Application of Random Matrix Theory in analyzing the Global Financial Indices during the 2008 Financial Crisis

Sood, Pranay (20239217)

Final Project Report
ELEC 5450: Random Matrix Theory and Applications

Department of Electronic and Computer Engineering
The Hong Kong University of Science and Technology

Abstract

This report focuses on the application and implementation of Random Matrix Theory (RMT) in analyzing and interpreting the global financial indices during the 2008 financial crisis. Using RMT, it is possible to know the empricial eigenvalue distribution and its corresponding distribution of the eigenvector components. As a result, comparisons will be made to analyze the financial data in three different periods which are before, during and after the 2008 financial crisis. Furthermore, new contributions will be introduced and implemented to analyze, clean and provide an interpretation and comparison of the financial data from the three different periods of the financial crisis.

I. INTRODUCTION

Financial markets are complicated systems wherein statistical approaches and concepts from mathematics and physics have been used to actively understand and investigate price trends in financial time series. The availability and access to large financial data has made it possible to apply and examine various techniques in order to understand and make comparisons between data and theory [1]. This has allowed mathematicians and physicists to introduce and apply the concept of RMT, in order to analyze and interpret the complex of nature of these financial systems. RMT makes it possible to evaluate and deal with the statistics behind the eigenvalues and eigenvectors of cross-correlation matrices of complex and disordered systems [1]. Given that we are dealing with large financial data, covariance-correlation matrices are of fundamental importance because it allows us to understand the relation between the many random variables that are present [2]. The unpredictable nature of the financial time series makes it difficult to extract the huge amount of economic information from it [3]. Through this RMT approach, it has made it possible to investigate cross-correlations that are present in the financial time series.

RMT was developed in the context of complex quantum systems where the precise nature of the interactions between the sub-units are unknown [4]. When dealing with large financial data and many financial markets, we will notice that there will be significant deviations from the RMT prediction due to the identification of non-random properties that are present in the complex system. Due to the presence of these deviations, we can use it to analyze the underlying interactions that are present in the financial time series [1]. The largest eigenvalue and its corresponding eigenvector components highlight the presence of a collective response that results in a stimulus in the market along with the financial time series. It is denoted that during the time of a financial crisis, the largest eigenvalue is significantly higher than that of any other period thus implying that there strong interactions present among the financial indices [5].

The infamous 2008 subprime mortgage global financial crisis is recognized as one of the worst financial crises since the Great Depression in 1930s. The crisis originated in the United States of America (USA) due to the debacle of the housing bubble in 2007 which was triggered by the default of a large number of mortgages [6]. Subprime mortgages were sold to borrowers with low credit scores. Since they have low credit scores, interest rates on subprime products were low but what borrowers did not realize that interest rates on such products were adjustable. This meant that for an initial amount of time, borrowers would have to pay low interest rates, but after that interest rates would rise significantly. This made it difficult for borrowers to repay their mortgages and eventually most of them defaulted on their loans which led to a high number of foreclosures [6]. Furthermore, these loans were transformed into pools known as Collateralized Debt Obligations (CDOs) which were purchased by many interested and foreign investors due to its high return and low risk thus creating a financial bubble. Once the default rates began to rise, the value of the CDOs dropped drastically and the foreign investors lost high amounts of money which resulted in their home countries getting affected in a negative manner. As a result, several financial institutions such as Lehman Brothers, Merill Lynch etc. experienced severe losses that diminished their financial strength. The interconnections between the financial institutions and global financial system resulted in the spreading of the crisis to the world.

This report is organized as followed with Section II focusing on the methodology of RMT implemented by the author of paper [5] in understanding the effect of the 2008 financial crisis by using the 30 global financial indices as the financial data and making comparisons of the three different periods which comprise of before, during and after the 2008 financial crisis. I will be generating the simulations done by the author of paper [5] with the 30 global financial indices data. Section III will focus on the critical analysis of paper [5] mainly focusing on the assumptions and arguments made by the author. Section IV will focus on my new contribution to the paper [5]. Firstly, I will be incorporating additional methodologies and concepts of RMT from the paper [4] in order to provide more perspective, analysis and interpretation of the financial data. Next, I will also implement the ”market mode removal” method as discussed in [4] and [2] in order to construct a cleaner correlation matrix of the financial data. Finally, I will also implement the concept of Markowitz Mean-Variance Portfolio Theory as discussed in [7] to design a portfolio of the global financial indices from different countries in order to determine the variance of daily returns of these portfolios during the three different periods of the 2008 financial crisis. In addition, I have also done all the simulations in Section IV. Finally, Section V will discuss the conclusion and proposed future plan of the report.

II. Discussion of Methodology Implemented and Results of [5]

In paper [5], the author has decided to implement RMT to identify the level of interactions and fluctuations between the Global Financial Indices. The author discusses the empirical eigenvalue distribution and its corresponding eigenvector components distribution of the global financial indices in three different periods which are before, during and after the 2008 financial crisis. In the following subsections, the methodology discussed in [5] is presented along with the author’s interpretation and understanding of the results.

A. Methodology

1) ***Data Analyzed****:* The data analyzed are the daily closing prices of the global financial indices of the 30 financial markets of the world from the period January 1, 2006, to December 31, 2011. The global Financial Indices used are as followed: DOW JONES (United States); S&P500 (United States); NASDAQ (United States); SPTSX (Canada); MEXBOL (Mexico); IBOV (Brazil); IPSA (Chile); COLCAP (Colombia); MERVAL (Argetina); SPBLPGPT (Peru); SX5E (Europe); UKX (United Kingdom); CAC (France); DAX (Germany); IBEX (Spain); FTSEMIB (Italy); AEX (Netherlands); OMX (Sweden); SMI (Switzerland); RTSI$ (Russia); NKY (Japan); HSI (Hong Kong); SHSZ300 (China); AS51 (Australia); KOSPI (South Korea); NIFTY (India) TWSE (Taiwan); JCI (Indonesia); FBMKLCI (Malaysia); STI (Singapore). The data has been collected from the Bloomberg Terminal. Some of the global financial indices collected for this report are different compared to the ones collected by the author of paper [5] however, it has been ensured that the same methodology as mentioned in paper [5] has been applied on all the above-mentioned global financial indices. The financial data has been divided into three periods whereby the period from January 1, 2006 to December 31, 2007 is considered as before the crisis, from January 1, 2008 to December 31, 2009 is considered as during the crisis and, from January 1, 2010 to December 31, 2011 is considered as after the crisis. Given that if a particular country is experiencing a holiday on a particular date and the market is closed, the entire day for all the global financial indices will be removed. As a result, the number of days the market was open before the crisis is 302 days, during the crisis is 301 days and after the crisis is 298 days.

2) ***Prices of Financial Indices as Stochastic Variables****:* Before examining the three different periods of the financial crisis, it is important for us to first characterize and understand the correlation properties of the financial data [8]. Let’s denote that $S_i(t)$ and $R_i(t)$ denote the closing prices and returns of the financial indices i (= 1, ..., N) at time t (= 1, ..., T) respectively. Therefore, the logarithmic return is defined as:

$$R_i(t) \equiv \ln(S_i(t+\Delta t)) - \ln(S_i(t)) \tag{1}$$

where the time lag is $\Delta t = 1$. Hence, the normalized logarithmic return of index i is defined as:

$$r_i(t) \equiv \frac{(R_i(t) - \langle R_i \rangle)}{\sigma_i} \tag{2}$$

where $\sigma_i \equiv \sqrt{\langle R_i^2 \rangle) - \langle R_i \rangle^2}$ is the standard deviation of $R_i(t)$ and $\langle \dots \rangle$ defines the time average over the period T. Therefore, the normalized return $r_i(t)$ has a zero mean and an unit variance.

