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

Wei Wei * Junfu Yin * Jinyan Li * Longbing Cao *

Abstract
Modeling high-dimensional dependence is widely studied to explore deep relations in multiple variables particularly useful for financial risk assessment. 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 asymmetry, upper and lower tail dependence between multiple variables. The paper proposes a partial regular vine copula model to relax these restrictions. The new model employs partial correlation to construct the regular vine structure, which is algebraically independent. This model is also able to capture the asymmetric characteristics among multiple variables by using two-parametric copula with flexible lower and upper tail dependence. Our 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 financial market.

1 Introduction

Learning dependeces 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 concern on the degree of dependence, however, few of them focus on the another important respect of dependence– the dependence structure, especially the asymmetric and tail dependence characteristics. Dependence structure studying plays a important role in financial area, especially in portfolio investment theory. Typical models of investment theory impose strong restriction on the dependence structure, which did not consider the asymmetric characteristics. It results in these typical models 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 2008 global financial crisis. It indicates that stock returns have stronger dependence during bear market than bull market, which results in 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 allow dependence and correlation to vary over time. It uses correlation/conditional correlation to capture the natural of dependence, and at the same time, it can build flexible structure to model complex high-dimensional dependencies structure. In order to model the asymmetric dependence with high-dimensional financial variables, it is essential to develop flexible dependence model with parametric copula families, which is suitable to multivariate data with various dependence structures. Hence, the model should have desired properties, which are described as follow:

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

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

(iii). Flexible rang 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 above conditions. Typically, multivariate Archimedean copula model has the structure with only narrow range of negative dependence [19]. The multivariate Gaussian copula model is not suitable to model the asymmetric characteristics, since (1) Gaussian copula does not have lower and upper tail dependence, and (2) the Gaussian assumption are not appropriate in the real world [2, 9]. The multivariate t copula model, which is studied by [8, 20], does not have flexible lower and upper tail dependence since t copula has same lower and upper tail dependence. Canonical vine or D vine copula model, such as [1, 4, 22], have wide range of dependence by choosing appropriate bivariate copula families. However, they does not have flexible dependence structure due to their structure assumptions. These assumptions imposed on dependence structure lead to their dependence structure may not reflect the actual dependence in high-dimensional data.

In order to fulfill the above needs, we propose a new partial correlation-based regular vine copula model with asymmetric dependence. 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

*Advanced Analytics Institute, University of Technology, Sydney. timajia@gmail.com, junfu.yin@gmail.com, jinyan.li@uts.edu.au, longbing.cao@uts.edu.au

the dependence structure. Hence, it can reflect the actual dependence structure of high-dimensional data. In addition, copula family with flexible lower and upper tail dependence are connecting with our new models, which ensure the new model have a wide range of lower and upper tail dependence.

The paper have these contributions: (1) The paper 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 independently from the established tree structure and bivariate copulas selection. Hence, our model did not impose any strong restriction on the dependence structure; (2) For all linking bivariate copula on the partial regular vine dependence structure, we use only the BB1, survival BB1, BB7 and survival BB7, 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 analyze the moving trends of lower and upper tail dependence with the high-dimensional data structure. In addition, we also analyze the trends of lower and upper tail dependence during dynamic period.

The rest of paper is structured as follow. Section 2 introduces the definition of copula and its related tail dependence. The regular vine theorem and related definitions are introduced in Section 3. Section 4 discuss how to construct the partial regular vine copula model and copula family selection. Section 5 discusses how to estimate parameters in partial regular vine copula and marginal distribution. Value at Risk, which is the popular evaluation method used in financial market, and its related tests are discussed in Section 7. Finally, Section 8 conclude the paper.

