ARTICLE IN PRESS
Physica A xxx (xxxx) xxx

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Composition of the first principal component of a stock index — A comparison between SP500 and VNIndex

Q. Nguyen [a,b,*], N.K.K. Nguyen [c]

[a] *Division of Computational Mathematics and Engineering, Institute for Computational Science, Ton Duc Thang University, Ho Chi Minh City, Viet Nam*
[b] *Faculty of Finance and Banking, Ton Duc Thang University, Ho Chi Minh City, Viet Nam*
[c] *Van Lang University, Ho Chi Minh City, Viet Nam*

HIGHLIGHTS

- We found that the components of the 1st eigenvector of a stock market cross-correlation matrix are not uniformly distributed and could be negative.
- Their distribution are compared between a developed market (the SP500 Index) and an frontier market (the VNIndex), the one in Vietnamese market (VNIndex) is broader and contains more negative values.
- We found a phenomenological linear dependence between the components of the 1st eigenvector and the average correlation of each stock, and derived an approximation equation.
- The portfolio constructed from the 1st eigenvector, the 1st Principle Components, is perfectly correlated to the Index itself in the case of SP500, but not in the case of VNIndex. We suggest that this portfolio should be called the Most Correlated Portfolio.
- We found that the largest weights of this portfolio are also the central hubs in a Minimum-Spanning-Tree (MST) analysis.

ARTICLE INFO

Article history:
Received 7 July 2018
Received in revised form 31 January 2019
Available online xxxx

Keywords:
Eigenvalue
Eigenvector
Cross-correlation stock return matrix
Principle Components
Correlated portfolio
Minimum-spanning tree

ABSTRACT

We analyzed the components of the eigenvector corresponding to the largest eigenvalue of a cross-correlation matrix of stock price change (the 1^{st} eigenvector) using the Principle Components (PCs) analysis method. We found that the components of the 1^{st} eigenvector are not uniformly distributed. In fact, they are proportional to the correlations of individual stocks and the 1st PC, and could be negative if the corresponding stock is "in average" negatively correlated to other stocks. By analyzing data from the world stock index in a developed country (SP500) and an emerging one such as Vietnam (VNIndex) from 2013 to 2017, we found negative components in the 1^{st} eigenvector in the VNIndex (and their distribution is also broader than that of SP500). Furthermore, we found that the largest components correspond to stocks of financial services (brokerage or investment advisory firms) firms in both indices. Those stocks are found to be center hubs in a Minimum Spanning Tree (MST) analysis and play an important role whenever a systemic breakdown happens. Finally, we found a phenomenological linear dependence between the components of the 1^{st} eigenvector and the average stock correlation in both indices, and derived an approximation of this dependency using empirical stylized facts of the data. This result provides an estimation of how the PC1 is composed.

* Corresponding author at: Division of Computational Mathematics and Engineering, Institute for Computational Science, Ton Duc Thang University, Ho Chi Minh City, Viet Nam
E-mail address: nguyenquang@tdtu.edu.vn (Q. Nguyen).

https://doi.org/10.1016/j.physa.2019.04.216
0378-4371/

1. Introduction

The interaction between stock price in a stock market is of great importance in studying the stock market systemic risk and/or finding an optimal trading strategy. One of the approaches is the study of the cross-correlation matrix **C**, whose elements $\mathbf{C}_{ij}$ are the correlation coefficients between the time-series price change of two stocks i and j. The structure and dynamic of the cross-correlation matrix is crucial for the understanding of the market as a complex system.

Several methods, mostly employed from physics statistics, have been used to analyze the cross-correlation matrix. For example, by using Random Matrix Theory (RMT), some common important trends have been found [1–11]: there exist a dominant eigenvalue, which is at least one order of magnitude larger than others, and several other large eigenvalues. In addition, the majority of eigen-spectrum falls within a tight range predicted by the RMT theory, $\lambda_- < \lambda < \lambda_+$, where the RMT bounds λ_- and λ_+ are positive and close to 1 (explicit formula is given in the next section). The eigenvector associated with the largest eigenvalue (hereby call the 1st eigenvector) composes a portfolio that is called market mode while eigenvectors associated with other large eigenvalues represent different economic sectors or groups of stocks that exhibit common trends. Many studies have focused on the market modes and examine its compositions, its inter- and intra–interactions or its temporal evolution. However, few groups have analyzed the market mode in detail.