*3) **Correlation Matrix**:* Let's denote $\boldsymbol{W}$ as a $N \times T$ rectangular matrix with the elements $\{W_{it} \equiv r_i(t)\}$ and $\boldsymbol{W}^T$ is the transpose of $\boldsymbol{W}$ [8]. Therefore, the equal time cross-correlation matrix $\boldsymbol{C}$ is defined as:

$$\mathbf{C} = \frac{1}{T}\mathbf{W}\mathbf{W}^T \tag{3}$$

$\boldsymbol{C}$ is a symmetric $N \times N$ with the diagonal elements equalling to a unity with the elements being:

$$C_{ij} \equiv \langle r_i(t) r_j(t) \rangle \tag{4}$$

C_{ij} represents the cross-correlation coefficients with the domain varying from: $-1 \leq C_{ij} \leq 1$ whereby, $C_{ij} = 1$ means perfect correlations, $C_{ij} = 0$ means that the pair of financial indices are uncorrelated, and $C_{ij} = -1$ means perfect anti-correlations [4]. Paper [5] constructs a probability density of $\boldsymbol{C}$ and also examines the average and standard deviation of the cross correlation coefficients of the log returns of the three different periods.

*4) **Random Matrix Theory Approach - Eigenvalue Distribution of the Correlation Matrix** :* Now let's say that $\{r_i(t)\}$ is replaced with sequences of random elements with zero mean, unit variance that are mutually uncorrelated as done in paper [4]. Therefore, a random cross-correlation matrix $\boldsymbol{R}$ that corresponds to $\boldsymbol{C}$ for real data is obtained which is defined as:

$$\mathbf{R} = \frac{1}{T}\mathbf{W}'\mathbf{W}'^T \tag{5}$$

where $\boldsymbol{W}'$ as a $N \times T$ matrix that comprises of columns of time series with zero mean and unit variance that are mutually uncorrelated. Now let us consider the asymptotic behaviour of the singular values that are present in large rectangular matrices. This is where the MP Law distribution (RMT Prediction) is introduced which is concerned with the emergent behaviours of large classes of random matrices in the asymptotic limit whereby the dimensions of the matrix tend to infinite [2]. Based from the statistical properties of RMT whereby in the limit $N \rightarrow \infty$ and $T \rightarrow \infty$, $Q = \frac{T}{N}$ (≥ 1) remains finite hence, the probability density function $\rho_{rm}(\lambda)$ of eigenvalues λ of random correlation matrix $\boldsymbol{R}$ is denoted by:

$$\rho_{rm}(\lambda) = \frac{Q}{2\pi}\frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda} \tag{6}$$

whereby λ ranges from: $\lambda_- \leq \lambda_i \leq \lambda_+$, given that λ_- and λ_+ represent the minimum and maximum eigenvalues of $\boldsymbol{R}$. This is given by:

$$\lambda_{\pm} = \left(1 \pm \frac{1}{\sqrt{Q}}\right)^2 = 1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}} \tag{7}$$

this limits the interval where probability density function differs from zero. In paper [5], the eigenvalues λ of the financial data that are being analyzed, corresponds to the global financial index. The empirical eigenvalue distribution of the global financial indices is denoted by $\rho(\lambda)$. On that basis, paper [5] determine the minimum (λ_{min}) and maximum (λ_{max}) eigenvalues before, during and after the 2008 financial crisis in order to interpret the correlation among the global financial indices.

*5) **Distribution of the Eigenvector Components:*** Let's denote that $\rho(u)$ is the distribution of the components of eigenvector u^k of cross-correlation matrix $\boldsymbol{C}$. The author of paper [5] has simply displayed the largest and second largest eigenvectors that correspond to the largest and second largest eigenvalues of the 30 global financial indices before, during and after the 2008 financial crisis.

*6) **Inverse Participation Ratio:*** When there is an increase in separation from the RMT upper bound λ_+, there is an increasing effect of randomness [4]. Therefore, the Inverse Participation Ratio (IPR) allows the quantification of the number of components that are participating significantly in each eigenvector thus indicating the degree of deviation from the RMT result in regards to the distribution of the eigenvector components. IPR of eigenvector u^k is denoted by:

$$I^k = \sum_{l=1}^{N} [u_l^k] \tag{8}$$

where u_k^l, l (= 1, ..., N), are the eigenvector components u^k [5]. Paper [5] has compared the IPR of the three different periods of the crisis in order to determine that which period of the crisis has a lower IPR.

B. Results

Upon following the procedures mentioned by author of paper [5], the following results were obtained with Figure 1 shows the probability density of cross-correlation matrix $\boldsymbol{C}$:

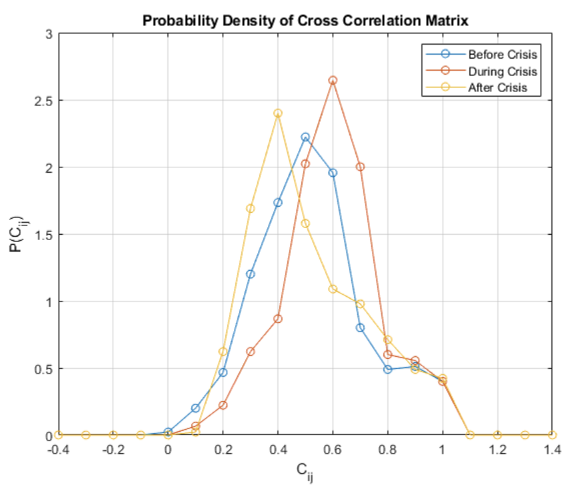

Fig. 1. Probability Density of the cross-correlation matrix $\boldsymbol{C}$ determined by using daily returns of 30 global financial indices before, during and after the 2008 financial crisis [5].

The average values of the cross-correlation coefficients before crisis is 0.5275, during the crisis is 0.5956, and after the crisis is 0.5171 of. The standard deviation of the cross-correlation coefficients before crisis is 0.2023, during the crisis is 0.1773, and after the crisis is 0.2089 of the global financial indices. Based from Figure 1 and the average and standard deviation of the cross-correlation coefficients, the author of paper [5] states that during the period of the financial crisis, the distribution of the cross-correlation coefficients is narrower with the average value being the highest and there is a decrease in the standard deviation thus being the lowest. The next results display the empirical eigenvalue distribution which is denoted by $\rho(\lambda)$. Given, $N = 30$ and $T = 301$ (before crisis), $T = 300$ (during crisis) and, $T = 297$ (after crisis) hence, $Q = \frac{301}{30} = 10.0333$ (before crisis), $Q = \frac{300}{30} = 10$ (during crisis) and, $Q = \frac{297}{30} = 9.9$ (after crisis). In addition, the values of $\lambda_+ = 1.7311$ and $\lambda_- = 0.4683$ (before crisis), $\lambda_+ = 1.7325$ and $\lambda_- = 0.4675$ (during crisis) and, $\lambda_+ = 1.7367$ and $\lambda_- = 0.4654$ (after crisis). The author of paper [5] states that the factor of Q and the values of λ_+ and λ_- indicate that the uncorrelated time series has eigenvalues that will be bounded between λ_+ and λ_- for the three different periods.

Figures 2 represents the empirical eigenvalue distribution $\rho(\lambda)$ of cross-correlation matrix $\boldsymbol{C}$ from all three different periods of the crisis. The minimum (maximum) eigenvalues, $\lambda_{min(max)} = 0.0087(16.3498)$ before the crisis, $\lambda_{min(max)} =$

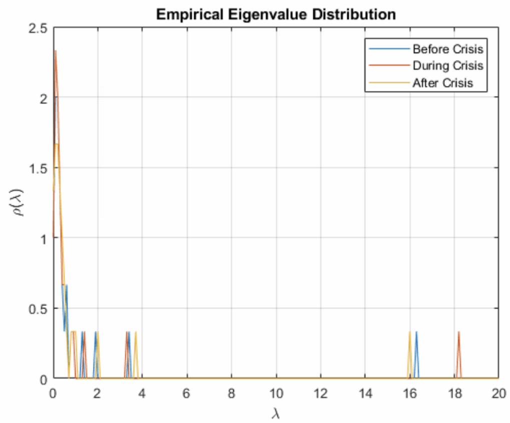

Fig. 2. Empirical eigenvalue distribution $\rho(\lambda)$ of cross-correlation matrix $\boldsymbol{C}$ from all three different periods of the crisis.

