Optimal Allocation of High Dimensional Assets through Canonical Vines

Wei Wei[1], Jinyan Li[1], Longbing Cao[1], Jingguang Sun[2], Chunming Liu[1], and Mu Li[1]

[1] Advanced Analytics Institute, FEIT, University of Technology Sydney, Australia
[2] Liaoning Technical University, China
{Wei.Wei-3,chunming.Liu}@student.uts.edu.au,
{Jinyan.Li,LongBing.Cao}@uts.edu.au, sunjinguang88@163.com,
muli60@gmail.com

Abstract. The widely used mean-variance criteria 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. It is a computationally challenging task to model the asymmetries and skewness of joint distributions of returns in high dimensional space due to their own complicated structural complexities. This paper proposes to use a new form of canonical vine to produce the complex joint distribution of asset returns. Then, we use the utility function of Constant Relative Risk Aversion 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 new 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.

Keywords: Canonical Vine, Mean Variance Criterion, Financial Return.

1 Introduction

Financial asset returns follow non-normal distributions—asymmetries and skewness very often exist in the distribution of financial asset returns such as in stock returns [2] and [3]. These asymmetries facts violate the traditional distribution assumption on financial asset returns, making the traditional mean variance analysis [10] unreasonable. Some previous studies [12], [13] and [15] had attempted to compare the expected utility obtained from the mean-variance criterion with the approximated utility obtained from the benchmark portfolios (those equally divided portfolios). It was found that the mean variance criterion had poor performance on analyzing 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 an normal distribution in line with the mean variance criterion assumption.

Arrow [4] laid down a theoretical foundation for the importance of using distributional asymmetries. He suggests that a desirable property of utility functions (including the Constant Relative Risk Aversion utility function) is the

J. Pei et al. (Eds.): PAKDD 2013, Part I, LNAI 7818, pp. 366–377, 2013.

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 direct impact on the skewness of the portfolio return. Therefore, while making the portfolio decision favorable 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 [12] and [13], had proposed models to construct the dependence structure with only two financial assets. 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 demanded to develop a model that can resolve difficulties in the high dimensional asset allocation.

There are three challenges in the 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, it is 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 big number of parameters which are increased 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, we propose to use a canonical vine based dependence model for an optimization of the high dimensional asset allocation. The new model can capture asymmetric and skew correlations in the dependence structure, and can optimize the dependent structure. To address the high dimensional issue, we employ the idea of partial correlation to construct the canonical vine in our dependence model. It can capture the most important correlations in the dependence structure, and can reduce the complexity of the dependence structure remarkably to make invertors understand comprehensively. In addition, we also employ the ARMA-Garch model for the estimate of marginal distribution to capture volatility clustering and fat tail in financial assets.

The main contribution made by this paper is the partial correlation based canonical vine. The canonical vine is optimized 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. Suppose there are 50 variables, the normal canonical vine will generate 1225 nodes. However, in our partial correlation based canonical vine, the number of nodes is only one tenth of that (around 267 nodes). In addition, our method can test hypotheses in parallel such as: (1) whether these asymmetries are predictable out of sample; and (2) whether we 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 the 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 paper, we first build the model which can capture all important asymmetries and skewness in the dependence structure, then we compare it with those models that do not consider the correlations and/or asymmetric dependence structure between financial assets.

The rest of the paper is organized as follows. Section 2 presents a short introduction to copula theories which is closely related to canonical vine. Section 3 describes the problem of optimal assets allocation in portfolio, including how to construct the optimal canonical vine and marginal distribution. Section 4 discusses how to evaluate the performance of our model with equally divided allocation and mean-variance criterion. In Section 5, we apply the optimal canonical vine to capture the dependence structures of two portfolios, and evaluate the performance of our model in comparison to the performance by the mean variance criterion and equally divided allocation method. Section 6 concludes the paper with a summary.

