

Econophysics of Stock Market Dynamics

Thesis

submitted to

Indian Institute of Science Education and Research Pune
in partial fulfillment of the requirements for the
BS-MS Dual Degree Programme

by

Hrishidev



Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road,
Pashan, Pune-411008, INDIA

April, 2020

Supervisor: Anirban Chakraborti

© Hrishidev 2020

All rights reserved

Certificate

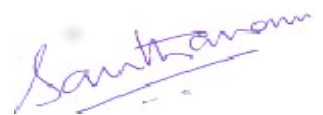
This is to certify that this dissertation entitled “Econophysics of Stock Market Dynamics ” towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Hrishidev at Indian Institute of Science Education and Research under the supervision of Anirban Chakraborti, Professor, School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi, during the academic year 2019-2020.



Signature of candidate
(Hrishidev)



Signature of supervisor
(Prof. Anirban Chakraborti)



Signature of expert
(Prof. M S Santhanam)

“All our knowledge begins with the senses, proceeds then to the understanding, and ends with reason. There is nothing higher than reason.” - Immanuel Kant.

Declaration

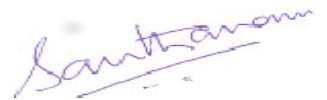
I hereby declare that the matter embodied in the report entitled “Econophysics of Stock Market Dynamics ” are the results of the work carried out by me at the Department of Physics, Indian Institute of Science Education and Research, Pune, under the supervision of Prof. Anirban Chakraborti , Jawaharlal Nehru University, New Delhi, and the same has not been submitted elsewhere for any other degree.



Signature of candidate
(Hrishidev)



Signature of supervisor
(Prof. Anirban Chakraborti)



Signature of expert
(Prof. M S Santhanam)

Acknowledgments

I would like to express my sincere gratitude towards my supervisor Prof. Anirban Chakraborti. Thank you for believing in me and always being there for me with your guidance, encouragement and patience in matters both academic and non-academic. I am indebted to Prof. M S Santhanam for his valuable inputs and comments during the course of this project. I want to thank my collaborators Dr. Hirdesh Kumar Pharasi and Dr. Kiran Sharma for assisting and guiding me in the course of writing this thesis by giving critical inputs and helping me to produce research quality figures. I would also like to thank Dr. Sunil Kumar for his pieces of advice. My appreciation also extends to Mr. Ashish Kumar, Mr. Rajdeep Haldar and my ETC laboratory mates, for the most stimulating discussions and debates in both subjects very closely related, and totally unrelated to the work presented in this thesis. This work would not have been possible if it were not for them and the constant support of my family (especially for my parents) and friends. And finally, I would like to thank IISER Pune and JNU; two totally different campuses with entirely different atmospheres, for moulding my mind and soul with their charms and challenges.

Abstract

The stock market is a fantastic example of a “complex system” that exhibits vibrant correlation patterns and behaviours of the price (or return) time series. The direct impact of it on the economic ecosystem of a country, as well as the abundant availability of structured data, makes this system interesting for empirical studies. This thesis presents a general and robust methodology to extract information about the “disorder” (or randomness) in the market and its eigen modes, using the entropy measure “eigen-entropy” H , computed from the eigen-centralities (ranks) of different stocks in the correlation-network. We have used correlation matrix constructed using the log-return of adjusted closing price of two different data sets containing stocks from United States of American S&P-500 index (USA) and Japanese Nikkei-225 index (JPN), spanning across a sufficiently long period of 32 years, to demonstrate its robustness. Further, the eigenvalue decomposition of the correlation matrix into partial correlation - market, group and random modes, and the relative-entropy measures computed from these eigen modes enabled us to construct a phase space, where the different market events undergo phase-separation and display “order-disorder” transitions. Our proposed methodology may help us to understand the market events and their dynamics, as well as find the time-ordering and appearances of the bubbles and crashes, separated by normal periods. We have studied the evolution of events around major crashes and bubbles (from historical records in USA and JPN). Furthermore, the relative entropy with respect to the market mode $H - H_M$, displayed “universal scaling” behavior with respect to the mean market correlation μ ; a data-collapse was observed when plotted in a linear-logarithmic scale, which suggested that the fluctuations and co-movements in price returns for different financial assets and varying across countries are governed by the same statistical law. In addition, our study may lead to a deeper and broader understanding of scaling and universality phenomena in complex systems, in general.

Contents

Abstract	xi
1 Introduction	1
1.1 Financial market as a complex system	1
1.2 Price return time series and correlation matrices	2
1.3 Objectives and outline	7
2 Brief review of correlation-based networks and entropy measures	9
2.1 Introduction	9
2.2 Random matrix theory (RMT) and applications	10
2.3 Network representations of correlations	13
2.4 Entropy measures and application	17
2.5 Remarks	25
3 Phase separation using eigen-entropy	27
3.1 Introduction	27
3.2 Monitoring eigen-entropy	29
3.3 Order-disorder transitions and phase separation.	36
3.4 Remarks	41

4 Discussions and outlook	47
Bibliography	70

Chapter 1

Introduction

1.1 Financial market as a complex system

A financial market is truly a spectacular example of a complex system that is generally composed of many constituents, which may be diverse in forms but largely interconnected, such that their strong inter-dependencies and emergent behavior change with time. Thus, it becomes almost impossible to describe the dynamics of the system through some simple mathematical equations, and so new tools and interdisciplinary approaches are needed [1, 2]. Hence, there has been a surge of efforts in using ideas from complexity theory [3, 4, 5, 6, 7] to explain and understand economic and financial markets. New insights and concepts, such as networks, systemic risk, tipping points, contagion and resilience have surfaced in the financial literature and may have the potential for better monitoring of the highly interconnected macroeconomic and financial systems and thus, may help anticipate future economic slowdowns or financial crises.

Financial markets have historically exhibited sharp and largely unpredictable drops at a systemic scale, which are termed “market crashes” [8, 9]. Such rapid changes or phase transitions (not in the strict thermodynamic sense of physics [10, 11]) may in some cases have been triggered by unforeseen stochastic events or exogenous shocks, or more often, they may have been driven by certain endogenous underlying processes [12, 13]. The recent global economic downturn in 2007-08 brought us both predicament and hope! Predicament, since the traditional theories in economics could not predict, nay even warn, the near complete

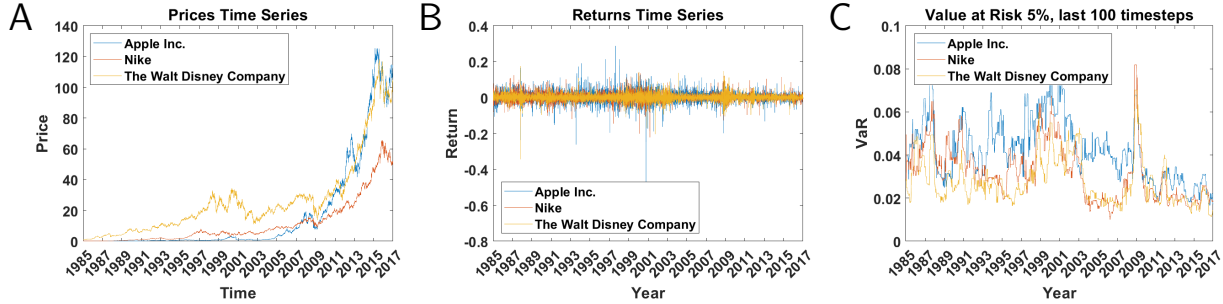


Figure 1.1: **Price, Return, and VaR time series.** Three stocks are chosen in random and their adjusted closed price, logarithmic returns and volatility (captured by VaR) are plotted in (A), (B), and (C) respectively for the time period of 32 years (1985-2017) of the American market S&P-500.

breakdown of the global financial system. And we are yet to recover from its long-lasting effects on the global economy. Hope, since one can now witness signs of change in economic and financial thinking, including the very fact that there is deeper (and less understood) link between macroeconomics and finance [14, 15], which certainly merits more attention.

1.2 Price return time series and correlation matrices

Looking at the time series of the asset prices that are in a financial market is the best way to study the dynamics of this ever-evolving complex system. The efficient market hypothesis says that it is impossible to beat the market consistently in the long term, which can be interpreted as the asset prices will contain the entirety of the available information about the market. Instead of using the price time series of each asset directly, we will be using the logarithmic returns instead. This log-returns time series of an asset tells us the fractional gain or loss of each day compared to the previous day, whereas the price time series; as the name suggests, says the price of the asset at the end of each day. Log-return time series of the i^{th} ($i = 1, 2, 3, \dots, N$ where N is the total number of stocks) stock at time t , return $r_i(t)$ can be calculated from the corresponding price time series $p_i(t)$ as

$$r_i(t) = \ln(p_i(t)) - \ln(p_i(t-1)) \quad (1.1)$$

The reason we are using returns $r_i(t)$ instead of prices is the stationarity of time series.

It has been observed that the mean and variance of the $p_i(t)$ changes a lot even over a short period, whereas the mean and variance of $r_i(t)$ somewhat remain constant over this time. This “wider sense” of stationarity is a prerequisite of many of the statistical techniques that we will be using in this study. These can be observed in the figure 1.1

1.2.1 Data description

For the empirical studies conducted in this thesis, we have used the adjusted closure price time series from the Yahoo finance database [16], for two countries: United States of America (USA) S&P-500 index and Japan (JPN) Nikkei-225 index, for the period 02-01-1985 to 30-12-2016, and for the stocks as follows:

- USA — 02-01-1985 to 30-12-2016 ($T = 8068$ days); Number of stocks $N = 194$;
- JPN — 04-01-1985 to 30-12-2016 ($T = 7998$ days); Number of stocks $N = 165$;

where we have included the stocks which are present in the indices for the entire duration. The sectoral abbreviations are as follows: **CD**–Consumer Discretionary; **CS**–Consumer Staples; **EG**–Energy; **FN**–Finance; **HC**–Health Care; **ID**–Industrials; **IT**–Information Technology; **MT**–Materials; **TC**–Telecommunication Services; and **UT**–Utilities.

The list of stocks (along with the sectors) for the two markets are given in the Table 4.3 and Table 4.4 in the appendix.

We have $T = 7897$ days of data for the Nikkei-225 index, whereas $T = 7998$ days of data for stocks. To resolve this, we add zero return entries corresponding to the missing days in the time series of JPN index for comparison.

Volatility index

We have used the daily closure volatility index (VIX) of the CBOE from Yahoo finance. It is a popular measure of the stock market’s expectation of volatility implied by S&P- 500 index options. It is calculated and disseminated on a real-time basis by the Chicago Board Options Exchange (CBOE). We have data for the period 02-01-1990 to 30-12-2016, for $T = 6805$ days.

For the JPN Nikkei-225 we use the *Garch*(1, 1) model to estimate the volatility index using the market index returns data (1985-01-08 to 2016-12-30) from Yahoo finance.

1.2.2 Stylized facts

Some empirical findings remain consistent to the point that one has to accept them as truths to which any theory should agree with. Such findings/patterns are referred to in economics as “stylized facts”. They are mostly qualitative because of their generality [17, 18]. In case of financial time series, which are a wide array of the log-returns of assets, there are some statistical stylized facts observed [18, 6]. Let us look at some of them in the following section.

Absence of autocorrelations

Autocorrelation measures the similarity of the lagged versions of time series with the original one. For asset return time series the autocorrelations are always insignificant. This observation is illustrated in figure 1.2. This lack of autocorrelation is indicative of the unpredictable behaviour of the returns. The frequency of the data (daily, weekly, or even intra-day) does not affect this observation.

Heavy/fat tails

Compared to a normal distribution, the probability of observing extreme events (large losses and large gains) is higher in the case of financial time series. When plotting the probability density function of the return values, the above feature manifests itself in the form of “fatter tails”. For demonstrating this fact, the probability distribution of the log-return values of three stocks, as well as the corresponding normal distributions, are plotted in the semi-log scale in figure 1.3. In statistics, this degree of peakedness, relative to the tails of a distribution is measured using the value of kurtosis. The kurtosis of the returns is larger (as shown in figure 1.3) than 3, which is the kurtosis of a normal distribution [19, 20]. This existence of fat/heavy tails implies that some rare extreme events cause much of the variance.

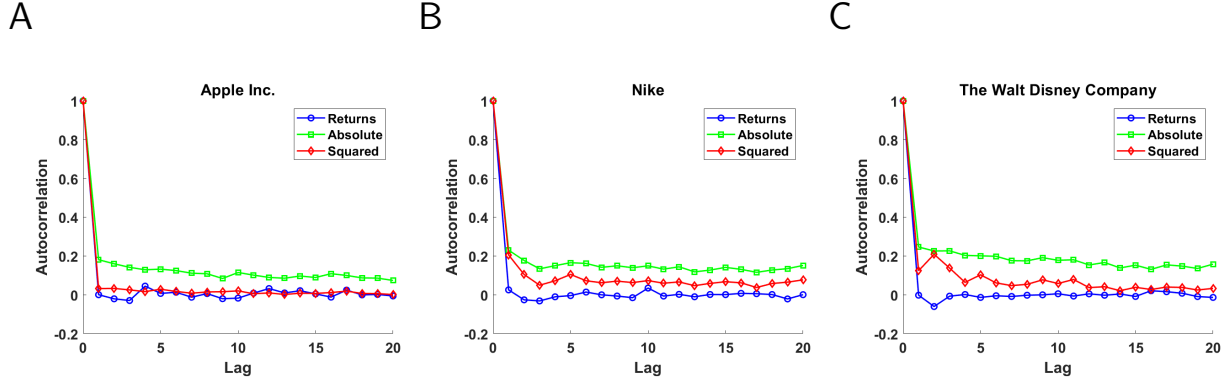


Figure 1.2: **Auto-correlation in financial time series.** Three stocks are chosen (Apple Inc.(A), Nike(B), The Walt Disney Company(C)) and their auto correlation in return time series (blue), absolute value of return time series (green), and square of return time series (red) are plotted.

Slow decay of autocorrelation in absolute returns

Unlike the return time series, the autocorrelation function of absolute value and the square of the return time series is significant and positive [21, 17]. These autocorrelations decay slowly with the increase in the time lag as a power law. Figure 1.2 illustrates this fact. This observation was first made by Taylor [22].

1.2.3 Correlation matrices

Given two time series (r_i and r_j) one can measure their equal-time Pearson correlation coefficient C_{ij} to quantify the strength of the linear relationships in their relative movements [23, 24] as follows:

$$C_{ij} = (\langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle) / \sigma_i \sigma_j \tag{1.2}$$

Here $\langle \dots \rangle$ represents the expectation value and σ_i, σ_j represents the standard deviations of r_i, r_j . C_{ij} 's values range from -1 to 1, where 1 corresponds to the maximum correlation, -1 corresponds to the maximum anti-correlation and 0 corresponds to uncorrelation. One can also look at C_{ij} as the cosine of angle between the T dimension vectors corresponding to the two time series, where T is the number of time steps.

If we generalize this to more than two time series, we can construct a matrix which has

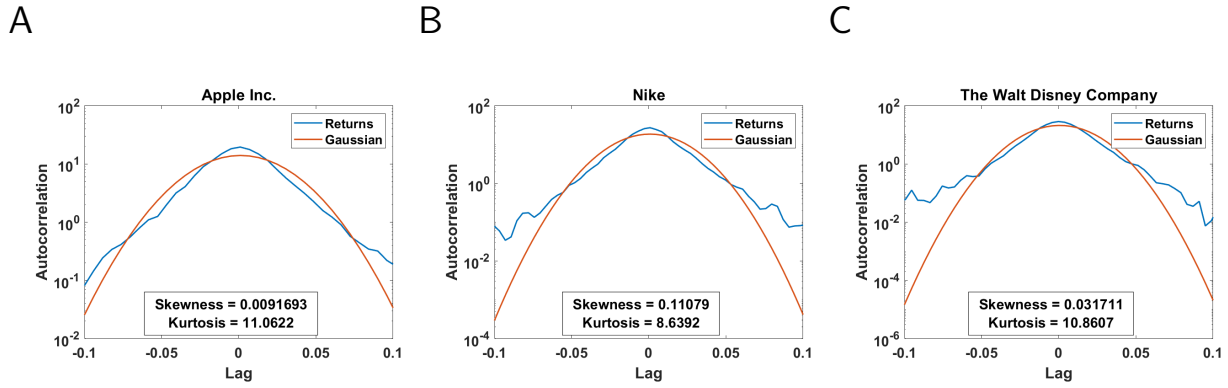


Figure 1.3: **Histogram of log-returns.** Three stocks are chosen (Apple Inc.(A), Nike(B), The Walt Disney Company(C)) and their distributions of market returns is plotted in semi-log scale coloured in blue. For comparison the PDF of a Gaussian distribution that has the same average and variance is also plotted coloured in red.

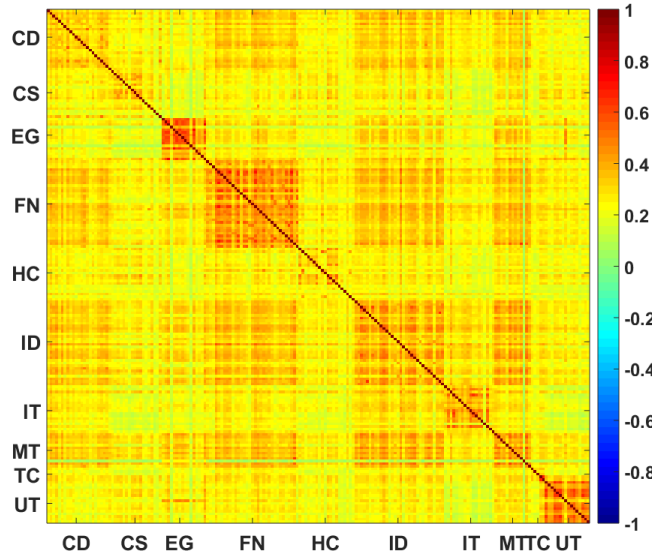


Figure 1.4: **Empirical correlation matrix.** Correlation matrix calculated from the return time series of 194 stocks of S&P-500 market for the period of 32 years (1985-2017). The stocks are arranged sectorwise which is creating the visible block structure across the diagonals. The complete list of stocks are given in the appendix B.