0.0049(18.2111) during the crisis, and $\lambda_{min(max)} = 0.0034(16.0497)$ after the crisis. Based from paper [5], the eigenvalue is highest during the crisis thus implying that there is a strong correlation between the global financial indices during the crisis. Largest eigenvalue after the crisis is the lowest which suggests that crisis has been overcome and hence indicating that the correlation between the global financial indices is less stronger. After the crisis has a lower largest eigenvalue compared to the before the crisis largest eigenvalue because the global market has become more stable and less interaction between the global financial indices as compared to before the crisis.

The author of paper [5] also compared the largest and second largest eigenvector components corresponding to the largest and second largest eigenvalues of the global financial indices before, during and after the financial crisis. According to the author on the basis of Figures 3 and 4, it is noted that the components of the eigenvector that carry negative sign change to positive sign after the crisis and components of eigenvector that carry positive sign during the crisis also have a change to the opposite sign [5]. Similar behaviour is observed between the components of eigenvector before and during the crisis. Furthermore, the eigenvector components that correspond to the eigenvalues that are present in or near the RMT prediction do not show any significant behaviour in the global financial indices [5].

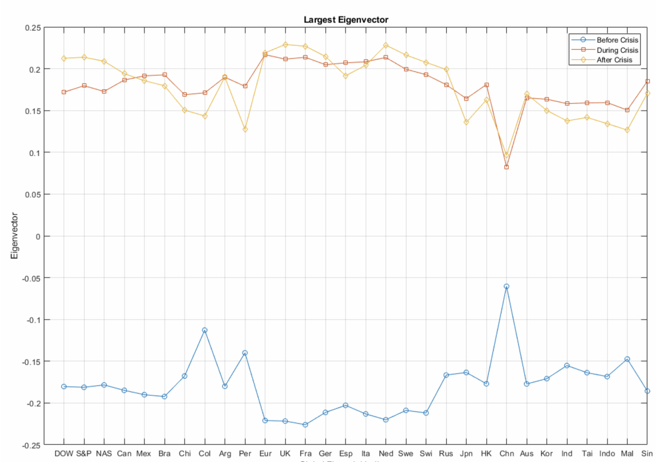

Fig. 3. Largest eigenvector components of the 30 global financial indices before, during, and after financial crisis.

Finally, the author of paper [5] has also displayed the results of the IPR from the three different periods of the crisis as shown in Figure 5. Based on Figure 5, the IPR of cross-correlation matrix $\boldsymbol{C}$ in the three different periods of the financial crisis have edges present in the eigenvalue spectrum hence, there is the presence of significant deviations. This suggests that all the indices are participating in the largest eigenvector. The largest IPR before the crisis is 0.5821, during the crisis is

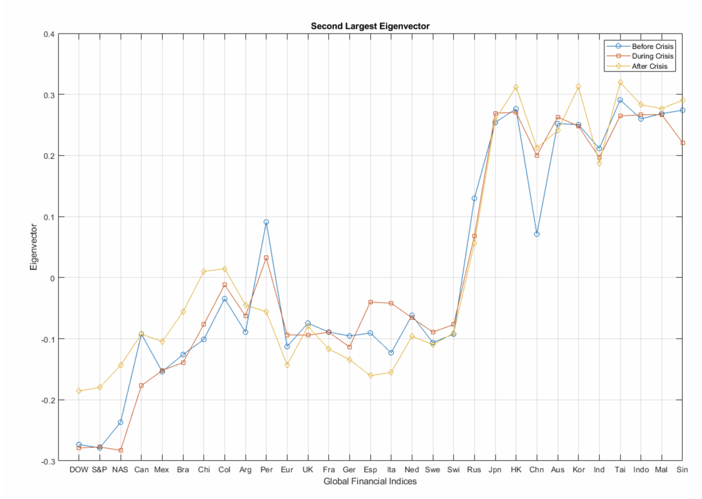

Fig. 4. Second largest eigenvector components of the 30 global financial indices before, during, and after financial crisis.

0.5817 and after the crisis is 0.5441. The author also states that after the crisis has a lower IPR because few global financial indices contribute to the distribution of the smallest eigenvector components after the crisis [5].

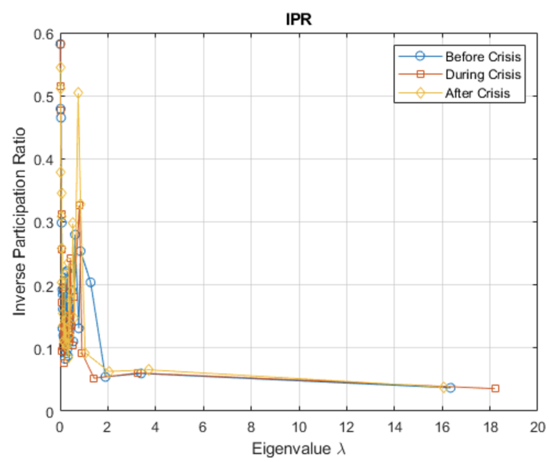

Fig. 5. IPR for the 30 global financial indices before, during and after the financial crisis.

III. Critical Analysis of [5]

The tasks and contributions done by the author of paper [5] has been quite significant in investigating the 2008 financial crisis. However, upon producing the simulations of the results of the paper, there seemed to be a lot of assumptions being made with the results and with certain sections of the paper being vague. Since I am new to the field of financial data, I found it a bit challenging to follow the methodology as proposed by the author. Some of the issues I had with the paper were regarding the deductions made by the paper [5] based on the average and standard deviation of the cross-correlation coefficients, maximum eigenvalues, largest and second largest components eigenvector components and the IPR. The results have just been mentioned with brief interpretation and devoid of in-depth analysis.

Furthermore, the author of the paper [5] has simply compared the results of before, during and after the financial crisis with the values of the results not being significantly different during these three periods. The values of the average and standard deviation of the cross-correlation coefficients, maximum eigenvalues and IPR are not significantly different in the three periods and hence, not the most ideal way to draw deduction and conclusions. Furthermore, upon comparing the largest and second largest eigenvector components, the author simply mentioned about the change in positive or negative sign of the largest and second largest eigenvector components of a particular country when there is a transition between one of the three different periods of crisis. The underlying issue over here that the author has not considered the situation sign ambiguity of the eigenvector components and as a result, the change of sign to derive conclusion regarding the crisis tends to be unimportant.

Therefore, it is important to perform more statistical analysis on the global financial indices data to draw upon more conclusions regarding the distribution of the eigenvalues and its corresponding distribution of the eigenvector components. This will make it possible to develop more in-depth comparisons and interpretations of the results. Therefore, I will apply the techniques introduced in the paper [4] on the financial data which is further discussed in Section IV. In addition, the author has not considered any cleaning techniques towards the financial data and this is one possible aspect that is further explored in Section IV. Perhaps, the implementation of the cleaning technique can improve the quality of the of the results for better interpretation.

Finally, one issue I had with paper [5] was that despite obtaining the results, these are simply observations made by the author. The author has failed to do any new contribution with the financial data and the results obtained. Given that the author is analyzing the financial data from three different periods of the financial crisis, the author could have done more in-depth analysis to find financial trends and information and this new-found methodology could have been used to analyze other financial data during other different periods. Therefore, the important thing to consider is that there is more than one way to analyze the global financial indices during the three different periods of financial crisis to obtain much more informative information and comparison.