2 Introducation to Copula and Tail Dependence

In the Section, we introduce the definitions of copula and its tail dependence. Due to the theorem in [21], an n-dimensional joint distribution can be decomposed into its n univariate marginal distributions and an n-dimensional copula function:

(2.1)
$$\begin{aligned} Let \quad \mathbf{x} \quad &\equiv \quad [x_1, ..., x_n]' \sim F, \quad with \quad x_i \sim F_i \\ then \quad \exists\, C \quad &: \quad [0,1]^n \to [0,1] \\ s.t. \quad F(y) \quad &= \quad C(F_1(x_1), ..., F_n(x_n)) \quad \forall \quad x \in \mathbb{R}^n \end{aligned}$$

Thus, copula C is the function that maps the univariate marginal distributions F_i to the joint distribution F. According the above representation of the joint cumulative distribution function, the representation of the joint probability distribution function is as following:

(2.2)
$$\begin{aligned} f(y_1, ..., y_n) \quad &= \quad c(F_1(y_1), ..., F_n(y_n)) \times \prod_{i=1}^{n} f_i(y_i) \\ where \quad \mathbf{c}(u_1, ..., u_n) \quad &= \quad \frac{\partial^n C(u_1, ..., u_n)}{\partial u_1, ..., \partial u_n} \end{aligned}$$

The representation is typically useful. Given any set of n univariate distribution $(F_1, ..., F_n)$ and any copula C, the function F in Equation (2.1) can give a valid joint distribution with marginal distributions $(F_1, ..., F_n)$. The ability to sperate dependence structure independently from marginal distribution allow researchers to focus on the dependence structure, ignoring the effect of marginal distribution.

One important copula-based dependence measurement is tail dependence coefficient, which indicates the dependence between extreme events. The extremal dependence of a multivariate distribution F can be described by various tail dependence parameters of its copula C. Suppose that 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 coefficients are defined as follow:

(2.3)
$$\begin{aligned} \lambda_L &= \lim_{u \to 0} Pr\{U_1 \leq u, ..., U_n \leq u \mid U_n \leq u\} \\ &= \lim_{u \to 0} \frac{C(u, ..., u)}{u} \\ \lambda_U &= \lim_{u \to 0} Pr\{U_1 > 1-u, ..., U_n > 1-u \mid U_n > 1-u\} \\ &= \lim_{u \to 0} \frac{\overline{C}(1-u, ..., 1-u)}{u} \end{aligned}$$

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 coefficient, and no upper tail dependence coefficient if $\lambda_U = 0$. Similarly, If λ_L exists and $\lambda_L \in (0, 1]$, then copula C has upper tail dependence coefficient, and no upper tail dependence coefficient if $\lambda_L = 0$. In addition, tail dependence is one of important properties that discriminate the different copula families. For example, Gumbel copula has only upper tail dependence coefficient, and Gaussian copula does not allow for any tail dependence coefficient. For simplicity, the tail dependence in the late of paper are short for tail dependence coefficient.

Frahm et al [11] proposed a non-parametric method to obtain the non-parametric estimator of lower and upper tail dependence by using 'Pickand's dependence function'. One simple nonparametric estimator of tail dependence is the log estimator. In the paper, the non-parametric method is used for roughly analyzing the tail dependence coefficient before we building the regular vine copula model.

3 Regular Vine Theory

In the Section, we introduce the regular vine theory and its related definitions. Vine theory is introduced by [3], which is one kind of graphical model. Let V, T, E, N be denoted as vine structure, trees, edges, nodes respectively. The regular vine and its related definition are given as follow:

DEFINITION 3.1. (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.*

DEFINITION 3.2. (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:*

(3.4)
$$U_{e_i}(k) = \{e \in E_{i-k} \,|\, \exists\, e_j \in E_j, j = 1-(k-1), ..., i-1 \text{ with } e \in e_{i-(k-1)} \in e_{i-(k-2)} \in ... \in e_{i-1} \in e_i\}$$

Then, the complete union of $e_i \in E_i$ is defined as

$$U_{e_i} = U_{e_i}(k) \tag{3.5}$$

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 3.3. *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*

$$D_e = U_a \cap U_b, \tag{3.6}$$

and the conditioned set (C_e) with e are

$$C_{e(a)} = U_a \backslash D_e \tag{3.7}$$

$$C_{e(b)} = U_b \backslash D_e \tag{3.8}$$

$$C_e = C_{e(a)} \cup C_{e(b)} = U_a \triangle U_b \tag{3.9}$$

The constraint set for e is

(3.10)
$$CV_e = \{(\{C_{e(a)}, C_{e(b)}\}, D_e) \,|\, i = 1, ..., n-1, e \in E_i, e = a, b\}$$

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 conditioning set D_e.