2 Related Work

Copula is a useful tool to model the non-normal distribution. It can capture the complicated correlation between variables, including linear or non-linear. According to Sklar's theorem [14], a copula function is defined to connect univariate functions to form a multivariate distribution function. The definition of a copula function is given by:

$$F(x_1, ..., x_n) = C(F_1(x_1), F_2(x_2), ..., F_n(x_n)) \tag{1}$$

where, $x = [x_1, x_2, ..., x_n]$ is a random variable vector, F is a joint distribution and $F_1, F_2, ..., F_n$ are the marginal distributions of the corresponding variables respectively. It shows that all multivariate distribution functions, and copula function can be used in conjunction with univariate distributions to construct multivariate distribution functions. The differential of Equation (1) is:

$$f(x_1, ..., x_n) = \prod_{i=1}^{n} f_i(x_i) \cdot c(F_1(x_1), F_2(x_2), ..., F_n(x_n)) \tag{2}$$

where, c is the density copula function and $f_i(x_i)$ is the density function of marginal distributions. Equation (2) shows that the joint density function includes two parts. One is the description of the marginal behavior of individual factor, which is marginal distributions. The other is the description of their dependence structure, which is copula function. It implies an important property that copula function can separate dependence structure from marginal distribution function. We can model the individual variables using whichever marginal distributions provide the best fit and then model the dependence structure by using the copula function. The useful property can help understand the complex dependence structure, and describing the complex dependence structure on a quantile scale. The deep instruction regarding with copula can be found in [11].

Patton [12] builds an asset allocation with copula. The portfolio is composed by two assets, evaluating the asset allocation with the investor's utility function. Jondeau and Rockinger [7] use 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. Sun et al. [15] 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-garch model is able to capture multi-dimensional co-movements among the indices.

3 The Portfolio Optimization Problem and Our New Method

3.1 CRRA Optimization Function

Suppose that hypothetical investors follow the class of Constant Relative Risk Aversion (CRRA) utility functions:

$$U(\gamma) = \begin{cases} (1-\gamma)^{-1} \cdot (P_0 R_{port})^{1-\gamma}, & \text{if}\gamma \neq 1 \\ log(P_0 R_{port}), & \text{if}\gamma = 1 \end{cases} \tag{3}$$

where γ is the risk aversion parameter, P_0 is the initial wealth and R_{port} is the portfolio return. In this paper, the value of risk aversion parameter is considered at four different levels, including $\gamma = 2, 5, 7$ and 10, as suggested by [5]. We use CRRA utility function to calculate the expectation return of the hypothetical investors as its prominence in the finance literature. 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 built a portfolio of returns. Our work is focused on the portfolio return with high dimensional assets, defined as

$$P_0 R_{port} = P_0 \cdot (1 + \sum_{i=1}^{n} \omega_i X_t) \tag{4}$$

where $X_t = x_{i,t}$ is the asset return at time t, and ω 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. We develop the density forecasts of the joint distribution $F_{1,t+1}, ..., F_{n,t+1}$ and the copula function C_{t+1}. Then, we use the forecast function to calculate the optimal weights ω^*_{t+1} for the portfolio. The optimal weights, ω^*_{t+1}, are found by maximizing the expected CRRA utility function:

$$\begin{aligned} \omega^*_{t+1} &= \arg\max_{\omega \in W} E_{F_{t+1}}[U(1 + \sum_{i=1}^{n} \omega_{i,t+1} X_{t+1})] \\ &= \arg\max_{\omega \in W} \int_{x_1} \int_{x_2} ... \int_{x_n} U(1 + \sum_{i=1}^{n} \omega_i x_i) \cdot f_{t+1}(x_1, x_2, ..., x_n) \cdot dx_1 \cdots dx_n \\ &= \arg\max_{\omega \in W} \int_{x_1} \int_{x_2} ... \int_{x_n} U(1 + \sum_{i=1}^{n} \omega_{i,t+1} x_{i,t+1}) \cdot f_{1,t+1}(x_1) \cdots f_{n,t+1}(x_n) \\ &\quad \cdot c_{t+1}(F_{1,t+1}(x_1), F_{2,t+1}(x_2), ..., F_{n,t+1}(x_n)) \cdot dx_1 \cdots dx_n \end{aligned} \tag{5}$$

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 conditional distribution of returns by using maximum likelihood estimation, and then optimize the portfolio weights via the predicted distribution of return. For the integral function, we use Monte Carlo replications to estimate the value of integral. For the optimal portfolio weights, we employ the *Broyden-Fletcher-Goldfarb-Shanno* (BFGS) algorithm to obtain the optimal weights.