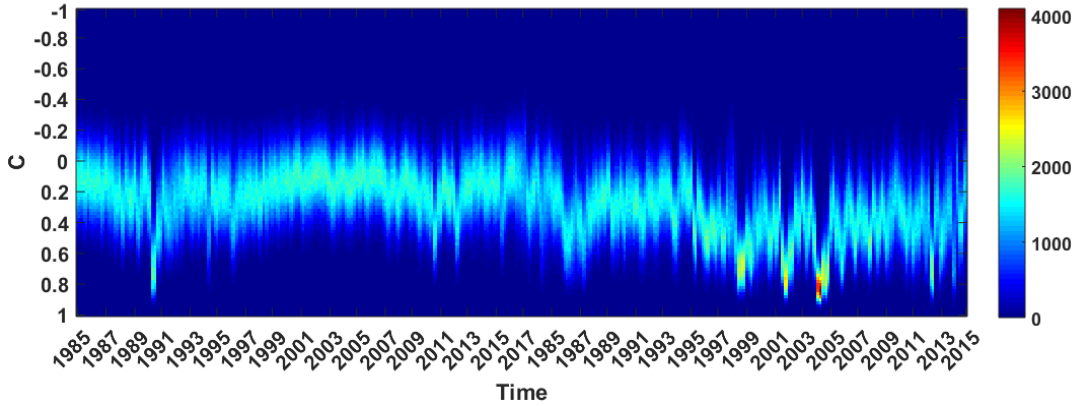


Figure 1.5: **Evolution of correlation coefficients' distribution.** Using rolling window of size 40 days and shift 20 days over a time period of 32 years (1985-2017), the time series of 194 stocks of S&P-500 are divided into 402 frames. For each frame the correlation matrix and correlation coefficients are calculated. After creating fixed bins of width 0.02 between -1 and 1, the number of correlation coefficients in each of these are stored as a column in the 194×402 matrix. This matrix is coloured based on the values of the elements to represent the temporal evolution of the distribution.

the pairwise Pearson correlations between all the stocks. Such a matrix is called a correlation matrix. It helps in summarizing a large amount of data so that one can observe patterns in the co-movements among the time series. Instead of working with the entirety of the time series, we can divide it into shorter windows (epochs) of equal length. This will allow us to monitor the evolution of the market as time progresses. If we consider windows of size M , which shifts by Δ in each step, for a time series of total length T , we have nearest integer greater than or equal to $(\frac{T-W}{\Delta})$ number of frames.

1.3 Objectives and outline

Our aim is to monitor the empirical time series and detect patterns using different methods borrowed from random matrix theory, information theory and complex networks, such that the insight gained can be applied practically. It is known that markets (and thereby the underlying network structures) behave very differently during “crashes and bubbles”, which are extreme events. Detection of bubbles and crashes and prediction of these rare events are a challenge. In the thesis, we would like to develop methods to study and characterise

these market events and possibly detect precursors of such events. We would like to study the continuously evolving network structures of stocks, and gain valuable insights so that we can apply them for developing better investment strategies and framing economic policies. The structure of the thesis is as follows.

Chapter 2: This chapter will act as a review of pre-existing work that deals with correlation networks of price time series and its entropy measures. First, we will discuss the random matrix theory (RMT) approach to characterize different eigen-modes of the correlation matrix. Secondly, we will compare various approaches that give us a network representation of a correlation matrix. And finally, we will look at ways in which people have tried to define and calculate different kinds of entropy measures from the correlation networks.

Chapter 3: In this chapter, we will study the financial market in the light of the entropy measure that we are proposing. First, we will motivate and explain the methodology that we are using to calculate this entropy measure. Then we will discuss the fascinating observations that emerged in this analysis, such as the phase separation of market events. And finally, we will check the effect of the variation of some parameters on the results, and its interpretation.

Chapter 4: This chapter will contain discussions and outlooks regarding the analysis done in the previous chapter. After a brief review of the essential points discussed in the previous chapter, this chapter will continue the discussion on a particular relative entropy measure that shows a universal scaling behaviour and will explore its connection with market risk and volatility.

The appendices contain the list of historical stock market events and the list of stocks whose data we have used in this study.

A major part of the text and figures that I present in this thesis will have an overlap with the manuscript that we have uploaded in the arXiv named “*Phase separation and universal scaling in markets: Fear and fragility*” [25].

Chapter 2

Brief review of correlation-based networks and entropy measures

2.1 Introduction

An interesting way to study the financial market has been in the form of a correlation-based network [26, 27, 28, 29]. For a given epoch, from the correlation matrix C_{ij} , one can infer an underlying network of stocks/assets which can be interpreted to be manifesting as the correlation patterns that we observe in the price movements. This allows us to obtain a representation of the temporal cross-section of the evolving market structure [30, 31] upon which various well-established tools of graph theory can be applied. This method has given new and useful insights into the underlying mechanisms and patterns that drive the overall behavior of this seemingly unpredictable complex system.

Along with obtaining the network representation, we will also be looking at two more things. Before inferring a network representation of the correlation matrix, we can model and characterize the correlation matrix using random matrix theory. The insights given by the random matrix theory approach are valuable in interpreting the correlation patterns and filtering unnecessary noise. Also, after obtaining the network representation, we will be looking at entropy measures that can be calculated from these networks. These entropy measures allow us to monitor the dynamics of the underlying network. In this chapter, we will be looking at and reviewing the tools that enabled academic endeavours of this philosophy

and will be reviewing such previous works which have utilized these tools in interesting ways [32, 33].

2.2 Random matrix theory (RMT) and applications

Eugene Wigner first proposed random matrix theory as an attempt to model the energy levels of a nuclear system using the eigenvalue distribution of a random matrix [34, 35]. One can consider the intricate interactions among nuclear constituents as random fluctuations in the context of R-matrix scattering theory and look at its eigenvalues. These eigenvalues of the random matrix could roughly estimate the energy levels of the nuclear system and the spacings between the energy levels of nuclei could be modelled using the spacing of eigenvalues of the random matrix. Vasiliki Plerou introduced the application of RMT in analysing financial correlation matrices in 1999 [36]. Since then it has found its home in econophysics. Researchers are using this framework for various applications, such as characterizing the random noise [37] and serving as a null model for community detection algorithms [38].

2.2.1 Eigenvalue decomposition of correlation matrices

RMT gives us a way to identify the portion of the correlation matrix from the eigenvalue decomposition, which accounts for the spurious correlations occurring due to T and N not being large enough. If one were to look at an empirical correlation matrix calculated from the return time series of stocks, they would see that in its eigenspectrum along with the eigenvalues of an analogous WOE (same T and N) there are some more existing beyond λ_+ (as seen in figure 2.1). These larger eigenvalues contain information about the group structures in the correlation matrix. The largest one, however, corresponds to the market mode, which represents the super-community comprising of the entire market. From this framework, one can look at the correlation matrix as a sum of three modes, each of which are constructed using the eigenvectors and eigenvalues that occupy different parts of the eigenspectrum. The market mode C_M corresponds to the largest eigenvalue; the group mode C_G corresponds to the eigenvalues that are greater than λ_+ but less than λ_{max} , and finally, the random mode C_R corresponds to the eigenvalues that are less than λ_+ [39, 40].

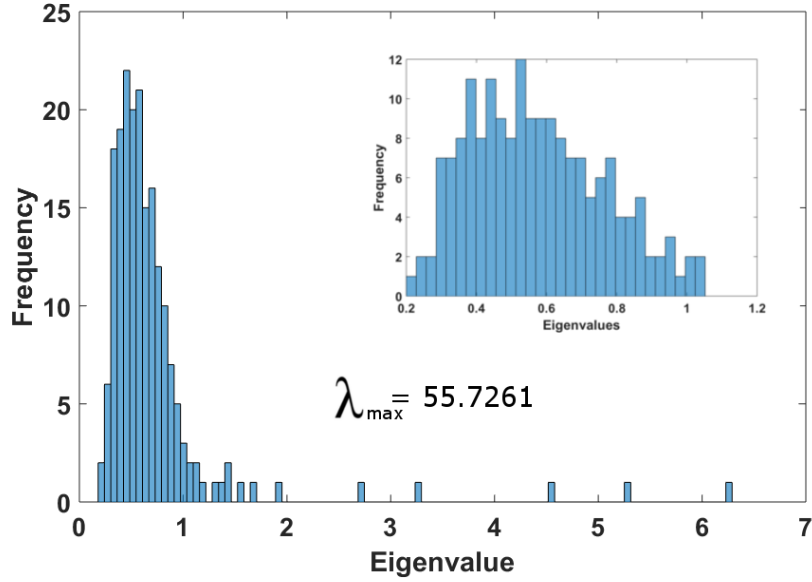


Figure 2.1: **Eigenvalue distribution of empirical correlation matrix.** Correlation matrix using the price return time series of 194 stocks of S&P-500 was calculated and the distribution of its eigenvalues is plotted. Because the largest eigenvalue is significantly larger than the rest it is excluded from the diagram, but its value is displayed for reference. The inset figure corresponds to eigenvalues of just the random mode of the matrix, which has a distribution similar to that of a Wishart orthogonal ensemble of comparable dimension.

If λ_i and v_i are the eigenvalues and eigenvectors of our correlation matrix C , we can reconstruct C as

$$C = \sum_{i=1}^N \lambda_i e_i e_i' \quad (2.1)$$

For convenience, let us sort the eigenvalues in descending order so that as the subscript/index i increases λ_i become smaller and smaller. Now we can define N_G as the number of eigenvalues which falls in the range $\lambda_+ \leq \lambda_i < \lambda_1$. Using these we can reconstruct the correlation matrix using the three modes as:

$$C = C_M + C_G + C_R = \lambda_1 e_1 e_1' + \sum_{i=2}^{N_G} \lambda_i e_i e_i' + \sum_{i=N_G+1}^N \lambda_i e_i e_i' \quad (2.2)$$

The market mode has information about the super-community of all/most of the stocks (in a way the market as a whole). It is highly related with the average market correlation. The group mode contains information about the sectoral correlations (internal and external).

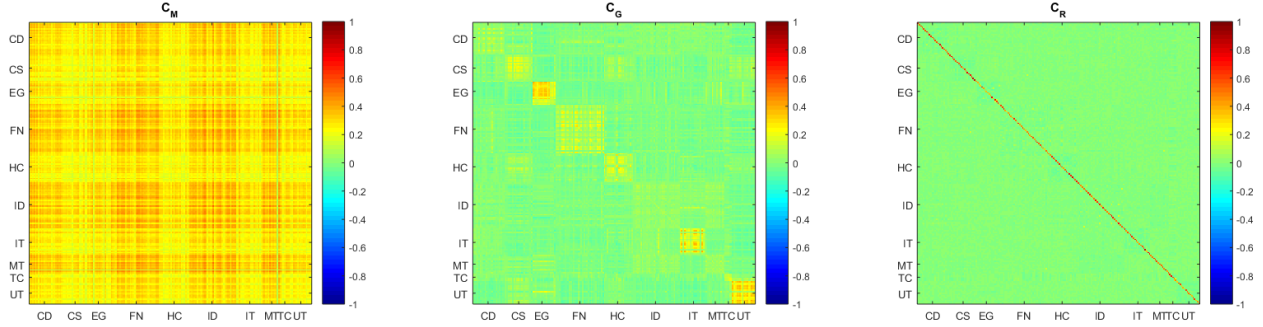


Figure 2.2: **Eigen-decomposition of empirical correlation matrix.** Using eigenvalues and eigenvectors, the entire correlation matrix given in the figure 1.4 is decomposed into market mode C_M , group mode C_G , and random mode C_R . For computing the group mode, $N_G = 20$ was taken. These three matrices are visualized using a color-map.

And the random mode is attributed to the noise (recent studies have looked deeper into the random mode and have found out that it is not that random after all, but it is out of the scope of this thesis). The eigen modes of an empirical cross correlation matrix illustrating this is given in figure 2.2.

2.2.2 Wishart orthogonal ensemble and its eigenspectrum

Let us consider an N dimensional B matrix which is constructed using N random time series of length T . Each of these random time series is uncorrelated white noise, which is defined to be having zero mean and a finite variance. From this, we can construct a Wishart matrix [41] W as:

$$W = \frac{1}{T}BB' \quad (2.3)$$

An ensemble of such many Wishart matrices is called a Wishart orthogonal ensemble (WOE). We can interpret W as a covariance matrix in the context of time series. As the name suggests, W will not have any correlations on average. The spectrum of W 's eigenvalues follow Marchenko - Pastur distribution [42]. For very large N and T and $Q = T/N$ the probability density function of these eigenvalues is defined as:

$$\bar{\rho}(\lambda) = \frac{Q}{2\pi\sigma^2} \frac{\sqrt{(\lambda_{\max} - \lambda)(\lambda - \lambda_{\min})}}{\lambda} \quad (2.4)$$

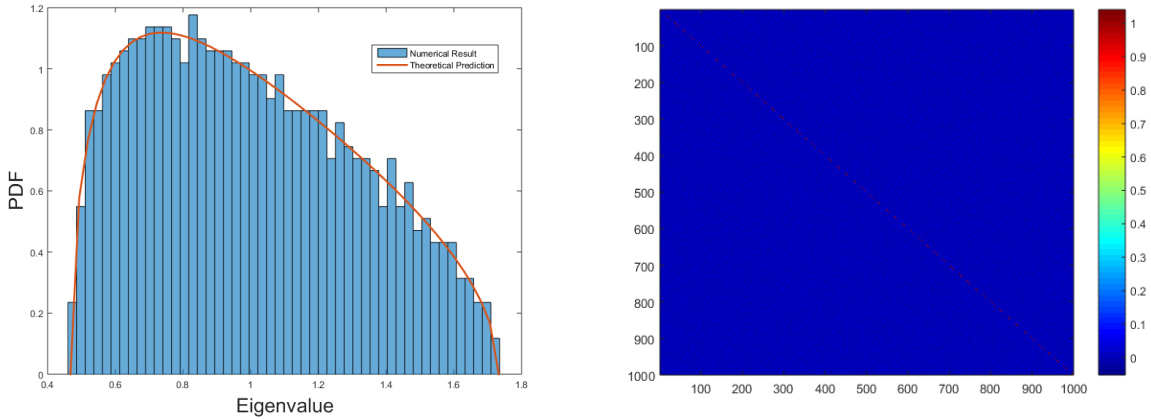


Figure 2.3: **Eigenvalue distribution of a Wishart matrix.** On left, the eigenvalue distribution of a single Wishart matrix computed from 1000 random time series of 10000 time steps is plotted. For comparison, the theoretical Marchenko - Pastur distribution corresponding to this system is also plotted in orange. If we repeat this process multiple times and take an ensemble average of all of them, the histogram will get smoothed out to match the theoretical one. On right the Wishart matrix is also shown.

Where σ is the variance of the coefficients of G while λ_{\min} and λ_{\max} are:

$$\lambda_{\min}^{\max} = \sigma^2 \left(1 \pm \frac{1}{\sqrt{Q}} \right)^2 \quad (2.5)$$

As shown in figure 2.3, one can verify these using computational methods. It is also interesting to notice that when $N > T$ the number of zero eigenvalues will be $N - T + 1$. By giving a small distortion to the matrix, it is possible to break the degeneracy of zero eigenvalues.

2.3 Network representations of correlations

We can approach the task of finding a network representation of the correlation structure in two different ways. First is to consider the market as a complete graph with some weights that represents the correlations between them. Based on the questions asked, one can choose the transformation of the correlation matrix that they want to use as the weights of the edges which will have different interpretations. And second is to impose some rule to determine

the existence of a link (edge) between a pair of stocks and construct a network with these un-weighted edges. For the networks constructed like this, the evolution of topological properties like degrees and clustering will reflect the market dynamics.

In his seminal work, Mantegna [43] introduced the ultrametric distance between return time series, which allowed one to embed the stocks in a metric space where the pairwise distances between the stocks reflect the magnitude of correlation or anti-correlation between them.

$$d_{ij} = \sqrt{\sum_{k=0}^T \left(\frac{x_{ik}}{|\vec{x}_i|} - \frac{x_{jk}}{|\vec{x}_j|} \right)^2} \quad (2.6)$$

Here \vec{x}_i is the normalized returns of the stocks and T is the length of the time window we are using. This measure of distance can be simplified to a transformation of the correlation matrix C_{ij} . Where C_{ij} and d_{ij} is related by the equation

$$d_{ij} = \sqrt{2(1 - C_{ij})} \quad (2.7)$$

These distances can serve as non negative edge weights in a complete graph of stocks whose properties like eigenvector centrality and betweenness centrality has a robust physical interpretation. Along with that having these distances allows one to apply techniques like minimum spanning tree (MST) and threshold graphs to obtain a filtered network representation of the correlation structure.

2.3.1 Minimum spanning trees

One way of getting a network representation of the correlation structure is using the concept of minimum spanning tree. Given an undirected and connected graph $G(E, V)$, it is possible to define at least one subgraph $H(E', V)$ such that all the vertices of G are there in H and the vertices of H (E'), is a subset of the vertices of G (E). The subgraphs which satisfy these conditions are called spanning trees. Out of these, the minimum spanning trees are the ones that minimize the sum of the weights of the edges [44]. There can be more than one MST for a graph. Considering the distance matrix (d_{ij}) computed from the correlation matrix (c_{ij}) as a complete graph with d_{ij} s as edge weights, we can find an MST for the market in every correlation frame [26]. This gives us a refined network representation of the

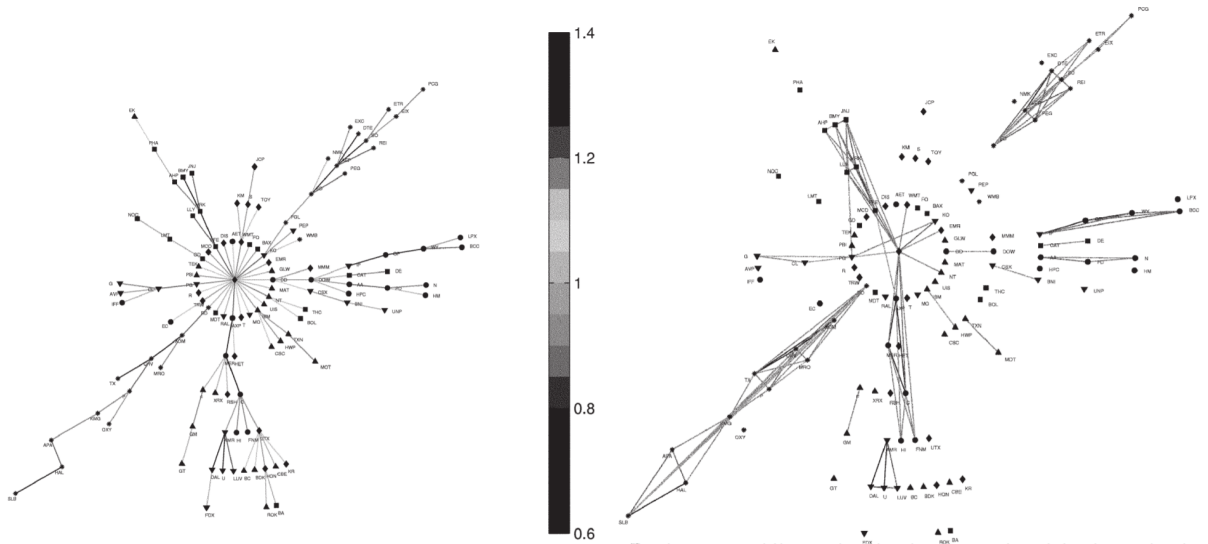


Figure 2.4: **Network representation of correlation matrix.** An MST, shown in left and an asset graph shown in right corresponding to the end date January 1, 1998 taken from the paper of Onnela, Chakraborti and Kaski [45].

correlation structure of the market.