IV. New Contribution

A. In-depth analysis of the Global Financial Indices

1) ***Empirical Eigenvalue Distribution:*** Using the techniques and knowledge I have obtained from paper [4], I decided to conduct a more in-depth and thorough statistical analysis of the global financial indices. Firstly, I decided to compare the eigenvalue distribution $\rho(\lambda)$ of cross-correlation matrix $\boldsymbol{C}$ with the MP Law distribution $\rho_{rm}(\lambda)$ from the three different periods of the crisis as shown in Figures 6, 7 and 8.

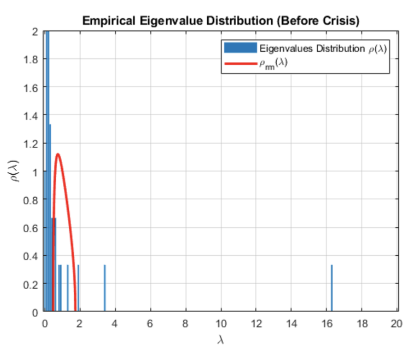

Fig. 6. Empirical eigenvalue distribution $\rho(\lambda)$ of cross-correlation matrix $\boldsymbol{C}$ and MP Law distribution $\rho_{rm}(\lambda)$ before the financial crisis.

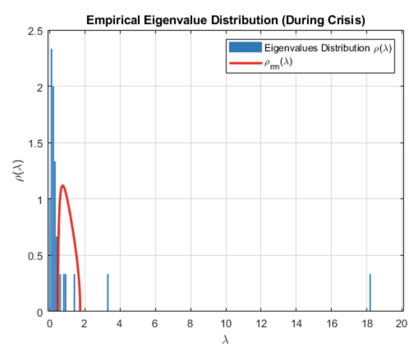

Fig. 7. Empirical eigenvalue distribution $\rho(\lambda)$ of cross-correlation matrix $\boldsymbol{C}$ and MP Law distribution $\rho_{rm}(\lambda)$ during the financial crisis.

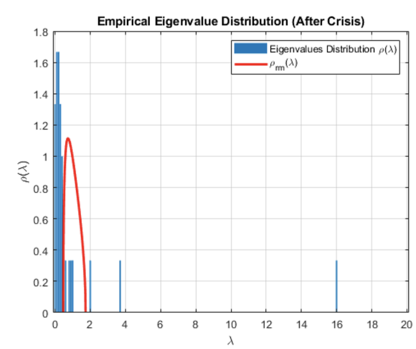

Fig. 8. Empirical eigenvalue distribution $\rho(\lambda)$ of cross-correlation matrix $\boldsymbol{C}$ and MP Law distribution $\rho_{rm}(\lambda)$ after the financial crisis.

The eigenvalues that fall inside the MP Law bound do not reflect useful information and can be regarded as noise meanwhile, the eigenvalues that lie outside the MP Law curve are considered as essential signals whereby the financial data is not purely random [2]. I also decided to generate a random cross-correlation matrix $\boldsymbol{R}$, as shown in equation (5). The values chosen for the MP Law distribution in this case are identical to the values of Q, λ_+ and λ_- which are from the after crisis period as mentioned above in Section II B.. In addition, I also decided to examine the situation whereby I shuffle the financial data in time series such that cross-correlation matrix $C_{shuffle}$ denotes the time series randomized data of cross correlation matrix $\boldsymbol{C}$ from the three different periods of the financial crisis.

Based on Figure 9, it can be seen that there no deviating eigenvalues as the eigenvalues of $\boldsymbol{R}$ converge to the MP Law distribution. This highlights that the eigenvalues in the bulk are basically noise [4]. Meanwhile, when the financial data

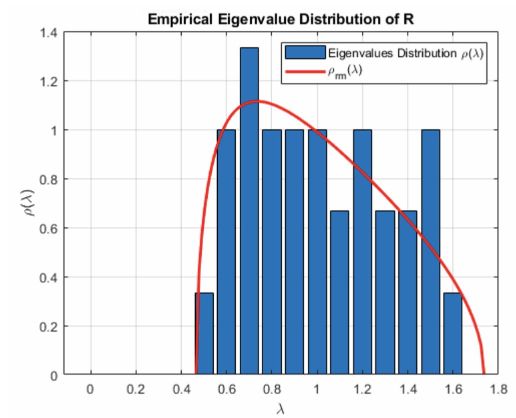

Fig. 9. Empirical Eigenvalue Distribution $\rho(\lambda)$ of random correlation matrix $\boldsymbol{R}$ with $N = 30$ from random generated uncorrelated time series with $T = 297$ in comparison to the MP Law distribution $\rho_{rm}(\lambda)$.

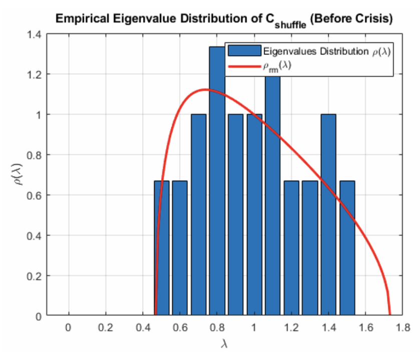

Fig. 10. Empirical eigenvalue distribution $\rho(\lambda)$ of cross-correlation matrix $C_{shuffle}$ and MP Law distribution $\rho_{rm}(\lambda)$ before the financial crisis.

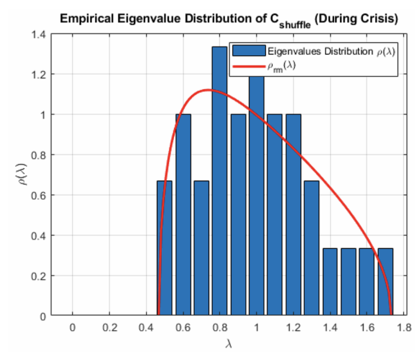

Fig. 11. Empirical eigenvalue distribution $\rho(\lambda)$ of cross-correlation matrix $C_{shuffle}$ and MP Law distribution $\rho_{rm}(\lambda)$ during the financial crisis.

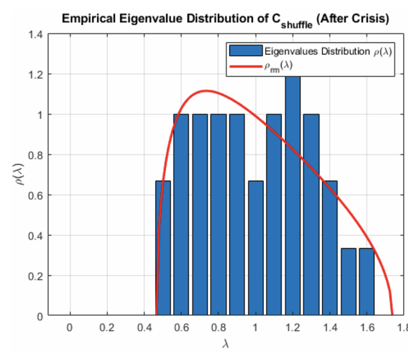

Fig. 12. Empirical eigenvalue distribution $\rho(\lambda)$ of cross-correlation matrix $C_{shuffle}$ and MP Law distribution $\rho_{rm}(\lambda)$ after the financial crisis.

is shuffled in time series [4], it can be seen from Figures 10, 11, and 12 that the eigenvalues converge towards the MP Law distribution. This is because when there is randomization of the time series of the returns, the equal-time correlations are being destroyed and hence, the cross-correlation matrix $C_{shuffle}$ behaves like a random noise matrix and the result is identical to Figure 9. Furthermore, the largest eigenvalues are the effect of equal-time correlations among global financial indices and as a result, there are no deviating eigenvalues [4].