3.2 Partial Correlation Based Canonical Vine: Our New Ideas

The key step in Equation 5 is to form the joint density function $f_{t+1}(x_1, ..., x_n)$ at time t+1. Equation (2) shows that the joint density function can be divided into two parts: the copula function c_{t+1} and the marginal distributions $f_{1,t+1}(x_1) \cdots f_{n,t+1}(x_n)$. In this Section, we discuss how to produce a copula density function. The construction of marginal distributions is explained in Section 3.3.

One way to build high-dimensional copula density function is to use canonical vine to build dependence structure as proposed by [1]. The basic scheme for modeling high-dimensional dependence structure with 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 canonical vine transforms one high dimensional dependence structure into multiple two-dimensional structures. However, one important issue of canonical vine is that if the variables are large in number, the canonical vine will become quit complex. In addition, the nodes of canonical vine will increase exponentially as the variables increase. Therefore, we propose a new partial correlation based canonical vine to model high dimensional dependence structure.

The principle for the new canonical vine construction is to capture the most important correlation in the dependence structure, meanwhile to decrease the number of nodes, and to reduce the complexity of dependence structure. That is, the new canonical vine can capture the most important correlation, and ignore the weak correlations. Following this principle, we develop a new algorithm to construct and optimize the canonical vine by using partial correlation. First, we construct the canonical vine by using partial correlation rather than by using bivariate conditional copula function. Then, we optimize the partial correlation based canonical vine by setting the absolute small value of partial correlation to zero to decrease number of nodes and reduce the complexity of canonical vine. The optimal partial correlation based canonical vine can be mapped into the canonical vine based on bivariate conditional copula, provided that pair copulas are from the elliptical copula family [8]. The algorithm to construct the optimal canonical vine is presented in Algorithm 1.

We take an example to explain how to construct and optimize the canonical vine. Suppose there are one comprehensive index and 6 stocks which denoted by M, A, B, C, D, E and F. The canonical vine will consist of 6 trees and 21 nodes in both the canonical vine structure based on partial correlation and conditional copula. All the trees and nodes are shown in Figure 1. Each node

Algorithm 1. Canonical Vine Construction and Optimization

Require: observations of n input variables

1: Calculate all values of partial correlation, and then allocate the smallest absolute value of partial correlation to the node in T_{n-1} (T_{n-1} is the bottom tree).
2: **for** $k = 1, ..., n-2$ **do**
3: If $T_i > T_k$, find an appropriate root variables in T_i which can minimize the function $\sum |\rho_{c:d}|$, where T_i indicates the ith tree and T_k is broken level tree;
4: If $T_i \leq T_k$, find an appropriate root variables in T_i which can minimize the function of $\sum log(1-\rho^2_{c;d})$;
5: **end for**
6: There will be $(n-2)-1$ canonical vines as $k = 1, ..., n-2$. Calculate the function $-log(D)$ of all of the canonical vines based on partial correlation, and choose the maximum value of the function as the 'best' canonical vine. (D is calculated in Equation 7);
7: For the 'best' canonical vine, the small absolute values of partial correlation, which are less than significance value τ, are set to zero;
8: The optimal canonical vine based on conditional copula is corresponding to the canonical vine based on partial correlation;
9: **return** The optimal canonical vine dependence structure.

can be allocated to one bivariate copula or one partial correlation. The first step is to build the partial correlation based canonical vine. The partial correlation of all variables can be obtained by using the following Equation:

$$\rho_{12;3,\ldots,n} = \frac{\rho_{12;3,\ldots,n-1} - \rho_{1n;3,\ldots,n-1} \cdot \rho_{2n;3,\ldots,n-1}}{\sqrt{1-\rho^2_{1n;3,\ldots,n-1}} \cdot \sqrt{1-\rho^2_{2n;3,\ldots,n-1}}} \tag{6}$$