Prim's [46] method is an example of such an algorithm that lets you find the minimum spanning tree from a given connected graph. Starting from a chosen vertex, it grows the tree by adding the edges with the smallest weight that are connected to the nodes which are already part of the tree. This is a greedy algorithm with a time complexity of $O((V + E)\log(V))$.

It has been observed that the topological properties of these minimum spanning trees have a relation with the risk measures of the market. For example, the average length of the edges of the MST is observed to be highly anti-correlated with the risk of the minimum risk Markowitz portfolio [47].

2.3.2 Asset graphs

Another way to deduce an underlying network structure from the correlations is to use a threshold on the correlation matrix or the distance matrix to obtain an adjacency matrix. With the thresholds θ and θ' on correlation matrix C and the distance matrix d respectively

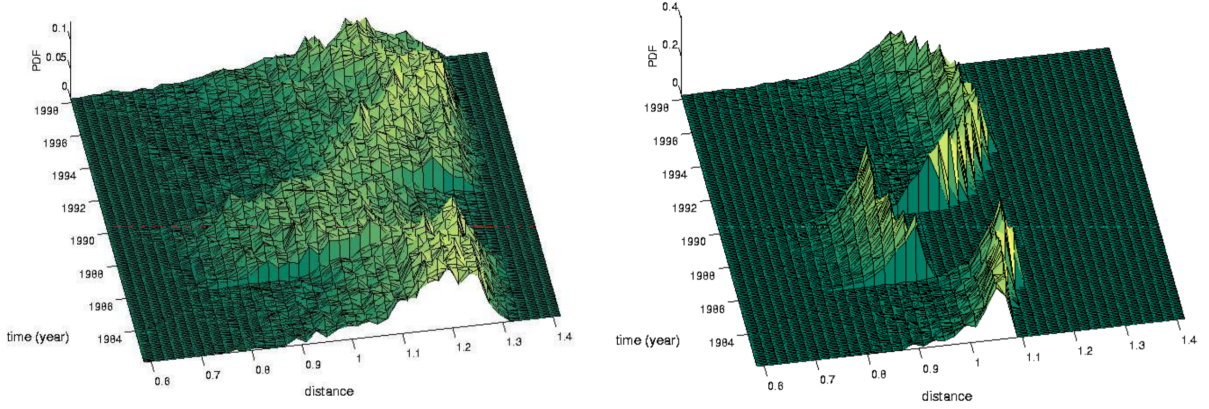


Figure 2.5: **The time evolution of distribution of distances.** The time evolution of the probability density of distances corresponding to the edges in the MSTs (left) and the asset graphs (right) are plotted. The figure is taken from the work of Onnela et. al [45].

(where $\theta' = \sqrt{2(1-\theta)}$), the relations between C , d and the adjacency matrix a can be written as:

$$a_{ij} = \begin{cases} 1, & \text{if } C_{ij} \geq \theta \\ 0, & \text{if } C_{ij} < \theta \end{cases} \quad (2.8)$$

or,

$$a_{ij} = \begin{cases} 1, & \text{if } d_{ij} \leq \theta' \\ 0, & \text{if } d_{ij} > \theta' \end{cases} \quad (2.9)$$

In the network representation that is obtained this way, the parameter θ controls network properties such as mean degree and mean clustering coefficient. Instead of fixing θ one can also fix the number of edges chosen. This is analogous to choosing a fixed number of smallest distances and drawing the graph using them as edges. The network representation of the correlation structure obtained this way is called an asset graph.

An example of such an asset graph constructed using $N - 1$ edges (equal number of edges as the MST) is given in figure 2.4. The evolution of the distributions of edge weights in both MST and asset graph is depicted in figure 2.5. Asset graph's distance distribution is characterized by a sharp cut off whereas the MST's distribution is more spread-out. A drastic shift towards the smaller distance values in the distribution can be observed during the financial crisis of 1987. An interesting observation we can make in figure 2.4 is that, in both the MST and the asset graph of the stocks of S&P-500, similar sectors are appearing together in close positions.

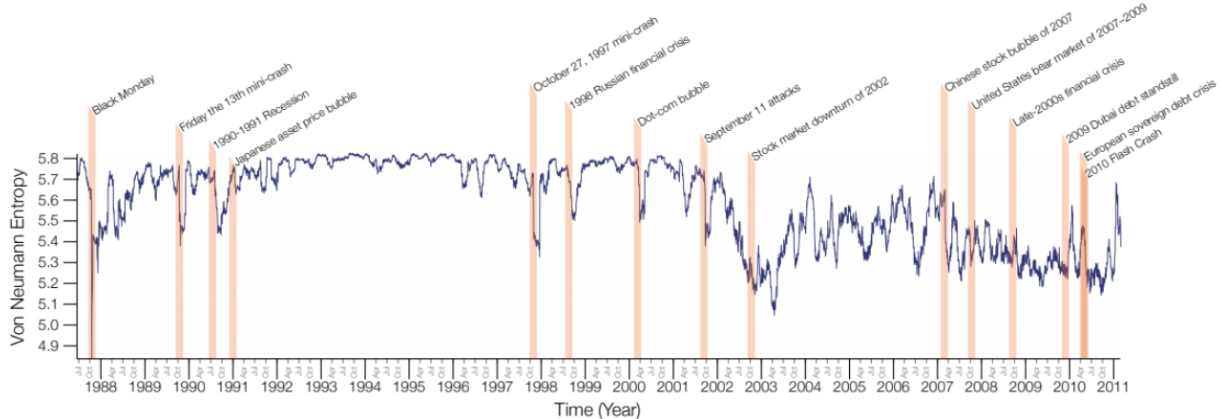


Figure 2.6: **Signatures of historical events in von Neumann entropy.** von Neumann entropy calculated for New York Stock Exchange data for the time period 1988 to 2011 and is plotted in this figure which is taken from the paper of Filipi N Silva [49]. The important historical events are marked on the timeline with a red shadow.

2.4 Entropy measures and application

The entropy is one of the fascinating quantities that one can define in the context of a complex system. Philosophically there are ontological motivations like the proposal of John Wheeler “to take information as the primitive component of reality from which other physical properties are derived” [48], as well as the usual epistemological motivations that drive one to define and measure the entropy of a complex system. The network representations of the correlation matrices provide an excellent framework to define various entropy measures that focus on different types of information flow in this system. In the case of financial markets, this entropy measures can be used to monitor the evolving network structure.

2.4.1 Von Neumann entropy

In quantum mechanics, given a density matrix, one can calculate the entropy of the system as defined by von Neumann [50], from its normalized eigenvalues. This definition of entropy could be extended to networks by considering its normalized Laplacian matrix to be a density matrix [51, 52, 53]. After constructing the adjacency matrix from the correlation matrix by applying a threshold, this normalized Laplacian matrix can be calculated using the degrees of the nodes d_i s as:

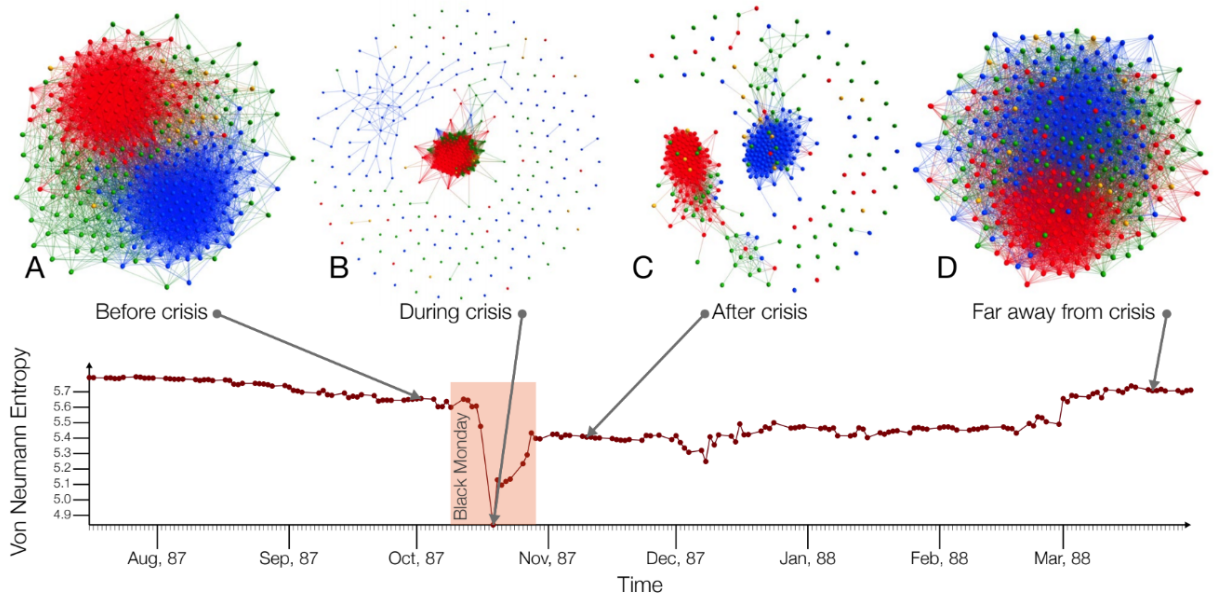


Figure 2.7: **Behaviour of von Neumann entropy during a financial crisis.** This diagram, taken from [49], depicts the evolution of the community structure in NYSE stocks during a financial crisis. In this particular case the event in observation is 1987 Black Monday crash.

$$\tilde{L}_{ij} = \begin{cases} 1, & \text{if } i = j \text{ and } d_i \neq 0 \\ -\frac{1}{\sqrt{d_i d_j}}, & \text{if } i \neq j \text{ and } (i, j) \in E \\ 0, & \text{otherwise} \end{cases} \quad (2.10)$$

From its eigenvalues, the von Neumann entropy is calculated by applying Shannon's formula [49]. It is given as:

$$S^\circ = - \sum_{i=1}^n \frac{\lambda_i}{n} \ln \frac{\lambda_i}{n} \quad (2.11)$$

The evolution of von Neumann entropy calculated like this is depicted in figure 2.6. During critical events in the market, one can observe that this entropy is decreasing. As we have observed in previous instances, the entire market starts behaving like a singular community during these events, which causes this behaviour. This dynamics is depicted in figure 2.7 where we can see the evolution of community structure in the market through the

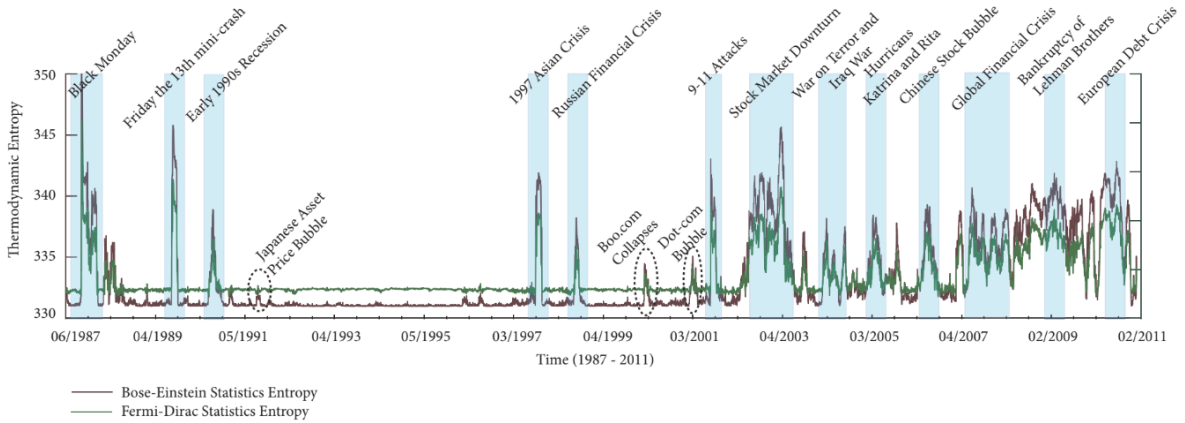


Figure 2.8: **Signatures of historical events in thermodynamic entropies.** Thermodynamic Entropy in NYSE (1987-2011) derived from Bose-Einstein and Fermi-Dirac statistics taken from [54]. Critical financial events are marked with light blue shade.

crisis period of the 1987 Black Monday crash. The nodes of the correlation network collapsed into a single community which in turn got captured by the entropy measure.

2.4.2 Thermodynamic entropy

If we take the normalized network Laplacian L as the Hamiltonian H , one can calculate the energy levels of the network by looking at the eigenvalue of L . Assuming these energy states are occupied by particles that follow quantum statistics, we can calculate an entropy by writing its partition function [55].

Based on the two kinds of occupation statistics of particles, one can get two different expressions for entropy. If one was to use Bose-Einstein statistics [56] which consider all the particles as indistinguishable and allows infinite occupation of a particular state, you would get the thermodynamic entropy S'_{BE} to be:

$$\begin{aligned}
 S'_{BE} &= \log Z - \beta \frac{\partial \log Z}{\partial \beta} \\
 &= \sum_{i=1}^V \log (1 - e^{\beta(\mu - \varepsilon_i)}) - \beta \sum_{i=1}^V \frac{(\mu - \varepsilon_i) e^{\beta(\mu - \varepsilon_i)}}{1 - e^{\beta(\mu - \varepsilon_i)}}
 \end{aligned} \tag{2.12}$$

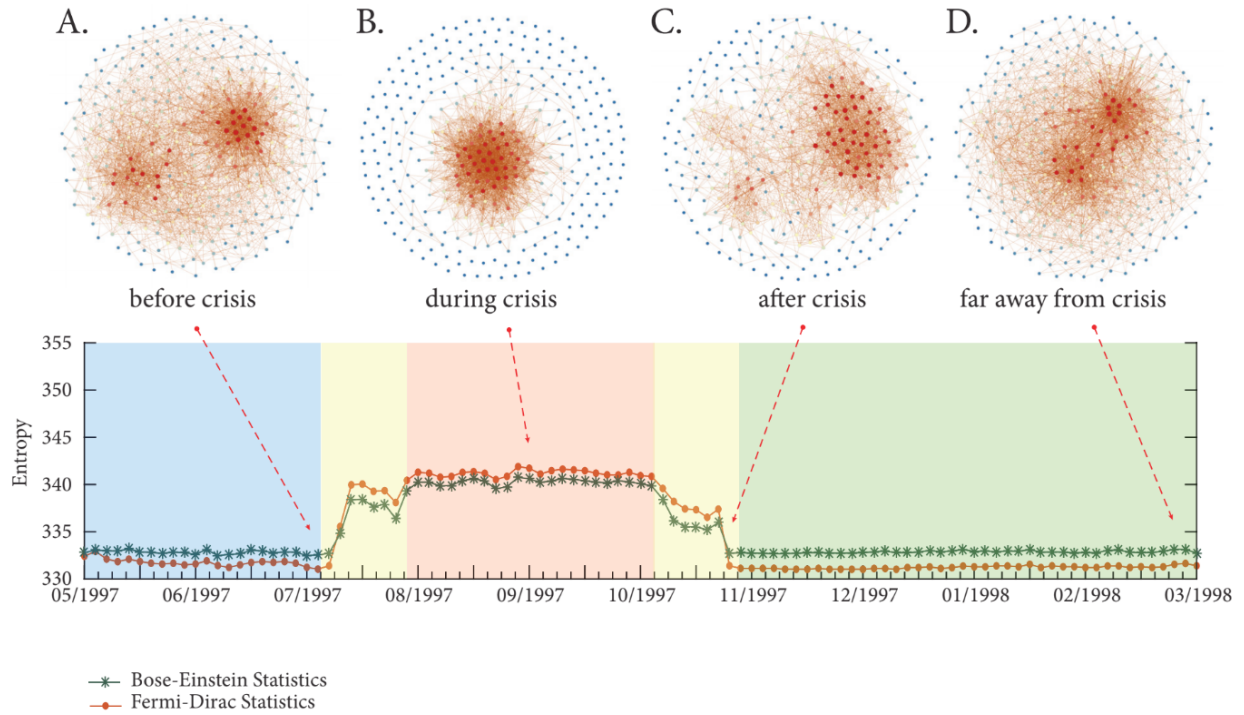


Figure 2.9: **Behaviour of thermodynamic entropy during a financial crisis.** During 1997 Asian financial crisis the evolution of von Neumann entropy is depicted in this figure that is taken from [54]. The change in the network structure (especially the connected components) across this critical event can be observed in the representations (A), (B), (C) and (D).

Here μ that is analogous to the chemical potential is a control parameter. But if you put the additional constraint that the number of particles in the energy level should be non-negative, the condition that $\mu \leq \min \varepsilon_i$.

On the other hand, if you assume that the particles are like indistinguishable fermions which obeys Pauli's exclusion principle and have Fermi-Dirac occupation statistics [57], the expression for the thermodynamic entropy S'_{FD} , will be:

$$\begin{aligned} S'_{FD} &= \log Z - \beta \frac{\partial \log Z}{\partial \beta} \\ &= \sum_{i=1}^V \log (1 + e^{\beta(\mu - \varepsilon_i)}) - \beta \sum_{i=1}^V \frac{(\mu - \varepsilon_i) e^{\beta(\mu - \varepsilon_i)}}{1 + e^{\beta(\mu - \varepsilon_i)}} \end{aligned} \quad (2.13)$$

Since only one particle can occupy a single energy state, the “chemical potential” is equal to the n^{th} energy level. That is $\mu = \varepsilon_n$.