*2) **Distribution of Eigenvector Components**:* If there are deviations present in the empirical eigenvalue distribution $\rho\lambda$ from the MP Law distribution $\rho_{rm}(\lambda)$, then these deviations are also present in the corresponding distribution of the eigenvector components. Therefore, the eigenvector distribution of random cross-correlation matrix $\boldsymbol{R}$ which conforms to the Gaussian Distribution that has a mean zero and unit variance is: [4]:

$$\rho_{rm}(u) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-u^2}{2}\right) \tag{9}$$

Since the $\rho(\lambda)$ of cross-correlation matrix $\boldsymbol{C}$ deviates from the MP Law distribution range $\rho_{rm}(\lambda)$, it can also be inferred that the the corresponding distribution of the eigenvector components of these eigenvalues, denoted by $\rho(u)$ will also deviate systematically from the RMT result $\rho_{rm}(u)$ [4]. Therefore, I decided to display the distribution of the largest eigenvector components corresponding to the largest eigenvalue of cross-correlation matrix $\boldsymbol{C}$ and compare it to $\rho_{rm}(u)$.

As it can be seen in Figures 13, 14 and 15, the distribution of the largest eigenvector components that correspond to the largest eigenvalue distribution deviate systematically from the $\rho_{rm}(u)$ thus suggesting the participation of majority of the indices in the largest eigenvector components during the three periods of the financial crisis.

In addition, I also decided to display the distribution of the largest eigenvector components corresponding to the largest

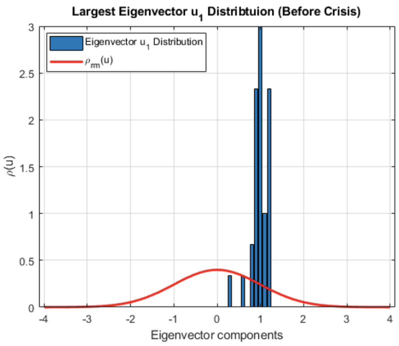

Fig. 13. Largest eigenvector components distribution $\rho(u)$ of cross-correlation matrix $\boldsymbol{C}$ and RMT result $\rho_{rm}(u)$ before the financial crisis.

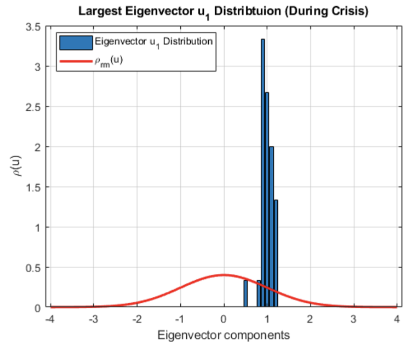

Fig. 14. Largest eigenvector components distribution $\rho(u)$ of cross-correlation matrix $\boldsymbol{C}$ and RMT result $\rho_{rm}(u)$ during the financial crisis.

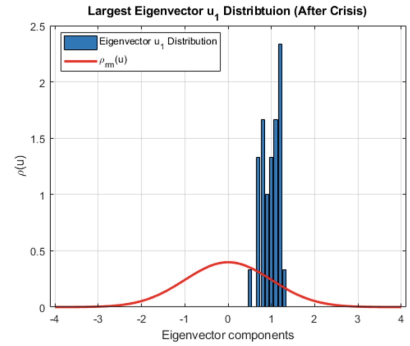

Fig. 15. Largest eigenvector components distribution $\rho(u)$ of cross-correlation matrix $\boldsymbol{C}$ and RMT result $\rho_{rm}(u)$ after the financial crisis.

eigenvalue of cross-correlation matrix $\boldsymbol{R}$ and cross-correlation matrix $C_{shuffle}$ from all the three different periods of the financial crisis and compare it to $\rho_{rm}(u)$ as shown in Figures 16, 17, 18 and 19.

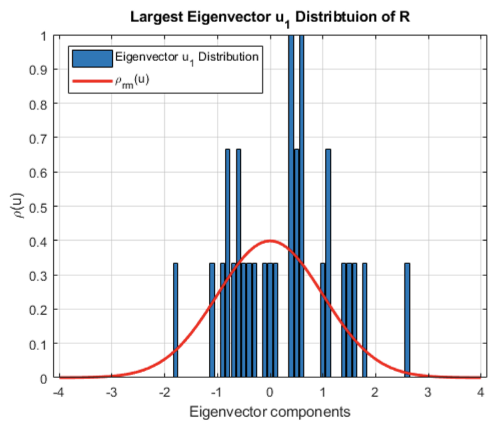

Fig. 16. Largest eigenvector components distribution $\rho(u)$ of $\boldsymbol{R}$ and RMT result $\rho_{rm}(u)$.

Based from all these four figures, the distribution of the largest eigenvector components of cross-correlation matrix $C_{shuffle}$ in all three different periods of financial crisis are quite identical to the distribution of the largest eigenvector components of cross-correlation matrix $\boldsymbol{R}$ as they conform to a Gausssian distribution of $mean = 0$ and $variance = 1$.

B. Implementation of Market Mode Removal

Now based from Figures 13, 14 and 15, it can be seen that majority of the global financial indices participate in the largest eigenvector components that correspond to the largest eigenvalue and this results in an influence which is common to all the indices. This is known as the market mode and the approach to remove this market mode can be done by implementing a simple linear regression method that is introduced in paper [4]. This method is able to remove the maximum eigenvalue and its affect on the remaining data. One method to model the influence that is common to all the indices is expressed below:

$$R_i(t) = \alpha_i + \beta_i M(t) + \varepsilon_i(t) \tag{10}$$

whereby R_i is the return of the index i, $M(t)$ represents the additive term that is similar for all the indices and explains the correlations between any pair of indices, $\langle \varepsilon(t) \rangle = 0$, α_i and β_i are the index-specific constants and $\langle M(t)\varepsilon(t) \rangle = 0$ [4]. Since the largest distribution of the eigenvector components $u^k_{largest}$ has a significant influence to all the indices, the term $M(t)$ can be approximated with $R^{largest}(t)$ and the parameters of α_i and β_i can be estimated by ordinary least square regression. After removing the contribution of $R^{largest}(t)$ from each times series of $R_i(t)$, it is possible to construct a new cross-correlation

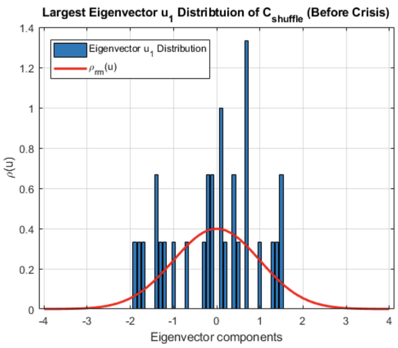

Fig. 17. Largest eigenvector components distribution $\rho(u)$ of cross-correlation matrix $C_{shuffle}$ and RMT result $\rho_{rm}(u)$ before the financial crisis.

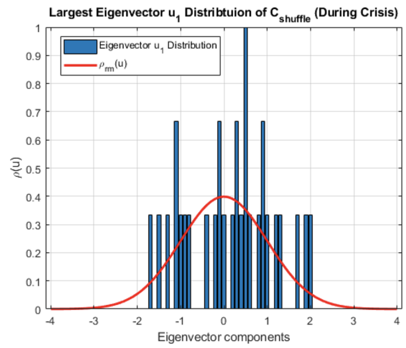

Fig. 18. Largest eigenvector components distribution $\rho(u)$ of cross-correlation matrix $C_{shuffle}$ and RMT result $\rho_{rm}(u)$ during the financial crisis.

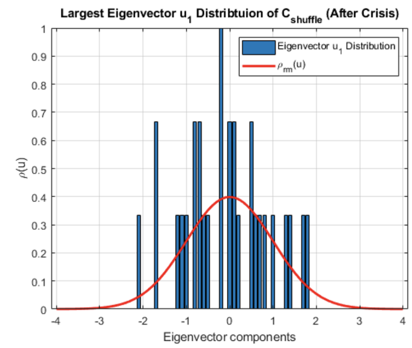

Fig. 19. Largest eigenvector components distribution $\rho(u)$ of cross-correlation matrix $C_{shuffle}$ and RMT result $\rho_{rm}(u)$ after the financial crisis.

matrix $\boldsymbol{C}$ from residuals $\varepsilon_i(t)$ for all the three different periods of the financial crisis. Figure 20 shows the probability density of cross-correlation matrix $\boldsymbol{C}$ after the removal of the market mode.