For these 7 variables, there are totally 21 partial correlations and 6 trees. T_i is defined as the ith tree in the paper. The smallest absolute value of these partial correlations is allocated to the root node in T_6 (the sixth tree shown in Figure 1) as the T_6 only has one node. Suppose the selected partial correlation in the T_6 is $\rho_{E,F;M,A,B,C,D}$. The variables in T_6 are variables E and F. The sets $c_6 = \{E, F\}$ and $d_6 = \{M, A, B, E, F\}$ are called conditioned set and conditioning set respectively. The next step is to chose the root variable in T_5. In T_5, there are two nodes which can be allocated as two partial correlations. The root variable of T_5 should be from d_6. If the selected root variable of T_5 is D, then it can generate two new conditioned sets: $c_5 = \{D, E\}$ and $c'_5 = \{D, F\}$. The corresponding conditioning set for c_5 and c'_5 is $d_5 = \{M, A, B, C\}$. The partial correlations allocated to the two nodes are $\rho_{D,E;M,A,B,C}$ and $\rho_{D,F;M,A,B,C}$. If we choose C as the root variable of T_5, the two new conditioned sets are: $c_5 = \{C, E\}$ and $c'_5 = \{C, F\}$. The corresponding conditioning set for c_5 and c'_5 is $d_5 = \{M, A, B, D\}$. The partial correlation allocated to the two nodes will be $\rho_{C,E;M,A,B,D}$ and $\rho_{C,F;M,A,B,D}$. If the selected root variable of T_5 is M, A or B , the processes of generating conditioned and conditioning set are similar to those root variable as C or D. However, we need to identify which variable is the most appropriate root variable in T_5.

We proposed a method to identify the appropriate root variable, which is called tree broken method. In the paper, we define that k is a tree-broken level. For trees beyond the kth tree ($T_i > T_k$), the appropriate root variable must minimize the value of function $\sum|\rho_{c:d}|$. For trees within the kth tree ($T_i \leq T_k$), the appropriate root variables must minimize the value of function $\sum log(1-\rho^2_{c;d})$. For example, if $k=3$, for T_1, T_2 and T_3 (the first, second and third trees), the appropriate root variables for these trees must minimize the value of function $\sum|\rho_{c:d}|$. For T_4, T_5 and T_6 (the fourth, fifth and sixth trees), the appropriate root variables for these trees must minimize the value of function $\sum log(1-\rho^2_{c;d})$. The parameter k can be chosen from $1, 2, 3, 4, 5$. Therefore, there should totally have 5 canonical vines. Then, the 'Best' canonical vine should maximize the value of function $-log(D)$, where D is the determinant of canonical vine which is calculated by using:

$$D = \prod_{\{i,j\}} (1-\rho^2_{i,j;d(i,j)}) \tag{7}$$

where $d(i,j)$ is the conditioning set excluding variable i,j. The corresponding conditioned set is i,j. The small absolute values of partial correlation in the 'Best' canonical vine, which is less than the significance value τ, is set to zero. Finally, the optimal canonical vine structure based on partial correlation is built. Since the canonical vine based on conditional vine has a similar structure, we can construct the optimal canonical vine based conditional copula by using the structure based on partial correlation.