Most of studies using RMT found that the components of the 1st eigenvector have same sign and/or they are uniformly distributed [7,12–14]. The corresponding portfolio represents the common interaction among stocks in the entire market. Other groups studied the temporal evolution of the largest eigenvalue and found that it becomes very large during market crashes [15,16]. In a similar work, Zheng et al. [17,18] analyzed the temporal dynamic of the sum of the largest eigenvalues of correlation matrix consisting of ten US sector indices to evaluate the systemic risk of the US economy. They found that a steep increase in this sum can be used as an indicator of systemic risk, in which, the largest eigenvalue accounts for nearly 60% of the total variation. In a separate work [14], Nguyen approximated the largest eigenvalue by $N \times \langle \rho \rangle$ where N is the number of stocks and $\langle \rho \rangle$ is the average of all correlation coefficients. Therefore, the increase in the average level of correlation in the entire market is also an equivalent indicator (see also [19]).

In this work, we study the components of the 1st eigenvector in detail from the Principle Component Analysis (PCA) point of view [20]. We show that the components of the 1st eigenvector are not necessary of the same sign and more importantly, they are not uniformly distributed. We will demonstrate that the jth component of the 1st eigenvector can be approximated by the average correlation of the jth stock with all other stocks, to a constant factor. This linear dependency has been found empirically in our previous work [21]. The fact that in many other studies the components of the 1st eigenvector are of the same sign is just an empirical fact: because most stocks in a financial market are, in average, positively correlated to all other stocks. However, in some period and in some market, some stocks may have more negative correlation coefficient (than positive) with other stocks and their corresponding components of the 1st eigenvector will have negative signs (if the others have positive signs). We will show negative components of the 1st eigenvector in the local market.

In general, stocks which have higher average correlation coefficients with other stocks will have higher components of the 1st eigenvector and vice versa. This suggests that the portfolio constructed from the components of the 1st eigenvector is related to the correlation degree of each stock, rather than being equally weighted or market capitalization weighted. Consequently, we suggest that the mode corresponding to the 1st eigenvector should be called the Most Correlated Portfolio. This is in accordance to the name of the 1st Principle Component (PC), the Most Correlated Components as suggested in [22]. The 1st PC is weighted by the components of the 1st eigenvector divided by stock standard deviations.

We also found that the most weighted stocks in 1st PC are usually financial service stocks (in both a developed market (US) and an emerging market (Vietnam)). Those stocks have high average correlation coefficients with the market and therefore have a big influence on the market. Finally, we found that those stocks play important roles in a minimum spanning tree (MST) analysis. MST is a simplified graph constructed from the correlation matrix and is usually used for a taxonomy of market sectors. We show that the most weighted stocks in 1st PC are also central hubs in a MST.

The paper is organized as follows: Section 2 describes the methodologies for constructing the Principle Components and their properties, as well as the data we use for demonstration. Empirical results and discussions are shown in Section 3. Finally, Section 4 concludes and gives some insights for practical applications.

2. Definitions and properties

Let N be the number of stocks in a market and consider random vector $\mathbf{x} = (x_i)_{i=\overline{1,N}}$ where x_i is the daily log-return of stock i, $x_i(t) = logP_i(t+1) - logP_i(t)$, where $P_t(t)$ represents the price of stock price at time t. We standardize $\mathbf{x}$ by $\mathbf{x}^* = \left(x_i^*\right)_{i=\overline{1,N}}$ where $x_i^* = \frac{x_i}{\sigma_i}$ with σ_i denotes the standard deviation of stock i's log-return.

The $N \times N$ cross-correlation matrix $\mathbf{C} = \left(c_{ij}\right)_{i,j=\overline{1,N}}$ of $\mathbf{x}$ is also the covariance matrix of $\mathbf{x}^*$, i.e

$$c_{ij} = \text{Cov}\left(x_i^*, x_j^*\right), \qquad i,j = \overline{1,N}. \tag{1}$$

Please cite this article as: Q. Nguyen and N.K.K. Nguyen, Composition of the first principal component of a stock index — A comparison between SP500 and VNIndex, Physica A (2019), https://doi.org/10.1016/j.physa.2019.04.216.

Principle components (with notations as presented in [23]). Let $\lambda_i, i = \overline{1,N}$ be eigenvalues of the cross-correlation matrix **C**. With financial data, we suppose that the rank of matrix **C** equals the number of its columns. Therefore, one gets N distinct eigenvalues which are ordered from λ_1 to λ_N decreasingly.