DEFINITION 3.4. (REGULAR VINE COPULA SPECIFICATION) *A regular vine copula specification on n variables is a multivariate distribution function is defined as $C = (V, B(V), \theta(B(V)))$*

(i). V is a vine structure on n variables;

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

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

Based on the definition of regular vine specification, the full specification of a regular vine copula has three components: the vine tree structure V, the copula family set $B(V)$, and the corresponding copula parameters $\theta(B(V))$. Then, there is a corresponding density distribution that realises the regular vine copula specification, which is given as follow:

(3.11)
$$f_{1:n}(\mathbf{x}|V, B, \theta) = \prod_{k=1}^{n} f_k(x_k) \cdot$$
$$\prod_{i=1}^{n-1} \prod_{e \in E_i} c_{e(a),e(b)|D_e}(F_{e(a)|D_e}(x_{e(a)}|\mathbf{x}_{D_e}), F_{e(b)|D_e}(x_{e(b)}|\mathbf{x}_{D_e}))$$

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 Regular Vine Copula Model with Asymmetric Dependence

4.1 Regular Vine Tree Structure Construction In the Section, we discuss how to build the vine tree structure V firstly, and then consider how to select the bivariate copula for all edges in vine tree structure. The regular vine tree structure is dependence structure which connects all bivariate copula together. For dependence structure with dimension < 5, the vine tree structure is either canonical vine or D vine. However, for dependence structure with high dimensional (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 impose restrictions on vine tree dependence structure, in which each variable connect to one variable in each tree. Hence, canonical vine has a star-like structure. D vine has the restriction that each variable links

Algorithm 1 Regular Vine Construction via Top-to-Bottom Strategy

Require: Observations of n variables

1: Calculate partial correlation $\rho_{x,y\,;\,I\backslash\{x,y\}}$ for all possible pair variables $\{x,y\}, x,y \in \{1,...,n\} = I$.
2: Find the Maximum Spanning Tree (MST), which can maximize the sum of absolute value of partial correlation ρ, such as: $max \sum |\rho_{x,y\,;\,I\backslash\{x,y\}}|$.
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 maximize the sum of absolute value of partial correlation, $max \sum |\rho_{e(p),e(q)\,;\,D_e}|$, where ρ is partial correlation.
7: **end for**
8: **return** Partial regular vine tree structure

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

We develop a new partial correlation-based algorithm to construct the regular vine, which is called partial regular vine. We consider to use partial correlation 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:

$$\rho_{1,2:3,...,n} = -\frac{K_{12}}{\sqrt{K_{11}K_{22}}} \tag{4.12}$$

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:

$$\rho_{1,2:3,...,n} = \frac{\rho_{1,2:3,...,n-1} - \rho_{1,n:3,...,n-1} \cdot \rho_{2,n:3,...,n-1}}{\sqrt{1-\rho^2_{1,n:3,...,n-1}}\sqrt{1-\rho^2_{2,n:3,...,n-1}}} \tag{4.13}$$

Obviously, $\rho_{1,2}$ is equal to correlation. Hence, when building the partial regular vine tree structure, the current vine tree structure is independently from the previous tree structure, and (2) As we discussed in Section 3, for copula selection, regular vine tree structure built by partial correlation can ensure that copula selection in current tree structure is independently from the copula selection result in previous tree. Hence, partial correlation-based regular vine tree structure ensure we focus on the dependence structure itself, ignoring the effect of different copula family selection.