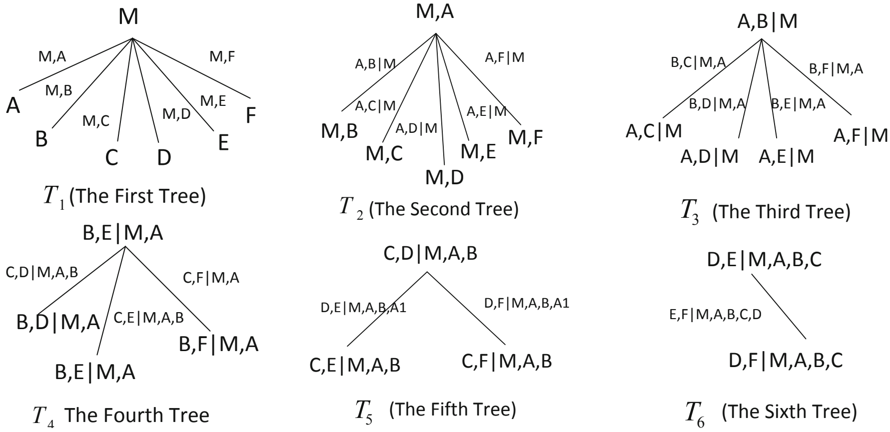

Fig. 1. Canonical Vine Trees

3.3 Marginal Models Specification

The second step of constructing joint probability density function is to build the marginal distribution for each asset return. In the paper, we choose the AR(1)-Garch(1,1) as the marginal distribution. Hansen and Lunde [6] provided evidence that it is difficult to find a volatility model which outperforms the Garch(1,1) model. We estimate the AR(1)-Garch(1,1) with skewed student t innovations. The reason is that the skewed student t innovations can capture and model the characteristics of financial asset return, such as volatility clustering and fat tail.

3.4 Parameter Estimation

We use a two-step procedure to estimate the canonical vine copula and marginal distributions. Taking a log of both sides in Equation 2, we can obtain:

$$\log f(x_1, ..., x_n) = \sum_{i=1}^{n} \log f_i + \log c[F_1(x_1), ..., F_n(x_n)] \tag{8}$$

The joint log-likelihood is equal to the sum of the marginal log-likelihoods and the canonical vine copula log-likelihood. Parameters can be estimated separately by optimizing the marginal log-likelihood and canonical vine copula log-likelihood in two steps.

3.5 Evaluation

In the paper, we use the final amount of the portfolio, the utility obtained by daily returns, to compare the assets allocation with different portfolio decisions. The final amount is the amount dollars obtained at the end of entire out of sample period (testing period). To compare the utility, we use opportunity cost, also called management fee or forecast premium, which is the amount that investor would pay to switch from the the equally divided portfolio to analyzed allocation. The benchmark of assets allocation is by equally divided portfolio, which means the weights of all assets are equal. We compare the performance of canonical vine with mean-variance criterion. Suppose that r_{port} is the optimal portfolio return obtained by canonical vine or mean-variance, and r^*_{port} is the return obtained from the equally divided portfolio. In other word, the opportunity cost is actually return which is added to the return obtained from equally divided portfolio, to make sure the investor be indifferent to the returns obtained from the analyzed model. Then, the opportunity cost Δ can be defined as:

$$U(1 + r_{port} + \Delta) = U(1 + r^*_{port}) \tag{9}$$

Equation (9) can be resolved via the Taylor approximation with CRRA utility function in [7].

4 Asymmetry Analysis with Case Study

4.1 Data and Model Specification

We used two financial asset portfolios in the evaluation of the performance of the newly 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 $Stoxx$50 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 3.3, AR(1)-Garch(1,1) is considered as the marginal distribution to capture the skewness. The Ljung Box Q test was used for checking the existence of residual autocorrelation for all of the stocks and indices. If the marginal distribution of any stock or index fails the Q test, we increased the value of p in the AR(p)-Garch(1,1) model until all pass the Q test. The results of the Ljung Box Q test are not listed in the paper due to page limit.

For the canonical vine, we choose the 'best' canonical vine that maximizes the value of function $-log(D)$. The root variables in T_1 are the comprehensive indices $S\&P500$ and $Stoxx50E$. This is reasonable as the comprehensive index has much stronger correlation than other stock prices very often. In the optimization of the canonical vine, we considered to use different values of τ (significance values), and compared with the 'unprune' canonical vine. The comparison between canonical vines are based on function $-log(D)$, where D is the determinant as mentioned in Section 3.2. The function is used to calculate the determinant of the partial correlation based canonical vine, and it can be then used for comparing the similarity of vine structures [9]. In the paper, all the nodes in the canonical vine are assigned to bivariate t copulas.