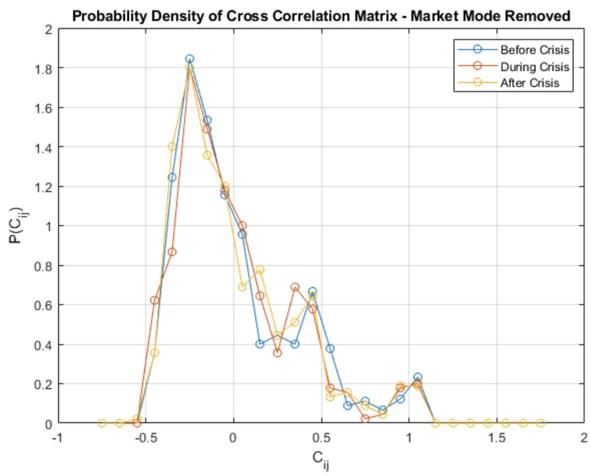

Fig. 20. Probability Density of the cross-correlation matrix $\boldsymbol{C}$ after the removal of the market mode before, during and after the 2008 financial crisis.

The average values of the cross-correlation coefficients before crisis is 0.0075, during the crisis is 0.0054, and after the crisis is 0.0016 of the global financial indices. By removing the largest eigenvalue, the average value of the cross-correlation coefficients has decreased significantly thus implying that large degree of the cross-correlations contained in cross-correlation matrix $\boldsymbol{C}$ are attributed to the largest eigenvalue and its corresponding largest eigenvector components [4]. As seen in Figure 20, the curves in all three different periods of financial crisis are almost identical with each other with minimal differences. Figures 21, 22 and 23 represent the empirical eigenvalue distribution of cross-correlation matrix $\boldsymbol{C}$ after the removal of the market mode from before, during and after the financial crisis respectively.

When comparing Figures 21, 22 and 23 with Figures 6, 7 and 8, more eigenvalues have started to converge towards the MP Law curve. This shows that by removing the market mode, the largest eigenvalue has been removed and the trace of the cross-correlation matrix $\boldsymbol{C}$ has been preserved. As a result, the smaller eigenvalues are not being pushed towards zero by the largest eigenvalue and hence, there is a higher portion of smaller eigenvalues that are converging towards the MP Law curve thus implying that these eigenvalues can be considered as noise. Furthermore, it can be seen in Figures 21, 22 and 23 that there are still deviating eigenvalues present in the three different periods of the financial crisis which can be used to provide essential information about the global financial indices.

The next step is to examine the distribution of the largest eigenvector components corresponding to the largest eigenvalue distribution after the removal of the market mode and comparing it to the RMT result $\rho_{rm}(u)$. This is shown in Figures 24, 25

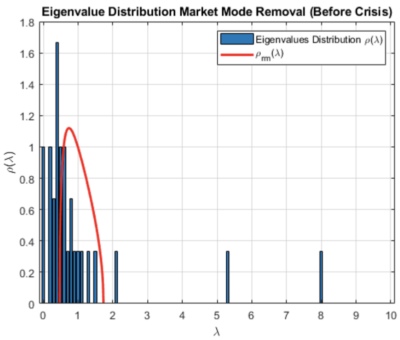

Fig. 21. Empirical eigenvalue distribution $\rho(\lambda)$ of cross-correlation matrix $\boldsymbol{C}$ after removal of market mode and MP Law distribution $\rho_{rm}(\lambda)$ before the financial crisis.

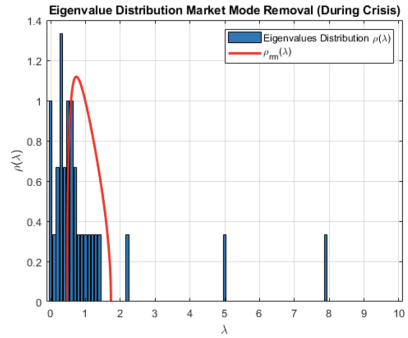

Fig. 22. Empirical eigenvalue distribution $\rho(\lambda)$ of cross-correlation matrix $\boldsymbol{C}$ after removal of market mode and MP Law distribution $\rho_{rm}(\lambda)$ during the financial crisis.

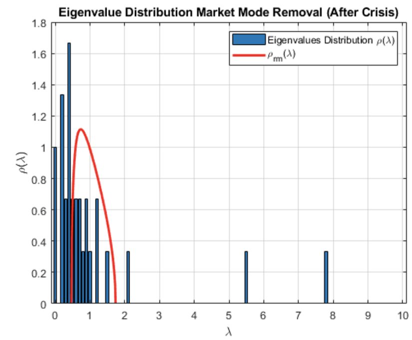

Fig. 23. Empirical eigenvalue distribution $\rho(\lambda)$ of cross-correlation matrix $\boldsymbol{C}$ after removal of market mode and MP Law distribution $\rho_{rm}(\lambda)$ after the financial crisis.

and 26.

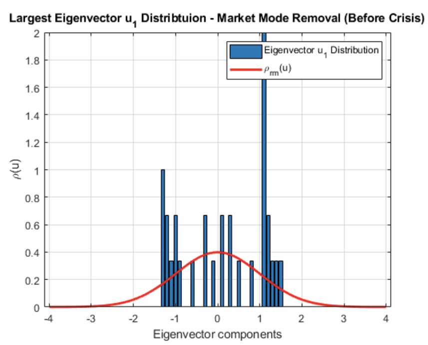

Fig. 24. Largest eigenvector components distribution $\rho(u)$ of cross-correlation matrix $\boldsymbol{C}$ after removal of market mode and RMT result $\rho_{rm}(u)$ before the financial crisis.

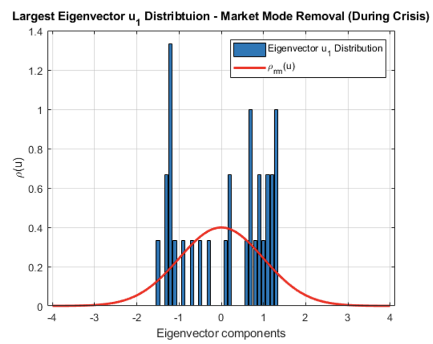

Fig. 25. Largest eigenvector components distribution $\rho(u)$ of cross-correlation matrix $\boldsymbol{C}$ after removal of market mode and RMT result $\rho_{rm}(u)$ during the financial crisis.

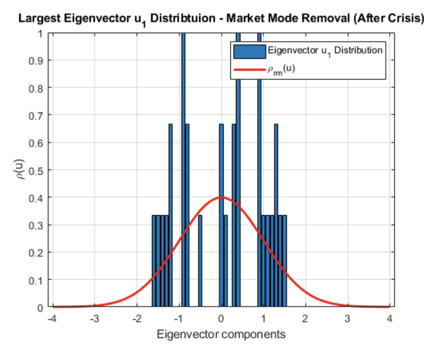

Fig. 26. Largest eigenvector components distribution $\rho(u)$ of cross-correlation matrix $\boldsymbol{C}$ after removal of market mode and RMT result $\rho_{rm}(u)$ after the financial crisis.