The historical evolution of this entropy in the stock market, as shown in figure 2.8 tells us that during a market crisis or a critical event, the entropy measures S'_{FD} and S'_{BE} both goes up. But as one can observe, the particular entropy which has higher values is different in different periods. This interplay happens because of the difference in the range of eigenvalue spectrum that these two measures focus on. S'_{BE} concentrate on the lesser Laplacian eigenvalues since the BE statistics favour placing more particles in smaller energy states, whereas S'_{FD} looks at more number eigenvalues since the particles are forced to occupy the higher ones due to Pauli's exclusion principle. From a network perspective, this means that the S'_{BE} reflects the overall network-structure by quantifying the extent of bi-partiality and how many connected components are there [58]. And S'_{FD} capture smaller and subtler deviations inside the structure of the network [59]. Figure 2.9 exposes these dynamics and characterizes it across a crisis period by tracking the evolution of the network structure as well as the entropies.

2.4.3 Structural entropy

Structural entropy is a novel measure proposed by Almog et al. [60], which is calculated from the community structure of the market. Similar to a diversity measure, the structural entropy

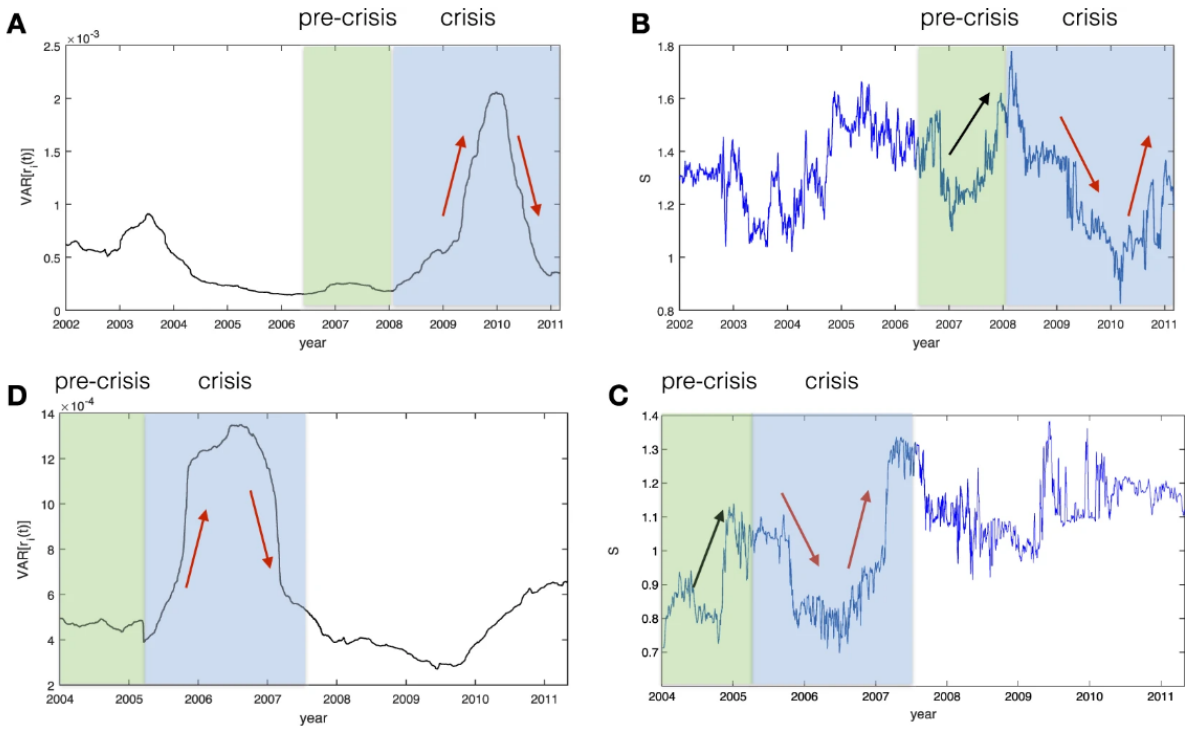


Figure 2.10: **Evolution of Structural Entropy during a financial crisis.** Evolution of volatility and structural entropy for FTSE-100 (A and B) and NIKKEI-225 (C and D) around the 2008 Lehmann brother's bankruptcy. The figure is taken from [60].

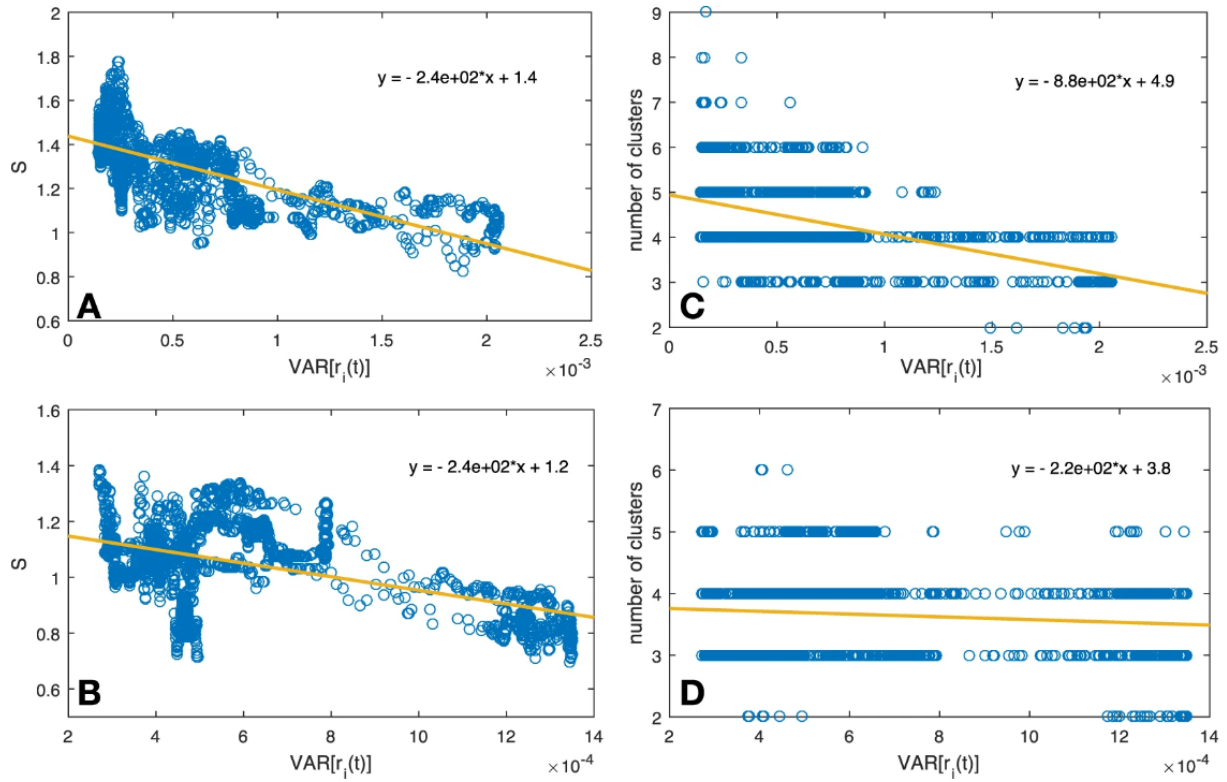


Figure 2.11: **Relation between S and VaR.** Plot of the relation between volatility (X axis) and structural entropy (Y axis) for FTSE-100 (A) and NIKKEI-225 (B). For reference we also plotted the relation between volatility (X axis) and the number of detected communities (Y axis) for FTSE-100 (C) and NIKKEI-225 (D). The figure is taken from [60].

compares the relative sizes of the communities which were found out using a community detection algorithm from the correlation matrix.

S is defined in the context of a network. The methodology behind its computation is as follows. Let G be a network with n number of nodes. We then determine the communities of the network by applying any community detection algorithm on G . As a result, we obtain an n -dimensional vector $v = (v_1, v_2, \dots, v_n)$, where each v_i , is the community to which node $i = 1, 2, 3 \dots n$ is assigned. Let k be the total number of communities detected (which will be $\max(v)$). Probability of extracting a node from a particular community j is given as:

$$P_j = c_j/n \quad (2.14)$$

where c_j is the size of community j . From this we compute S by applying Shannon entropy formula to probabilities:

$$S \equiv H(P) = - \sum_{j=1}^k P_j \log P_j \quad (2.15)$$

A significant part of this method is the detection of communities in the correlation structure. One way to approach this step is to deduce an underlying network structure using a threshold on the correlation or the distance matrix to obtain an adjacency matrix. Since the value of the threshold has a huge effect on the resulting network structure, this method is hence not much robust. Instead of the threshold method one can directly find the community structure of the market from the correlation matrix C using the community detection algorithm proposed by Garlaschelli [38]. This algorithm utilizes random matrix theory to filter out noise and take out only the C_G which has the information about the community structure.

One of the intriguing outcomes of continuously monitoring the structural entropy was its high negative correlation with the market volatility as shown in figure 2.11. This is consistent with our prior understanding that the market volatility is positively correlated with the average market correlation (μ). As μ increase, the market starts behaving more like a singular community which decreases the diversity in the community structure and as a result, decreases the structural entropy. This was best observed during the financial crisis of 2007-2008 (figure 2.10). S increased initially as the market approached the crisis, and during the crisis, it decreased first and then increased analogous to the movement of the volatility

which did the opposite.

2.5 Remarks

In this chapter, we have discussed the methods that people have used to obtain a network representation and entropy measures of these networks. The motivation behind these measures is that by following them, one can get insights about the current structure of the underlying correlation network of the financial market. But all of these entropy measures suffer from a general problem. They are too much reliant on the process of obtaining the network structure from the correlation matrix. And this process loses information no matter what method you are using. As an example, let us consider the two methods that we have discussed. If we are using minimum spanning trees as the filtration method, the network representation we get will contain no loops. On the opposite, if we apply a threshold method, the network representation that we obtain will contain loops, but the components will be disconnected. One can attempt to combine these two methods by superimposing the two resultant network representations on top of each other. Still, even this will have a dependency or arbitrariness on the threshold that we are using.

We are proposing a new entropy measure that can be calculated directly using the entire correlation matrix. This measure solves the problem of information loss due to the filtration process that the other methods rely on, by considering the correlation matrix as a complete graph with weighted edges. And also, our method allows us to incorporate our understandings from the random matrix theory into this network representation by enabling us to calculate the entropy of the partial correlation structures (market mode, group mode, and random mode).

Chapter 3

Phase separation using eigen-entropy

3.1 Introduction

Recently, Pharasi et. al [61, 62] used the tools of random matrix theory to determine market states and confirmed that during a market crash all the stocks start behaving similarly and the whole market begins to act like a single huge cluster or community. In contrast, during a bubble period, a particular sector gets overpriced or over-performs, causing accentuation of disparities among the various sectors or communities. The eigenvalue decomposition of a correlation matrix into partial correlations; market, group and random modes [62], enables identification of dominant stocks (influential leaders) and sectors (communities). The correlation-based network of leaders and communities changes with time, especially during market events like crashes, bubbles, etc. Thus, if one were able to monitor the evolution of this network structure continuously [61, 60, 63], one would be able to acquire useful insights that would help in developing better investment strategies, manage risk and stress-test the global financial system.

Here, our aim is to extract information about the “disorder” (or randomness) in the market and its eigenmodes, using the entropy measure – *eigen-entropy* [64], computed from the *eigen-centralities* (ranks) [65, 31] of different stocks in the correlation-network. The relative-entropy measures computed for these eigenmodes enable us to construct a phase space, where the different market events undergo phase-separation (akin to many physical or biophysical phenomena [66, 67, 68, 69, 70]) and display “order-disorder” transitions as in

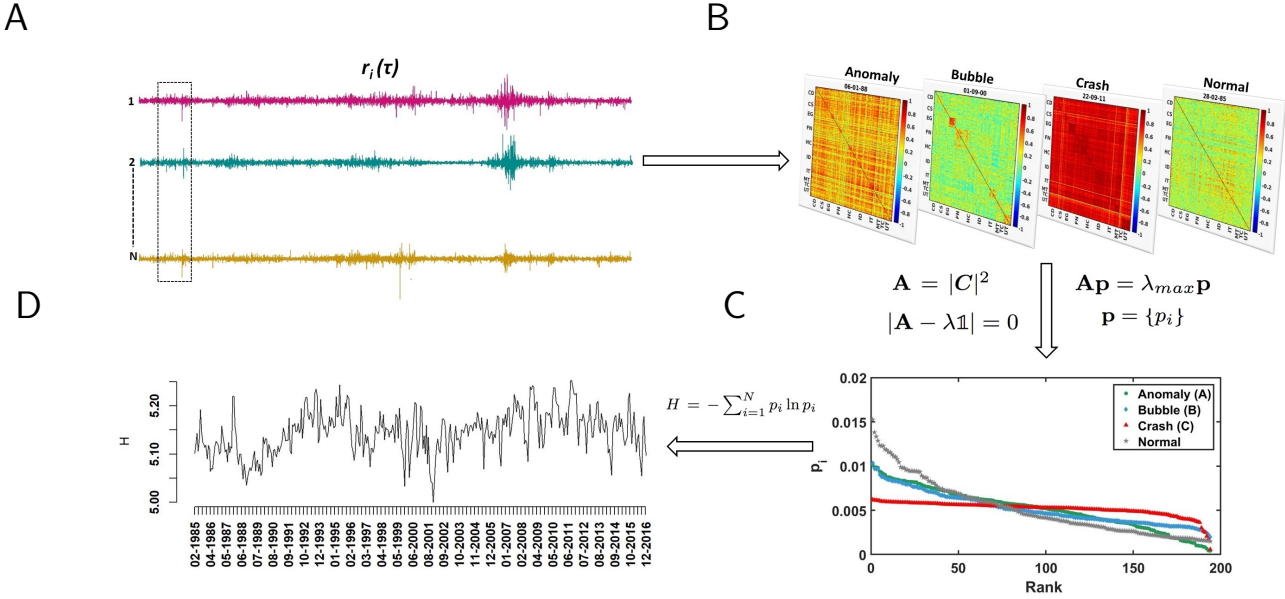


Figure 3.1: **Schematic diagram for computation of eigen-entropy from market returns.** (A) Return time-series plots for three arbitrarily chosen stocks (out of a total N stocks), with a chosen epoch (of size M days) ending on day τ for the computation of Pearson correlation coefficients. (B) Four chosen cross-correlation matrices $C(\tau)$: anomaly (06/01/1988), bubble (01/09/2000), crash (22/09/2011) and normal (28/02/1985) periods, in the S&P-500 market; the stocks are arranged according to their sectors. The sectoral abbreviations are: **CD**–Consumer Discretionary; **CS**–Consumer Staples; **EG**–Energy; **FN**–Financial; **HC**–Health Care; **ID**–Industrials; **IT**–Information Technology; **MT**–Materials; **TC**–Telecommunication Services; and **UT**–Utilities. We define $\mathbf{A} = |C|^2$ (matrix element-wise) and use the characteristic equation $|\mathbf{A} - \lambda \mathbf{1}| = 0$ to compute the eigenvalues $\{\lambda_1 \dots \lambda_N\}$; we denote the maximum eigenvalue as λ_{max} and the eigenvector corresponding to the maximum eigenvalue as \mathbf{p} , such that $\mathbf{A} \mathbf{p} = \lambda_{max} \mathbf{p}$. The normalized eigenvector has components: $\mathbf{p} = \{p_i\}$, that are known as eigen-centralities. (C) The ranked (sorted) eigen-centralities $\{p_i\}$ of the normalized eigenvector corresponding to the maximum eigenvalue are plotted, for the anomalous (green circles), type-1 (blue diamonds), crash (red triangles) and normal (grey stars) periods of the financial market. (D) Eigen-entropy ($H = -\sum_{i=1}^N p_i \ln p_i$), evaluated from the correlation matrices using a rolling epoch of $M = 40$ days and a shift of $\Delta = 20$ days, is plotted for the 32-year period 1985-2016.

critical phenomena in physics [10, 11]. This type of behavior has never been recorded for financial markets, and is distinct from the two-phase behavior in financial markets reported earlier by Plerou et al. [71]. One of the relative entropy measures displays “universal scaling” behavior with respect to the mean market correlation; there is a data collapse, which suggests that the fluctuations in price returns for different financial assets, varying across countries, economic sectors and market parameters, are governed by the same statistical law. This apparent universal behavior may motivate us to do further research as to determine which market forces are responsible for driving the market or are important for determining the price co-movements and correlations. Further, a functional of the relative entropy measure acts as a good market indicator, as it can gauge the market “fragility” (minimum risk of the market portfolio) and the “market fear” (volatility index). This new and simple methodology helps us to better understand market dynamics and characterize the events in different phases as anomalies, bubbles, crashes, etc. that display intriguing phase separation and universal scaling behavior. In addition, this may lead to a foundation for understanding scaling and universality in a broader context, and providing us with altogether new concepts not anticipated previously. This methodology may be generalized and used in other complex systems to understand and foresee tipping points (similar to market crashes and bubbles) and fluctuation patterns.

3.2 Monitoring eigen-entropy

3.2.1 Eigenvector centrality

Generally, for any given graph $G := (N, E)$ with $|N|$ nodes and $|E|$ edges, let $A = (a_{i,j})$ be the adjacency matrix, such that $a_{i,j} = 1$, if node i is linked to node j , and $a_{i,j} = 0$ otherwise. The relative centrality, p_i , score of node i can be defined as: $p_i = \frac{1}{\lambda} \sum_{v \in M(i)} p_j = \frac{1}{\lambda} \sum_{j \in G} a_{i,j} p_j$, where $M(i)$ is a set of the neighbors of node i and λ is a constant. With a small mathematical rearrangement, this can be written in vector notation as the eigenvector equation:

$$\mathbf{A}\mathbf{p} = \lambda\mathbf{p}.$$

In general, there can exist many different eigenvalues λ for which a non-zero eigenvector solution exists. We use the characteristic equation:

$$|\mathbf{A} - \lambda \mathbf{1}| = 0$$

to compute the eigenvalues $\{\lambda_1 \dots \lambda_N\}$. However, the additional requirement that all the entries in the eigenvector be non-negative ($p_i \geq 0$) implies (by the Perron–Frobenius theorem) that only the maximum eigenvalue (λ_{max}) results in the desired centrality measure. The i^{th} component of the related eigenvector then gives the relative *eigen-centrality* score of the node i in the network. However, the eigenvector is only defined up to a common factor, so only the ratios of the centralities of the nodes are well defined. To define an absolute score one must *normalise* the eigenvector, such that the sum over all nodes N is unity, i.e., $\sum_{i=1}^N p_i = 1$. Furthermore, this can be generalized so that the entries in \mathbf{A} can be any matrix with real numbers representing the connection strengths. For correlation matrices $\mathbf{C}(\tau)$, in order to enforce the Perron–Frobenius theorem, we work with $\mathbf{A} = |\mathbf{C}|^n$, where n is any positive integer (we have used $n = 2$ in this study).

