Tail Dependence among Multiple Variables by Using Partial Regular Vine. *in* 'Proceedings of the SIAM International Conference on Data Mining Conference (**SDM14**)' , pp.776-784.

- Junfu Yin, Zhigang Zheng, Longbing Cao, Yin Song, **Wei Wei** (2013), Efficiently Mining Top-K High Utility Sequential Patterns. *in* 'Proceedings of the 13th IEEE International Conference on Data Mining (**ICDM13**)', pp. 1259-1264.

- **Wei Wei**, Jinyan Li, Longbing Cao, Jingguang Sun, Chunming Liu, Mu Li (2013), Optimal Allocation of High Dimensional Assets through Canonical Vines. *in* 'Proceedings of the 17th pacific-asia conference on knowledge discovery and data mining (**PAKDD13**)' , pp. 366-377.

- **Wei Wei**, Xuhui Fan, Jinyan Li, Longbing Cao (2012), Model the complex dependence structures of financial variables by using canonical vine. *in* 'Proceedings of the 21st ACM international conference on Information and knowledge management (**CIKM12**)', ACM, pp, 1382-1391.

Papers to be Submitted/Under Review

- Zhigang Zheng, **Wei Wei**, Chunming Liu, Wei Cao, Longbing Cao, and Maninder Bhatia (2014) An Effective Contrast Sequential Pattern Mining Approach on Taxpayer Behavior Analysis, *World Wide Web Journal*.

- **Wei Wei**, Longbing Cao, Jinyan Li, and Philip S. Yu (2014) High-dimensional Dependence Modelling through Truncated Partial Regular Vine Copula, to be submit the 14 th IEEE International Conference on Data Mining (**ICDM14**).

Research Reports of Industry Projects

- **Wei Wei**, Zhigang Zheng, Maninder Bhatia. Self-finaliser analysis (debt cases) of activity statement – data mining modelling and evaluation report , Debt Collection and Optimisation Project, Australian Taxation Office, Dec 2012.

- Zhigang Zheng, **Wei Wei**, Maninder Bhatia. Self-finaliser analysis (due date cases) of activity statement – data mining modelling and evaluation report , Debt Collection and Optimisation Project, Australian Taxation Office, May 2013.

- Zhigang Zheng, **Wei Wei**, Wei Cao, Maninder Bhatia. Action response model of activity statement – data mining modelling and evaluation report, Debt Collection and Optimisation Project, Australian Taxation Office, Nov 2013.

- **Wei Wei**, Zhigang Zheng and Wei Cao. Self-finalising Lodgment Analysis Project – data mining modelling and evaluation report, Lodgements Collection and Optimisation Project, Australian Taxation Office, May 2014.

Abstract

Big data applications increasingly involve high-dimensional and sophisticated dependence structures in complex data. Modelling high-dimensional dependence, that is, the dependence between a set of high-dimensional variables, is a critical but challenging issue in many applications including social media analysis and financial markets. A typical example concerns the interplay of financial variables involved in driving complex market movements. A particular problem is understanding the dependence between high-dimensional variables with tail dependence and asymmetric characteristics which appear widely in financial markets. Typically, existing methods,such as the Bayesian logic program, relational dependency networks and relational Markov networks, build a graph to represent the conditional dependence structure between random variables. These models aim at high-dimensional domains, and have the advantage of learning latent relationships from data. However, they tend to force the local quantitative part of the model to take a simple form such as the discretized form of the data when multivariate Gaussian or its mixtures cannot capture the data in the real world. The complex dependencies between high-dimensional variables are difficult to capture.

In statistics and finance, the copula has been shown to be a powerful tool for modelling high-dimensional dependencies. The copula splits the multivariate marginal distributions from dependence structures, so that the specification of dependence structures can be investigated independently of the marginal distributions. It can provide a flexible mechanism for modelling real world distributions that cannot be handled well by graphical models. Thus,

researchers have tried to combine copula and probability graphical models, such as the tree-structured copula model and copula Bayesian networks. These copula-based models aim to resolve the limitations of discretizing data, but they impose assumptions and restrictions on the dependence structure.These assumptions and restrictions are not appropriate for dependence modelling among financial variables.

In order to address these research limitations and challenges, this thesis proposes the use of the truncated partial correlation-based canonical vine copula, partial correlation-based regular vine copula and truncated partial correlation-based regular vine copula to model the dependence of high-dimensional variables. Chapter 3 introduces a new partial correlation-based canonical vine to identify the asymmetric and non-linear dependence structures of asset returns without any prior dependence assumptions. To simplify the model while maintaining its merit, a partial correlation-based truncation method is proposed to truncate the canonical vine. The truncated partial correlation-based canonical vine copula is then applied to construct and analyse the dependence structures of European stocks as a case study.

Chapter 4 introduces the truncated partial correlation-based regular vine copula to explore the relations in multiple variables. Very often, strong restrictions are applied on a dependence structure by existing high-dimensional dependence models. These restrictions disabled the detection of sophisticated structures such as the upper and lower tail dependence between multiple variables. A partial correlation-based regular vine copula model may relax these restrictions. The partial correlation-based regular vine copula model employs a partial correlation to construct the regular vine structure, which is algebraically independent. This model is able to capture the asymmetric characteristics among multiple variables by using a two-parametric copula with flexible lower and upper tail dependence. The method is tested on a cross-country stock market data set to analyse the asymmetry and tail dependence in the dynamic period.

Chapter 5 proposes a novel truncated partial correlation-based regular

vine copula model which can capture more flexible dependence structures without making pre-assumptions about the data. Specifically, the model employs a new partial correlation to build the dependence structures via a bottom-up strategy. It can identify important dependencies and information among high-dimensional variables, truncating the irrelevant information to significantly reduce the parameter estimate time. The in-sample and out-of-sample performance of the model are examined by using the data in currency markets over a period of 17 years.

Chapter 6 discusses how to resolve the high-dimensional asset allocation problem through a partial correlation-based canonical vine. Typically, the mean-variance criteria which is widely used in asset allocation, is actually not the optimal solution for asset allocation as the joint distribution of asset returns are distributed in asymmetric ways rather than in the assumed normal distribution. The partial correlation-based canonical vine can resolve the issue by producing the asymmetric joint distribution of asset returns in the utility function. Then, the utility function is then used for determining the optimal allocation of the assets. The performance of the model is examined by using data in both European and United State stock markets.

In summary, this thesis proposes three dependence models, including one canonical vine and two regular vines. The three dependence models, which do not impose any dependence assumption on the dependence structure, can be used for modelling different high-dimensional dependencies, such as asymmetry or tail dependencies. All of these models are examined by the datasets in the real world, such as stock or currency markets. In addition, the partial correlation-based canonical vine is used to resolve optimisation allocation of assets in stock markets. This thesis works to show that there is great potential in applying copula to model complex dependence, particular in modelling time-varying parameters, or in developing efficient vine copula simplification methods.

Chapter 1

Introduction

1.1 Background

Uncertainty, which may come, for example from the lack of knowledge or noise in data, cannot be ignored. It may be the result of a regulatory requirement to optimise the industrial process. Consequently, various methods to deal with uncertainty have been developed, including trying to predict the uncertainty in a system. One important element is to measure and model uncertainty before predicting it. First of all, to analyse any model, its input factors have to be identified. The process of input identification is not very easy since it may require approaching experts in given study areas and using their previous knowledge to reasonably determine the input parameters. Once the input parameters have been identified, their probability distribution can be identified, and uncertainty may be predicted based on distributions. However, most of the existing theories, tools, systems in statistics, data mining, and machine learning, are based on the assumption of the *Independence and Identical Distribution* (IID) of all input factors. The assumption may work well for those simple problems with weakened and avoidable relations and heterogeneity. The IID assumption, however, is unreasonable in the real world. The dependence between input factors may significantly affect the output of modelling and result in the whole analysis being unrealistic if

this dependence is not considered. Therefore, the dependence modelling of input factors becomes fundamental to research in many fields.

The example which is used in (Kurowicka & Joe 2011) can explain why it is very important to model dependence for financial area. Suppose \$1000 is invested for five years, and the five-year return is:

$$r_{5-year} = 1000(1 + r_1)(1 + r_2)(1 + r_3)(1 + r_4)(1 + r_5)$$

where r_1, r_2, r_3, r_4, r_5 are denoted as the interest rate for those five years. The interest rates are unknown due to the uncertainty. The distribution can be found by investigating their historical data. For simplicity, it can be assumed that they are uniformly distributed between 0.05 and 0.15. To find the final return after five years, it is necessary to build the joint distribution of interest rates. If it is assumed that interest rates are independent of each other, their joint distribution is then a product of marginal distribution dependence, which can be easily calculated. However, if some sort of dependence between interest rates is recognised, then the joint distribution can be built given marginal distribution and dependence.

There are various measurements of dependence in the literature, including Pearson's product moment correlation, Spearman's rank correlation, Kendall's tau, and etc, (Joe 1997, Mari & Kotz 2001). These various measurements of dependence are not mutually independent. These measurement of dependence are not the best, since they either rely on a normal distribution assumption or do not have a wide range of dependence in a parametric family, or provide both positive or negative dependence. These measurements cannot handle the task mentioned in the above example.

Copula can be used for the purpose. The theory of copula is the mathematical language to describe the dependence modelling of input factors. It represents a natural tool for modelling dependence with uniform margins (Joe 1997, Nelsen 1999). These allow for the separate examination of the dependence structure and marginal distribution. They do not rely on the normal assumption, and also do not ensure variables are characterised by

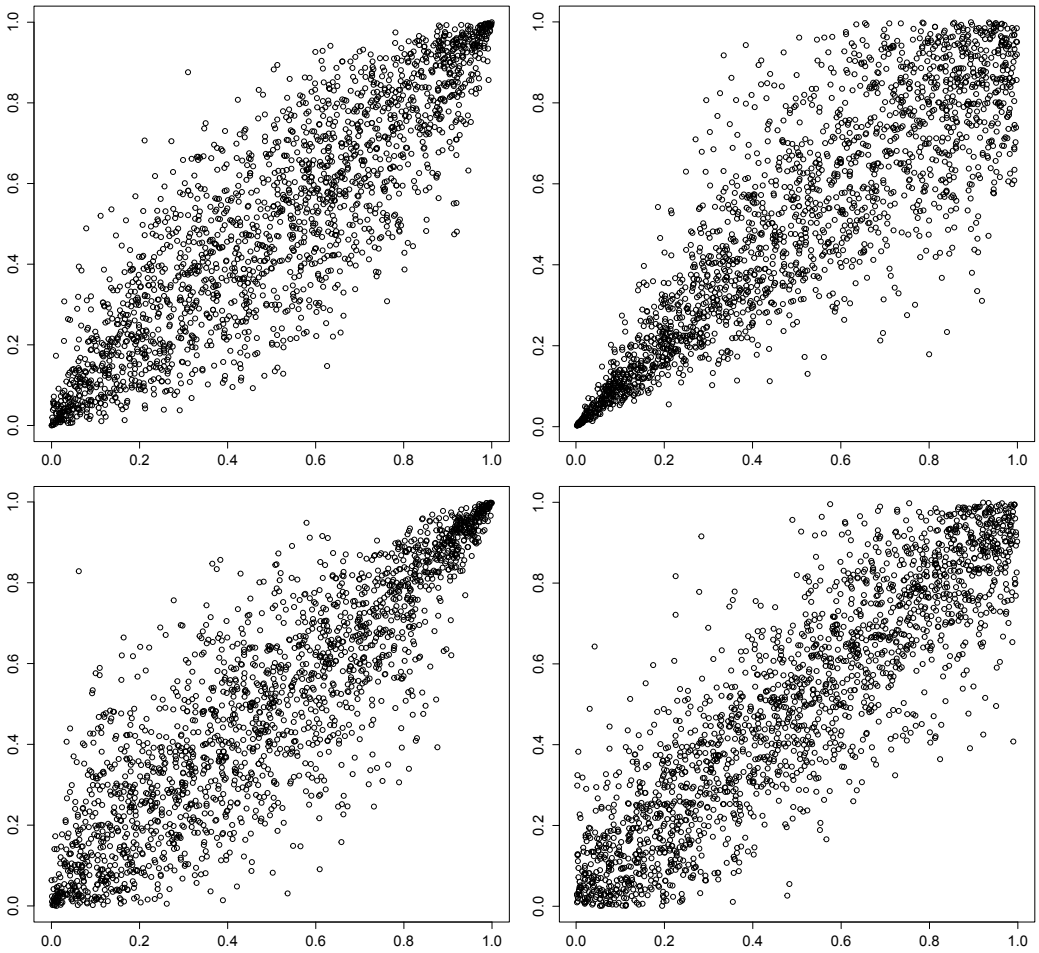

Figure 1.1: The scatter plot of various copulas with correlation 0.7

the same parametric family of the univariate distribution. Different bivariate copula have the ability to model various features of a joint distribution. Figure 1.1 shows the scatter plot for Gaussian (top left), Clayton (top right), Gumbel (bottom left) and Frank (bottom right) copulas. For the same correlation (0.7), the four copula families show different distributions.

Copula has been found to be most successful in bivariate dependence modelling, since a large number of bivariate copula families are accumulated in the literature which can be found in (Joe 1997, Nelsen 1999). These bivariate copulas can be used for a wide range of dependence and tail dependence,

including positive and negative dependence.

The high dimensional dependence modelling, however, is a fundamental research problem in the economics domain and it is useful for a wide range of applications. It is believed that economics is much more complex than physics. In physics, parameters, and phenomena are independent. However, most phenomena are interrelated in economics. To describe economics phenomena numerically, the dependence has to be taken into account. For example, there are almost three thousand stock listed on the New York Stock Exchange. Investors are now exploring an increasing number of asset classes. A portfolio manager finds it challenging work to select investments. The price of all of these assets are constantly changing in response to the anticipation of future performance and news. The movements in price are not independent (Engle 2009). If they were independent, then it could be possible to form a portfolio with negligible volatility, which defied common sense. The dependence structure is a key characteristic in the portfolio choice problem, since it is instrumental in determining and measuring risk. Hence, it is very important to recognise that movements in all asset prices are dependent, and all elements are interconnected in a general equilibrium system. Modelling the dependence structures of high dimensional assets is a very difficult task since the dependence structure exists in hierarchical and horizontal coupling relations.

Suppose there are $\{1, 2, ..., n\}$ assets, and then the ith asset price at time t is $P_{t,i}$. The asset return $r_{t,i}$ is calculated as $r_{t,i} = log(P_{t,i}/P_{t-1,i})$. According to the perspective, a multivariate return matrix can be made as follows:

$$multivariate\ return\ matrix\ = \begin{pmatrix} r_{1,t} & r_{2,t} & \cdots & r_{n,t} \\ r_{1,t-1} & r_{2,t-1} & \cdots & r_{n,t-1} \\ \vdots & \vdots & \ddots & \vdots \\ r_{1,t-k} & r_{2,t-k} & \cdots & r_{n,t-k} \end{pmatrix}$$

where $k < t$. Based on common sense, $r_{1,t}$ is affected by $\{r_{1,t-1}, ..., r_{1,t-k}\}$. In addition, the asset price r_1 is also affected by $\{r_2, ..., r_n\}$. The intra-

dependence is the relationship within one column of the above matrix, while how the assets interact is embodied among the rows of the multivariate return matrix, indicated as inter-dependence. The inter-dependence is very easy to explain, since it is common sense that the asset return is determined by its historical data. The intra-dependence has been demonstrated by the 2008 global financial crisis. That 2008 financial turmoil was originated from the subprime mortgage market in the United States (US), and it quickly spread to every part in the US and global financial system. The bankruptcy of Lehman Brothers in September, 2008 marked a peak point of the crisis – the failure of the fourth largest investment company was the largest bankruptcy in US history. The European financial market was heavily impacted by the collapse of Lehman Brothers.

It is very a challenging job to model the complex high dimensional dependence structure in financial fields. One reason is that each asset return has its characteristics, which are called stylised facts. The important stylised facts are given as follow (Andersen 2009):

(i). Daily returns have little or no exploitable conditional mean predictability;

(ii). The variance of daily returns greatly exceeds the mean;

(iii). Daily returns are not normally distributed;

(iv). Even after standardising daily returns by a dynamic variance model, the standardised daily returns are not normally distributed;

(v). Positive and negative returns of the same magnitude may have different impacts on the variance;

In addition, there are some characteristics in intra-dependence. The dependence between assets returns is asymmetric. For example, stock returns will have a stronger correlation in a bear market downturn than in a bull market. They may decrease together in a bear market, however, they may not increase together in a bull market.

In order to model the dependence in the financial area, a high dimensional copula may be a good choice. Compared with the successfully bivariate dependence modelling through copula, however, little research of the high dimensional dependence modelling can be found in the literature. The most important reason is that many of these bivariate copulas have no straightforward multivariate extension, which leads to the number of high dimensional copulas being fairly limited. Only a limited number of bivariate copulas can be extended to a multivariate version, including multivariate Gaussian copula, multivariate t copula and multivariate Archimedean copula. It should be noted that not all bivariate Archimedean copulas have a corresponding multivariate version (Nelsen 1999). These existing multivariate copula can only handle a limited dependence structure. For example, the multivariate Archimedean copula model has the structure with only a narrow range of negative dependence (McNeil & Nešlehová 2009). The multivariate Gaussian copula model is not suitable to model the asymmetric characteristics, since (i) the Gaussian copula does not have lower and upper tail dependence, and (ii) the Gaussian assumption is not appropriate in the real world (Abdous, Genest & Rémillard 2005, Fang, Fang & Kotz 2002). The multivariate t copula model, which is studied by (Demarta & McNeil 2005, Nikoloulopoulos, Joe & Li 2009), does not have flexible lower and upper tail dependence since the t copula has the same lower and upper tail dependence. In addition, another reason is that the high dimensional dependence structure is susceptible to the curse of dimensionality with limited computational capacity since the complexity of the dependence structure increases exponentially as the dimensions grow. For example, for a n-dimensional dependence structure, the number of estimate parameters is up to $n(n-1)$.

In summary, the main problems of modelling the high dimensional dependence includes: (1) the dependence structure is asymmetric; (2) each individual variable has fat tail, and it does not follow normal distribution; (3) typical model cannot capture both lower and upper tail dependence, and; (4) the structure is susceptible to the curse of dimensionality.

1.2 Research Issues

Based on the aforementioned current research limitations, this thesis is of particular interest in the following research issues:

1.2.1 RI 1: High Dimensional Data

The existing research in terms of copula-based dependence modelling mainly focuses on bivariate studies. However, the cases in the real world often refer to high dimensional data, which is much more complicated than bivariate data. For example, in the portfolio optimisation theory, it is advised that the investor should buy a large number of assets, in order to minimise the risk. It is obvious that when the investor only buys stock and currency, there is a high probability of losing money during an economic recession since there is a strong correlation between the stock market and foreign exchange market. If the money is invested in several markets, such as stock, currency, property, derivatives or gold, the risk of losing money can be greatly reduced.

1.2.2 RI 2: Dependence Structure Assumptions

The existing research in terms of dependence structure generally has some assumptions, which are made on the domain or prior knowledge. For example, the canonical vine autoregressive model assumes that the stock returns of different markets are independent (The details of the model and its assumption can be found in Chapter 3, Section 3.2). The reason is that the current research methodologies cannot handle the dependence structure since it is too complex without any assumption in terms of the high dimensional dependence structure. From the data-driven perspective, it is not correct to make assumptions on the dependence structure. Prior knowledge may not reflect the true distribution of data, and it may result in misunderstanding. If the assumption of structure is based on incorrect prior knowledge, the dependence structure will not reflect the true distribution of data. Typically, a large number of models assume the distribution follows the Gaussian distribution.

However, it has been demonstrated that these Gaussian assmuption models are not appropriate in the real world (Abdous et al. 2005, Fang et al. 2002) since a large number of data do not follow the Gaussian distribution.

1.2.3 RI 3: Dependence Structure Truncation and Optimisation

It is very important to truncate, simplify, or optimise the high dimensional dependence structure. For high dimensional data, the complexity of the dependence structure increases exponentially as the number of dimensions grow. For example, the number of nodes in a canonical vine dependence structure is $n(n-1)/2$, where n is the number in terms of the dimension of data. When these nodes are related to two-parametric bivariate copulas, the number of estimate parameters is $n(n-1)$, which may cause a very large computational burden.

1.2.4 RI 4: High Dimensional Dependence Evaluation

The quantitative research targets dependence modelling and evaluation via exposing the log-likelihood, statistical significance, mutual information or the relationships among variables. For the dependence modelling applications, the focus is not only on the methodologies and techniques targets, but also on industrial application and evaluation. For a case study in the real world, it is much more important to implement the industrial evaluation, in order to ensure the models are being used appropriately in the real world scenarios.

1.3 Research Contributions

To address the above research issues related to high dimensional dependence modelling, this thesis makes the specific contributions below:

- It proposes the algorithm to construct the canonical vine by using a

partial correlation via a bottom-up strategy, in order to model the dependence with a flexible dependence assumption (Chapter 3, RI 2);

- It proposes the truncated and simplified algorithms for a partial correlation-based canonical vine, which can significantly reduce the number of parameters and estimate time. It ensures the partial correlation-based canonical vine copula model can be effectively applied in high dimensional financial data(Chapter 3, RI 3) ;

- It proposes the algorithm to construct the regular vine by using partial correlation via a top-bottom strategy (Chapter 4, RI 2);

- It applies two-parametric bivariate copulas in the partial correlation-based regular vine, and then analyse the tail dependence trends of cross-country markets in the dynamic periods (Chapter 4, RI 4);

- It proposes the algorithm to construct the regular vine by using partial correlation via a bottom-up strategy (Chapter 5, RI 2);

- It proposes the truncated algorithm for a partial correlation-based regular vine, which can effectively reduce the number of parameters and decrease the parameter estimate time (Chapter 5, RI 3);

- It applies these models in high-dimensional data in various markets, such as European stock markets (Chapter 3, RI 1 and RI 4), cross-country stock markets (Chapter 4, RI 1) and foreign exchange rate markets (Chapter 5, RI 1).

- It evaluates these models through Value at Risk, which is widely used in the financial area (Chapter 3, 4 and 5, RI 4);

- It builds the utility function via a partial correlation-based canonical vine, to address the issue of assets allocation with high dimensional financial variables in the European and United States stock markets (Chapter 6, RI 1, 2, 3 and 4).

In summary, the above work forms a comprehensive solution, which is a collection of copula-based high dimensional dependence modelling approaches to address the major research issues listed in Section 1.2.

1.4 Thesis Structure

The thesis is structured as follows:

Chapter 2 provides the literature review in relation to the definition of copula and its bivariate families and various copula-based models. The parameter estimate methods and related inference are reviewed. This is followed by a review of the graphical theory and the regular vine. The relationship between the graphical probability model and the regular vine are also reviewed. In addition, the time series theory, partial correlation and model comparison tests, which constitute the foundation of this thesis, are reviewed. Finally, the applications of the copula based-model in the financial area are reviewed.

Chapter 3 presents the partial correlation-based canonical vine copula model. The truncation and simplification methods of the canonical vine are proposed in the chapter to greatly reduce the number of parameters and estimate time. The model is then applied to construct and analyse dependence structures of European stocks as case studies. Its performance is evaluated by measuring a portfolio of Value at Risk, a widely used risk management measure. In comparison to a very recent canonical vine model (canonical vine autoregressive), the experimental results demonstrate that the model has much better quality in terms of the Value at Risk, and it provides insightful knowledge for investors to control and reduce the aggregation risk of the portfolio.

Chapter 4 presents the partial correlation-based regular vine via the top-bottom strategy. This model is able to capture the asymmetric characteristics among multiple variables by using two-parametric copula with flexible lower and upper tail dependence. The method is tested on a cross-country stock market data set to analyse the asymmetry and tail dependence. The

high prediction performance is examined by the Value at Risk, which is a commonly adopted evaluation measure in the financial market.

Chapter 5 proposes a novel truncated partial correlation-based regular vine copula model which can capture more-flexible dependence structures without making pre-assumptions about the data. Specifically, the model employs a partial correlation-based method to build the dependence structures. It can identify important dependencies and information among high dimensional variables, truncating the irrelevant information to reduce the parameter estimation time. The model is then applied to construct the dependence structures of 17 currency markets over 17 years as a case study. The model's in-sample performance is evaluated via a standard model selection criteria Vuong test, and the out-of-sample performance is evaluated by Value at Risk, a widely used industrial benchmark. The extensive experiment results show that the model and its intrinsic design significantly outperform industry baselines, and provide financially interpretable knowledge and profound insights into the high dimensional dependence structures of complex financial variables.

Chapter 6 proposes the new utility function, in which the partial correlation-based canonical vine is used to produce the complex joint distribution of asset returns. The utility function is used to determine the optimal allocation of the assets. The importance of using the asymmetries information is assessed by comparing the performance of a portfolio based on the mean-variance criteria and that of a portfolio based on the truncated partial canonical vine. The results show that the investors using the forecasts of these asymmetries can make better portfolio decisions than those who ignore the asymmetries information.

Chapter 7 concludes the thesis and outlines the scope for future work.

Figure 1.2 shows the research profile of this thesis.

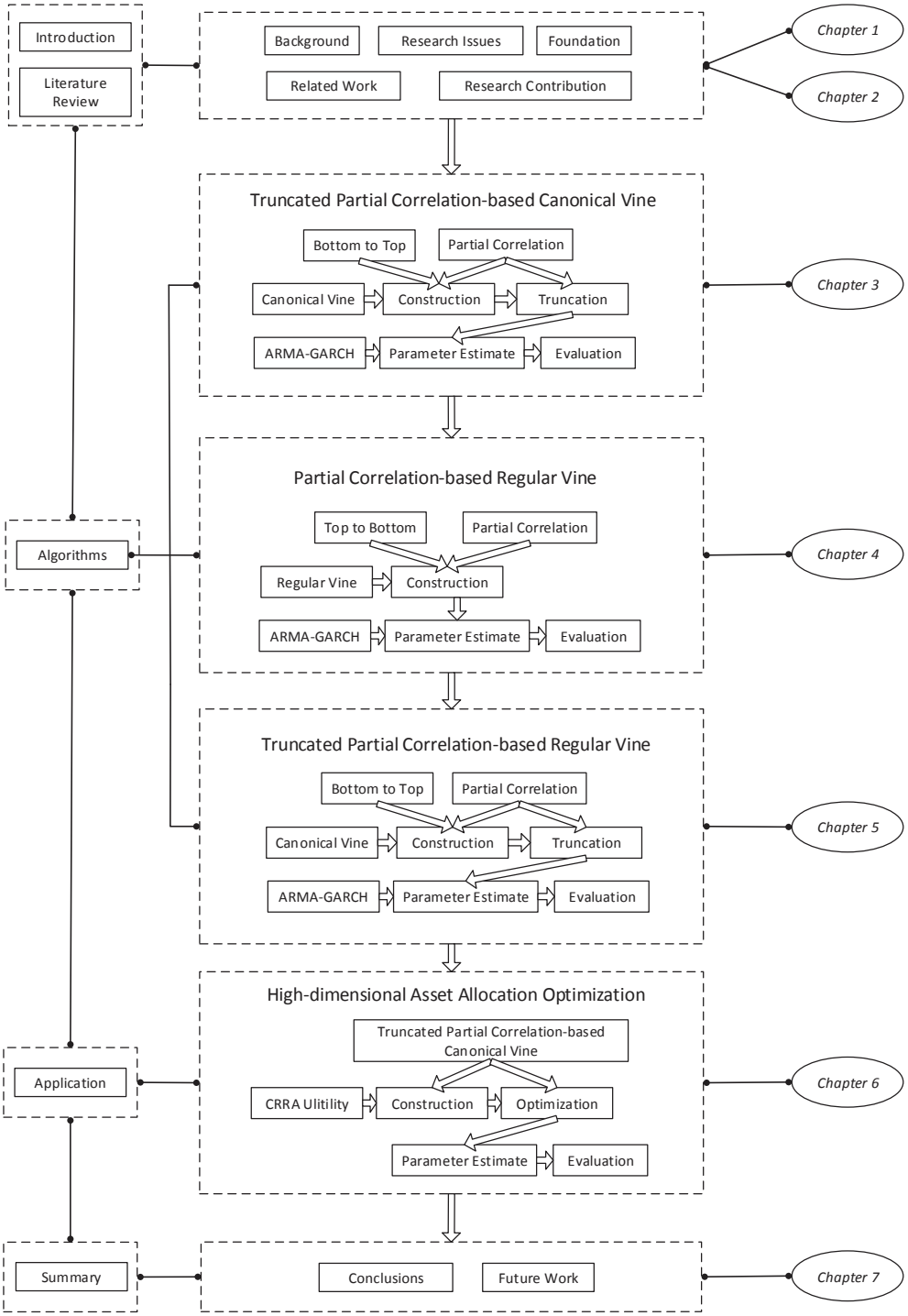

Figure 1.2: The profile of work in this thesis

Chapter 2

Literature Review and Foundation

This chapter reviews the related work of copula, including the definition of copula, the bivariate copula families, copula-based models, the regular vine theory, the foundation theory and applications in the financial area. The definition of copula and its families are introduced in Section 2.1, and then various copula-based models are reviewed in Section 2.2. Section 2.3 presents the multivariate copula models. The graphical theory and regular vine models are presented in Section 2.4. Section 2.5, 2.6 and 2.7 introduces the foundation employed throughout the thesis, including the partial correlation, time series analysis and model comparison tests. Section 2.8 presents the applications in the financial area. Finally, Section 2.9 summarises this chapter.

2.1 Copula

As a first introduction to copula, an important result in the literature is recalled. Let $\mathbf{x} = (x_1, ..., x_n)$ be n-dimensional random variables with joint density function $f(x_1, ..., x_n)$ and cumulative distribution function (cdf) $F(x_1, ..., x_n)$. Further, let $F_1(x_1), ..., F_n(x_n)$ be the corresponding (strictly

increasing, continuous) marginal distributions of $x_1, ..., x_n$. A copula, which is itself a multivariate cdf, is the link that connects the marginal distribution to a cdf.

Definition 2.1 (Copula) *An n-dimensional copula C is a multivariate cumulative distribution function on* $[0,1]^n$*,*

$$C \ : \ [0,1]^n \to [0,1],$$

with univariate uniform distributed margins.

According to Sklar's theorem (Sklar 1959), there exists a unique copula C so that:

$$F(x_1, ..., x_n) = C(F_1(x_1), ..., F_n(x_n)) \tag{2.1}$$

In fact, it assumes continuous marginals $F_1, ..., F_n$, C is unique. This is because Equation (2.1) can be converted to

$$C(u_1, ..., u_n) = F(F_1^{-1}(u_1), ..., F_n^{-1}(u_n)) \tag{2.2}$$

where $F_1^{-1}, ..., F_n^{-1}$ are the inverse distribution functions of the margins. The above representation of the joint cumulative distribution function (cdf) implies the following representation of the joint probability distribution function (pdf):

$$\begin{aligned} f(x_1, ..., x_n) &= c(F_1(x_1), ..., F_n(x_n)) \cdot \prod_{i=1}^{n} f_i(x_i) \\ where \quad c(x_1, ..., x_n) &= \frac{\partial C(x_1, ..., x_n)}{\partial x_1, ..., \partial x_n} \end{aligned} \tag{2.3}$$

where c is the copula density, and $(f_1(x_1), ..., f_n(x_n)$ and f are the corresponding marginal distributions of $(F_1(x_1), ..., F_n(x_n))$ and the joint cumulative distribution of F respectively. The usefulness of the above representation is the converse version Sklar's theorem: given any distribution $(F_1, ..., F_n)$

and any copula C, the function F, which is defined by Equation (2.1), defines a valid joint distribution with marginal distributions. The marginal distributions and copula function do not need belong to same family of distribution, which indicates that they can be symmetric or skewed, continuous or discrete, fat-tailed or thin-tailed. For example, it can combine a Gaussian distributed variable with an exponentially distribution variable via a Clayton copula, and obtain a valid joint bivariate distribution. The ability of copula to combine marginal distributions with a copula function allows research to focus on the work of modelling the dependence structure, leaving the task of marginal distributions.

One thing which should be noted is that in Equation (2.3) when all $(x_1, ..., x_n)$ are all independent random variables, then f is equal to the product of the marginal densities. In that case, the copula c is equal to unity across its support.

2.1.1 Elliptical Copulas

Elliptical copulas are copulas generated by elliptical distributions using the inversion Equation (2.2). The detail can be found in (Owen & Rabinovitch 1983).

Definition 2.2 (Elliptical Distribution) *The n-dimensional random vector $\boldsymbol{x}$ has an elliptical distribution iff the density function $f(\boldsymbol{x})$ has the representation*

$$f(\boldsymbol{x}) = c_n|\Sigma|^{-\frac{1}{2}} g((\boldsymbol{x}-\boldsymbol{\mu})'\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})) \tag{2.4}$$

with some constant $c_n \in \mathbb{R}$, mean vector $\boldsymbol{\mu} \in \mathbb{R}^{n\times n}$ symmetric positive definite, and some function g is independent of n.

The most famous example of an elliptical distribution is the multivariate normal distribution with $c_n = (2\pi)^{-n/2}$ and $g(s) = e^{-\frac{1}{2}s}\ \forall s > 0$. According

Equation (2.2), the multivariate Gaussian copula is given as follows:

$$c(\mathbf{u}) = \Phi_R(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_n)) \tag{2.5}$$

where Φ and Φ^{-1} is denoted as the standard normal cdf and the inverse of the standard normal cdf respectively. Φ_R is denoted as the multivariate standard normal cdf with symmetric positive definite correlation matrix $R \in [-1, 1]^n$. Thus, the density is given by:

$$c(u) = |R|^{-\frac{1}{2}} e^{\frac{1}{2}\mathbf{x}'(I_n - R^{-1}\mathbf{x})} \tag{2.6}$$

where $\mathbf{x} = (x_1, ..., x_n)' \in \mathbb{N}$ with $x_i = \Phi^{-1}(u_i), i = 1, ..., n$

In addition, another widely used elliptical copula is the multivariate t copula, which is presented by Demarta and McNeil (2005). It is derived from the multivariate t distribution with constant $c_n = (\pi n)^{-n/2}\Gamma(\frac{\nu+n}{2})/\Gamma(\frac{\nu}{2}))$ and $g(s) = (1 + \frac{s}{\nu})^{-(\nu+n)/2}\ \forall s > 0$. Thus, the multivariate t copula is defined as follows:

$$C(\mathbf{u}) = t_{R,\nu}(t_\nu^{-1}(u_1), ..., t_\nu^{-1}(u_n)) \tag{2.7}$$

where $t_{R,\nu}$ is denoted as the cdf of the multivariate standard t distribution with correlation matrix $R \in [-1, 1]^n$ and $\nu > 0$ degree of freedom, and t_ν^{-1} is the inverse of the cdf t_ν of the unvariate standard t distribution with ν degrees of freedom.

2.1.2 Archimedean Copulas

Archimedean copulas have three advantages, which lead to a large number of references on the copulas. These advantages consist of: (1) the ease with which they can be constructed; (2) the great variety of families of copulas which belong to this class; and (3) the many nice properties, such as symmetry and associativity, possessed by the members of the class. The detail refers to Nelsen (1999).

Theorem 2.1 *Let $\varphi : [0,1] \to [0,\infty]$ be a continuous strictly decreasing function such that $\varphi(0) = \infty$ and $\varphi(1) = 0$ and let φ^{-1} denote the inverse of φ such that it is completely monotonic. Then*

$$C(\boldsymbol{u}) = \varphi^{-1}(\varphi(u_1) + ... + \varphi(u_n)) \tag{2.8}$$

is a copula.

The copula C in Theorem 2.1 is called the n-dimensional Archimedean copula with generator φ. In the bivariate case, the assumption of complete monotonicity and $\varphi(0) = \infty$ are not necessary, when the pseudo-inverse $\varphi^{[-1]}$ of a convex generator φ is considered instead of φ^{-1}. The pseudo-inverse is defined by:

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t), & 0 \leq t \leq \varphi(0), \\ 0, & \varphi(0) \leq t \leq \infty. \end{cases}$$

Theorem 2.1 provides an easy method to construct multivariate copulas of arbitrary dimension. However, since most commonly used generators depend on one or at the most two parameters, there are only one or two parameters to model the dependency of n random variables, which are the limiting factors of Archimedean copulas. The limitation results in these Archimedean copulas are very appealing in the bivariate case.

2.1.3 Bivariate Copula Families

The bivariate copula families are introduced as follow. The detail can be be found in (Joe 1997, Nelsen 1999).

Gaussian Copula

The Gaussian copula belongs to the elliptical copula family. The bivariate Gaussian copula with correlation parameter $\rho \in (-1, 1)$ is defined as

$$C(u_1, u_2) = \Phi_\rho(\Phi^{-1}(u_1), \Phi^{-1}(u_2)), \tag{2.9}$$

where Φ_ρ denotes the bivariate standard normal cdf with correlation ρ. The correspond density is given as follows:

$$c(u_1, u_2) = \frac{1}{\sqrt{1-\rho^2}} e^{-\frac{\rho^2(x_1^2+x_2^2)-2\rho x_1 x_2}{2(1-\rho^2)}}, \tag{2.10}$$

where $x_1 = \Phi^{-1}(u_1)$ and $x_2 = \Phi^{-1}(u_2)$. The bivariate Gaussian copula belongs to the class of elliptical copulas. For $\rho \to 1$, the Gaussian copula exhibits complete positive, and for $\rho \to -1$ complete negative dependence.

t Copula

The t copula belongs to the elliptical copula family The t copula is the two-parametric copula (Demarta & McNeil 2005).The density of the bivariate t copula with parameters $\rho \in (-1, 1)$ and $\nu > 0$ is given by:

$$c(u_1, u_2) = \frac{2}{2\pi dt_\nu(x_1) dt_\nu(x_1)\sqrt{1-\rho^2}} (1 + \frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{\nu(1-\rho^2)})^{-\frac{\nu+2}{2}}, \tag{2.11}$$

$$\text{where} \quad dt_\nu(x_1) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\pi\nu}} (1 + \frac{\pi^2}{\nu})^{-\frac{\nu+1}{2}}$$

is the density of an univariate t distribution with ν degree of freedom with $x_1 = t_\nu^{-1}(u_1)$ and $x_2 = t_\nu^{-1}(u_2)$, where $t_\nu^{-1}(\cdot)$ denotes the inverse distribution function of an univariate t distribution with ν degrees of freedom. For $\rho \to 1$ the t copula exhibits complete positive and for $\rho \to 1$ complete negative dependence.

Clayton Copula

The Clayton copula (see (Nelsen 1999)) is an Archimedean copula with generator $\varphi(t) = \frac{1}{\theta}(t^\theta - 1)$ and it is therefore given by:

$$C(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}} \tag{2.12}$$

and the density

$$c(u_1, u_2) = (1+\theta)(u_1 u_2)^{-1-\theta}(u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}-2} \tag{2.13}$$

where $\theta \in (-1, \infty) \backslash \{0\}$. For $\theta \to 0$, the Clayton copula exhibits independence and for $\theta \to \infty$ complete positive dependence.

Gumbel Copula

The Gumbel copula (see (Nelsen 1999)) is an Archimedean copula with generator $\varphi(t) = (-log\, t)^\theta$, which is given by:

$$C(u_1, u_2) = e^{-((-log\, u_1)^\theta + (-log\, u_2)^\theta)^{\frac{1}{\theta}}}, \tag{2.14}$$

where $\theta \geq 1$. The corresponding density is given by:

$$\begin{aligned} c(u_1, u_2) = & \frac{C(u_1, u_2)}{u_1 u_2} \times \frac{(log\, u_1 log\, u_2)^{\theta-1}}{((-log\, u_1)^\theta + (-log\, u_2)^\theta)^{2-\frac{1}{\theta}}} \\ & \times [((-log\, u_1)^\theta + (-log\, u_2)^\theta)^{\frac{1}{\theta}} + \theta - 1] \end{aligned} \tag{2.15}$$

where $\theta \in [1, \infty)$. The Gumbel copula is only applicable to model positive dependence. For $\theta = 1$, the Gumbel copula exhibits independence and for $\theta \to \infty$ complete positive dependence.

Frank Copula

The Frank copula (see (Nelsen 1999)) is an Archimedean copula with generator $\varphi(t) = -log[\frac{e^{-\theta t}-1}{e^{-\theta}-1}]$, and the corresponding distribution function is given by:

$$C(u_1, u_2) = -\frac{1}{\theta} log[1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1}], \tag{2.16}$$

where $\theta \in (-\infty, \infty) \backslash \{0\}$. The density is given by:

$$c(u_1, u_2) = \theta(e^{-\theta} - 1) \frac{e^{-\theta(u_1+u_2)}}{[e^{-\theta} - 1 + (e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)]^2} \tag{2.17}$$

For $\theta \to \infty$, the Frank copula exhibits complete positive, and for $\theta \to -\infty$ complete negative dependence.

Joe Copula

The Joe copula (see (Nelsen 1999)) is another Archimedean copula with generator $\varphi(t) = -log[1-(1-t)^\theta]$, and it is therefore given by:

$$C(u_1, u_2) = 1 - [(1-u_1)^\theta + (1-u_2)^\theta - (1-u_1)^\theta(1-u_2)^\theta]^{\frac{1}{\theta}}, \tag{2.18}$$

with $\theta > 1$, and the following density:

$$\begin{aligned} c(u_1, u_2) =& [(1-u_1)^\theta + (1-u_2)^\theta - (1-u_1)^\theta(1-u_2)^\theta]^{\frac{1}{\theta}-2}(1-u_1)^{\theta-1}(1-u_2)^{\theta-1} \\ & \times [\theta - 1 + (1-u_1)^\theta + (1-u_2)^\theta - (1-u_1)^\theta(1-u_2)^\theta] \end{aligned} \tag{2.19}$$

For $\theta \to \infty$, the Joe copula exhibits negative dependence.

Clayton-Gumbel Copula (BB1)

The Clayton-Gumbel copula is a two-parametric Archimedean copula and can regarded as a generalisation of the one-parametric Clayton and Gumbel families, which have been studied by Joe (1997). To avoid confusion, it can simply be called the BB1 copula. The generator is $\varphi(t) = (t^{-\theta} - 1)^\delta$, which yields:

$$C(u_1, u_2) = [1 + [(u_1^{-\theta} - 1)^\delta + (u_2^{-\theta} - 1)^\delta]^{\frac{1}{\delta}}]^{-\frac{1}{\theta}} \tag{2.20}$$

where $\theta > 0$ and $\delta \geq 1$. The corresponding density is given by:

$$\begin{aligned} c(u_1, u_2) = & \left\{1 + [(u_1^{-\theta} - 1)^\delta + (u_2^{-\theta} - 1)^\delta]^{\frac{1}{\delta}}\right\}^{-\frac{1}{\theta}-2} \\ & \times [(u_1^{-\theta} - 1)^\delta + (u_2^{-\theta} - 1)^\delta]^{\frac{2}{\delta}-2} \\ & \times \left\{\theta\delta + 1 + \theta(\delta - 1)[(u_1^{-\theta} - 1)^\delta + (u_2^{-\theta} - 1)^\delta]^{-\frac{1}{\delta}}\right\} \\ & \times (u_1^{-\theta} - 1)^{\delta-1} u_1^{-\theta-1} (u_2^{-\theta} - 1)^{\delta-1} u_2^{-\theta-1} \end{aligned} \tag{2.21}$$

The BB1 copula is similar to Clayton and Gumbel copulas, and positive dependence is onbtained for $\theta \to \infty$ and $\delta \to \infty$. BB1 is Clayton copula when $\delta = 1$, and Gumbel copula for $\theta \to 0$.

Joe-Clayton Copula (BB7)

Similar to the BB1 copula, the Joe-Clayton copula is a two-parametric generalisation of the corresponding one-parametric copula families, which have been studied by Joe (1997) . It can simply be called the BB7 copula. The generator is defined as $\varphi(t) = [1-(1-t)^{\theta}]^{-\delta} - 1$, and the copula distribution function is given by:

$$C(u_1, u_2) = 1 - [1 - [(1-(1-u_1)^{\theta})^{-\delta} + (1-(1-u_2)^{\theta})^{-\delta}]^{-\frac{1}{\delta}}]^{\frac{1}{\theta}} \tag{2.22}$$

for $\theta \geq 1$ and $\delta > 0$. The corresponding density is

$$\begin{aligned}
c(u_1, u_2) =& (-\frac{1}{\theta})(\frac{1}{\delta} - 1) \cdot h^{\frac{1}{\theta}-2} du_1 h \cdot du_2 h \; - \; \frac{1}{\theta} \cdot h^{\frac{1}{\theta}-1} du_1 u_2 h \\
&\text{where} \\
h =& 1 - ((1-(1-u_1)^{\theta})^{-\delta} - (1-(1-u_2)^{\theta})^{-\delta} - 1)^{\frac{1}{\delta}} \\
du_1 h =& -\theta((1-(1-u_1)^{\theta})^{-\delta} - (1-(1-u_2)^{-\delta} - 1)^{\frac{1}{\delta}-1} \\
& (1-(1-u_1)^{\theta})^{-\delta-1}(1-u_1)^{\theta-1} \\
du_2 h =& -\theta((1-(1-u_1)^{\theta})^{-\delta} - (1-(1-u_2)^{-\delta} - 1)^{\frac{1}{\delta}-1} \\
& (1-(1-u_2)^{\theta})^{-\delta-1}(1-u_2)^{\theta-1} \\
du_1 u_2 h =& \frac{1}{\delta}(-\frac{1}{\delta} - 1)((1-(1-u_1)^{\theta})^{-\delta-1}(1-u_1)^{\theta-1} \\
& - (1-(1-u_2)^{\theta})^{-\delta} - 1)^{\frac{1}{\delta}-2} du_1 S \cdot du_2 S \\
du_1 S =& -\theta\delta(1-(1-u_1)^{\theta})^{-\delta-1}(1-u_1)^{\theta-1} \\
du_2 S =& -\theta\delta(1-(1-u_2)^{\theta})^{-\delta-1}(1-u_2)^{\theta-1}
\end{aligned} \tag{2.23}$$

It is similar to Joe and Clayton copulas, the positive dependence is obtained for $\theta \to \infty$ and $\delta \to \infty$. Furthermore, the BB7 copula is Clayton copula when $\theta = 1$, and Joe copula for $\delta \to 0$.

2.1.4 The Tail Dependence of Copulas

The tail dependence and measurements are reviewed in this section. It is very simple and easy to measure the dependence structure via the linear correlation coefficient when assuming normality. However, if more flexible models for dependence structure are considered, other dependence measures and methods may be considered. Joe (1997) and Nelsen (1999) introduced a large number of dependence measures and methods. including: Spearman's rank correlation, Kendall's tau (detail can be found in Appendix A) and tail dependence. The empirical analysis are often familiar with the sample version based on ranked data. For instance, the rank correlation is useful for providing information on the sign of the dependence between variables. Suppose X_1 and X_2 are continuous-valued, and X follows the copula model suggested by Equation (2.1). Let random vector $(U_1, U_2) := (F_1(x_1), F_2(X_2))$ with standard uniform marginal distribution. In many cases, high values of X_1 and X_2 exhibit different levels, or even directions of dependence that lower the values of X_1 and X_2. The phenomenon is called asymmetric dependence. Nelsen (1999) suggested that if X_1 and X_2 are continuous-valued, the dependence properties between U_1 and U_2, where U_1 and U_2 are the corresponding distribution of X_1 and X_2 respectively. In that case, measures of asymmetric dependence are often based on the conditional probabilities, which are given as follows:

$$\begin{aligned} \lambda_{1,2}^{up}(\alpha) =& P(U_1 > \alpha | U_2 > \alpha) \\ \lambda_{1,2}^{low}(\alpha) =& P(U_1 < \alpha | U_2 < \alpha) \end{aligned} \tag{2.24}$$

where $0 < \alpha < 1$. Then, the limits of Equation (2.24) are called the upper and lower tail dependencies (Joe 1997), which are given by

$$\begin{aligned} \lambda_{1,2}^{up}(\alpha) &= \lim_{\alpha \to 1} \lambda_{1,2}^{up}(\alpha) \\ \lambda_{1,2}^{low}(\alpha) &= \lim_{\alpha \to 0} \lambda_{1,2}^{low}(\alpha) \end{aligned} \tag{2.25}$$

Tail dependence, which is a measure of dependence between extreme events, is a key feature to tell various copula families. The extremal dependence of a multivariate distribution F can be described by the various tail dependence parameters of its copula C. Generally, for random vector $(U_1, ..., U_n) := (F_1(x_1), ..., F_n(X_n))$ with standard uniform marginal distribution, the lower and upper tail dependence are defined as follows:

$$\begin{aligned} \lambda_L &= \lim_{u \to 0} Pr\{U_1 \leq u, ..., U_n \leq u \mid U_n \leq u\} \\ &= \lim_{u \to 0} \frac{C(u, ..., u)}{u} \\ \lambda_U &= \lim_{u \to 0} Pr\{U_1 > 1 - u, ..., U_n > 1 - u \mid U_n > 1 - u\} \\ &= \lim_{u \to 0} \frac{\overline{C}(1 - u, ..., 1 - u)}{u} \end{aligned} \tag{2.26}$$

where $\overline{C}$ is denoted as the survival function of C. If λ_U exists and $\lambda_U \in (0, 1]$, then copula C has upper tail dependence, and no upper tail dependence if $\lambda_U = 0$. Similarly, If λ_L exists and $\lambda_L \in (0, 1]$, then copula C has lower tail dependence , and no lower tail dependence if $\lambda_L = 0$. In addition, tail dependence is one of the most important properties that differentiate the different copula families. For example, the Gumbel copula has only upper tail dependence coefficient, and the Gaussian copula does not allow for any tail dependence coefficient.

The lower and upper tail dependencies for the above bivariate copulas is given in Table 2.1.

2.2 Copula-based Models

This section reviews the copula-based models which are proposed in the literature. The majority of applications of the copula-based model are in the financial fields, so this section focuses on such models. A large number of copula-based models are reviewed firstly, and then the parameter estimate of copula-based models, inference and the goodness of fitting tests are reviewed.

Table 2.1: The Tail Dependence of the Bivariate Copula Family

	Lower Tail Dependence	Upper Tail Dependence
Gaussian	-	-
t	$2t_{\nu+1}(-\sqrt{\nu+1}\sqrt{\frac{1-\rho}{1+\rho}})$	$2t_{\nu+1}(-\sqrt{\nu+1}\sqrt{\frac{1-\rho}{1+\rho}})$
Clayton	$2^{-1/\theta}$	-
Gumbel	-	$2^{-1/\theta}$
Frank	-	-
Joe	-	$2-2^{1/\theta}$
BB1	$2^{-1/(\theta\delta)}$	$2-2^{1/\delta}$
BB7	$2-2^{1/\delta}$	$2-2^{1/\theta}$

θ and δ are parameters of the corresponding copula family.

2.2.1 Time-varying Copula Models

The motivation, which is to consider the conditional dependence structure as it varies through time, is that Andersen (2006) provided an abundant amount of evidences that the conditional volatility of the economic time series changes throughout time. Therefore, before specifying a function for a time-varying conditional copula model, it is essential to test for the presence of time-varying dependence. A simple test, which is mentioned in (Patton 2006a), is used for a break in rank correlation at some specified point in the sample. It is used for examining the exchange currency rate. This test is very easy to implement, however, it needs a researcher to have a prior knowledge of when a break in the dependence structure may occur. Andrews (1993) proposed another test to examine the time-varying dependence structure. The second test can be used for a break in the rank correlation coefficient at some unknown date. A variety of statistics test is available for the test. Engle (1982) proposed a *Autoregressive Conditional Heteroscedasticity* (*ARCH*) LM-based test for the time-varying dependence structure test. The *ARCH LM*-based test seeks for autocorrelation in a measure of dependence, which is captured by an autoregressive-type model, rather than

looking for discrete one-time breaks in the dependence structure in the previous two tests. Rémillard (2010) used a test for one-time changes in the copula at some time in the sample period. The break time can be known or unknown in his framework. The results show that the test statistics are similar to Kolmogorov-Smirnov statistics, which are compared with the empirical copula, before or after break date. Gaißer, et al. (2010) considered using a test for a change in the dependence structure by looking for a change in Hoeffding's ϕ^2 dependence measure.

Observation Driven Models

The observation driven models, in which the copula functional form is fixed and its parameters are set to vary over time, are proposed in the literature, such as Patton (2004, 2006a) Jondeau and Rockinger (2006*a*), Ausin and Lopes (2010), Christoffersen, et al. (2012) and Creal, et al. (2013).