The algorithm to construct the partial regular vine tree structure is given in Algorithm 1. Based on the partial correlation definition, the partial correlation is equal to correlation in the first tree T_1. In the paper, we use Kendall's tau τ 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, we firstly calculate partial correlation ρ for all possible pair variables. Then, we employ the Maximum Spanning Tree (MST) to find the vine tree structure in T_1. Typically, the Algorithm of Prim is used to producing a Minimum S-panning Tree [7]. 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, we can build a large number of regular vine tree structures, we choose the structure which can maximize the sum of absolute value of partial correlation ρ. Once the first tree structure is identified, we starting building following trees, which employs similar strategy 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 3. 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 3. 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 copulas. However, the theorem in [3] 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, we can construct partial regular vine structure firstly, and then map it into the conditional correlation based regular vine dependence structure. Then, we can fit the whole structure with various copulas. We can remove the limitation of partial regular vine and fit the vine tree structure with various copula, other than elliptical copula.

4.2 Bivariate Copula Family Selection Once the vine tree structure is identified, the next step is to choose appropriate bivariate copula for all edges. According to the theory in [15], 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 we select the copula with symmetric lower and upper tail dependence, then it is a reflection of symmetric dependence structure. If we want to model the asymmetric dependence, it is better to choose copula with various lower and upper tail dependence. Currently, There are a huge of copula families, which have various tail dependencies. The detail of tail dependence of copula families are listed in Table 1. For one-parametric copula, Gaussian and Frank copulas do not have any tail dependence, Clayton and Joe copulas have only lower tail dependence,

Table 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	$2^{-1/(\phi\delta)}$	$2-2^{1/\delta}$
BB6	-	$2-2^{1/(\delta\phi)}$
BB7	$2-2^{1/\delta}$	$2-2^{1/\phi}$
S.BB7	$2-2^{1/\phi}$	$2-2^{1/\delta}$
S.BB1	$2-2^{1/\delta}$	$2^{-1/(\phi\delta)}$
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, $\mu(\nu,\phi)=\left(-\sqrt{\nu+1}\sqrt{\frac{1-\phi}{1+\phi}}\right)$.

and Gumbel copula has only upper tail dependence. For two-parametric copula, 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 their have various lower and upper tail dependence, which can vary independently from 0 to 1.

5 Marginal Distribution Specification and Parameter Estimate

In the Section, we discuss the marginal distribution specification and parameter estimate. According to the Equation (2.2), the multivariate joint density function has two parts, one is multivariate copula mentioned in above section, another part is the marginal distributions. For financial data, ARMA(1,1)-GARCH(1,1) model is best choice for the marginal distribution [16] [10]. Typically, let $X_t(t=0,1,...,Z)$ be a time series of the price on a financial asset, such as stock market index. Then the return of financial asset can be defined as $log(X_t/X_{t-1})$. Suppose there are n assets with returns $r_{t,1},...,r_{t,n}$. The estimation of partial regular vine copula model can be proceed in two steps. In the first step, we select the appropriate marginal distribution of variables (i.e.financial asset), which is univariate distribution. Due to the character of financial assets, such as volatility cluster, a common choice is ARMA(1,1)-GARCH(1,1) with Student-t innovations, which is defined as follow:

(5.14)
$$\begin{aligned} r_{t,j} &= c_j+\varphi_j r_{t-1,j}+\gamma_j\varepsilon_{t-1,j}+\varepsilon_{t,j},\\ \varepsilon_{t,j} &= \sigma_{t,j}\cdot e_{t,j}\\ \sigma^2_{t,j} &= \omega_j+\alpha_j\varepsilon^2_{t-1,j}+\beta_j\sigma^2_{t-1,j}\end{aligned}$$