In the sequence, let $\mathbf{U}_i$ be the eigenvector associated with eigenvalue λ_i such that $\|\mathbf{U}_i\|$, defined as the scalar product $\mathbf{U}_i'\mathbf{U}_i$, equals 1. Here, $'$ denotes transpose.

Let **A** be the orthogonal matrix whose ith column is the ith eigenvector $\mathbf{U}_i$.

Let $\mathbf{z} = \mathbf{A}'\mathbf{x}^*$, it is easy to see that the ith component of vector **z** is:

$$z_i = \mathbf{U}_i'\mathbf{x}^*, \quad i = \overline{1,N}. \tag{2}$$

z_i is called the ith principle component (PC) of **x**.

From [23], we mention other important properties of PCs that will be used in this work:

- For all principle components, the variance of each of them equals the corresponding eigenvalue. i.e.

$$\mathrm{Var}\,(z_i) = \lambda_i, \quad i = \overline{1,N}. \tag{3}$$

- For all linear transformations $\mathbf{U}\mathbf{x}^*$ of vector $\mathbf{x}^*$ such that **U** is a unit vector, the first principle component z_1 is the one that has the largest variance, i.e.

$$\max_{\|\mathbf{U}\|=1} \mathrm{Var}\left(\mathbf{U}\mathbf{x}^*\right) = \mathrm{Var}\,(z_1)\,. \tag{4}$$

- For all non-zero linear transformations $\mathbf{U}\mathbf{x}^*$ of vector $\mathbf{x}^*$, the first principle component z_1 is the one that maximize the sum of square of the Pearson correlation coefficients with each of stock returns, i.e.

$$\max_{y=\mathbf{U}\mathbf{x}^*, y\neq 0} \sum_{i=1}^{N} \left(r_y^i\right)^2 = \sum_{i=1}^{N} \left(r_{z_1}^i\right)^2 \tag{5}$$

  where r_y^i and $r_{z_1}^i$ is the Pearson correlation coefficients between random variables y and z_1 with x_i, respectively.

Spectral decomposition of cross-correlation matrix **C** *(see proof in [22]).* :

$$\mathbf{C} = \sum_{i=1}^{N} \lambda_i \mathbf{U}_i \mathbf{U}'_i \tag{6}$$

In addition, the component $u_{ij}\left(j = \overline{1,N}\right)$ of eigenvector $\mathbf{U}_i$ is called a loading of the ith PC, which is equal to the correlation coefficient between x_i and the jth PC divided by $\sqrt{\lambda_i}$.

For empirical illustration, we collect data of the SP500 index component stocks from the US market and the VNIndex component stocks from the Vietnamese stock market, for a period of 5 years from 2013-01-01 to 2017-12-31. The data of the VNIndex are taken from 'CafeF' (http://cafef.vn) and the data of the NYSE are from 'Yahoo Finance' (http://finance.yahoo.com). We limit to the companies that have at least 60% of trading days during that period, and an average daily volume of higher than 1000 shares/day. In consequence, we have 482 and 274 stocks for the SP500 and VNIndex, respectively. However, there are still days when there was no negotiation of a stock. In such case, the log-return of that day and the next day will not be defined. The pairwise correlation coefficients between any two stocks is computed using only complete pairs of log-return on those stocks. (Using this method, the resulting cross-correlation matrix **C** could be not positive semi-definite, however, the data will be more reliable)

3. Empirical results and discussions

3.1. The elements of $\mathbf{U}_1$ and the 1st principle component

We compute the sampling correlation matrix **C**, the eigenvalues $\lambda_1 \ldots \lambda_N$, the eigenvectors $\mathbf{U}_1 \ldots \mathbf{U}_N$ and the sampling PCs according to the method described in the previous section.

Fig. 1a illustrates the distribution of correlation coefficients of **C** and its eigen-spectrums (inserted) for SP500. Correlation coefficients are rather positive and distributed around a mean of 0.283. As found in other previous studies, there exists the largest eigenvalue of value 144 (upper inserted), which is an order of magnitude higher than the others. There are also several large eigenvalues that are separated from the bulk of eigenvalues bounded by values λ_- and λ_+, predicted by the RMT theory (lower inserted). In contrast, the correlation coefficient distribution of VNIndex is narrower with smaller mean (0.082) and more negative value as seen in Fig. 1b. As a result, the largest eigenvalue is of 33.1. Furthermore, there are fewer large eigenvalues other than the largest one which deviates from the RMT bulk.