Patton (2004) proposed time-varying copula models to capture dynamic moments (up to the fourth order), and then constructed a time-varying dependence structure that allows for a different dependence structure in various financial markets. Patton (2006a) strengthened time-varying copula model via the conditional copula theory. The time-varying copulas and structural breaks are combined in this model. It allows for analysing the dynamic conditional dependence structure, and using conditional copulas for multivariate density modelling. Jondeau and Rockinger (2006*a*) proposed a *Generalized Autoregressive Conditional Heteroscedasticity* ($GARCH$)-copula based methodology to model the dynamic dependence between financial returns. It models a time-varying correlation for Gaussian and t copulas in three different ways, including *Dynamic Conditional Correlations* (DCC) correlations and regime switching correlations. Since estimating the univariate distributions at first and then the joining distribution, the parameters in dependence can easily be rendered conditional and time varying. Ausin and Lopes (2010) developed a Bayesian methodology to make inference and predication in the $GARCH$-copula based model. A one-step Bayseian procedure

is used to ensure that all parameters are estimated simultaneously using the whole likelihood function. Then, dynamic copulas are modelled to capture the time evolution of the dependence structure. Christoffersen, et al. (2012) developed a *Dynamic Asymmetric Copula* (*DAC*) model which allows for asymmetric and dynamic tail dependence. It is a generalised model based on the flexible *DCC* model of Engle (2002) and Tse and Tsui (2002). The *DAC* model allows for asymmetric trends in dependence structure and deviation from multivariate normality. It then models the joint distribution using dynamic copulas to capture the nonlinear dependence in financial markets. Creal, et al. (2013) proposed a class of observation drive time series models, which is called *Generalized Autoregressive Score*. The scaled score of likelihood functions is used for updating parameters over time. The framework is illustrated by introducing a dynamic copula function for the multivariate point processes with time-varying parameters.

Stochastic Copula Models

Hafner and Manner (2012) developed dynamic copula models, in which the copula parameters are driven by an independent stochastic process rather than by the observation as in the *DCC* model. This dynamic copula model consists of the Gaussian copula with stochastic correlation process, thus, it can be viewed as a generalisation of multivariate stochastic volatility models which are introduced by Kim et al. (1998). Manner and Segers (2011) produced significant research regarding the correlation mixtures of elliptical copulas (e.g. Gaussian and t copulas) in dynamic copula models. They found that both penultimate and asymptotic tail dependence are much larger than for ordinary elliptical copulas with the same unconditional correlation.

Regime Switch Copula Models

Another point of interest in relation to time-varying copula models is that both the degree and the type of dependence varies over time. These models allow for a number of states, and each is characterised by a different copula

that may be from the same family but still allow for different parameters.

Rodriguez (2007) presented time-varying copula models, in which the conditional copula allows the functional form of the copula to change over time. The dependence structure is modeled by a mixture of copulas with parameters which vary over time according to a Markov switching model. Two classes of copulas are studied, one is a finite mixture of Frank, Gumber and Clayton copulas, which ensures that the asymmetries in tail dependence are captured, and the other is the bivariate t copula, which exhibits symmetric tail dependence. Okimoto (2008), Chollete, etal., (2009) and Garcia and Tsafack (2011) made similar studies by building a regime switching copula model and Markov switching model. Markwat, et al. (2009) developed a framework that allows for both changes in the structure of the degree of dependence. It models the changes in the degree and structure as switches between regimes. The regime processes are latent and follow the first-order Markov chains.

Other Copula Models

Giacomini, et al. (2009) proposed copula models with adaptively estimated time varying parameters, which are free from the usual normality assumption. When employing a semi-parametric approach, the time-varying copula parameters are selected locally via an adaptive estimation under the assumption of local homogeneity. The adaptive estimation does not make a structure break, since it can ensure parameters vary smoothly from one value to another.

Guegan and Zhang (2010) developed the time-varying copula model, in which copula functions vary over time. Accordly, they compare the parametric copula to a nonparametric estimation of copula density by using the goodness-of-fit test. In such a way, they can measure the variance of the dependence structure. Then, due to changes in the variance of the dependence structure, the model can determine whether both copula families and parameters change or whether it is only the parameters that change.

Dias and Embrechts (2010) presented a time-varying copula model. They introduced a dynamic specification for the correlation between two variables which is time-varying and modelled independently from marginal distribution by using the Fisher transformation.

Harvey (2010) proposed some statistical procedures for tracking copula probabilities over time and investigated how these procedures changed with real data. Since the distribution function of copula is not specified, the parameters are estimated by using filters which are designed for a binary time series. Busetti and Harvey (2011) did a further study employing the same method. They developed tests for changes in different part of copulas, as well as overall tests for changing dependence. These test statistics are constructed from time series of indicator variables and their asymptotic distribution, which have power against breaks at unknown points as well as against gradual changes.

Hafner and Reznikova (2010) presented a semi-parame-tric time-varying copula model, where the marginal distribution is specified as the parametric $GARCH$-type process and the dependence parameters of the copula are allowed to change through time in a nonparametric way.

2.2.2 Estimate and Inference

In the subsection, the estimate procedures of copula-based models are reviewed. It focuses on the continuous random variables. The assumption of continuity is not always required, however, it simplifies some of the presentation. At first, the *Maximum Likelihood Estimator* (MLE) is described as follows:

Let $\{x_{1t}, x_{2t}, ..., x_{nt}\}$ be the data matrix. According to Equation (2.3), the expression for the log-likelihood functions is given by:

$$l(\boldsymbol{\theta}) = \sum_{t=1}^{T} logc(F_1(x_{1t}), F_2(x_{2t}), ..., F_n(x_{nt})) + \sum_{t=1}^{T}\sum_{j=1}^{n} logf_j(x_{jt}) \qquad (2.27)$$

where $\boldsymbol{\theta}$ is the set of all parameters of both margins and copulas.

Hence, given a set of marginal probability distribution functions (pdf) and a copula function, the maximum likelihood estimators are obtained by:

$$\hat{\boldsymbol{\theta}}_{MLE} = \arg\max_{\boldsymbol{\theta}\in\Theta} l(\boldsymbol{\theta}) \tag{2.28}$$

The maximum likelihood method may be very computationally intensive, especially in the case of a high dimension, since it is necessary to estimate jointly the parameters of the margins and the parameters of the dependence structure represented by the copula function. Equation (2.27) is composed of two positive terms: one consisting of the copula density and its parameters, and one involving the margins and all parameters of the copula density. According to the assumptions of copula models, the estimation procedure and approach to inference results in parametric, semi-parametric and non-parametric methods.

Parametric Estimate

Joe and Xu (1996) proposed that these set of parameters can be estimated in two stages:

As a first stage, they estimate the margins' parameters $\boldsymbol{\theta}^m$ by performing the estimation of the univariate marginal distributions:

$$\hat{\boldsymbol{\theta}^m} = \arg\max_{\boldsymbol{\theta}^m} \sum_{t=1}^{T}\sum_{j=1}^{n} log f_j(x_{jt};\boldsymbol{\theta}^m) \tag{2.29}$$

The second stage is to perform the estimate of the copula parameters $\boldsymbol{\theta}^c$ given $\boldsymbol{\theta}^m$;

$$\hat{\boldsymbol{\theta}^c} = \arg\max_{\boldsymbol{\theta}^c} \sum_{t=1}^{T} log\, c(F_1(x_{1t}), F_2(x_{2t}), ..., F_n(x_{nt}); \boldsymbol{\theta}^c, \hat{\boldsymbol{\theta}}^m) \tag{2.30}$$

The two-stage estimation method is called *Inference for Margins* (IFM), which is defined as the vector:

$$\hat{\boldsymbol{\theta}}_{IFM} = (\hat{\boldsymbol{\theta}}^m, \hat{\boldsymbol{\theta}}^c)' \tag{2.31}$$

Let l, l_j^m and l^c be denoted as the entire log-likelihood function, the log-likelihood of the jth margin, and the log-likelihood for the copula function respectively. Thus, the IFM estimator is the solution of:

$$(\frac{\partial l_1^m}{\partial \theta_1^m}, \frac{\partial l_2^m}{\partial \theta_2^m}, ..., \frac{\partial l_n^m}{\partial \theta_n^m}, \frac{\partial l^c}{\partial \boldsymbol{\theta}^c}) = \mathbf{0}' \tag{2.32}$$

while the MLE comes from solving:

$$(\frac{\partial l}{\partial \theta_{(1)}^m}, \frac{\partial l}{\partial \theta_{(2)}^m}, ..., \frac{\partial l}{\partial \theta_{(n)}^m}, \frac{\partial l}{\partial \boldsymbol{\theta}^c}) = \mathbf{0}' \tag{2.33}$$

The two estimators, in general, are not equivalent. Generally, IFM is known as the *Multi-stage Maximum Likelihood*. The multi-stage maximum likelihood is asymptotically less efficient than the one-stage MLE. There is a trade off between the computability and asymptotic relative efficiency of estimators. Joe (1997) pointed out that the IFM method is highly efficient compared to the MLE. Joe (2005) implemented simulation studies, and suggested that the loss is not great in many cases. IFM is much easier for the total number of parameters exceeds 15, and it is convenient for a comparison of the various copula with the same set of univariate margins. Patton (2006b) also arrived at similar conclusions.

Semi-parametric Estimate

Semi-parametric copula-based models generally employ a non-parametric model for marginal distribution and a parametric model for copula function. The univariate marginal distributions are estimated non-parametrically in the first stage, such as by the empirical density function or their scaled versions. The estimate of copula parameters is then usually conducted via MLE in the second stage. The method can be described as follows:

1. Estimate the margins using the empirical distribution (without assumption on the parametric form for each of them), i.e. $\hat{F_i}(x_{it}), i = 1, ..., n$;

2. Estimate the copula parameters via MLE;

$$\hat{\boldsymbol{\theta}}^c = \arg\max_{\boldsymbol{\theta}^c} \sum_{t=1}^{T} log\, c(F_1(x_{1t}), F_2(x_{2t}), ..., F_n(x_{nt}); \boldsymbol{\theta}^c) \tag{2.34}$$

In this case, it is called *Canonical Maximum Likelihood* (CML) or *Pseudo Maximum Likelihood* (PML) by Genest et al. (1995) and Klaassen, et al. (1997). Genest et al. (1995) investigated the properties of the semi-parametric method for estimating the dependence parameters in a family of multivariate distribution. The result presents that the estimator obtained by PML, is consistent and asymptotically normal. Shih and Louis (1995) studied the properties of both parametric and semi-parametric estimation procedures for bivariate copulas. They concluded that both parametric and semi-parametric estimators are efficient in terms of independence Further the parameter estimates in the margins have high efficiency and are robust in relation to misspecification of the dependence structure.

Compared with the fully parametric estimation procedure, the copula likelihood function in the semi-parametric method depends on the finite-dimensional parameters $F_1, ..., F_n$ and the marginal distribution parameters. Thus, the standard maximum likelihood methods cannot be applied in this case. Chen and Fan (2006a) proposed a method for estimating the asymptotic covariance matrix. However, the asymptotic normal distribution can be obtained under certain condition. The proof is provided by Chan et al. (2009).

Another result suggested by Chen and Fan (2006b) is that the asymptotic variance of the MLE of the copula parameters depends on the estimate error in the empirical distribution function, but not the estimated parameters in the marginal distribution. It ensures that the model is estimated for conditional means and variance through calculating the standardised residuals and ignoring the estimate error from the mean and variance models.

The semi-parametric estimate method is only applied for constant conditional copula models. Rémillard (2010) showed that for time-varying conditional copula models, the estimate error from the models for the conditional

means and variance affects the asymptotic distribution of the estimators of the copula. In addition, the semi-parametric method is only applied in cases where the marginal distribution of the standardised residuals are estimated nonparametrically. For parametric marginal distribution, the estimate error from models will affect the distribution of the estimator of the copula. Chen (2006) proposed a sieve maximum likelihood estimation procedure for a broad class of semi-parametric multivariate distribution models. The methods achieves full efficiency.

Nonparametric Estimate

In the nonparametric case, the majority of the literature assumes that the conditional distribution is constant and it is estimated via the empirical distribution function. Genest and Rivest (1993) studied the fully non-parametric estimation procedure with the IID data for the Archimedean copulas. While selecting an appropriate Archimedean copula to provide a suitable representation of the dependence structure between two variables, the key to the estimation procedure is a one-dimensional empirical distribution function that can be constructed whether the uniform representation of two variables is Archimedean or not , and this is independent from their marginal distribution. Genest (2011) developed a new rank-based estimator for Archimedean copulas. The estimation procedures are based on a reconstruction of the radial part of the simplex distribution from the Kendall distribution, which is obtained through the multivariate probability integral transformation of data. Genest and Segers (2009) developed a rank-based estimation method for the extreme-value copula with a Pickands dependence function. Scaillet and Fermanian (2002) studied the nonparametric estimation method for time series data. Sancetta and Satchell (2004) introduced a new family of copula based on Bernstein polynomials. The Bernstein copula can be used as an appropriate known or unknown copula. For unknown Bernstein copulas, nonparametric estimation procedures are developed via the empirical Bernstein copula. Ibragimov (2009) studied the characteristics of the higher-

order Markov process via the copula based approach. The nonparametric estimation of the copula function is used for the construction higher-order Markov process with arbitrary one-dimensional margins, which is in line with dependence assumption.

Other Estimate Methods

Some estimation methods, other than maximum likelihood estimation, are used in the literature. Ghoudi and Rémillard (2004) presented a unified treatment of inference procedures based on pseudo-observations in the multivariate setting. The parameters of a given copula family are mapped to a dependence measure. Genest et al. (2011) presented an estimation method based on the inversion of two multivariate extensions of Kendall's tau, which are developed by Kendall and Smith (1940) and Joe (1990) respectively. Then, the performance of the estimators from the inversion of two versions of Kendall's tau are compared in the context of copula models through simulations.

Oh and Patton (2013) proposed an estimation of the parameters of a semiparametric copula-based model via a simulated method of moments (SMM) type approach. Both IID and time series data were considered.

Bayesian estimation of copula models are also considered in the literature. As discussed before, copula-based models usually consider using two-stage estimation procedure, *Inference for Margins* (IFM), for the parametric copula-based model. An alternative to this circumstance is to construct an inference from the joint posterior evaluated in the Monte Carlo manner, with a Gibbs style sampling scheme. This is discussed by Pitt et al. (2006), Silva and Lopes (2008) and Ausin and Lopes (2010).

2.2.3 Goodness-of-Fit Testing and Model Selection

Goodness-of-Fit Tests

The motivation of *Goodness-of-Fit* (*GoF*) tests are that multivariate models constructed via a parametric copula are subject to model misspecification. The *GoF* tests are then applied to find the evidence that the copula is misspecified. For example, the copula function in a model may be different from the unknown true copula. According to different model assumptions, inference for *GoF* tests are divided into two categories, parametric and semi-parametric. The case of non-parametric margins combined with a time-varying conditional copula has not been considered in the literature.

The widely used two *GoF* tests for copula-based models are Kolmogorov-Smirnov (KS) and Cramér-von Mises (CvM) (Rémillard 2010). Both of two tests are based on comparing the fitting copula cumulative distribution function to the empirical copula. The *GoF* tests based on an empirical copula depends on the assumption that the true conditional copula is constant. In other words, it is inappropriate for the time-varying copula. In addition, an alternative is considered in the literature, which is based on Rosenblatt's transformation, which is discussed by Diebold et al. (1999) and Rémillard (2010). In the approach, the data is transformed first, and then KS and CvM are applied to the transformed data. The *GoF* tests based on Rosenblatt's transformation can be used in both constant and time-varying copula-based models.

For fully parametric copula-based models, *GoF* tests are straightforward. Chen (2007) presented a moment-based test for copula, which can detect copula misspecification in various directions. Chen (2011) studied the *GoF* tests for multivariate distributions, with the assumption of moment conditions. Diebold et al. (1999) studied the *GoF* tests via Rosenblatt's transform, which is discussed above.

For semi-parametric copula-based models, Rémillard (2010) used *GoF* tests for time series data, and the results show that the asymptotic distribu-

tion of *GoF* tests are unaffected by the estimation of marginal distribution parameters.

Model Selection Tests

Another issue of *GoF* tests is that of model selection tests. *GoF* tests seek to compare the fitted copula model with the unknown true copula. The model selection tests are employed to identify the best model from a given set of competing specifications. The model selection tests are undertaken using the in-sample data, or using the out-of-sample data. The treatment of the two cases is different, due to the in-sample or out-of-sample data.

For in-sample comparisons, the nested copula models can be accomplished via a likelihood ratio or a Wald test, in which the null hypothesis is that the smaller model is correct and the alterative one is that the bigger model is correct. The in-sample comparison of non-nested, fully parametric copula-based model can be conducted by using the test of Voung (1989) for IID data. Rivers and Vuong (2002) generalised the Vuong test in several important directions. Firstly, they sought to allow for semi-parametric copula-based models. Secondly, they sought to allows for a variety of estimation methods and a variety of evaluation metrics. In addition, their results greatly simplify the test procedure for fully parametric copula-based models. Chen and Fan (2006a) studied the similar tests for the semi-parametric copula-based model, which allows for infinite-dimensional nuisance parameters in model selection. The assumption in tests is that the conditional copula is constant and the corresponding tests of the time-varying case have not been studied in the literature.

For the out-of-sample comparison, the motivation is that it is a very important aspect of the evaluation of economic forecasts, i.e. the predictive ability plays an important role in the evaluation of econometric models. A useful way to compare multivariate density forecasts is to compare their out-of-sample log-likelihood value, which is introduced by Diks (2010). The difference of the out-of-sample evaluation of predictive models depends not

only on whether models are parametric or semi-parametric, but also on the treatment of the parameter estimation error in the forecasts. Giacomini and White (2006) presented a comparison framework for both nested or non-nested copula models. It can even compare the same model in different estimation methods, e.g., one-stage MLE or two-stage IFM. The West (2006) developed a similar approach for the comparison, but it can only be applied to fully parametric moles, or non-nested models.

2.3 Multivariate Copula Models

Most of these copula-based models discussed above are limited to the bivariate cases. To build a high-dimensional copula is then a natural next step. However, it is not an easy task. There are several schemes which seek to construct a high-dimensional copula in the literature, including the multivariate elliptical copula, multivariate Archimedean copula and pair copula construction. In this section, two high-dimensional construction schemes, the multivariate elliptical copula and the multivariate Archimedean copula, are reviewed.

2.3.1 Multivariate Elliptical Copulas

The study of multivariate elliptical multivariate distributions was launched by Fang et al. (2002). They extended the meta-Gaussian family of distributions that was proposed by Krzysztofowicz and Kelly (1996) by using a wider class of continuous distributions with given margins. In place of the normal distribution, elliptically contoured distributions are considered as the basic framework and meta-Gaussian distributions are extended to meta-elliptical distributions.

The t copula, which is studied by Embrechts et al. (1997) and Fang et al. (2002), represents the dependence structure implicit in a multivariate t distribution. Breymann et al. (2003) suggested that the empirical fit of t copula is generally superior to that of the Gaussian copula, which presents the de-

pendence structure of the multivariate normal distribution. The reason is the ability of the t copula to better capture the phenomenon of dependent extreme values, which is often observed in financial data. Demarta and McNeil (2005) studied the t copula through extreme value theory, and then derived the t-EV copulas. They brought together the t copula, particularly with regard to its extremal properties, to present some extensions of the t copula that follow on from the representation of the multivariate t distribution as a mixture of multivariate normals. Nikoloulopoulos et al.(2009) introduced the tail dependence and conditional tail dependence functions to derive the t-EV copulas. This tail dependence functions, are particularly effective for the tail dependence analysis of multivariate t copulas. Two limiting cases are studied: one is that the t-EV copulas yield the Hsler-Reiss distribution as the degree of freedom goes to infinity, and the other is the Marshall-Olkin distribution as the degree of freedom goes to zero.

Abdous et al. (2005) suggested that there are two serious limitations to modelling association with multivariate elliptical copulas. First, no value of r corresponds to independence, unless the copula is actually that of the normal. Second, except in that special multivariate Gaussian case, it is not entirely clear under which circumstances elliptically contoured structures of association meet the concept of positive quadrant dependence. The latter, however, represents a bare minimum in many applications.

2.3.2 Multivariate Archimedean Copulas

The class of the Archimedean copula is a class that has attracted particular interest due to the numerous properties which make them simple to analyse. The most common multivariate extension is the exchangeable multivariate Archimedean copula (EAC) (Nelsen 1999, Joe 1997). The EAC is extremely restrictive, allowing the specification of only one distribution parameter, regard less of dimension. The more flexible multivariate Archimedean copula extensions are then constructed: the fully nested Archimedean construction (FNAC) and the partially nested Archimedean construction (PNAC)

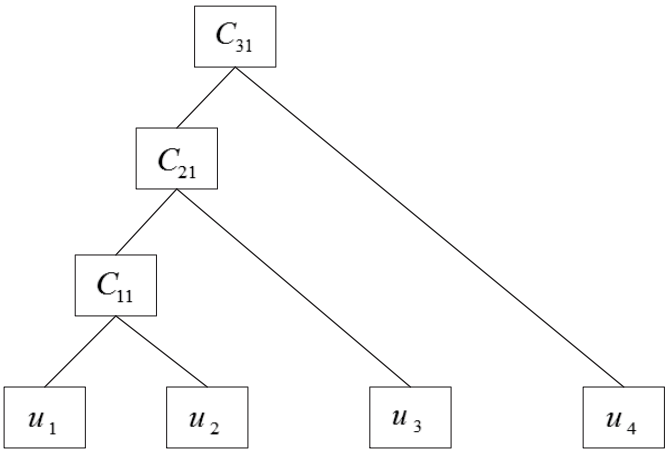

Figure 2.1: Fully nested Archimedean construction

are discussed by (Embrechts, Lindskog & McNeil 2003, McNeil 2008), and the hierarchically nested Archimedean construction (HNAC) is studied by (McNeil 2008, Savu & Trede 2010).

Exchangeable Archimedean Copulas (EAC)

The typical way to define a multivariate Archimedean copula is EAC, which is defined as

$$C(u_1, u_2, ..., u_n) = \varphi^{-1}\{\varphi(u_1) + ... + \varphi(u_n)\} \tag{2.35}$$

where the function φ is a decreasing function knowns as the generator of the copula and φ^{-1} denotes its inverse. The detail refers to (Nelsen 1999). The copula in Equation (2.35) suffers from a very restricted dependence structure, since all k-dimensional marginal distributions ($k < n$) are identical. For the more flexible multivariate Archimedean copula, the following three nested Archimedean copula are developed.

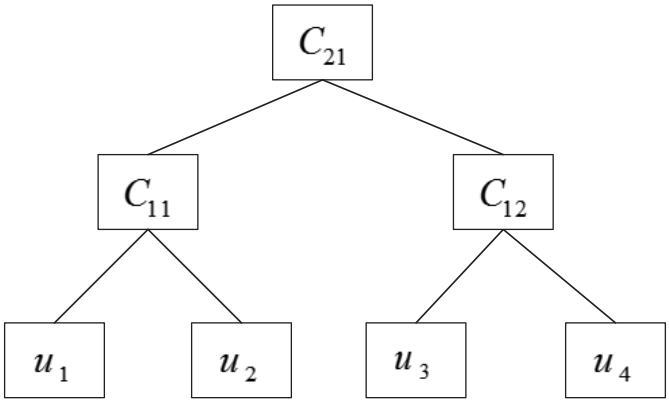

Figure 2.2: Partially nested Archimedean construction

Fully Nested Archimedean Construction (FNAC)

The scheme of Fully Nested Archimedean Construction (FNAC) is to add the dimension step by step. The corresponding expression for the 4-dimensional case is written as:

$$C(u_1, u_2, u_3, u_4) = C_{31}(u_4, C_{21}(u_3, C_{11}(u_1, u_2))) \tag{2.36}$$

It allows for the free specification of 3 copulas and corresponding distribution parameters. The FNAC is a construction of partial exchangeability and there are some technical conditions that needs to become satisfied to be a proper 4-dimensional copula. Figure 2.1 shows the corresponding figure. The two pairs (u_1, u_3) and (u_2, u_3) both have copula C_{21}. The three pairs (u_1, u_4), (u_2, u_3) and (u_3, u_4) all have copula C_{31}.

Partially Nested Archimedean Construction (PNAC)

An alternative multivariate extension is the *Partial Nested Archimedean Construction* (*PNAC*), which is originally proposed by (Joe 1997). It is then discussed by (Whelan 2004, McNeil, Frey & Embrechts 2010). The

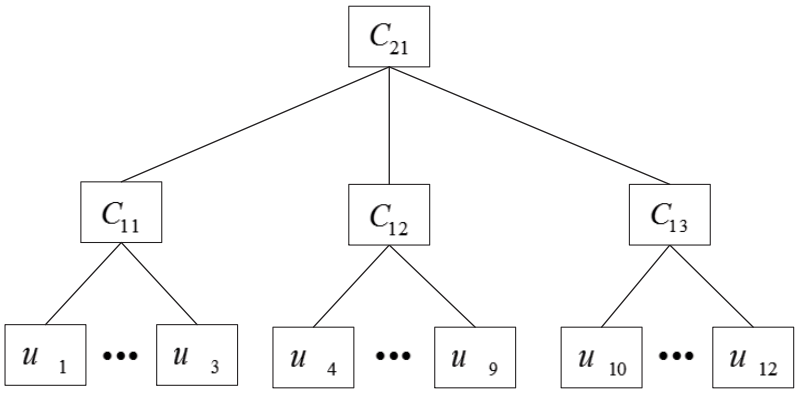

Figure 2.3: Hierarchically nested Archimedean construction

4-dimensional expression is given as follows:

$$C(u_1, u_2, u_3, u_4) = C_{21}(C_{11}(u_1, u_2), C_{12}(u_3, u_4)) \tag{2.37}$$

Figure 2.2 illustrates the structure graphically. Firstly, the two pairs (u_1, u_2) and u_3, u_4 are coupled with copula C_{11} and C_{12}. These two copulas are coupled with a third copula C_{21}. The resulting copula is exchangeable between among u_1,u_2, u_3 and u_4. Thus, it combines the EAC and FNAC.

Hierachically Nested Archimedean Construction (HNAC)

The HNAC was originally proposed by (Joe 1997), and then discussed by (Whelan 2004). Savu and Trede (2006) first developed the idea in full generality.

The difference between HNAC and PNAC is that a copula at a specific level in the hierarchical structure, does not have to be the bivariate copula. Figure 2.3 shows an example of a 12-dimensional case. At the first level, there are three copulas. The C_{11} is a 3-dimensional EAC joining the variables u_1,u_2 and u_3. The copula C_{12}, is a 6-dimensional EAC joining the variables u_4, u_5, u_6, u_7, u_8 and u_9. The copula C_{13}, is a 3-dimensional EAC joining the variable u_{10}, u_{11} and u_{12}. At the second level, the three copulas from the first

level are joined by C_{21}, which is a 3-dimensional EAC. Thus, the HNAC is a partially exchangeable copula. The 12-dimensional copula can be expressed as:

$$C(u_1, u_2, ..., u_{12}) = C_{21}(C_{11}(u_1, u_2, u_3), C_{12}(u_4, ..., u_9), C_{13}(u_{10}, u_{11}, u_{12})) \tag{2.38}$$

2.4 Vine Copula Models

This section firstly introduces the pair copula construction, which is to motivate the regular vine. Then, it presents the essential definitions of graph theory, which is the foundation of regular vine theory. After that, the definition of the regular vine is reviewed, and then the regular vine modelling is discussed, including the dependence structure selection, bivariate copula selection and parameter estimate. At last, the relationship between the regular vine and the graphical probability model are reviewed.

2.4.1 Pair Copula Construction

To model high-dimensional data which exhibits non Gaussian dependency via the copula, the general choice is either elliptical copulas, suggested by Frahm et al. (2003) or Archimedean copulas indicated by Joe (1997). However, it assumes that there is a similar dependence pattern among all pairs of variables. A large number of applications, such as Aas et al. (2009) , Fischer et al. (2009) and Dissmann (2013), showed that the restriction is not satisfied.

Bedford and Cooke (2002) firstly proposed the idea of constructing multivariate distributions using only two-dimensional copulas as building blocks. This was then explicitly discussed by Aas (2009). The corresponding decomposition of a multivariate copula into a bivariate copulas is called *Pair Copula Construction* (*PCC*). A 3-dimensional variable example is used to explain the pair copula construction (*PCC*).

Consider a 3-dimensional random vector $\mathbf{X} = (X_1, X_2, X_3)$ with joint density function f and univariate densities f_1, f_2, and f_3. According to the definition of conditional densities, the 3-dimensional joint density can be obtained as follows:

$$f(x_1, x_2, x_3) = f_3(x_3) f(x_2|x_3) f(x_1|x_2, x_3) \tag{2.39}$$

According to Sklar's theorem (Sklar 1959) and Equation (2.3):

$$f(x_1, x_2, x_3) = c_{123}(F_1(x_1), F_2(x_2), F_3(x_3)) f_1(x_1) f_2(x_2) f_3(x_3) \tag{2.40}$$

where c_{123} is the density of a 3-dimensional copula. In the bivariate case, this yields

$$f(x_2, x_3) = c_{23}(F_2(X_2), F_3(X_3)) f_2(x_2) f_3(x_3) \tag{2.41}$$

for a bivariate copula density c_{23}. Hence,

$$\begin{aligned} f(x_2|x_3) =& \frac{f(x_2, x_3)}{f_3(x_3)} \\ =& c_{23}(F_2(x_2), F_3(x_3)) f_2(x_2) \end{aligned} \tag{2.42}$$

Similarly, for the 3-dimensional case, it can be decomposed

$$\begin{aligned} f(x_1|x_2, x_3) =& \frac{f(x_1, x_3|x_2)}{f(x_3|x_2)} \\ =& c_{13|2}(F(x_1|x_2), F(x_3|x_2)) f(x_1|x_2) \end{aligned} \tag{2.43}$$

where $c_{13|2}$ is the copula density for $f(x_1, x_3|x_2)$ with margins $F_{1|2}$ and $F_{3|2}$. Then, $f(x_1|x_2)$ in Equation (2.42) is decomposed by:

$$f(x_1|x_2, x_3) = c_{13|2}(F(x_1|x_2), F(x_3|x_2)) c_{12}(F_1(x_1), F_2(x_2)) f_1(x_1) \tag{2.44}$$

All decompositions are combined, and then put them into Equation (2.39), which yields:

$$\begin{aligned} f(x_1, x_2, x_3) =& c_{12}(F_1(x_1), F_2(x_2))c_{23}(F_2(x_2), F_3(x_3))c_{13|2}(F(x_1|x_2), F(x_3|x_2)) \\ & \times f_1(x_1)f_2(x_2)f_3(x_3) \end{aligned} \tag{2.45}$$

Then, a trivariate copula density can be constructed using only bivariate copulas, which is given by:

$$\begin{aligned} c_{123}(F_1(x_1), F_2(x_2), F_3(x_3)) =& c_{12}(F_1(x_1), F_2(x_2))c_{23}(F_2(x_2), F_3(x_3)) \\ & \times c_{13|2}(F(x_1|x_2), F(x_3|x_2))) \end{aligned} \tag{2.46}$$

One thing that needs to be noted is that the decomposition is not unique. In Equation (2.39), the variables can be permuted in $3! = 6$ ways. In addition, there is very important assumption of *PCC*. It assumes that that the pair copula $c_{13|2}$ in Equation (2.43) is independent of the conditioning variables X_2, i.e.,

$$c_{13|2}(F(x_1|x_2), F(x_3|x_2); x_2) = c_{13|2}(F(x_1|x_2), F(x_3|x_2)) \tag{2.47}$$

The assumption is necessary to construct flexible models. The simplifying assumption reduces the specification of a *PCC* to choose the bivariate copula families and their parameters. It means that the dependence between variables and the copula function itself can be neglected. Furthermore, the parameters estimation and inference can be performed more easily. Hobæk (2010) showed that it is a good appropriation to the correct decomposition.

Generally, a n-dimensional random vector $\mathbf{X} = (X_1, ..., X_n)$ with joint density f can be decomposed into:

$$f(\mathbf{x}) = f_n(x_n)f(x_{n-1}|x_n)f(x_{n-2}|x_{n-1}, x_n)...f(x_1|x_2, ..., x_n) \tag{2.48}$$

According to arguments in the above 3-dimensional example, each term in Equation (2.48) can be decomposed into marginal densities and bivariate copulas by using the general formula

$$f(x|\boldsymbol{\nu}) = c_{x\nu_j|\boldsymbol{\nu}_{-j}}(F(x|\boldsymbol{\nu}_{-j}), F(\nu_j|\boldsymbol{\nu}_{-j}))f(x|\boldsymbol{\nu}_{-j}) \tag{2.49}$$

where $\boldsymbol{\nu}$ denotes the m-dimensional vector, and ν_j is an arbitrary component of $\boldsymbol{\nu}$. $\boldsymbol{\nu}_{-j}$ denotes the $(m-1)$-dimensional vector $\boldsymbol{\nu}$ excluding ν_j.

The pair copula can be applied to transformed variables, which are marginal conditional distribution of the $F(x|\boldsymbol{\nu})$. According to (Joe 1996), the following is obtained:

$$F(x|\boldsymbol{\nu}) = \frac{\partial C_{x\nu_j|\boldsymbol{\nu}_{-j}}(F(x|\boldsymbol{\nu}_{-j}), F(\nu_j|\boldsymbol{\nu}_{-j})))}{\partial F(\nu_j|\boldsymbol{\nu}_{-j})} \tag{2.50}$$

where $C_{x\nu_j|\boldsymbol{\nu}_{-j}}$ is a bivariate copula distribution function.

There is no unique pair copula construction of n-dimensional random variables. The number of possible decompositions increases significantly with increasing dimensions n. Aas et al. (2009) showed that there are more than 240 different constructions for a 5-dimensional density. Thus, it is necessary to find an appropriate way to describe such models. The regular vine resolves the issue, which is discussed in the following section.

2.4.2 Graph Theory

As discussed above, it is essential to find a way to classify different pairs of copula constructions. The most widely researched copulas arising from *PCC* are the vine copulas. These vine copulas admit a graphical representation called a *Regular Vine* (R vine), which consists of a sequence of trees, and each edge of which is associated with a certain pair copula in the corresponding *PCC*. *Vine* is used since the induced dependence structure is visualised to resemble a grape vine. Thus, the graph theory is reviewed before discussing the regular vine.

The essential theories regarding model construction are reviewed in this thesis. Further information refer to Diestel (2000) and Harris et al. (2008).

A graph is a collection of nodes connected by edges, which is defined by:

Definition 2.3 (Graph) *Let N be an arbitrary set and E be a two dimensional subset of all possible combinations of N*

$$E \in \{\{n_1, n_2\} \mid n_1, n_2 \in N\} \tag{2.51}$$

Elements of N are called nodes, elements of E are called edges and the tuple $G = (N, E)$ is a graph. The numbers of a node $\upsilon \in N$ is the degree of υ denoted by $d(\upsilon)$.

Two nodes are *connected* if and only if there is an edge linked to both of the nodes. The graph defined above is usually referred to as undirected, since the order of nodes correspond to an edge that is arbitrary.

Definition 2.4 (Path, Cycle) *A path in a graph $G = (N, E)$ is a sequence of nodes $(n_1, n_2, ..., n_k) \in N_k, k \geq 2$ such that from each of its nodes there is an edge to the next node in the sequence. For example,*

$$\{n_i, n_{i-1} \in E\} \tag{2.52}$$

for $i = 1, ..., k-1$. A path always has a first node, called its start node (n_1), and a last node (n_k), called its end node. the other nodes in the path are internal nodes.

A cycle is a path such that the start node and the end node are the same $n_1 = n_k$.

The choice of the start node in a cycle is arbitrary.

Definition 2.5 (Connected) *Let $G = (N, E)$ be a graph, two nodes $a, b \in N, a \neq b$ are connected if there is a path from a to b.*

A graph is called connected if every node is connected to every other node.

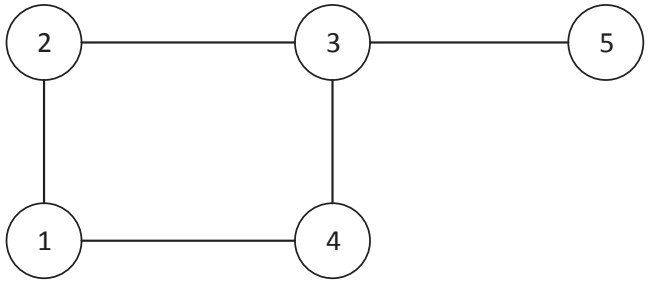

Figure 2.4: An example graph with 5 nodes and 5 edges

Definition 2.6 (Degree) *The degree of a node is the number of its neighbours. Let* $n \in N$ *then* $deg(n) = \#\{e \in E \mid n \in e\}$.

Tree is special class of graph, which is defined by:

Definition 2.7 (Tree) *A tree is a graph* $T = (N, E)$ *that is connected and has no cycles.*

Theorem 2.2 *Let* $T = \{N, E\}$ *be a graph, following characterisations are equivalent.*

(i). T *is a tree,*

(ii). Any two nodes in T *are connected by an unique path,*

(iii). T *is connected and has* $\#(N - 1)$ *edges.*

Example 2.3 *Suppose the nodes are 1 to 5,* $N = \{1, 2, 3, 4, 5\}$*, and then the edges set are* $E = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{3, 5\}, \{1, 4\}\}$*. Figure 2.4 shows the graphical representation of* N *and* E*. Here the cycles represent the nodes, and the lines represent different edges. Two nodes are connected by a line if and only if there is an edge in* E *with two nodes. Figure 2.4 is a connected graph, since there is a path from every node to every other node, i.e., there are no isolated nodes. Figure 2.5 is an example of a cycle. There is a path from 1 to 1 by* $1 \to 2 \to 3 \to 4 \to 1$*. Figure 2.6 is show an example tree, which satisfies all the requirements in Definition 2.7.*

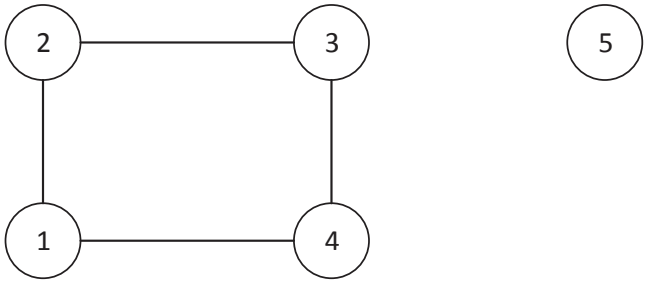

Figure 2.5: An example graph with a cycle

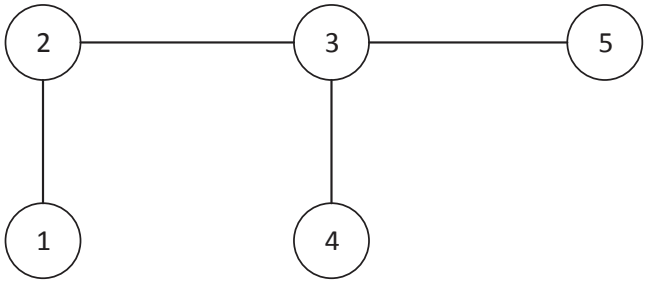

Figure 2.6: An example tree with 5 nodes and 4 edges

2.4.3 Regular Vine

Regular Vine (R Vine), introduced by (Bedford & Cooke 2002), is a special case of graphic model $Vines$. V, T, E and N are denoted as vine, trees, edges, nodes respectively. A vine is a nested set of connected trees $V = T_1, ..., T_{n-1}$. The edges of a tree T_j are the nodes from tree T_{j+1}, where $j = 1, ..., n-1$. A *Regular Vine* is a special case of $Vine$, in which two edges in tree T_j are joined by an edge in tree T_{j+1} only if these edges share a common node. The formal definition appears below:

Definition 2.8 (Regular Vine) *V is a regular vine on n variables if*

(i). T_1 is a tree with nodes $N_1 = 1, ..., n$ and a set of edges denoted by E_1;

(ii). For $j = 2, ..., n-1$, T_j is a tree with nodes $N_j = E_{j-1}$ and edge set E_j;

(iii). ***(Proximity condition)*** *For $j = 2, ..., n-1$ and $a, b \in E_j$, $\#(a \triangle b) = 2$, where $\triangle$ denotes the symmetric difference operator and $\#$ denotes the cardinality.*

R vine on n variables is a nested set of $n-1$ trees such that the edges of tree j become the nodes of tree $j+1$. The proximity condition ensures that if two nodes in T_j have a common node in T_{j+1}, the two nodes are connected by one edge in T_j.

Property (iii) in Definition 2.8 expresses the fact that two nodes are adjacent in tree T_j only if the corresponding edges in tree T_{j-1} are adjacent, i.e., they share a common node. A example of an R Vine is shown in Figure 4.1.

The number of possible regular vines on n variables is still large, since they produce $n!/2\prod_{i=1}^{n-3} i!$ vine dependencies. Hence, two special cases of regular vine are recently developed, *Canonical Vine* (C vine) and *Drawable Vine* (D vine). The two special cases impose additional restrictions, and hence limit effectively the number of different models. For n variables, there are $n!/2$ different C vines or D vines. C vine and D vine are defined by Kurowicka and Cooke (2006*c*), which are given as follow:

Definition 2.9 (Canonical Vine, Drawable Vine) *A regular vine is called a*

(i). canonical vine if each $T_j, j = 1, ..., n-1$ is start, i.e., if each tree T_j has a unique node of degree $n-j$, the root node.

(ii). drawable vine if T_1 is a path, i.e., if each node in T_1 has a degree of at most 2.

The first tree T_1 of a D vine determines all higher order trees $T_2, ..., T_{n-1}$ due to the proximity condition. The example of the 7-variable canonical vine is given in Figure 3.5.

The complete union, constraint, conditioning and conditioned set of an edge are defined as:

Definition 2.10 (Complete Union) *For any* $e_i \in E_i, i \leq n-1, k = 2, ..., i$, *the subset* $U_{e_i}(k)$ *of* $E_{i-k} = N_{i-k+1}$ *is defined by:*

$$U_{e_i}(k) = \{e \in E_{i-k} \,|\, \exists\, e_j \in E_j, j = 1-(k-1), ..., i-1 \\ with\ e \in e_{i-(k-1)} \in e_{i-(k-2)} \in ... \in e_{i-1} \in e_i\} \tag{2.53}$$

Then, the *Complete Union* of $e_i \in E_i$ is defined as

$$U_{e_i} = U_{e_i}(k) \tag{2.54}$$

Thus, U_{e_i} is a set of all nodes in N_i that are connected by the edges e_i. By definition, $U_{e_i}(1) = e_i$.

Definition 2.11 (Conditioning, Conditioned and Constraint Sets) *For* $e = \{a, b\} \in E_i, a, b \in E_{i-1}, i = 1, ..., n-1$, *the conditioning set* (D_e) *with edge* e *is*

$$D_e = U_a \cap U_b, \tag{2.55}$$

and the Conditioned Set (C_e) *with* e *are*

$$C_{e(a)} = U_a \backslash D_e \tag{2.56}$$

$$C_{e(b)} = U_b \backslash D_e \tag{2.57}$$

$$C_e = C_{e(a)} \cup C_{e(b)} = U_a \triangle U_b \tag{2.58}$$

The constraint set for V *is*

$$CV = \{(\{C_{e(a)}, C_{e(b)}\}, D_e) \,|\, i = 1, ..., n-1, e \in E_i, e = a, b\} \tag{2.59}$$

The edge e can be written as $\{C_e | D_e\}$, where the conditioning set D_e is shown to the right of "|", and the conditioned set C_e to the left. $\{U_a \backslash D_e\}$ is the set which includes all variables in the set U_a, but excludes the variables in the set D_e.

Definition 2.12 (M-Child, M-Descendant) *Let the edge f be an element of edge k, and f be an m-child of k. If f is reachable from k via the membership relation: $f \in f_1 \in ... \in k$, f is an m-descendant of k.*

In Definition 2.12, the constraint sets for its two m-children can be written as follows:

$$\begin{aligned} CV_a &= \{(C_{e(a)}, D_e) \,|\, i = 1, ..., n-1, e \in E_i, e = a, b\} \\ CV_b &= \{(C_{e(b)}, D_e) \,|\, i = 1, ..., n-1, e \in E_i, e = a, b\} \end{aligned} \tag{2.60}$$

It can be concludes that the constraint sets of two m-children are indexed by the different elements in the conditioned set C_e of edge k.

Example 2.4 *Suppose there is a regular vine on 4 variables. In Tree T_1, $N_1 = \{1, 2, 3, 4\}$, $E_1 = \{12, 23, 24, 34\}$. Then, in Tree T_2, $N_2 = E_1$, and $E_2 = \{\{1, 2\}, \{2, 3\}; \{2, 3\}, \{2, 4\}; \{2, 4\}, \{3, 4\}\} = \{13|2, 34|2, 23|4\}$. For edge $e = 13|2$ in the tree T_2, the corresponding complete union are $U_a = \{1, 2\}$ and $U_b = \{2, 3\}$. The conditioning set is $D_e = \{1, 2\} \cap \{2, 3\} = \{2\}$. The conditioned set is $C_{e(a)} \cup C_{e(b)} = \{1, 3\}$, where $C_{e(a)} = \{1, 2\} \backslash \{2\} = \{1\}$ and $C_{e(b)} = \{2, 3\} \backslash \{2\} = \{3\}$. The corresponding constraint set is $\{\{C_{e(a)}, C_{e(b)}\}, D_e\} = \{(\{1, 3\}, 2)\}$.*

Generally, a regular vine copula model has three components, the vine tree structure, the copula family for each edge in the tree structure and the corresponding dependence parameters for each pair copula. In the following sections, the three components are reviewed respectively.

2.4.4 Vine Tree Structure Selection Strategies

In the literature, two construction strategies have been proposed, including top down strategies using maximal spanning tree algorithms with various weights and the sequence Bayesian tree selection.

Top Down Strategy Using Maximal Spanning Tree Algorithms

Top down strategy indicates that it starts with selection of the first tree and continues tree by tree until the last tree. A straightforward solution is the strategy using the maximal spanning tree algorithm, which is proposed by Dissmann et al. (2013). The method assumes that a higher weight induces a better fit to the chosen characteristics. Then, the weights can be estimated sequentially by using a sequential estimation approach. Given weights, the Algorithm of Prim (2001) can be applied to select the tree structure which maximises the sum of weights in each tree.

A key step in the solution is to select the weights in the strategy. Possible choices consist of:

(i). the absolute empirical Kendall's τ as proposed by Dissmann et al. (2013) and Czado et al. (2012);

(ii). the AIC of each pair copula as proposed by Almeida and Czado (2012);

(iii). the p-value of a copula goodness-of-fit tests and variants as proposed by Czado et al. (2012).

For the first weight, the motivation is to capture the strongest pairwise dependencies in the data. In a regular vine, the task is to select the strongest pairwise dependencies for the first tree. The most common dependency measure is Kenall's τ, which can capture non-linear dependencies and be invariant under the monotone transformation of the margins. Hence, those variables that can maximise the sum of absolute value of Kendall's τ among all pairs, form a tree. The empirical estimates are used, since the true Kendall's τ is unknown.

When it is useful to find a good fitting of the regular vine to the data, the selection of copula family can be implemented in the way of fitting the corresponding observations well. Then, the most prominent goodness-of-fit measure is the *Akaike Information Criterion* (AIC). However, it does not allow for assigning statistical significance, i.e., the p-value corresponding to

a statistical goodness-of-fit tests. The selection of the pair copula from a set of bivariate copula families is separated in terms of the parameters for each pair of variables. Then, the corresponding AIC is calculated, and the copula family with the lowest AIC is selected.

Since the AIC based weights have drawbacks, the third weights can be used for resolving the issue. The performance of the sequential estimation procedure depends on the selection of a pair copula term for the corresponding pair of pseudo data values. Hence, a copula goodness-of-fit measure is considered for the selection. Genest and Rémillard (2008) presented a computationally expensive parameteric bootstrap procedure for the p-value of GoF tests. Kojadinovic and Yan (2011) and Kojadinovic et al. (2011) provided a multiplier approach to obtain approximate p-value much faster.

Sequential Bayesian Tree Selection

A reversible jump *Markov Chain Monte Carlo* (MCMC) based approach is proposed by Min and Czado (2010). It is an extension of the ordinary MCMC to sample from discrete-continuous posterior distribution. The sampling algorithm is a generalisation of the Metropolis-Hastings algorithm (Green 1995). The approach is to obtain a sequential estimate of the the posterior distribution of the regular vine tree structure, the copula family and their corresponding parameters. It extends the pair copula construction of Aas (2009) by using bivariate t copulas. Modeling the prior density function which favours sparse models can serve to guard against selecting models with runaway complexity. On the other hand, the use of non-informative flat priors allows for tree-by-tree maximum likelihood estimation of the regular vine tree structure, the copula family and the corresponding parameters.

2.4.5 Parameter Estimate

To estimate the parameters for a given a regular vine tree structure and bivariate copula families, the classical method is to use the maximum likelihood estimation. As discussed in Section 2.2.2, a two-stage estimation method can

be considered. Firstly, the marginal parameters are either estimated parametrically or non-parametrically, and then these parameters are fixed to their estimated value in the estimation of copula parameters. The first approach is called *Inference for Margins*, which is discussed in (Joe & Xu 1996), and the second one is called *Maximum Pseudo Likelihood* (MPL) which is presented by Genest et al. (1995). However, even when the estimate of marginal and copula parameters is separated, the joint maximum likelihood estimation of regular vine copula parameters are still very challenging. For example, for an n dimensional vine, its decomposition consists of $n(n-1)/2$ bivariate copulas with the corresponding parameters. Thus, Aas et al. (2009) developed a sequential approach. It starts with the copula of the first tree, then proceeds tree-wise and estimates the parameters of the copulas in a tree by fixing the parameters of copulas in all previous trees. The sequential estimation provides a much faster approach for estimating copula parameters. The sequential approach only includes estimation of bivariate copulas, and thus it is computationally much simpler than the joint maximum likelihood estimation of all parameters at once. Hobæk (2013) investigated the asymptotic properties of the sequential approach. A comparison study of estimators for regular vine copulas is taken by Hobæk (2012). In addition, if the joint maximum likelihood estimates are desired, the sequential method can be used to obtain starting values for the numerical optimisation, which is discussed by Whittaker (2009).

2.4.6 Copula Families Selection

For a given regular vine tree structure, it is necessary to consider how to select the copula from a set of families. Typical criteria for copula selection from a given set of families is information criteria, such as AIC, which is discussed by Manner (2007). For a general regular vine, the selection of a pair copula depends on the choices for the copulas in a previous tree. Since a joint selection seems infeasible due to a lager number of combinations, one typical approach proceeds tree by tree as proposed in the sequential estimation

method. The copula families are selected at first, and then estimated. Given the selected and estimated copula of the previous tree, the copula families of the next tree are selected. The copula selection procedure usually coincides with most vine tree selection strategies. However, the sequential copula selection strategy accumulates in the selection and thus the final model needs to be carefully checked and compared to alternative models. The test for the non-nested model comparison, which is discussed in Section 2.2.3, may be used.

2.4.7 Vine Optimisation

One important issue for the vine copula is the dimensions curse. The computational effort to estimate all parameters grows exponentially with the dimensions. If the vine copula model wants to fit these high dimensional structure (greater than 20 variables), then the vine copula model may need to be optimised. Brechmann et al. (2012) presented a statistical approach to either the truncated or simplified vine copula model. For a regular vine copula, all pair copulas at higher than a certain level, are replaced by the independence copula or the Gaussian copula. The most appropriate level is identified by AIC, BIC and the likelihood-ratio based tests.

2.4.8 Probability Graphical Models and Copula

Probability graphical model, which is a marriage between probability and graph theory, is a general purpose framework aimed at high-dimensional modelling. These models generally combine a qualitative graph structure that encodes independencies and local quantitative parameters to represent multivariate densities. The joint density has a decomposable form that can have the corresponding intuitive graph structure. It allows for efficient methods for model selection (structure learning), marginal and posterior computation, and estimation. Thus, the probability graphical models dominate the machine learning community, and had a significant impact upon

many fields, such as machine vision, natural language processing and bioinformatics. For example, concentration inequalities for dependent random variables (Kontorovich, Ramanan et al. 2008), or via coupling (Chazottes, Collet, Külske & Redig 2007), Bayesian logic program (De Raedt & Kersting 2008), Markov logic networks (Richardson & Domingos 2006), relational dependency networks (Neville & Jensen 2007) first-order probabilistic languages (Milch & Russell 2007), statistical predicate invention (Kok & Domingos 2007),and relational Markov networks (Getoor & Taskar 2007). The latent factor models with dependency structure in the latent space is studied in (He, Qi, Kavukcuoglu & Park 2012). A set of probabilistic dependencies is learned in (Gao & Suzuki 2003) to express the relations between the headwords of each phrase.

However, the probability graphical model has limitations in the context of real word scenarios. The probability graphical models are conceptually general, and the considerations in the real world scenarios always force the local quantitative part of the model to take a simple form. For example, most graphical models have the Gaussian assumption. When the Gaussian assumption cannot capture the data well, the majority of models will discrete data firstly, and then take advantage of the existing methods to resolve it. Therefore, the probability graphical model is difficult to handle with real world data, especially in terms of financial variables which have non-Gaussian distributions and strong asymmetry.

Copula has shown to be a powerful tool for modelling dependencies. Copula splits the multivariate marginal distributions from dependence structures, so that the specification of dependence structures can be studied independently from the marginal distributions. It can provide a flexible mechanism for modelling real world distributions that cannot be handled well by graphical models. Thus, some researchers start to combine the copula and probability graphical models, including tree-structured models by (Kirshner 2007), nonparametric belief Bayesian networks (Kurowicka & Cooke 2005, Hanea, Kurowicka, Cooke & Ababei 2010), and copula Bayesian

networks (Elidan 2010).