Effect of changing the value of n is shown in figure 3.11. As observed the values of eigen-entropy H differ with the variation of the power n of correlation matrices. The variation is because, with the increase in power, the dissimilarities in the elements of the correlation matrix are amplified, which will then, in turn, changes the centrality of the matrix. For very high powers the transformed correlation matrices will act like an adjacency matrix with very high values (close to 1s) and very low values (close to 0s). It is also interesting to note that, depending on the problem, we can decide the range of correlations to focus on by adjusting the power of the elements of the correlation matrix.

One may argue that the information about the anti-correlations is lost during this process. But if we consider just the existence of a correlation (doesn't matter whether its positive or negative) as the edge between two time series, the squared values or the absolute values of the correlations indicate the weights of that link.

3.2.2 Eigen-entropy using C , C_M , and C_{GR}

Following the tradition in information theory, we propose a new measure, the eigen-entropy $H = -\sum_{i=1}^N p_i \ln p_i$, since all the normalised eigen-centralities are non-negative ($p_i \geq 0$)

and $\sum_{i=1}^N p_i = 1$, as explained above. The eigen-entropy can be described as a measure of disorder, or the degree of randomness in the matrix $\mathbf{A} = |\mathbf{C}|^2$; higher the eigen-entropy, higher is the disorder in the matrix; the highest being in the case of WOE, where $H \sim \ln N$. The detailed schematic diagram of the methodology is given in figure 3.1 and the calculated time series are plotted in the figure 3.3.

The eigen-entropies may be computed (see Methods) from the full correlation C , market mode C_M and group-random mode C_{GR} . Figure 3.3 (A) and (C) show the evolution of market returns $r(\tau)$, mean market correlations $\mu(\tau)$, and different eigen-entropies $H(\tau)$, $H_M(\tau)$, and $H_{GR}(\tau)$ (shown in different colors; see legend), for S&P-500 and Nikkei-225 markets, respectively. The vertical dashed lines correspond to some indicative dates for type-1 events (blue) and crashes (red) (see Table 4.2). These eigen-entropies can then be used for the characterization of market events, such as bubbles and crashes.

We have used the eigenvalue decomposition of the correlation matrices into a market mode C_M , group mode C_G and a random mode C_R and a composite group and random modes C_{GR} . From such a decomposition, it is also possible to reconstruct the correlation matrix as aggregates of the contributions of modes C_M, C_G , & C_R or C_M & C_{GR} as we have shown before. For empirical matrices (especially the ones using shorter window size), it is very difficult to determine the exact value of λ_+ and hence figure out N_G , for which the eigenvectors from 2 to N_G would describe the sectoral dynamics. Here, we choose $N_G = 20$ arbitrarily for the correlation decomposition (figure 3.2), corresponding to the 20 largest eigenvalues after the largest one.

In order to avoid the arbitrariness, we prefer the following decomposition:

$$C = C_M + C_{GR} \quad (3.1)$$

$$= \lambda_1 e_1 e_1' + \sum_{i=2}^N \lambda_i e_i e_i'. \quad (3.2)$$

Figure 3.2 shows the eigenvalue decompositions of the correlation matrices, for (A) normal, (B) anomalous, (C) bubble, (D) crash periods of the financial market, corresponding to the frames in figure 3.1, and in addition (E) shows the results for a random matrix taken from a Wishart orthogonal ensemble (WOE), where we have denoted the different matrices as: full correlation C , market mode C_M , group mode C_G , random mode C_R , group-random

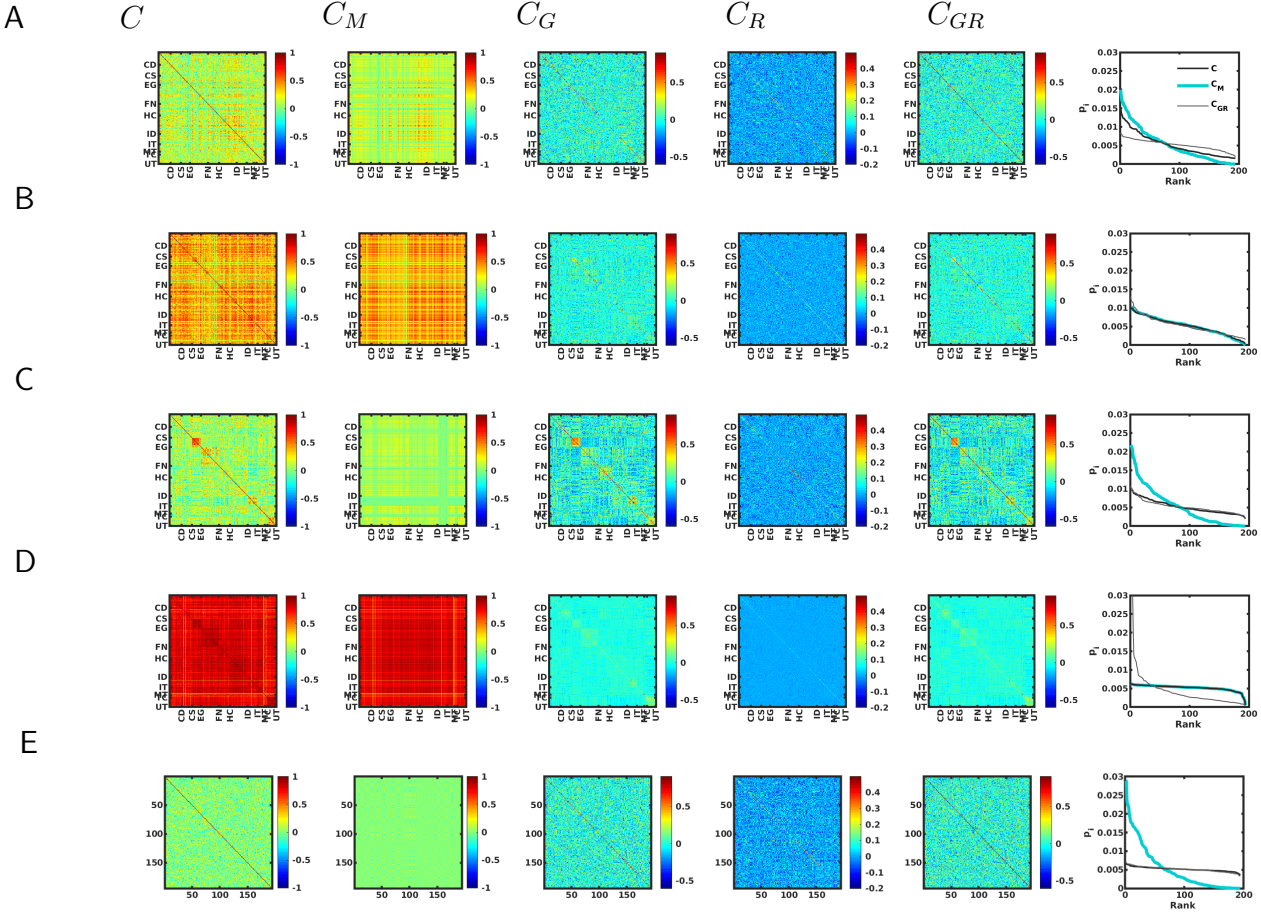


Figure 3.2: **Eigenvalue decomposition of the correlation matrices.** For (A) normal, (B) anomalous, (C) bubble, (D) crash periods of the financial market, as in figure 3.1, and (E) random matrix taken from uncorrelated WOE. (Left to right) Plots showing the correlation matrices: full C , market mode C_M , group mode C_G , random mode C_R , group-random mode C_{GR} and the ranked eigen-centralities (p_i) of the different correlation modes: full (C in black curve), market mode (C_M in turquoise curve) and group-random mode (C_{GR} in grey curve). Interestingly, for a normal period, the three curves are distinct and there are hierarchies in ranks in all curves; for the market anomaly, all the three curves almost coincide; for the bubble period, the curves corresponding to the full and the group-random modes coincide while there is a strict hierarchy in the eigen-centralities of the market mode; for crash period, the curves corresponding to the full and the market modes coincide while there is a strict hierarchy in the eigen-centralities of the group-random mode; and for the WOE, once again the curves corresponding to the full and the group-random modes coincide while there is a strict hierarchy in the eigen-centralities of the market mode. This feature is then exploited in characterizing the market events into anomalies, bubbles, crashes, normal periods, etc. with the help of the corresponding entropy functions as in figure 3.3 and figure 3.6.

mode C_{GR} and displayed the results in figure 3.2 (*Left to Right*). The last column shows the results for the ranked eigen-centralities (p_i) of the different correlation modes: full (C in black curve), market mode (C_M in turquoise curve) and group-random mode (C_{GR} in grey curve). Interestingly, for a normal period, the three curves are distinct and there are hierarchies in ranks in all curves; for the market anomaly, all the three curves almost coincide; for the bubble period, the curves corresponding to the full and the group-random modes coincide while there is a strict hierarchy in the eigen-centralities of the market mode; for crash period, the curves corresponding to the full and the market modes coincide while there is a strict hierarchy in the eigen-centralities of the group-random mode; and for the WOE, once again the curves corresponding to the full and the group-random modes coincide while there is a strict hierarchy in the eigen-centralities of the market mode. This feature is then exploited in characterizing the anomalies, bubbles, crashes and normal periods in the market, with the help of the corresponding entropy functions as explained below and displayed in figure 3.3 and figure 3.6.

As a standard, we can use a Wishart orthogonal ensemble(WOE) to represent a correlation matrix computed from a set of totally random time series. The eigen-entropy for a WOE can be calculated from the ensemble average of the principal eigenvector components, which intuitively will be equal to $\frac{1}{N}$ because all the nodes will have equal importance in a totally random market. So for a totally random case, the eigen-entropy will be $H = \log(N)$. figure 3.4 (A) shows the plot of sorted eigen-centralities p_i against rank, computed from the normalized eigenvectors corresponding to the maximum eigenvalues for 1000 independent realizations of a Wishart orthogonal ensemble (WOE). Filled black squares represent the mean eigen-centralities computed from 1000 independent realizations of the WOE, that serves as a reference (the maximum disorder or randomness) in the market correlation with $N = 194$. figure 3.4 (B) shows the plot of the variation of eigen-entropy H as a function of system size (correlation matrix size) N , where each point represents a mean computed from 1000 independent realizations of an uncorrelated WOE. The theoretical curve (red dash) shows the variation $\sim \ln N$. figure 3.4 (C) shows the histograms of the eigen-centralities p_i for typical anomaly (06/01/1988) (green circles), Dot-com burst (01/09/2000) (blue diamonds), crash (22/09/2011) (red triangles) and normal (28/02/1985) (grey stars) and WOE (black squares). One can observe that unlike the ideal case there is an observable spread in the centrality values. This is happening due to the process of sorting that is destroying the information about the node that each of the values belong to. If it were not sorted and the information about the nodes was preserved, all the values will average out to $1/n$. It is still

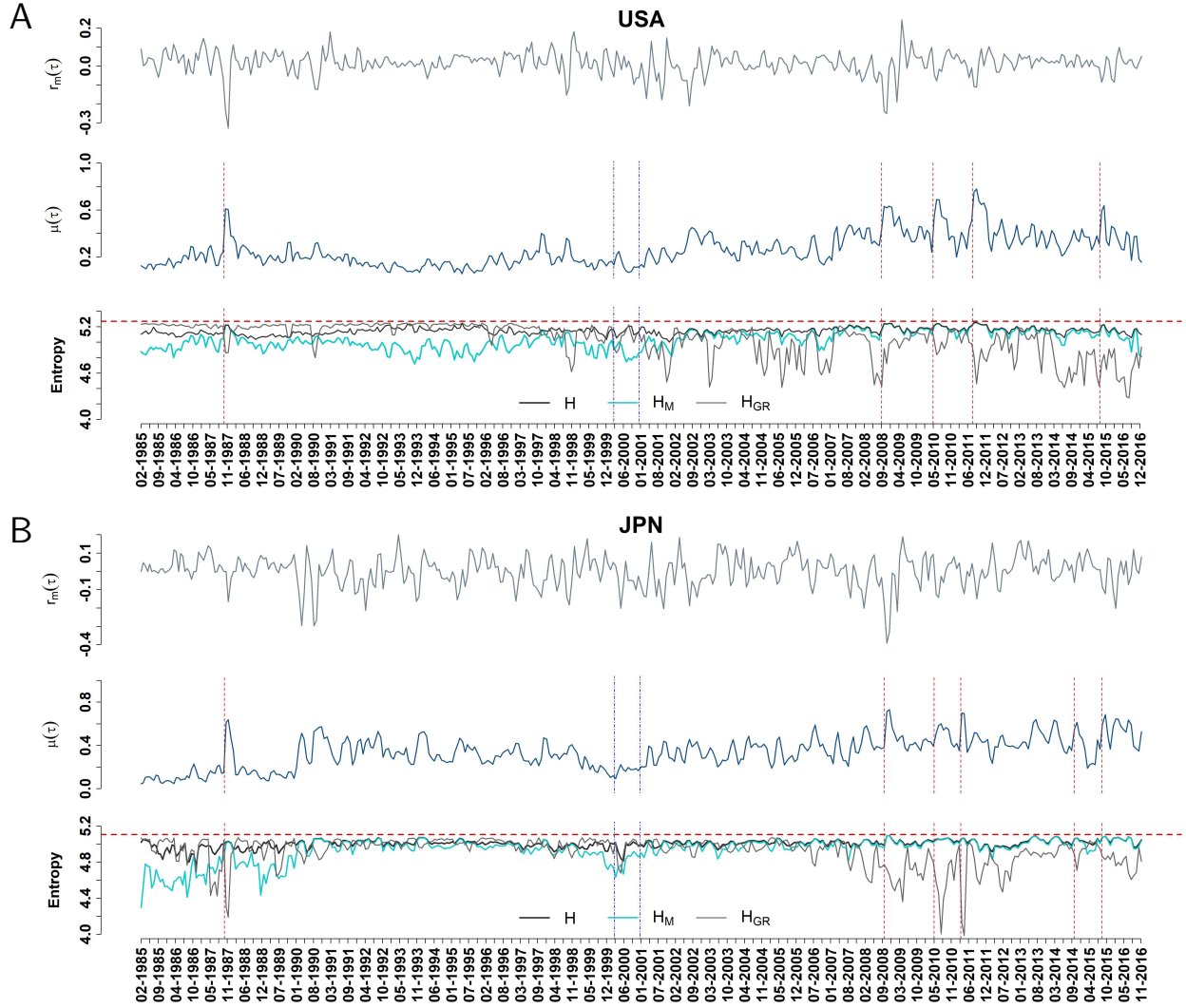


Figure 3.3: **Evolution of market returns ($r(\tau)$), mean market correlations ($\mu(\tau)$), and eigen-entropies.** The eigen-entropies are computed from the full correlation, market mode and group-random mode (shown in different colors; see legend), for (A) S&P-500 and (B) Nikkei-225 markets.

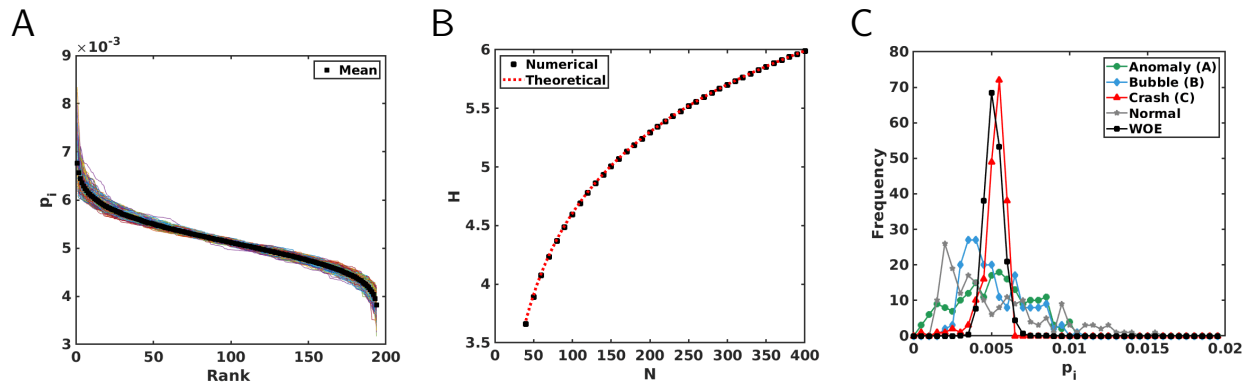


Figure 3.4: **Eigen-centralities (ranks) and eigen-entropy.** (A) Plots of sorted eigen-centralities p_i against rank, computed from the normalized eigenvectors corresponding to the maximum eigenvalues for 1000 independent realizations of a Wishart orthogonal ensemble (WOE). Filled black squares represent the mean eigen-centralities computed from 1000 independent realizations of the WOE, that serves as a reference (the maximum disorder or randomness) in the market correlation with $N = 194$. (B) Plot showing the variation of eigen-entropy H as a function of system size (correlation matrix size) N , where each point represents a mean computed from 1000 independent realizations of a WOE. The theoretical curve (red dash) shows the variation $\sim \ln N$. (C) Histograms of the eigen-centralities p_i for typical anomalous (green circles), bubble (blue diamonds), crash (red triangles) and normal (grey stars) and WOE (black squares).

important to do the process of sorting because the deviation from the ideal value signifies the spurious correlations which also will be present in the real correlation matrix.

3.3 Order-disorder transitions and phase separation.

Using the three entropies that we calculated, it is possible to construct a phase space through which the market moves as it evolves with time. We can embed the frames (Each corresponding to a period of time) in this phase space to characterize the regions that they are occupying and to look for order-disorder transitions. We used a *rolling mean* and *rolling standard deviation* (with a window size of 40 days), and computed the *standardized values* of eigen-entropies H^{Std} , H_M^{Std} and H_{GR}^{Std} . The figures 3.8 and 3.9 show the 3D-plots of the standardized values of eigen-entropy (H^{Std}) corresponding to the full (along z-axis), eigen-entropy (H_{GR}^{Std}) corresponding to the group-random (x-axis), and eigen-entropy (H_M^{Std}) corresponding to the market mode (along y-axis), for S&P-500 and Nikkei-225 markets, respectively. The sequence of frames display the “order-disorder” transitions in case of the events given in the table 4.2. Even though H_{GR} is becoming very low in all of the observed events, the dynamics of H and H_M is different for different events. This can be seen in figure 3.7. This observation hints towards the underlying categorization of the events that I will be discussing next.