We now focus on the elements $\left\{u_{1j}\right\}$ of the 1st eigenvector $\mathbf{U}_1$ corresponding to the largest eigenvalue and the first PC $z_1 = \mathbf{U}_1'\mathbf{x}^*$.

We plot the distribution of $\left\{u_{1j}\right\}$ together with the corresponding stock's market capitalization in a two-dimensional graph for both SP500 and VNIndex, as shown in Fig. 2. For a better comparison of $\left\{u_{1j}\right\}$ between two markets, we scale

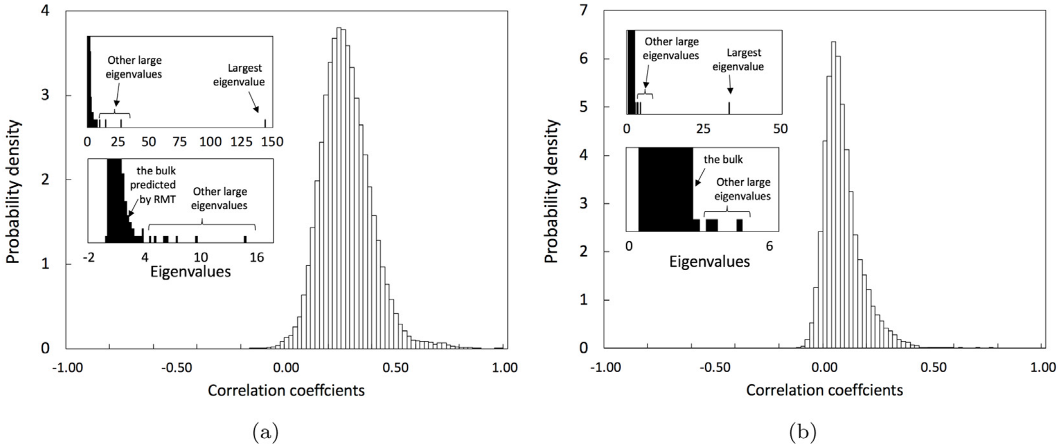


Fig. 1. Probability density of the cross-correlation matrix **C** and their eigen spectrums (upper insert: all eigenvalues; lower insert: all eigenvalues excluding the largest) for (a) SP500 and (b) VNIndex.

Fig. 2. The components $\{u_{1j}\}$ (scaled by $\sqrt{\lambda_1}$) against market capitalization for individual stocks j in the SP500 (US) and VNIndex (VN).

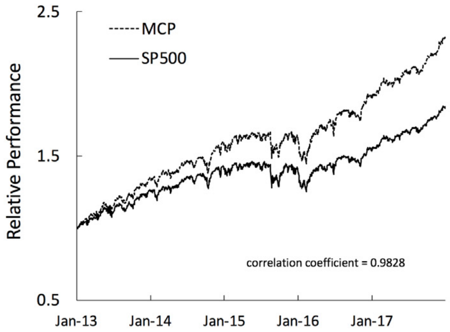
those components by $\sqrt{\lambda_1}$ of their corresponding market. We found that although the range of $\{u_{1j}\}$ for SP500 is small in comparison to that of other eigenvectors (not shown), it is not constant even if we account for statistical noise due to the finite time T. The range of $\{u_{1j}\}$ for VNIndex is clearly broader, suggesting that they are not statistically equal. Note that u_{1j} can be negative if the correlation between the corresponding stock and the z_1 is negative (or the stock is negatively correlated with many other stocks in the market). For SP500 we do not find any negative value during the analyzing time-frame, but we found one stock with negative u_{1j} for the VNIndex (ticker SSC). This is a stock with very low liquidity, and has many consecutively non-trading days (but still satisfy the condition of having over 60% of trading days during the analyzing period,).

Furthermore, we found that even though the market capitalization change over 3 orders of magnitude among the SP500 stocks (and 5 orders of magnitude among the VNIndex stocks), no significant trend of $\{u_{1j}\}$ with respect to market capitalization was found. Furthermore, the top $\{u_{1j}\}$ stocks are rather mid-cap and are financial service stocks in both markets. However, this observation may depends on the index's composition.