The tree-structured model is the first model that combines the probability graphical framework and copula in the machines learning community. The model generalises Darsow's Markovian operator (Darsow, Nguyen, Olsen et al. 1992), and allows for the construction of high-dimensional copulas via a composition of (unconditional) bivariate copulas. One advantage of the tree-structured model is that it needs to only estimate the bivariate cases. It imposes an independence assumption on the tree structure, but it can be relaxed by allowing for a mixture of all trees construction which is efficiently learned using a compactly represented prior. The Bayesian refinement is presented by Silva and Gramacy (2009), which allows for 10s of variables.

The nonparametric (distribution free) belief Bayesian networks use the Bayesian networks structure to encode a decomposition of the joint distribution, which aim to overcome the limitation of simple vines (i.e. the canonical vine). In principle, the construction can be used with any copula for which the specified conditional rank correlation can be realised. However, in practice, only elliptical copula families can be easily carry out.

Copula Bayesian networks use Bayesian networks to encode independencies that are assumed in the dependence structure. The local conditional density is parameterised differently via a proper normalisation of a joint local copula over a variable and its parents in the graph. Compared with tree-structured models, copula Bayesian networks reduces to the tree construction suggested by Kirshner (2007). It is possible to estimate the parameters of the entire model to ensure the preservation of the univariate marginals when using only Gaussian copulas. Thus, the models are equivalent to the nonparametric belief Bayesian networks when using local Gaussian copulas.

Liu et al. (2009) used a nonparanormal method to tackle the problem of structure learning in the complementing representation of undirected graphs. This model is specially focused on a Gaussian copula. It provided theoretical guarantees of consistency when the data was generated from the model, as well as risk consistency guarantees when the samples arose from different

distributions.

All of these models aim at handling high-dimensional dependence modelling. Naturally, different models impose different assumptions and restrictions. The most significant difference between the probability graphical framework and the vine copula framework is that the probability graphical models use conditional independence in the tree structure, rather than the conditional dependence assumed in the vine copula framework. The method applies to the previously unstudied regime of nonparametric estimation in high-dimensions when the number of parameters exceeds that of the samples.

2.5 Model Comparison Test – Vuong Test

In this section, one important test, the Vuong test (Vuong 1989) is reviewed. This is used for comparing various non-nested vine copula models. The test is based on the likelihood ratio and the *Kullback-Leibner information criterion* ($KLIC$). The $KLIC$ is a measure for the distance between two statistical models and is defined by:

$$KLIC := E_0[log\, h_0(Y_i|x_i)] - E_0[log\, f(Y_i|x_i, \hat{\beta}] \tag{2.61}$$

where $h_0(\cdot|\cdot)$ is the unknown true conditional probability function of Y_i given x_i. Here E_0 is the expected value under the true model and $\hat{\beta}$ is the estimator (vector) for the parameter (vector) β in the model $f(Y_i|x_i, \hat{\beta})$, which has not to be the true model. The model with the minimum KLIC is the best one.

The goal is to compare two models with the probability function $f_1(Y_i|x_{i,1}, \hat{\beta}_1)$ and $f_2(Y_i|x_{i,2}, \hat{\beta}_2)$. If the model 1 is better than the model 2, it is true that:

$$E_0[log\frac{f_1(Y_i|x_{i,1}, \hat{\beta}_1)}{f_2(Y_i|x_{i,2}, \hat{\beta}_2)}] > 0 \tag{2.62}$$

According to the likelihood ratio discussed above, Vuong (1989) developed the following statistics to compare the two models:

$$m := (m_1, ..., m_n)^t$$
$$\text{where} \quad m_i := log \frac{f_1(Y_i|x_{i,1}, \hat{\beta}_1)}{f_2(Y_i|x_{i,2}, \hat{\beta}_2)}$$

for $i = 1, ..., n$ and the expected value is:

$$E_0[m] = \mu_0^m = (\mu_1^m, ..., \mu_n^m)^t$$

The null-hypothesis of the Vuong test is

$$H_0 : \mu_0^m \quad \text{versus} \quad H_1 : \mu \leq 0$$

where μ_0^m is know. Additionally he defined the test statistics

$$\nu := \frac{\sqrt{n}(\frac{1}{n}\sum_{i=1}^n m_i)}{\sqrt{\frac{1}{n}\sum_{i=1}^n (m_i - \hat{m})^2}} \tag{2.63}$$

where $\hat{m} = \frac{1}{n}\sum_{i=1}^n m_i$. It shows that under H_0, statistics ν converges in distribution to a standard normal distribution, i.e.,

$$\nu \xrightarrow{D} N(0, 1)$$

The drawback of the Vuong test is that it does not account for the number of parameters in the models, which may differ between the two models. Thus, Vuong suggests correcting the log-likelihood ratio with the correction term if either the Akaike's information criterion or Schwarz's Bayesian information criterion (Vuong 1989) apply,

$$\text{Akaike's information criterion :} \quad p - q$$

or

$$\text{Schwarz's Bayesian information criterion :} \quad \frac{p}{2n} log\, n - \frac{q}{2n} log\, n$$

where p, q are the number of parameters of model 1 and model 2 respectively. For example, the log-likelihood ratio with the Schwarz's Bayesian information criterion is given by:

$$log\, f_1(Y_i|x_{i,1}, \hat{\beta}_1) - log\, f_2(Y_i|x_{i,2}, \hat{\beta}_2) - (\frac{p}{2n} log\, n - \frac{q}{2n} log\, n)$$

The quality of test is measured by the corresponding p-value. Since the null hypothesis is $H_0 : \mu_0^m \quad vs. \quad H_1 : \mu \leq 0$, and the hypothesis are rejected at the significance level of $\alpha\%$ if $|\nu| \geq \Phi^{-1}(1 - \frac{\alpha}{2})$, i.e., the smallest α for which the null hypothesis can be rejected is

$$\alpha = 2\Phi(-|\nu|)$$

where Φ is the distribution function of the standard normal distribution. Therefore, the corresponding p-value of the Vuong test is:

$$p - value = 2\Phi(-|\nu|)$$

2.6 Partial Correlation

In the section, the basic concepts of partial correlation, which are used throughout the thesis are reviewed. The definition of partial correlation are reviewed firstly, and then the relationship between partial correlation and conditional correlation are discussed.

2.6.1 Partial Correlation Definition

The section introduces the partial correlation definition and concepts, which are very important for building vine copula models in the thesis. Bedford and Cooke (2002) defined the partial correlation, which is given by:

Definition 2.13 (Partial Correlation) *Let $X_1, X_2, ..., X_n$ be random variables. The partial correlation of X_1 and X_2 given $X_3, ..., X_n$ is*

$$\rho_{12;3,...,n} = \frac{\rho_{12;3,...,n-1} - \rho_{1n;3,...,n-1} \cdot \rho_{2n;3,...,n-1}}{\sqrt{1 - \rho^2_{1n;3,...,n-1}} \cdot \sqrt{1 - \rho^2_{2n;3,...,n-1}}} \tag{2.64}$$

If $X_1, ..., X_n$ follow a joint normal distribution with variance covariance matrix of full rank, the partial correlation corresponds to the conditional correlation. From the definition, the partial correlation can be computed from the correlation by iterating Equation (2.64).

Lemma 2.5 *if $x, y, z \in (-1, 1)$ and*

$$w = z((1 - x^2)(1 - y^2))^{\frac{1}{2}} + xy \tag{2.65}$$

then $w \in (-1, 1)$.

Lemma 2.5 shows that as long as the partial correlations have been chosen strictly between -1 and 1, it always give a partial correlation lying between -1 and 1. The proof of lemma refers to Bedford and Cooke (2002).

2.6.2 Partial Correlation and Conditional Correlation

Suppose the variance-covariance matrix of $\mathbf{Y} = (Y_1, ..., Y_m)$ is positive definite. The partial variance-covariance matrix for $\mathbf{X} = (X_1, X_2)$ is defined by:

$$\Sigma_{\mathbf{XX;Y}} = \begin{bmatrix} \sigma_{11;\mathbf{Y}} & \sigma_{12;\mathbf{Y}} \\ \sigma_{21;\mathbf{Y}} & \sigma_{22;\mathbf{Y}} \end{bmatrix}$$

It can be calculated as $\Sigma_{\mathbf{XX;Y}} = \Sigma_{\mathbf{XX}} - \Sigma_{\mathbf{XY}}\Sigma^{-1}_{\mathbf{YY}}\Sigma_{\mathbf{YX}}$ by partitioning the variance-covariance matrix of $(\mathbf{X,Y})$ into

$$V\left(\begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix}\right) = \begin{bmatrix} \Sigma_{\mathbf{XX}} & \Sigma_{\mathbf{XY}} \\ \Sigma_{\mathbf{YX}} & \Sigma_{\mathbf{YY}} \end{bmatrix}$$

where $\Sigma_{\mathbf{XX}}$ is 2×2, $\Sigma_{\mathbf{XY}}$ is $2 \times m$ and $\Sigma_{\mathbf{YY}}$ is $m \times m$.

The partial correlation is then rewritten by:

$$\rho_{12;\mathbf{Y}} = \frac{\sigma_{12;\mathbf{Y}}}{\sqrt{\sigma_{11;\mathbf{Y}}\sigma_{22;\mathbf{Y}}}}$$

The partial variance or covariance given Y can be considered as the variance or covariance between residuals of projections of X_1 and X_2 on the linear space spanned by $\mathbf{Y}$,

$$\sigma_{12;\mathbf{Y}} = cov(X_1 - \hat{X}_1(\mathbf{Y}), X_2 - \hat{X}_2(\mathbf{Y}))$$

Thus, $\hat{\mathbf{X}}(\mathbf{Y}) = E(\mathbf{X}) + \Sigma_{\mathbf{XY}}\Sigma_{\mathbf{YY}}^{-1}(\mathbf{Y} - E(\mathbf{Y}))$ is the projection of $\mathbf{X}$, which is the conditional expectation of $\mathbf{X}$ given $\mathbf{Y}$.

In a similar way, the conditional covariance of X_1 and X_2 given $\mathbf{Y}$ is defined by:

$$cov(X_1, X_2 \,|\, \mathbf{Y}) = E((X_1 - E(X_i \,|\, \mathbf{Y}))(X_2 - E(X_2 \,|\, \mathbf{Y}) \,|\, \mathbf{Y}))$$

The conditional covariance matrix is defined by:

$$\Sigma_{\mathbf{XX}\,|\,\mathbf{Y}} = \begin{bmatrix} \sigma_{11\,|\,\mathbf{Y}} & \sigma_{12\,|\,\mathbf{Y}} \\ \sigma_{21\,|\,\mathbf{Y}} & \sigma_{22\,|\,\mathbf{Y}} \end{bmatrix}$$

Hence, the conditional correlation is

$$\rho_{12\,|\,\mathbf{Y}} = \frac{\sigma_{12\,|\,\mathbf{Y}}}{\sqrt{\sigma_{11\,|\,\mathbf{Y}}\sigma_{22\,|\,\mathbf{Y}}}}$$

Theorem 2.6 *For any vectors* $\boldsymbol{X} = (X_1, X_2)$ *and* $\boldsymbol{Y} = (Y_1, ..., Y_m)$*, the following two conditions are equivalent.*

(i). $E(\boldsymbol{X} \,|\, \boldsymbol{Y} = \boldsymbol{\alpha} + \boldsymbol{B}\boldsymbol{Y}$ *for a vector* $\boldsymbol{\alpha}$ *and a matrix* $\boldsymbol{B}$ *,*

(ii). $\Sigma_{\boldsymbol{XX};\boldsymbol{Y}} = E(\Sigma_{\boldsymbol{XX}\,|\,\boldsymbol{Y}})$ *.*

The proof of the Theorem 2.6, which can be found in (2004), is omitted here. Lawrance (1976) shows that the property (i) implies property (ii) for the case when $\mathbf{Y}$ is a scalar variable.

Corollary 2.7 *For any random vectors* $\boldsymbol{X} = (X_1, X_2)$ *and* $\boldsymbol{Y} = (Y_1, ..., Y_m)$, *if there exists a vector* $\boldsymbol{\alpha}$ *and a matrix* $\boldsymbol{B}$, *such that*

$$E(\boldsymbol{X} \mid \boldsymbol{Y}) = \boldsymbol{\alpha} + \boldsymbol{B}\boldsymbol{Y}$$

and $\rho_{12 \mid \boldsymbol{Y}}$ *does not depend on* $\boldsymbol{Y}$, *then* $\rho_{12;\boldsymbol{Y}} = \rho_{12 \mid \boldsymbol{Y}}$.

Proof 2.8 *According to Theorem 2.6, if* $E(\boldsymbol{X} \mid \boldsymbol{Y})$ *is a linear function of* $\boldsymbol{Y}$, *then* $\rho_{12;\boldsymbol{Y}} = E(\rho_{12 \mid \boldsymbol{Y}})$. *the assertion of the corollary holds true, since* $\rho_{12 \mid \boldsymbol{Y}}$ *is independent of* $\boldsymbol{Y}$.

The elliptical distribution is a natural generalisation of the multivariate normal distribution. Corollary 2.7 is true for this distribution.

Example 2.9 (Elliptical Distribution) *The elliptical distribution is a family of distributions whose characteristic functions are given by:*

$$\Psi(\boldsymbol{t}) = e^{i\boldsymbol{t}^T\boldsymbol{\mu}}\phi(\boldsymbol{t}^T\Sigma\boldsymbol{t})$$

for some scalar function ϕ *(see example in (Fang, Kotz & Ng 1990) . This family is denoted by* $EC_n(\boldsymbol{\mu}, \Sigma, \phi)$. *According to the corollary indicated by Cambanis et al. (1981), if* $(\boldsymbol{X}, \boldsymbol{Y}) \stackrel{d}{=} C_n(\boldsymbol{\mu}, \Sigma, \phi)$, *then*

$$E(\boldsymbol{X} \mid \boldsymbol{Y}) = E(\boldsymbol{X}) + \Sigma_{\boldsymbol{XY}}\Sigma_{\boldsymbol{YY}}^{-1}(\boldsymbol{Y} - E(\boldsymbol{Y}))$$

and

$$\Sigma_{\boldsymbol{XX} \mid \boldsymbol{Y}} = s(\boldsymbol{Y})\Sigma^\star$$

where s *is a function and the matrix* $\Sigma^\star$ *is independent of the value of* $\boldsymbol{Y}$. *The conditional distribution is also elliptical. It shows that Corollary 2.7 is true for the elliptical distribution and the partial correlation is identical to the conditional correlation.*

In summary, it shows that the partial correlation is equal to the corresponding conditional correlation for elliptical distribution. This conclusion is very important, since it is the foundation of building models.

2.7 Time Series Analysis

As discussed in Section 2.1, the copula has two parts, the copula function and the marginal distributions. After having concentrated on the copula function, it turns to the modelling of the margins. The time series are considered as marginal distributions since the time series are suitable to the financial applications. The analysis through time series can effectively remove serial dependence among observations and obtain the IID data, which are used as the input for copula models. This section provides background on the time series, and introduces some basic concepts, including stationarity, white noise and autovariance function. Then, the definition of of $ARMA$ models are given. After that, the $ARCH$ and $GARCH$ models for the time series models with the heteroscedastic variance are discussed. The details of the definitions and models can be found in Brockwell and Davis (2009, 2002).

2.7.1 Basic Definition

The definition of a stochastic process is reviewed at first as it is the basic component of time series analysis.

Definition 2.14 (Stochastic Processes) *A stochastic process is a family of random variables $(X_t)_{t\in\mathfrak{T}}$ defined on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$. $\mathfrak{T}$ is the time domain.*

The time series $(x_t)_{t\in\mathfrak{T}}$ is then a realisation of the family of random variables $(X_t)_{t\in\mathfrak{T}}$. The term, time series is used for both the data and the process of which it is the realisation.

Definition 2.15 (Autocovariance Function) *If $(X_t)_{t\in\mathfrak{T}}$ is a process such that $E[X_t^2] < \infty$ for eahc $t \in \mathfrak{T}$, then the autocovariance function $\gamma_X(\cdot,\cdot)$ of (X_t) is defined by*

$$\gamma_X(r,s) = cov(X_r, X_s) \quad r,s \in \mathfrak{T} \tag{2.66}$$

The autocovariance function summarises important information about the dependency of an infinite collection of random variables. Another concept is stationarity, which is the foundation of time series analysis.

Definition 2.16 (Stationarity) *The time series $(X_t)_{t\in\mathfrak{T}}$ with index set $\mathfrak{T} = \mathbb{Z}$ is said to be stationarity if*

- $E[X_t^2] < \infty$ *for all* $t \in \mathbb{Z}$,
- $E[X_t] = m$ *for some* $m \in \mathbb{R}$, *for all* $t \in \mathfrak{T}$,
- $\gamma_X(s,t) = \gamma_X(s+h, t+h)$ *for all* $s, t \in \mathfrak{T}$ *and all* $h \in \mathbb{Z}$.

According to the definition, it follows that $\gamma_X(r,s) = \gamma_X(r-s,0)$ for all $r, s \in \mathbb{Z}$. Hence, for a stationary process (X_t), it defines

$$\gamma_X(h) = \gamma_X(h,0) = Cov(X_{t+h}, X_t) \quad for\ all\ t, h \in \mathbb{Z} \tag{2.67}$$

$\gamma_X(h)$ is called the autocovariance function of the stationary stochastic process $(X_t)_{t\in\mathbb{Z}}$. One of the simplest kinds of time series is when the X_t are IID with zero mean and variance σ^2. The process is called the *White Noise* (WN) process, which is defined as follow:

Definition 2.17 (White Noise) *Let $(X_t)_{t\in\mathfrak{T}}$ be a stationary stochastic process with $E[Z_t] = 0$ for all $t \in \mathbb{Z}$ and the autocovariance function*

$$\gamma_Z(h) = \begin{cases} \sigma^2, & for \quad h = 0 \\ 0, & for \quad h \neq 0 \end{cases} \tag{2.68}$$

for $\sigma^2 > 0$. Then $(Z_t)_{t\in\mathbb{Z}}$ is called white noise with mean 0 and variance σ^2, i.e., $(Z_t)_{t\in\mathbb{Z}} \sim WN(0, \sigma^2)$.

Definition 2.18 (Backward Shift Operator) *Let $(X_t)_{t\in\mathfrak{T}}$ be a stochastic process. Then the backward shift operator B is defined by*

$$BX_t = X_{t-1} \quad \forall\, t \in \mathbb{Z} \tag{2.69}$$

The power of B are defined iteratively by $B^j X_t = X_{t-j}$.

2.7.2 The ARMA Models

The *Autoregressive Moving Average* ($ARMA$) models are reviewed as follows:

Definition 2.19 (The ARMA(p,q) Process) *Let* $(X_t)_{t\in\mathfrak{T}} \sim WN(0,\sigma^2)$ *for some* $\sigma^2 > 0$ *be a white noise on* $(\Omega,\mathcal{F},\mathcal{P})$. *Let* $p,q \in \mathbb{N}_0$ *and* $\Phi_1,...,\Phi_p$, $\Theta_1,...,\Theta_q \in \mathbb{R}$. *Then any stationary time series* $(X_t)_{t\in\mathfrak{T}}$ *on* $(\Omega,\mathcal{F},\mathcal{P})$ *satisfying* $E[X_0] = 0$ *and*

$$X_t - \Phi_1 X_{t-1} - ... - \Phi_p X_{t-p} = Z_t + \Theta_1 Z_{t-1} + ... + \Theta_q Z_{t-q} \quad for t \in \mathbb{Z} \tag{2.70}$$

is called an ARMA(p,q) process w.r.t $(X_t)_{t\in\mathfrak{T}}$. *A solution to Equation (2.70) is called causal if* $X_t = f(Z_t, Z_{t-1}, ...)$ *with* f *an appropriate measurable function.*

$(X_t)_{t\in\mathfrak{T}}$ *is an ARMA(p,q) process with mean* μ *if* $(X_t - \mu)_{t\in\mathbb{Z}}$ *is and ARMA(p,q) process,* $\mu \in \mathbb{R}$.

Equation (2.70) can be rewritten symbolically in a more compact form using the backshift operator

$$\Phi(B)X_t = \Theta(B)Z_t, \quad for\, all\, t \in \mathbb{R} \tag{2.71}$$

where Φ and Θ are the polynomials

$$\begin{aligned} \Phi(z) &= 1 - \Phi_1 z - ... - \Phi_q z^p, \\ \Theta(z) &= 1 + \Theta_1 z + ... + \Theta_q z^q. \end{aligned} \tag{2.72}$$

The polynomials Φ and Θ will be referred to as the autoregressive and moving average polynomials. Below is an essential result that proves the existence of a unique solution to the ARMA equation.

Theorem 2.10 *If* $\Phi(z) \neq 0$ *for all* $z \in \mathbb{C}$ *such that* $|z| = 1$, *then the ARMA has the unique stationary solution,*

$$X_t = \sum_{j=-\infty}^{\infty} \Psi_j Z_t - j, \quad foralllt \in \mathbb{Z}, \tag{2.73}$$

where the coefficients Ψ_j *are determined by*

$$\Phi(z)\Theta(z)^{-1} = \sum_{j=-\infty}^{\infty} \Psi_j z^j = \Psi(z), \quad with\ r^{-1} < |z| < r \tag{2.74}$$

for an $r > 1$.

The proof and further explanation refers to (Brockwell & Davis 2002).

2.7.3 The GARCH Models

The $ARMA$ models are based on the white noise with variance σ^2. However, sometimes especially in finance, it can be observed that time series exhibits features which are not in line with the behaviour of $ARMA$ processes with IID noise. These may be non-stationary. They have clusters of high volatility and clusters of low volatility, the magnitude of the change changes over time, i.e. the variance is conditional. It is in contrast to the constant variance of noise in the $ARMA$ process. To model the behaviour, Bollerslev (1986) introduced the *Generalised Autoregressive Conditional Heteroscedasticity* ($GARCH$) model, which is a generalised form of the *Autoregressive Conditional Heteroscedasticity* ($ARCH$) models introduced by Engle (1982).

Definition 2.20 (The GARCH(p,q) Process) *The process* $(\varepsilon_t)_{t\in\mathbb{Z}}$ *is called a* $GARCH(p,q)$ *process if for every* t,

$$\varepsilon_t = \sigma_t Z_t \quad and \quad \sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{2.75}$$

and

- $Var(\varepsilon - t|F_{t-1}) = \sigma_t^2,\ \varepsilon_t = \sigma_t Z_t \quad and \quad Z_t\ is\ IID.$ *(strong GARCH)*
- $Var(\varepsilon - t|F_{t-1}) = \sigma^2$ *(semi-strong GARCH), or*
- $P(\varepsilon_t^2|1, \varepsilon_{t-1}, \varepsilon_{t-2}, ..., \varepsilon_{t-1}^2, \varepsilon_{t-2}^2...) = \sigma_t^2$ *(weak GARCH).*

with $p \geq 0, q > 0, \omega > 0, \alpha_i \geq 0, i = 1, ..., q, \beta_j \geq 0, j = 1, ..., p$, *where* F_t *is the information set at time* t.

Here, it assumes the strong $GARCH$ assumption is fulfilled. Using the backward shift operator L, Equation (2.75) can be transformed to

$$\begin{aligned} \sigma_t^2 &= \omega + A(L)\varepsilon_t^2 + B(L)\sigma_t^2 \\ \text{where} \quad A(L) &= \alpha_1 L + \alpha_2 L^2 + ... + \alpha_q L^q \\ B(L) &= \beta_1 L + \beta_2 L^2 + ... + \beta_p L^p \end{aligned} \tag{2.76}$$

Theorem 2.11 *The GARCH(p,q) process (2.75) is stationary with* $E[\varepsilon_t] = 0, Var(\varepsilon_t) = \omega(1 - A(1) - B(1))^{-1}$ *and* $Cov()\varepsilon_t, \varepsilon_s = 0$ *for* $t \neq s$, *if and only if* $A(1) + B(1) < 1$.

An equivalent description of the GARCH(p,q) is given by

$$\begin{aligned} \varepsilon_t^2 &= \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2, \\ \varepsilon_t &= \sigma_t Z_t \end{aligned} \tag{2.77}$$

where $Z(t)_{t\in\mathbb{Z}}$ is and IID sequence with zero mean and unit variance independent of $\varepsilon_{t-k}, k \geq 1$ for all t.

2.7.4 The ARMA(1,1)-GARCH(1,1) Model

In the application, two models are combined. The $ARMA$(1,1)-$GARCH$(1,1) is used as an example to explain the $ARMA$ process with $GARCH$ noise.

$$\begin{aligned} X_t &= \mu + \Phi_1 X_{t-1} + \varepsilon_t + \Theta_1 \varepsilon_{t-1}, \\ \sigma^2 &= \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \\ \varepsilon_t &= \sigma_t Z_t \end{aligned} \tag{2.78}$$

where $(Z_t)_{t\in\mathbb{Z}}$ is the IID sequence with 0 zero and unit variance. Then, the standardised residuals of the model are given by:

$$\hat{Z}_t = \frac{1}{\hat{\sigma}_t}(X_t - \hat{\mu} - \hat{\Phi}_1 X_{t-1} - \hat{\Theta}_1 \hat{\sigma}_{t-1}\hat{Z}_{t-1}), \tag{2.79}$$

where $\hat{\mu}, \hat{\Phi}_1, \hat{\Theta}_1$ and $\hat{\sigma}_t$ are the estimates of μ, Φ_1, Θ_1 and σ_t respectively.

Generally, the choice of error distribution for $(Z_t)_{t\in\mathbb{Z}}$ is standard normal distribution, i.e., $Z_t \sim N(0,1)$. However, data in financial area often exhibits not only volatility clustering, but also negative skewness and heavy tails. Thus, the Student-t distribution (Bollerslev 1987), skewed Student-t distribution (Hansen 1994) or the Normal Inverse Gaussian distribution (Andersson 2001) are also considered as the error distribution.

2.7.5 The Ljung Box Test

To test the goodness of fit of the estimated $ARMA$-$GARCH$ models for a given time series, Ljung and Box (1978) developed a test, which is called the Ljung Box (LB) test. This test examines the independence of the residuals. If there is no dependence among these residuals, one can could be regard them as observations of independent random variables and use them as the input of the copula function. The test addresses this issue by examining the hypotheses:

$$H_0 : (\hat{Z}_t)_{t=1,\ldots,n} \textit{ is white noise} \quad \textit{against} \quad H_1 : \textit{not } H_0$$

where $\hat{Z}_t$ are the standardised residuals, which is defined in Equation (2.79).

For lags $k = 1, ..., n-1$. The test of the statistics of the Ljung Box test is defined as follows:

$$\hat{r}_k = \frac{\sum_{t=k+1}^{n} \hat{Z}_t \hat{Z}_{t-k}}{\sum_{t=1}^{n} \hat{Z}_t^2} \tag{2.80}$$

The procedure is a one-at-a-time test, i.e., the significance level applies to the autocorrelations considered individually. Ljung and Box (1978) develop

a test to evaluate the jointly autocorrelation of the first m lags, where $1 \leq m \leq n-1$. The corresponding test statistics is given by:

$$\hat{Q}(\hat{r}) = n(n+1)\sum_{k=1}^{m}(n-k)^{-1}\hat{r}_k^2 \tag{2.81}$$

Under the null hypothesis (2.7.5), $\hat{Q}(\hat{r})$ asymptotically follows a χ^2 distribution with $m-p$ degree of freedom, where p is the number of parameters in the choose model. For example,the $ARMA(1,1)$-$GARCH(1,1)$ model with Student-t error distribution, uses 6 parameters.

2.8 Applications in Financial Market

One of the main fields of application for the copula-based model has been in the field of financial economics. Since the empirical evidence shows that dependence between most important assets return is non-normal. For example, non-normal dependence suggests it is shown in stock markets where two asset returns exhibit greater correlation during market downturns than during market upturns. The evidence of asset returns against normality is firstly found by Mills (1927). However, a larger number of studies show that the normal copula is not suitable for recent asset returns. Erb et al. (1994), Longin and Solnik (2001), Ang and Chen (2002), Ang and Bekaert (2002), Bae et al. (2003) show evidence that asset returns exhibit non-normal dependence which is not consistent with a normal copula. The evidence has wide-implications for financial decision-making, in risk management, derivative contracts and portfolio decisions. Last, the vine copula models are reviewed, which are generally applied to high-dimensional cases.

2.8.1 Risk Managements

One important application of the copula in economics and finance is risk management. The main measure of risk management is Value at Risk(VaR), and other measures to estimate the probability of large losses, which results

in a demand for flexible models of dependence between sources of risk. Hull and White (1998) studied the VaR for a collection of non-normal variables. Cherubini and Luciano (2001) study the VaR of portfolio by using copula methods to perform the pairwise analysis of the dependence structure of losses of desks. Embrechts et al. (2003) and Embrechts and Höing (2006) presented similar studies for VaR by investigating various scenarios of dependencies via the copula approach. Rosenberg and Schuermann (2006) presented a copula approach for joint risk distribution, including market, credit and optional risk. Mcneil et al. (2010) gave a clear and detailed treatment of the copula and risk management.

2.8.2 Derivative Contracts

In derivatives markets, non-normal dependence is an important characteristic for pricing. Any contract with two or more underlying assets has a price that is affected by both the strength and the shape of the dependence between assets. Cherubini et al. (2004) provide detailed approaches in derivative and option pricing via the copula. Rosenberg (2003), Bennett and Kennedy (2004), Goorbergh et al. (2003) and Salmon and Schleicher (2006) presented copula approaches for option pricing. Taylor and Wang (2004) and Hurd et al. (2005) considered the use of observed derivatives prices to find the implied copula of the underlying assets.

Since the booming market in credit derivatives, such as credit default swaps and collateralised debt obligations, and the fact that these assets consist of multiple underlying sources of risks leading to a large number of applications for copula in credit risks. Li (2000) introduced the copula approach in credit risk application, and then applied copula in finance. Frey et al. (2001), Schönbucher and Schubert (2001), Giesecke (2004) then considered copula approaches could be used for default risk.

2.8.3 Portfolio Decision Problems

Another important issue for dependence between risky assets is the portfolio decision, which seeks to find portfolio weights that can maximise the investors's expected utility. Hence, it requires a predictive multivariate distribution for assets being considerered. Under the multivariate normality assumption, the optimal portfolio weights only depend on the first two moments of the assets, and linear correlation adequately comprises the necessary dependence information which is required for an optimal portfolio decision. However, when the joint distribution of assets does not follow normal distribution, and when the utility is not quadratic in wealth, the optimal portfolio weights will require a specification of the entire conditional distribution of returns. Patton (2004) used the time-varying copula model for a bivariate equity portfolio problem. Jondeau and Rockinger (2006*b*) used the Taylor series to calculate the expected utility. An obvious advantage of the method is that it remains operational even if a large number of assets are involved. It can also be observed that under the moderate non-normality, the mean-variance criterion provides a good approximation to the expected utility maximisation. Hong et al. (2007) presented a copula method for the investment decision consisting of 11 equity portfolios under 'disappointment aversion' preferences. Sun et al. (2008) proposed a copula ARMA-GARCH model to predict the co-movements of six German equity market indices at high frequency. It was found that the copula $ARMA\text{-}GARCH$ model is able to capture multi-dimensional co-movements among the indices. Christoffersen and Langlois (2011) used the copula-based model for the portfolio decision including 4 common equity market factors. Garcia and Tsafack (2011) considered a portfolio decision problem of 4 assets from two countries.

2.8.4 Financial Contagion

Financial contagion is a phenomenon, where the crisis that occurs in one market can lead to a problem in other markets beyond what would be ex-

pected on the basis of fundamental linkages between the markets. A classical example of financial contagion is the 2008 global financial crisis. The main difficulty in financial contagion is that a baseline level of dependence between the markets has to be established before it can be asserted that the dependence increased during a period of crisis. Rodriguez (2007) developed a Markov switching copula model for the financial contagion problem. Arakelian and Dellaportas (2009) considered the financial contagion problem via the Markov chain Monte Carlo-based copula approach.

2.8.5 Regular Vine Copula Applications

The vine copula model has been successfully applied in higher dimensions. The vine copula have been extended to the regime switching copulas (2009), the factor copula model (2013), and the multivariate option pricing (2000). In addition, discrete data via vine copula has been studied by Panagiotelis and Czado (2012), and constructing non Gaussian distribution on directed acyclic graphs via pair copula construction is presented by Bauer (2012).

2.8.6 Other Applications

There are a large number of applications using the copula approach, which do not fit into any of the above categories. Breymann et al. (2003) and Dias and Embrechts (2010) presented the copula-based model for financial assets using intra-daily data which was sampled at different frequency. Granger et al. (2006) defined a common factor for bivariate time series by using the copula approach. Hu (2006) used a mixture of copula to separate the degree of dependence from the shape of dependence. Brendstrup and Paarsch (2007) studied multi-object auctions via a semi-parametric copula-based model. Bartram et al. (2007) developed a time-varying copula-based model for financial market integration between 17 European stock market indices. Heinen and Rengifo (2007) developed a copula-based model for multivariate counts which exhibit discreteness, overdispersion and both auto and

cross-correlation. Bouyé and Salmon (2009) used copula for quantile regressions. Dearden et al. (2008) and Bonhomme and Robin (2009) considered the dynamics problem in earning data using the copula approach. Lee and Hong (2009) proposed a copula-based multivariate *GARCH* model with uncorrelated dependent errors. Patton (2009), Dudley and Nimalendran (2011) and Kang et al. (2010) studied the dependence structure between hedge funds and other assets via copula. Zimmer (2012) studied the recent US housing crisis using copula.

2.9 Summary

This chapter introduces the copula and related work. Section 2.1 provides the definition of copula from a mathematical perspective, and presents various copula families and their tail dependence. Section 2.2 introduces the copula-based models and their inference. It presents the measure method of dependence at first, and introduces a large number of time-varying copula-based models. Following this, the parameter estimate and inference are presented. The goodness of fitting tests and model selection criteria are also presented. Section 2.3 presents the multivariate copula models, including the multivariate Gaussian copula and the multivariate Archimedean copula models. Section 2.4 introduced the vine copula models. The definition of vine is firstly presented, and then the regular vine is defined. The three most important components of the regular vine, which are the vine tree structure, copula family selection and parameter estimate, are then proposed and lastly the optimisation of the regular vine is introduced. The relationship between the regular vine and the graphical probability models are discussed. Section 2.5, 2.6 and 2.7 present the model comparison tests, partial correlation and time series analysis respectively. Section 2.8 reviews applications in a large number of financial fields, such as risk management and portfolio decision problems. According to the above work related to the copula, it can be concluded that the main gap between the existing methodologies and important

research issues of the copula are: (1) the existing methods are not capable of dealing with high-dimensional data; (2) these models, such as multivariate Archimedean copulas, have made strong assumptions on the dependence structure, and; (3) the existing methodologies do not develop the truncation methods to optimise the structure.

Chapter 3

Modelling the Complex Dependence Structures by Using Truncated Partial Canonical Vine

3.1 Introduction

Modelling the complex dependence structures of financial variables is a fundamental research problem in the financial domain and it is useful for a wide range of applications including price prediction and risk measurement. Its extreme importance has been partially demonstrated in the 2007 global financial crisis. That particular financial turmoil originated from the subprime mortgage market in the United States (US), and it quickly spread to every cell in the US and global financial system. The bankruptcy of the Lehman Brothers in September, 2008 marked a peak point of the crisis. Indeed, the failure of the fourth largest investment company was the largest bankruptcy in US history. The European financial market was heavily impacted by the collapse of the Lehman Brothers as the European financial market was not isolated. If early precautionary measures had been taken, according to

fundamental understandings of the global financial dependence, some of the crisis could have been avoided.

There are various hierarchical and horizontal coupling relations in the stock markets (Cao, Ou, Yu & Wei 2010). European stocks are used as an example to explain the concept of dependence between two financial variables. ^STOXX-50E is a composite index of European stocks, composed of 50 European stocks from six countries. Siemens and Allianz are two listed companies in Germany; France Telecom and Societe Generale GRP are two listed companies from France. It can be easily understood that the price index of Siemens is directly dependent on its national index i.e. Germany. Although sometimes not easily visible, its price index is also dependent on the composite index ^STOXX50E, the national index of France, the price index of Allianz, that of France Telecom, or that of Societe Generale GRP (see the dashed lines in Figure 3.1).

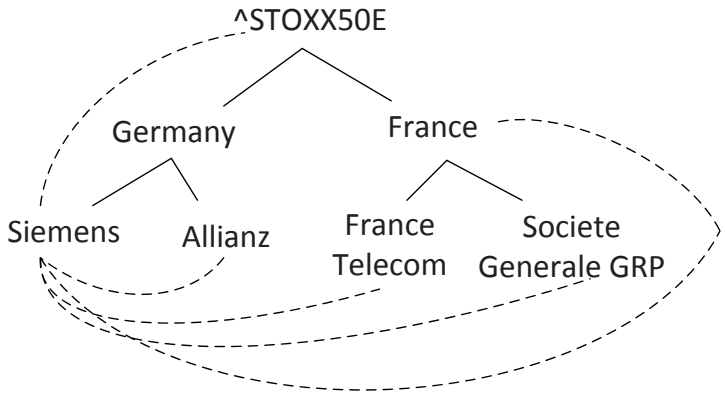

Figure 3.1: An example of dependence structure

Thus, the price of an individual stock can be affected not only by the composite index and its national index, but also by other national indices and the price of individual stocks from other countries. So, the price index of two stocks even from different countries should be dependent on each other as it is assumed to be independent by past models. The concept of this comprehensive dependence structure without prior independence assumptions

can give investors an insightful understanding of the financial market. In particular, it is useful for portfolio investors to control the aggregation risk (Alexander 2009) of the portfolio with higher precision. The classical portfolio investment theorem shows that investors will focus more on the portfolio of the stock return than on single stock returns as they can significantly reduce the aggregation risk while investing in a portfolio of stocks. It is very important for the portfolio investor to identify the dependence structure among assets returns. If investors can accurately capture and model the dependence structure among these returns, they can have a deeper understanding of assets returns, and then predict the trends of the assets return more accurately. In addition, the dependence model include much richer information, which can be used for measuring the risk and accurately implementing risk management.

To model the price and composite index in stock markets, the price's log difference of every two consecutive trading days is usually taken, which are called returns. Early techniques in modelling the dependence of returns in stock markets is by the Pearson's correlation method. It uses the average deviation from mean, ignoring the small or large returns as well as negative and positive returns. However, it is unable to explain the asymmetric correlation of stock markets as shown in (Patton 2004). Another classical idea in modelling dependence is through the use of the Capital Asset Pricing Model ($CAPM$). $CAPM$ belongs to the family of factor models, in which $CAPM$ is the simplest version with only one factor to control the market. The normality assumption makes $CAPM$ inappropriate in modelling returns of stocks, as the returns of stock markets are not normally distributed.

The copula model is a powerful tool in modelling the dependence structure for the returns of stock markets. This is mostly attributed to the fact that the copula model can separate the dependence structure from the marginal distribution. Therefore, the selection of copula functions is not constrained by the choice of marginal distributions. Another key point about the copula model is its consideration of both the dependence of the portfolio of stock

returns as well as correlations of individual stock returns at the same time. Recently, Heinen and Valdesogo (2009) proposed a new copula model, named the *Canonical Vine Autoregressive* (*CAVA*) model, which introduces three different levels of variables: market indices, sector indices and individual stock returns. It integrates these different levels of stock returns as a whole to conduct dependence analysis. However, one limit of the *CAVA* model is that it imposes two independence assumptions on the dependence structure. With this restriction on dependence structures, it is definitely not a perfect model to understand the complex relationships of financial variables.

The challenging task to investigate the dependence structure in stock returns contains three other difficulties. Firstly, it is difficult to build an appropriate model to describe and capture the dependence. The dependence structure will be very much complicated if the data dimension is quite large. If the dependence is constructed by using the canonical vine (which is discussed in detail in Section 3.2), the number of nodes increases exponentially as the number of variables grows. It is difficult to construct and optimise the model when the variables are quite large. Second, financial variables have their own characteristics, which are called stylised facts. The detail of all stylised facts can be found in (Andersen 2009). In the section, the three most important stylised facts are discussed, including volatility clustering, fat tails and asymmetry. Volatility clustering refers to the variance of returns, namely low values of volatility tend to be followed by low values; and high values of volatility are followed by high values (Andersen 2009). The fat tail means that extreme values in stock return occur more frequently than the implied numbers by normal distribution. Asymmetry means that positive and negative returns of the same magnitude have different impacts on the variance. For example, stock returns will have stronger correlation in bear market downturns than in bull markets (Erb, Harvey & Viskanta 1994). Third, daily returns are not normally distributed. Even after the daily returns have been standardised by the dependence model, the standardised residuals are still not normally distributed. Therefore, a new idea in mod-

elling the dependence structure is needed to consider both the dependence between stock returns as well as the stylised facts of the financial time series at the same time.

To fulfill this need, this chapter proposes a new partial correlation-based canonical vine copula model, called the *Partial Canonical Vine*(PCV) Model. The PCV model can capture various and forms of important hierarchical and horizontal dependence. In order to address the high-dimensionality issue, an idea of partial correlation is used to optimise the canonical vine structure. The canonical vine, when it is constructed and optimised by the partial correlation method, is able to model high-dimensional dependence structures. It is capable of maintaining the most important dependence, while reducing the complexity of the dependence structure remarkably, especially for high-dimensional input. In addition, an $ARMA$-$GARCH$ model with skewed Student-t distribution is used for marginal models to capture the stylised facts in financial variables, such as volatility clustering and fat tail.

The main contribution made by this work is the new partial correlation method to construct and truncate the canonical vine in the model PCV. It is practically needed to truncate the partial canonical vine as the number of parameters in the canonical vine is quite large for high dimensional data. The truncated canonical vine can greatly reduce the number of parameters and simplify the canonical vine structure. The truncated partial canonical vine can capture the most important dependencies of data, without a great effect on the structure of the partial canonical vine. In addition, compared with $CAVA$ model, the partial correlation based canonical vine model, PCV, does not impose any independence assumption on the structure and totally eliminate any prior bias for modelling the dependence structure in stock returns. PCV model not only summarises the various dependence with a single number, but also contains much information which is helpful for a deeper understanding of dependence structure. For example, the PCV model can be used to determine the correlation between two stock returns as well as multivariate stock returns. It can be also used to calculate the conditional

copula of stock returns, such as the correlation of two stock returns conditioning on the market indices. Further, the rich information in the PCV model can be utilised to compute risk measures such as Value at Risk.

The rest of this chapter is organised as follows. Section 3.2 provides a short introduction to the $CAVA$ model. Section 3.3 describes the framework of the partial canonical vine in detail, including the construction, truncation and simplification of the canonical vine, and copula family selection. Section 3.4 discusses the selection of marginal distributions. Section 3.5 provides simulation steps for calculating the portfolio of Value at Risk. Section 3.6 applies the partial canonical vine to capture the dependence of returns in European stock markets, and evaluates the model by comparing it with the performance of $CAVA$. Finally, Section 3.7 summarises this chapter.

3.2 Canonical Vine Autoregressive Model

Recently, Heinen and Valdesogo (2009) proposed a *Canonical Vine Autoregressive* ($CAVA$) model based on the canonical vine. Suppose that there are four stock returns (r_{A1},r_{A2},r_{B1} and r_{B2}) from two different sectors (S_A and S_B). The return of market is r_M, and the returns of sector A and B are r_A and r_B respectively. The independence assumptions imposed by the $CAVA$ model are:

- the returns of stock are only dependent on their own sector returns conditioned on the market, but they are independent of all the other sector returns. It means that r_{A1} and r_{A2} are independent to r_B conditioned on r_M, and r_{B1} and r_{B2} are independent to r_A conditioned on r_M. This leads to $c_{r_A,r_{B1}|r_M}(\cdot,\cdot) = 1$, $c_{r_A,r_{B2}|r_M}(\cdot,\cdot) = 1$, $c_{r_B,r_{A1}|r_M}(\cdot,\cdot) = 1$, and $c_{r_B,r_{A2}|r_M}(\cdot,\cdot) = 1$;

- the returns of sector conditioned on market are independent of each other. It means that r_A and r_B conditioned on r_M are independent. This leads to $c_{r_A,r_B|r_M}(\cdot,\cdot) = 1$

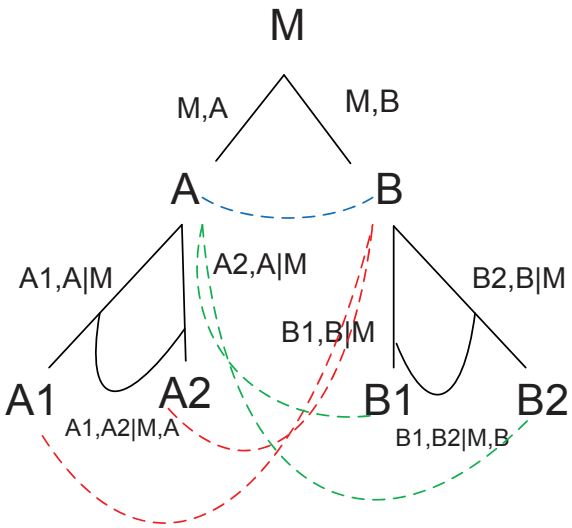

Figure 3.2: CAVA dependence structure

Thus, the following joint density function is obtained by:

$$\begin{aligned} &f(r_M, r_A, r_B, r_{A1}, r_{A2}, r_{B1}, r_{B2}) = \\ &f(r_M) \cdot f(r_A) \cdot f(r_B) \cdot f(r_{A1}) \cdot f(r_{A2}) \cdot f(r_{B1}) \cdot f(r_{B2}) \\ &c(r_M, r_A, r_B, r_{A1}, r_{A2}, r_{B1}, r_{B2}) \end{aligned} \tag{3.1}$$

where $c(r_M, r_A, r_B, r_{A1}, r_{A2}, r_{B1}, r_{B2})$ are the copula function, which can be decomposed into the following Equation:

$$\begin{aligned} &c(r_M, r_A, r_B, r_{A1}, r_{A2}, r_{B1}, r_{B2}) = \\ &c_{r_M, r_A}(F(r_M), F(r_A)) \cdot c_{r_M, r_B}(F(r_M), F(r_B)) \\ &c_{r_M, r_{A1}}(F(r_M), F(r_{A1})) \cdot c_{r_M, r_{A2}}(F(r_M), F(r_{A2})) \\ &c_{r_M, r_{B1}}(F(r_M), F(r_{B1})) \cdot c_{r_M, r_{B2}}(F(r_M), F(r_{B2})) \\ &c_{r_A, r_{A1}|r_M}(F(r_A, |r_M), F(r_{A1}|r_M)) \\ &c_{r_A, r_{A2}|r_M}(F(r_A, |r_M), F(r_{A2}|r_M)) \\ &c_{r_{A1}, r_{A2}, r_{B1}, r_{B2}|r_M, r_A, r_B}(\cdot, \cdot) \end{aligned} \tag{3.2}$$

The dependence structure modelled by $CAVA$ is described in Figure 3.2, where a dashed line indicates that the correlation is assumed to be independent. It is a biased assumption, not an assumption based on data analysis. It is implausible.

3.3 Partial Canonical Vine Model

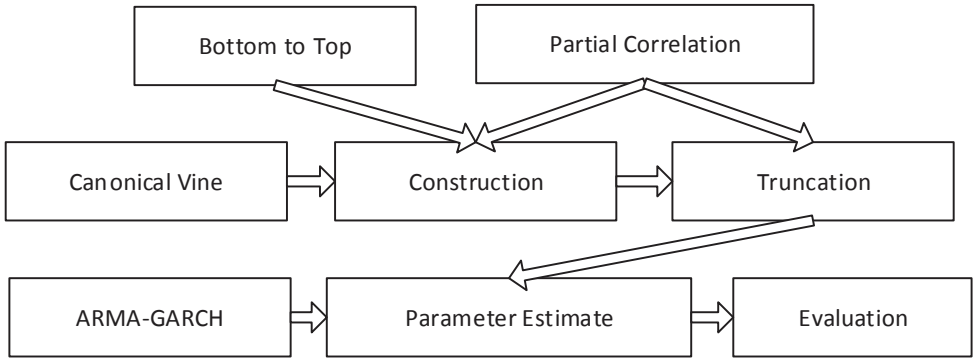

Figure 3.3: PCV model work flow

Figure 3.3 shows the flow chart of the partial canonical vine model. It consists of two separated parts: the canonical vine and marginal distributions. For the part of the canonical vine, the first step is to construct the partial correlation-based canonical vine, and then to truncate and optimise the canonical vine. For the second part, the $ARMA$-$GARCH$ model with the skew Student-t error distribution is used as the marginal distribution. The parameters of both parts are estimated by maximising the likelihood estimation. A parameterised partial canonical vine is therefore capable of measuring the risk or trends of the financial market.

3.3.1 Canonical Vine

As highlighted in Section 3.1, the partial canonical vine model is centered on the canonical vine. It is constructed by a large number of conditional pair copula functions, making it flexible and powerful for modelling the complex dependence structures of high-dimensional financial variables. Typically, the canonical vine density is given by:

$$\begin{aligned} &c(F_1(x_1), F_2(x_2), ..., F_n(x_n)) = \\ &\prod_{j=1}^{n-1}\prod_{i=1}^{n-j} c_{j,j+i|1,...j-1}(F(x_j|x_1, ..., x_{j-1}), F(x_{j+i}|x_1, ..., x_{j-1})) \end{aligned} \tag{3.3}$$

where, $c(\cdot,\cdot)$ is the pair copula density function. It does not impose any independence assumption on the partial canonical vine model, which is more appropriate in modelling the high-dimensional dependence structure of stock returns. The above example in Section 3.1 is used to describe the main ideas of the partial canonical vine model.

Suppose that there is one market M with four stocks in two sectors (A and B). The return of M is denoted by r_M, returns of sectors A and B are denoted by r_A and r_B respectively. The returns of stocks r_{A1} and r_{A2} belong to sector A, and the returns of stocks r_{B1} and r_{B2} belong to sector B. Mapping to the example discussed in Introduction, M (Market) stands for ˆSTOXX50E (market index), A and B stand for Germany and France respectively. $A1$, $A2$, $B1$ and $B2$ stand for Siemens, Allianz, France Telecom and Societe Generale GRP respectively.

The joint density function of the partial canonical vine is given by:

$$\begin{aligned} &f(r_M, r_A, r_B, r_{A1}, r_{A2}, r_{B1}, r_{B2}) \\ =&c(r_M, r_A, r_B, r_{A1}, r_{A2}, r_{B1}, r_{B2}) \cdot \prod f(\cdot) \end{aligned} \tag{3.4}$$

where $\prod f(\cdot) = f(r_M)\cdot f(r_A)\cdot f(r_B)\cdot f(r_{A1})\cdot f(r_{A2})\cdot f(r_{B1})\cdot f(r_{B2})$ are marginal distributions, which will be described in Section 3.4. $c(r_M, r_A, r_B, r_{A1}, r_{A2}, r_{B1}, r_{B2})$

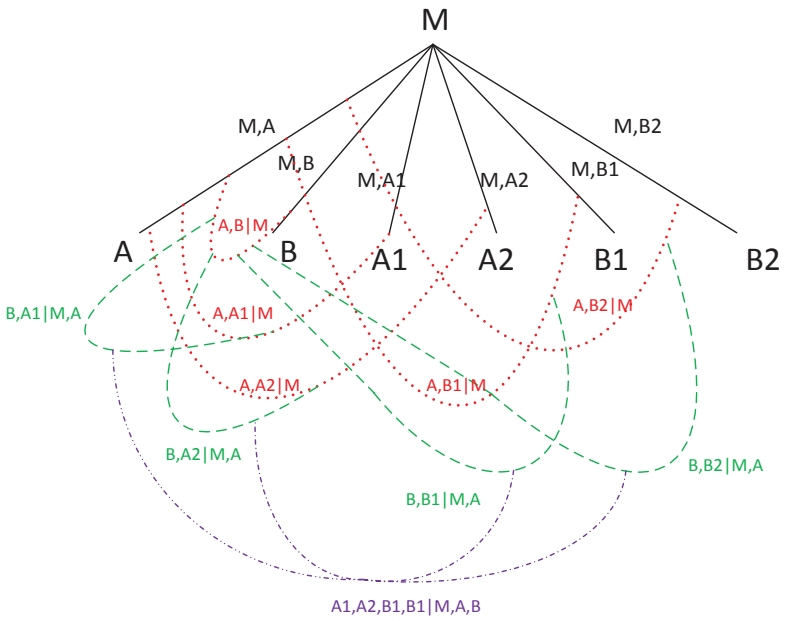

Figure 3.4: PCV dependence structure

is the copula function defined by Equation (3.5):

$$
\begin{aligned}
& c(r_M, r_A, r_B, r_{A1}, r_{A2}, r_{B1}, r_{B2}) = \\
& c_{r_M,r_A}(F(r_M), F(r_A)) \cdot c_{r_M,r_B}(F(r_M), F(r_B)) \\
& c_{r_M,r_{A1}}(F(r_M), F(r_{A1})) \cdot c_{r_M,r_{A2}}(F(r_M), F(r_{A2})) \\
& c_{r_M,r_{B1}}(F(r_M), F(r_{B1})) \cdot c_{r_M,r_{B2}}(F(r_M), F(r_{B2})) \\
& c_{r_A,r_B|r_M}(F(r_A|r_M), F(r_B|r_M)) \\
& c_{r_A,r_{A1}|r_M}(F(r_A, |r_M), F(r_{A1}|r_M)) \\
& c_{r_A,r_{A2}|r_M}(F(r_A, |r_M), F(r_{A2}|r_M)) \\
& c_{r_A,r_{B1}|r_M}(F(r_A, |r_M), F(r_{B1}|r_M)) \\
& c_{r_A,r_{B2}|r_M}(F(r_A, |r_M), F(r_{B2}|r_M)) \\
& c_{r_B,r_{A1}|r_M,r_A}(F(r_B, |r_M, r_A), F(r_{A1}|r_M, r_A)) \\
& c_{r_B,r_{A2}|r_M,r_A}(F(r_B, |r_M, r_A), F(r_{A2}|r_M, r_A)) \\
& c_{r_B,r_{B1}|r_M,r_A}(F(r_B, |r_M, r_A), F(r_{B1}|r_M, r_A)) \\
& c_{r_B,r_{B2}|r_M,r_A}(F(r_B, |r_M, r_A), F(r_{B2}|r_M, r_A)) \\
& c_{r_{A1},r_{A2},r_{B1},r_{B2}|r_M,r_A,r_B}(\cdot, \cdot)
\end{aligned}
\tag{3.5}
$$

where c_{r_M,r_A} stands for the copula between the returns of market and the returns of sector A, $c_{r_M,r_{A1}}$ is the copula between the return of market and the return of stock 1 in sector A. $c_{r_{A1},r_{A2},r_{B1},r_{B2}|r_M,r_A,r_B}$ represents a four-dimensional conditional copula, meaning the dependence of the four stocks conditioned on the market and the two sectors. The term $c_{r_{A1},r_{A2},r_{B1},r_{B2}|r_M,r_A,r_B}$ $(\cdot,\cdot)$ can be decomposed into 6 conditional pair copulas.