We compute the relative-entropies $H - H_M$, $H - H_{GR}$, and $H_M - H_{GR}$, starting from the eigen-entropies corresponding to the full correlation, market mode and group-random mode, respectively. We then use these new variables to characterize and identify the different market events as crashes, normal periods and three more different kinds of interesting events. Of these interesting events specifically in type-1,2, many frames corresponds to periods in which there are recorded calamity in the market like Burst of the Dot-com bubble or loss due to Katrina-Rita hurricane. Anomalies on the other hand are interesting due to its fascinating centrality distribution across the three modes. Figure 3.6 (B) and (D) show the 2D-plots of the phase space using relative-entropies $H - H_M$, and $H - H_{GR}$, for S&P-500 and Nikkei-225 markets, respectively. As evident, the epochs (event frames) clearly undergo “phase separation” – segregate into different market events: anomalies (green), type-1 (light blue), type-2 (blue), crashes (red) and normal (grey). The results can be compared to the benchmarks of WOE (see figures 3.4 and 4.2) for both USA and JPN. For the first

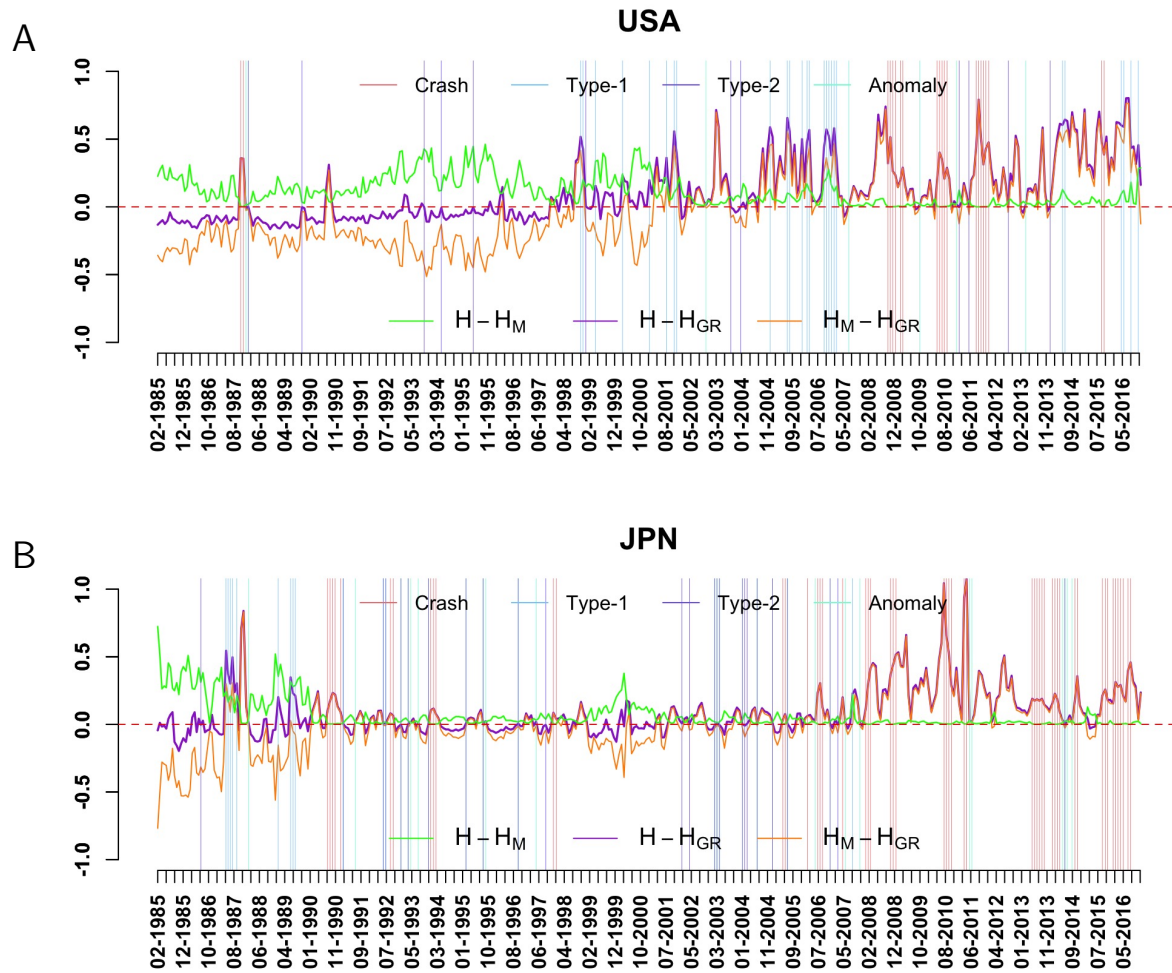


Figure 3.5: **Evolution of relative-entropies.** For (A) S&P-500 and (B) Nikkei-225 markets, the relative-entropies $H - H_M$, $H - H_{GR}$, & $H_M - H_{GR}$ are evaluated from the full, market and group-random mode to characterize and identify the different market events as anomalies, type-1, type-2, crashes and normal periods.

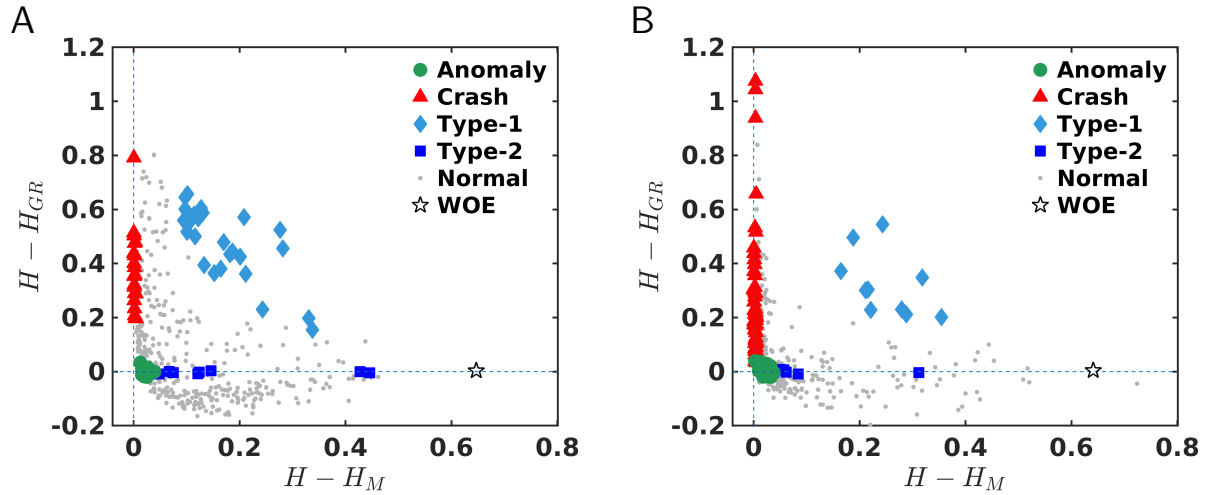


Figure 3.6: **Phase separation.** The 2D-plots of the phase space using relative-entropies $H - H_M$ and $H - H_{GR}$, for (A) S&P-500, and (B) Nikkei-225 markets. The event frames show “phase separation” – segregation of different market events: anomalies (green), type-1 (light blue), type-2 (blue), crashes (red) and normal (grey). The black stars correspond to the benchmarks WOE in both USA and JPN.

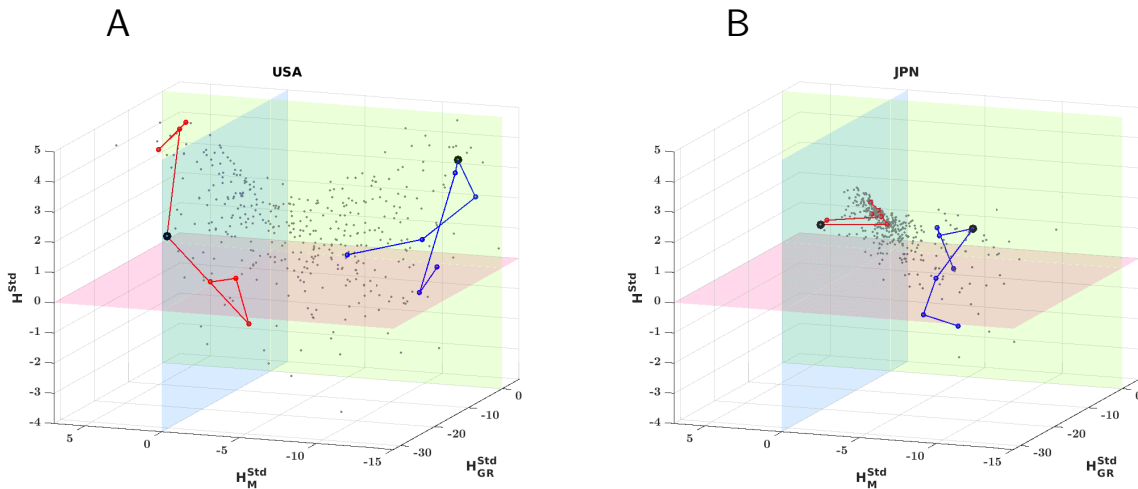


Figure 3.7: **Order-disorder transitions around critical events.** The 3D-plots of the standardized values of eigen-entropy corresponding to the full correlation matrix H^{Std} (along z-axis), eigen-entropy corresponding to the group-random mode H_{GR}^{Std} (x-axis), and eigen-entropy corresponding to the market mode H_M^{Std} (along y-axis), for (A) S&P-500, and (B) Nikkei-225 markets. The sequence of seven frames display the “order-disorder” transitions around the main events (in black filled circle) – in case of bubble bursts (Dot-com in USA and JPN; shown in blue) and crashes (Lehman Brothers in USA and Fukushima in JPN; shown in red).

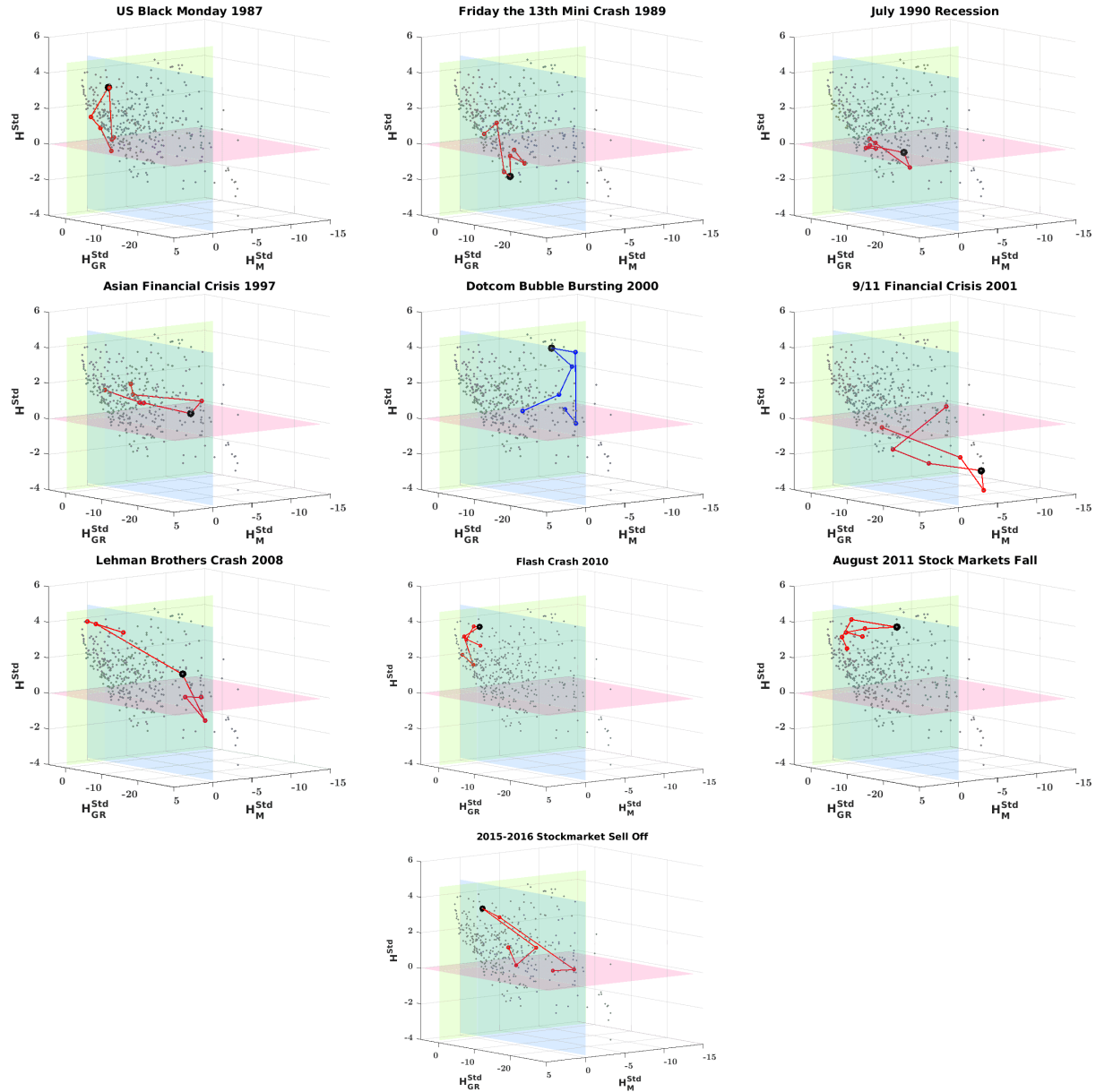


Figure 3.8: **Evolution around the important events in USA market.** Eigen-entropy H calculated from the correlation matrices: (full) C , market mode C_M and group-random mode C_{GR} for all the frames (epoch $M = 40$ days and shift $\Delta = 20$ days) over a period of 1985-2016 of USA (S&P-500). After standardizing the variables with moving average and moving standard deviation, each frame (grey dot) is embedded in a 3-D space with axes H^{std} , H_M^{std} and H_{GR}^{std} . Eleven important events with seven frames around those events (three before and three after the event) were taken from the history and shown in the plots. The critical events are connected with red lines and the Dot-com bubble burst is connected with blue. The frame containing the important event is marked with black circle for better visibility.

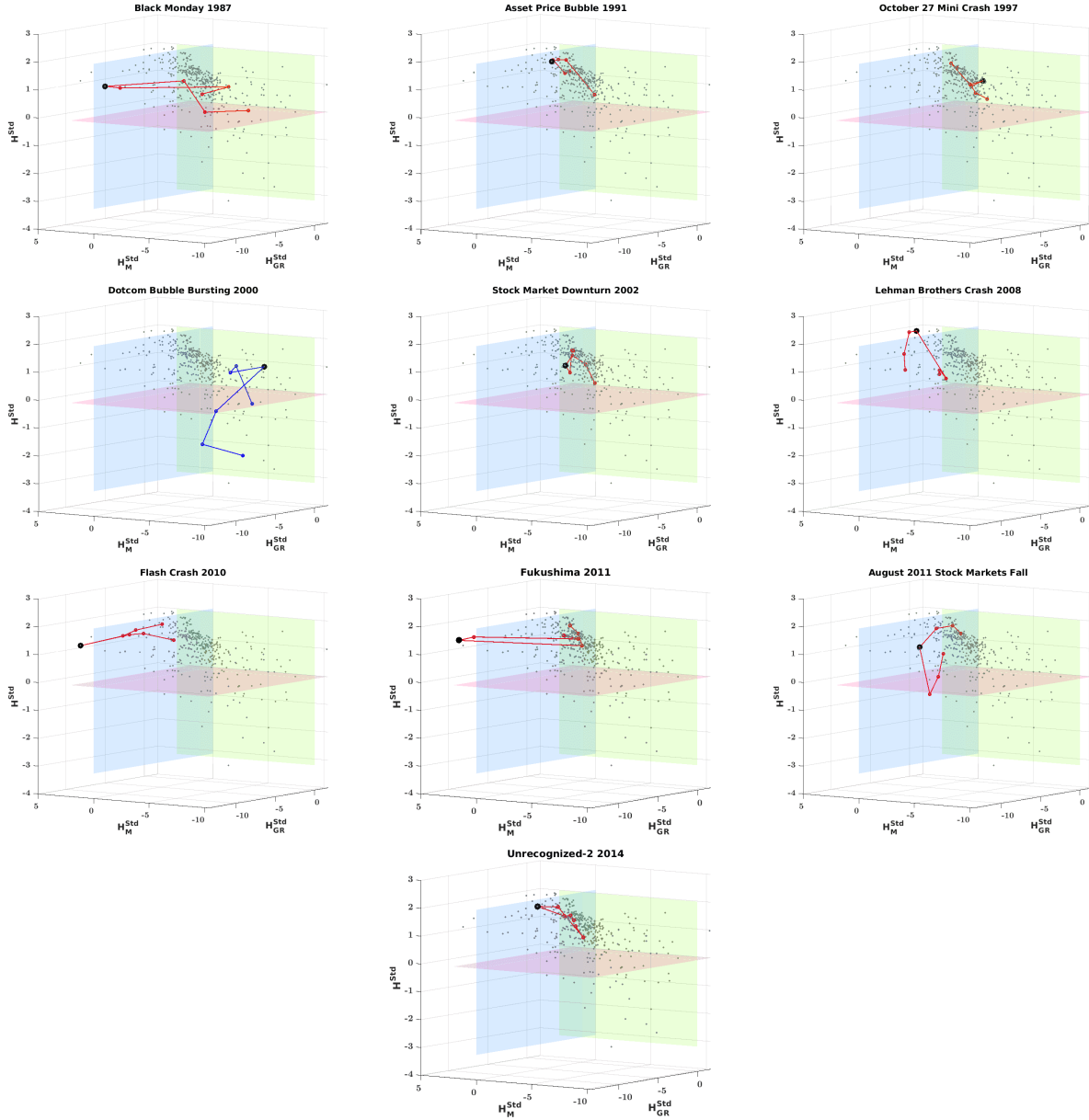


Figure 3.9: **Evolution around the important events in JPN market.** Eigen-entropy H calculated from the correlation matrices: (full) C , market mode C_M and group-random mode C_{GR} for all the frames (epoch $M = 40$ days and shift $\Delta = 20$ days) over a period of 1985-2016 of JPN (Nikkei-225). Three co-ordinates axes H^{std} , H_M^{std} and H_{GR}^{std} are the standardized variables, same as figure 3.8. Plots show thirteen important events from the history. The critical events are connected with red lines and the Dot-com bubble burst is connected with blue. The frame containing the important event is marked with black circle for better visibility.

time, we have been able to display such a phenomenon in the context of financial markets, which can be extremely significant for characterization and prediction of market events. The characterized events (corresponding to figure 3.6 (B) and (D)) are then indicated as vertical lines in the time-evolution plots in figure 3.6 (A) and (C). Interestingly, we find that many anomalies occur just around the major crashes, and intriguing patterns appear around the type-1 and type-2 events also.