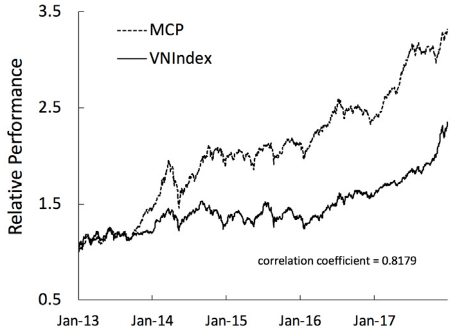
About the principal component z_1, according to the properties described in Section 2, z_1 is the highest variance and therefore the riskiest portfolio that one can build, given the unit norm condition. It is also the variable that has the highest sum of squared correlation coefficients with all other stock returns and therefore, it correlates the best with the overall market. As a matter of fact, [22] suggested that the 1st PC, should be called the Most Correlated Components and in the financial framework, we suggest replacing the name "market mode" or "market portfolio" with the "Most Correlated Portfolio" (MCP).

We simulate the equity performance of the MCP and compare to that of the Indices, as shown in Fig. 3. By the second property of the 1st PC given above (4) , MCP is the variable that has the highest variance among the linear transformations

Please cite this article as: Q. Nguyen and N.K.K. Nguyen, Composition of the first principal component of a stock index — A comparison between SP500 and VNIndex, Physica A (2019), https://doi.org/10.1016/j.physa.2019.04.216.



(a)



(b)

Fig. 3. Relative performance of simulated most correlated portfolio (blue) vs. Index from 2013 to the end of 2017: (a) SP500 and (b) VNIndex. The correlation coefficient between the MCP and Index is of 0.9828 for US and 0.8179 for VN.

of the original vector **x** (under the unit norm condition of the transforming vector). That means MCP is the riskiest portfolio in the market that we can have given the unit norm condition. From the portfolio management point of view, this portfolio should be avoid (or reduced) if one is risk-averse and/or seeks for beta-neutral strategy, but this portfolio could be considered if one is bench-marked by the overall market performance. Our result shows that for the SP500, the MCP portfolio is quite close to the SP500 index, with a correlation of 0.982. However, this correlation coefficient is only 0.8179 for the Vietnamese market. In another word, their is a significant deviation between Most Correlated Portfolio and the Index portfolio, which is market capitalization weighted. This result can also be predicted given the broad distribution of $\mathbf{U}_1$ as shown in Fig. 2. In conclusion, for an emerging market like the Vietnamese one, beside the Index, investors should also consider the MCP, which has high degree of correlation with any diversified portfolio, and does not always match the movement of the Index.

Another important property of the cross-correlation matrix PCs is that they are scale invariant: the PC portfolios are unchanged if the off-diagonal elements are multiplied by a positive factor: if there is a sudden change in the market resulting in a shift in the correlation matrix, then the Most Correlated Portfolio is unchanged. Meredith and Millsap [24] derived Property A6 in [22] independently and noted that optimizing the multiple correlation criteria (or sum of squared covariances) gives a scale-invariant method (as does Property A5 in [22]).

3.2. Estimation of elements of $\mathbf{U}_1$ (the loadings of the 1st PC)

One question is how we estimate the elements of $\mathbf{U}_1$, the loadings of the 1st PC, and the MCP's weights. In fact, each u_{1j}, a loading of $\mathbf{U}_1$, is the correlation coefficients between z_1 and stock j divided by a factor of $1/\sqrt{\lambda_1}$. By property 3 of the PCs that we mention in previous section, z_1 has highest sum square of correlation coefficients with all stocks. Given that most correlation coefficients are positive, intuitively, the stock that correlates well with other stocks in the market will consequently have a high correlation with z_1, and therefore has a higher value in the components of eigenvector $\mathbf{U}_1$.

However, this property does not help: u_{1j} are proportioned to the correlations of the z_1 and stocks, but in order to calculate the z_1 we do need those weights. In [21], it was found empirically that the loadings of the 1st PC are approximately linearly depending on the average correlation coefficients of each stock. Using the spectral decomposition of correlation matrix **C** (6), here we will demonstrate this approximation from the Principle Components point of view.