This dependence structure is displayed in Figure 3.4, where the term (M, A) stands for the correlation between the market and sector A measured by the copula function c_{r_M,r_A}. The term $(A, B|M)$ means the conditional correlation between sector A and sector B conditioned on M, measured by the conditional copula function $c_{r_A,r_B|r_M}$. In the example, the canonical vine consists of 6 trees and 21 nodes. All trees and nodes (explained in Section 3.2) are shown as Figure 3.5. If the dependence is modelled by *CAVA* which imposes two independence assumptions, then the following copula functions $c_{r_A,r_B|r_M}(\cdot,\cdot)$, $c_{r_A,r_{B1}|r_M}(\cdot,\cdot)$, $c_{r_A,r_{B2}|r_M}(\cdot,\cdot)$, $c_{r_B,r_{A1}|r_M,r_A}(\cdot,\cdot)$, $c_{r_B,r_{A2}|r_M,r_A}(\cdot,\cdot)$, $c_{r_B,r_{B1}|r_M,r_A}(\cdot,\cdot)$ and $c_{r_B,r_{B2}|r_M,r_A}(\cdot,\cdot)$ are all ignored in $CAVA$. However, the partial canonical vine does not make any independence assumption in the structure. The conditional independence or dependence between financial variables are judged by data instead.

Considering the characteristics of the canonical vine dependence structure, the key step to construct a canonical vine is to determine the root variables for all trees. The example in Section 3.3.2 is used to explain the root variable. In Figure 3.5, from the first tree to the sixth tree, the corresponding root variables are M, A, B, $A1$, $A2$, $B1$, $B2$ respectively. When the root variable of each tree is identified, the whole canonical vine structure is then determined. Nodes in a canonical vine are defined as the relations in this chapter. Each tree has a different number of nodes, and each nodes can only be allocated to one conditional pair copula. In Figure 3.4, for the first tree, there are 6 nodes, and each nodes is allocated to one bivariate copula, including c_{r_M,r_A}, c_{r_M,r_B},... and $c_{r_M,r_{B2}}$. For the sixth tree, there is only one node, which is allocated to one conditional copula $c_{r_{B1},r_{B2}|r_M,r_A,r_B,r_{A1},r_{A2}}$. It

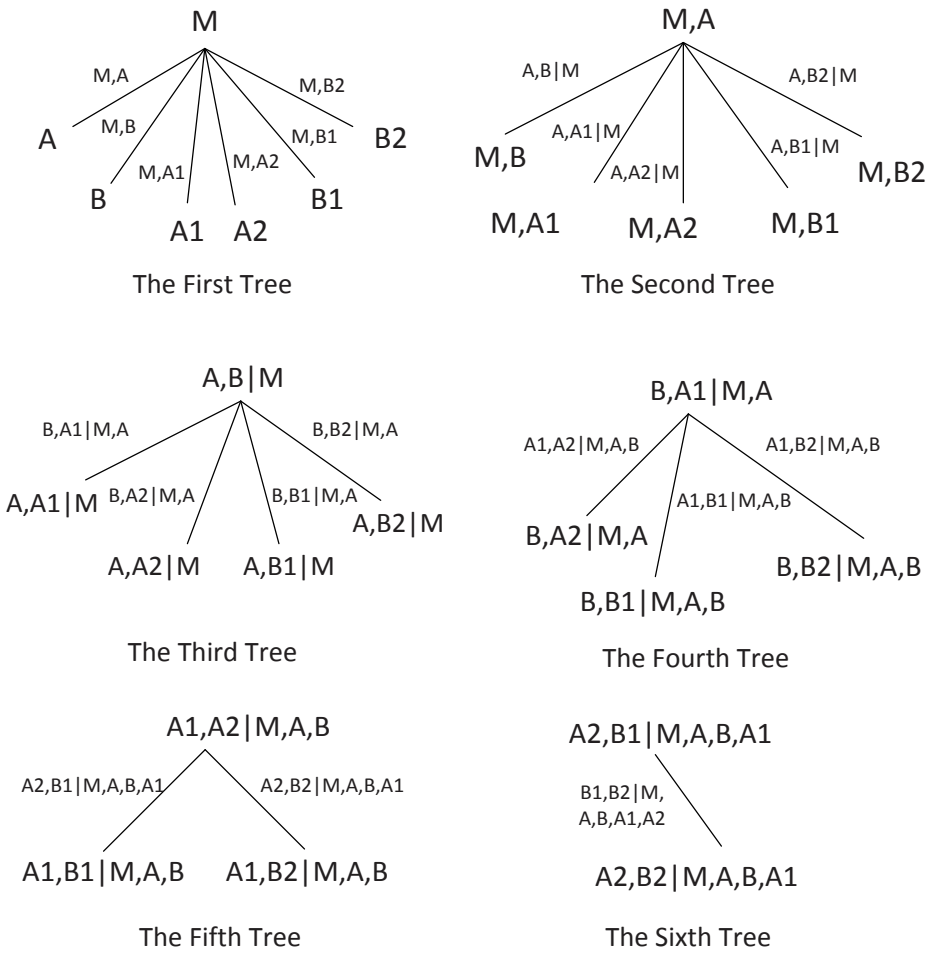

Figure 3.5: The canonical vine dependence structure with 7 variables

is obvious that the number of nodes in a canonical vine will increase exponentially as the number of variables increases. Since each node has to be allocated to one conditional copula, the number of parameters will double for the two-parameter conditional pair copulas. For ease of comprehensibility, it is essential to simplify and optimise the canonical vine.

In the partial canonical vine model, the principle of construction is to capture the most important dependence, and then to truncate it to reduce the complexity of the model without significantly affecting or changing the original structure. It means that the truncated canonical vine should capture

the most important dependence, ignoring the weak dependence. Based on this principle, a partial correlation method is used to constructe the canonical vine dependence structure. In $CAVA$, the construction of the canonical vine is based on the conditional copula. The new method is to build a canonical vine via partial correlation, which is much easier than the conditional canonical vine for constructing and truncating.

3.3.2 Partial Canonical Vine Specification

Generally, let $V^C_{\rho_e}$, T, E, N be denoted as partial correlation based canonical vine dependence structures, trees, edges and nodes respectively. A canonical vine V_{pc} on n variables is a nested set of connected trees $V^C_{\rho_e} = \{T_1, ..., T_{n-1}\}$, where the edges of tree j are the nodes of tree $j+1, j = 1, ..., n-2$. Each tree T_j and have a unique node with $n-j$ edges. The nodes with $n-j$ edges in T_j are called the *root* nodes. It can see that each tree T_j has one root node that connects to all other nodes.

Definition 3.1 (Partial Canonical Vine Specification) *If $V^C_{\rho_e}$ is a canonical vine on n variables, and $e \in E_i$, then a complete partial correlation specification is a canonical vine with a partial correlation p_e specified for each edge e. A distribution satisfies the complete partial correlation specification if , for any edge $e = \{a, b\}$ in the vine, the partial correlation is equal to ρ_e, with the conditioned set C_e and the conditioning set D_e.*

Thus, a Partial Canonical Vine Specification $(V^C_{\rho_e}, B(V^C_{\rho_e}), \theta(B(V^C_{\rho_e}))$ *on n variables, is a multivariate distribution function:*

(i). $V^C_{\rho_e}$ is a partial correlation based canonical vine structure on n variables;

(ii). $B(V^C_{\rho_e}) = \{C_{e(a),e(b)|D_e}\,;\, e \in E_i, i = 1, ..., n-1\}$ is the set of $n(n-1)/2$ bivariate copulas; and

(iii). $\theta(B(V^C_{\rho_e})) = \{\theta_{e(a),e(b);D_e} \,|\, e \in E_i, i = 1, ..., n-1\}$ is the set of parameters, corresponding to the copula in $B(V^C_{\rho_e})$.

According to the above definition, the edge e in the partial canonical vine can be rewritten as $C_e : D_e$, or $C_{e(a)}, C_{e(b)}, ; D_e$ when $e = \{a, b\}$, where C_e and D_e are the corresponding conditioning and conditioned sets for the edge e. The full specification of a partial canonical vine copula has three components: the canonical vine tree structure V_{pc}, the copula family set $B(V_{\rho_e}^C)$, and the corresponding copula parameters $\theta(B(V_{\rho_e}^C))$.

As discussed in Section 2.6.2, the partial correlation is equal to the conditional correlation for elliptical family. The following theorem provides the one-to-one relationship between the partial correlation and joint normal distribution, which is given by:

Theorem 3.1 *Given any complete partial correlation vine specification for normal random variables $X_1, ..., X_n$, there is a unique joint normal distribution for $X_1, ..., X_n$ satisfy ing all the partial correlation specifications.*

The proof can be found in (Bedford & Cooke 2002). Theorem 3.1 provides the foundation to construct the multivariate joint distribution via partial correlation. The multivariate joint distribution can be decomposed into many pair copulas via the canonical vine, thus, it considers using partial correlation to construct a canonical vine.

3.3.3 Dependence Structure Construction

This section discusses how to build the canonical vine dependence structure $V_{\rho_e}^C$ via partial correlation. The reason for using partial correlation is that it is not easy to obtain the conditional copula with high-dimensional data, but partial correlations can be easily obtained from the correlation matrix. As discussed in Section 2.6.2, for the elliptical family, the partial correlation is equal to conditional correlation. It ensures that the canonical vine dependent structure can be built via partial correlations. The method to construct a canonical vine based on partial correlation is described in Algorithm 3.1.

The example in Section 3.3.2 is used to explain how to construct a canonical vine based on the partial correlation. There is one market variable (M),

Algorithm 3.1 Partial Canonical Vine Construction

Require: pseudo-observations of n variables

1: Calculate all values of partial correlation, and then allocate the smallest absolute value of partial correlation to the node of last tree.
2: **for** $k = 1, ..., n-2$ **do**
3: **for** $j = 1, ..., n-3$ **do**
4: **if** $T_j > T_k$ **then**
5: Find appropriate root variables for each tree which can minimise the function $\sum |\rho_{C_{e(a)},C_{e(b)}\,;\,D_e}|$.
6: **else**
7: Find appropriate root variables for each tree which can minimise the function of $\sum log(1-\rho^2_{C_{e(a)},C_{e(b)}\,;\,D_e})$.
8: **end if**
9: **end for**
10: **end for**
11: There will be $(n-2)-1$ canonical vines as $k = 1, ..., n-2$. The best canonical vine is the one that can maximise the function $-log(Demt)$, where $Demt = \prod_{e(a),e(b)\in E(V_{pc})} (1-\rho^2_{C_{e(a)},C_{e(b)}\,;\,D_e})$
12: **return** The partial canonical vine dependence structure $V^C_{\rho_e}$

two sector variables (A,B), and four stocks ($A1$,$A2$,$B1$,$B2$), totally 7 variables. There will be 6 trees and 21 nodes in both the canonical vine structure based on the partial correlation and conditional copula. Each node can be allocated to one bivariate copula or one partial correlation. For constructing the canonical vine based on conditional copula, the partial correlation-based canonical vine is built at first. The partial correlation can be obtained via the Equation (2.13). For these 7 variables, there are in total 21 partial correlations, including $\rho_{M,A;B,A1,A2,B1,B2}$,$\rho_{M,B;A,A1,A2,B1,B2}$,...,$\rho_{B1,B2;M,A,B,A1,A2}$. The smallest absolute value of these partial correlations is allocated to the root node of the last tree (the sixth tree in Figure 3.5) as the last tree only has one node. Suppose the selected partial correlation in the last tree is

$\rho_{B1,B2;M,A,B,A1,A2}$. The variables in the last tree are variables $B1$ and $B2$. The sets $c_7 = \{B1, B2\}$ and $d_7 = \{M, A, B, A1, A2\}$ are called the conditioned set and conditioning set respectively. For the selection of the root variable of the second to last tree (the fifth tree in Figure 3.5), there are two nodes which can be allocated as two partial correlations. It has to select one root variable for the second to last tree (the fifth tree) from d_7 and generate 2 new conditioned sets. If the selected root variable of the second to last tree is $A2$, then the two new conditioned sets are $c_6 = \{A2, B1\}$ and $c_6' = \{A2, B2\}$. The corresponding conditioning set for c_6 and c_6' is $d_6 = \{M, A, B, A1\}$. The partial correlations allocated to the two nodes are $\rho_{A2,B1;M,A,B,A1}$ and $\rho_{A2,B2;M,A,B,A1}$. If the chosen root variable is $A1$, the two new conditioned sets will be $c_6 = \{A1, B1\}$ and $c_6' = \{A1, B2\}$. The corresponding conditioning set for c_6 and c_6' is $d_6 = \{M, A, B, A2\}$. The partial correlation allocated to the two nodes are $\rho_{A1,B1;M,A,B,A2}$ and $\rho_{A1,B2;M,A,B,A2}$. When the selected root variable of the second to last tree is M, A or B , the process is similar to $A1$ and $A2$.

A method is proposed to identify the appropriate root variable, which is called the tree broken method. Suppose k is a tree-broken level. For trees beyond the kth tree ($T_j > T_k$), the appropriate root variable must minimise the value of function $\sum |\rho_{C_{e(a)},C_{e(b)}\,;\,D_e}|$. For trees within the kth tree ($T_j \leq T_k$), the appropriate root variables must minimise the value of function $\sum log(1 - \rho^2_{C_{e(a)},C_{e(b)}\,;\,D_e})$. For example, suppose k is 3 in the example. For the first, second and third trees, the selected root variables for each tree must minimise the value of function $\sum |\rho_{C_{e(a)},C_{e(b)}\,;\,D_e}|$. For the fourth, fifth and sixth trees, the chosen root variables for each tree must minimise the value of function $\sum log(1 - \rho^2_{C_{e(a)},C_{e(b)}\,;\,D_e})$. The parameter k can choose different values, such as $k = 1, 2, 3, 4, 5$. Therefore, there should totally have 5 canonical vines. The 'best' canonical vine should maximise the value of function of $-log(Demt)$, where $Demt$ is the determinant of canonical vine.

The determinant ($Demt$) can be calculated by using:

$$Demt = \prod_{e(a),e(b)\in E(V_{pc})} (1 - \rho^2_{C_{e(a)},C_{e(b)}\,;\,D_e}) \tag{3.6}$$

3.3.4 Copula Families Selection

Once the canonical vine dependence structure $V^C_{\rho_e}$ is identified, the next step is to choose the bivariate copula $B(V^C_{\rho_e})$ for each edge of all of the trees. Since the partial correlation is equal to the conditional correlation for elliptical family only, the Gaussian or t copulas can be considered. Gaussian is not a good option, since a large number of evidence shows that the distribution in financial fields does not obey normal assumption, which is discussed in Chapter 2. In addition, the Gaussian copula does not have tail dependence. Thus, the t copula, which has symmetric upper and lower tail dependence, is chosen as the bivariate copula family given the partial canonical vine dependence structure $V^C_{\rho_e}$.

3.3.5 Vine Tree Structure Truncation and Simplification

The number of the parameters is very large for higher dimensional data. For example, if a n-variable canonical vine with bivariate t copula (two parameters) is used as building blocks, the number of parameters of the multivariate copula functions is $n(n-1)$. It is important to reduce the parameters of the multivariate copula function. The idea is that if the edges of those absolute values are less than the specified significant value ρ_{sign} which is between 0 and 1 (i.e., $0 < \rho_{sign} < 1$), then they are replaced with conditional independence copulas or bivariate Gaussian copulas. For the conditional independence copula case, it can be called the *Truncated Partial Canonical Vine* ($TPCV$), and for the bivariate Gaussian copula case it can be called the *Simplified Partial Canonical Vine* ($SPCV$).

Before discussing the truncated partial canonical vine, the conditional independence copula is first introduced. For a canonical vine with 3 variables, the general expression is given by:

$$\begin{aligned} f(x_1, x_2, x_3) =& f(x_1) \cdot f(x_2) \cdot f(x_3) \\ =& c_{12}(F(x_1), F(x_2)) c_{23}(F(x_2), F(x_3)) \\ =& c_{13|2}(F(x_1|x_2), F(x_3|x_2)) \end{aligned} \tag{3.7}$$

It assumes that X_1 and X_3 are independent given X_2, then it leads to $c_{13|2}(F(x_1|x_2), F(x_3|x_2)) = 1$. Thus, Equation (3.7) can be written as follow:

$$f(x_1, x_2, x_3) = f(x_1) \cdot f(x_2) \cdot f(x_3) c_{12}(F(x_1), F(x_2)) c_{23}(F(x_2), F(x_3)) \tag{3.8}$$

Generally, for any vector of variables $\mathbf{V}$ and two variables X_1 and X_2, it holds that X_1 and X_2 are conditional independent given $\mathbf{V}$ if and only if

$$C_{x_1, x_2|\mathbf{v}}(F_{x|\mathbf{v}}(x|\mathbf{v}), F_{y|\mathbf{v}}(y|\mathbf{v})) = 1 \tag{3.9}$$

Thus, according to Equation (3.9), the conditional independence copula is equal to 1, without any parameters. The conditional independence copula is used to replace the edges with weak correlation indicated by low absolute value of partial correlation in the vine. Thus, the number of estimation parameters can be greatly reduced. Some dependencies in the vine may be lost, however, they are very weak. Thus, they will not affect the performance of the partial canonical vine model. The algorithm to truncate the partial canonical vine is given in Algorithm 3.2.

Another alternative method is that these edges are replaced with bivariate Gaussian copulas. The parameters of bivariate Gaussian copulas can be quickly estimated due to the Gaussian assumption. The actual dependencies may not be reflected when using the bivariate Gaussian copulas, since the Gaussian assumption is not suitable to real world data. However, like the

Algorithm 3.2 Partial Canonical Vine Truncation

Require: The n-variable partial canonical vine dependence structure $V_{\rho_e}^C$ and ρ_{sign}

1: Calculate the all partial correlations based on the canonical vine dependence structure $V_{\rho_e}^C$;
2: **for** $k = 1, ..., n-1$ **do**
3: In tree T_j, find all edges, in which their absolute value of partial correlation are less than ρ_{sign}, e.g. $\rho_{C_{e(a)},C_{e(b)};D(e)} < \rho_{sign}$;
4: For edges found in Step 3, replace the corresponding edges with conditional independence copulas;
5: **end for**
6: **return** The truncated partial canonical vine dependence structure

Algorithm 3.3 Partial Canonical Vine Simplification

Require: The n-variable partial canonical vine dependence structure $V_{\rho_e}^C$ and ρ_{sign}

1: Calculate the all partial correlations based on the canonical vine dependence structure $V_{\rho_e}^C$;
2: **for** $k = 1, ..., n-1$ **do**
3: In tree T_j, find all edges, in which their absolute value of partial correlation are less than ρ_{sign}, e.g. $\rho_{C_{e(a)},C_{e(b)};D(e)} < \rho_{sign}$;
4: For edges found in Step 3, replace the corresponding edges with Gaussian copula;
5: **end for**
6: **return** The simplified partial canonical vine dependence structure

conditional independence copula, these replacement edges are weak correlation, and they are not able to affect the performance of the partial canonical vine. The algorithm to simplify the partial canonical vine is given in Algorithm 3.3.

3.3.6 Parameter Estimate

Due to the large number of parameters, a two-stage estimation method is considered to estimate parameters. Firstly, the parameters of the univariate margin are estimated, then the parameters of the copula function are estimated given fixed parameters of margins. The method, which is discussed in Section 2.2.2, is called *Inference for Margins* (IFM). For high-dimensional data, the IFM is more efficient than the MLE. The loss of efficiency for the IFM is not great, which is studied by Joe (2005).

Suppose n variable are observed at $\mathfrak{T}$ time points, then let $\mathbf{x}_i = (x_{i,1}, ..., x_{i,\mathfrak{T}}$, $i = 1, ..., n)$ is denoted as the ith observation vector in the data set. Here, it assumes that $\mathfrak{T}$ observations of each variable are independent over time. The assumption is not limiting, since the univariate time-series model (i.e., $ARMA$-$GARCH$ model) can be fitted to the margins, and the residuals, which are independent over time, are used for analysing the dependence structure.

Let parameters $\boldsymbol{\theta}^c = \{\theta_{1,1}, ..., \theta_{j,i}$, for $i = 1, ..., n, j = 1, ..., n-1\}$. The corresponding bivariate copulas are $C_{j,j+i|1...j-1}(u_j, u_{j+i})$. Then, let $\theta^c_{j,i}$ be the set of parameters in the copula density $c_{j,j+i|1...j-1}(F_{j|1:2...j-1}, F_{j+i|1:2...j-1})$. According to Equation (3.3), the log-likelihood of the canonical vine is given by:

$$\sum_{j=1}^{n-1}\sum_{i=1}^{n-j}\sum_{t=1}^{\mathfrak{T}} log[c_{j,j+1|1...j-1}\{F_{j|j-1:1...j-2}(x_{j,t}|\mathbf{x}_t^{(j-1)}), F_{j+i|j-1:1...j-2}(x_{j+i,t}|\mathbf{x}_t^{(j-1)})\}] \tag{3.10}$$

where $\mathbf{x}_t^{(j-1)} = (x_{1,t}, ..., x_{j-1,t})$. In general, for each copula in Equation (3.10), there is at least one parameter to be estimated. The number depends on which bivariate copula is chosen. In the section, there are two parameters to be estimated, since the t copula family is chosen.

Aas et al. (2009) proposed an algorithm to evaluate the likelihood for a canonical vine. The algorithm can be applied to estimate the log-likelihood

Algorithm 3.4 Parameter Estimate of Partial Canonical Vine

Log-likelihood = 0 ;
for $i = 1, ..., n$ **do**
  $\boldsymbol{\upsilon}_{0,i} = \boldsymbol{\upsilon}_i$ (vectorized over t) ;
end for
for $j = 1, ..., n-1$ (tree level j) **do**
  for $i = 1, ..., n-j$ **do**
    log-likelihood = log-likelihood + $L_{j,j+i}(\boldsymbol{\upsilon}_{j-1,1}, \boldsymbol{\upsilon}_{j-1,i+1}, \theta^c_{j,i})$;
  end for
  if $j == n-1$ **then**
    Stop;
  end if
  for $i = 1, ..., n-j$ **do**
    $\boldsymbol{\upsilon}_{j,i} = C_{j+i|j:1...j-1}(\boldsymbol{\upsilon}_{j-1,1}; \theta^c_{j,i})$ (vectorized over t) ;
  end for
end for

function of the partial canonical vine.

The Algorithm 3.4 is to evaluate the log-likelihood for the partial canonical vine. The out for-loop corresponds to the outer sum in Equation (3.10), which corresponds to the tree level of the canonical vine, where $\boldsymbol{\upsilon}_{j,i}$ is $\upsilon = F(x_{i+j,t|x_{1,t},...,x_{j,t}})$. The for-loop consists in turn of two other for-loops. The first for-loop corresponds to the sum over i in Equation 3.10. For the second for-loop, the conditional distribution function needed for the next run of the outer for-loop are computed. In addition, $L(\mathbf{x}, \boldsymbol{\upsilon}, \boldsymbol{\theta}^c)$ is the log-likelihood of the chosen bivariate copula with parameter set $\boldsymbol{\theta}^c$ given the data vector $\mathbf{x}$ and $\boldsymbol{\upsilon}$, which is given by:

$$L(\mathbf{x}, \boldsymbol{\upsilon}, \boldsymbol{\theta}^c) = \sum_{t=1}^{\mathfrak{T}} log(c(x_t, \upsilon_t, \theta^c)) \tag{3.11}$$

A good starting values of parameters can maximise the log-likelihood function. Aas (2009) provides a sequential estimation method, which can be

used as the starting value of the Algorithm 3.4. The method is described as follows:

i. Estimate the parameters of the copula in tree 1 from the data;

ii. Compute observations for tree 2 using the copula parameters from tree 1 and the conditional distributions ;

iii. Estimate the parameters of the copula in tree 2 using the observation from step ii ;

iv. Compute observation for tree 3 using copula parameters at tree 2 and the conditional distribution;

v. Estimate the parameters of the copula in tree iii using the observations from step iv ;

vi. Repeat step iii and step iv until all parameters are estimated.

3.4 Marginal Models Specification

Volatility clustering is one of the most important characteristics of stocks returns, referring as it does to the variance of returns. It says that low values of volatility tend to be followed by low values, and that high values of volatility are followed by high values. As $ARMA$ representation of $GARCH$ models can capture this characteristic of volatility clustering, the $ARMA$-$GARCH$ model is chosen as the marginal distribution of the partial canonical vine to capture the volatility of daily returns.

3.4.1 ARMA-GARCH Model

Typically, let $X_t(t = 0, 1, ..., \mathfrak{T})$ be a time series of the prices of a financial asset, such as the stock market index. The return of a financial asset can be defined as:

$$r_t = log(X_t/X_{t-1}) \tag{3.12}$$

Then, the $ARMA(p,q) - GARCH(1,1)$ model is defined as follows:

$$\begin{aligned} r_t =& \mu + \sum_{i=1}^{p} \Phi_i r_{t-i} + \sum_{j=1}^{q} \Theta_j \varepsilon_{t-j} + \varepsilon_t \\ \sigma_t^2 =& \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \\ \varepsilon_t =& \sigma_t Z_t \end{aligned} \tag{3.13}$$

where $\omega > 0, \alpha \geq 0, \beta \geq 0, \alpha + \beta < 1$, r_t is the actual financial asset return and σ_t is the volatility of return on day t. The constraint $\alpha + \beta < 1$ is to maintain the $GARCH$ weak stationary. Z_t is the error as discussed in the following subsections. In general, a $GARCH$(1,1) model with three parameters is adequate to fit the financial time series. Hansen and Lunde (1994) provides evidence that it is difficult to find a volatility model which outperforms the $GARCH$ (1, 1) model. In this chapter, $GARCH$ (1,1) is used to interpret how the $GARCH$ models capture volatility clustering. The $GARCH$ coefficient is to be constant and positive. Given the value of ω, it is obvious that a small value of σ_{t-1}^2 will result in a small value of σ_t^2, and a large value of σ_{t-1}^2 will result in a large value of σ_t^2. Therefore, $AR(1)$-$GARCH(1,1)$ is selected for the marginal distributions in a partial canonical vine. Typically, the $AR(1)$-$GARCH(1,1)$ is given by:

$$\begin{aligned} r_t =& \mu + \Phi r_{t-1} + \varepsilon_t \\ \sigma_t^2 =& \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \\ \varepsilon_t =& \sigma_t Z_t \end{aligned} \tag{3.14}$$

3.4.2 Error Distribution

The simplest choice of Z_t is the standard normal distribution, such as $Z_t \sim N(0,1)$. The $ARMA$-$GARCH$ with normal distribution errors indicates a symmetric distribution for observations $\{x_t\}$. However, the distribution of the financial time series is not normal with negative skewness and excess kurtosis. It also has been demonstrated that the distribution of financial time series has fat tails beyond normal distribution (Andersen 2009). Fat

tail means that extreme values occur more frequently than those produced by a normal distribution. The $ARMA$-$GARCH$ with normal error cannot take account of the asymmetry in the distribution of financial data. Therefore, the normal distribution should be excluded. It is more appropriate to use a distribution which has fatter tails than Gaussian distribution. The most commonly used fat tail distributions for fitting the $ARMA$-$GARCH$ model are the skew Student-t distribution. Hansen (1994) proposed the skew Student-t distribution to fit the financial time series. Its density function is as follows:

$$f_{Skewt}(Z_t) = \frac{2\gamma}{1+\gamma^2}[f_t(\gamma Z_t)I(Z_t < 0) + f_t(\frac{Z_t}{\gamma})I(Z_t > 0)] \tag{3.15}$$

where $I(\cdot)$ is the indicator function , $\gamma > 0$ and $f_t(\cdot)$ is the density of the Student-t distribution with v degrees of freedom. When $\gamma = 1$, the skew Student-t distribution becomes Student-t distribution. The skew Student-t distribution is to skew the symmetric Student-t distribution by combining two differently scaled halves of the symmetric distributions. The advantage of skew Student-t distribution is that it can fit the fat-tailed data very well as it has two tails behaving as polynomials. Therefore, the skew Student-t distribution is selected as the error distribution in the application.

3.5 Portfolio of Value at Risk: A Widely Used Measure for Risk Management

Value at Risk (VaR) is a widely used risk measurement on a specific portfolio of stock returns (Alexander 2009). The performance of models can be evaluated by estimating the Value at Risk. A good model can generate good estimates of VaR forecasting. The quality of VaR forecasting generated by different models can be judged by using backtesting methods, including the unconditional coverage test (Kupiec 1995) and the conditional coverage test (Christoffersen 1998). These tests are based on likelihood ratios and the null

hypothesis is that VaR should exhibit a conditional or unconditional coverage equal to the normal significance level α (0.05 in general). The details of these tests can be found in (Guermat & Harris 2002). The portfolio of stock returns at time t is denoted by:

$$r_{t,portfolio} = \sum_{i=1}^{n} \mu_i r_{t,i} \tag{3.16}$$

where r_i is the return for $i = 1, ..., n$ and μ_i is the weight.

The process for computing a Value at Risk is as follows:

(i). Fit $AR(1)$-$GARCH(1,1)$ with skewed Student-t error distribution with returns by using Equation (3.14)

Then, the standardised residuals is obtained by:

$$\hat{Z}_{t,j} = \frac{r_{t,j} - \hat{\mu}_j - \hat{\Phi}_j r_{t-1,j}}{\hat{\sigma}_{t,j}} \tag{3.17}$$

(ii). The ex-ante garch variance forecast for $j = 1, ..., n$ can be computed as follows:

$$\hat{\sigma}^2_{t+1,j} = \hat{\omega}_j + \hat{\alpha}_j \hat{\varepsilon}^2_{t,j} + \hat{\beta}_j \hat{\sigma}^2_{t,j} \tag{3.18}$$

(iii). The standardised residuals obtained from $AR(1)$-$GARCH(1,1)$ are transformed to approximately uniform data $\mathbf{u}_j = u_{1,j}, ..., u_{t,j}$ by using the skewed Student-t cumulative probability function;

(iv). Fit a partial canonical vine structure with approximately uniform data $\mathbf{u}_j$ and estimate copula parameters;

(v). Use the fitted regular vine structure with estimated copula parameters to simulate a sample for each financial return variable, i.e.,$v_{t+1,j}$;

(vi). Transfer the sample to standard residuals by using the inverse Student-t cumulative probability distribution functions with parameters obtained in Step (i), and obtain the simulated standardised residuals, i.e., $\hat{Z}_{t+1,j}$;

(vii). Calculate the one day forecast return and variance for each financial variable by using the estimated $AR(1)$-$GARCH(1,1)$ which is calculated in Step (i), i.e.,

$$\hat{r}_{t+1,j} = \mu_j + \hat{\Phi}_j r_{t,j} + \hat{\varepsilon}_{t+1,j} \tag{3.19}$$

(viii). The portfolio return is calculated by using Equation (3.16). Then, Steps from (iv) to (vii) are repeated for T times (e.g. $T = 10000$). Then, the 99%, 95%, and 90% VaR forecast is determined by taking the corresponding 1%, 5% and 10% quantiles of the portfolio return forecast respectively.

The backtestings are used to evaluate the performance of the portfolio of VaR forecasting by using the partial canonical vine. If the VaR forecasting is accurate, the VaR forecast should possess two properties. The first property is that the exceedances have to occur independently. The second property is that the proportion of exceedances should approximately be equal to the significance level α. The hit variable of ex-post exceedances is given by:

$$I_t = \begin{cases} 1, & \text{if} \quad r_{portfolio,t} < VaR_t(1-\alpha); \\ 0, & \text{otherwise.} \end{cases} \tag{3.20}$$

where $r_{t,portfolio}$ is the ex-post observed portfolio return at time t. If the VaR forecast is accurate, I_t should be equal to the significance level α of backtesting.

In addition, the quality of VaR forecasting can be judged by backtesting methods. Typically, backtesting methods consist of unconditional, independent and conditional coverage tests (Guermat & Harris 2002, Kupiec 1995, Christoffersen 1998). LR_{UC}, LR_{CC} and LR_{IC} are denoted as the log-likelihood ratio of unconditional, conditional and independent coverage. Consider the indicator sequence of violations $I_t, t = 1, ..., \mathfrak{T}$, the LR_{UC} is defined

by:

$$\begin{aligned} LR_{UC} =& -2log\left(\frac{L(q; I_1, ..., I_{\mathfrak{T}})}{L(\hat{\pi}; I_1, ..., I_{\mathfrak{T}})}\right) \\ =& -2log\left(\frac{q_m(1-q)^{\mathfrak{T}-m}}{\hat{\pi}^m(1-\hat{\pi})^{\mathfrak{T}-m}}\right) \sim \chi_1^2 \end{aligned} \tag{3.21}$$

where m is the number of violations and $\hat{\pi} = m/\mathfrak{T}$.

Then, LR_{IC} is defined as follow:

$$\begin{aligned} LR_{IC} =& -2log\left(\frac{L(\hat{\pi}_2; I_1, ..., I_{\mathfrak{T}})}{L(\hat{\pi}_{01}, \hat{\pi}_{11}; I_1, ..., I_{\mathfrak{T}})}\right) \\ =& -2log\left(\frac{\hat{\pi}_2^{n_{01}+n_{11}}(1-\hat{\pi}_2)^{n_{00}+n_{01}}}{\hat{\pi}_{01}^{n_{01}}(1-\hat{\pi}_{01})^{n_{00}}\hat{\pi}^{n_{11}}(1-\hat{\pi}_{11})^{n_{10}}}\right) \sim \chi_1^2 \end{aligned} \tag{3.22}$$

where $\pi_{ij} = Pr(I_t = j|\mathfrak{T}_t - 1 = i), n_i j = \mathfrak{T}\hat{\pi}_{ij}$ is the corresponding frequency, and $\hat{\pi} = n_{01} + n_{11}/n_{00} + n_{10} + n_{01} + n_{11}$.

The LR_{CC} combines the previous two test, which is defined by:

$$LR_{CC} = -2log\left(\frac{L(q; I_1, ..., I_{\mathfrak{T}})}{L(\hat{\pi}_{01}, \hat{\pi}_{11}; I_1, ..., I_{\mathfrak{T}})}\right) \sim \chi_2^2 \tag{3.23}$$

3.6 Dependence Analysis on European Stock Returns

3.6.1 Data and Marginal Model Specification

To evaluate the performance of truncated and non-truncated PCV on real financial data, the log daily returns of 45 stocks from ^STOXX50E, a Europe-wide composite index, corresponding to the market variable (M) mentioned in the example 3.3.2 are used in the experiment. The sector variables are the five national leading stock indices corresponding to the home country of the chosen stocks such as ^FCHI, ^GDAXIP, ^AEX, FTSEMIB.MI and ^IBEX, namely the national indices of France, Germany, Netherland, Spain

Table 3.1: Indices and Stocks

Indices	France	Germany	Netherlands	Spain	Italy
^STOXX50E	ACA.PA	ALV.DE	AGN.AS	BBVA.MC	ENEL.MI
^AEX	AI.PA	BAYN.DE	INGA.AS	IBE.MC	ENI.MI
^IBEX	ALO.PA	DAI.DE	PHIA.AS	REP.MC	G.MI
^FTSEMIB.MI	BN.PA	DB1.DE		SAN.MC	ISP.MI
^GDAXIP	BNP.PA	DBK.DE		TEF.MC	TIT.MI
^FCHI	CA.PA	DTE.DE			
	CS.PA	EOAN.DE			
	DG.PA	MUV2.DE			
	FP.PA	RWE.DE			
	VIV.PA	SIE.DE			
	FTE.PA	SAP.DE			
	GLE.PA				
	GSZ.PA				
	MC.PA				
	OR.PA				
	SAN.PA				
	SGO.PA				
	SU.PA				
	UL.PA				

and Italy. Some of these national leading stock indices are referred to as the sector variables A and B by the example of Section 3.3.2. There 51 variables involved in the experiment in total. The data was downloaded from yahoo finance (http://finance.yahoo.com), it spans 970 trading days from 22/05/2006 to 30/06/2010. The stocks and indices in the experiment are listed in Table 3.1. The returns of these indices and stocks are calculated by taking the log difference of the prices on every two consecutive trading days.

The experiments and obtained descriptive statistics of the data are shown in Table 3.2. Skewness (Skew) is used to measure of asymmetry of the finan-

cial time series, and Kurtosis (Kurt) is used to measure the "peakedness". All of the variables have positive skewness, excluding the national indices ^AEX and ^IBEX. All of the variables exhibit an excess kurtosis, indicating that they do not follow any normal distribution. This would suggest that different and appropriate models should be selected for their marginal distributions. As described in the method Section 3.4.1, $AR(1)$ - $GARCH(1,1)$ is considered as the marginal distribution model to capture the asymmetry. The Ljung Box (LB) test (McLeod & Li 1983) is then used for examining the existence of the residual autocorrelation for each time series. It can be conducted so that the Ljung Box test for each marginal model will ensure that residuals do not have autocorrelation in all of the return data. The Ljung Box test results of the composition index ^STOXX50E and the five national leading indices are shown in Table 3.3. The 45 stocks are not listed. In Table 3.3, p is the corresponding$p$-value. The results indicate that the six indices do not have autocorrelation as all of the p-values are bigger than the significance level 0.05.

Table 3.2: Descriptive Statistics for the Indices

	^STOXX50E	^GDAXIP	^FCH
Min	-0.081	-0.0975	-0.947
Max	0.1211	0.1080	0.1059
Skew	0.1178	0.2168	0.7650
Kurt	9.8754	10.1641	9.5348
	^AEX	^IBEX	FTSEMIB.MI
Min	-0.0959	-0.0959	-0.0860
Max	0.1003	0.1023	0.1088
Skew	-0.1723	-0.1330	0.0014
Kurt	10.0045	8.9012	9.1447

Table 3.3: Results of Ljung Box Test

	^STOXX50E	^GDAXIP	^FCH
statistics	12.7841	14.6496	11.4877
p-value	0.8865	0.7961	0.9326
	^AEX	^IBEX	FTSEMIB.MI
statistics	20.2718	19.0624	18.4021
p-value	0.4447	0.5178	0.5609

3.6.2 The Truncation and Simplification Analysis of Partial Canonical Vine

In this section, the performance of the truncated, simplified partial canonical vine with the non-truncated partial canonical vine are compared and analysed.

Table 3.5 provide the comparison results among truncated, simplified and non-truncated partial canonical vines, including the total number of parameters, estimation time (seconds) [1] and log-likelihood. Gau and Ind refer to the bivariate Gaussian copula and condition independence copulas respectively. If $\rho_{sign} = 0$, it means that the model is a non-truncated or non-simplified model, since no partial correlation is less than or equal to 0. The corresponding figures are shown in Figure 3.6. In Figure 3.6, the number of parameters of $TPCV$ decrease greatly as the significant values increase. The parameter number of $SPCV$ decrease at the beginning, and then is maintained at a certain level after that. The estimation time of both the $TPCV$ and $SPCV$ models reduces greatly. However, compared with $SPCV$, $TPCV$ uses less estimate time. The reason is that the number of parameters of $TPCV$ is less than the number of parameter of $SPCV$ given the same significant value ρ_{sign}. For the log-likelihood, it shows a very small

[1] The estimate time (seconds) is calculated via parameters estimation in partial canonical vine, excluding the estimation of parameters in marginal distributions. The computation is on a Linux cluster computer with 6 Inter Xeon CPU (3.74GHz) and 12G memory.

gap between the $SPCV$ and the non-truncated one, which indicates that the log-likelihood of $SPCV$ is very close to the non-truncated partial canonical vine. The log-likelihood of $TPCV$ decreases smoothly at the beginning, and then deeply reduces after a certain point. According to the analysis of Figure 3.6, the significant value 0.1 is a good choice for the truncated canonical vine. The reason is that the log-likelihood is close to the non-truncated one, but the number of parameters and the estimate time is greatly reduced. In addition, $TPCV$ with significant value 0.4 is selected as the reference, since the log-likelihood decreases sharply after the point value 0.4.

Thus, the following models are used in the case study:

- $TPCV_{0.1}$: Truncated partial canonical vine, built by Algorithm 3.1, and then truncated by Algorithm 3.2 with truncation value 0.1(i.e. $\rho_{sign} = 0.1$);
- $TPCV_{0.4}$; Truncated partial canonical vine, built by Algorithm 3.1, and then truncated by Algorithm 3.2 with truncation value 0.4(i.e. $\rho_{sign} = 0.4$);
- PCV: Partial canonical vine, built by Algorithm 3.1, non-truncated models;
- $CAVA$: Canonical Vine Autoregressive model, built by (Heinen & Valdesogo Robles 2009).

To compare the performance among above four models, t copulas are used for all four model candidatures.

3.6.3 Experiments on Value at Risk

The performance of the model was further evaluated by measuring the Value at Risk (VaR) forecasting. The backtesting tests discussed in Section 3.5 are used for validating VaR forecasting generated by different models. A moving window of 970 observations, corresponding to appropriately 4 years of trading

Table 3.4: Determinants and Numbers of Pair Copulas

ρ_{sign}	0	0.01	0.02	0.03	0.04
$-log(Demt)$	29.02	29.02	28.97	28.88	28.74
No. nodes	1275	1081	906	755	643
ρ_{sign}	0.05	0.06	0.07	0.08	0.09
$-log(Demt)$	28.52	28.31	28.04	27.81	27.43
No. nodes	534	465	399	358	307
ρ_{sign}	0.1	0.15	0.20	0.25	0.30
$-log(Demt)$	27.10	25.10	23.65	22.59	21.82
No. nodes	271	139	90	69	59

days from 22/05/2006 to 30/6/2010, was used to construct the models. The test period was from 01/07/2010 to 01/03/2012 with 500 observations of trading days. The three level of VaR: 90%, 95%, and 99% VaR are calculated. The process is explained in Section 3.5. The VaR of portfolio returns of 51 variables (45 stocks, 5 leading national indices and 1 composite index) in the test period was calculated by using the method in Section 3.5. The models $TPCV_{0.1}$ and $TPCV_{0.4}$ are compared with PCV and $CAVA$.

It is important to examine whether VaR forecastings are consistent with the subsequently realised return given a significant level. Kupiec (1995) introduced a percentage of failure of unconditional coverage to examine the quality of VaR. Christoffersen (1998) proposed a more complete test to examine the conditional coverage and independence of VaR. These tests were used as backtesting for examining the quality of VaR here. The exceedances result is presented in Table 3.6, where PoF stands for the percentage of failure. It can been seen that $TPCV_{0.1}$ performs very well as the two levels of PoF are in the level 95% and 90%, lower than α. At the level 99%, the exceedances slightly increase in comparison to those expected. Compared with $TPCV_{0.1}$, the non-truncated partial canonical vine has a similar performance. However, the $CAVA$ model does not have a good performance, as all of the three levels of exceedances are higher than the expected one. The

performance of $TPCV_{0.4}$ is worse than $CAVA$. The corresponding display of these VaR forecastings is shown in Figure 3.7.

Table 3.7 presents statistics of a conditional coverage test, unconditional coverage test and an independent coverage test, where LR_{CC}, LR_{UC}, and LR_{IC} are short for the likelihood ratio of conditional coverage, likelihood ratio of unconditional coverage and likelihood ratio of independent coverage respectively. The values in the parentheses are the corresponding p-value of these tests. The smaller the statistics are, the bigger the corresponding p-value is, indicating that the VaR forecast is more accurate. If the p-value is bigger than 0.05, it means that the null hypothesis of backtesting can be accepted, and the VaR forecasts are accurate and reliable. For $TPCV_{0.1}$, it can been seen that the VaR forecasting at all of the three levels are sufficiently accurate, as the null hypotheses of the conditional, unconditional and independent cannot be rejected according to these tests. The non-truncated PCV has similar performance in backtesting. It is evident that the model based on the optimal canonical vine performs as good as the non truncated partial canonical vine, suggesting that these weak correlations can be ignored without affecting or changing the whole dependence structure. It means that the truncation method for constructing and optimising the canonical vine is excellent and reliable. On the other hand, the tests on $CAVA$ suggest that VaR forecasting at all of the three levels is not accurate, since the null hypotheses of conditional, unconditional and independent at all of the three levels are rejected. The reason is that $CAVA$ imposes two independence assumptions to ignore some strong correlations in the canonical vine. $TPCV_{0.4}$ has similar performance with $CAVA$, and the reason is that the number of nodes is too small to capture the most important dependencies in the canonical vine tree structure.

Similar analysis for the partial canonical vine under different significance value (ρ_{sign}) is listed in Table 3.4. The results show that for those $PCVs$ with a ρ_{sign} greater than 0.1, their VaR performance is not accepted. For those PCV models with a ρ_{sign} less than 0.1, their performance is as good

as the non truncated partial canonical vine, however, the number of nodes is quite large. When the ρ_{sign} is equal to 0.1, the nodes are small enough, and the performance is accepted.

3.7 Summary

This chapter proposes a new partial canonical vine, as truncated based on partial correlation, to model the complex dependence structures of financial variables. It is useful for analysing the complicated dependence structure of a large portfolio of European stocks. Compared with *CAVA*, the partial canonical vine shows a much better performance in producing the VaR forecastings of portfolio returns. The truncated partial canonical vine , which greatly reduces the number of parameters and the computation burden, shows a similar performance of VaR forecastings as that of the non-truncated partial canonical vine model although with a size much smaller.