Table 2 shows all of the partial correlations in the 'unprune' canonical vine. All of these partial correlations are used with absolute value as our 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 $Stoxx50E$ has 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 $S\&P500$ is strictly selected from the least correlations the each other. The result shown in Table 1 implies a similar conclusion. Table 1 shows the determinant and number of nodes under various values of τ. The value of function $-log(D)$ means the strength of canonical vine. It is observed that the portfolio of $S\&P500$ has more number of nodes in 'unprune' canonical vine than the portfolio of $Stoxx50E$ as $S\&P500$ has 51 variables, compared with 50 variables in $Stoxx50E$. However, the portfolio of $Stoxx50E$ has more nodes than the portfolio of $S\&P500$ in each corresponding level of τ. It suggests that $Stoxx50E$ has stronger correlations. Under the case of the optimal canonical vine ($\tau = 0.1$), for the both portfolios, the optimal canonical vine has less number of nodes and parameters, namely only one tenth of those of the 'unprune' canonical vine. It means that the one tenth nodes can contribute the majority of the dependence structure. Other nodes contribute a little in the dependence structure.

4.2 Experiment Results and Analysis

The performance of our model was 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. We evaluated the performance of our model with 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. We considered to compare the performance of our model with the mean variance criterion and the equally divided allocation to understand whether our model is useful.

Table 3 shows the results related to the opportunity costs and the final amounts for the two portfolios $S\$P500$ and $Stoxx50E$. Table 3 provides strong evidence that our canonical vine based model is the best at all levels γ for both portfolios. In detail, for the portfolio $S\&P500$, we compare the two canonical vines with $\tau = 0.1$ and 0.05. There is no obvious difference between these two canonical vines, indicating that the two canonical vines implement a similar forecasting of the samples. However, the number of parameters in the canonical vine (τ=0.1) is only half of those in the canonical vine ($\tau = 0.05$). The canonical vine($\tau = 0.1$) 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 a basis of mean variance analysis, the final profit would be less than those on the basis of equally divided allocation. Therefore, the mean variance criterion is not useful. Considering the good performances of canonical vine model, the mean variance criterion, which is caused by the normal distributions, cannot catch the features of asymmetry and skewness of 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 to the mean variance criterion is that $Stoxx50E$ has stronger asymmetry and skewness than those 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.

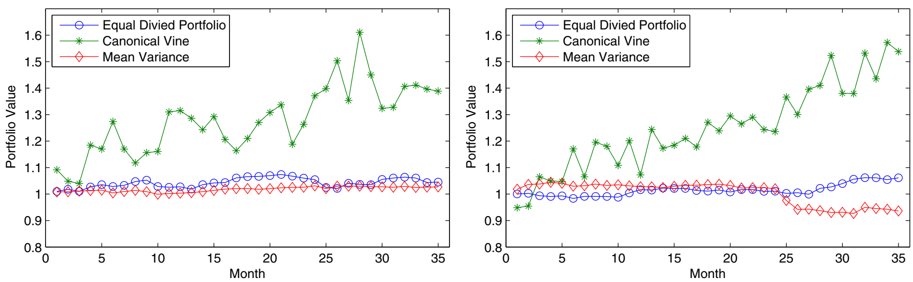

Fig. 2. Portfolio values over 35 months for the canonical vine ($\tau = 0.1$), mean variance, and equally divided allocation, $\gamma = 2$. The left one is for $S\&P500$, and the right one is for $Stoxx50E$.

Table 1. Determinants and Numbers of Nodes for Portfolio $S\&P500$ and $Stoxx50E$

τ	unprune	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.20
$Stoxx50E$												
$-log(D)$	28.36	28.35	28.31	28.24	28.09	27.87	27.66	27.39	27.15	26.92	26.46	22.74
No. nodes	1225	1059	875	758	638	526	456	392	349	317	267	85
$S\&P500$												
$-log(D)$	13.22	13.21	13.16	13.05	12.91	12.70	12.41	12.09	11.81	11.45	11.16	9.44
No. nodes	1275	1044	847	678	560	457	362	285	235	186	153	62

Table 2. Partial Correlations in Canonical Vine for Portfolios $S\&P500$ and $Stoxx50E$

	Min	Max	Mean	25%Quantile	50%Quantile	75%Quantile
$S\&P500$	0.0002	0.5281	0.0572	0.0144	0.0340	0.0641
$Stoxx50E$	0.0004	0.8702	0.0787	0.0171	0.0422	0.0929