We calculate $\bar{c}_i$, the average of correlation coefficients between stock j and others as follows:

$$\bar{c}_i = \frac{1}{N-1}\sum_{j\neq i} c_{ij} \tag{7}$$

From the 4th property of the PCs and by taking into account the observed eigen-spectrum that λ_1 is one order of magnitude larger than other eigenvalues, and given all eigenvectors having a unit norm, we suppose that the 1st term of the right hand side of (6) will dominate all other terms and therefore:

$$c_{ij} = \sum_{k=1}^{N} \lambda_k u_{ki} u_{kj} \approx \lambda_1 u_{1i} u_{1j} \tag{8}$$

by summing over all stocks, we have

$$\bar{c}_i = \frac{1}{N-1}\sum_{j\neq i} c_{ij} = \frac{1}{N-1}(\sum_{j=1}^{N} c_{ij} - 1) \approx (\lambda_1/N)(\sum_{j=1}^{N} u_{1j})u_{1i} \tag{9}$$

where in (9) we use (8) and the fact that $N \gg 1$ and $c_{ii} = 1$.

Please cite this article as: Q. Nguyen and N.K.K. Nguyen, Composition of the first principal component of a stock index — A comparison between SP500 and VNIndex, Physica A (2019), https://doi.org/10.1016/j.physa.2019.04.216.

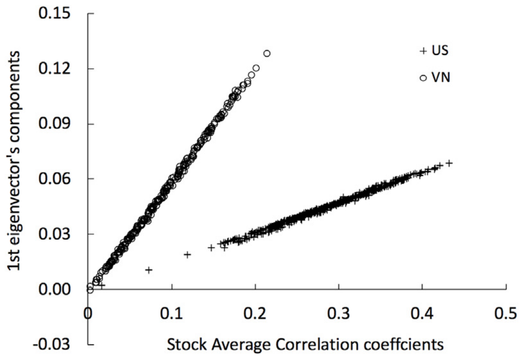
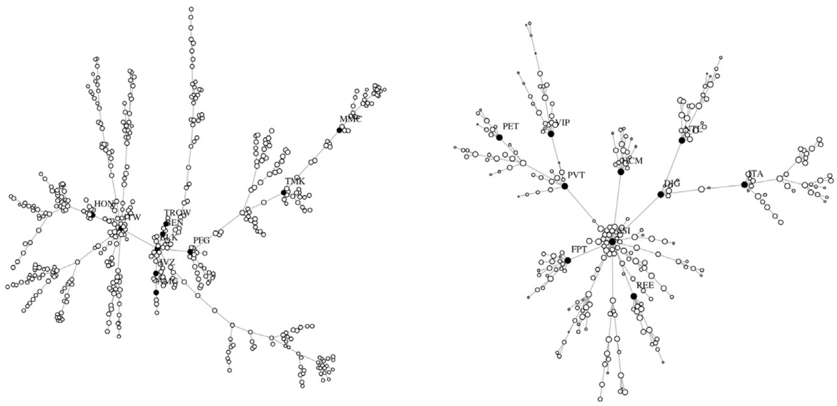


Fig. 4. Relationship between the first eigenvector's components u_{1i}'s and the average correlation coefficient $\overline{c}_i$ of the corresponding stocks for SP500 and VNIndex.



(a) (b)

Fig. 5. Topology of the MST with node size as log value of the components of 1st eigenvector for (a) SP500 Index and (b) VNIndex (tickers of the stocks in the top 10 components are shown and their corresponding nodes' colors are in black). One found that high 1st components stocks are also central nodes of the MST. However, when considering the other node's size there are difference in two market: sizes are relatively uniform in the US tree and more diverse in the VN tree. It does agree with the results in Section 3.1.

In such case, we will have an approximate linear dependence between $\overline{c}_i$ and u_{1i}. Fig. 4 shows the empirical results of this dependence for the SP500 and VNIndex. A good linear fit is found between u_{1i} and the mean of correlation coefficients of stock i with all other stocks for both indices. As we supposed previously, a stock that has high (positive) correlation with other stocks will have large element's value in $\mathbf{U}_1$ and a high weight in MCP. This phenomenological relation was found in the Vietnamese market previously [21], and to our knowledge, is the first time reported for the US market in this work.

3.3. The implication for the structure of the minimum spanning tree (MST) network

One of the common methods to analyze the topography of the stock market is the minimum spanning tree (MST) constructed from $\mathbf{C}$ [25–31] for some examples. The tree's nodes represent stocks and its links are placed from the highest correlation coefficient c_{ij}, in the condition that no loop is found. In consequence, a loop-free network is established with N nodes and $N-1$ links.