Table 3.5: Truncated and Simplified Partial Canonical Vine Analysis

ρ_{sign}	No. Parameters		Estimation Time		Log-likelihood	
	Gau	Ind	Gau	Ind	Gau	Ind
0	1753		17736.23		30157.23	
0.01	1717	1549	14836.62	13771.91	30077.08	29652.81
0.02	1687	1354	13070.86	11156.90	30010.27	29254.79
0.03	1644	1140	10976.36	8399.20	29903.54	28908.65
0.04	1618	987	9621.40	6533.48	29848.93	28695.89
0.05	1588	865	8308.09	5207.68	29741.34	28489.97
0.06	1560	751	7205.76	4099.29	29680.13	28342.98
0.07	1538	669	7835.68	3462.03	29626.86	28191.73
0.08	1513	579	6882.76	3245.45	29619.52	28090.70
0.09	1494	521	6444.99	2795.42	29568.22	27982.74
0.1	1479	472	5872.44	2294.04	29501.07	27796.20
0.11	1456	410	6224.03	2101.07	29484.17	27658.33
0.12	1443	377	5098.59	1606.01	29400.44	27466.45
0.13	1428	332	4694.03	1308.68	29345.36	27142.47
0.14	1411	293	4155.37	996.27	29280.00	26928.20
0.15	1401	270	4592.50	706.05	29340.87	26825.75
0.16	1391	245	4301.66	585.15	29316.88	26579.64
0.17	1382	226	4115.33	705.85	29293.53	26406.71
0.18	1376	212	4004.73	668.08	29275.89	26258.00
0.19	1369	198	3807.06	565.07	29257.09	26069.38
0.2	1364	184	3673.35	511.02	29244.35	25865.96
0.21	1355	164	3519.16	418.93	29225.49	25575.29
0.22	1352	157	3500.87	283.14	29203.94	25351.59
0.23	1348	148	2846.83	243.77	29179.12	25127.67

Continued on next page

0.24	1344	139	2306.98	167.37	29159.01	24906.06
0.25	1340	131	3091.25	146.71	29139.76	24626.02
0.26	1337	125	3062.49	135.07	29127.76	24429.67
0.27	1337	125	3057.43	136.40	29127.76	24429.67
0.28	1336	123	3053.04	132.12	29125.36	24347.60
0.29	1335	121	3060.17	128.82	29124.32	24281.09
0.3	1335	121	3050.37	131.44	29124.32	24281.09
0.31	1334	119	3020.40	126.70	29115.19	24239.01
0.32	1334	118	3024.80	124.60	29115.48	24152.56
0.33	1332	114	2812.87	108.90	29105.68	23827.74
0.34	1332	114	2846.86	107.12	29105.68	23827.74
0.35	1331	112	2800.81	124.67	29092.99	23661.18
0.36	1329	108	2744.71	90.70	29079.59	23348.55
0.37	1329	108	2768.22	91.31	29079.59	23348.55
0.38	1329	108	2765.77	92.48	29079.59	23348.55
0.39	1329	108	2784.19	92.11	29079.59	23348.55
0.4	1328	106	2755.80	90.03	29068.01	23172.38
0.41	1324	98	2713.48	82.19	28996.58	22343.52
0.42	1323	96	2730.19	81.72	28977.24	22133.08
0.43	1321	92	2723.99	79.27	28944.16	21702.16
0.44	1316	82	2662.26	67.63	28865.60	20527.42
0.45	1314	78	2676.69	64.49	28831.70	20005.14
0.46	1310	70	2649.22	71.47	28761.86	19015.21
0.47	1310	70	2649.77	71.86	28761.86	19015.21
0.48	1310	70	2619.84	71.40	28761.86	19015.21
0.49	1307	64	2652.34	66.88	28707.18	18118.04
0.5	1307	64	2617.13	65.77	28707.18	18118.04
0.51	1304	58	2571.59	59.58	28644.90	17135.11
0.52	1304	58	2569.45	59.29	28644.90	17135.11
0.53	1302	54	2564.49	56.27	28621.74	16414.99

Continued on next page

0.54	1302	54	2540.26	54.95	28621.74	16414.99
0.55	1300	50	2558.70	53.08	28587.12	15615.88
0.56	1297	44	2161.31	47.13	28529.89	14383.73
0.57	1293	36	2137.89	39.53	28443.08	12678.80
0.58	1292	34	2149.45	38.01	28418.83	12235.56
0.59	1291	32	2141.55	36.33	28394.38	11799.33
0.6	1287	24	2111.82	28.60	28283.16	9862.87
0.61	1287	24	2136.32	28.82	28283.16	9862.87
0.62	1285	20	2119.50	24.86	28235.09	8804.88
0.63	1285	20	2118.12	24.90	28235.09	8804.88
0.64	1283	16	2166.24	21.10	28156.28	7660.45
0.65	1282	14	2099.66	20.03	28149.24	7055.75
0.66	1282	14	2098.28	19.80	28149.24	7055.75
0.67	1280	10	2096.25	16.30	28077.71	5747.10
0.68	1280	10	2136.72	16.70	28077.71	5747.10
0.69	1280	10	2136.82	16.64	28077.71	5747.10
0.7	1280	10	2182.55	17.16	28077.71	5747.10
0.71	1280	10	2131.52	17.44	28077.71	5747.10
0.72	1280	10	2211.48	17.03	28077.71	5747.10
0.73	1280	10	2211.36	17.48	28077.71	5747.10
0.74	1280	10	2229.68	16.54	28077.71	5747.10
0.75	1280	10	2160.90	16.90	28077.71	5747.10
0.76	1279	8	2148.83	14.81	28065.68	4814.00
0.77	1278	6	2123.78	12.07	28043.30	3854.32
0.78	1277	4	2480.75	10.30	28019.58	2873.42
0.79	1277	4	2472.16	9.77	28019.58	2873.42
0.8	1277	4	2501.70	9.92	28019.58	2873.42

Table 3.6: Number of Exceedances of VaR Forecast

Model	VaR	α	Expected	Actual	POF
$TPCV_{0.1}$	99%	1%	5	7	1.4%
	95%	5%	25	21	4.2%
	90%	10%	50	41	8.2%
$TPCV_{0.4}$	99%	1%	5	16	3.2%
	95%	5%	25	43	8.6%
	90%	10%	50	96	19.2%
$CAVA$	99%	1%	5	13	2.6%
	95%	5%	25	39	7.8%
	90%	10%	50	82	16.4%
PCV	99%	1%	5	7	1.4%
	95%	5%	25	20	4%
	90%	10%	50	40	8%

Table 3.7: The Results of VaR Backtesting

Model	VaR	LR_{CC}	LR_{UC}	LR_{IC}
$TPCV_{0.1}$	99%	3.81	0.72	0.91
		(0.149)	(0.397)	(0.44)
	95%	2.01	0.71	0.74
		(0.367)	(0.399)	(0.39)
	90%	2.16	0.88	0.77
		(0.340)	(0.77)	(0.379)
$TPCV_{0.4}$	99%	13.21	15.47	6.32
		(0.001)	(0.000)	(0.012)
	95%	13.75	15.04	5.68
		(0.001)	(0.000)	(0.017)
	90%	14.60	16.18	6.03
		(0.001)	(0.000)	(0.014)
$CAVA$	99%	11.96	10.99	5.16
		(0.003)	(0.001)	(0.023)
	95%	11.17	10.19	3.98
		(0.004)	(0.001)	(0.046)
	95%	11.73	10.45	4.64
		(0.003)	(0.001)	(0.031)
PCV	99% (0.149)	3.80	0.72	0.77
		(0.397)	(0.381)	
	95%	2.00	0.71	0.84
		(0.367)	(0.399)	(0.361)
	90%	2.28	0.77	0.84
		(0.320)	(0.379)	(0.361)

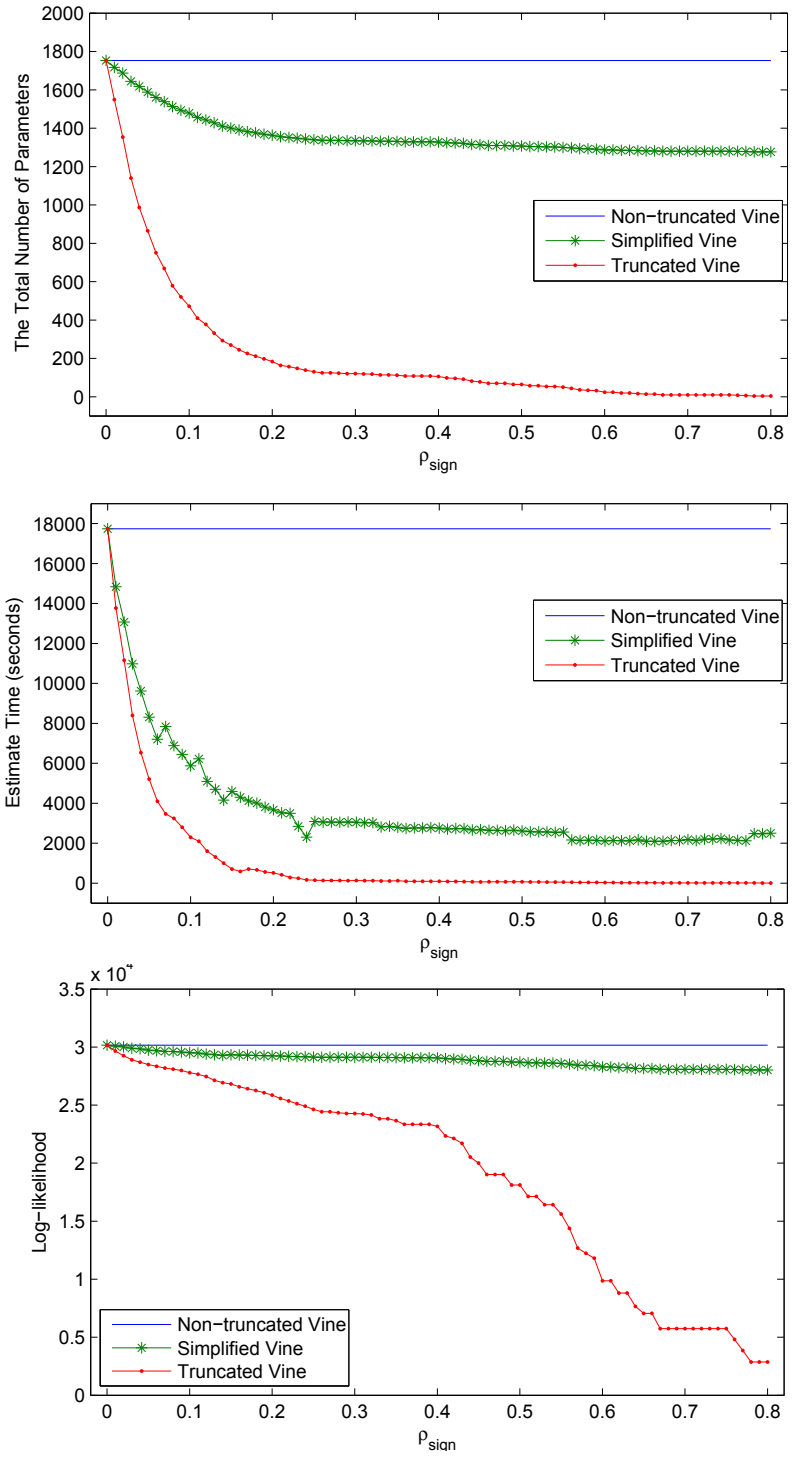

Figure 3.6: The results of truncated and simplified partial canonical vine against ρ_{sign}

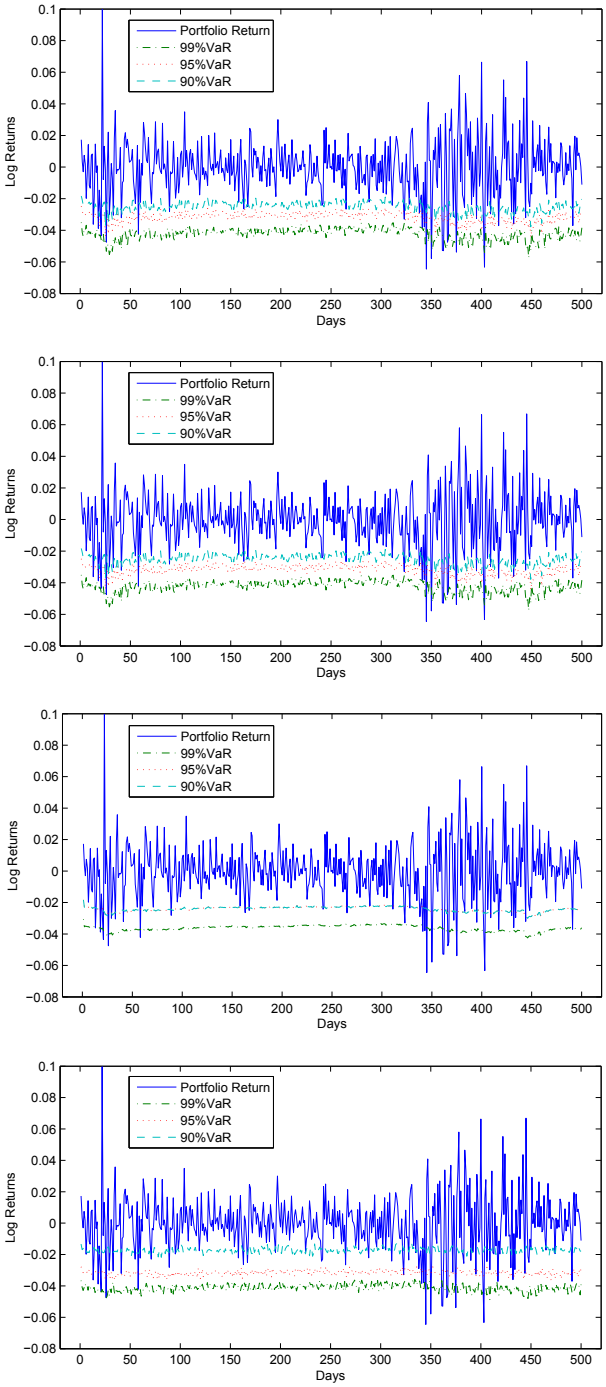

Figure 3.7: The Value at Risk forecasting of portfolio returns among models. From top to bottom figures, they are truncated PCV with ρ_{sign}=0.1, truncated PCV ρ_{sign}=0.4, non-truncated PCV and CAVA.

Chapter 4

Modeling Asymmetry and Tail Dependence among Multiple Variables by Using Partial Regular Vine

4.1 Introduction

Learning dependencies among high-dimensional variables, has been widely studied and applied in a large number of areas, such as social media and financial markets. Existing studies focus on the degree of dependence, however, few of them focus on the other important aspect of dependence–the dependence structure, especially the asymmetric and tail dependence characteristics. Dependence structure studying plays an important role in the financial area, especially in portfolio investment theory. Typical models of investment theory impose strong restriction on the dependence structure, which does not consider the asymmetric characteristics. It results in models which do not reflect the scenario in real world. For example, in the cross-country stock market, the typical investment theory suggests portfolio diversification. However, it is useless when all stocks tend to fall as the

market falls, which is demonstrated in the 2008 global financial crisis. It indicates that stock returns have stronger dependence during bear markets than bull markets, which means that stock returns may fall together, rather than boom together.

Recently, copula based dependence modelling emerges as a promising tool. Copula based dependence modelling is free of the linear correlation restriction, and allows dependence and correlation to vary over time. It uses the correlation/conditional correlation to capture the natural dependence, and at the same time, it can build flexible structures to model complex high-dimensional dependencies structures. In order to model the asymmetric dependence with high-dimensional financial variables, it is essential to develop a flexible dependence model with parametric copula families, which are suitable to multivariate data with various dependence structures. Hence, the model should have desired properties, which are described as follows:

(i). Flexible dependence structure, without imposing any assumptions or restrictions;

(ii). Wide range of dependence, allowing for both positive and negative dependencies;

(iii). Flexible range of tail dependence, allowing for various lower and upper tail dependencies;

(iv). Computationally feasible estimation for the joint density functions.

The existing multivariate copula models with parametric families did not satisfy all of the above conditions. Typically, the multivariate Archimedean copula model has the structure with only a narrow range of negative dependence (McNeil & Nešlehová 2009). The multivariate Gaussian copula model is not suitable to model the asymmetric characteristics, since (1) the Gaussian copula does not have lower and upper tail dependence, and (2) the Gaussian assumptions are not appropriate in the real world (Abdous et al. 2005, Fang et al. 2002). The multivariate t copula model, which is studied by (Demarta

& McNeil 2005, Nikoloulopoulos et al. 2009), does not have flexible lower and upper tail dependence since the t copula has the same lower and upper tail dependence. Canonical vine or D vine copula model, such as (Aas, Czado, Frigessi & Bakken 2009, Aas & Berg 2009), have a wide range of dependence by choosing appropriate bivariate copula families. However, they do not have a flexible dependence structure due to their structurale assumptions. These assumptions which are imposed on the dependence structure lead to a dependence structure which may not reflect the actual dependence in high-dimensional data.

In order to fulfill the above needs, a new partial correlation-based regular vine copula model with asymmetric dependence is proposed in this chapter. The new model can capture asymmetric dependence in high-dimensional data. The new model employs regular vine theory to construct the dependence structure, in which it does not impose any strong restriction on the dependence structure. Hence, it can reflect the actual dependence structure of high-dimensional data. In addition, the copula family with flexible lower and upper tail dependence connects with the new models, which ensure the new model has a wide range of lower and upper tail dependence.

This chapter has these contributions: (1) The chapter develops a new partial correlation based algorithm to construct the regular vine structure, which is called partial regular vine. The partial regular vine can uniquely determine the correlation matrix and be algebraically independent. It indicates that dependence structure constructed via partial correlation is more flexible, since the current tree structure is independent from the established tree structure and bivariate copulas selection. Hence, the model does not impose any strong restriction on the dependence structure; (2) For all linking bivariate copulas on the partial regular vine dependence structure, only the BB1, survival BB1, BB7 and survival BB7 are used, since these copula have both lower and upper tail dependence that can range independently from 0 to 1, and; (3) In the literature, it is the first time to analyse the moving trends of lower and upper tail dependence with the high-dimensional data

structure. In addition, the trends of lower and upper tail dependence during the dynamic period are also analysed.

The rest of this chapter is structured as follows. The partial correlation-based regular vine model is introduced in Section 4.2. Section 4.3 discusses how to estimate parameters in the partial regular vine copula and marginal distribution. Value at Risk, which is the popular evaluation method used in the financial market, and its related tests are discussed in Section 4.4. Section 4.5 shows the case study results in cross-country stock markets. Finally, Section 4.6 concludes the chapter.

4.2 Regular Vine Copula Model with Asymmetric Dependence

4.2.1 Regular Vine Copula Specification

The regular vine theory and its related definitions have been introduced in Section 2.4.3. Thus, the regular vine copula model specification is presented as follows:

Definition 4.1 (Regular Vine Copula Specification) *A regular vine copula specification on n variables is a multivariate distribution function is defined as* $C = (V^R_{\rho_e}, B(V^R_{\rho_e}), \theta(B(V^R_{\rho_e})))$

(i). $V^R_{\rho_e}$ *is a vine structure on n variables;*

(ii). $B(V^R_{\rho_e}) = \{C_{e(a),e(b)|D_e} | e_i \in E_i, i = 1, ..., n-1\}$ *is the set of* $n(n-1)/2$ *copula families; and*

(iii). $\theta(B(V^R_{\rho_e})) = \{\theta_{e(a),e(b)|D_e} | e \in E_i, i = 1, ..., n-1\}$ *is the set of parameters, corresponding to the copula family in* $B(V^R_{\rho_e})$.

Based on the definition of regular vine specification, the full specification of a regular vine copula has three components: the vine tree structure $V^R_{\rho_e}$, the copula family set $B(V^R_{\rho_e})$, and the corresponding copula parameters

$\theta(B(V^R_{\rho_e}))$. Then, there is a corresponding density distribution that realises the regular vine copula specification, which is given as follows:

$$f_{1:n}(\mathbf{x}|V^R_{\rho_e}, B, \theta) = \prod_{k=1}^{n} f_k(x_k) \cdot \prod_{i=1}^{n-1} \prod_{e \in E_i} c_{e(a),e(b)\,;\,D_e}(F_{e(a)|D_e}(x_{e(a)}\,;\,\mathbf{x}_{D_e}), F_{e(b);D_e}(x_{e(b)}\,;\,\mathbf{x}_{D_e})) \tag{4.1}$$

where $\mathbf{x} = (x_1, x_2, ..., x_n)$, $e = a, b \in E$ and $\mathbf{x}_{D_e}$ stands for the variables in D_e. f_i is denoted as the density function of the corresponding F for $i = 1, ..., n$. The corresponding density function of multivariate vine copula can be factorized in terms of many bivariate copulas, hence, various vine tree structures V can be constructed. For n-dimensional regular vine, there are $(n-1)$ bivariate copulas at tree level 1, and $(n-2)$ bivariate copulas at tree level 2. Typically, there are $(n-l)$ bivariate copulas in tree level l for $l = 2, ..., n-1$.

4.2.2 Regular Vine Tree Structure Construction

The construction of vine tree structure $V^R_{\rho_e}$ is discussed at first, and then the selection of bivariate copula for all edges in the vine tree structure is discussed. The regular vine tree structure is a dependence structure which connects all bivariate copula together. For a dependence structure with dimension < 5, the vine tree structure is either the canonical vine or D vine. However, for a dependence structure with high dimensions (dimension ≥ 5), there are three dependence structures, including regular, canonical or D vine. The canonical and D vine are two boundary cases of regular vine. The canonical vine imposes restrictions on the vine tree dependence structure, in which each variable connect to one variable in each tree. Hence, the canonical vine has a star-like structure. D vine has the restriction that each variable links to no more than two variables, which results in a flat-path-like structure. Regular vine, which does not impose any assumption or restriction on the vine structure, can reflect the actual dependence structure of the high-dimensional data set.

Hence, for high-dimensional data, regular vine is better than canonical vine or D vine.

A new partial correlation-based algorithm is developed to construct the regular vine, which is called the *partial regular vine*. The partial correlation is used to produce the regular vine tree structure, since (1) The partial correlation is obtained directly from the data, without knowing any structure or parametric assumption. The partial correlation ρ is defined by:

$$\rho_{1,2:3,\ldots,n} = -\frac{K_{12}}{\sqrt{K_{11}K_{22}}} \tag{4.2}$$

where K_{ij} is denoted as the (i,j) cofactor of the correlation matrix. The partial correlation can be computed from correlation with the following recursive formula:

$$\rho_{1,2:3,\ldots,n} = \frac{\rho_{1,2:3,\ldots,n-1} - \rho_{1,n:3,\ldots,n-1} \cdot \rho_{2,n:3,\ldots,n-1}}{\sqrt{1-\rho^2_{1,n:3,\ldots,n-1}}\sqrt{1-\rho^2_{2,n:3,\ldots,n-1}}} \tag{4.3}$$

Obviously, $\rho_{1,2}$ is equal to correlation. Hence, when building the partial regular vine tree structure, the current vine tree structure is independent from the previous tree structure, and (2) As discussed in Section 4.2.1, for copula selection, regular vine tree structure built by partial correlation can ensure that copula selection in current tree structure is independent from the copula selection result in the previous tree. Hence, the partial correlation-based regular vine tree structure ensures that it can focus on the dependence structure itself, ignoring the effect of the different copula family selection.

The algorithm to construct the partial regular vine tree structure is given in Algorithm 4.1. Based on the partial correlation definition, the partial correlation is equal to the correlation in the first tree T_1. In this chapter, Kendall's tau τ is used to measure the correlation between any two variables, since it can measure dependence independently of the assumed distribution. Hence, in T_1, $\rho_{1,2} = \tau_{1,2}$. To build the first tree T_1, the partial correlations ρ for all possible pair variables are calculated. Then, the *Maximum Spanning Tree* (MST) is employed to find the vine tree structure in T_1. Typically, the Algorithm of Prim is used to produce a Minimum Spanning Tree (Cormen,

Algorithm 4.1 Regular Vine Construction via Top-to-Bottom Strategy

Require: Observations of n variables

1: Calculate partial correlation $\rho_{x,y\,;\,I\setminus\{x,y\}}$ for all possible pair variables $\{x,y\}, x,y \in \{1,...,n\} = I$.

2: Find the *Maximum Spanning Tree* (MST), which can maximise the sum of absolute value of partial correlation ρ, such as: $max \sum |\rho_{x,y\,;\,I\setminus\{x,y\}}|$.

3: **for** $j = 2,...,n-1$ **do**

4: In T_j, based on the structure in T_{j-1}, find all possible edges $\{e(p), e(q)\,;\,D_e\}$ which are part of tree T_j, where $e = \{p,q\} \in I$, and $\{p,q\} \notin \{x,y\}$.

5: Ensure that these edges satisfy the proximity condition in Definition 1;

6: Choose MST which can maximise the sum of absolute value of partial correlation, i.e., $max \sum |\rho_{e(p),e(q)\,;\,D_e}|$, where ρ is partial correlation.

7: **end for**

8: **return** Partial regular vine tree structure

Leiserson, Rivest, Stein et al. 2001). However, Algorithm of Prim can work in both ways. That means that Algorithm of Prim can also produce the Maximum Spanning Tree. By using the MST, a large number of regular vine tree structures will be built, and the structure which can maximise the sum of absolute value of partial correlation ρ can be selected. Once the first tree structure is identified, the following trees are built and they employ a similar strategy to that used in the first tree building. For tree building from T_2 to T_{n-1}, all edges must satisfy the proximity condition mentioned in Section 4.2.1. The partial regular vine is obtained by assigning a partial correlation ρ with a value chosen arbitrarily in the interval (-1,1) to each edge e that is defined in Section 4.2.1. Therefore, the most important advantage of the partial regular vine is that it can uniquely determine the correlation matrix and algebraically independent. The limitation of partial regular vine is that the partial regula vine structure V is built on elliptical

Table 4.1: The Tail Dependence of Copula Family

	Lower Tail Dependence	Upper Tail Dependence
Gaussian	-	-
t	$2t_{\nu+1}(\mu(\nu,\phi))$	$2t_{\nu+1}(\mu(\nu,\phi))$
Gumbel	-	$2^{-1/\phi}$
Frank	-	-
Clayton	$2^{-1/\phi}$	-
Joe	-	$2-2^{1/\phi}$
BB1	$2^{-1/(\phi\delta)}$	$2-2^{1/\delta}$
S.BB1	$2-2^{1/\delta}$	$2^{-1/(\phi\delta)}$
BB6	-	$2-2^{1/(\delta\phi)}$
BB7	$2-2^{1/\delta}$	$2-2^{1/\phi}$
S.BB7	$2-2^{1/\phi}$	$2-2^{1/\delta}$
BB8	-	$2^{-1/\phi}$ when $\delta=1$

S.BB1 and S.BB7 are survival BB1 and BB7 copula respectively. ϕ and δ are parameters of the corresponding copula family. For t copula, $\mu(\nu,\phi)=\left(-\sqrt{\nu+1}\sqrt{\frac{1-\phi}{1+\phi}}\right)$.

copulas. However, the theorem in (Bedford & Cooke 2002) indicates that a partial regular vine structure can provide a bijective mapping from $(-1,1)^{\binom{n}{2}}$ into the set of positive definition matrices with 1's on the diagonal. Therefore, the partial regular vine structure is constructed firstly, and then mapped into the conditional correlation based regular vine dependence structure. Then, the whole structure is fitted with various copulas. The limitation of the partial regular vine can be removed, and the tree structure can be fitted with various copulas, other than the elliptical copula.

4.2.3 Bivariate Copula Family Selection

Once the vine tree structure is identified, the next step is to choose an appropriate bivariate copula for all edges. According to the theory in (Joe, Li &

Nikoloulopoulos 2010), if the multivariate uniform vector $\mathbf{U} = (U_1, ..., U_n) = (1 - U_1, ..., 1 - U_n)$, then $\mathbf{U}$ is a reflection of symmetry. If the copula density function $\mathbf{C} = c(u_1, ..., u_n) = c(1 - u_1, ..., 1 - u_n)$, then the vine is a reflection symmetric dependence structure. It means that if the copulas with the symmetric lower and upper tail dependence are selected, then it is a reflection of symmetric dependence structure. If modelling the asymmetric dependence, it is better to choose a copula with various levels of lower and upper tail dependence. Currently, there are a huge of copula families, which have various tail dependencies. The detail in relation to the tail dependence of copula families are listed in Table 5.1. For one-parametric copula, Gaussian and Frank copulas do not have any tail dependence, Clayton and Joe copulas have only lower tail dependence, and the Gumbel copula has only upper tail dependence. For two-parametric copulas, the t copula has symmetric upper and lower tail dependence, which reflects the symmetric dependence. BB1, S.BB1 BB7 and S.BB7 copulas have different lower and upper tail dependencies, where S.BB1 and S.BB7 copula are short for survival (rotated 180 degree) BB1 and BB7 copulas respectively. BB6 and BB8 copulas have only upper tail dependence. To capture the asymmetric characteristics, the BB1, S.BB1, BB7 and S.BB7 copulas should be the best choice since they have various levels of lower and upper tail dependence, which can vary independently from 0 to 1.

4.3 Marginal Distribution Specification and Parameter Estimate

The marginal distribution specification and parameter estimate are discussed in this Section. According to Equation (2.27), the multivariate joint density function has two parts, the first is the multivariate copula mentioned in the above section, the second is the marginal distribution. For financial data, $ARMA(1,1)$-$GARCH(1,1)$ model is the best choice for the marginal distribution (Jondeau & Rockinger 2006*a*, Fantazzini 2009). Typically, let

$X_t(t = 0, 1, ..., Z)$ be a time series of the price on a financial asset, such as stock the market index. Then the return of financial assets can be defined as $log(X_t/X_{t-1})$. Suppose there are n assets with returns $r_{t,1}, ..., r_{t,n}$. The estimate of partial regular vine copula model can be processed in two steps. In the first step, the appropriate marginal distribution of the variable (i.e. the financial asset) is selected. Due to the character of financial assets, such as the volatility cluster, a common choice is $ARMA(1,1)$-$GARCH(1,1)$ with Student-t innovations, which is defined as follows:

$$\begin{aligned} r_{j,t} &= c_j + \Phi_j r_{j,t-1} + \Theta_j \varepsilon_{j,t-1} + \varepsilon_{j,t}, \\ \varepsilon_{j,t} &= \sigma_{j,t} \cdot Z_{j,t} \\ \sigma^2_{j,t} &= \omega_j + \alpha_j \varepsilon^2_{j,t-1} + \beta_j \sigma^2_{j,t-1} \end{aligned} \tag{4.4}$$

where $j = 1, ..., n, t = 1, ..., \mathfrak{T}$ and $Z_{j,t}$ is the innovations which follow Student-t distribution. Let $\theta^m_j = (c_j, \Phi_j, \Theta_j, \omega_j, \alpha_j, \beta_j)$ be denoted as the parameter set of marginal distribution. Let θ^c be denoted as the parameters of the multivariate copula functions. The multivariate joint log-likelihood is given by:

$$\begin{aligned} L(\theta^m_1, ..., \theta^m_n, \theta^c) &= \sum_{t=1}^{\mathfrak{T}} log f(r_{1,t}, ..., r_{n,t}; \theta^m_1, ..., \theta^m_n, \theta^c) \\ &= \sum_{t=1}^{\mathfrak{T}} \log c(F_1(r_{1,t}), ..., F_n(r_{n,t}); \theta^c) \\ &\quad + \sum_{t=1}^{\mathfrak{T}} \sum_{j=1}^{n} \log f_j(r_{j,t}; \theta^m_j) \end{aligned} \tag{4.5}$$

where the multivariate $c(\cdot\,; \theta^c)$ is denoted as the regular vine model. Maximum Equation (5.3) is possible. However, it is time consuming when n is large. $Inference\ Functions\ for\ Margins$ (IFM) method (detail can be found in Section 2.2.2 and (Joe 2005)) is used to resolve the issue. The IFM is two-step estimate method, which can efficiently estimate the parameters. In the first step, the marginal distribution $ARMA(1,1)$-$GARCH(1,1)$

is employed to filter the financial returns and the univariate parameters $\theta_j^m = (c_j, \Phi_j, \Theta_j, \omega_j, \alpha_j, \beta_j)$ are derived. In the second step, the joint log-likelihood in Equation (5.3) is maximised over copula parameters θ^c, and the univariate parameters $(c_j, \Phi_j, \Theta_j, \omega_j, \alpha_j, \beta_j)$ are fixed at the estimated value in the first step. It means that the joint log-likelihood is reduced to the equation which consists of only copula parameters due to the fact that the parameters of the log-likelihood are fixed.

4.4 Value at Risk– A widely Used Evaluation in Financial Market

Value at Risk (VaR) is a probabilistic metric of market risk and is an industrial golden benchmark for measuring market risk. VaR at the level $(1-\alpha)$ is defined by

$$VaR_t(1-\alpha) = -inf\{c \in \mathbb{R} : P(r_t \leq c|\digamma_{t-1})| \geq (1-\alpha)\} \tag{4.6}$$

where $\digamma_{t-1}$ represents the past information at time $t-1$. For a good model, it is capable of producing a high quantity of VaR. Given a set of financial returns, such as stock indices, the portfolio returns can be defined as:

$$r_{t,portfolio} = \sum_{j=1}^{n} \mu_j r_{j,t} \tag{4.7}$$

Suppose the current time is t, the process for computing VaR forecasting at time $t+1$ is given as follows:

(i). Fit $ARMA(1,1)$-$GARCH(1,1)$ with Student-t innovations with returns by using Equation (4.4) Then, the standardised residuals is obtained by:

$$\hat{Z}_{j,t} = \frac{r_{t,j} - \hat{c}_j - \hat{\Phi}_j r_{t,j-1} - \hat{\Theta}_j \hat{\sigma}_{t-1,j} \hat{Z}_{t-1,j}}{\hat{\sigma}_{t,j}} \tag{4.8}$$

(ii). The ex-ante $GARCH$ variance forecast for $j = 1, ..., n$ can be computed as follows:

$$\hat{\sigma}^2_{t+1,j} = \hat{\omega}_j + \hat{\alpha}_j \hat{\varepsilon}^2_{t,j} + \hat{\beta}_j \hat{\sigma}^2_{t,j} \tag{4.9}$$

(iii). The standardised residuals obtained from $ARMA(1,1)$-$GARCH(1,1)$ are transformed to approximately uniform data $\mathbf{u}_j = u_{1,j}, ..., u_{t,j}$ by using the Student-t cumulative distribution function;

(iv). Fit a regular vine structure with approximately uniform data $\mathbf{u}_j$ and estimate the parameters of the copula;

(v). Use the fitted regular vine structure with estimated copula parameters to simulate a sample for each financial return variable, i.e.,$v_{t+1,j}$;

(vi). Transfer the sample to standard residuals by using the inverse Student-t cumulative probability distribution functions with parameters obtained in Step (i), and then obtain the simulated standardised residuals, i.e., $\hat{e}_{t+1,j}$;

(vii). Calculate the one day forecast return and variance for each financial variable by using the estimated $ARMA(1,1)$-$GARCH(1,1)$ which is calculated in Step (i), i.e.

$$\hat{r}_{t+1,j} = c_j + \hat{\Phi}_j r_{t,j} + \hat{\Theta}_j \hat{\varepsilon}_{t,j} + \hat{\varepsilon}_{t+1,j} \tag{4.10}$$

(viii). The portfolio return is calculated by using Equation (5.6). Then, Steps from (iv) to (vii) are repeated for N times (e.g. $N = 10000$). Then, the 99%, 95%, and 90% VaR forecasting is determined by taking the corresponding 1%, 5% and 10% quantiles of the portfolio return forecast respectively.

To validate the VaR forecasting, the test of ex-post exceedance is used, which is defined at time t as:

$$I_t = \begin{cases} 1, & \text{if} \quad r_{t,portfolio} < VaR_t(1-\alpha); \\ 0, & \text{otherwise.} \end{cases} \tag{4.11}$$

where $r_{t,portfolio}$ is the ex-post observed portfolio return at time t. If the VaR forecasting is accurate, I_t should be equal to the significance level α. In addition, the quality of VaR forecasting can be judged by backtesting methods, including unconditional, independent and conditional coverage tests, which are presented in Section 3.5.

4.5 Case Study

The daily log-return data of 8 major European indices are considered in the experiment, including the Athen Index Composite (GD.AT), ATX(ˆATX), Euronext BEL-20 (BFX), CAC40 (ˆFCHI), DAX(ˆGDAXI), FTSE 100 (ˆFTSE), SMI (ˆSSMI), and AEX (AEX.AS), where symbols are in the corresponding parenthesis. The number indicates the following indices of the European stock market: $v1$=GD.AT, $v2$=ˆATX, $v3$=ˆBFX, $v4$=ˆFCHI, $v5$=ˆGDAXI, $v6$=ˆFTSE, $v7$=ˆSSMI and $v8$=AEX.AS. The eight major indices cover the majority of European stock, which reflects the general trading situation of the European stock market. In particular, during period from 01/03/2006 to 28/12/2012, there were in total 1682 observations for the European indices. All the data was downloaded from FRB St. Louis (http://research.stlouisfed.org).

4.5.1 Non-parametric Dependence Analysis

Before building the model to fit the data, a non-parametric method [1] is implemented to analyse the lower and upper tail dependence. The results are shown in Table 4.2. For a total of 56 pairs, 46 pairs have a strong upper tail

[1]Frahm et al (Frahm, Junker & Schmidt 2005) proposed a non-parametric method to obtain the non-parametric estimator of lower and upper tail dependence by using 'Pickand's dependence function (Pickands 1981)'. One simple nonparametric estimator of tail dependence is the log estimator, which is denoted by:

$$\begin{aligned}\hat{\lambda}_{lower} =& 2 - \lim_{u^*\to 0} \frac{log\left(1-2(1-u^*)+T_{-1}\sum_{t=1}^{T}\mathbf{1}\{U_1\le 1-u^*,...,U_n\le 1-u^*\}\right)}{log(1-u^*)}\\ \hat{\lambda}_{upper} =& 2 - \lim_{u^*\to 0} \frac{log\left(T_{-1}\sum_{t=1}^{T}\mathbf{1}\{U_1\le 1-u^*,...,U_n\le 1-u^*\}\right)}{log(1-u^*)}\end{aligned} \tag{4.12}$$

For extreme value estimation, a threshold u^* need to be chosen for estimation, which can affect both the lower and upper tail dependence. The selection of threshold u^* is the actual trade-off variance in the estimator against bias (Frahm et al. 2005). In the chapter, the non-parametric method is used for roughly analysing the tail dependence coefficient before building the regular vine copula model.

Table 4.2: Non-parametric Tail Dependence Analysis

	$v1$	$v2$	$v3$	$v4$	$v5$	$v6$	$v7$	$v8$
$v1$		0.33	0.29	0.24	0.16	0.26	0.21	0.29
$v2$	0.18		0.23	0.28	0.36	0.21	0.43	0.25
$v3$	0.24	0.45		0.52	0.53	0.53	0.53	0.49
$v4$	0.11	0.34	0.46		0.63	0.60	0.48	0.61
$v5$	0.17	0.33	0.39	0.73		0.45	0.54	0.57
$v6$	0.18	0.31	0.51	0.58	0.50		0.39	0.41
$v7$	0.11	0.28	0.37	0.29	0.36	0.36		0.47
$v8$	0.14	0.37	0.46	0.52	0.55	0.48	0.29	

The values above (below) the diagonal are corresponding upper (lower) tail dependencies.

dependence, which indicates that their upper tail dependence is larger than their lower tail dependence. In addition, only 11 pairs have a small gap (less than 0.1) between the lower and upper tail dependence. These descriptive statistics indicate that for most financial returns, they have a stronger upper tail dependence than a lower tail dependence. Due to the large gap between the lower and upper tail dependence, it seems that the two kinds of tail dependencies are significantly different. Therefore, the regular vine copula model with asymmetric dependence can be used to check whether the two kind of tail dependencies are significantly different.

4.5.2 Regular Vine Copula Specification and Tail Dependence Analysis

Each index returns is fit with univariate $ARMA(1,1)$-$GARCH(1,1)$ with Student-t innovations. The tests of Box and Pierce (BP) (Box & Pierce 1970) and Ljoung and Box (LB) (Ljung & Box 1978) are employed for checking the autocorrelation of standardised residuals. Table 4.3 shows the result of the two tests, which indicates that there are no autocorrelations left for all

Table 4.3: Results of BP and LP Tests

j	Z_j (BP)	Z_j^2(BP)	Z_j(LP)	Z_j^2(LP)
GD.AT	0.573	0.199	0.573	0.199
^ATX	0.500	0.113	0.499	0.113
^FCHI	0.798	0.150	0.798	0.149
^GDAXI	0.319	0.315	0.318	0.315
^FTSE	0.993	0.152	0.993	0.152
^SSMI	0.766	0.656	0.766	0.655
AEX.AS	0.223	0.713	0.222	0.713

Z_j and Z_j^2 are standardised residuals and squared standardised residuals respectively from $ARMA$-$AGRCH$ fits. The values in the corresponding columns are the p-values for both the BP and LP tests.

indices in the standardise residuals e_j and squared standardised residuals e_j^2 (all p values > 0.05). Then, the standardised residuals are used as an argument of the partial regular vine copula.

The next step is to build the partial regular vine copula model. The vine tree structure V is obtained by using Algorithm 4.1. The Figure 4.1 shows the full tree structure that is built by Algorithm 4.1. in Figure 4.1, there are two main blocks in Tree 1. One is $v2$, which connects to three variables, $v1$, $v3$ and $v7$. Another is $v4$, which connects to four indices, such as $v5$, $v6$, $v7$ and $v8$. Once the structure is identified, the next step is to choose the copula for each edge. According to the analysis in Section 4.5.1, the bivariate copulas which can provide flexible lower and upper tail dependence are most appropriate to build the partial vine copula model with asymmetric dependence. Based on the Section 4.2.3, BB1, S.BB1, BB7 and S.BB7 copula can provide both lower and upper tail dependence. Therefore, the BB1, BB7, S.BB1 and S.BB7 copulas are used to build the regular vine copula model with asymmetric dependence to capture the asymmetric characteristics. In

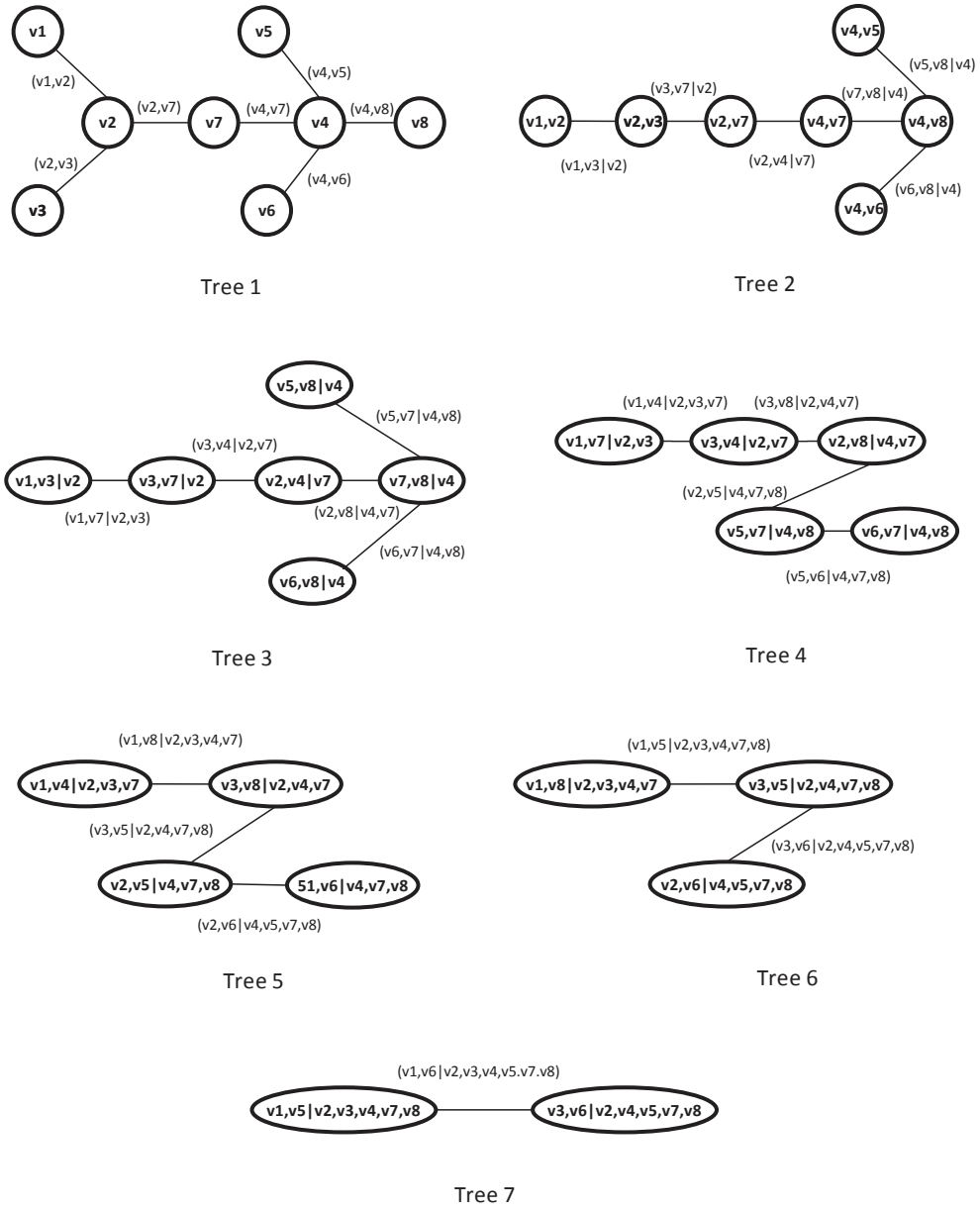

Figure 4.1: The partial regular vine tree structure

order to compare the performance of various copula, only one copula family (e.g., BB1) is used to fit the partial regular tree structure, which allow for easily assessing the performance of each copula family.

The Table 4.4 shows the tail dependence in Tree 1 of Figure 4.1 during

the period from 2006 to 2012. The non-parametric and t copula results are listed as a reference. The results show that the lower tail dependence of pairs in Tree 1 is less than their corresponding upper tail dependence. Various bivariate copulas provide different results. However, they show the same conclusion that the lower tail dependence is less than the upper one.

In order to investigate the tail dependence and its movement trends, a fixed period (e.g., one year) is used as the investigation period of tail dependence. Moving windows of 890 observations corresponding to approximately 2.5 years of daily observations, from 6/02/2007 to 28/12/2012 are used in the experiment. The partial regular vine copula is re-estimated daily in moving windows to produce the tail dependence of the investigation period. While estimating the tail dependence of the investigation period over the moving windows, the vine tree structure (mentioned in Figure 4.1) is used as the partial regular vine. For copula selection, the mixed copula candidatures (including BB1, BB7, S.BB1 and S.BB7) are used to fit the vine tree structure. The selection criteria is based on the AIC, which means the copula candidatures are selected with the smallest value of AIC. In order to find the movement trend in different time periods, the 12, 24 or 36 months are used as the investigation period. The results of pair $\{v1, v4\}$ in tree 1 are shown in Figure 4.2. The top 2, middle 2 and bottom 2 figures in Figure 4.1 use the 12, 24, 36 months as the investigation periods respectively. The left 3 figures show the lower and upper tail dependence, and the right 3 figures indicate the corresponding difference between lower and upper tail dependence. The gap between the lower and upper tail dependence has sharply increase since January 2009. The gap in the short investigation period (12 months) is larger than those in the long investigation period (24 or 36 months). It indicates that the difference between the lower and upper tail dependence is more significant in the short investigated period than in the long one. However, the difference decreases over the length of the longer investigation period. For other pairs, these similar conclusions can be reached, which are shown by Figures 4.3 to 4.8.

4.5.3 Value at Risk Forecasting

The predictable performance of the model can be examined via Value at Risk (VaR) forecasting. Moving windows from 04/01/2007 to 28/12/2012, totally 1417 observations, corresponds to approximately 4 years of trading days. A training period from 04/01/2006 to 28/12/2006 with 264 observations is approximately 1 year of trading days. Then, the model is re-estimated daily to produce the one day ahead VaR forecasting. While re-estimating the regular vine copula model, the partial regular vine tree structure shown in Figure 4.1 is used. Various copulas are fitted in the regular vine copula model, in order to compare performance capabilities and assess whether the model with asymmetric dependence is better than the one with symmetric dependence. In addition, the canonical vine and D vine tree structure are fitted with copulas, in order to compare the performance and find whether the regular vine tree structure is better than canonical vine or D vine.

Table 4.5 shows the backtesting results of the partial regular vine, canonical vine and D vine with various copulas. The regular vine, canonical vine or D vine with the copula which are displayer in the second row of Table 4.5 are fitted with different copulas. BB1, S.BB1, BB7 and S.BB7 copulas have flexible lower and upper tail dependence, which is a reflection of asymmetric dependence. The t copula has symmetric lower and upper tail dependence, to reflect the symmetric dependence. The Clayton copula has only lower tail dependence, and the Gumbel copula has only upper tail dependence. The BB1 and S.BB1 copulas have the best performance, followed by the BB7 and S.BB7 copulas. The model with t copula is better than the Clayton, Gumbel and BB6 copulas which have only one tail dependence. Figure 4.9 shows the corresponding VaR forecastings that are produced by the regular vine with the BB1 and t copulas.

In conclusion, the results of VaR forecasting indicate that (1) the partial regular vine copula with asymmetric dependence is better than the one with symmetric lower and upper tail dependence, and (2) the models with two tail dependencies are better than those with only one tail dependence.

4.6 Summary

It is a very challenging task to model high-dimensional and asymmetric dependence. Existing research has made only partial progress regarding high-dimensional asymmetric dependence modelling. This chapter has proposed a partial correlation-based regular vine copula model to address this challenging issue. It has been demonstrated by analysing the asymmetric dependence in cross-country stock markets.

Table 4.4: Tail Dependence Analysis by Using Various Copula

	Non-para*		BB1		S.BB1**		t		BB7		S.BB7**	
	λ_L	λ_U	λ_L	λ_U	λ_L	λ_U	λ_L	λ_U	λ_L	λ_U	λ_L	λ_U
$\{v4, v6\}$	0.50	0.60	0.63	0.76	0.71	0.78	0.54	0.54	0.76	0.82	0.77	0.82
$\{v4, v5\}$	0.73	0.63	0.74	0.81	0.78	0.82	0.71	0.71	0.83	0.85	0.82	0.87
$\{v4, v8\}$	0.52	0.61	0.71	0.79	0.77	0.81	0.62	0.62	0.80	0.84	0.81	0.84
$\{v4, v7\}$	0.29	0.48	0.60	0.69	0.66	0.75	0.45	0.45	0.71	0.75	0.72	0.75
$\{v2, v7\}$	0.28	0.43	0.40	0.55	0.50	0.67	0.18	0.18	0.54	0.62	0.57	0.61
$\{v2, v3\}$	0.45	0.23	0.45	0.61	0.56	0.69	0.37	0.37	0.61	0.68	0.63	0.67
$\{v1, v2\}$	0.18	0.33	0.28	0.41	0.39	0.62	0.25	0.25	0.41	0.48	0.46	0.44

* Non-para means that the tail dependence coefficient is calculated via non-parametric method;

** S.BB1 and S.BB7 are the survival BB1 and BB7 copula respectively.

Table 4.5: The Backtesting Results of Value at Risk Forecasting

		Partial Regular Vine							Canonical Vine			D Vine		
	$1-\alpha$	BB1	S.BB1	BB7	S.BB7	t	Clayton	Gumbel	BB1	S.BB1	BB7	BB1	S.BB1	BB7
PoF	99%	14	14	14	14	14	14	14	14	14	14	14	14	14
		14	14	13	13	15	15	21	15	14	15	13	14	14
	95%	70	70	70	70	70	70	70	70	70	70	70	70	70
		72	73	73	74	71	70	81	74	74	76	75	76	74
	90%	141	141	141	141	141	141	141	141	141	141	141	141	141
		134	131	140	137	130	136	136	135	132	145	138	141	140
LR_{uc}	99%	0.002	0.002	0.100	0.100	0.048	0.354	3.74	0.048	0.032	0.248	0.100	0.062	0.100
		(0.964)	(0.964)	(0.751)	(0.751)	(0.826)	(0.552)	(0.053)	(0.826)	(0.858)	(0.618)	(0.751)	(0.803)	(0.751)
	95%	0.020	0.068	0.068	0.145	0.051	0.224	1.466	0.145	0.145	0.385	0.220	0.068	0.145
		(0.899)	(0.794)	(0.794)	(0.703)	(0.821)	(0.636)	(0.264)	(0.703)	(0.703)	(0.309)	(0.488)	(0.273)	(0.225)
	90%	0.473	0.919	0.023	0.175	1.101	1.101	0.258	0.357	0.753	0.385	0.423	0.917	0.175
		(0.492)	(0.338)	(0.880)	(0.676)	(0.294)	(0.294)	(0.612)	(0.550)	(0.385)	(0.535)	(0.338)	(0.338)	(0.676)
LR_{cc}	99%	0.282	0.282	0.341	0.341	0.369	0.859	4.434	0.369	0.282	0.569	0.641	0.382	0.541
		(0.869)	(0.869)	(0.843)	(0.843)	(0.831)	(0.651)	(0.109)	(0.831)	(0.869)	(0.831)	(0.467)	(0.171)	(0.869)
	95%	1.436	2.599	2.599	2.479	1.946	1.729	2.662	1.685	1.785	2.351	1.436	2.530	1.436
		(0.488)	(0.273)	(0.273)	(0.290)	(0.378)	(0.421)	(0.264)	(0.431)	(0.410)	(0.309)	(0.488)	(0.826)	(0.488)
	90%	1.467	2.316	0.862	0.837	1.613	1.997	1.023	1.633	2.008	1.374	2.316	3.447	1.387
		(0.480)	(0.314)	(0.650)	(0.658)	(0.446)	(0.369)	(0.600)	(0.442)	(0.366)	(0.503)	(0.314)	(0.826)	(0.763)
LR_{ic}	99%	0.280	0.280	0.241	0.241	0.321	0.505	0.694	0.321	0.250	0.321	0.541	0.320	0.441
		(0.597)	(0.597)	(0.624)	(0.624)	(0.571)	(0.477)	(0.405)	(0.571)	(0.617)	(0.571)	(0.462)	(0.572)	(0.507)
	95%	1.417	2.531	2.531	2.334	1.895	1.505	1.196	1.540	1.640	1.966	1.217	2.531	2.834
		(0.234)	(0.112)	(0.112)	(0.127)	(0.169)	(0.220)	(0.274)	(0.215)	(0.200)	(0.161)	(0.270)	(0.112)	(0.092)
	90%	0.994	1.397	0.839	0.662	0.513	0.896	0.765	1.276	1.255	0.989	1.397	2.530	1.212
		(0.319)	(0.237)	(0.360)	(0.416)	(0.474)	(0.344)	(0.382)	(0.259)	(0.263)	(0.320)	(0.237)	(0.112)	(0.271)

The POF is percentage of failure. The first row shows the expected number of exceedances, and the following row is the actual number of exceedances. LR_{uc}, LR_{ic} and LR_{cc} are short for the likelihood ratio of conditional, independent and unconditional coverage respectively. The first row shows the value, while the corresponding p value is given the parenthesis in the following row. The critical value of LR_{uc} or LR_{ic} is 3.841, while the critical value of LR_{cc} is 5.991.