3.4 Remarks

Effects of the variation of the epoch size M and shift Δ

The continuous monitoring of the market can be done by dividing the total time series data into smaller epochs of size M . The corresponding correlation matrices generated from these smaller epochs are used for calculating the eigen-entropy H . In figure 3.10, we investigate the effects of the variation of parameters, epoch size M and shift Δ .

We observe that either the increase in the epoch M or shift Δ makes the time series plot of H more smooth (less fluctuations), and vice versa. The choice of these parameters are thus arbitrary to some extent, depending on the research questions and time scale we are interested.

Effect of the variation in the powers of correlation matrices $|C|^n$

Instead of taking the square of individual elements of the correlation matrix C , to make all the elements non-negative, we can also use the even powers or the odd powers of absolute values to accomplish the same. The effect of the same is shown in the figure 3.11. As observed the values of eigen-entropy H differ with the variation of the power n of correlation matrices. This is due to the fact that with the increase in power, the dissimilarities in the elements of the correlation matrix are amplified which will then in turn changes the centrality of the matrix. For very high powers the transformed correlation matrices will act like an adjacency matrix with very high values (close to 1s) and very low values (close to 0s).

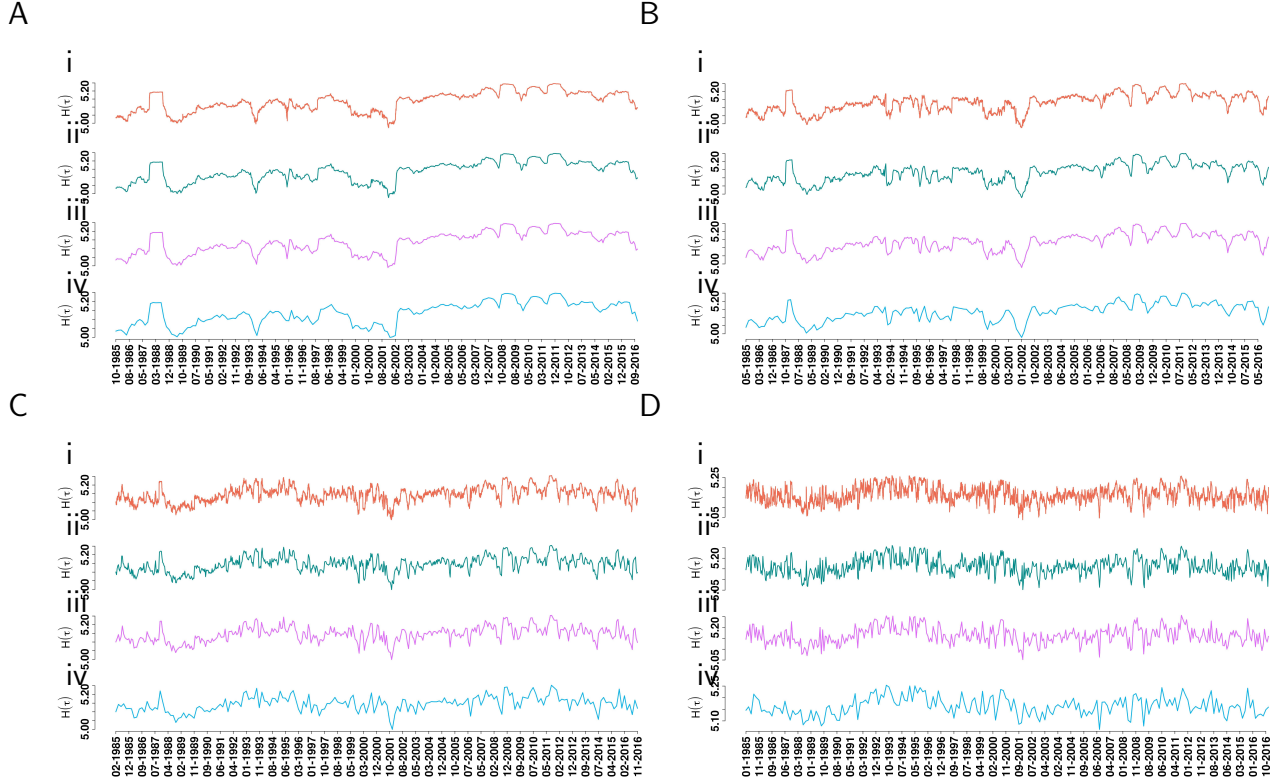


Figure 3.10: **Effects of epoch size M and shift Δ on the time series of eigen-entropy H .** The evolution of eigen-entropy H is calculated from correlation matrices corresponding to four different time epochs (A) $M = 200$, (B) $M = 100$, (C) $M = 40$, and (D) $M = 20$ days and each with four different shifts (i) $\Delta = 1$ day, (ii) $\Delta = 10$ days, (iii) $\Delta = 20$ days, and (iv) $\Delta = 40$ days over a period of 1985-2016. The fluctuations (local) of the eigen-entropy H are smoothened (smaller) for bigger shifts Δ .

It is also interesting to note that, depending on the problem, we can decide the range of correlations to focus on by adjusting the power of the elements of the correlation matrix.

3.4.1 Comparison with structural entropy

The structural entropy depends on the communities of the network and quantifies the ‘structural diversity’. One finds that the evolution of structural entropy may provide information about extreme events in the financial market, e.g., crises, bubbles, etc.

Following the prescription given in Almog et al. [60], the structural entropy may be calculated from the normalized sizes of the “communities” detected in the market after

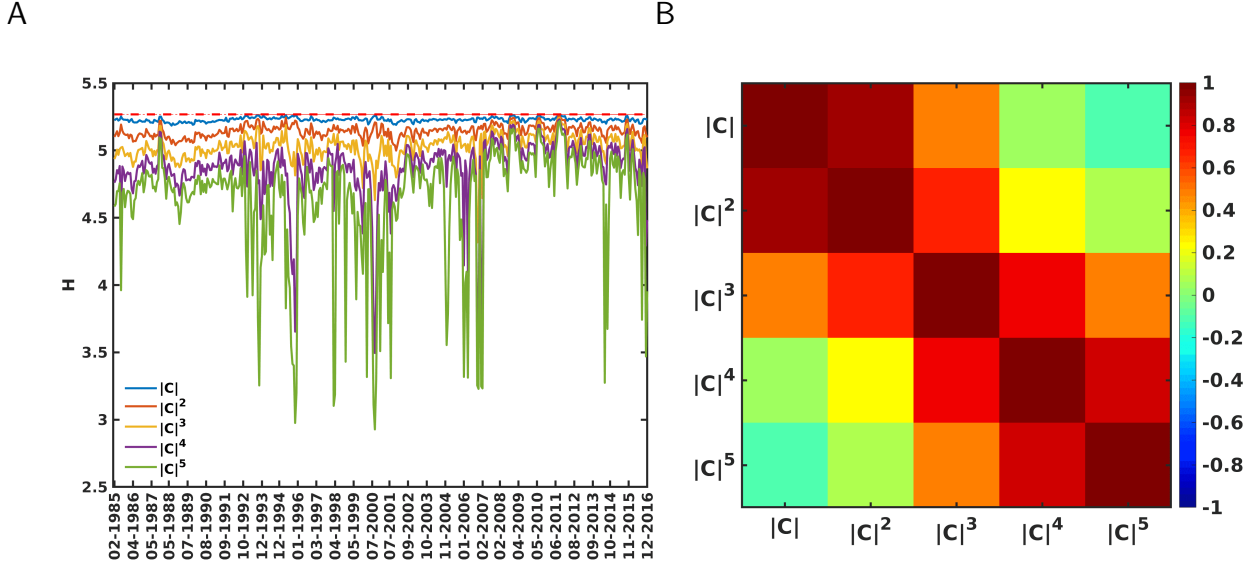


Figure 3.11: **Comparison of the variation of n for $|C|^n$.** The eigen-entropy H is calculated for different powers n of correlation matrix C by raising the elements of C to even powers or the absolute value of C to odd powers. (A) shows the time series of the eigen-entropies H of the correlation matrices of epoch $M = 40$ days and $\Delta = 20$ days for five different powers upto $n = 5$. The correlations among these five time series of eigen-entropy H is shown in (B).

applying a community detection algorithm [38].

In figure 3.12, we compare the eigen-entropy H measure with the structural entropy S using the community detection algorithm, where they obtain a modularity matrix directly from a correlation matrix, by applying random matrix theory tools and separating out just the group mode. The advantage of this method is that a modularity matrix can be supplied directly to a community detection algorithm, without using any arbitrary threshold.

When one compares the two entropy measures, it is evident that the structural entropy is very sensitive to the community detection algorithm (different algorithms yield different community structures). Even the community detection algorithm, which involves identifying the group mode from the correlation matrix is not easy because the boundary (determined by the eigenvalues of the correlation matrix) between the random mode and the group mode, is not distinct (and often arbitrary). In this way, our eigen-entropy measure has an advantage that it is uniquely determined and non-arbitrary (and also has less computational complexity). Also, during a market crash, the structural entropy S behaves differently from the eigen-entropy H , as the market starts behaving like a single (huge) super-community.

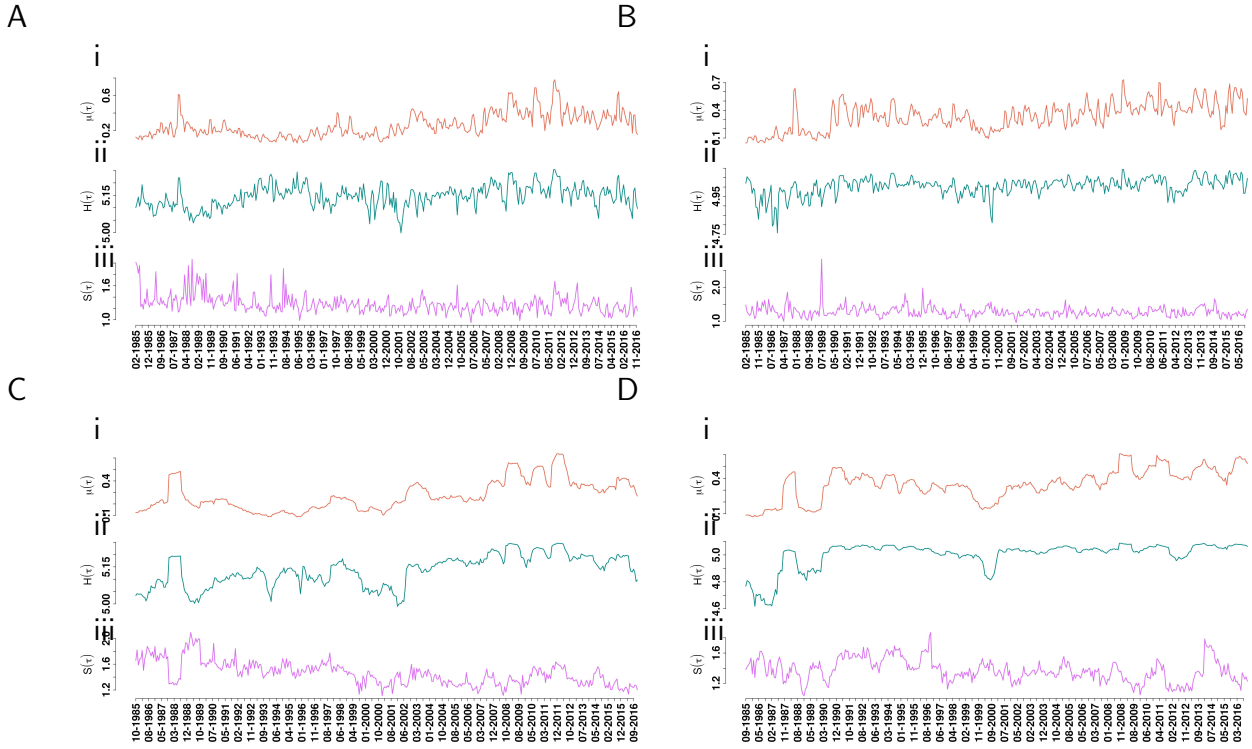


Figure 3.12: **Comparison of eigen-entropy H and structural entropy S .** Evolution of (i) average correlation μ , (ii) eigen-entropy H , and (iii) structural entropy S : (A) and (B) $M = 40$ days epoch and $\Delta = 20$ days shift for USA and JPN, respectively, and (C) and (D) $M = 200$ days epoch and $\Delta = 20$ days shift for USA and JPN, respectively.

So, during a crash S (measure of diversity) decreases in contrast to H (measure of disorder or randomness) that increases.

Chapter 4

Discussions and outlook

Here, we developed a general and robust methodology to extract information about the “disorder” (or randomness) in the market and its eigen modes, using the entropy measure – *eigen-entropy*, computed from the *eigen-centralities* (ranks) of different stocks in the correlation-network. We have used two different data sets of the stock markets USA S&P-500 and JPN Nikkei-225, spanning across a sufficiently long period of 32 years, to demonstrate its robustness.

We showed that the eigen-entropy is a simple yet robust prescription to quantify the disorder in a financial market. The methodology does not have any arbitrary thresholds. Further, the relative-entropy measures computed for these eigen modes enabled us to construct a “phase space”, where the different market events undergo “phase-separation” and display “order-disorder” transitions. The crashes occupy the region in the phase space, where $H - H_M \simeq 0$. During the crashes, the H and H_M almost touch the maximum disorder, $\ln N$ (corresponding to the random WOE). The events like “Dotcom bubble bursting” appear in the $H - H_{GR} \simeq 0$ axis. The events lying far away from the origin and axes are happening during bubble formation periods. The events lying close to the origin are like anomalies happening right before or right after major crashes. This type of phase-separation behavior in financial markets is being reported for the first time. Thus, we have here laid a clear prescription for characterizing the market events as anomalies, bubbles, crashes, etc. using the relative entropy measures. It was not well-understood how and when bubbles form and when they burst. Our proposed methodology may help us to understand the market events

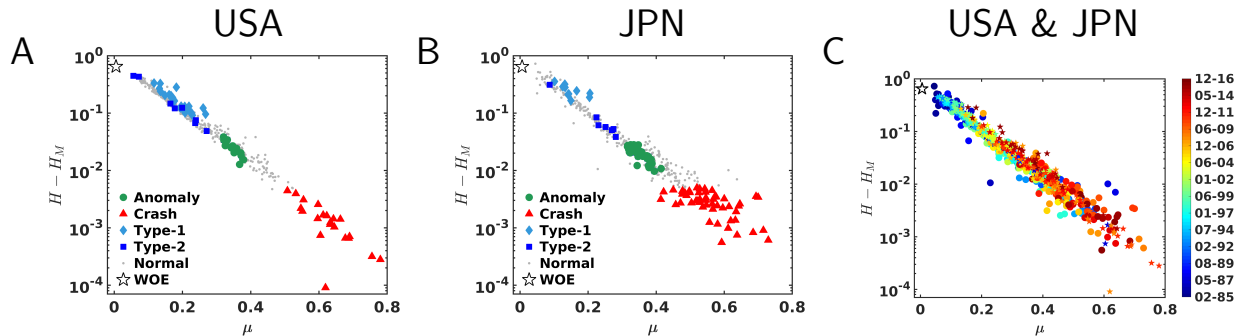


Figure 4.1: **Scaling behavior of the relative entropy** $H - H_M$. Plot of $H - H_M$ versus mean market correlation μ in linear-logarithmic scale for (A) S&P-500, and (B) Nikkei-225. The events are seen to lie on a straight line and the market event frames segregate into different portions: anomalies (green), type-1 (light blue), type-2 (blue), and crashes (red), interspersed by the normal events (grey). In (C), we see a data-collapse for both markets (USA in pink; JPN in purple) on a single curve, indicating a “universal scaling” behavior.

and their dynamics, as well as find the time-ordering and appearances of the bubbles (formations or bursts) and crashes, separated by normal periods. We have studied the evolution of events around major crashes and bubbles (from historical records in USA and JPN; see 4.2). Of course, further studies are required.

We reiterate that our eigen-entropy measure has an advantage that it is uniquely determined and non-arbitrary (and also has less computational complexity). When one compared (for details, see figure 3.12) our methodology with structural entropy [60], it is evident that the structural entropy is very sensitive to the community detection algorithm (different algorithms yield different community structures). Even the community detection algorithm [38], which involves identifying the group mode from the correlation matrix is not easy because the boundary (determined by the eigenvalues of the correlation matrix) between the random mode and the group mode, is not distinct (and often arbitrary).

Also, we have observed from the evolution of the entropy measures (H , H_M and H_{GR}) that the market behavior changes radically after 2000 (USA) and 1990 (JPN) corroborating to the findings of our earlier work [61], where we had found that the markets have “states” with different mean market correlation and market volatility.