Similarly, upon comparing Figures 24, 25 and 26 with Figures 13, 14 and 15, the largest distribution of the eigenvector components corresponding to the largest eigenvalue distribution in three different periods of the financial crisis are quite identical to the distribution of the largest eigenvector components of cross-correlation matrix $\boldsymbol{R}$ as they conform to a Gausssian distribution of $mean = 0$ and $variance = 1$. This shows that not only majority of the global financial indices are participating in the largest eigenvector components, but these global financial indices are also contributing to the other eigenvector components and hence, there is no significant deviation of the largest distribution of the eigenvector components. In addition, I decided to follow the author of paper [5] and display the largest and second largest eigenvector components after the removal of the market mode of the 30 global financial indices before, during and after the financial crisis. On the basis of Figures 27 and 28, the results do agree with the observation of the author regarding the change in sign for certain global financial indices (i.e. US, Eur, UK, Fra, Ind, Indo etc.) where there is a transition from period of the financial crisis to another period. When compared to Figures 3 and 4, Figures 27 and 28 do provide a better indication to the observation and interpretation of the author of paper. However, using the distribution of the largest and second largest eigenvector components and examining the sign change as a way examine the crisis in the three different periods is negligible because eigenvector components are subjected to sign ambiguity thus the interpretation being unimportant.

Finally, Figure 29 displays the IPR of the gloabl financial indices after the removal of the market mode. Similar to Figure 5, there are edges present in the eigenvalue spectrum thus implying the presence of deviations. This shows that even after the removal of the market mode, there are still majority of the indices that are participating in the largest eigenvector components. However, the size of the edge of the IPR after the removal of the market mode is significantly smaller compared

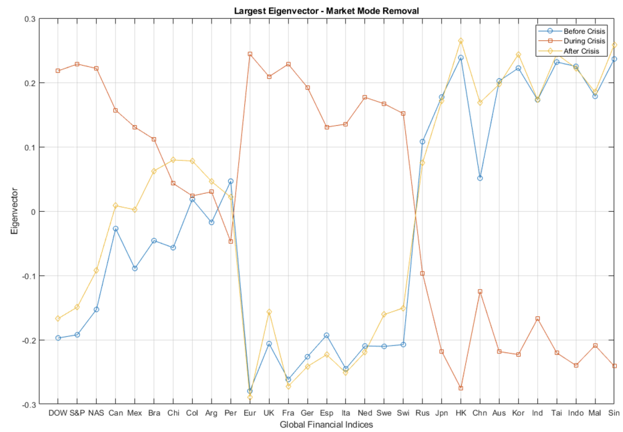

Fig. 27. Largest eigenvector of 30 global financial indices before, during, and after 2008 financial crisis after the removal of market mode.

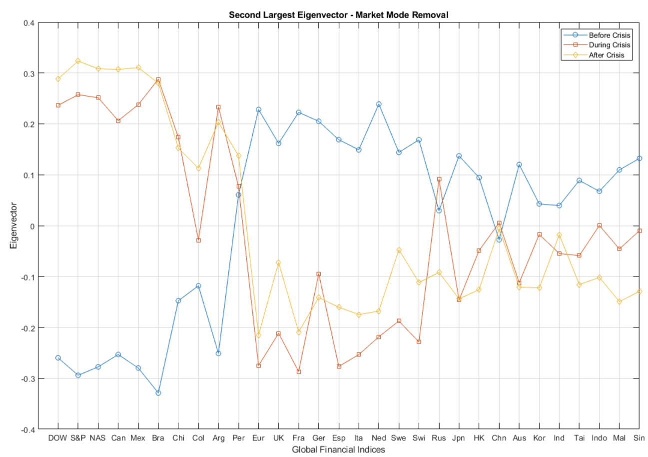

Fig. 28. Second largest eigenvector of 30 global financial indices before, during, and after 2008 financial crisis after the removal of market mode.

to the IPR in Figure 5. This is because after the removal of the market mode, majority of indices are not only contributing to the largest eigenvector components but they are also contributing to the other eigenvector components meanwhile, when the market mode was not removed, there was a heavier participation of the indices on the largest eigenvector components and fewer participation of the indices towards the smaller eigenvector components.

After the implementation of the market mode removal, it can be seen that there is more cleaning done on the financial data. Given that the author of paper [5] would have incorporated this methodology, it would have been easier to determine which eigenvalue and its corresponding eigenvector components can be regarded as noise and which deviating eigenvalue and corresponding eigenvector components would provide more useful and essential information in the three different periods of the financial crisis. However, even after the implementation of the market mode removal, the financial data is being cleaned and observations are being made about the changes in the three different periods of the financial crisis.

C. Implementation of Markowitz Mean-Variance Portfolio Theory

Instead of simply making observations about the eigenvalues and its corresponding eigenvector components, there is a possible method to observe a unique trend during the three different periods of the financial crisis given that this methodology can be applied to other global financial indices data from different periods and can be used as a predictive model. The method that I thought would be suitable to implement was the Markowitz mean-variance portfolio theory as discussed in [9]. In this case, a portfolio will be made out of financial indices respective to their country. This is because global financial indices

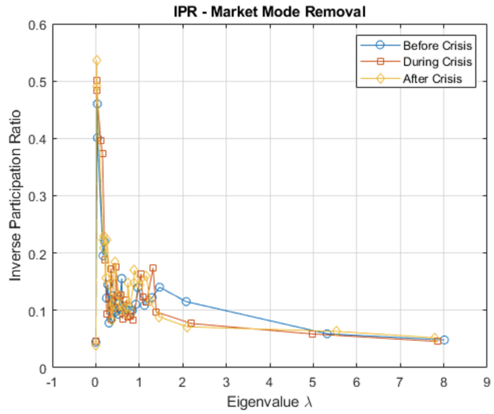

Fig. 29. IPR for the 30 global financial indices before, during and after the crisis after the removal of market mode and IPR of $\boldsymbol{R}$.

comprise of constituents that represent the stock indices of a large number of companies. Given that a portfolio contains a global financial index, it is basically storing the stock indices of a large number of companies of that particular country. When designing the portfolio, the idea is to have an expected return that is maximized at a particular level of risk. Given that there are 30 global financial indices, 27 individual portfolios have been designed as three of the indices are from the USA.

By implementing the Markowitz mean-variance portfolio theory, it is possible to model the rate of returns on the assets as random variables [9]. Given a random vector m, m represents an individual global financial index corresponding to its respective country. In total, there are 27 random vectors m as we are dealing with 27 countries. $cov(x) = \Sigma$ and p is a portfolio of each individual random vector m. Hence, I will be analyzing 27 different p.

Now given that the covariance matrix is Σ, under the Markowitz mean-variance portfolio theory, a portfolio p needs to be designed such that it can minimize the risk $R(p)$ which is denoted by:

$$p_{opt|\Sigma} = \frac{\Sigma^{-1}m}{e^T\Sigma^{-1}m} \tag{11}$$

Furthermore, the idea is to choose an optimal portfolio weighting factor and regarding the Markowitz portfolio theory, the optimal set of weights is one where $e^T m = 1$ whereby $m = (1,\dots,1)^T$. This allows the portfolio to have an acceptable baseline expected rate of return with minimal volatility [9]. The main motivation behind the Markowitz portfolio theory is to determine the variance of the daily return which is calculated as

$$R^2(p) = p^T\Sigma p \tag{12}$$

it represents the squared investment risk that is associated with the portfolios [9].

The next step is to design an estimator that makes it possible to compute the return for a whole period of time. Let's denote w as the window size of the training data in order to calculate the variance of the daily return. Given that the length of the time series for the global financial indices data is approximately 300, I have decided use $w = 30, 60, 90, 120$ for the portfolio period (training window) in order to compute the variance of daily return.