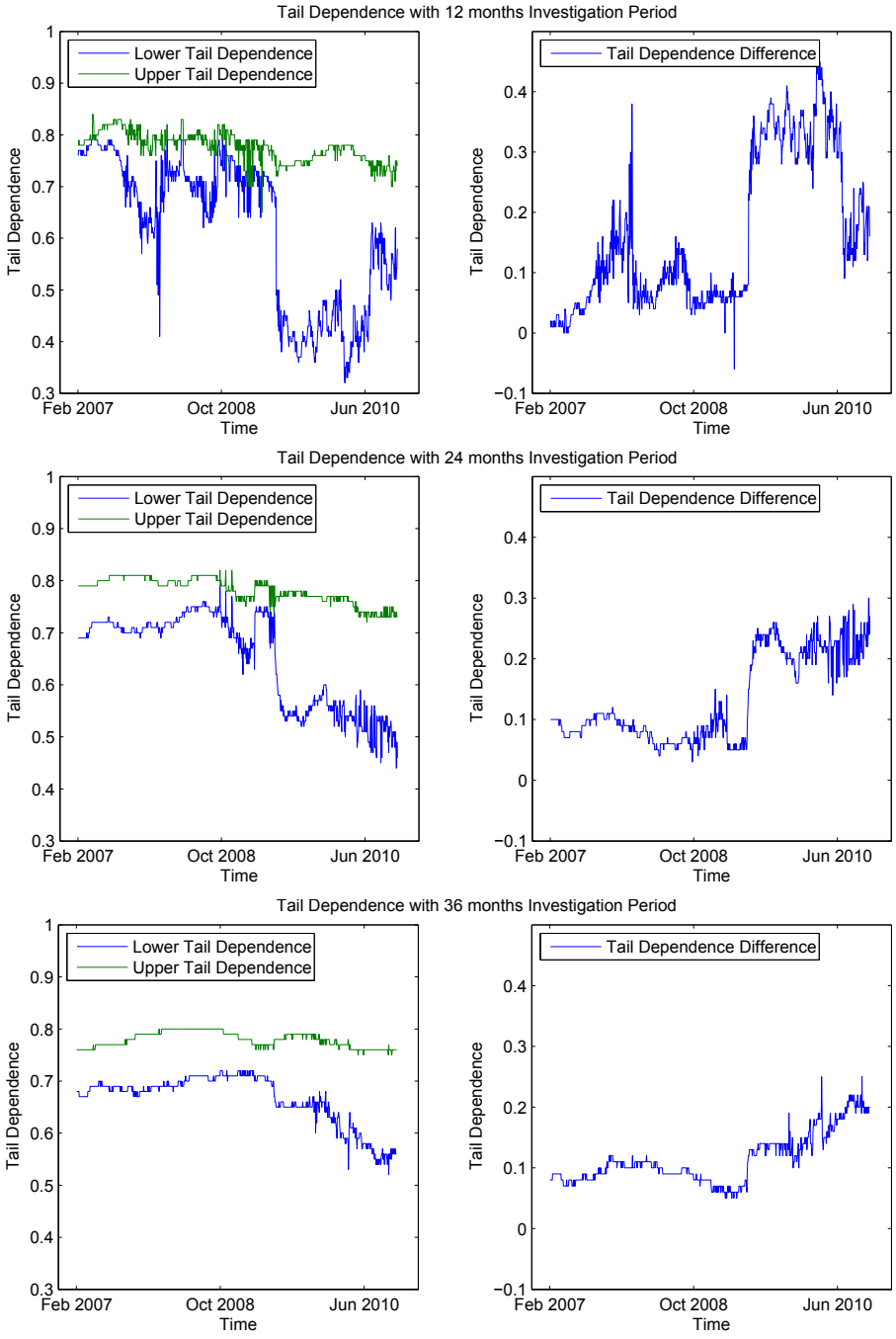

Figure 4.2: The lower and upper tail dependence of variables $v4$ and $v6$

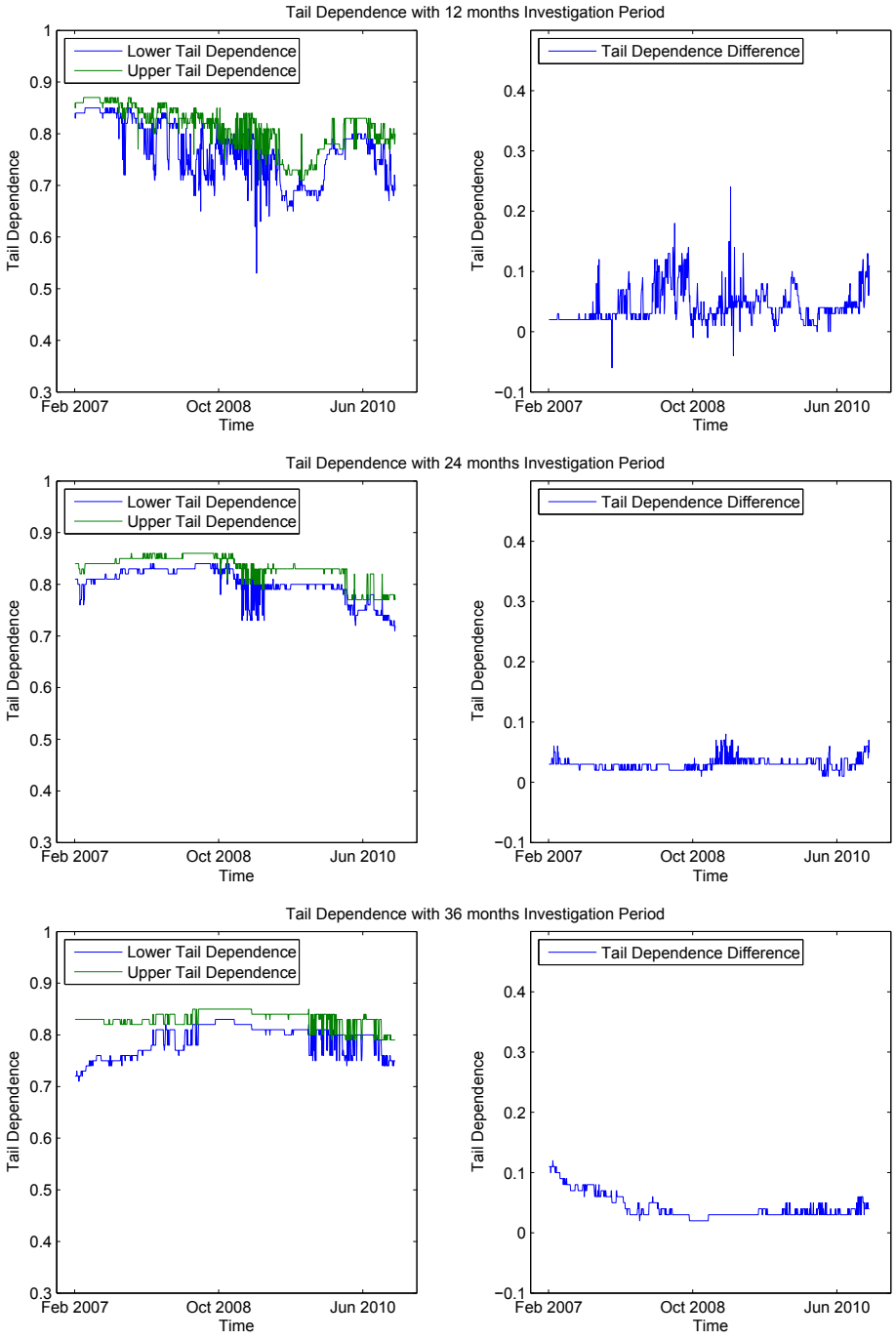

Figure 4.3: The lower and upper tail dependence of variables $v4$ and $v5$

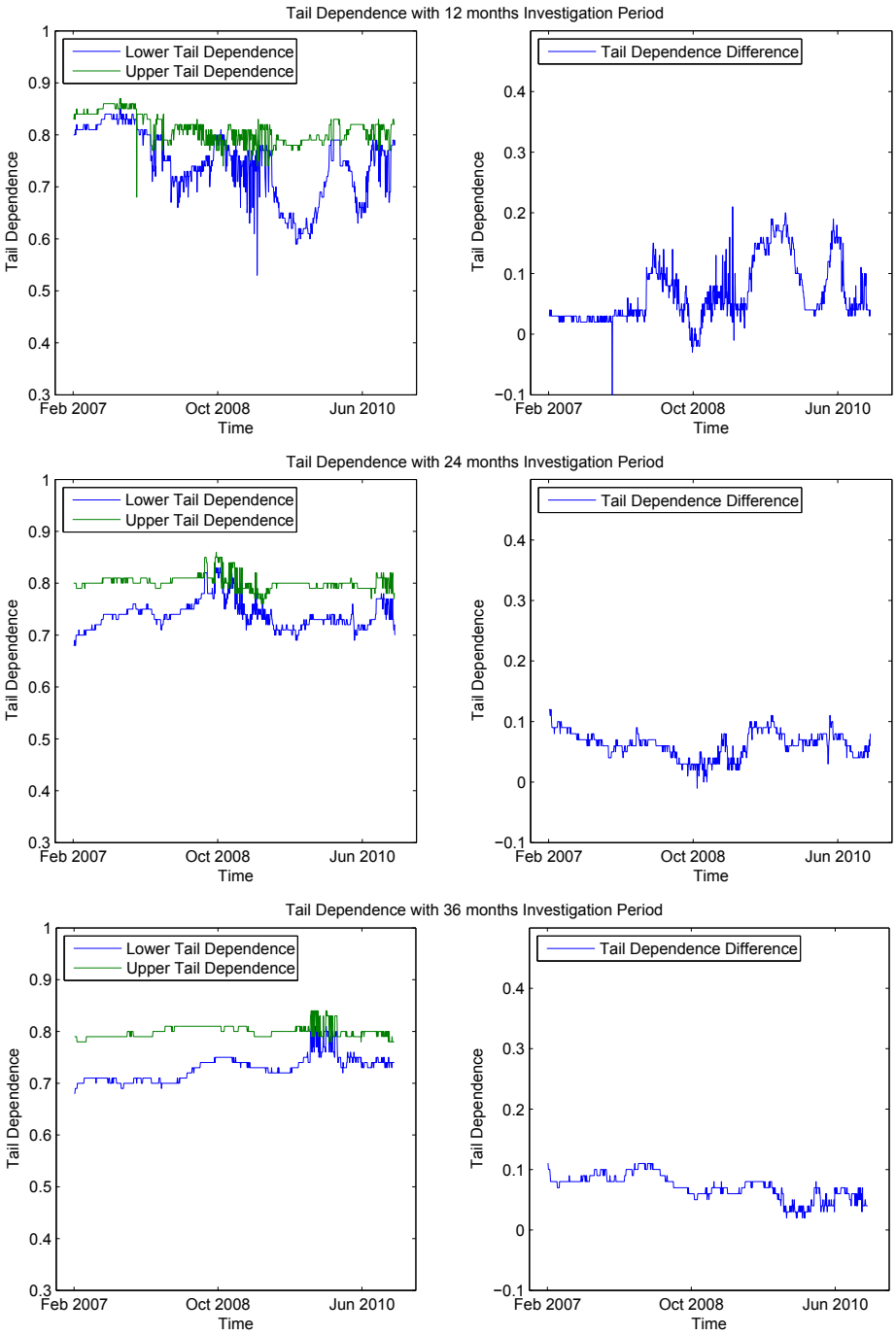

Figure 4.4: The lower and upper tail dependence of variables $v4$ and $v8$

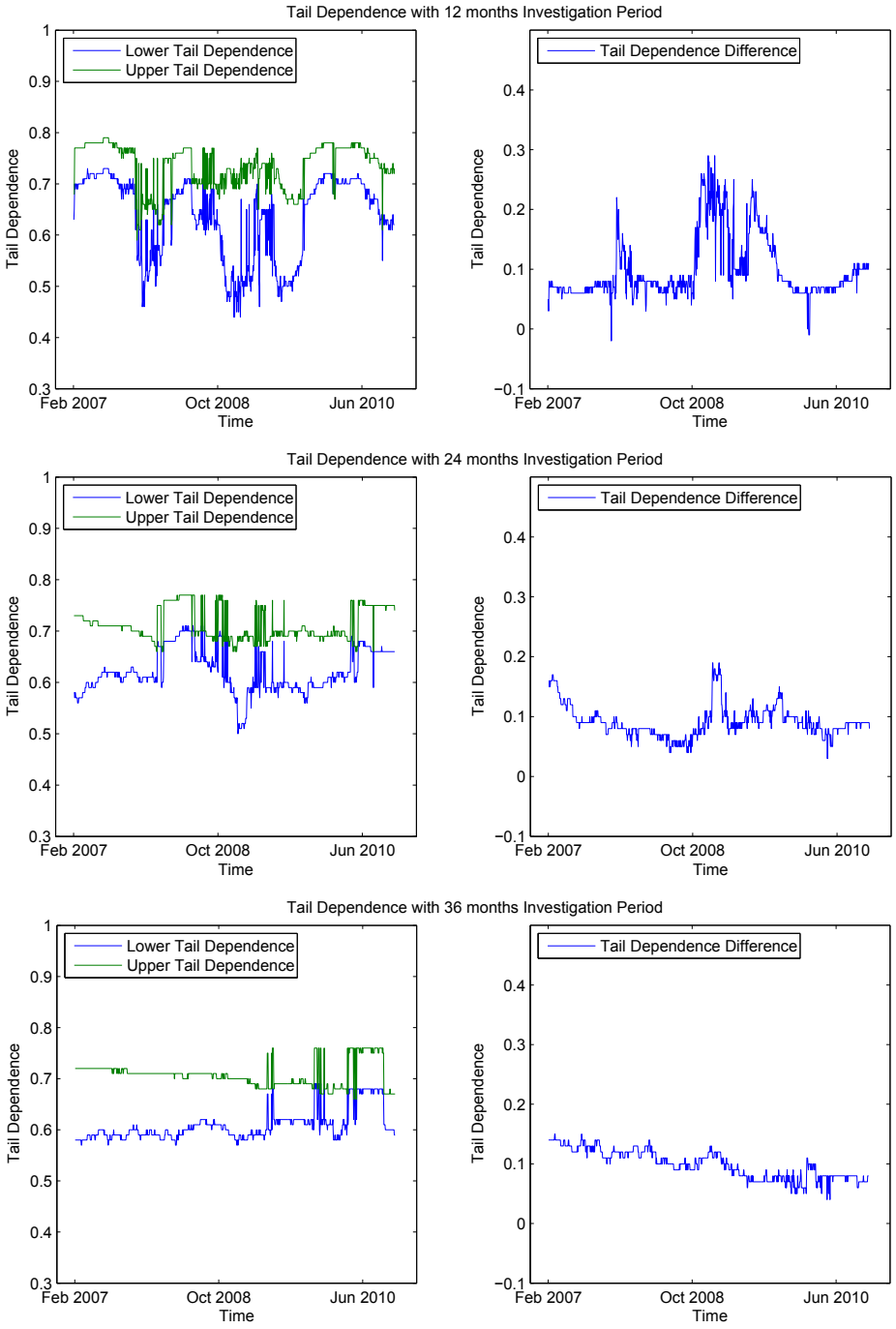

Figure 4.5: The lower and upper tail dependence of variables $v4$ and $v7$

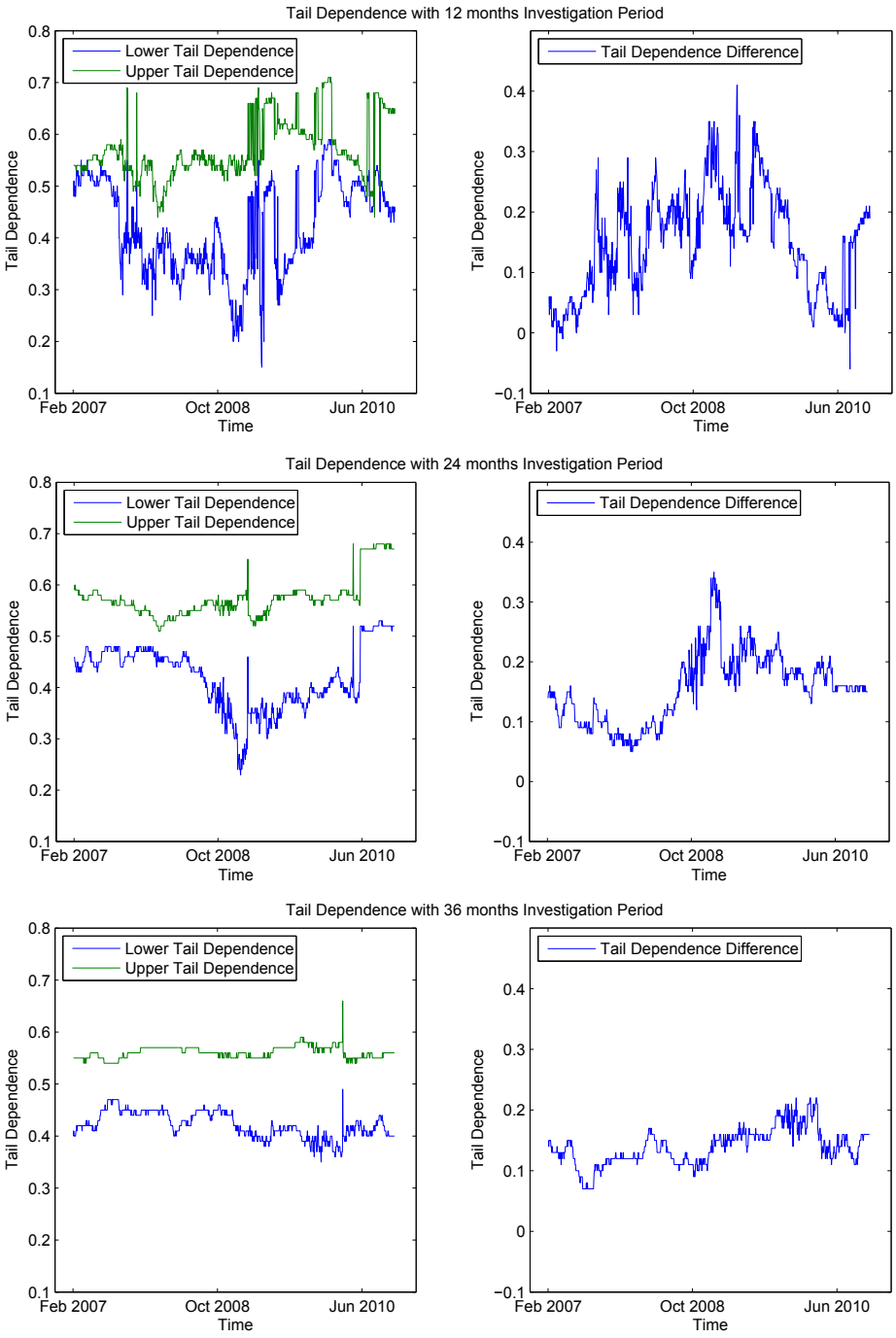

Figure 4.6: The lower and upper tail dependence of variables $v2$ and $v7$

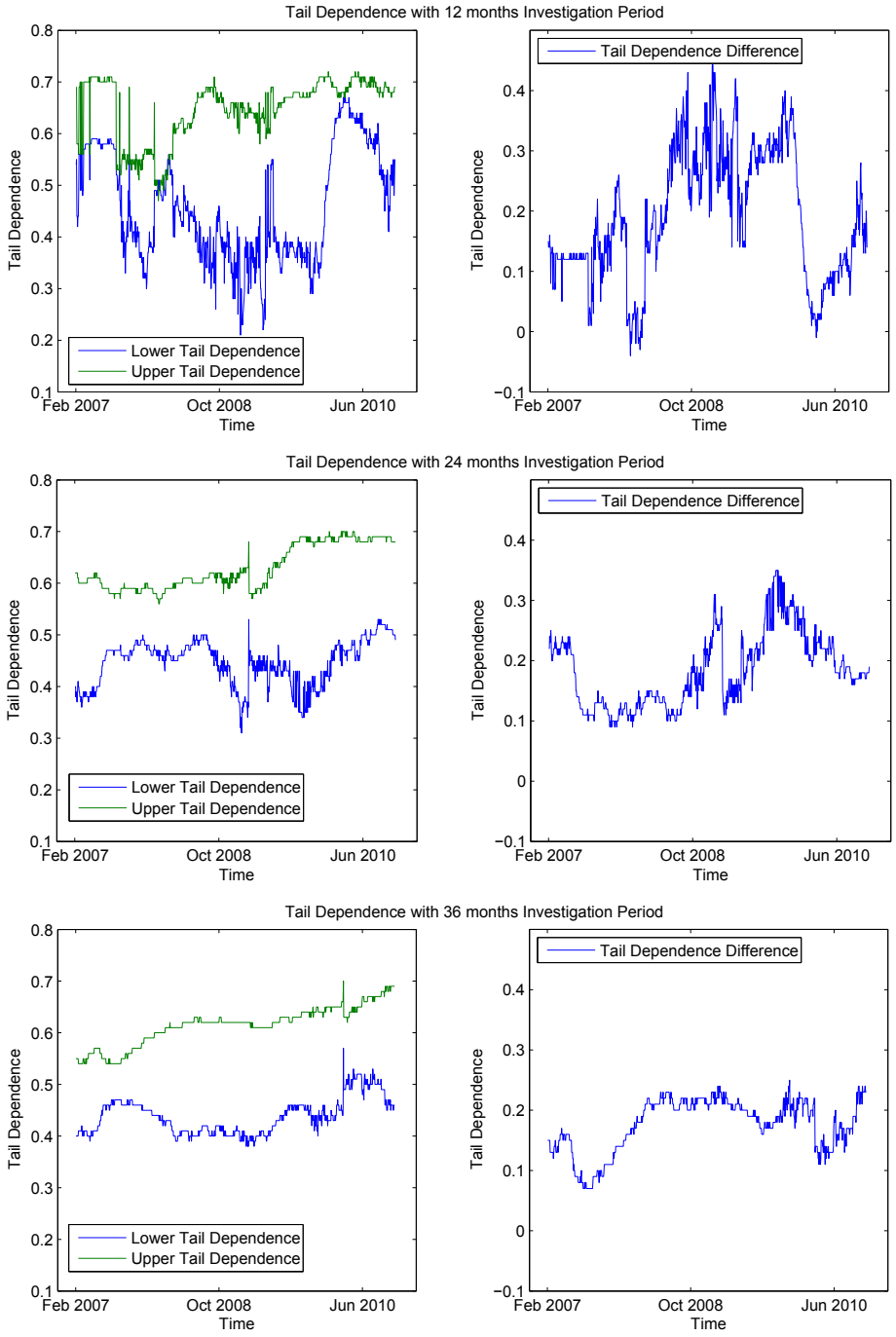

Figure 4.7: The lower and upper tail dependence of variables $v2$ and $v3$

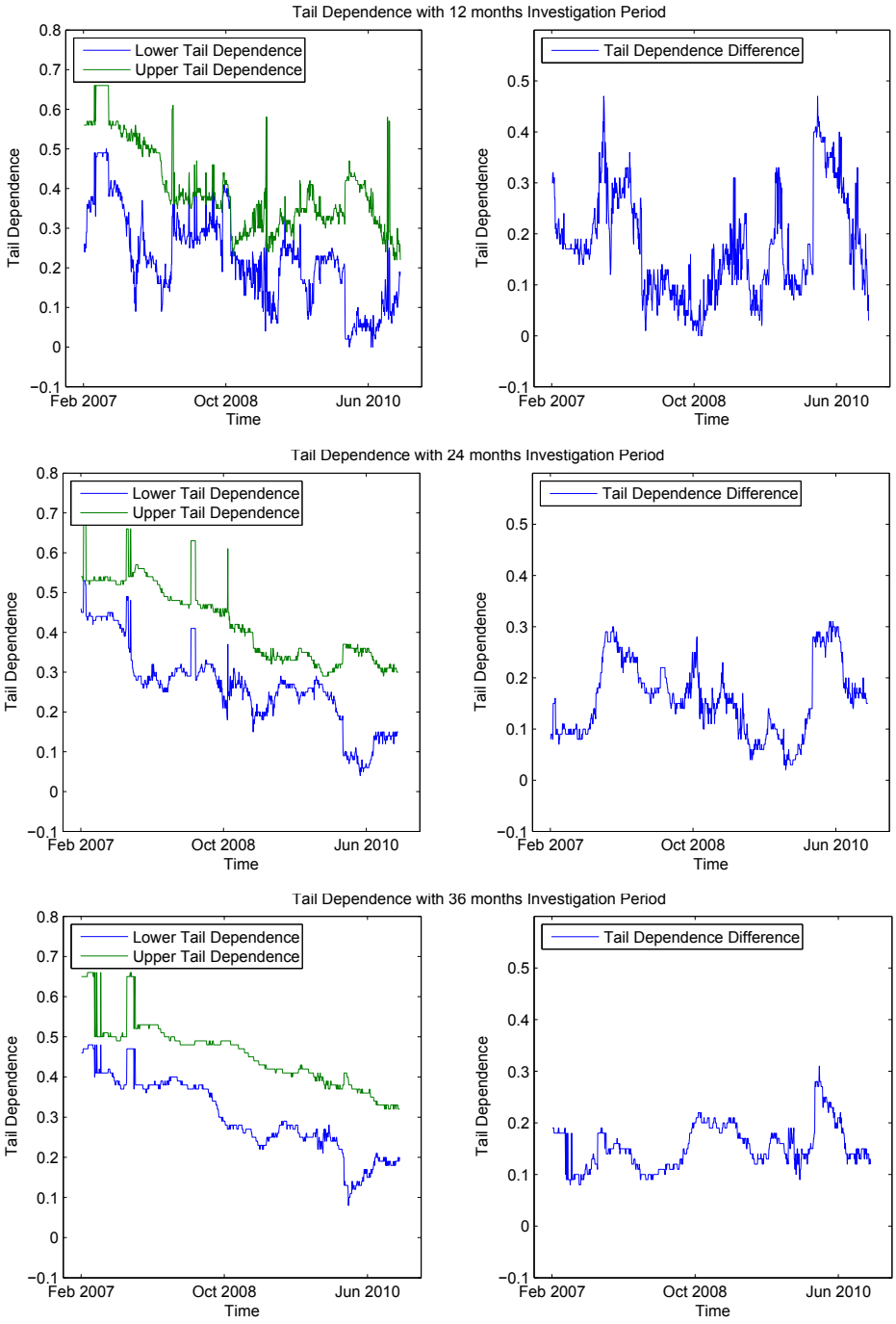

Figure 4.8: The lower and upper tail dependence of variables $v2$ and $v1$

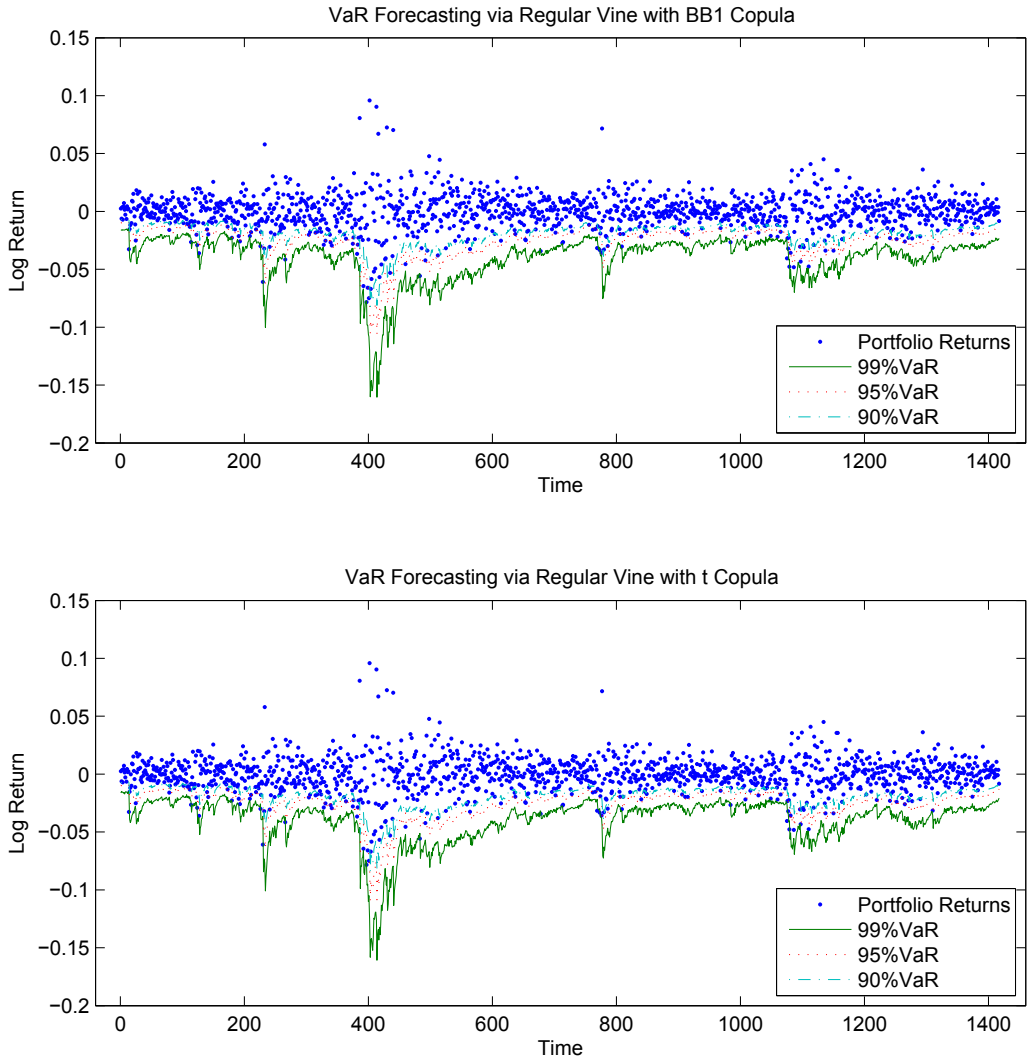

Figure 4.9: The VaR forecasting of portfolio returns

Chapter 5

High-dimensional Dependence Modelling through Truncated Partial Regular Vine Copula

5.1 Introduction

Learning about high-dimensional dependence, that is, the dependence between a set of high-dimensional variables, is a critical but challenging issue in many applications including social media analysis and financial markets. A particular problem is understanding the dependence between high-dimensional variables with fat tail and asymmetric characteristics, which appear widely in such areas as financial markets. In the currency market, for example, the asymmetric responses of central banks to currency exchange rate movements often result in asymmetric dependence on the currencies as well. For instance, if the Japanese government wants to maintain export competitiveness with the United States (US) when British products are also exported to the US, the central bank of Japan will ensure the matching depreciation of the Japanese Yen against the US dollar whenever the British Pound is depreciated against the US dollar. Similar situations increasingly occur in recognized big data applications, in which high dimensional variables

are dependent on one another..

While there is an increasingly recognized need for effective techniques for modelling high-dimensional dependence, it is very challenging to properly model the complex dependencies between large scale variables. The challenges lie in the fact that the dependencies between high-dimensional variables interweave with different forms and complexities, such as the volatility clustering and asymmetry phenomena in the currency market, which are difficult to capture adequately. In addition, high-dimensional dependence modelling with limited computational resources can be susceptible to the curse of dimensionality.

Probabilistic graphical models (Pearl 1988), such as Bayesian logic program (De Raedt & Kersting 2008), relational dependency networks (Neville & Jensen 2007) and relational Markov networks (Getoor & Taskar 2007), build a graph to represent the conditional dependence structure between random variables. For example, the latent factor models with a dependency structure in the latent space are studied in (He et al. 2012). A set of probabilistic dependencies are learned in (Gao & Suzuki 2003) to identify the relationships between the headwords of each phrase. These models are aimed at high-dimensional domains, and have the advantage of learning latent relationships from data. However, as discussed in (Elidan 2013), they tend to force the local quantitative part of the model to take a simple form such as the discretized form of the data when multivariate Gaussian or its mixtures cannot capture the data in the real world. The complex dependence between high-dimensional variables is hard to be capture.

In statistics and finance, the copula has been shown to be a powerful tool for modelling high-dimensional dependencies. The copula splits the multivariate marginal distributions from dependence structures, so that the specification of the dependence structures can be investigated independently of the marginal distributions. It can provide a flexible mechanism for modelling real world distributions that cannot be handled well by graphical models. Thus, researchers have tried to combine copula and probability

graphical models. The tree-structured copula (TSC) model (Kirshner 2007) was first presented in the combination of a copula and a graphical model, following which the sparse undirected copula-based (SUC) model (Liu, Lafferty & Wasserman 2009) and the copula Bayesian networks (CBN) (Elidan 2010) were developed. These copula-based models aim to resolve the limitations of discretizing data, but they impose assumptions and restrictions on the dependence structure. For example, TSC imposes an independence assumption on the tree structure, and SUC does not consider general (non-Gaussian) cases. CBN has a conditional independence assumption in the dependence structure. These assumptions and restrictions are not appropriate for dependence modelling among financial variables.

Another typical copula-based framework for dependence modelling is the vine copula model (Bedford & Cooke 2002). The vine copula model builds on successive conditioning, and uses bivariate copulas as building blocks to construct multivariate distributions. Compared with CBN, the most important difference is that vine models use conditional dependence to replace the conditional independence assumed in CBN. Thus, vine copula models provide a flexible mechanism for modelling real world distributions that cannot be handled well by probability graphical models. Recently, most studies of vine copula models have focused on the canonical vine and D vine models since they make an assumption on the dependence structure for simplifying the dependence construction, such as (Aas et al. 2009). The canonical vine assumes that one variable connects to all other variables, which results in a star-like structure, and D vine assumes that one variable can only connect to no more than 2 variables, which leads to a path-like structure. The dependence structures satisfying their assumptions may not be consistent with the actual dependence in real world scenarios. However, the generalized vine model, which is called regular vine, is very difficult to construct since it does not make any assumption on the dependence structure. In (Dissmann, Brechmann, Czado & Kurowicka 2013), a sequential method is proposed to build a regular vine, but its subsequent structures and bivariate copula se-

lections are heavily dependent on established structures and selections. In addition, the computational burden for high-dimensional data is huge. In (Brechmann, Czado & Aas 2012), a statistical test-based truncation method is presented for regular vine. It heavily depends on the result of the statistical test, which may lead to a scenario in which the model cannot completely reflect the inherent dependence.

It is thus a challenging job to build a proper multivariate model with the following desired properties: (i) A flexible dependence structure that does not impose assumptions or restrictions; (ii) An appropriate truncation method to ensure the model can be effectively applied to high-dimensional data; (iii) A wide range of tail dependencies, allowing for a variety of lower and upper tail dependencies, and; (iv) An estimation of joint density functions that is computationally feasible. The existing multivariate copula models or vine models, including canonical vine and D vine, do not satisfy all of these conditions. The Archimedean copula-based multivariate models have a structure that has only a narrow range of negative dependence (McNeil & Nešlehová 2009). The multivariate Gaussian copula model is not suitable for modelling the asymmetric characteristics, because (i) the Gaussian copula does not have lower and upper tail dependence, which is a measure of dependence between extreme events (Nelsen 1999), and (ii) the Gaussian assumption is not appropriate in the real world (Fang et al. 2002). The multivariate t copula model, which was studied by (Nikoloulopoulos et al. 2009), does not have flexible lower and upper tail dependence, since t copula has only symmetric tail dependence. In addition, multivariate elliptical copulas and multivariate Archimedean copulas are difficult to expand to high-dimensional data since their desinence structure is fixed.

Building on the power of the copula, this work proposes a new truncated partial correlation-based model called *truncated partial regular vine* ($TPRV$), to address the above issues. Our models capture actual asymmetric dependence in high-dimensional data by incorporating regular vine theory into the construction of the dependence structures, which does not

impose assumptions on the dependence structures. In addition, our models are embedded with the copula family which supports a wide range of flexible lower and upper tail dependence. The proposed truncated partial correlation-based regular vine has the following advantages: (i) The $TPRV$ model can uniquely determine the correlation matrix and be algebraically independent. This indicates that a dependence structure constructed via partial correlation is more flexible, since the current tree structure is independent of the established tree structure and bivariate copula selection. In addition, by employing regular vine, our models do not impose assumptions or restrictions on the complex dependence structures. (ii) The $TPRV$ model is constructed by a new bottom-up construction strategy, which ensures that the weakest correlations are at the bottom, and the strongest correlations appear at the top of dependence structures. (iii) There are $n(n-1)/2$ parameters for n variable regular vines. The $TPRV$ model employs a new truncation method, which effectively decreases the number of parameters and reduces the computation burden without losing the most important information inherent in the dependence. (iv) The $TPRV$ model selects bivariate copulas from a large number of mixed copula families rather than only using the elliptical copula family, which ensures a wide range of lower and upper tail dependencies.

The rest of this chapter is structured as follows. Section 5.2 presents the truncated partial regular vine model in detail, including truncated vine tree structure building, bivariate copula selection, marginal distribution specification and parameter estimation. Evaluation methods are discussed in Section 5.3 for verifying both the in-sample and out-of-sample performance of high-dimensional financial variables. Section 5.4 shows the case study results in currency markets. Section 5.5 concludes this chapter.

5.2 Truncated Partial Regular Vine

5.2.1 Partial Regular Vine Specification

The reason of using partial correlation to construct the regular vine is discussed firstly, and the construction of regular vine tree structure, selecting bivariate copulas and estimate the parameters are discussed in following. There are four main reasons why using partial correlation: (1) According to Definition 2.13, the partial correlation can be obtained directly from the data set, without knowledge of assumed structure, bivariate copulas or corresponding copula parameters; (2) A bottom-up strategy can be employed to build the regular vine by using partial correlation, which ensures that the weakest correlation is always at bottom; (3) In regular vine tree dependence structure, the structure of the current tree (excluding the first tree) does not depend on the structure of previous trees and the corresponding bivariate copulas and parameters, and (4) The partial correlation-based regular vine can be easily truncated, which ensures that it is applicable to high-dimensional data. Hence, the partial correlation-based regular vine is much more flexible than a typical regular vine. For simplicity, the partial correlation-based regular vine is called *partial regular vine* (PCV).

Let $V^R_{\rho_e}$ be denoted as partial correlation based-regular vine dependence structure, and then a complete partial correlation based-regular vine is defined below.

Definition 5.1 *(Partial Regular Vine Specification). A complete partial correlation based-regular vine specification is a regular vine with a partial correlation ρ_e specified for each edge e. A distribution satisfies the complete partial correlation specification, if and only if for any edge $e = \{a, b\}$ in the vine, the partial correlation of the variables in $C_{e,a}$ and $C_{e,b}$ given the variables in D_e, is equal to ρ_e.*

Thus, a Partial Canonical Vine Copula Specification $(V^R_{\rho_e}, B(V^R_{\rho_e}),$ $\theta(B(V^R_{\rho_e})))$ *on n variables is a multivariate distribution function:*

(i). $V^R_{\rho_e}$ *is a partial correlation-based regular vine structure on* n *variables;*

(ii). $B(V^R_{\rho_e}) = \{C_{e(a),e(b)\,;\,D_e} \,|\, e \in E_i, i = 1, ..., n-1\}$ *is the set of* $n(n-1)/2$ *bivariate copulas; and*

(iii). $\theta(B(V^R_{\rho_e})) = \{\theta_{e(a),e(b)\,;\,D_e} \,|\, e \in E_i, i = 1, ..., n-1\}$ *is the set of parameters, corresponding to the bivariate copulas in* $B(V^R_{\rho_e})$.

According to the above definition, the edge e in a partial regular vine can be written as $\{C_e\,;\,D_e\}$ or $\{C_{e(a)}, C_{e(b)}\,;\,D_e, e = \{a, b\}\}$, where C_e, D_e are conditioned and conditioning sets respectively. For simplicity, the edge e will be written as $\{e(a), e(b)\,;\,D_e, e = \{a, b\}\}$ in this work. A complete partial regular vine copula consists of three components: the partial regular vine tree structure V_{ρ_e}, the bivariate copula set $B(V^R_{\rho_e})$, and the corresponding bivariate copula parameters $\theta(B(V^R_{\rho_e}))$. In the following sections, the three components are discussed respectively.

5.2.2 Partial Regular Vine Tree Structure Construction

The regular vine on n variables shares several important properties (see details in (Kurowicka & Cooke 2006*b*)):

(1). There are $(n-1)$ trees and $\binom{n}{2}$ edges in total;

(2). Each pair appears once as a conditioned set of an edge;

(3). There are $(j-1)$ and $(j+1)$ variables in the conditioning sets and constraint sets of an edge of the jth tree respectively;

(4). If two or more nodes have the same constraint sets, they are the same node;

(5). If variable i is a member of the conditioned set of an edge e in a regular vine, then i is a member of the conditioned set of exactly one of the m-child of e, and the conditioning set of an m-child is a subset of D_e.

According to the above properties, two lemmas, which are important for constructing the partial regular vine tree structure, are derived. The two lemmas are given as follows:

Lemma 5.1 *Let* $I \in \{1, ..., n\}$, $x_1, x_2, y_1, y_2 \in I$ *and* $x_1 \neq x_2$, *the nodes of* T_j *be* $N_1 = \{x_1, y_1 \,;\, I \backslash \{x_1, x_2, y_1\}\}$ *and* $N_2 = \{x_2, y_2 \,;\, I \backslash \{x_1, x_2, y_2\}\}$. *For a regular vine on* n *variables, nodes* N_1 *and* N_2 *have a common m-child. If* $y_1 \neq y_2$, *the common m-child is* $\{y_1, y_2 \,;\, I \backslash \{x_1, x_2, y_1, y_2\}\}$.

Proof 5.2 *According to Definition 2.12, each node has two m-children. For* N_1, *the constraint set* CV_{x_1} *of its m-children are* $\{x_1, I \backslash \{x_1, x_2, y_1\}\}$ *and* $\{y_1, I \backslash \{x_1, x_2, y_1\}\}$. *For* N_2, *the constraint set* CV_{x_2} *of its m-children are* $\{x_2, I \backslash \{x_1, x_2, y_2\}\}$ *and* $\{y_2, I \backslash \{x_1, x_2, y_2\}\}$. $\{y_1, I \backslash \{x_1, x_2, y_1\}\}$ *and* $\{y_2, I \backslash \{x_1, x_2, y_2\}\}$ *are equal, but indexed by different variables in a conditioned set. According to Property (4),* N_1 *and* N_2 *have a common m-child. If* $y_1 \neq y_2$, y_1 *and* y_2 *should be in the conditioned set of the m-child.*

Example 5.3 *Suppose a regular vine has two nodes,* $N_1 = \{\{1, 3\} \,;\, \{4, 5, 6\}\}$ *and* $N_2 = \{\{2, 4\} \,;\, \{3, 5, 6\}\}$, *if the variable* 3 *is not equal to the variable* 4, N_1 *and* N_2 *have a common m-child, which is* $\{\{3, 4\} \,;\, \{5, 6\}\}$.

Lemma 5.4 *For a regular vine on* n *variables,* $j = 2, ..., n-1$, *the edge* e *in* T_j *has only two constraint sets of m-children in* T_{j-1}, *which are indexed by different variables in a conditioned set.*

Proof 5.5 *Suppose there are three identical constraint sets indexed by different variables in a conditioned set, according to Property (4), nodes with the same constraint sets should be the same node. Based on Property (5), the variables in the conditioned set will still be in the conditioned set of its m-children. This means that the node will have three variables in its conditioned set, which violates Property (3) and the proximity condition in the regular vine definition. Therefore, one edge has only two constraint sets which are indexed by different variables in a conditioned set.*

Example 5.6 *Suppose one edge in tree* T_j *is* $\{\{1,2\};\{3,4\}\}$. *In* T_{j-1}, *the constraint sets of its m-children consists of* $\{1,\{3,4\}\}$ *and* $\{2,\{3,4\}\}$. *Only two identical constraint sets are indexed by different variables in the conditioned set, not the three identical constraint sets.*

According to the above properties of regular vine and the two lemmas, the partial regular vine is built by using Algorithm 5.1. Then, the following theorem can be proved.

Theorem 5.7 *Algorithm 5.1 produces a partial regular vine.*

Proof 5.8 *According to Lemma 5.1, Steps 1 and 2 are evidently true since the two edges in* T_{n-2}, $\{x, pt(x)\,;\, I\backslash\{x,y,pt(x)\}$ *and* $\{y, pt(y)\,;\, I\backslash\{x,y,pt(y)\}$ *have a common m-child* $\{I\backslash\{x,y\}\}$ *in* T_{n-3}. *Lemma 1 ensures that the two edges have a common m-child. Step 5 produces* $2(n-j)$ *constraint sets of m-children of edges in* T_j. *These constraint sets are indexed by variables in the conditioned sets. Step 6 ensures that the duplicated constraint sets are removed. At this stage, some constraint sets are equal but are indexed by different variables in the conditioned sets. Step 7 employs Lemma 5.1 to ensure that the same constraint sets are in one edge. After Step 7, all the different constraint sets are obtained. For these different constraint sets, the Lemma 5.4 is applied to ensure that the selection of partners will satisfy the definition of regular vine. For example, for the constraint set* $CV = \{1,\{2,3,4\}\}$ *in* T_k, *the indexed variable in the conditioned set is* $\{1\}$. *All possible edges in* T_{k-1} *are* $\{\{1,2\};\{3,4\}\}$, $\{\{1,3\};\{2,4\}\}$ *or* $\{\{1,4\};\{2,3\}\}$. *The constraint sets of m-children of all possible edges in* T_{k-2} *which do not include variable* $\{1\}$ *are* $\{2,\{3,4\}\}$, $\{3,\{2,4\}\}$ *and* $\{4,\{2,3\}\}$. *These constraint sets are the same because they have the same variables. However, they are indexed by different variables in conditioned sets. If there exist* $\{2,\{3,4\}\}$ *and* $\{3,\{2,4\}\}$ *which are from other edges in* T_{k-2}, $\{\{1,4\};\{2,3\}\}$ *are ;not selected due to edges for which the Lemma 5.4. The reason is that the edge* $\{\{1,4\};\{2,3\}\}$ *in* T_{k-1} *consists of constraint set* $\{4,\{2,3\}\}$. *If* $\{\{1,4\};\{2,3\}\}$ *is chosen, there will be three identical constraint sets, which are indexed by different*

Algorithm 5.1 Tree Structure Construction via Bottom-Top Strategy

Require: Observations of n variables

1: Calculate all partial correlation values. Put the smallest absolute value of partial correlation to the edge in T_{n-1}. Let $I = \{1, ..., n\}$ and $\{x, y\} \in I$. Then, in T_{n-1}, $E_{n-1} = \{x, y\,;\, I\backslash\{x, y\}\}$. The constraint set for its m-children are $CV_x = \{x, I\backslash\{x, y\}\}$ and $CV_y = \{y, I\backslash\{x, y\}\}$.

2: Find the partners of x and y to minimise the function $\sum|\rho_{e(x),e(pt(x))\,;\,D(e)}|$ and $\sum|\rho_{e(y),e(pt(y))\,;\,D(e)}|$ where $pt(x)$ and $pt(y)$ are partners of x and y respectively. Then, in T_{n-2}, $E_{n-2} = \{\{x, pt(x)\,;\, I\backslash\{x, y, pt(x)\}\}, \{y, pt(y)\,;\, I\backslash\{x, y, pt(y)\}\}\}$.

3: **for** $k = 1, ..., n-2$ **do**

4: **for** $j = 1, ..., n-3$ **do**

5: Obtain all constraint sets of m-children in T_j;

6: Remove constraint sets for which $CV_p = CV_q$, $p \neq q$;

7: Select the partners of variables in conditioned sets by using Lemma 5.1;

8: Ensure they satisfy Lemma 5.4;

9: **if** $T_j > T_k$ **then**

10: for left constraint sets, select their partners of variables in conditioned sets which minimise the function $\sum|\rho_{e(p),e(q)\,;\,D_e}|$;

11: **else**

12: for left constraint sets, select their partners of variables in conditioned sets which minimise the function $\sum log(1-\rho^2_{e(p),e(q)\,;\,D_e})$;

13: **end if**

14: **end for**

15: **end for**

16: This produces a total of $n-3$ partial regular vine tree structures. The 'best' one is the vine tree structure which maximises the function $-log(Demt)$, where $Demt$ is a determinant of vine (Kurowicka & Cooke 2006*b*),

$$Demt = \prod_{e=\{a,b\}\in E(V_{\rho_e})} (1-\rho^2_{e(a),e(b)\,;\,D_e}); \tag{5.1}$$

17: **return** The partial regular vine tree structure

variables. This can result in the violation of the definition and its property of regular vine. Lemma 5.4 can ensure avoidance of the circumstance and maintain the regularity of regular vine.

To obtain the 'best' partial regular vine tree structure, it is essential to catch as many correlations as possible while building the vine tree structure. In Algorithm 5.1, a broken tree method is used to obtain the 'best' vine tree structure. If $T_j > T_k$, the selection of partners for variables in the conditioned set has to minimise the function $\sum |\rho_{e(p),e(q)\,;\,D_e}|$, which guarantees that the weak partial correlations are assigned to the edges in T_j. If $T_j < T_k$, the selection of partners for variables in the conditioned set has to minimise the function $\sum log(1-\rho^2_{e(p),e(q)\,;\,D_e})$, which ensures that the strong partial correlations are in the edges of T_j. In this way, $(n-3)$ partial regular vine tree structures are obtained. The 'best' one is obtained through calculating the value of $-log(Demt)$.

5.2.3 Partial Regular Vine Tree Structure Truncation

The number of parameters increases exponentially as the dimension increases. For a n-variable partial regular vine with bivariate t copula (two-parametric copula), the number of parameters is $n(n-1)$. This may result in a huge computational burden and be time consuming. Hence, it is necessary to reduce the number of parameters by truncating the partial regular vine. The conditional independence copula, which is discussed in (Aas et al. 2009), is equal to 1, such as

$$C^{independence}_{e(a),e(b)\,|\,D_e} = 1$$

The partial regular vine tree structure can be truncated by using conditional bivariate independence copula to replace those edges that have a low absolute value of partial correlation. This ensures to retain the important dependencies indicated by strong correlations, truncating the useless dependencies indicated by weak correlations.

In the truncation procedure, those edges for which the absolute value of

partial correlations are less than a specified truncation value ρ_{trun} between 0 and 1 (e.g., 0.1), are replaced with conditionally independent copula. Then, based on the specified truncation value ρ_{trun}, the partial regular can be truncated. The process of truncating the partial regular vine tree structure is given in Algorithm 5.2.

Algorithm 5.2 Partial Regular Vine Tree Structure Truncation

Require: The n-variable partial canonical vine tree structure $V^R_{\rho_e}$ and truncation value ρ_{trun}

1: Calculate the corresponding partial correlations based on the partial regular vine tree structure $V^R_{\rho_e}$;
2: **for** $j = 1, ..., n-1$ **do**
3: In tree T_j, find all edges for which the absolute value of partial correlation is less than ρ_{trun}, e.g. $\rho_{e(p),e(q)\,;\,D(e)} < \rho_{trun}, e = \{p, q\}$;
4: For those edges found in Step 4, replace all bivariate copulas with conditional independence copulas;
5: **end for**
6: **return** The truncated partial regular vine tree structure

5.2.4 Bivariate Copula Selection

Once the partial regular vine tree structure V_{ρ_e} is identified, the next step is to select bivariate copulas for each edge in all trees. As discussed in Section 2.6.2, the partial correlation is equal to its corresponding conditional correlation for the elliptical family. This means that our partial regular vine tree structure is built based on an elliptical copula family (i.e., Gaussian or t copulas). However, according to the following theorem, the limitation of partial correlation can be removed by mapping the partial regular vine tree structure to typical regular vine via conditional correlation.

Theorem 5.9 *For any regular vine on n variables, there is one-to-one correspondence between the set of $n \times n$ positive definite correlation matrices and the set of partial correlation specification of the vine.*

Table 5.1: The Tail Dependence of Bivariate Copulas

	Lower Tail Dependence	Upper Tail Dependence
Gaussian	–	–
t	✓	✓
Gumbel	–	✓
Frank	–	–
Clayton	✓	–
Joe	–	✓
S.Gumbel	✓	–
S.Clayton	–	✓

The check mark (✓) indicates the the bivariate has lower/upper tail dependence, and the dash mark (−) means there is no corresponding tail dependence. The bivariate t copula has symmetric lower and upper tail dependence, which indicates that the lower tail dependence is equal to the upper tail dependence.

The proof of Theorem 5.9 can be referred to (Bedford & Cooke 2002), which is omitted here. It shows that there is a one-to-one relationship between the partial regular vine specification and the correlation matrix, which ensures that it can map the partial regular vine tree structure to the typical conditional correlation based-regular vine tree structure. Then the bivariate copulas can be selected from a large number of copula family candidates, rather than the elliptical copula family. Hence, the limitation can be removed while selecting the bivariate copulas.

For a given partial regular vine tree structure V_{ρ_e} and a set of bivariate copula candidates, the AIC-based selection criterion is used to select an appropriate bivariate copula with the highest p-value of the copula goodness of fit test on Cramer-von Mises statistics (Brechmann et al. 2012). The tail dependence is a key feature for identifying various copula families. Table 5.1

provides the tail dependence for popular bivariate copulas (see (Nelsen 1999) for details). In this work, the following bivariate copula families are used in the experiment, including Gaussian, t, Clayton, Gumbel, Frank and Joe. The survival bivariate copula [1] (including Survival Clayton and Gumbel copulas) are also considered in this work, which covers a wide spectrum of dependence. In these bivariate copulas, only the t copula is two-parametric, the others are one-parametric.

5.2.5 Marginal Distribution Specification and Parameter Estimation

As discussed in Section 2.2.2, a copula has two parts: the copula function and the marginal distributions. Here the financial variables are used as a real world example. For the finance use of the partial regular vine copula model, the volatility models (i.e. $ARMA$-$GARCH$ models) is used as the margins. Typically, let $X_t(t = 0, 1, ..., \mathfrak{T})$ be a time series of the prices of a financial asset, such as the stock market index. The return of financial asset can be defined as $r_t = log(X_t/X_{t-1})$. If there are n assets with returns $r_{t,1}, ..., r_{t,n}$, The appropriate marginal distribution of individual variables (i.e., returns of financial variables) is chosen firstly. Due to the character of financial variables, such as volatility cluster, a common choice is ARMA(1,1)-GARCH(1,1) with Student-t error distribution, which is defined as follows:

$$\begin{aligned} r_{t,j} &= \eta_j + \Psi_j r_{t-1,j} + \Theta_j \varepsilon_{t-1,j} + \varepsilon_{t,j}, \\ \sigma^2_{t,j} &= \omega_j + \alpha_j \varepsilon^2_{t-1,j} + \beta_j \sigma^2_{t-1,j} \\ \varepsilon_{t,j} &= \sigma_{t,j} \cdot Z_{t,j} \end{aligned} \tag{5.2}$$

where $j = 1, ..., n, t = 1, ..., \mathfrak{T}$, and $Z_{t,j}$ is the error which follows the Student-t distribution. Let $\theta^m_j = (\eta_j, \Psi_j, \Theta_j, \omega_j, \alpha_j, \beta_j)$ and θ^c be denoted as the

[1] The survival copula is the copula with survival functions, i.e., $\overline{C}(x_1, x_2) = x_1 + x_2 - 1 + C(1 - x_1, 1 - x_2)$, where $\overline{C}$ is the survival copula. See details in (Nelsen 1999).

parameter set of margins and the parameters of copula function respectively. The multivariate joint log-likelihood is given by:

$$\begin{aligned} L(\theta_1^m, ..., \theta_n^m, \theta^c) &= \sum_{t=1}^{\mathfrak{T}} logf(r_{t,1}, ..., r_{t,n}; \theta_1^m, ..., \theta_n^m, \theta^c) \\ &= \sum_{t=1}^{\mathfrak{T}} \log c(F_1(r_{t,1}), ..., F_n(r_{t,n}); \theta^c) \\ &\quad + \sum_{t=1}^{\mathfrak{T}} \sum_{j=1}^{n} \log f_j(r_{t,j}; \theta_j^m) \end{aligned} \tag{5.3}$$

where the multivariate $c(\cdot\,; \theta^c)$ is denoted as the partial regular vine. It is possible to maximise Equation (5.3) by using the Maximum Log-likelihood Estimate (MLE). However, it is time consuming when n is large. $Inference\ Functions\ for\ Margins$ (IFM) method (detail can be found in Section 2.2.2 and (Joe 2005)) is then used to resolve the issue. IFM is a two-step estimation method, which efficiently estimates the parameters of a model. In the first step, the margins $ARMA(1,1)$-$GARCH(1,1)$ with Student-t error distribution are employed to filter the financial returns and the univariate parameters $\theta_j^m = (\eta_j, \Psi_j, \Theta_j, \omega_j, \alpha_j, \beta_j)$ are derived. In the second step, the joint log-likelihood in Equation (5.3) is maximised over copula parameters θ^c given the fixed parameters of margins θ_j^m. This means that the joint log-likelihood is reduced to the equation which consists of only copula parameters because the parameters of log-likelihood are fixed.

5.3 Evaluation Methods

In this chapter, the high-dimensional dependence of financial variables are used as an example to illustrate the above proposed truncated partial regular vine copula. In this section, different evaluation methods are discussed, which consist of in-sample evaluation and out-of-sample evaluation. Two frameworks are used to evaluate the performance of the proposed models.

One is to measure in-sample performance by examining whether a model fits the data well. The other is to measure the out-of-sample performance to evaluate the prediction performance of the model.