Due to the MST's construction algorithm, stocks that have high (positive) correlations with other stocks will have a tendency to own more connections and become hubs in the MST. We confirm this by plotting Fig. 5 describing the MST with node sizes as an increasing function of the components of $\mathbf{U}_1$.

We found that the most weighted stocks in MCP are also highly connected ones in MST. Those stocks play an important role in a breakdown process of a network [32], and are worth to protect firstly. This observation gives us another view about the 1st PC; it contains the most information about the MST structure. One can directly analyze the overall temporal dynamic of MST by that of the 1st PC. Last but not least, PCs analysis has more information than MST analysis because it conserves the whole data of a cross-correlation matrix. Other eigen modes excluding the largest ones also contain rich information and will be considered in our future works.

4. Conclusions

We re-examine the market mode constructed from the eigenvector associated with the largest eigenvalue of a cross-correlation matrix of stock returns. We insist that those components are not uniformly distributed as being suggested by previous studies but depend on the average correlation of the individual stocks in the remaining market. Therefore, the 1st principle component portfolio constructed from the 1st eigenvector should be called the Most Correlated Portfolio instead of market portfolio or market mode. We found that the highest weighted stocks in this portfolio are usually financial services firms (brokerage or investment advisory), in both the US and Vietnamese market. While the US Most Correlated Portfolio performs similarly to the market Index SP500 (therefore closes to the market portfolio), that of Vietnam does not. In the latter case, one of the conclusion is that investing in the index portfolio may not result in a performance that represents the market as a whole. Finally, we suggest that a deep analysis of the components of the MPC could provide an alternative way to study the strong common interaction between components in a complex system such as the financial market. In particular, their temporal dynamic and feedback can shed light to the mechanism and evolution of a financial crisis [33].

Acknowledgments

We would like to acknowledge Dr. V. L. Nguyen for helpful discussions and Dr. D. H. Tong for reviewing the manuscript. We are also grateful to the three anonymous referees for showing us numerous suggestions that have helped us to revise the manuscript carefully. N.K.K. Nguyen is supported by Vietnam National University HoChiMinh City (VNU-HCM) under grant number B2017-42-01.