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

(5.15)
$$\begin{aligned} L(\theta^m_1,...,\theta^m_n,\theta^c) &= \sum_{t=1}^{Z} logf(r_{t,1},...,r_{t,n};\theta^m_1,...,\theta^m_n,\theta^c)\\ &= \sum_{t=1}^{Z}\log c(F_1(r_{t,1}),...,F_n(r_{t,n});\theta^c)\\ &\quad+\sum_{t=1}^{Z}\sum_{j=1}^{n}\log f_j(r_{t,j};\theta^m_j)\end{aligned}$$

where the multivariate $c(\cdot\,;\theta^c)$ is denoted as the regular vine model. Maximum the Equation (5.15) is possible. However, it is time consuming when n is large. We use Inference Functions of Margins (IFM) method [14] to resolve the issue. 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 filtered the financial returns and the univariate parameters $\theta^m_j=(c_j,\varphi_j,\gamma_j,\omega_j,\alpha_j,\beta_j)$ are derived. In the second step, the joint log-likelihood in Equation (5.15) is maximized over copula parameters θ^c, and the univariate parameters $(c_j,\varphi_j,\gamma_j,\omega_j,\alpha_j,\beta_j)$ is fixed at the estimated value in the first step. It means that the joint log-likelihood is reduced to the equation which consist of only copula parameters due to parameters of log-likelihood is fixed.

6 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

(6.16)
$$VaR_t(1-\alpha)=-inf\{c\in\mathbb{R}:P(r_t\le c|\mathcal{F}_{t-1})|\ge(1-\alpha)\}$$

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

(6.17)
$$r_{portfolio,t}=\sum_{i=1}^{n}\mu_i r_{i,t}$$

Suppose the current time is time t, we calculate the Value at Risk forecasting at time $t+1$. The process for computing VaR is given as follow:

(i). Fit ARMA(1,1)-GARCH(1,1) with Student t innovations with returns by using the Equation (5.14) Then, the standardized residuals is obtained by :

(6.18)
$$\hat{e}_{t,j}=\frac{r_{t,j}-\hat{c}_j-\hat{\varphi}_j r_{t-1,j}-\hat{\gamma}_j\hat{\sigma}_{t-1,j}\hat{e}_{t-1,j}}{\hat{\sigma}_{t,j}}$$

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

$$\hat{\sigma}^2_{t+1,j} = \hat{\omega}_j + \hat{\alpha}_j \hat{\varepsilon}^2_{t1,j} + \hat{\beta}_j \hat{\sigma}^2_{t,j} \quad (6.19)$$

(iii). The standardized 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 Student-t cumulative distribution function;

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

$$\hat{r}_{t+1,j} = c_j + \hat{\varphi}_j r_{t,j} + \hat{\gamma}_j \hat{\varepsilon}_{t,j} + \hat{\varepsilon}_{t+1,j} \quad (6.20)$$

(viii). The portfolio return is calculated by using Equation (6.17). Then, we repeat from Steps (iv) to (vii) for N times (e.g. $N = 10000$). 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.

To validate the VaR forecast, we consider use the test of ex-post exceedance, which is defined at time t as:

$$I_t = \begin{cases} 1, & \text{if} \quad r_{portfolio,t} < VaR_t(1-\alpha); \\ 0, & \text{otherwise.} \end{cases} \quad (6.21)$$

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. In addition, the quality of VaR forecasting can be judged by backtesting methods, including unconditional, independent and conditional coverage tests, which are presented in [13].

7 Case Study

We study the daily log-return data of 8 major European indices, including 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 Europeans 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 reflect the most trading situation of European stock market. In particular, the period we used is from 01/03/2006 to 28/12/2012, totally 1682 observations for European indices. All the data was downloaded from FRB St. Louis (http://research.stlouisfed.org).

Table 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	

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

7.1 Non-parametric Dependence Analysis Before building our model to fit the data, we perform a non-parametric method to analyse the lower and upper tail dependence, which is mentioned in Section 2. The results are shown in Table 2. We can see that for total 56 pairs, 46 pairs have a strong upper tail dependence, which indicates that their upper tail dependence are larger than lower tail dependence. In addition, only 11 pairs have a small gap (less than 0.1) between lower and upper tail dependence. These descriptive statistics indicate that for most financial returns, they have a stronger upper tail dependence than lower tail dependence. Due to the large gap between lower and upper tail dependence, it seems that the two kind of tail dependencies are significantly different. Therefore, vine copula model with asymmetric dependence can be used for checking whether the two kind of tail dependencies are significantly different.