5.3.1 In-Sample Evaluation Methods

Classic model evaluation criteria such as Akaike's information criterion (AIC), the Bayesian information criterion of Schwarz (BIC), and the consistent AIC of Bozdogan (CAIC) are used in the in-sample performance evaluation.

The model comparison test are also taken, such as the Vuong test (Vuong 1989). To compare models R_1 and R_2 with their estimated parameters θ_1 and θ_2, the standardised sum υ of the log difference of their point-wise likelihoods $m_j := log[\frac{R_1(x_j|\theta_1)}{R_2(x_j|\theta_2)}]$ for observations $\mathbf{x} = (x_1, ..., x_N)' \in \mathbb{R}^{N\times n}$ is calculated. Under fairly general regularity conditions, υ is shown to be asymptotically standard normal, resulting in the following test – the Vuong test statistics (υ) at level α are defined as:

$$statistics := \upsilon = \frac{(1/n)\sum_{j=1}^{n} m_j}{\sqrt{\sum_{j=1}^{n}(m_j - \bar{m})^2}} \tag{5.4}$$

Here, Φ^{-1} is denoted as the inverse of the standard normal distribution function. If $\upsilon > \Phi^{-1}(1-\alpha/2)$, it prefers model R_1 to model R_2 at the level α. If $\upsilon < -\Phi^{-1}(1-\alpha/2)$, the model R_2 is chosen over model R_1. If, however, $|\upsilon| \leq \Phi^{-1}(1-\alpha/2)$, no decision can be made between the two models. The Vuong test does not take the effect of parameters into account. Hence, the Vuong test with Akaike and Schwarz corrections, which correspond to the penalty terms in the AIC and BIC respectively, are used.

5.3.2 Out-of-Sample Evaluation Method

To evaluate the out-of-sample performance, Value at Risk (VaR) is used, since it is a probabilistic metric measuring market risk and is an industrial golden benchmark for measuring market risk. VaR at the level $(1-\alpha)$ is

defined by

$$VaR_t(1-\alpha) = -inf\{c \in \mathbb{R} : P(r_t \leq c|\digamma_{t-1})| \geq (1-\alpha)\} \tag{5.5}$$

where $\digamma_{t-1}$ represents the past information at time $t-1$. A good model is expected to produce a high quantity of VaR. Given a set of financial returns, the portfolio returns can be defined as:

$$r_{portfolio,t} = \sum_{j=1}^{n} \mu_j r_{j,t} \tag{5.6}$$

where $r_{j,t}$ is financial return at time t for $i = 1, ..., n$, and μ_j is the weight, where $\sum_{j=1}^{n} \mu_j = 1$.

Suppose the current time is t, the process for computing VaR forecasting at time $t+1$ is given as follows:

(i). Fit $ARMA(1,1)$-$GARCH(1,1)$ with the Student-t error distribution with returns by using Equation (5.2). The standardised residuals are obtained by:

$$\hat{Z}_{j,t} = \frac{r_{j,t} - \hat{\eta}_j - \hat{\Psi}_j r_{j,t-1} - \hat{\Theta}_j \hat{\sigma}_{j,t-1} \hat{e}_{j,t-1}}{\hat{\sigma}_{j,t}} \tag{5.7}$$

(ii). The ex-ante GARCH variance forecast (see detail in (Andersen 2009)) for $j = 1, ..., n$ can be computed as follows:

$$\hat{\sigma}^2_{j,t+1} = \hat{\omega}_j + \hat{\alpha}_j \hat{\varepsilon}^2_{j,t} + \hat{\beta}_j \hat{\sigma}^2_{j,t} \tag{5.8}$$

(iii). The standardised residuals obtained from $ARMA(1,1)$-$GARCH(1,1)$ are transformed to approximately uniform data $\mathbf{u}_j = u_{1,j}, ..., u_{t,j}$ by using the Student-t cumulative distribution function;

(iv). Fit the partial regular vine tree structure with approximately uniform data $\mathbf{u}_j$ and estimate copula parameters;

(v). Use the fitted regular vine structure with estimated copula parameters to simulate a sample for each financial return variable, i.e.,$v_{t+1,j}$;

(vi). Transfer the sample to standard residuals by using the inverse Student-t cumulative probability distribution functions with parameters obtained in Step (i), and obtain the simulated standardised residuals, i.e., $\hat{Z}_{j,t+1}$;

(vii). Calculate the one day ahead forecasting return and variance for each financial variable by using the estimated $ARMA(1,1)$-$GARCH(1,1)$ which is calculated in Step (i), i.e.,

$$\hat{r}_{j,t+1} = \hat{\eta}_j + \hat{\Psi}_j r_{j,t} + \hat{\Theta}_j \hat{\varepsilon}_{j,t} + \hat{\varepsilon}_{j,t+1} \tag{5.9}$$

(viii). The portfolio return is calculated by Equation (5.6). Steps from (iv) to (vii) are repeated for N times (e.g. $N = 10000$). The 99%, 95%, and 90% VaR forecasts are determined by taking the corresponding 1%, 5% and 10% quantiles of the portfolio return forecast respectively.

The quality of VaR forecasting can be judged by backtesting. Typically, backtesting methods consist of unconditional, independent and conditional coverage tests (Guermat & Harris 2002). In addition, another test for the quality of VaR forecasting is ex-post exceedance (Andersen 2009), which is defined at time t as:

$$I_t = \begin{cases} 1, & \text{if} \quad r_{portfolio,t} < VaR_t(1-\alpha); \\ 0, & \text{otherwise.} \end{cases} \tag{5.10}$$

where $r_{portfolio,t}$ is the ex-post observed portfolio return at time t. If the VaR forecasting is accurate, I_t should be less than or equal to the significance level α of the backtesting.

5.4 Experimental Results

5.4.1 Data and Marginal Distributions Specification

The real-world financial data sets are used to evaluate the in-sample and out-of-sample performance. They are the returns of 17 currency exchange

Table 5.2: Ljung-Box Test Results

j	BRL	CAD	CNY	KRW	HKD	INR
Z_j	0.23	0.51	0.91	0.95	0.78	0.09
Z_j^2	0.10	0.15	0.97	0.38	0.97	0.16
j	JPY	MXN	NOK	CHF	SGD	SEK
Z_j	0.83	0.30	0.94	0.95	0.82	0.94
Z_j^2	0.35	0.37	0.35	0.41	0.73	0.46
j	TWD	AUD	EUR	NZD	GBP	
Z_j	0.06	0.61	0.97	0.89	0.66	
Z_j^2	0.96	0.67	0.42	0.27	0.43	

rates against USD. These trading currencies are EUR, GBP, NOK, SEK, CHF, CAD, KRW, HKD, JPY, INR, MXN, AUD, NZD, CNY, BRL, SGD and TWD. They represent major currencies in the global market and can be arranged into portfolios. For the in-sample performance test, these observations from 01/08/1999 to 12/01/2007, a total of 418 weekly returns, are used for evaluation. Observations from 19/09/2005 to 23/06/2013, a total 310 weekly returns are used for out-of-sample testing. All the data was downloaded from FRB St. Louis (http://research.stlouisfed.org).

The raw returns are fitted with univariate ARMA(1,1)-GARCH(1,1) models with the Student-t error distribution. The Ljung-Box (LB) tests (Ljung & Box 1978) are used to remove the autocorrelation among these financial returns. The results of LB tests are shown in Table 5.2, which give the corresponding p-values for all returns. It suggests that there are no autocorrelation left for all returns in the standardise residuals Z_j and squared standardised residuals Z_j^2 (all p-values > 0.05). The standardised residuals are transferred to uniform data by using the empirical probability integral transformation, which is actually the input of partial regular vine.

5.4.2 Partial Regular Vine Truncation Analysis

The following vine copula models are used in the experiments which cover a wide spectrum of vines to evaluate the proposed copula model:

- $TPRV_{0.1}$: Truncated partial regular vine, built by Algorithm 5.1, and then truncated by Algorithm 5.2 with truncation value 0.1;
- PRV: Partial regular vine, built by Algorithm 5.1, non-truncated;
- PCV: Partial Canonical vine, built by Algorithm 3.1;
- $RSeq$: Regular vine, built by (Dissmann et al. 2013);
- $CSeq$: Canonical vine, built by (Dissmann et al. 2013).

Considering the tail dependence of bivariate copula families, these bivariate copula families are considered for the five models, including t, Clayton, Gumbel, Frank, Joe, S.Clayton and S.Gumbel, where S.Clayton and S.Gumbel are the survival version of Clayton and Gumbel copula respectively.

In general, the selection of ρ_{trun} is determined by the characteristics of data and domain knowledge. Table 5.3 shows the result of the total number of edges against various truncation values (ρ_{trun}) in the training data set. If $\rho_{trun} = 0$, it refers to the non-truncated vine, since no partial correlation is less than 0. The total number of edges decreases when the truncation value increases. To investigate the performance of the truncated partial regular vine, an experiment is implemented to compare the performance between truncated and non-truncated partial regular vines. The truncated partial regular vines are used, with truncation value ρ_{trun} from 0.01 to 0.4, separated by 0.01. The training data set described in Section 5.4.1 is used to examine the performance of non-truncated and truncated vines. The results are shown in Figure 5.1. The top two figures show the total parameters and estimation

Table 5.3: The Total Number of Edges in Truncated Partial Regular Vine against Truncation Value

ρ_{trun}	0	0.01	0.02	0.03	0.04	0.05	0.06
Total No.	136	121	111	101	87	80	74
ρ_{trun}	0.07	0.08	0.09	0.1	0.15	0.2	0.3
Total No.	70	66	63	57	36	26	17

time [2] of truncated vine against non-truncated vine, and the bottom two figures give the log-likelihood value and determinant of partial regular vine. The x-axis of all four figures is ρ_{trun}. The dashed line indicates the non-truncated vine as reference. The number of parameters and the estimation time of the truncated vine decreases significantly as ρ_{trun} increases. However, the log-likelihood and determinant (discussed in Algorithm 5.1) do not show a great reduction. It can be concluded that the truncated vine efficiently reduces the number of parameters and estimation time, and at the same time retains the important dependencies, since the log-likelihood and determinant decrease smoothly and slowly.

Considering the characteristics of the currency data (time series) and the result shown in Table 5.3 and Figure 5.1, the truncated partial regular vine $\rho_{trun} = 0.1$ is chosen in our case study. The $TPRV_{0.1}$ in the following sections'indicates a truncated partial regular vine whose truncation value is 0.1. Table 5.4 provides a detailed comparison between $TPRV_{0.1}$ and non-truncated PRV models. The number of parameters and the estimation time of $TPRV_{0.1}$ are only half of PRV. However, the difference of log-likelihood and determinant is not very significant, which indicates that $TPRV_{0.1}$ retains important dependencies but truncates the unnecessary information.

The comparison between partial regular vine model and the model with copula-based probability graphical models (copula Bayesian networks) in our

[2]The estimation time refers to only the parameters of partial regular vine, excluding the parameters of margins. The computation is on a Linux cluster computer with 6 Inter Xeon CPU (3.74GHz) and 12G memory.

Table 5.4: The Comparison between $TPRV_{0.1}$ and PRV

	$TPRV_{0.1}$	PRV
Total Parameters	185	79
Estimation Time [1]	5979	2357
Log-likelihood	6013	5913
Determinant	7.796	7.624

[1] The estimation time is seconds.

case study, because those models are lack of efficient sampling methods. Another reason is the difficulty of obtaining the one day ahead forecasting returns for computing VaR forecasting.

5.4.3 In-Sample Performance Analysis

Table 5.5 presents the results with the bivariate copula selection and model selection criteria. It shows that PRV takes the highest log-likelihood, followed by $TPRV$, PCV, $RSeq$ and $CSeq$ models. For the model selection criteria, the small value of AIC, BIC or CAIC indicates a good model. The three criteria show that $TPRV_{0.1}$ is the best of these five models, because PRV has the highest log-likelihood, but it is penalized as a result of the number of parameters. Compared with PCV, $RSeq$ and $CSeq$ models, $TPRV_{0.1}$ uses fewer parameters and obtains the higher log-likelihood.

Table 5.6 shows the Vuong test results with statistics and the corresponding p-values in the parenthesis. The results are shown without correction and with Akaike and Sch-warz corrections respectively. The Vuong test statistical value without correction cannot indicate which of the $TPRV_{0.1}$ and PRV models is better. However, the statistical values with Akaike and Schwarz corrections suggest that $TPRV_{0.1}$ is better than PRV, since PRV is penalized due to its large number of parameters. Thus, it indicates that $TPRV_{0.1}$ is better than PRV, since $TPRV$ uses fewer parameters to capture the most important dependencies. Comparison with the other three models shows that

Table 5.5: Results of Bivariate Copula Selection and Model Selection

	$TPRV_{0.1}$	PRV	PCV	$RSeq$	$CSeq$
t	22	49	48	41	48
Clayton	1	11	7	9	6
Gumbel	1	3	5	7	3
Frank	21	42	6	51	42
Joe	1	4	41	1	2
S.Clayton [1]	4	16	19	6	12
S.Gumbel [1]	7	11	13	4	12
Total Families	57	136	136	136	136
Total Parameters	79	185	184	160	173
Log-likelihood	5913	6013	5598	5578	5562
AIC	-11669	-11656	-10828	-10836	-10778
BIC	-11620	-11542	-10714	-10737	-10671
CAIC	-11541	-11357	-10530	-10577	-10498

[1] S.Clayton and S.Gumbel are survival versions of Clayton and Gumbel copulas respectively.

$TPRV_{0.1}$ is the best in all three Vuong tests.

In summary, the in-sample performance tests show that $TPRV_{0.1}$ outperforms the other four models.

5.4.4 Out-of-Sample Performance Analysis

The out-of-sample performance is evaluated by the Value at Risk (VaR) forecasting performance. The Multivariate $ARMA$-$GARCH$ (MAG) model [3] is used as reference, since MAG is the benchmark used in the financial area.

[3] The Multivariate $ARMA$-$GARCH$ model used in the case study is multivariate $ARMA$(1,1)-$GARCH$(1,1) with Student-t error distribution; see detail in (Andersen 2009).

Table 5.6: Vuong Test Results

	PRV	PCV	$RSeq$	$CSeq$
No Corr	-0.3130	6.2331	6.6841	6.1411
	(0.7543)	(0.0000)	(0.0000)	(0.0000)
Akaike Corr	2.6653	7.9304	8.5481	7.9852
	(0.0077)	(0.0000)	(0.0000)	(0.0000)
Schwarz Corr	9.5028	11.8271	12.8276	12.2190
	(0.0000)	(0.0000)	(0.0000)	(0.0000)

No Corr, Akaike Corr or Schwarz Corr are short for Vuont test without correction, and with Akaike and Schwarz corrections respectively. The table shows results of Vuong tests at the 5% level, testing the R_1 model ($TPRV_{0.1}$) against the models R_2 (including RPV, PCV, $RSeq$ and $CSeq$ models). If the statistical value is greater than 1.96, it favors the model R_1. If it is less than -1.96, the model R_2 is chosen. If between -1.96 and 1.96, no conclusion is made.

Table 5.7 presents the statistics of a conditional, independent and unconditional coverage and the percentage of exceedance failure. A large p-value indicates that the VaR forecasting is accurate and reliable; however, the p-value should at least be greater than 0.05. The percentage of failure for exceedance should be less than or equal to α. The results in Table 5.7 show that $TPRV_{0.1}$ achieves the best performance of all six models. The other five models fail in one or more tests since the p-value is less than 0.05. Compared to the other five models, our model demonstrates better performance since it has higher p-value for all three levels in all tests. Hence, the out-of-sample performance is in line with that of the in-sample tests. The corresponding VaR forecasting results are shown in Figure 5.2. The VaR forecasting generated by $TPRV_{0.1}$ accurately predicts the volatility. In addition, the result in

Table 5.7 suggests that the Multivariate $ARMA(1,1)$-$GARCH(1,1)$ (MAG) cannot deal with the high-dimensional data, since it can only capture the linear dependence, not the complex dependence shown in high-dimensional data.

5.5 Summary

This chapter presents a truncated partial regular vine model to resolve the issue without imposing restrictions, and on the dependence structures. The model's merit has been demonstrated through an analysis of the complicated structures of portfolios in currency markets. Both in-sample and out-of-sample performance evaluation results highly outperform other methods from statistic and risk evaluation perspectives. It is found that our model $TPRV_{0.1}$ captures important (including asymmetric) dependencies with the least parameters.

Table 5.7: Backtesting Results of Value at Risk for Currencies

	$1-\alpha$	PoF [1]	LR_{UC} [2]	LR_{CC} [2]	LR_{IC} [2]
$TPRV_{0.1}$	99%	2	0.458	0.484	0.026
		0.65%	(0.499)	(0.785)	(0.872)
	95%	12	0.035	0.305	0.270
		3.87%	(0.852)	(0.859)	(0.270)
	90%	27	0.089	0.455	0.366
		8.71%	(0.765)	(0.796)	(0.545)
PRV	99%	2	0.458	0.484	0.026
		0.65%	(0.499)	(0.785)	(0.872)
	95%	15	0.075	0.505	0.430
		4.84%	(0.784)	(0.777)	(0.512)
	90%	29	0.354	0.678	0.324
		9.35%	(0.552)	(0.712)	(0.569)
PCV	99%	3	6.008	6.271	0.263
		0.97%	(0.014)	(0.043)	(0.608)
	95%	15	5.808	7.089	1.281
		4.84%	(0.016)	(0.029)	(0.258)
	90%	30	6.128	9.845	3.717
		9.68%	(0.013)	(0.007)	(0.054)

[1] The PoF is the percentage of exceedance failure. The first row shows the exceedance number, and the second row gives the corresponding percentage;

[2] LR_{UC}, LR_{IC} and LR_{CC} are short for the likelihood ratio of unconditional, independent and conditional coverage respectively. The first row in each cell shows the statistics value, while the corresponding p-value is given in the corresponding parenthesis. The critical value of LR_{UC} or LR_{IC} is 3.841, while the critical value of LR_{CC} is 5.991.

Table 5.8: Backtesting Results of Value at Risk for Currencies

	$1-\alpha$	PoF [1]	LR_{UC} [2]	LR_{CC} [2]	LR_{IC} [2]
$RSeq$	99%	2	3.571	5.816	2.245
		0.65%	(0.059)	(0.055)	(0.134)
	95%	14	4.876	6.157	1.281
		4.52%	(0.027)	(0.046)	(0.258)
	90%	28	7.073	7.306	0.233
		9.03%	(0.008)	(0.026)	(0.629)
$CSeq$	99%	3	6.007	7.471	7.073
		0.97%	(0.014)	(0.027)	(0.289)
	95%	16	7.471	8.752	1.281
		5.16%	(0.006)	(0.013)	(0.258)
	90%	32	7.073	10.789	3.716
		10.32%	(0.008)	(0.005)	(0.054)
MAG	99%	6	2.152	9.553	12.651
		1.94%	(0.014)	(0.027)	(0.289)
	95%	22	6.231	31.802	65.324
		7.10%	(0.044)	(0.000)	(0.000)
	90%	38	4.079	22.248	52.673
		12.26%	(0.043)	(0.000)	(0.000)

[1] The PoF is the percentage of exceedance failure. The first row shows the exceedance number, and the second row gives the corresponding percentage;

[2] LR_{UC}, LR_{IC} and LR_{CC} are short for the likelihood ratio of unconditional, independent and conditional coverage respectively. The first row in each cell shows the statistics value, while the corresponding p-value is given in the corresponding parenthesis. The critical value of LR_{UC} or LR_{IC} is 3.841, while the critical value of LR_{CC} is 5.991.

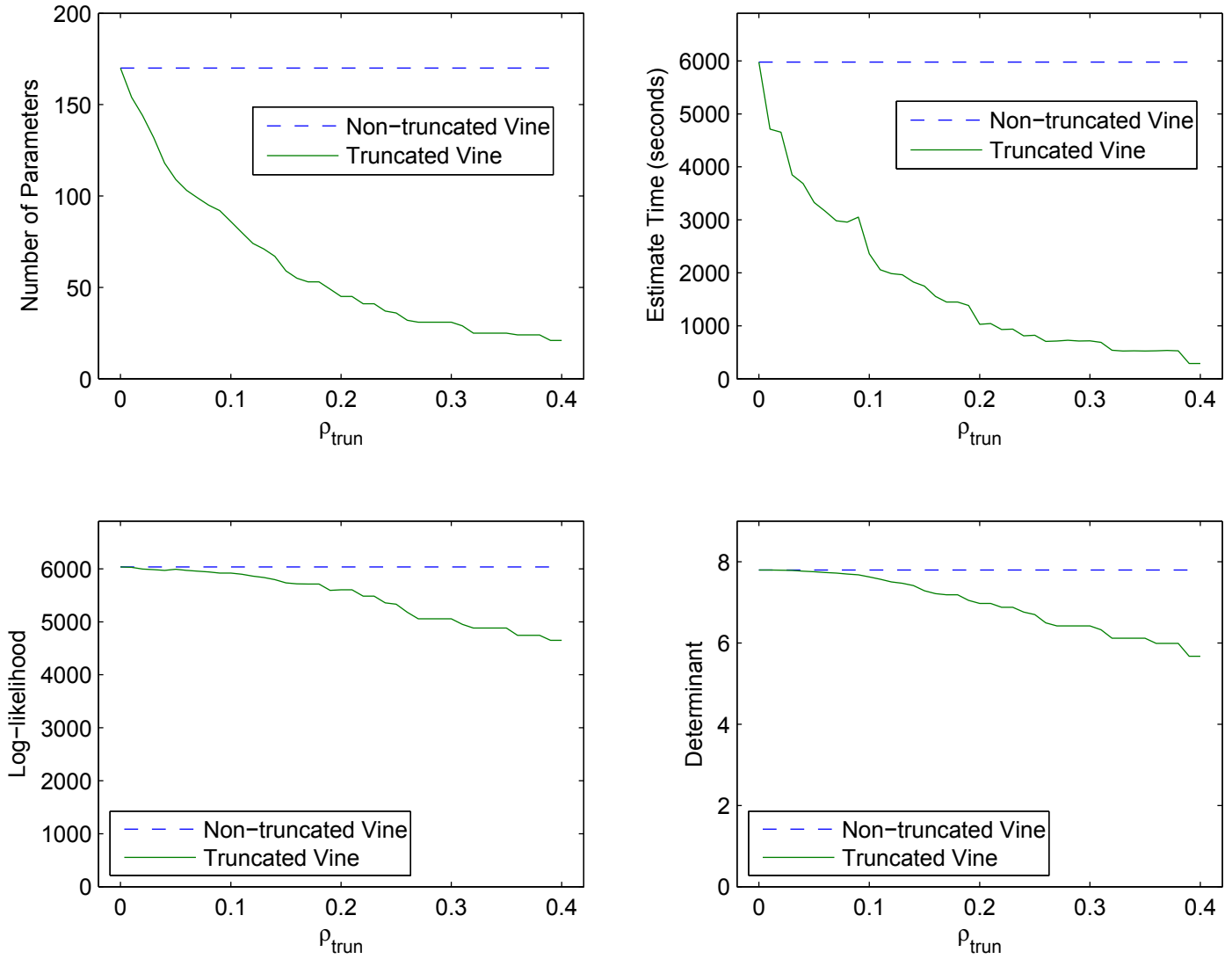

Figure 5.1: Performance analysis of truncated partial regular vines

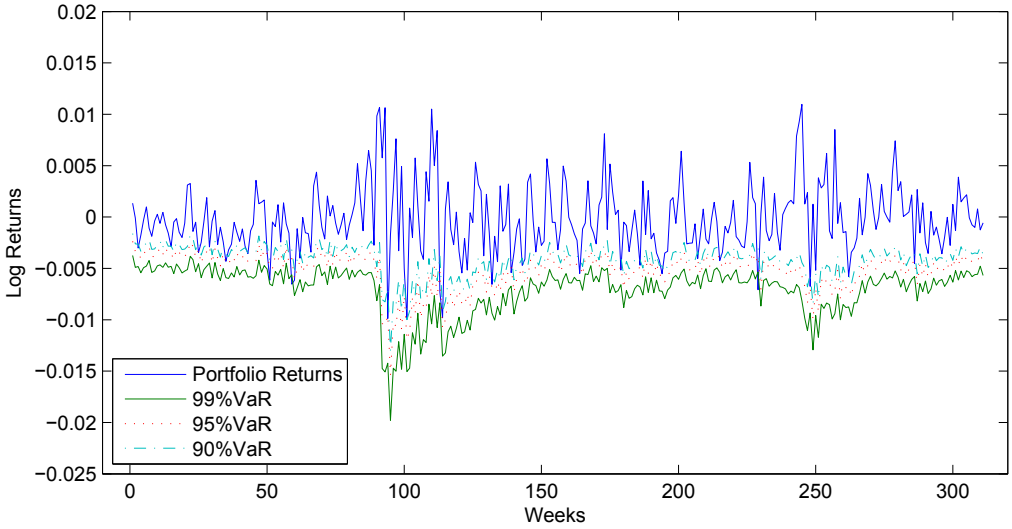

Figure 5.2: The VaR forecast of portfolio returns by using $TPRV_{0.1}$

Chapter 6

Optimal Allocation of High Dimensional Assets through Truncated Partial Canonical Vines

6.1 Introduction

Financial asset returns follow non-normal distributions—asymmetries and skewness very often exists in the distribution of financial asset returns such as in stock returns (Ang & Bekaert 2002) and (Ang & Chen 2002). These asymmetry facts violate the traditional distribution assumption on financial asset returns, making the traditional mean-variance analysis (Markowitz 2012) unreasonable. Some previous studies (Patton 2004), (Riccetti 2012) and (Sun, Rachev, Stoyanov & Fabozzi 2008) have attempted to compare the expected utility obtained from the mean-variance criterion with the approximated utility obtained from the benchmark portfolios (those equally divided allocation portfolios). It has been found that the mean-variance criterion had poor performance in relation to analysing the skew and asymmetric portfolios. The mean-variance criterion was good only at the portfolio that consists

of riskless assets as riskless assets are driven by a normal distribution in line with the mean-variance criterion assumption.

Arrow (1971) laid down a theoretical foundation for the importance of using distributional asymmetries. He suggests that a desirable property of utility functions, such as the *Constant Relative Risk Aversion* utility function is the non-increasing absolute risk aversion. It means that under the non-increasing absolute risk aversion, investors may have a preference for positively skewed portfolios. Asymmetries in the dependence structure have a direct impact on the skewness of the portfolio return. Therefore, while making the portfolio decision favourable to risky assets, it is also essential to consider the existence of and the impact between asymmetries and skewness. In the past, some studies, such as (Patton 2004, Jondeau & Rockinger 2006*b*, Hong, Tu & Zhou 2007, Sun et al. 2008, Christoffersen & Langlois 2011)(detail refer to 2.8.3), have proposed models to construct the dependence structure with only two financial assets. However, they focus on only two or three asset allocations. That is far away from the need of investors. Investors and trading agents generally purchase tens of risky assets, rather than two assets in order to reduce aggressive risk. Therefore, it is important to develop a model that can resolve the difficulties in the high dimensional asset allocation.

There are three challenges in high dimensional asset allocation. First, as discussed above, the correlations between financial assets are asymmetric, rather than normal. With high dimensional input, the dependence structure becomes extremely complex. Indeed, it is difficult to capture and model all of these correlations between assets. Second, it is important to obtain the joint probability density function. However, the high dimensional joint distribution function has a large number of parameters which increases exponentially as the data dimensions increase. Third, each individual asset has its own characteristics, such as volatility clustering and fat tail. It is a challenging task to combine these characteristics into the dependence structure.

To fulfill this need, a truncated partial canonical vine-based dependence model is proposed for an optimisation of the high dimensional asset alloca-

tion. The new model can capture asymmetric and skew correlations in the dependence structure. To address the high dimensional issue, an idea of partial correlation is employed to construct and truncate or simplify the canonical vine dependence structure in the model. In addition, it can capture the most important correlations in the dependence structure, and truncate the useless dependencies(information), which can effectively reduce the complexity of the dependence structure. Furthermore, the $ARMA$-$GARCH$ model is used for marginal distribution to capture volatility clustering and fat tail in financial assets.

The main contribution made by this work is the truncated partial correlation based canonical vine called truncated *Partial Canonical Vine* (PCV) for short. The truncated PCV is optimised to be suitable to high dimensional data input as it can remarkably reduce the number of nodes and parameters and simplify the canonical vine structure. For example, for a 50-variable typical canonical vine model dependence model, it will generate 1225 parameters for dependence structure. However, in the truncated partial canonical vine, the number of parameters is only one tenth of that (around 267 parameters). In addition, the method can test hypotheses in parallel such as: (1) whether these asymmetries are predictable out of sample; and (2) whether it can make better portfolio decisions by using the forecasts of these asymmetries. If the answer to any of these questions is 'yes', then the asymmetries are very important for high dimensional asset allocation. Finding models to fit in-sample data very well without considering the asymmetries and skewness does not imply that it will result in a better out-of -sample portfolio decisions. In the chapter, the model which can capture all important asymmetries and skewness in the dependence structure is built firstly, and then it is compared with those models that do not consider the correlations and/or asymmetric dependence structure between financial assets.

The rest of this chapter is organised as follows. Section 6.2 describes the problem of optimal assets allocation in portfolio, including how to construct the optimal canonical vine and marginal distribution. Section 6.3 discusses

how to evaluate the performance of the model with equally divided allocation and mean-variance criterion. In Section 6.4, it applies the truncated partial canonical vine to capture the dependence structures of two portfolios, and evaluate the performance of the model in comparison to the performance by the mean-variance criterion and equally divided allocation method. Finally, Section 6.5 concludes this chapter with a summary.

6.2 The Portfolio Optimisation Problem and Partial Canonical Vine

6.2.1 CRRA Optimisation Function

Suppose that hypothetical investors follow the class of *Constant Relative Risk Aversion* ($CRRA$) utility functions:

$$U(\gamma) = \begin{cases} (1-\gamma)^{-1} \cdot (P_0 R_{port})^{1-\gamma}, & \text{if} \quad \gamma \neq 1 \\ log(P_0 R_{port}), & \text{if} \quad \gamma = 1 \end{cases} \tag{6.1}$$

where γ is the risk aversion parameter, P_0 is the initial wealth and R_{port} is the portfolio return. In this chapter, the value of risk aversion parameter is considered at four different levels, including $\gamma = 2, 5, 7$ and 10, as suggested by (Campbell, Koedijk & Kofman 2002). $CRRA$ utility function is used to calculate the expected return of the hypothetical investors. If the results are obtained by using the $CRRA$ utility function, then the methods or algorithms are used as a conservative estimate of the other possible results or gains by using other more sensitive utility functions.

The next step is to build a portfolio of returns. The work is focused on the portfolio return with high dimensional assets, which is defined as;

$$P_0 R_{port} = P_0 \cdot (1 + \sum_{i=1}^{n} \omega_i r_{i,t}) \tag{6.2}$$

where $x_{i,t}$ is the asset return at time t, and ω_i is the proportion of wealth for each asset i. Generally, the initial wealth P_0 is set to zero as it does

not affect the choice of weights. Suppose the joint distribution is F_t, with the associated marginal distribution $F_{1,t}, ..., F_{n,t}$, and copula C_t. The density forecasts of the joint distribution $F_{1,t+1}, ..., F_{n,t+1}$ and the copula function C_{t+1} are developed. Then, the forecast function is used to calculate the optimal weights ω^*_{t+1} for the portfolio. The optimal weights, ω^*_{t+1}, are found by maximising the expected $CRRA$ utility function:

$$\begin{aligned}
\omega^*_{t+1} &= \arg\max_{\omega \in W} E_{F_{t+1}}[U(1 + \sum_{i=1}^{n} \omega_{i,t+1} x_{i,t+1})] \\
&= \arg\max_{\omega \in W} \int_{x_{1,t+1}} \int_{x_{2,t+1}} ... \int_{x_{n,t+1}} U(1 + \sum_{i=1}^{n} \omega_i x_i) \\
&\quad \cdot f_{t+1}(x_1, x_2, ..., x_n) \cdot dx_1 \cdots dx_n \\
&= \arg\max_{\omega \in W} \int_{x_{1,t+1}} \int_{x_{2,t+1}} ... \int_{x_{n,t+1}} U(1 + \sum_{i=1}^{n} \omega_{i,t+1} x_{i,t+1}) \\
&\quad \cdot f_{1,t+1}(x_1) \cdots f_{n,t+1}(x_n) \\
&\quad \cdot c_{t+1}(F_{1,t+1}(x_1), F_{2,t+1}(x_2), ..., F_{n,t+1}(x_n)) \cdot dx_1 \cdots dx_n
\end{aligned} \tag{6.3}$$

where $W_{t+1} = \{(\omega_{1,t+1}, ..., \omega_{n,t+1}) \in [0,1]^n : \sum_{i=1}^{n} \omega_i \leq 1\}$ for the short sales constrained investors. The investors will estimate the model of the conditional distribution of returns by using maximum likelihood estimation, and then optimise the portfolio weights via the predicted distributions of return. For the integral function, Monte Carlo replications are used to estimate the value of integral. For the optimal portfolio weights, the *Broyden-Fletcher-Goldfarb-Shanno* algorithm is employed to obtain the optimal weights.

6.2.2 Truncated Partial Canonical Vine

The key step in Equation (6.3) is to form the joint density function $f_{t+1}(x_1, ..., x_n)$ at time $t+1$. Equation (2.3) shows that the joint density function can be divided into two parts: the copula function c_{t+1} and the marginal distributions $f_{1,t+1}(x_1) \cdots f_{n,t+1}(x_n)$. This Section discusses how to produce a copula

density function. The construction of marginal distributions is explained in the following Section 6.2.3.

One way to build high-dimensional copula density function is to use the canonical vine to build the dependence structure as proposed by (Aas et al. 2009). The basic scheme for modelling high-dimensional dependence structure with the canonical vine is to decompose multivariate density functions into many conditional pair copulas. These pair copulas are bivariate copulas in one time. The model based on the canonical vine transforms one high dimensional dependence structure into multiple two-dimensional structures. However, one important issue of the canonical vine is that if the variables are large in number, the canonical vine will become quite complex. In addition, the number of bivariate copulas and their corresponding parameters of canonical vine will increase exponentially as the variables increase. Therefore, the partial correlation-based canonical vine is developed to model the high dimensional dependence structure.

The principle for new canonical vine construction is to capture the most important correlation in the dependence structure while at the same time decreasing the number of nodes, and reducing the complexity of the dependence structure. That is, the new canonical vine can capture the most important correlation, and ignore the weak correlations. Following this principle, the partial correlation based canonical vine is suitable, since (1) it is built via partial correlation with the bottom to top strategy, which can ensures the weak correlations are put to the edges in bottom trees, and (2) according to Algorithm 3.2, it can be easily truncated, which ensure that it can be used in high-dimensional cases. Please refer to Section 3.3.3 for the construction of the partial correlation based canonical vine dependence structure.

Once the partial correlation based canonical vine dependence structure is identified, the next step is to choose the bivariate copula for each edge of trees. The bivariate t copulas are selected, since the t copulas have symmetric tail dependence (see detail in Section 2.1.3).

6.2.3 Marginal Distribution Specification and Parameters Estimate

Marginal Distribution Specification

Typically, $ARMA$(p,q)-$GARCH$(1,1) models are used as marginal distributions, since the stock returns (time series data) are examined in the empirical analysis. Generally, AR(1)-$GARCH$(1,1) is suitable for all of these stock returns. Please refer to Section 2.7, Section 3.4 and (Andersen 2009) for detail in relation to the $ARMA$-$GARCH$ models.

Suppose $X_t(t = 0, 1, ..., \mathfrak{T})$ be a time series of the price on a financial asset, such as the stock market index. Then the return r_t of financial asset can be defined as $log(X_t/X_{t-1})$. There are n assets with returns $r_{t,1}, ..., r_{t,n}$. The AR(1)-$GARCH$(1,1) is defined as follows:

$$\begin{aligned} r_{t,j} &= c_j + \Phi_j r_{t-1,j} + \varepsilon_{t,j}, \\ \varepsilon_{t,j} &= \sigma_{t,j} \cdot Z_{t,j} \\ \sigma^2_{t,j} &= \omega_j + \alpha_j \varepsilon^2_{t-1,j} + \beta_j \sigma^2_{t-1,j} \end{aligned} \tag{6.4}$$

where $j = 1, ..., n, t = 1, ..., \mathfrak{T}$ and $Z_{t,j}$ is the error distribution. In the chapter, it follows skewed Student-t distribution.

Parameter Estimate

In the high-dimensional cases, the number of parameters is very large. As discussed in Section 2.2.2, it is necessary to employ the two-stage procedure. Typically, as shown in Equation (3.10), the log-likelihood has two parts, including the log-likelihood of copula and the log-likelihood of margins. Let $\boldsymbol{\theta}^m$ and $\boldsymbol{\theta}^c$ be the parameters set of the marginal distributions and copula function respectively. For example, the parameters of jth marginal distribution is $\theta^m_j = (\mu_j, \Phi_j, \Theta_j, \alpha_j, \beta_j)$ The parameters of marginal distributions ($\boldsymbol{\theta}^m$) are estimated at first, and then, the parameters of the copula functions $\boldsymbol{\theta}^c$ are estimated given $\boldsymbol{\theta}^m$. The two-stage procedure can effectively reduce the computation burden.

6.3 Evaluation Methods

The final amount of the portfolio, the utility obtained by daily returns, are used to compare the assets allocation with different portfolio decisions. The final amount is the amount in dollars obtained at the end of the entire out of sample period (testing period). To compare the utility between different portfolio decisions, the opportunity cost, also called management fee or forecast premium, is calculated in the experiment. The opportunity cost is the amount that investor would pay to switch from the the equally divided allocation portfolio to analysed allocation. The benchmark of assets allocation is by equally dividing the allocation portfolio, which means the weights of all assets are equal. The performance of the partial canonical vine is compared with mean-variance criterion and the equally divided allocation. Suppose that r_{port} is the optimal portfolio return obtained by partial canonical vine or mean-variance, and r^*_{port} is the return obtained from the equally divided allocation portfolio. In other words, the opportunity cost is actually the return which is added to the return obtained from the equally divided allocation portfolio, to make sure the investor is indifferent to the returns obtained from the analysed model. Then, the opportunity cost Δ can be defined as:

$$U(1 + r_{port} + \Delta) = U(1 + r^*_{port}) \tag{6.5}$$

Equation (6.5) can be resolved via the Taylor approximation with the $CRRA$ utility function in (Jondeau & Rockinger 2006*b*). For example, the opportunity cost Δ can be rewritten as follows:

$$\begin{aligned}\Delta =&(\mu^*_{port} - \mu_{port}) - \frac{\gamma}{2}(m^{*2}_{port} - m^2_{port}) + \frac{\gamma(\gamma+1)}{3!}(m^{*3}_{port} - m^3_{port}) \\ &- \frac{\gamma(\gamma+1)(\gamma+2)}{4!}(m^{*4}_{port} - m^4_{port})\end{aligned} \tag{6.6}$$

where μ^*_{port} and μ_{port} represents the mean return obtained from the equally divided allocation portfolio and the analysed portfolio respectively, and m represents the non central moments: $m^i_{port} = mean[r^i_{port}]$, where $i = 1, 2, 3$ *and* 4. Equation (6.6) calculates the opportunity cost. The opportuni-

ty cost can be used to compare the performance between different portfolio decisions.

6.4 Case Study

6.4.1 Data and Model Specification

Two financial asset portfolios are used in the evaluation of the performance of the proposed model. One portfolio composes a comprehensive index *S&P* 500, and 50 stocks from 10 industries. The other portfolio consists of a comprehensive index *Stoxx50* Euro, five national leading stock indices, and 44 stocks. All the data are downloaded from Yahoo Finance (http://finance.yahoo.com). The data in the both portfolios span 1200 trading days from 01/10/2004 to 31/07/2009. In the data pre-processing step, these returns and indices are calculated by taking the log difference of the prices on every two consecutive trading prices.

As described in Section 6.2.3, $AR(1,1)$-$GARCH(1,1)$ is considered as the marginal distribution to capture the skewness. The Ljung Box test (McLeod & Li 1983) was used for checking the existence of residual autocorrelation for all of the stocks and indices. Table 6.1 shows all of the partial correlations in the non-truncated partial canonical vine. All of these partial correlations are used with the absolute value as the study is focused on the extent of correlations rather than positive/negative correlations. It can be seen that $Stoxx50E$ has a value larger than $S\&P500$ at all levels. It indicates that the stocks in the portfolio of $Stoxx50E$ have much stronger correlations than those in the portfolio of $S\&P500$. This is understandable as the portfolio of $S\&P500$ is built by using 500 stocks from 10 industries. The stocks in the portfolio $S\&P500$ are strictly selected from the least correlations to each other.

Then, the following models are used in the case studies.

- $TPCV_{0.05}$: Truncated partial canonical vine, built by Algorithm 5.1,

Table 6.1: The Partial Correlation Analysis for Portfolios $S\&P500$ and $Stoxx50E$

	S&P500	Stoxx50E
Min	0.0002	0.0004
Max	0.5281	0.8702
Mean	0.0572	0.0787
25%Quantile	0.0144	0.0171
25%Quantile	0.034	0.0422
75%Quantile	0.0641	0.0929

and then truncated by Algorithm 3.2 with truncation value 0.05 ($\rho_{sign} = 0.05$);

- $TPCV_{0.1}$; Truncated partial canonical vine, built by Algorithm 5.1, and then truncated by Algorithm 3.2 with truncation value 0.01($\rho_{sign} = 0.1$);

- MV: Mean-variance model, built by (Markowitz 2012);

- ED: Equally Divided Allocation.

Figure 6.1 presents the result of the truncated and simplified partial canonical vine for two portfolios $S\&P500$ respectively. The truncated partial canonical vines are using conditional independence copulas to replace the weak correlations (refer to Algorithm 3.2). The dash lines in all figures indicate the non-truncated/simplified partial canonical vine. In Figure 6.1, the total number of the truncated vine decreases sharply, and then goes slowly after the significant value 0.14. However, the simplified vine decreases at the beginning, and then remains at a certain level. The estimate time for both the truncated and simplified vine reduces greatly. The truncated vine uses less time than the simplified one with the same significant value ρ_{sign}. For log-likelihood, the gap between the simplified vine and the non-truncated

Table 6.2: The Comparison between Truncated, Simplified and Non-truncated Vines

	ρ_{sign}	Total Parameters	Estimate * Time	Log likelihood
$S\&P500$				
Simplified PCV	0.05	1521	6401	15478
	0.1	1408	3172	15247
Truncated PCV	0.05	826	2889	14911
	0.1	338	597	12874
Non-truncated PCV		1641	10055	15669
$Stoxx50E$				
Simplified PCV	0.05	1590	8983	24255
	0.1	1439	6196	24040
Truncated PCV	0.05	1049	6141	23702
	0.1	518	1811	22625
Non-truncated PCV		1729	14158	24537

* The estimate time is seconds.

vine is very small. The log-likelihood of truncated vine decreases smoothly at the beginning, and then reduces quickly. A similar conclusion can be reached for the portfolio $Stoxx50E$ via Figure 6.2. Thus, according to the compared results of Figure 6.1 and Figure 6.2, the truncated canonical vine with significant value 0.05 and 0.1, such as $TPCV_{0.05}$ and $TPCV_{0.1}$ are used in the case studies. The details of the two models are shown in Table 6.2.

6.4.2 Experiment Results and Analysis

The performance of truncated partial canonical vines are evaluated by measuring the opportunity cost. A moving window of 1200 observations, approximately 5 years of daily returns from 01/10/2004 to 31/07/2009, was used to

construct the model. The test period was from 01/08/2010 to 01/03/2012 with 730 observations of daily returns. The performance of models is evaluated in the two portfolios: the European stock markets $Stoxx50E$ and United Stated stock markets $S\$P500$. All the portfolio decisions are re-balanced at the end of every month, and no cost is assumed for the re-balancing. The performance of models ($TPCV_{0.05}$ and $TPCV_{0.1}$) are compared with the mean-variance criterion (MV) and the equally divided allocation (ED), to understand whether new models are useful.

Table 6.3 and 6.4 shows the results related to the opportunity costs, utilities and the final amounts for the two portfolios $S\$P500$ and $Stoxx50E$ respectively. The utilities are calculated by using the $CRRA$ utility functions with the sample returns. The two tables provide strong evidence that the two truncated partial canonical vine models are the best at all levels γ for both portfolios. In detail, for the portfolio $S\&P500$, the performance of two truncated partial canonical vines $TPCV_{0.05}$ and $TPCV_{0.1}$ are compared. There is no obvious difference between these two truncated partial canonical vines, indicating that the two vines implement a similar forecasting of the samples. However, the number of parameters in $TPCV_{0.1}$ is only half of those in $TPCV_{0.1}$. $TPCV_{0.1}$ is sufficient to model the dependence structure.

The mean-variance criterion has a poor performance as the opportunity cost is negative at all of the levels γ. It indicates that if investors conduct assets allocation on the basis of mean-variance analysis, the final profit would be less than those on the basis of the equally divided allocation. Therefore, the mean-variance criterion is not useful. Considering the good performances of the canonical vine model, the mean-variance criterion, which is caused by the normal distributions, cannot catch the features of asymmetry and skewness within these stocks and indices. For the portfolio $Stoxx50E$, the mean-variance criterion is not useful either. The performance by the mean-variance criterion in $Stoxx50E$ is worse than those in $S\&P500$. The final amount is even less than one, suggesting that investors will obtain loss if they allocate the assets under the mean-variance criterion. The trouble with

the mean-variance criterion is that $Stoxx50E$ has stronger asymmetry and skewness than the criterion in $S\&P500$. However, the two canonical vines in $Stoxx50E$ perform better than those in $S\&P500$ as the opportunity cost is large at all level γ.

Figure 6.3 shows the portfolio values obtained from the truncated partial canonical vines, the mean-variance criterion and equally divided allocations at the end of each month. The trend obtained from the canonical vine in $Stoxx50E$ shows stronger increasing trends and less volatility than those in $S\&P500$. It indicates that the canonical vine has a better performance in $Stoxx50E$. For the trend obtained from the mean-variance criterion, there is no obviously different from the trends from the equally divided allocation in $S\&P500$. However, the trend of mean-variance is worse than those of the equally divided allocation. It can be concluded that the mean-variance criterion performs a bit better in $S\&P500$ than in$Stoxx50E$.

In summary, it can be concluded that:

(i). The truncated partial canonical vine has a better performance in high dimensional assets portfolios of strong asymmetry and skewness,

(ii). The mean-variance criterion does not have a good performance in a high dimensional assets portfolio that has asymmetry and skewness,

(iii). Compared with the equally divided allocation, the new partial canonical vine has a better performance

However, the performance by the mean-variance criterion has no obvious difference, or is even worse if the high dimensional assets portfolio has strong asymmetry and skewness.

6.5 Summary

This chapter proposes a truncated partial canonical vine to resolve the issue of high-dimensional asset allocations. To address the touch computational

issues caused by high dimensional assets, the partial correlation technique is employed to reduce the complexity of the dependence structure to make investors understand the model easily. The experimental results and analysis have shown that the truncated partial canonical vine has a better performance for portfolios of strong asymmetry and skewness in comparison to the mean-variance criterion, which is currently a widely used method.