Table 3. Opportunity Costs and Final Amounts for $S\&P500$ and $Stoxx50E$

	$S\&P500$				$Stoxx50E$			
Relative Risk Aversion (γ)	2	5	7	10	2	5	7	10
Opportunity Cost								
Canonical Vine ($\tau = 0.1$)	1.33%	3.21%	5.64%	7.32%	1.91%	4.19%	6.14%	10.01%
Canonical Vine ($\tau = 0.05$)	1.31%	3.18%	5.56%	7.21	1.89%	4.17%	6.12%	9.95%
Mean Variance Criterion	-2.36%	-3.62%	-1.24%	-1.19%	-2.55%	-6.52%	-3.20%	-1.40%
Final Amount								
Canonical Vine ($\tau = 0.1$)	1.39	1.41	1.46	1.50	1.53	1.58	1.62	1.64
Canonical Vine ($\tau = 0.05$)	1.40	1.42	1.47	1.50	1.54	1.60	1.62	1.65
Mean Variance Criterion	1.02	1.00	1.04	1.04	0.93	0.80	0.91	0.99
Equally Divided Allocation	1.05	1.05	1.05	1.05	1.06	1.06	1.06	1.06

Figure 2 shows the portfolio values obtained from our canonical vine, the mean variance criterion and equally divided allocations at the end of each month. The trend obtained from our 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. We can find that the mean variance criterion performs a bit better in $S\&P500$ than in$Stoxx50E$.

Overall, we can see that: (i) The new 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 high dimensional assets portfolio that has asymmetry and skewness; and (iii) Compared with the equally divided allocation, our new canonical vine has a better performance. However, the performance by the mean variance criterion has no obvious difference, or even worse if the high dimensional assets portfolio has strong asymmetry and skewness.

5 Conclusion

The paper proposed a canonical vine based model to optimize the asset allocations of high dimensional assets. To address the touch computational issues caused by high dimensional assets, we employed the partial correlation technique to reduce the complexity of the dependence structure to make invertors understand the model easily. Our experimental results and analysis have shown that the canonical vine model has a better performance for portfolios of strong asymmetry and skewness in comparison to the mean variance criterion, a current widely used method. As a future work, we will extend partial correlation based canonical vine to work with more copula families.

Acknowledgments. This work is sponsored in part by Australian Research Council Discovery Grant (DP1096218, DP130102691) and ARC Linkage Grant (LP100200774).

References

1. Aas, K., Berg, D.: Models for construction of multivariate dependence–a comparison study. The European Journal of Finance 15(7-8), 639–659 (2009)
2. Ang, A., Bekaert, G.: International asset allocation with regime shifts. Review of Financial Studies 15(4), 1137–1187 (2002)
3. Ang, A., Chen, J.: Asymmetric correlations of equity portfolios. Journal of Financial Economics 63(3), 443–494 (2002)
4. Arrow, K.: Essays in the theory of risk-bearing, pp. 90–120. Markham Publishing, Chicago (1971)
5. Campbell, R., Koedijk, K., Kofman, P.: Increased correlation in bear markets. Financial Analysts Journal, 87–94 (2002)
6. Hansen, B.E.: Autoregressive conditional density estimation. International Economic Review 35(3), 705–730 (1994)
7. Jondeau, E., Rockinger, M.: Optimal portfolio allocation under higher moments. European Financial Management 12(1), 29–55 (2006)
8. Kurowicka, D., Cooke, R.: Uncertainty analysis with high dimensional dependence modelling. John Wiley & Sons (2006)
9. Kurowicka, D., Cooke, R., Callies, U.: Vines inference. Brazilian Journal of Probability and Statistics 20, 103–120 (2006)
10. Markowitz, H.: Portfolio selection. The Journal of Finance 7(1), 77–91 (2012)
11. Nelsen, R.: An introduction to copulas. Springer (2006)
12. Patton, A.: On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. Journal of Financial Econometrics 2(1), 130–168 (2004)
13. Riccetti, L.: A copula–garch model for macro asset allocation of a portfolio with commodities. Empirical Economics, 1–22 (2012)
14. Sklar, A.: Fonctions de répartition à n dimensions et leurs marges. Publ. Inst. Statist. Univ. Paris 8(1) (1959)
15. Sun, W., Rachev, S., Stoyanov, S., Fabozzi, F.: 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), 42–76 (2008)