Figures 30, 31, 32 and 33 display the variance of daily returns $R^2(p)$ with different values of w for the three different periods of the financial crisis. An interesting observation that can be made in all the results with different w is that the variance of daily returns for majority of the countries is significantly higher during the period of the financial crisis as compared to before and after the financial crisis. This shows that if an investor of a particular country were to have a portfolio during the time of the crisis, the portfolio of the individual country would comprise of much more risky and unstable stocks belonging to a large number of companies. As a result, with high variance of daily return, it indicates that the portfolio of the investor is very risky as compared to before and after the financial crisis.

Therefore, the variance of daily return computed from the Markowitz mean-variance portfolio can be used as a financial crisis indicator to compare the global financial indices of the different countries in order to make logical comparisons between different periods and to establish more solid interpretation of the trends obtained from the financial data.

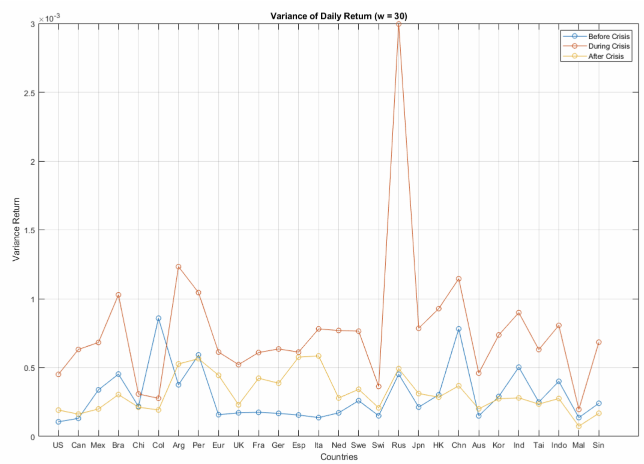

Fig. 30. Variance of Daily Return given $w = 30$.

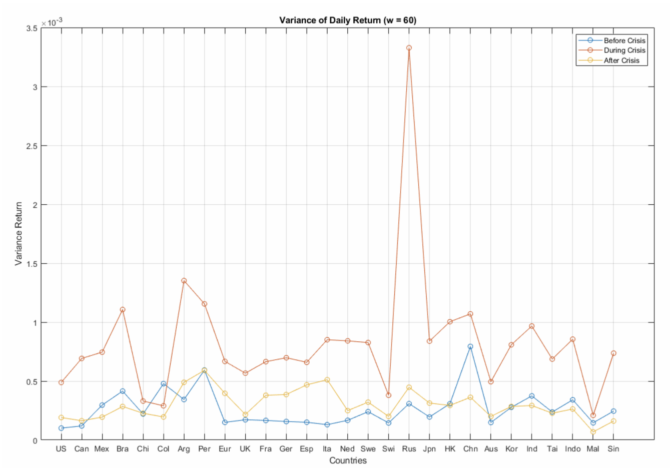

Fig. 31. Variance of Daily Return given $w = 360$.

V. Conclusion and Future Work

In this report, the method introduced by paper [5] was discussed to understand the application of RMT in analyzing the global financial indices during the 2008 financial crisis. The methodology implemented by the author of paper [5] was discussed along with the results and simulations I performed and obtained by following the paper [5]. I got the opportunity to critically analyze the paper whereby the author had simply mentioned the results without any in-depth analysis and interpretation behind the results. In addition, the author had not considered any techniques to clean the financial data. Therefore, using the knowledge of paper [4], I decided to a more in-depth analysis of the global financial indices data to draw out more conclusions and understanding behind the data. Furthermore, I decided to implement the market mode removal technique to further clean the financial data and draw out more interpretation. Finally, I wanted to incorporate my knowledge of the Markowitz mean-variance portfolio theory on the global financial indices to provide a much more interpretative means of understanding the financial data and situation of a particular country in three different periods of the financial crisis.

The future work includes:

1) Using the above mentioned methodologies and contributions on different financial crises such as The Black Monday (1987), Burst of the dot-com bubble and September 11th (2001).
2) Analyze the current financial data using the Markowitz mean-variance portfolio model to observe the trend of the variance of the daily returns of the global financial indices of the countries.

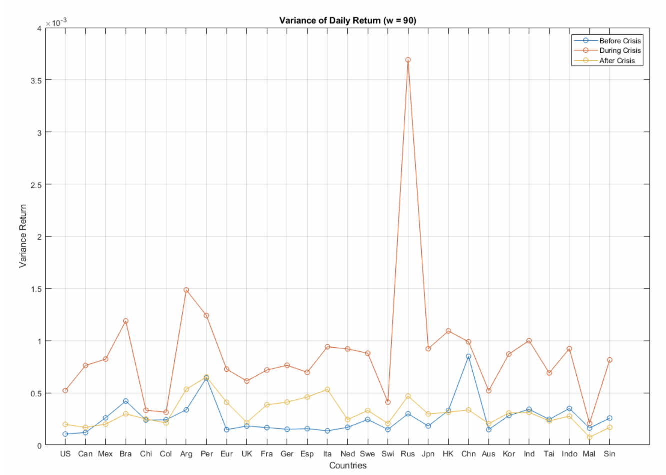

Fig. 32. Variance of Daily Return given $w = 90$.

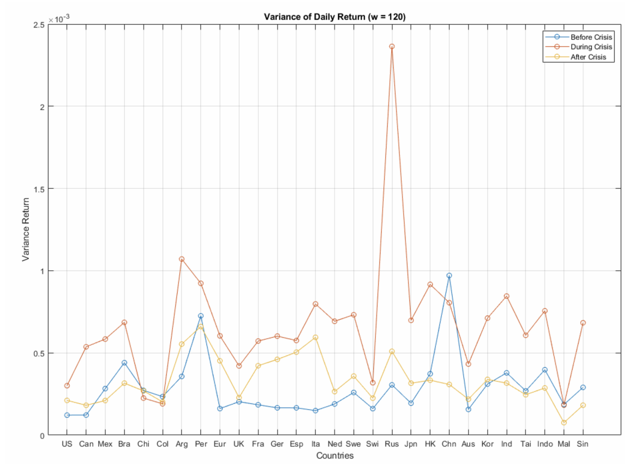

Fig. 33. Variance of Daily Return given $w = 120$.

3) Further improve the Markowitz mean-variance portfolio theory model or introduce a new model that can be used as a means of financial crisis indicator.

References

[1] S. Kumar and N. Deo, "Analysing correlations after the financial crisis of 2008 and multifractality in global financial time series," *Pramana*, vol. 84, no. 2, pp. 317–325, 2015.

[2] M. S. Wang, "Seeking gold in sand: financial applications of random matrix theory in stock market data," 2016.

[3] "Random Matrix Theory Approach to Financial Markets." `http://shodhganga.inflibnet.ac.in/bitstream/10603/35236/12/12_chapter%2002.pdf`. Accessed: 2019-04-28.

[4] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, T. Guhr, and H. E. Stanley, "Random matrix approach to cross correlations in financial data," *Physical Review E*, vol. 65, no. 6, p. 066126, 2002.

[5] A. Nobi, S. E. Maeng, G. G. Ha, and J. W. Lee, "Random matrix theory and cross-correlations in global financial indices and local stock market indices," *Journal of the Korean Physical Society*, vol. 62, no. 4, pp. 569–574, 2013.

[6] L. S. Junior and I. D. P. Franca, "Correlation of financial markets in times of crisis," *Physica A: Statistical Mechanics and its Applications*, vol. 391, no. 1-2, pp. 187–208, 2012.

[7] J. P. Bouchaud and M. Potters, "Financial applications of random matrix theory: a short review," 2009.

[8] Y. Nakayama and H. Iyetomi, "Random matrix theory of dynamical cross correlations in financial data," *Progress of Theoretical Physics Supplement*, vol. 179, pp. 60–70, 2009.

[9] "Markowitz Mean-Variance Portfolio Theory." `https://sites.math.washington.edu/~burke/crs/408/fin-proj/mark1.pdf`. Accessed: 2019-04-28.