Table 6.3: The Results of the Opportunity Costs, Final Amounts and Utilities for $S\&P500$

	$S\&P500$			
Relative Risk Aversion (γ)	2	5	7	10
Opportunity Cost				
$TPCV_{0.1}$	1.33%	3.21%	5.64%	7.32%
$TPCV_{0.05}$	1.31%	3.18%	5.56%	7.21%
Mean Variance Model	-2.36%	-3.62%	-1.24%	-1.19%
Utilities				
$TPCV_{0.1}$	-0.719	-0.063	-0.017	-0.005
$TPCV_{0.05}$	-0.714	-0.061	-0.017	-0.003
Mean Variance Model	-0.980	-0.250	-0.132	-0.078
Equally Divided Allocation	-0.952	-0.206	-0.124	-0.072
Final Amount				
$TPCV_{0.1}$	1.39	1.41	1.46	1.50
$TPCV_{0.05}$	1.40	1.42	1.47	1.50
Mean Variance Model	1.02	1.00	1.04	1.04
Equally Divided Allocation	1.05	1.05	1.05	1.05

Table 6.4: The Results of the Opportunity Costs, Final Amounts and Utilities for $Stoxx50E$

	$Stoxx50E$			
Relative Risk Aversion (γ)	2	5	7	10
Opportunity Cost				
$TPCV_{0.1}$	1.91%	4.19%	6.14%	10.01%
$TPCV_{0.05}$	1.89%	4.17%	6.12%	9.95%
Mean Variance Criterion	-2.55%	-6.52%	-3.20%	-1.40%
Utilities				
$TPCV_{0.1}$	-0.6654	-0.040	-0.009	-0.001
$TPCV_{0.05}$	-0.0649	-0.038	-0.009	-0.001
Mean Variance Model	-1.075	-0.610	-0.293	-0.122
Equally Divided Allocation	-0.943	-0.198	-0.117	-0.066
Final Amount				
$TPCV_{0.1}$	1.53	1.58	1.62	1.64
$TPCV_{0.05}$	1.54	1.60	1.62	1.65
Mean Variance Model	0.93	0.80	0.91	0.99
Equally Divided Allocation	1.06	1.06	1.06	1.06

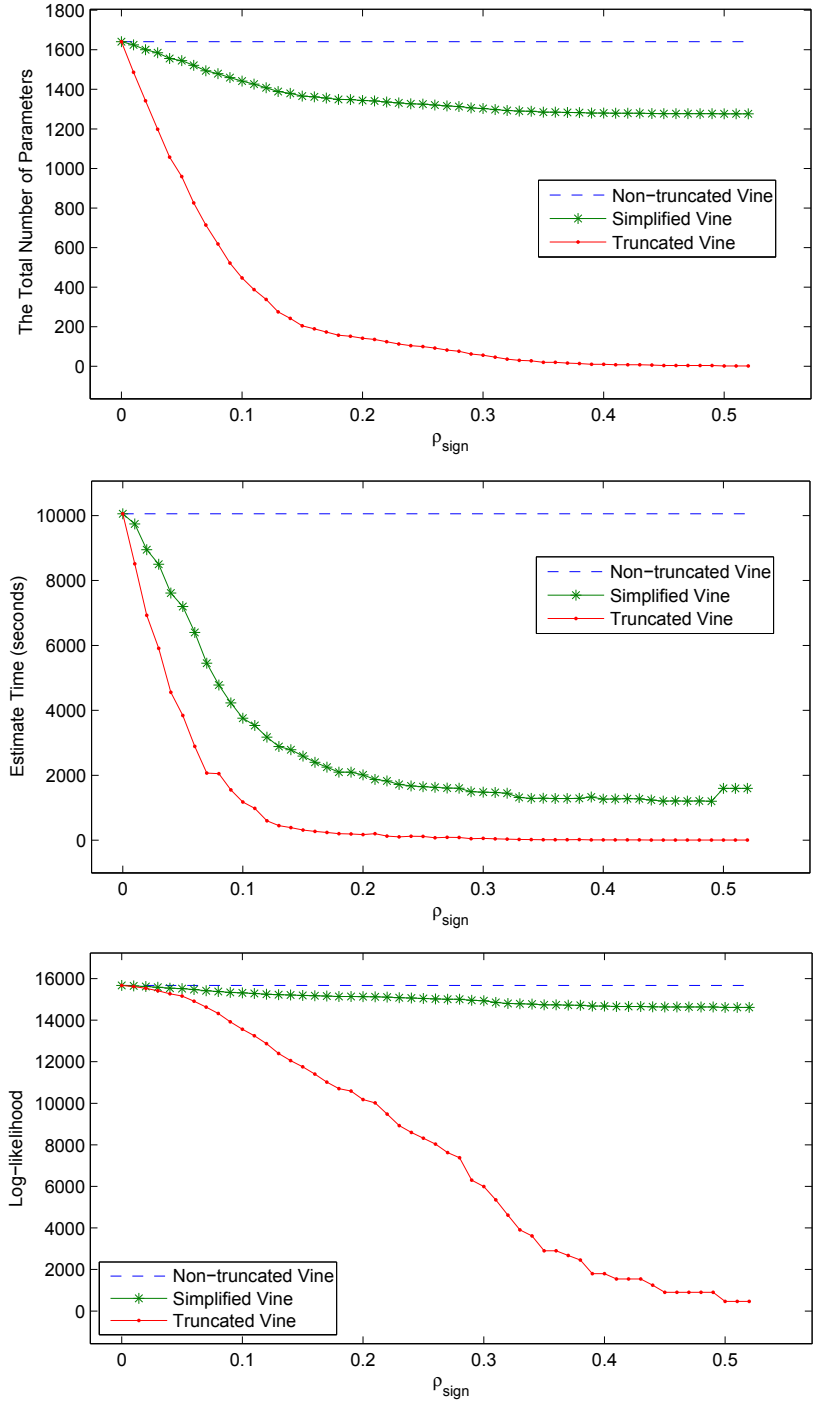

Figure 6.1: The results of truncated and simplified partial canonical vine against ρ_{sign} for *S&P500*

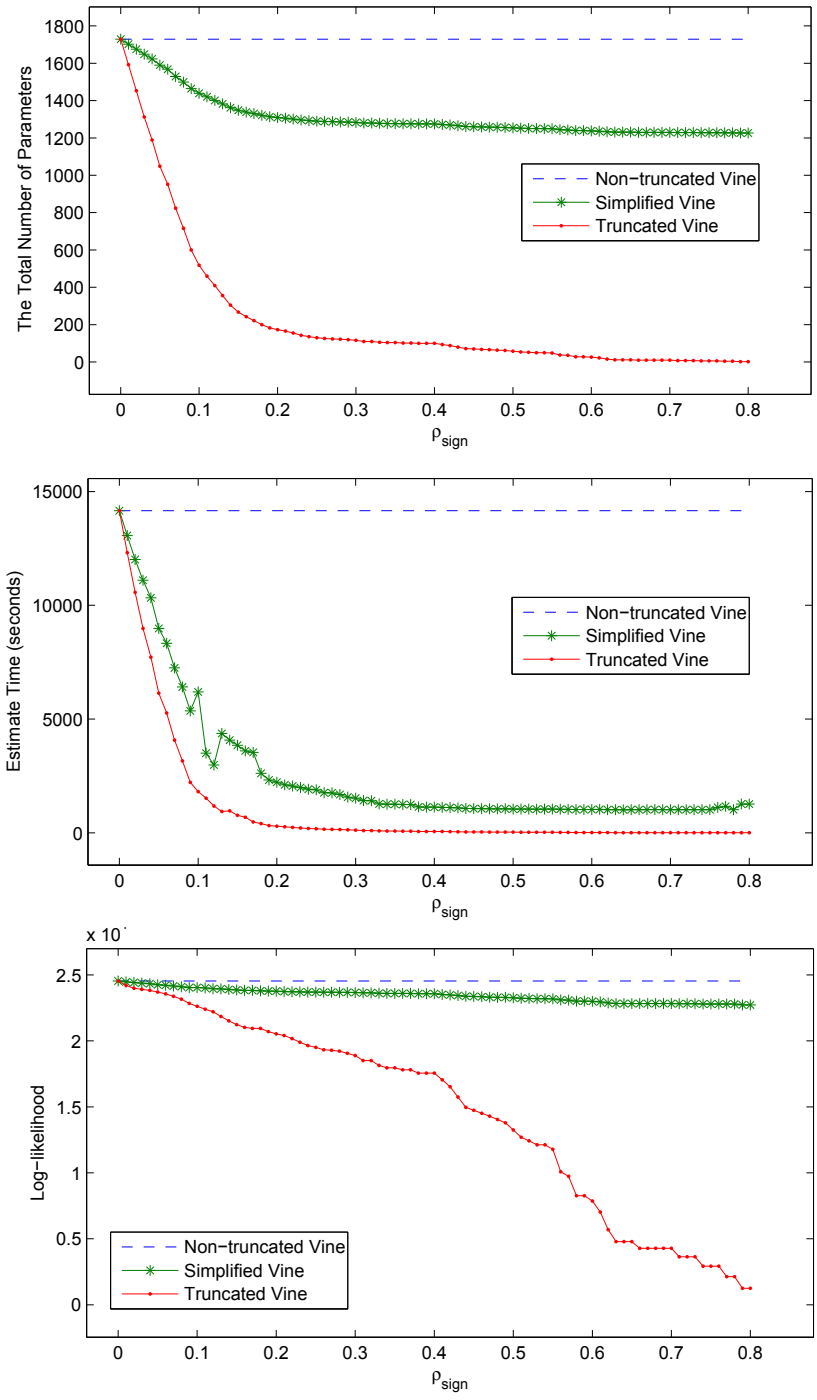

Figure 6.2: The results of truncated and simplified partial canonical vine against ρ_{sign} for *Stoxx50E*

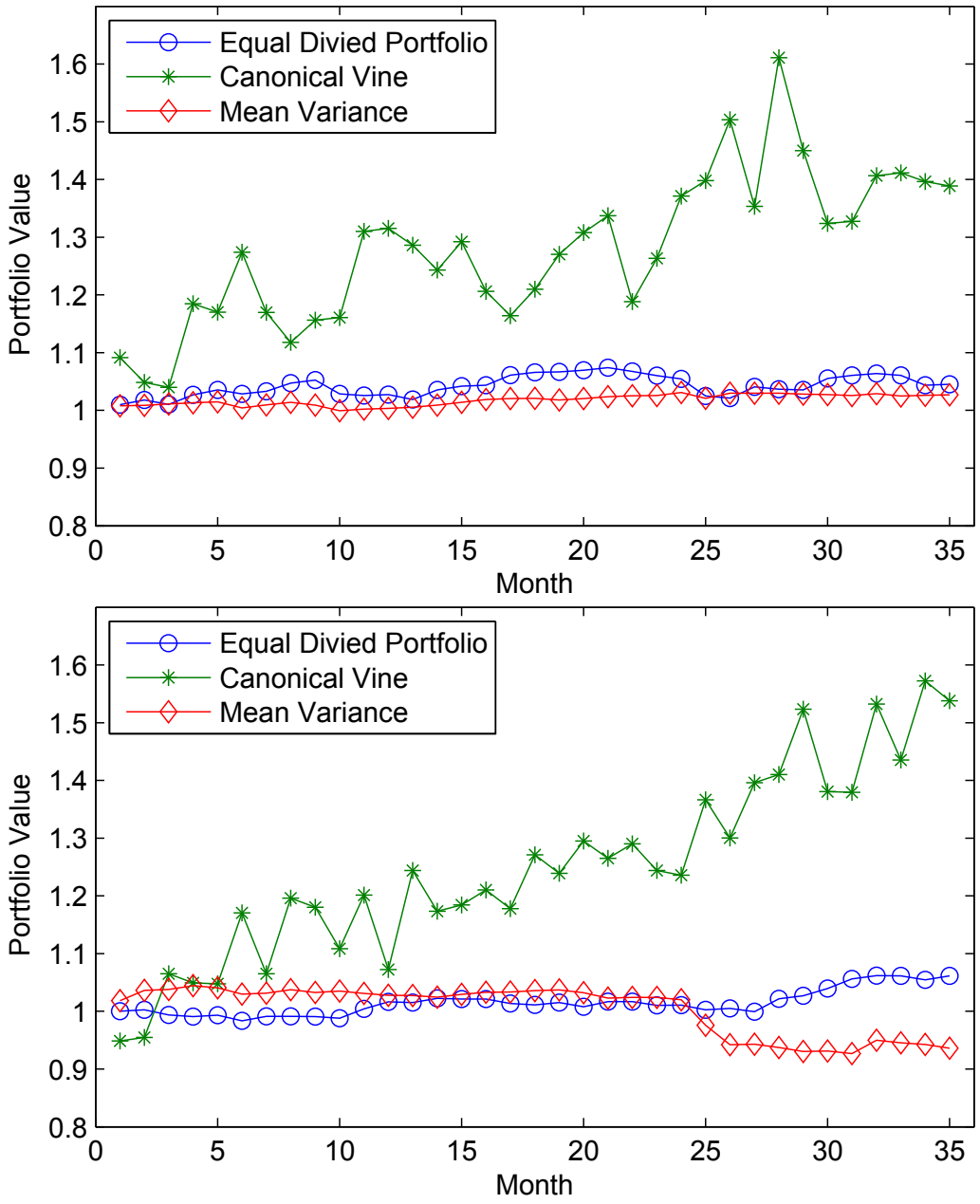

Figure 6.3: The portfolio values over 35 months for the truncated partial canonical vine ($TPCV_{0.1}$), mean-variance, and equally divided allocation, $\gamma = 2$. The up one is for $S\&P500$, and the bottom one is for $Stoxx50E$

Chapter 7

Conclusions and Future Work

7.1 Conclusions

In conclusion, this thesis has proposed techniques to address four important issues in high-dimensional dependence modelling, including high dimensional data (RI 1), dependence structure assumption (RI 2), dependence structure truncation and optimization (RI 3) and high dimensional dependence evaluation (RI 4).

In particular, Chapter 3 has addressed the research issues RI 1, 3 and 4. It presents the truncated partial correlation-based canonical vine copula. The partial correlation is employed to build the canonical vine dependence structure by using bottom to top construction strategy. Then, it develops a truncation method, which sets bivariate copulas with weak correlation to the conditional independence copula. Compared with the non-truncated canonical vine, the truncated one can still maintain the most important dependence but many unimportant nodes are removed to simplify the canonical vine structure. Thus, the truncation method can greatly reduce the computation burden of the canonical vine with high-dimensional data, without losing any important dependence. The truncated partial correlation-based canonical vine is applied to construct and analyze dependence structures of European stocks as case studies. Its performance is evaluated by measuring

a portfolio of Value at Risk, a widely used risk management measure. In comparison to a very recent canonical vine autoregressive model, the truncated partial correlation-based model has a much better quality of Value at Risk, providing insightful knowledge for investors to control and reduce the aggregation risk of the portfolio.

Chapter 4 has addressed the research issues RI 1 and 2. It presents the partial correlation-based regular vine to model the asymmetric, upper and lower tail dependence between multivariate variables. The partial correlation-based regular vine can uniquely determine the correlation matrix and be algebraically independent. It indicates that dependence structure constructed via partial correlation is more flexible, since the current tree structure is independent from the established tree structure and bivariate copulas selection. In addition, for all linking bivariate copula on the partial regular vine dependence structure, only the BB1, survival BB1, BB7 and survival BB7 are used since these copula have both lower and upper tail dependence that can range independently from 0 to 1. The partial correlation-based regular vine copula is tested on a cross-country stock market data set to analyse the asymmetry and tail dependence. The high prediction performance is examined by the Value at Risk, and the results show that the partial regular vine copula with asymmetric dependence and two tail dependence has a better performance in predicting Value at Risk.

Chapter 5 has addressed the research issues RI 1 and 2 and 3. It proposes the truncated partial correlation-based regular vine with bottom to top building strategy to model high-dimensional dependence. It can identify important dependencies and information among high-dimensional variables, truncating the irrelevant information to reduce the parameter estimation time significantly. The truncated partial correlation-based regular vine can capture actual asymmetric dependence in high-dimensional data by incorporating regular vine theory into the construction of the dependence structures, which does not impose assumptions on the dependence structures. In addition, the model is embedded with the copula family which supports a wide

range of flexible lower and upper tail dependence. The model is constructed by a new bottom-up construction strategy, which ensures that the weakest correlations are at the bottom, and the strongest correlations appear at the top of dependence structures. The model is applied to construct the dependence structures of 17 currency markets over 17 years as a case study. The model's in-sample performance is evaluated via a standard model selection criteria Vuong test, and the out-of-sample performance is evaluated by Value at Risk, a widely used industrial benchmark. The extensive experiment results show that our model and its intrinsic design significantly outperforms industry baselines, and provide financially interpretable knowledge and profound insights into the high-dimensional dependence structures of complex financial variables.

Chapter 6 addresses the research issue RI 1, 3 and 4. It proposes the constant relative risk aversion utility function through the partial-correlation based-canonical vine, which aims to resolve the assets allocation optimisation issue. To address the high dimensional issue, a truncated partial correlation-based canonical vine is built. Thus, the utility function can model the asymmetries and skewness of joint distributions of assets in high dimensional space. The importance of using the asymmetries information is assessed by comparing the performance of a portfolio based on the typical mean-variance criteria and that of a portfolio based on the truncated partial correlation-based canonical vine. The results show that the investors using the forecasts of these asymmetries can make better portfolio decisions than those who ignore the asymmetries information.

7.2 Future Work

The exploration in this thesis shows that there is great potential in applying vine copula models to complex dependence modelling. In future research, potential directions can be explored from both a theoretical and applied perspective. These tasks include:

(i). **Partial correlation-based D vine**: This thesis has implemented the partial correlation-based canonical vine and regular vine. The canonical vine and D vine are two special case of the regular vine. D vine, however, may be more efficient for those with a path-like dependence structure.

(ii). **Model simplification**: This thesis has proposed the truncated methods for both the partial correlation-based canonical vine and regular vine. However, the methodology is mainly dependent on the domina knowledge. From the data-driven perspective, it is not enough. Is it possible to define the significant value of truncation method through the data information prior to modelling dependence? For a given vine, how do we select the vine (canonical vine or regular vine) which has all of the specified conditional independence statement? Is it possible to find a more efficient conditional independence test that may allow us to find all conditional copulas that can be set to conditional independence?

(iii). **Vine copula and graphical models**: The relationship between vine copula models and graphical models, such as the Bayesian copula network or chain graphs are investigated in this thesis. However, whether the combination of the vine copula and the probability graphical models are better than the individual model is of particular importance. For example, is it possible to combine the hidden Markov model with the vine copula, since they are both from the graphical models?

(iv). **Vine copula with time-varying parameters**: The models with time-varying parameters are of great interest in the financial field. Vine can provide the flexibility to model the dependence in time. If allowing for the parameters to vary over time, the predictive power of the vine copula model may grow significantly. The task, however, is very challenging since it will add extra parameters into the vine copulas, and make the model much more complex.

Appendix A

Dependence Measurements

In the section, two different dependence measures are introduced, including the Pearson correlation and Kendall's tau (τ).

A.1 Pearson Correlation

Classically, the Pearson correlation, which is also called produce moment correlation or correlation coeddicient, is the most common measure of dependence, and especially in finance field. It is a skalar measure for the linear correlation of two random variables.

Definition A.1 (Pearson Correlation) *The Pearson correlation of two random variables X_1 and X_2 with finite expectations $E(X_1)$ and $E(X_2)$ and corresponding finite variances $\sigma^2_{X_1}$ and $\sigma^2_{X_2}$ is*

$$Corr_{pearson}(X,Y) = \frac{cov(X,Y)}{\sqrt{\sigma^2_X \sigma^2_Y}} = \frac{E[XY] - E[X]E[Y]}{\sqrt{\sigma^2_X \sigma^2_Y}} \qquad \text{(A.1)}$$

X and Y are uncorrelated if $Corr_{Pearson}(X,Y) = 0$.

The Pearson correlation can have values between -1 and $+1$, e.g., $Corr_{Pearson} \in [-1,1]$. For normal distributed random variables, the Pearson correlation is very useful since it fully specifies the dependence between them. However,

for general distribution, it can only measure the linear dependence. In addition, Kurowicka and Cooke (Kurowicka & Cooke 2006*c*) present some some disadvantages of the Pearson correlation, which are given as follow :

(i). The Pearson correlation is not defined if the expectation or variance of X and Y are not finite, such as Cauchy distribution ;

(ii). The Pearson correlation is not invariant under non-linear strictly increasing transformations ;

(iii). The possible values of the Pearson correlation depend on marginal distribution;

To model the dependence and marginal distribution separately, the above disadvantages (ii) and (iii) are unpleasing, since they are influencing each other. The Kendall' tau τ can overcome these shortcomings.

A.2 Kendall's Tau

Kendall's tau (τ) is a rank correlation, which indicate that it does not depend directly on the values, instead it is a function of relations between realizations.

Definition A.2 (Kendall's Tau) *Let (X_1, Y_1) and X_2, Y_2 be two independent pairs of random variables with joint cdf F and marginal distributions F_X and F_Y. Kendall's tau is given by*

$$\tau(X,Y) = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1, Y_2) < 0] \quad \text{(A.2)}$$

In case X and Y independent, $\tau(X,Y) = 0$. In terms of copula, the Kendall's tau is expressed as follow:

$$\tau(X,Y) = 4 \int \int_{[0,1]^2} C(u_1, u_2) - 1 \quad \text{(A.3)}$$

for continuous X and Y with $u_1 = F_X$ and $u_2 = F_Y$. For the empirical version of Kendall's tau, Nelsen (Nelsen 1999) defines *concordant* and *discirdant* pair of variables. Two continuous random variables X and Y are said to form a concordant pair if large values of one are associated with larhe valyes of the other one. By the discordance are the large values of one random variable associate with small values of the other respectively. For X, Y and their pairs of observation (x_i, y_i) and (x_j, y_j),

(i) concordant, if $(x_i - x_j)(y_i - y_j) > 0$;

(ii) discordant, if $(x_i - x_j)(y_i - y_j) < 0$.

Further let N_{conc} and N_{disc} denote the number of concordant pairs and the number of discordant pairs respectively. Let $N_{extra-x}$ and $N_{extra-y}$ denote the number of tied pairs among all pairs in the x and y respectively. Then, the empirical Kendall's tau can be estimated from an underlying data set by:

$$\hat{\tau}(X, Y) = \frac{N_{conc} - N_{disc}}{\sqrt{N_{conc} + N_{disc} + N_{extra-x}}\sqrt{N_{conc} + N_{disc} + N_{extra-y}}} \tag{A.4}$$

A high value in the numerator means that most pairs are concordant, suggesting that two two rankings are consistent and $\hat{\tau}$ will tale a value close 1. One the other hand, if most pairs are discordant, $\hat{\tau}$ will take a value close -1.

Appendix B

The Determinant of Partial Correlation-based Canonical Vine and Regular Vine

One important property of partial correlation is that the product of 1 minus the square partial correlation equals to the determinant of the correlation matrix. Before discussing the important property of partial correlation, the multivariate correlation is reviewed firstly,which is defined by:

Definition B.1 (Multiple Correlation) *The multiple correlation* $R_{1\{2,...,n\}}$ *of variables* X_1 *with respect to* $X_2, ..., X_n$ *is*

$$1 - R^2_{1\{2,...,n\}} = \frac{Demt}{K_{11}} \tag{B.1}$$

where demt is determinant of the correlation matrix, and K_{11} *is denoted as the* $(1,1)$*th cofactor of the correlation matrix.*

Kendall and Stuar (Kendall & Stuart 1961) show that $R_{1\{2,...,n\}}$ is non-negative and satisfies:

$$1 - R^2_{1\{2,...,n\}} = (1-\rho^2_{1n})(1-\rho^2_{1,n-1\,;\,n})(1-\rho^2_{1,n-2\,;\,n-1,n})...(1-\rho^2_{1,2\,;\,3...n}) \tag{B.2}$$

It is shown that $R_{1\{2,...,n\}}$ is invariant under permutation of $\{2,...,n\}$ and

$$Demt = (1 - R_{1\{2,...,n\}})(1 - R_{2\{3,...,n\}})...(1 - R_{n-1\{n\}}) \tag{B.3}$$

Obviously, $R_{n-1\{n\}} = \rho_{n-1,n}$.

Then, the following theorem is presented and proved. The detail refers to (Kurowicka & Cooke 2006*a*).

Theorem B.1 *Let Demt be the determinant of the n-dimensional correlation matrix (Demt* > 0*). For any partial correlation vine,*

$$Demt = \prod_{i=1}^{n-1} \prod_{e \in E_i} (1 - \rho^2_{C_{e(a)}, C_{e(b)}\,;\, D_e}) \tag{B.4}$$

where $C_e, e = \{a, b\}$*, and* D_e *are conditioned and conditioning sets of edge e.*

Proof B.2 *Let* $\{1, 2 \,|\, 3, ..., n\}$ *denote the constraint of the single edge of the bottom tree* T_{n-1}*. Then, all m-descendants of the edge containing the variable 1 are collected. Due to the property of regular vine, variable 1 only can occurs only in the conditioning set of its m-descendent edges, and the conditioning set of an m-child is a subset of the conditioning set of its m-parent. Therefore, variable 1 occurs exactly once with every ither variable in the conditioned set of some edges. The constraints of the edges containing variable 1 can be rewritten as follow :*

$$\{1, 2 \,|\, 3, ..., n\}, \{1, 3 \,|\, 4, ..., n\}, ..., \{1, n-1 \,|\, , n\}, \{1, n\}$$

The corresponding partial correlations with these m-descendent edges are

$$\rho_{1,2\,;\,3,...,n}, \rho_{1,3\,;\,4,...,n}, \cdots, \rho_{1,n-1\,;\,n}, \rho_{1,n}$$

they are exactly the terms occurring in Equation (B.3). Hence, the terms in the product on the right hand side of Equation (B.4) containing these partial correlations are replaced by $1 - R_{1\{2,...,n\}}$*. Equation (B.3) is invariant under permutation of* $\{2, ..., n\}$*. Remove variable 1 and edges containing 1. These are just the edges whose constrains are give above. The subvine over*

variables $\{2, ..., n\}$ *are obtained. The variable 2 is in the conditioned set of the top edge of the subvine. The same argument re-indexing* $\{3, ..., n\}$ *are applied. With the re-indexing, the produce of terms in Equation (B.4) are replaced as following:*

$$(1 - \rho^2_{2,3\,;\,4,...,n})(1 - \rho^2_{2,4\,;\,5,...,n})...(1 - \rho^2_{2,n})$$

by $1 - R_{2\{3,...,n\}}$. *Finally, Equation (B.4) can be obtained by proceeding in this way.*

Appendix C

Model Selection Criteria

To select appropriate model, researchers have to balance sensitivity, which indicates having enough parameters to adequately model the relationships among variables in the population, and specificity, which suggests not overfitting a mode or suggesting nonexistent relationships. The typical, most common used model selection criterion is log-likelihood functions with simple penalties. They consist of Akaike's Information Criterion (AIC) (Akaike 1992), the Bayesian Information Criterion (BIC) (Schwarz et al. 1978) and Bozdogan consistent AIC (CAIC) (Bozdogan 1987). These criteria include a goodness-of-fit term plus a penalty to control overfitting and provide a standardized way to balance sensitivity and specificity. They can be expressed by the following equation;

$$-2L + A_n p$$

where L is the log-likelihood, p is the number of parameters in the model, and A_n is some constant or some function of the sample size n. The A_n of AIC, BIC and CAIC are 2, $log(n)$ and $log(n+1)$ respectively. The detail and comparison of these criterion refer to (Dziak, Coffman, Lanza & Li 2012).

Appendix D

List of Symbols

The following list is neither exhaustive nor exclusive, but may be helpful.

$P_{i,t}$	The price of asset i at time t
$r_{i,t}$	The return of asset i at time t
u	uniform data between 0 and 1
C	The copula function
c	The copula density
cdf	The cumulative distribution function
pdf	The probability distribution function
$\boldsymbol{\theta}^m$	The set of parameters of marginal distributions
$\boldsymbol{\theta}^c$	The set of parameters of copula function
Φ	The standard normal cumulative distribution function
Φ^{-1}	The inverse of standard normal cumulative distribution function
$\rho_{1,2;3,\ldots,n}$	The partial correlation between variable 1 and variable 2 given variable 3,...,n

$C(e)$ or C_e	The conditioned set for an edge e
$D(e)$ or D_e	The conditioning set for an edge e
CV	The constraint set for vine
U_e	The complete union for edge e
$V^C_{\rho_e}$	The partial cationical vine tree structure
$B(V^C_{\rho_e})$	The corresponding bivariate copulas for $V^C_{\rho_e}$
$\theta(B(V^C_{\rho_e}))$	The corresponding parameters of bivariate copulas for $V^C_{\rho_e}$
$V^R_{\rho_e}$	The partial regular vine tree structure
$B(V^R_{\rho_e})$	The corresponding bivariate copulas for $V^R_{\rho_e}$
$\theta(B(V^R_{\rho_e}))$	The corresponding parameters of bivariate copulas for $V^R_{\rho_e}$
AIC	Akaike's Information Criterion
$ARMA$	Autoregressive Moving Average
BIC	Schwarz's Bayesian Information Criterion
$GARCH$	Generalized Autoregressive Conditional Heteroscedasticity
IFM	Inference For the Margins
IID	Independence and Identical Distribution
MLE	Maximum Likelihood Estimator
VaR	Value at Risk
Z	The standardised residuals

Bibliography

Aas, K. & Berg, D. (2009), 'Models for construction of multivariate dependence–a comparison study', *The European Journal of Finance* **15**(7-8), 639–659.

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

Abdous, B., Genest, C. & Rémillard, B. (2005), Dependence properties of meta-elliptical distributions, *in* 'Statistical modeling and analysis for complex data problems', Springer, pp. 1–15.

Akaike, H. (1992), Information theory and an extension of the maximum likelihood principle, *in* 'Breakthroughs in statistics', Springer, pp. 610–624.

Alexander, C. (2009), *Market Risk Analysis, Value at Risk Models*, Vol. 4, John Wiley & Sons.

Almeida, C. & Czado, C. (2012), 'Efficient bayesian inference for stochastic time-varying copula models', *Computational Statistics & Data Analysis* **56**(6), 1511–1527.

Andersen, T. G. (2009), *Handbook of financial time series*, Springer.

Andersen, T. G., Bollerslev, T., Christoffersen, P. F. & Diebold, F. X. (2006), 'Volatility and correlation forecasting', *Handbook of economic forecasting* **1**, 777–878.

Andersson, J. (2001), 'On the normal inverse gaussian stochastic volatility model', *Journal of Business & Economic Statistics* **19**(1), 44–54.

Andrews, D. W. (1993), 'Tests for parameter instability and structural change with unknown change point', *Econometrica: Journal of the Econometric Society* pp. 821–856.

Ang, A. & Bekaert, G. (2002), 'International asset allocation with regime shifts', *Review of Financial studies* **15**(4), 1137–1187.

Ang, A. & Chen, J. (2002), 'Asymmetric correlations of equity portfolios', *Journal of Financial Economics* **63**(3), 443–494.

Arakelian, V. & Dellaportas, P. (2009), Contagion tests via copula threshold models, Technical report, Department of Statistics, Athens University of Economics and Business.

Arrow, K. (1971), 'Essays in the theory of risk-bearing', *Chicago: Markham Publishing* pp. 90–120.

Ausin, M. C. & Lopes, H. F. (2010), 'Time-varying joint distribution through copulas', *Computational Statistics & Data Analysis* **54**(11), 2383–2399.

Baba, K., Shibata, R. & Sibuya, M. (2004), 'Partial correlation and conditional correlation as measures of conditional independence', *Australian & New Zealand Journal of Statistics* **46**(4), 657–664.

Bae, K.-H., Karolyi, G. A. & Stulz, R. M. (2003), 'A new approach to measuring financial contagion', *Review of Financial studies* **16**(3), 717–763.

Bartram, S. M., Taylor, S. J. & Wang, Y.-H. (2007), 'The euro and european financial market dependence', *Journal of Banking & Finance* **31**(5), 1461–1481.

Bauer, A., Czado, C. & Klein, T. (2012), 'Pair-copula constructions for non-gaussian dag models', *Canadian Journal of Statistics* **40**(1), 86–109.

Bedford, T. & Cooke, R. M. (2002), 'Vines–a new graphical model for dependent random variables', *The Annals of Statistics* **30**(4), 1031–1068.

Bennett, M. N. & Kennedy, J. E. (2004), 'Quanto pricing with copulas', *The Journal of Derivatives* **12**(1), 26–45.

Bollerslev, T. (1986), 'Generalized autoregressive conditional heteroskedasticity', *Journal of econometrics* **31**(3), 307–327.

Bollerslev, T. (1987), 'A conditionally heteroskedastic time series model for speculative prices and rates of return', *The review of economics and statistics* pp. 542–547.

Bonhomme, S. & Robin, J.-M. (2009), 'Assessing the equalizing force of mobility using short panels: France, 1990–2000', *The Review of Economic Studies* **76**(1), 63–92.

Bouyé, E. & Salmon, M. (2009), 'Dynamic copula quantile regressions and tail area dynamic dependence in forex markets', *The European Journal of Finance* **15**(7-8), 721–750.

Box, G. E. & Pierce, D. A. (1970), 'Distribution of residual autocorrelations in autoregressive-integrated moving average time series models', *Journal of the American Statistical Association* **65**(332), 1509–1526.

Bozdogan, H. (1987), 'Model selection and akaike's information criterion (aic): The general theory and its analytical extensions', *Psychometrika* **52**(3), 345–370.

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

Brendstrup, B. & Paarsch, H. J. (2007), 'Semiparametric identification and estimation in multi-object, english auctions', *Journal of Econometrics* **141**(1), 84–108.

Breymann, W., Dias, A. & Embrechts, P. (2003), 'Dependence structures for multivariate high-frequency data in finance', *Quantitative Finance* **3**(1), 1–14.

Brockwell, P. J. & Davis, R. A. (2002), *Introduction to time series and forecasting*, Vol. 1, Taylor & Francis US.

Brockwell, P. J. & Davis, R. A. (2009), *Time series: theory and methods*, Springer.

Busetti, F. & Harvey, A. (2011), 'When is a copula constant? a test for changing relationships', *Journal of Financial Econometrics* **9**(1), 106–131.

Cambanis, S., Huang, S. & Simons, G. (1981), 'On the theory of elliptically contoured distributions', *Journal of Multivariate Analysis* **11**(3), 368–385.

Campbell, R., Koedijk, K. & Kofman, P. (2002), 'Increased correlation in bear markets', *Financial Analysts Journal* pp. 87–94.

Cao, L., Ou, Y., Yu, P. S. & Wei, G. (2010), Detecting abnormal coupled sequences and sequence changes in group-based manipulative trading behaviors, *in* 'Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining', ACM, pp. 85–94.

Chan, N.-H., Chen, J., Chen, X., Fan, Y. & Peng, L. (2009), 'Statistical inference for multivariate residual copula of garch models', *Statistica Sinica* **19**(1), 53.

Chazottes, J.-R., Collet, P., Külske, C. & Redig, F. (2007), 'Concentration inequalities for random fields via coupling', *Probability Theory and Related Fields* **137**(1-2), 201–225.

Chen, X. & Fan, Y. (2006a), 'Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification', *Journal of econometrics* **135**(1), 125–154.

Chen, X. & Fan, Y. (2006b), 'Estimation of copula-based semiparametric time series models', *Journal of Econometrics* **130**(2), 307–335.

Chen, X., Fan, Y. & Tsyrennikov, V. (2006), 'Efficient estimation of semiparametric multivariate copula models', *Journal of the American Statistical Association* **101**(475), 1228–1240.

Chen, Y.-T. (2007), 'Moment-based copula tests for financial returns', *Journal of Business & Economic Statistics* **25**(4), 377–397.

Chen, Y.-T. (2011), 'Moment tests for density forecast evaluation in the presence of parameter estimation uncertainty', *Journal of Forecasting* **30**(4), 409–450.

Cherubini, U. & Luciano, E. (2001), 'Value-at-risk trade-off and capital allocation with copulas', *Economic notes* **30**(2), 235–256.

Cherubini, U., Luciano, E. & Vecchiato, W. (2004), *Copula methods in finance*, John Wiley & Sons.

Chollete, L., Heinen, A. & Valdesogo, A. (2009), 'Modeling international financial returns with a multivariate regime-switching copula', *Journal of financial econometrics* **7**(4), 437–480.

Christoffersen, P., Errunza, V., Jacobs, K. & Langlois, H. (2012), 'Is the potential for international diversification disappearing? a dynamic copula approach', *Review of Financial Studies* **25**(12), 3711–3751.

Christoffersen, P. F. (1998), 'Evaluating interval forecasts', *International economic review* pp. 841–862.

Christoffersen, P. & Langlois, H. (2011), 'The joint dynamics of equity market factors', *Risk Management eJournal* **3**, 128.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., Stein, C. et al. (2001), *Introduction to algorithms*, Vol. 2, MIT Press Cambridge.

Creal, D., Koopman, S. J. & Lucas, A. (2013), 'Generalized autoregressive score models with applications', *Journal of Applied Econometrics* **28**(5), 777–795.

Czado, C., Schepsmeier, U. & Min, A. (2012), 'Maximum likelihood estimation of mixed c-vines with application to exchange rates', *Statistical Modelling* **12**(3), 229–255.

Darsow, W. F., Nguyen, B., Olsen, E. T. et al. (1992), 'Copulas and markov processes', *Illinois Journal of Mathematics* **36**(4), 600–642.

De Raedt, L. & Kersting, K. (2008), *Probabilistic inductive logic programming*, Springer.

Dearden, L., Fitzsimons, E., Goodman, A. & Kaplan, G. (2008), 'Higher education funding reforms in england: The distributional effects and the shifting balance of costs', *The Economic Journal* **118**(526), 100–125.

Demarta, S. & McNeil, A. J. (2005), 'The t copula and related copulas', *International Statistical Review* **73**(1), 111–129.

Dias, A. & Embrechts, P. (2010), 'Modeling exchange rate dependence dynamics at different time horizons', *Journal of International Money and Finance* **29**(8), 1687–1705.

Diebold, F. X., Hahn, J. & Tay, A. S. (1999), 'Multivariate density forecast evaluation and calibration in financial risk management: high-

frequency returns on foreign exchange', *Review of Economics and Statistics* **81**(4), 661–673.

Diestel, R. (2000), *Graph Theory*, Springer-Verlag Berlin and Heidelberg GmbH & Company KG.

Diks, C., Panchenko, V. & Van Dijk, D. (2010), 'Out-of-sample comparison of copula specifications in multivariate density forecasts', *Journal of Economic Dynamics and Control* **34**(9), 1596–1609.

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

Dos Santos Silva, R. & Lopes, H. F. (2008), 'Copula, marginal distributions and model selection: a bayesian note', *Statistics and Computing* **18**(3), 313–320.

Dudley, E. & Nimalendran, M. (2011), 'Margins and hedge fund contagion', *Journal of Financial and Quantitative Analysis* **46**(05), 1227–1257.

Dziak, J. J., Coffman, D. L., Lanza, S. T. & Li, R. (2012), Sensitivity and specificity of information criteria, Technical report, The Methodology Center and Department of Statistics, Penn State, The Pennsylvania State University.

Elidan, G. (2010), Copula bayesian networks., *in* 'Advances in Neural Information Processing Systems', pp. 559–567.

Elidan, G. (2013), 'Copulas in machine learning', *Copulae in Mathematical and Quantitative Finance* pp. 39–60.

Embrechts, P. & Höing, A. (2006), 'Extreme var scenarios in higher dimensions', *Extremes* **9**(3-4), 177–192.

Embrechts, P., Höing, A. & Juri, A. (2003), 'Using copulae to bound the value-at-risk for functions of dependent risks', *Finance and Stochastics* **7**(2), 145–167.

Embrechts, P., Klüppelberg, C. & Mikosch, T. (1997), *Modelling extremal events: for insurance and finance*, Vol. 33, Springer.

Embrechts, P., Lindskog, F. & McNeil, A. (2003), 'Modelling dependence with copulas and applications to risk management', *Handbook of heavy tailed distributions in finance* **8**(1), 329–384.

Engle, R. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation', *Econometrica: Journal of the Econometric Society* pp. 987–1007.

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

Engle, R. (2009), *Anticipating correlations: a new paradigm for risk management*, Princeton University Press.

Erb, C. B., Harvey, C. R. & Viskanta, T. E. (1994), 'Forecasting international equity correlations', *Financial analysts journal* pp. 32–45.

Fang, H.-B., Fang, K.-T. & Kotz, S. (2002), 'The meta-elliptical distributions with given marginals', *Journal of Multivariate Analysis* **82**(1), 1–16.

Fang, K.-T., Kotz, S. & Ng, K. W. (1990), *Symmetric Multivariate and Related Distributions Monographs on Statistics and Applied Probability*, Vol. 1, Chapman and Hall London.

Fantazzini, D. (2009), 'The effects of misspecified marginals and copulas on computing the value at risk: A monte carlo study', *Computational Statistics & Data Analysis* **53**(6), 2168–2188.

Fischer, M., Köck, C., Schlüter, S. & Weigert, F. (2009), 'An empirical analysis of multivariate copula models', *Quantitative Finance* **9**(7), 839–854.

Frahm, G., Junker, M. & Schmidt, R. (2005), 'Estimating the tail-dependence coefficient: Properties and pitfalls', *Insurance: Mathematics and Economics* **37**(1), 80–100.

Frahm, G., Junker, M. & Szimayer, A. (2003), 'Elliptical copulas: applicability and limitations', *Statistics & Probability Letters* **63**(3), 275–286.

Frey, R., McNeil, A. J., McNeil, A. J. & McNeil, A. J. (2001), Modelling dependent defaults: Asset correlations are not enough!, Technical report, ETH, Eidgenössische Technische Hochschule Zürich, Department of Mathematics.

Gaißer, S., Ruppert, M. & Schmid, F. (2010), 'A multivariate version of hoeffdings phi-square', *Journal of Multivariate Analysis* **101**(10), 2571–2586.

Gao, J. & Suzuki, H. (2003), Unsupervised learning of dependency structure for language modeling, *in* 'Proceedings of the 41st Annual Meeting on Association for Computational Linguistics-Volume 1', Association for Computational Linguistics, pp. 521–528.

Garcia, R. & Tsafack, G. (2011), 'Dependence structure and extreme comovements in international equity and bond markets', *Journal of Banking & Finance* **35**(8), 1954–1970.

Genest, C., Ghoudi, K. & Rivest, L.-P. (1995), 'A semiparametric estimation procedure of dependence parameters in multivariate families of distributions', *Biometrika* **82**(3), 543–552.

Genest, C., Nešlehová, J. & Ben Ghorbal, N. (2011), 'Estimators based on kendall's tau in multivariate copula models', *Australian & New Zealand Journal of Statistics* **53**(2), 157–177.

Genest, C., Nešlehová, J. & Ziegel, J. (2011), 'Inference in multivariate archimedean copula models', *TEST* **20**(2), 223–256.

Genest, C. & Rémillard, B. (2008), Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models, *in* 'Annales de lInstitut Henri Poincaré: Probabilités et Statistiques', Vol. 44, pp. 1096–1127.

Genest, C. & Rivest, L.-P. (1993), 'Statistical inference procedures for bivariate archimedean copulas', *Journal of the American statistical Association* **88**(423), 1034–1043.

Genest, C. & Segers, J. (2009), 'Rank-based inference for bivariate extreme-value copulas', *The Annals of Statistics* **37**(5B), 2990–3022.

Getoor, L. & Taskar, B. (2007), *Introduction to statistical relational learning*, MIT Press.

Ghoudi, K. & Rémillard, B. (2004), 'Empirical processes based on pseudo-observations. ii. the multivariate case', *Asymptotic methods in stochastics* **44**, 381–406.

Giacomini, E., Härdle, W. & Spokoiny, V. (2009), 'Inhomogeneous dependence modeling with time-varying copulae', *Journal of Business & Economic Statistics* **27**(2).

Giacomini, R. & White, H. (2006), 'Tests of conditional predictive ability', *Econometrica* **74**(6), 1545–1578.

Giesecke, K. (2004), 'Correlated default with incomplete information', *Journal of Banking & Finance* **28**(7), 1521–1545.

Goorbergh, R., Genest, C. & Werker, B. (2003), Multivariate option pricing using dynamic copula models, Technical report, Tilburg University, Center for Economic Research.

Granger, C. W., Teräsvirta, T. & Patton, A. J. (2006), 'Common factors in conditional distributions for bivariate time series', *Journal of Econometrics* **132**(1), 43–57.

Green, P. J. (1995), 'Reversible jump markov chain monte carlo computation and bayesian model determination', *Biometrika* **82**(4), 711–732.

Guegan, D. & Zhang, J. (2010), 'Change analysis of a dynamic copula for measuring dependence in multivariate financial data', *Quantitative Finance* **10**(4), 421–430.

Guermat, C. & Harris, R. D. (2002), 'Robust conditional variance estimation and value-at-risk', *Journal of Risk* **4**, 25–42.

Hafner, C. M. & Manner, H. (2012), 'Dynamic stochastic copula models: Estimation, inference and applications', *Journal of Applied Econometrics* **27**(2), 269–295.

Hafner, C. M. & Reznikova, O. (2010), 'Efficient estimation of a semiparametric dynamic copula model', *Computational Statistics & Data Analysis* **54**(11), 2609–2627.

Hanea, A., Kurowicka, D., Cooke, R. M. & Ababei, D. (2010), 'Mining and visualising ordinal data with non-parametric continuous bbns', *Computational Statistics & Data Analysis* **54**(3), 668–687.

Hansen, B. E. (1994), 'Autoregressive conditional density estimation', *International Economic Review* pp. 705–730.

Harris, J., Hirst, J. L. & Mossinghoff, M. J. (2008), *Combinatorics and graph theory*, Springer.

Harvey, A. (2010), 'Tracking a changing copula', *Journal of Empirical Finance* **17**(3), 485–500.

He, Y., Qi, Y., Kavukcuoglu, K. & Park, H. (2012), Learning the dependency structure of latent factors, *in* 'Advances in Neural Information Processing Systems', pp. 2375–2383.

Heinen, A. & Rengifo, E. (2007), 'Multivariate autoregressive modeling of time series count data using copulas', *Journal of Empirical Finance* **14**(4), 564–583.

Heinen, A. & Valdesogo Robles, A. (2009), Asymmetric capm dependence for large dimensions: the canonical vine autoregressive model, Technical report, Université catholique de Louvain, Center for Operations Research and Econometrics (CORE).

Hobæk Haff, I. (2012), 'Comparison of estimators for pair-copula constructions', *Journal of Multivariate Analysis* **110**, 91–105.

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

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

Hong, Y., Tu, J. & Zhou, G. (2007), 'Asymmetries in stock returns: Statistical tests and economic evaluation', *Review of Financial Studies* **20**(5), 1547–1581.

Hu, L. (2006), 'Dependence patterns across financial markets: a mixed copula approach', *Applied financial economics* **16**(10), 717–729.

Hull, J. C. & White, A. D. (1998), 'Value at risk when daily changes in market variables are not normally distributed', *The Journal of Derivatives* **5**(3), 9–19.

Hurd, M., Salmon, M. & Schleicher, C. (2005), Using copulas to construct bivariate foreign exchange distributions with an application to the ster-

ling exchange rate index, Technical report, Centre for Economic Policy Researchs (CEPR).

Ibragimov, R. (2009), 'Copula-based characterizations for higher order markov processes', *Econometric Theory* **25**(03), 819–846.

Joe, H. (1990), 'Multivariate concordance', *Journal of multivariate analysis* **35**(1), 12–30.

Joe, H. (1996), 'Families of m-variate distributions with given margins and m (m-1)/2 bivariate dependence parameters', *Lecture Notes-Monograph Series* pp. 120–141.

Joe, H. (1997), *Multivariate models and dependence concepts*, Vol. 73, Springer.

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

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

Joe, H. & Xu, J. J. (1996), The estimation method of inference functions for margins for multivariate models, Technical report, University of British Columbia.

Jondeau, E. & Rockinger, M. (2006*a*), 'The copula-garch model of conditional dependencies: An international stock market application', *Journal of international money and finance* **25**(5), 827–853.

Jondeau, E. & Rockinger, M. (2006*b*), 'Optimal portfolio allocation under higher moments', *European Financial Management* **12**(1), 29–55.

Kang, B. U., In, F., Kim, G. & Kim, T. S. (2010), 'A longer look at the asymmetric dependence between hedge funds and the equity market', *Journal of Financial and Quantitative Analysis* **45**(3), 763.

Kendall, M. G. & Smith, B. B. (1940), 'On the method of paired comparisons', *Biometrika* **31**(3/4), 324–345.

Kendall, M. G. & Stuart, A. (1961), *The advanced theory of statistics*, Charles Griffin.

Kim, S., Shephard, N. & Chib, S. (1998), 'Stochastic volatility: likelihood inference and comparison with arch models', *The Review of Economic Studies* **65**(3), 361–393.

Kirshner, S. (2007), Learning with tree-averaged densities and distributions, *in* 'NIPS', pp. 761–768.

Klaassen, C. A. & Wellner, J. A. (1997), 'Efficient estimation in the bivariate normal copula model: normal margins are least favourable', *Bernoulli* **3**(1), 55–77.

Kojadinovic, I. & Yan, J. (2011), 'A goodness-of-fit test for multivariate multiparameter copulas based on multiplier central limit theorems', *Statistics and Computing* **21**(1), 17–30.

Kojadinovic, I., Yan, J. & Holmes, M. (2011), 'Fast large-sample goodness-of-fit tests for copulas', *Statistica Sinica* **21**(2), 841.

Kok, S. & Domingos, P. (2007), Statistical predicate invention, *in* 'Proceedings of the 24th international conference on Machine learning', ACM, pp. 433–440.

Kontorovich, L. A., Ramanan, K. et al. (2008), 'Concentration inequalities for dependent random variables via the martingale method', *The Annals of Probability* **36**(6), 2126–2158.

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

Krzysztofowicz, R. & Kelly, K. (1996), A meta-gaussian distribution with specified marginals, Technical report, University of Virginia.

Kupiec, P. H. (1995), 'Techniques for verifying the accuracy of risk measurement models', *The Journal of Derivatives* **3**(2), 73–84.

Kurowicka, D. & Cooke, R. (2005), 'Distribution-free continuous bayesian belief nets', *Modern Statistical and Mathematical Methods in Reliability* **10**, 309–323.

Kurowicka, D. & Cooke, R. (2006*a*), 'Completion problem with partial correlation vines', *Linear algebra and its applications* **418**(1), 188–200.

Kurowicka, D. & Cooke, R. (2006*b*), 'High-dimensional dependence modelling', *Uncertainty Analysis with High Dimensional Dependence Modelling* pp. 81–130.

Kurowicka, D. & Cooke, R. M. (2006*c*), *Uncertainty analysis with high dimensional dependence modelling*, Wiley.

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

Lawrance, A. (1976), 'On conditional and partial correlation', *The American Statistician* **30**(3), 146–149.

Lee, T.-H. & Long, X. (2009), 'Copula-based multivariate garch model with uncorrelated dependent errors', *Journal of Econometrics* **150**(2), 207–218.

Leiserson, C. E., Rivest, R. L., Stein, C. & Cormen, T. H. (2001), *Introduction to algorithms*, The MIT Press.

Li, D. X. (2000), 'On default correlation: a copula function approach', *The Journal of Fixed Income* **9**(4), 43–54.

Liu, H., Lafferty, J. & Wasserman, L. (2009), 'The nonparanormal: Semiparametric estimation of high dimensional undirected graphs', *The Journal of Machine Learning Research* **10**, 2295–2328.

Ljung, G. M. & Box, G. E. (1978), 'On a measure of lack of fit in time series models', *Biometrika* **65**(2), 297–303.

Longin, F. & Solnik, B. (2001), 'Extreme correlation of international equity markets', *The Journal of Finance* **56**(2), 649–676.

Luciano, E. & Cherubini, U. (2000), Multivariate option pricing with copulas, *in* 'EFMA 2001 Lugano Meetings'.

Manner, H. (2007), Estimation and model selection of copulas with an application to exchange rates, Technical report, Maastricht University, Maastricht Research School of Economics of Technology and Organization (METEOR).

Manner, H. & Segers, J. (2011), 'Tails of correlation mixtures of elliptical copulas', *Insurance: Mathematics and Economics* **48**(1), 153–160.

Mari, D. D. & Kotz, S. (2001), *Correlation and dependence*, Vol. 2, World Scientific.

Markowitz, H. (2012), 'Portfolio selection', *The journal of finance* **7**(1), 77–91.

Markwat, T., Kole, H. & Dijk, D. v. (2009), Time variation in asset return dependence: Strength or structure?, Technical report, Erasmus Research Institute of Management (ERIM).

McLeod, A. I. & Li, W. K. (1983), 'Diagnostic checking arma time series models using squared-residual autocorrelations', *Journal of Time Series Analysis* **4**(4), 269–273.

McNeil, A. J. (2008), 'Sampling nested archimedean copulas', *Journal of Statistical Computation and Simulation* **78**(6), 567–581.

McNeil, A. J., Frey, R. & Embrechts, P. (2010), *Quantitative risk management: concepts, techniques, and tools*, Princeton university press.

McNeil, A. J. & Nešlehová, J. (2009), 'Multivariate archimedean copulas, d-monotone functions and l_1-norm symmetric distributions', *The Annals of Statistics* pp. 3059–3097.

Milch, B. & Russell, S. (2007), First-order probabilistic languages: Into the unknown, *in* 'Inductive Logic Programming', Springer, pp. 10–24.

Mills, F. C. (1927), *The Behavior of Prices*, National Bureau of Economic Research, Inc.

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

Nelsen, R. B. (1999), *An introduction to copulas*, Springer.

Neville, J. & Jensen, D. (2007), 'Relational dependency networks', *The Journal of Machine Learning Research* **8**, 653–692.

Nikoloulopoulos, A. K., Joe, H. & Li, H. (2009), 'Extreme value properties of multivariate t copulas', *Extremes* **12**(2), 129–148.

Oh, D. H. & Patton, A. J. (2013), 'Simulated method of moments estimation for copula-based multivariate models', *Journal of the American Statistical Association* **108**(502), 689–700.

Okimoto, T. (2008), 'New evidence of asymmetric dependence structures in international equity markets', *Journal of Financial and Quantitative Analysis* **43**(03), 787–815.

Owen, J. & Rabinovitch, R. (1983), 'On the class of elliptical distributions and their applications to the theory of portfolio choice', *The Journal of Finance* **38**(3), 745–752.

Panagiotelis, A., Czado, C. & Joe, H. (2012), 'Pair copula constructions for multivariate discrete data', *Journal of the American Statistical Association* **107**(499), 1063–1072.

Patton, A. J. (2004), 'On the out-of-sample importance of skewness and asymmetric dependence for asset allocation', *Journal of Financial Econometrics* **2**(1), 130–168.

Patton, A. J. (2006a), 'Modelling asymmetric exchange rate dependence', *International economic review* **47**(2), 527–556.

Patton, A. J. (2006b), 'Estimation of multivariate models for time series of possibly different lengths', *Journal of applied econometrics* **21**(2), 147–173.

Patton, A. J. (2009), 'Are market neutral?hedge funds really market neutral?', *Review of Financial Studies* **22**(7), 2495–2530.

Pearl, J. (1988), *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Publishers Inc.

Pickands, J. (1981), Multivariate extreme value distributions, *in* 'Proceedings 43rd Session International Statistical Institute', Vol. 2, pp. 859–878.

Pitt, M., Chan, D. & Kohn, R. (2006), 'Efficient bayesian inference for gaussian copula regression models', *Biometrika* **93**(3), 537–554.

Rémillard, B. (2010), 'Goodness-of-fit tests for copulas of multivariate time series', *Social Science Research Network* .

Riccetti, L. (2012), 'A copula–garch model for macro asset allocation of a portfolio with commodities', *Empirical Economics* pp. 1–22.

Richardson, M. & Domingos, P. (2006), 'Markov logic networks', *Machine learning* **62**(1-2), 107–136.

Rivers, D. & Vuong, Q. (2002), 'Model selection tests for nonlinear dynamic models', *The Econometrics Journal* **5**(1), 1–39.

Rodriguez, J. C. (2007), 'Measuring financial contagion: A copula approach', *Journal of Empirical Finance* **14**(3), 401–423.

Rosenberg, J. V. (2003), 'Non-parametric pricing of multivariate contingent claims', *The Journal of Derivatives* **10**(3), 9–26.

Rosenberg, J. V. & Schuermann, T. (2006), 'A general approach to integrated risk management with skewed, fat-tailed risks', *Journal of Financial economics* **79**(3), 569–614.

Salmon, M. & Schleicher, C. (2006), Pricing multivariate currency options with copulas, Technical report, Warwick Business School, Finance Group.

Sancetta, A. & Satchell, S. (2004), 'The bernstein copula and its applications to modeling and approximations of multivariate distributions', *Econometric Theory* **20**(3), 535–562.

Savu, C. & Trede, M. (2006), Hierarchical archimedean copulas, *in* 'International conference on high frequency finance, Konstanz, Germany, May'.

Savu, C. & Trede, M. (2010), 'Hierarchies of archimedean copulas', *Quantitative Finance* **10**(3), 295–304.

Scaillet, O. & Fermanian, J.-D. (2002), 'Nonparametric estimation of copulas for time series', *International Center for Financial Asset Management and Engineering Research Paper* (57).

Schönbucher, P. J. & Schubert, D. (2001), Copula-dependent default risk in intensity models, Technical report, Working paper, Department of Statistics, Bonn University.

Schwarz, G. et al. (1978), 'Estimating the dimension of a model', *The annals of statistics* **6**(2), 461–464.

Shih, J. H. & Louis, T. A. (1995), 'Inferences on the association parameter in copula models for bivariate survival data', *Biometrics* pp. 1384–1399.

Silva, R. & Gramacy, R. B. (2009), Mcmc methods for bayesian mixtures of copulas, *in* 'International Conference on Artificial Intelligence and Statistics', pp. 512–519.

Sklar, M. (1959), *Fonctions de répartition à n dimensions et leurs marges*, Université Paris 8.

Sun, W., Rachev, S., Stoyanov, S. V. & Fabozzi, F. J. (2008), 'Multivariate skewed student's t copula in the analysis of nonlinear and asymmetric dependence in the german equity market', *Studies in Nonlinear Dynamics & Econometrics* **12**(2).

Taylor, S. J. & Wang, Y.-h. (2004), Option prices and risk neutral densities for currency cross-rates, Technical report, Mimeo, Department of Accounting and Finance, Lancaster University.

Tse, Y. K. & Tsui, A. K. (2002), 'A multivariate garch model with time-varying correlations', *Journal of Business and Economic Statistics* **20**, 351–362.

Vuong, Q. H. (1989), 'Likelihood ratio tests for model selection and non-nested hypotheses', *Econometrica: Journal of the Econometric Society* pp. 307–333.

West, K. D. (2006), 'Forecast evaluation', *Handbook of economic forecasting* **1**, 99–134.

Whelan, N. (2004), 'Sampling from archimedean copulas', *Quantitative finance* **4**(3), 339–352.

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

Zimmer, D. M. (2012), 'The role of copulas in the housing crisis', *Review of Economics and Statistics* **94**(2), 607–620.