7.2 Regular Vine Copula Specification and Tail Dependence Analysis Each index returns are fit with univariate ARMA(1,1)-GARCH(1,1) with Student-t innovations. The tests of Box and Pierce (BP) [5] and Ljoung and Box (LB) [17] are employed for checking the autocorrelation of standardised residuals. Table 3 shows the result of the two tests, which indicates there are no autocorrelation left for all 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. We can obtain the vine tree structure V by using

Table 3: Results of BP and LP Tests

j	e_j (BP)	e_j^2(BP)	e_j(LP)	e_j^2(LP)
GD.AT	0.573	0.199	0.573	0.199
ˆATX	0.500	0.113	0.499	0.113
ˆFCHI	0.798	0.150	0.798	0.149
ˆGDAXI	0.319	0.315	0.318	0.315
ˆFTSE	0.993	0.152	0.993	0.152
ˆSSMI	0.766	0.656	0.766	0.655
AEX.AS	0.223	0.713	0.222	0.713

e_j and e_j^2 are standardised residuals and squared standardised residuals respectively from Arma-garch fits. These values in corresponding columns are the p value for BP and LP tests.

Table 4: Tail Dependence Analysis by Using Various Copula

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

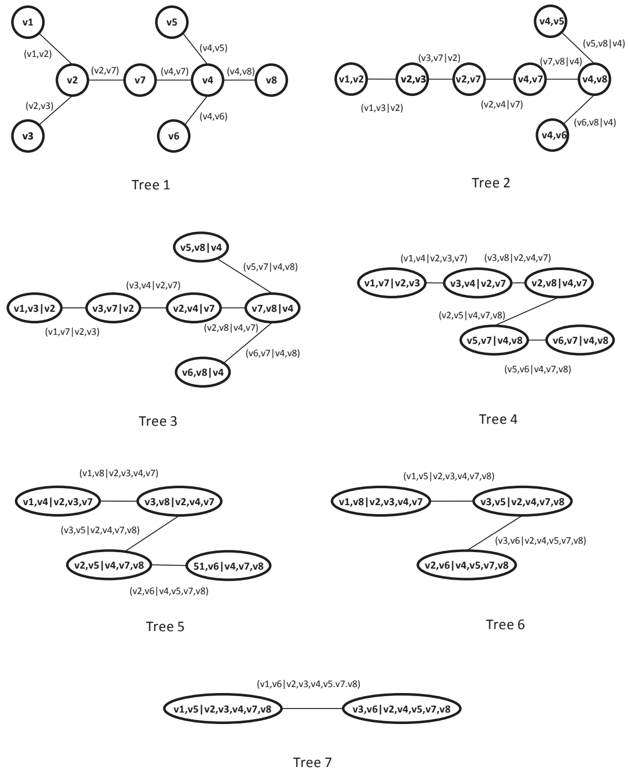

Figure 1: The Partial Regular Vine Tree Structure

Algorithm 1. The Figure 1 shows the full tree structure that is built by our algorithm. in Figure 1, we can see that 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 7.1, the bivariate copula 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, BB1, S.BB1, BB7 and S.BB7 copula can provide both lower and upper tail dependence. Therefore, we consider to use the BB1, BB7, S.BB1 and S.BB7 copula to build vine copula model with asymmetric dependence to capture the asymmetric characteristics. In order to compare the performance of various copula, we use only one copula family (e.g., BB1) to fit the partial regular tree structure, which allow us to easily assess the performance of each copula family.

The Table 4 shown the tail dependence in Tree 1 of Figure 1 during the period from 2006 to 2012. The non-parametric and t copula results are listed as reference. The result shows that lower tail dependence of pairs in Tree 1 are less than their corresponding upper tail dependence. Various bivariate copula provide different results. However, they show same conclusion that lower tail dependence are less than upper one.