References

[1] J.P. Bouchaud, M. Potters, Financial applications of random matrix theory: a short review, in: The Oxford Handbook of Random Matrix Theory, Oxford, 2009, pp. 824–850.
[2] M. Potters, J.P. Bouchaud, L. Laloux, Financial applications of random matrix theory: Old laces and new pieces, Acta Phys. Polon. B 36 (2005) 2767–2784.
[3] B. Rosenow, V. Plerou, P. Gopikrishnan, L.A.N. Amaral, H.E. Stanley, Application of random matrix theory to study cross-correlations of stock prices, Int. J. Theor. Appl. Finance 3 (2000) 399–403.
[4] A. Utsugi, K. Ino, M. Oshikawa, Random matrix theory analysis of cross correlations in financial markets, Phys. Rev. E 70 (2004) 026110.
[5] D.Y. Kenett, Y. Shapira, E. Ben-Jacob, RMT Assessments of the market latent information embedded in the stocks' raw, normalized, and partial correlations, J. Probab. Statist. (2009).
[6] V. Plerou, P. Gopikrishnan, B. Rosenow, L. Nunes Amaral, H. Stanley, Universal and nonuniversal properties of cross correlations in financial time series, Phys. Rev. Lett. 83 (1999) 1471–1474.
[7] V. Plerou, P. Gopikrishnan, B. Rosenow, L.A.N. Amaral, T. Guhr, H.E. Stanley, Random matrix approach to cross correlations in financial data, Phys. Rev. E 65 (2002) 066126.
[8] J. Shen, B. Zheng, Cross-correlation in financial dynamics, Europhys. Lett. 86 (2009) 48005.
[9] L. Lalouxs, J.P. Bouchaud, Random matrix theory and financial correlations, Int. J. Theor. Appl. Finance 3 (2000) 391–397.
[10] L. Laloux, P. Cizeau, J.P. Bouchaud, M. Potters, Noise dressing of financial correlation matrices, Phys. Rev. Lett. 83 (1999) 1467–1470.
[11] T. Qiu, B. Zheng, G. Chen, Financial networks with static and dynamic thresholds, New J. Phys. 12 (043057) (2010).
[12] P. Gopikrishnan, B. Rosenow, V. Plerou, H.E. Stanley, Quantifying and interpreting collective behavior in financial markets, Phys. Rev. E 64 (2001) 035106.
[13] R.K. Pan, S. Sinha, Collective behavior of stock price movements in an emerging market, Phys. Rev. E 76 (2007) 046116.
[14] Q. Nguyen, One-factor model for the cross-correlation matrix in the Vietnamese stock market, Physica A 392 (2013) 2915–2923.
[15] S. Drożdż, F. Grümmer, A.Z. Górski, F. Ruf, J. Speth, Dynamics of competition between collectivity and noise in the stock market, Physica A 287 (3-4) (2018) 440–449.
[16] T. Preis, D.Y. Kenett, H.E. Stanley, D. Helbing, E.B. Jacob, Quantifying the behavior of stock correlations under market stress, Sci. Rep. 2 (752) (2012).
[17] Z. Zheng, B. Podobnik, L. Feng, B. Li, Changes in cross-correlations as an indicator for systemic risk, Sci. Rep. 2 (888) (2012).
[18] D.J. Fenn, M.A. Porter, S. Williams, M. McDonald, N.F. Johnson, N.S. Jones, Temporal evolution of financial-market correlations, Phys. Rev. E 84 (026109) (2011).
[19] L. Sandoval, I.D.P. Franca, Correlation of financial markets in times of crisis, Physica A 391 (2012) 187–208.
[20] M. Billio, M. Getmansky, W.A. Lo, L. Pelizzon, Measuring systemic risk in the finance and insurance sectors, NBER working paper, 16223.
[21] H.T. Nguyen, P.N.U. Tran, Q. Nguyen, An analysis of eigenvectors of a stock market cross-correlation matrix, in: Econometrics for Financial Applications. ECONVN 2018. Studies in Computational Intelligence, Vol. 760, Springer, Cham, 2018, pp. 504–513.
[22] J. Cadima, A scale-invariant component analysis: MCCA, Technical report 4, Departamento de Matem atica, Instituto Superior de Agronomia, Universidade Tecnica de Lisboa.
[23] I.T. Jolliffe, Principal Component Analysis.
[24] W. Meredith, R.E. Millsap, On component analysis, Psychometrika 50 (1985) 495–507.
[25] G. Bonanno, F. Lillo, R.N. Mantegna, High-frequency cross-correlation in a set of stocks, Quant. Finance 1 (2001) 96–104.
[26] J. Onnela, A. Chakraborti, K. Kaski, Dynamics of market correlations: Taxonomy and portfolio analysis, Phys. Rev. E 68 (056110) (2003) http://dx.doi.org/10.1103/PhysRevE.68.056110.
[27] C. Coronnello, M. Tumminello, F. Lillo, S. Micciche, R.N. Mantegna, Sector identification in a set of stock return time series traded at the london stock exchange, Acta Phys. Polon. B 36 (9) (2005) 2653–2679.
[28] A.Z. Gorski, S. Drozdz, J. Kwapien, Scale free effects in world currency exchange network, Eur. Phys. J. B 66 (2008) 91.
[29] J. Wooseok, L. Junghoon, C. Woojin, Currency crises and the evolution of foreign exchange market: Evidence from minimum spanning tree, Physica A 390 (2011) 707.
[30] G.J. Wang, C. Xie, F. Han, B. Sun, Similarity measure and topology evolution of foreign exchange markets using dynamic time warping method: Evidence from minimal spanning tree, Physica A 391 (2012) 4136.

[31] G. Bonanno, G. Caldarelli, F. Lillo, S. Micciche', N. Vandewalle, R.N. Mantegna, Networks of equities in financial markets, Eur. Phys. J. B 38 (2004) 363–371.

[32] N.K.K. Nguyen, Q. Nguyen, Resilience of stock cross-correlation network to random breakdown and intentional attack, in: Econometrics for Financial Applications. ECONVN 2018. Studies in Computational Intelligence, Vol. 760, Springer, Cham, 2018, pp. 553–561.

[33] S. Begusic, Z. Kostanjcar, D. Kovac, H.E. Stanley, B. Podobnik, Information feedback in temporal networks as a predictor of market crashes, to be appeared in Complexity.

Please cite this article as: Q. Nguyen and N.K.K. Nguyen, Composition of the first principal component of a stock index — A comparison between SP500 and VNIndex, Physica A (2019), https://doi.org/10.1016/j.physa.2019.04.216.