In order to investigate the tail dependence and its movement trend, we use fixed period (e.g., one year) as the investigation period of tail dependence. We then use a moving windows of 890 observations corresponding to approximately 2.5 years of daily observations, from 6/02/2007 to 28/12/2012. The partial regular vine copula are re-estimated daily in moving windows to produce tail dependence of investigation period. While estimating the tail dependence of investigation period over the moving windows, we use the vine tree structure mentioned in Figure 1 as partial regular vine. For copula selection, we use mixed copula candidatures (including BB1, BB7, S.BB1 and S.BB7) to fit the vine tree structure. The selection criteria is based on the AIC, which means we choose copula candidatures with smallest value of AIC. In order to find the movement trend in different length of period, we use 12, 24 or 36 months as the investigation period. The result of pair $\{v1, v4\}$ in tree 1 is shown in Figure 7.2. The top 2, middle 2 and bottom 2 figures in Figure 1 use the 12, 24, 36 months as the investigation period respectively. The left 3 figures show the lower and upper tail dependence, and the right 3 figures indicate

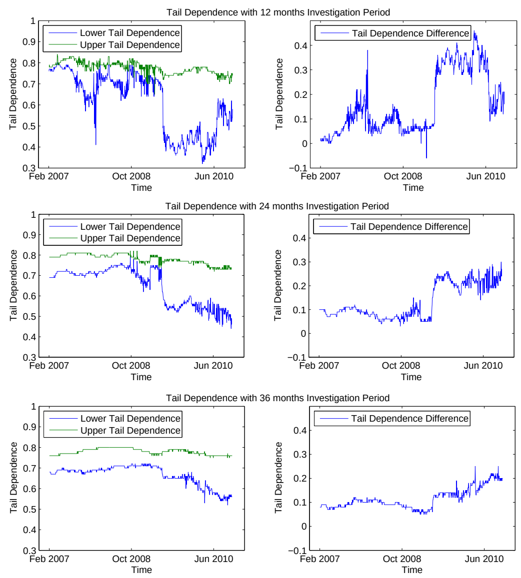

Figure 2: Lower and Upper Tail Dependence

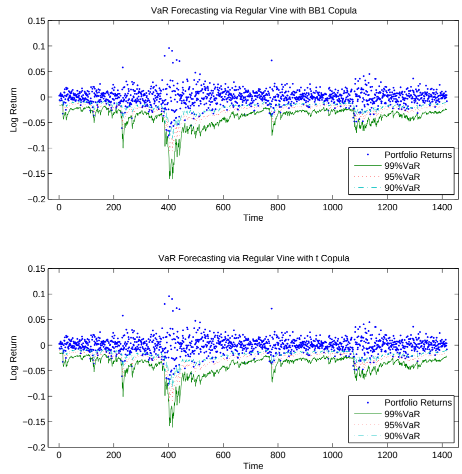

Figure 3: The VaR Forecasting of Portfolio Returns

the corresponding difference between lower and upper tail dependence. We can see that the gap between lower and upper tail dependence has a sharp increase since January 2009. The gap in short investigation period (12 months) is larger than those in long investigation period (24 or 36 months). It indicates that the difference between lower and upper tail dependence is more significant in short investigated period than in long one. However, the difference is decreasing over the length of investigation period increasing. For other pairs, which are omitted due to the limitation of pages, we can find similar conclusion.

7.3 Value at Risk Forecasting The predictable performance of model can be examined via Value at Risk (VaR) forecasting. We use a moving windows from 04/01/2007 to 28/12/2012, totally 1417 observations, corresponding to approximately 4 years of trading days. A training period from 04/01/2006 to 28/12/2006 with 264 observations, appropriately 1 year of trading days. Then, the model are re-estimated daily to produce the one day ahead VaR forecasting. While re-estimating the regular vine copula model, we use the partial regular vine tree structure shown in Figure 1. Various copulas are fit in the regular vine copula model, in order to compare the performance and conclude whether model with asymmetric dependence are better than those with symmetric dependence. In addition, we consider to fit the canonical vine and D vine tree structure with copula, in order to compare the performance and find whether regular vine tree structure are better than canonical vine or D vine.

Table 5 shows the backtesting results of partial regular vine, canonical vine and D vine with various copula. We fit regular vine, canonical vine or D vine with the copula indicated in second row of Table 5. 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. Clayton copula has only lower tail dependence, and 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 Clayton, Gumbel and BB6 copulas which have only one tail dependence. Figure 3 shows the corresponding VaR forecasting that produced by regular vine with 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 those with symmetric lower and upper tail dependence, and (2) the models with two tail dependencies are better than those with only one tail dependence.

8 Conclusion and Future Research

It is a very challenging task for modelling high-dimensional and asymmetric dependence. Existing research made only part progress regarding with the high-dimensional asymmetric dependence modelling. This work has proposed a partial correlation-based regular vine copula model to address this challenging issue. It has been demonstrated by analyzing the asymmetric dependence in cross-country stock markets. Since the total parameters of partial regular vine copula model may be increasing quadratically, we may optimize our model by having conditional independence copula or Guassian copula in the edges in which partial correlations are weak, and then apply it to data sets of a large number of dimensionalities ($>$ 100 variables).

Table 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.

References

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

[2] Abdous, B., Genest, C., Rémillard, B.: Dependence properties of meta-elliptical distributions. In: Statistical modeling and analysis for complex data problems, pp. 1–15. Springer (2005)

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

[4] Berg, D., Aas, K.: Models for construction of multivariate dependence: A comparison study. The European Journal of Finance 15(7), 639–659 (2009)

[5] Box, G.E., Pierce, D.A.: Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. Journal of the American Statistical Association 65(332), 1509–1526 (1970)

[6] Boyer, B., Gibson, M., Loretan, M.: Pitfalls in tests for changes in correlations. FRB International Finance Discussion Paper (597) (1997)

[7] Cormen, T., Leiserson, C., Rivest, R., Stein, C.: Introduction to algorithms. MIT press, third edn. (2009)

[8] Demarta, S., McNeil, A.J.: The t copula and related copulas. International Statistical Review 73(1), 111–129 (2005)

[9] Fang, H.B., Fang, K.T., Kotz, S.: The meta-elliptical distributions with given marginals. Journal of Multivariate Analysis 82(1), 1–16 (2002)

[10] Fantazzini, D.: 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 (2009)

[11] Frahm, G., Junker, M., Schmidt, R.: Estimating the tail-dependence coefficient: Properties and pitfalls. Insurance: Mathematics and Economics 37(1), 80–100 (2005)

[12] Garcia, R., Tsafack, G.: Dependence structure and extreme comovements in international equity and bond markets. Journal of Banking & Finance 35(8), 1954–1970 (2011)

[13] Guermat, C., Harris, R.D.: Robust conditional variance estimation and value-at-risk. Journal of Risk 4, 25–42 (2002)

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

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

[16] Jondeau, E., Rockinger, M.: The copula-garch model of conditional dependencies: An international stock market application. Journal of international money and finance 25(5), 827–853 (2006)

[17] Ljung, G.M., Box, G.E.: On a measure of lack of fit in time series models. Biometrika 65(2), 297–303 (1978)

[18] Longin, F., Solnik, B.: Extreme correlation of international equity markets. The Journal of Finance 56(2), 649–676 (2001)

[19] McNeil, A.J., Nešlehová, J.: Multivariate archimedean copulas, d-monotone functions and l_1-norm symmetric distributions. The Annals of Statistics pp. 3059–3097 (2009)

[20] Nikoloulopoulos, A.K., Joe, H., Li, H.: Extreme value properties of multivariate t copulas. Extremes 12(2), 129–148 (2009)

[21] Sklar, M.: Fonctions de répartition à n dimensions et leurs marges. Université Paris 8 (1959)

[22] Wei, W., Fan, X., Li, J., Cao, L.: Model the complex dependence structures of financial variables by using canonical vine. In: Proceedings of the 21st ACM international CIKM. pp. 1382–1391. ACM (2012)