Furthermore, the relative entropy $H - H_M$ displayed “universal scaling” behavior with respect to the mean market correlation μ ; a data-collapse was observed when plotted in a linear-logarithmic scale, which suggested that the fluctuations and co-movements in price

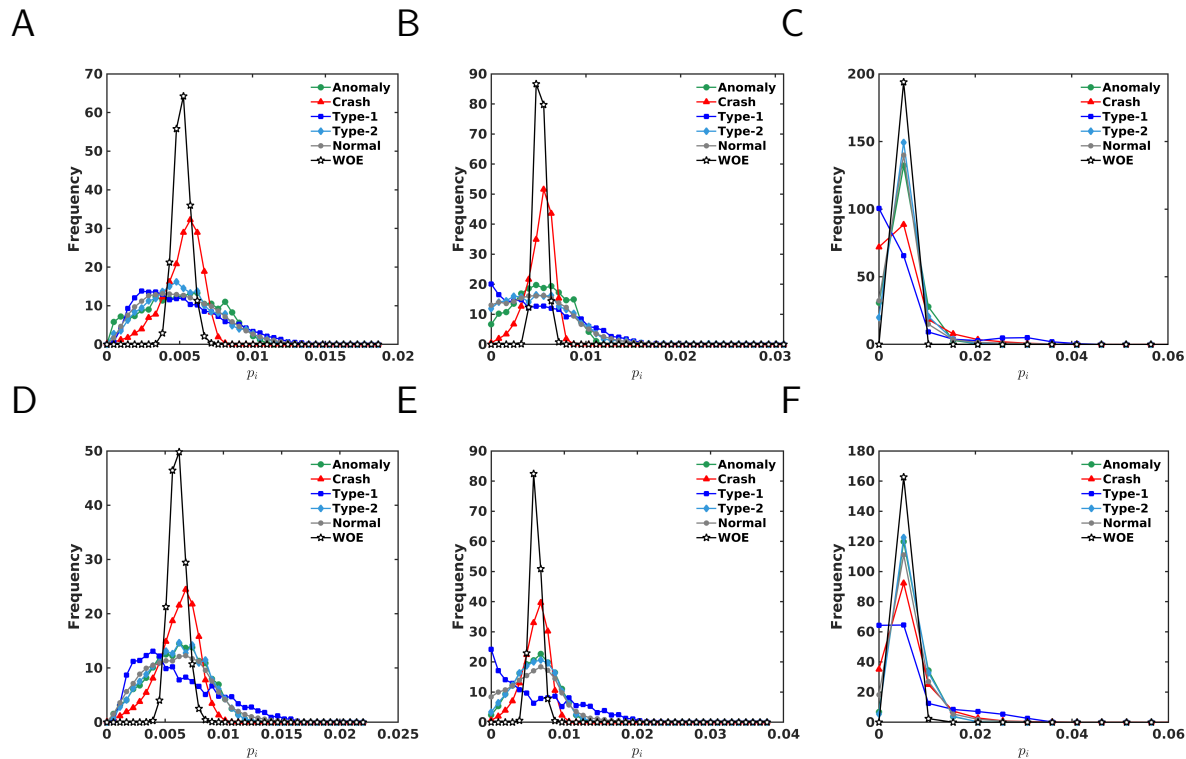


Figure 4.2: **Averaged distributions of the eigen-centralities, showing self-averaging properties.** Histograms of the eigen-centralities p_i for anomalies (green circles), type-1 (light blue diamonds), type-2 (blue squares), crash (red triangles) and normal (grey stars) and WOE (black squares), averaged over the respective ensembles for USA (top row) and for JPN (bottom row). Histograms are evaluated using (A and D) full correlation matrices C and decomposed correlation matrices of (B and E) market mode C_M , and (C and F) group and random mode C_{GR} .

returns for different financial assets and varying across countries are governed by the same statistical law. Also, the functional $-\ln(H - H_M)$ acted as a good indicator, as it could gauge market fragility (captured by the minimum risk of the market Markowitz portfolio) and the market fear (captured by the empirical volatility index). This can be important for managing risk and regulating the markets.

A “universal scaling” behavior is exhibited by the relative entropy $H - H_M$. The relative entropy $H - H_M$ versus mean market correlation μ is plotted in figure 4.1 (A) for S&P-500, (B) for Nikkei-225; the events are seen to lie on a straight line in a linear-logarithmic scale and the market event frames segregate into different portions: anomalies (green), type-1 (light blue), type-2 (blue), and crashes (red), interspersed by the normal events (grey). In figure 4.1 (C) the data for both markets collapse on a single curve, which indicates universal

Table 4.1: Values in cross-correlogram of the entropy measures with market parameters

	μ	H	H_M	H_{GR}	$H - H_M$	$-\ln(H - H_M)$	$H - H_{GR}$	VIX	ρ
μ	1	0.361	0.892	-0.392	-0.804	0.987	0.463	0.471	0.619
H	0.361	1	0.313	0.151	0.043	0.387	0.031	0.091	0.164
H_M	0.892	0.313	1	-0.283	-0.935	0.897	0.343	0.336	0.506
H_{GR}	-0.392	0.151	-0.283	1	0.354	-0.314	-0.983	-0.058	-0.41
$H - H_M$	-0.804	0.043	-0.935	0.354	1	-0.799	-0.35	-0.319	-0.471
$-\ln(H - H_M)$	0.987	0.387	0.897	-0.314	-0.799	1	0.389	0.505	0.59
$H - H_{GR}$	0.463	0.031	0.343	-0.983	-0.35	0.389	1	0.076	0.445
VIX	0.471	0.091	0.336	-0.058	-0.319	0.505	0.076	1	0.633
ρ	0.619	0.164	0.506	-0.41	-0.471	0.59	0.445	0.633	1

scaling behavior normally seen in many physical systems [72, 73]. This implies that the co-movements in price returns for different financial assets and varying across countries, are governed by the same statistical law.

Very interestingly, this functional $-\ln(H - H_M)$ also acts as a good gauge for the market fragility (minimum risk of the market portfolio) and the market fear (volatility index). figure 4.1 displays the cross-correlogram of the mean market correlation, functional $-\ln(H - H_M)$, the minimum risk (market fragility) and volatility index (market fear) and many other indicators. In addition, we would like to extent this methodology to other complex systems like brain, environment, etc., to investigate the existence of similar scaling laws and phase separations that will lead to a deeper and broader understanding of complex systems.

Appendix A

Table 4.2: List of major crashes and bubbles for USA and JPN markets and their characterization [74, 75, 76, 77, 78]. All the events are plotted in figures 3.8 and 3.9.

Important Stock Market Events			
Sl. No	Major crashes and bubbles	Period Date	Region Affected
1	Black Monday	19-10-1987	USA,JPN
2	Friday the 13th Mini Crash	13-10-1989	USA
3	Early 90s Recession	1990	USA
5	Mini Crash Due To Asian Financial Crisis	27-10-1997	USA
6	Lost Decade	2001-2010	JPN
7	9/11 Financial Crisis	11-09-2001	USA,JPN
8	Stock Market Downturn Of 2002	09-10-2002	JPN,USA
9	US Housing Bubble	2005-2007	USA
10	Lehman Brothers Crash	16-09-2008	USA,JPN
11	DJ Flash Crash	06-05-2010	USA,JPN
12	Tsunami/Fukushima	11-03-2011	JPN
13	August 2011 Stock Markets Fall	08-08-2011	USA,JPN
14	Chinese Black Monday and 2015-2016 Sell Off	24-08-2015	USA

Appendix B

Table 4.3: List of all stocks of USA market (S&P-500) considered for the analysis. The first column has the serial number, the second column has the abbreviation, the third column has the full name of the stock, and the fourth column specifies the sector as given in the S&P-500.

S.No.	Code	Company Name	Sector	Abbrv
1	CMCSA	Comcast Corp.	Consumer Discretionary	CD
2	DIS	The Walt Disney Company	Consumer Discretionary	CD
3	F	Ford Motor	Consumer Discretionary	CD
4	GPC	Genuine Parts	Consumer Discretionary	CD
5	GPS	Gap Inc.	Consumer Discretionary	CD
6	GT	Goodyear Tire & Rubber	Consumer Discretionary	CD
7	HAS	Hasbro Inc.	Consumer Discretionary	CD
8	HD	Home Depot	Consumer Discretionary	CD
9	HRB	Block H&R	Consumer Discretionary	CD
10	IPG	Interpublic Group	Consumer Discretionary	CD
11	JCP	J. C. Penney Company, Inc.	Consumer Discretionary	CD
12	JWN	Nordstrom	Consumer Discretionary	CD
13	LEG	Leggett & Platt	Consumer Discretionary	CD
14	LEN	Lennar Corp.	Consumer Discretionary	CD
15	LOW	Lowe's Cos.	Consumer Discretionary	CD
16	MAT	Mattel Inc.	Consumer Discretionary	CD
17	MCD	McDonald's Corp.	Consumer Discretionary	CD
18	NKE	Nike	Consumer Discretionary	CD
19	SHW	Sherwin-Williams	Consumer Discretionary	CD
20	TGT	Target Corp.	Consumer Discretionary	CD
21	VFC	V.F. Corp.	Consumer Discretionary	CD
22	WHR	Whirlpool Corp.	Consumer Discretionary	CD
23	ADM	Archer-Daniels-Midland Co	Consumer Staples	CS
24	AVP	Avon Products, Inc.	Consumer Staples	CS

25	CAG	Conagra Brands	Consumer Staples	CS
26	CL	Colgate-Palmolive	Consumer Staples	CS
27	CPB	Campbell Soup	Consumer Staples	CS
28	CVS	CVS Health	Consumer Staples	CS
29	GIS	General Mills	Consumer Staples	CS
30	HRL	Hormel Foods Corp.	Consumer Staples	CS
31	HSY	The Hershey Company	Consumer Staples	CS
32	K	Kellogg Co.	Consumer Staples	CS
33	KMB	Kimberly-Clark	Consumer Staples	CS
34	KO	Coca-Cola Company (The)	Consumer Staples	CS
35	KR	Kroger Co.	Consumer Staples	CS
36	MKC	McCormick & Co.	Consumer Staples	CS
37	MO	Altria Group Inc	Consumer Staples	CS
38	SYZ	Sysco Corp.	Consumer Staples	CS
39	TAP	Molson Coors Brewing Company	Consumer Staples	CS
40	TSN	Tyson Foods	Consumer Staples	CS
41	WMT	Wal-Mart Stores	Consumer Staples	CS
42	APA	Apache Corporation	Energy	EG
43	COP	ConocoPhillips	Energy	EG
44	CVX	Chevron Corp.	Energy	EG
45	ESV	Enscopl	Energy	EG
46	HAL	Halliburton Co.	Energy	EG
47	HES	Hess Corporation	Energy	EG
48	HP	Helmerich & Payne	Energy	EG
49	MRO	Marathon Oil Corp.	Energy	EG
50	MUR	Murphy Oil Corporation	Energy	EG
51	NBL	Noble Energy Inc	Energy	EG
52	NBR	Nabors Industries Ltd.	Energy	EG
53	SLB	Schlumberger Ltd.	Energy	EG
54	TSO	Tesoro Corp	Energy	EG
55	VLO	Valero Energy	Energy	EG
56	WMB	Williams Cos.	Energy	EG
57	XOM	Exxon Mobil Corp.	Energy	EG
58	AFL	AFLAC Inc	Financials	FN
59	AIG	American International Group, Inc.	Financials	FN
60	AON	Aon plc	Financials	FN
61	AXP	American Express Co	Financials	FN
62	BAC	Bank of America Corp	Financials	FN

63	BBT	BB&T Corporation	Financials	FN
64	BEN	Franklin Resources	Financials	FN
65	BK	The Bank of New York Mellon Corp.	Financials	FN
66	C	Citigroup Inc.	Financials	FN
67	CB	Chubb Limited	Financials	FN
68	CINF	Cincinnati Financial	Financials	FN
69	CMA	Comerica Inc.	Financials	FN
70	EFX	Equifax Inc.	Financials	FN
71	FHN	First Horizon National Corporation	Financials	FN
72	HBAN	Huntington Bancshares	Financials	FN
73	HCN	Welltower Inc.	Financials	FN
74	HST	Host Hotels & Resorts, Inc.	Financials	FN
75	JPM	JPMorgan Chase & Co.	Financials	FN
76	L	Loews Corp.	Financials	FN
77	LM	Legg Mason, Inc.	Financials	FN
78	LNC	Lincoln National	Financials	FN
79	LUK	Leucadia National Corp.	Financials	FN
80	MMC	Marsh & McLennan	Financials	FN
81	MTB	M&T Bank Corp.	Financials	FN
82	PSA	Public Storage	Financials	FN
83	SLM	SLM Corporation	Financials	FN
84	TMK	Torchmark Corp.	Financials	FN
85	TRV	The Travelers Companies Inc.	Financials	FN
86	USB	U.S. Bancorp	Financials	FN
87	VNO	Vornado Realty Trust	Financials	FN
88	WFC	Wells Fargo	Financials	FN
89	WY	Weyerhaeuser Corp.	Financials	FN
90	ZION	Zions Bancorp	Financials	FN
91	ABT	Abbott Laboratories	Health Care	HC
92	AET	Aetna Inc	Health Care	HC
93	AMGN	Amgen Inc	Health Care	HC
94	BAX	Baxter International Inc.	Health Care	HC
95	BCR	Bard (C.R.) Inc.	Health Care	HC
96	BDX	Becton Dickinson	Health Care	HC
97	BMJ	Bristol-Myers Squibb	Health Care	HC

98	CAH	Cardinal Health Inc.	Health Care	HC
99	CI	CIGNA Corp.	Health Care	HC
100	HUM	Humana Inc.	Health Care	HC
101	JNJ	Johnson & Johnson	Health Care	HC
102	LLY	Lilly (Eli) & Co.	Health Care	HC
103	MDT	Medtronic plc	Health Care	HC
104	MRK	Merck & Co.	Health Care	HC
105	MYL	Mylan N.V.	Health Care	HC
106	SYK	Stryker Corp.	Health Care	HC
107	THC	Tenet Healthcare Corp	Health Care	HC
108	TMO	Thermo Fisher Scientific	Health Care	HC
109	UNH	United Health Group Inc.	Health Care	HC
110	VAR	Varian Medical Systems	Health Care	HC
111	AVY	Avery Dennison Corp	Industrials	ID
112	BA	Boeing Company	Industrials	ID
113	CAT	Caterpillar Inc.	Industrials	ID
114	CMI	Cummins Inc.	Industrials	ID
115	CSX	CSX Corp.	Industrials	ID
116	CTAS	Cintas Corporation	Industrials	ID
117	DE	Deere & Co.	Industrials	ID
118	DHR	Danaher Corp.	Industrials	ID
119	DNB	The Dun & Bradstreet Corpora- tion	Industrials	ID
120	DOV	Dover Corp.	Industrials	ID
121	EMR	Emerson Electric Company	Industrials	ID
122	ETN	Eaton Corporation	Industrials	ID
123	EXPD	Expeditors International	Industrials	ID
124	FDX	FedEx Corporation	Industrials	ID
125	FLS	Flowserve Corporation	Industrials	ID
126	GD	General Dynamics	Industrials	ID
127	GE	General Electric	Industrials	ID
128	GLW	Corning Inc.	Industrials	ID
129	GWW	Grainger (W.W.) Inc.	Industrials	ID
130	HON	Honeywell Int'l Inc.	Industrials	ID
131	IR	Ingersoll-Rand PLC	Industrials	ID
132	ITW	Illinois Tool Works	Industrials	ID
133	JEC	Jacobs Engineering Group	Industrials	ID
134	LMT	Lockheed Martin Corp.	Industrials	ID
135	LUV	Southwest Airlines	Industrials	ID
136	MAS	Masco Corp.	Industrials	ID
137	MMM	3M Company	Industrials	ID
138	ROK	Rockwell Automation Inc.	Industrials	ID

139	RTN	Raytheon Co.	Industrials	ID
140	TXT	Textron Inc.	Industrials	ID
141	UNP	Union Pacific	Industrials	ID
142	UTX	United Technologies	Industrials	ID
143	AAPL	Apple Inc.	Information Technology	IT
144	ADI	Analog Devices, Inc.	Information Technology	IT
145	ADP	Automatic Data Processing	Information Technology	IT
146	AMAT	Applied Materials Inc	Information Technology	IT
147	AMD	Advanced Micro Devices Inc	Information Technology	IT
148	CA	CA, Inc.	Information Technology	IT
149	HPQ	HP Inc.	Information Technology	IT
150	HRS	Harris Corporation	Information Technology	IT
151	IBM	International Business Machines	Information Technology	IT
152	INTC	Intel Corp.	Information Technology	IT
153	KLAC	KLA-Tencor Corp.	Information Technology	IT
154	LRCX	Lam Research	Information Technology	IT
155	MSI	Motorola Solutions Inc.	Information Technology	IT
156	MU	Micron Technology	Information Technology	IT
157	TSS	Total System Services, Inc.	Information Technology	IT
158	TXN	Texas Instruments	Information Technology	IT
159	WDC	Western Digital	Information Technology	IT
160	XRX	Xerox Corp.	Information Technology	IT
161	AA	Alcoa Corporation	Materials	MT
162	APD	Air Products & Chemicals Inc	Materials	MT
163	BLL	Ball Corp	Materials	MT
164	BMS	Bemis Company, Inc.	Materials	MT
165	CLF	Cleveland-Cliffs Inc.	Materials	MT
166	DD	DuPont	Materials	MT
167	ECL	Ecolab Inc.	Materials	MT
168	FMC	FMC Corporation	Materials	MT
169	IFF	Intl Flavors & Fragrances	Materials	MT
170	IP	International Paper	Materials	MT
171	NEM	Newmont Mining Corporation	Materials	MT
172	PPG	PPG Industries	Materials	MT
173	VMC	Vulcan Materials	Materials	MT
174	CTL	CenturyLink Inc	Telecommunication Services	TC
175	FTR	Frontier Communications Corporation	Telecommunication Services	TC
176	S	Sprint Nextel Corp.	Telecommunication Services	TC

177	T	AT&T Inc	Telecommunication Services	TC
178	VZ	Verizon Communications	Telecommunication Services	TC
179	AEP	American Electric Power	Utilities	UT
180	CMS	CMS Energy	Utilities	UT
181	CNP	CenterPoint Energy	Utilities	UT
182	D	Dominion Energy	Utilities	UT
183	DTE	DTE Energy Co.	Utilities	UT
184	ED	Consolidated Edison	Utilities	UT
185	EIX	Edison Int'l	Utilities	UT
186	EQT	EQT Corporation	Utilities	UT
187	ETR	Entergy Corp.	Utilities	UT
188	EXC	Exelon Corp.	Utilities	UT
189	NEE	NextEra Energy	Utilities	UT
190	NI	NiSource Inc.	Utilities	UT
191	PNW	Pinnacle West Capital	Utilities	UT
192	SO	Southern Co.	Utilities	UT
193	WEC	Wec Energy Group Inc	Utilities	UT
194	XEL	Xcel Energy Inc	Utilities	UT

Table 4.4: List of all stocks of Japan market (Nikkei-225) considered for the analysis. The first column has the serial number, the second column has the abbreviation, the third column has the full name of the stock, and the fourth column specifies the sector as given in the Nikkei-225.

S.No.	Code	Company Name	Sector	Abbrv
1	S-8801	mitsui fudosan co., ltd.	Capital Goods	CG
2	S-8802	MITSUBISHI ESTATE CO., LTD.	Capital Goods	CG
3	S-8804	TOKYO TATEMONO CO., LTD.	Capital Goods	CG
4	S-8830	SUMITOMO REALTY & DEVELOPMENT CO., LTD.	Capital Goods	CG
5	S-7003	MITSUI ENG. & SHIPBUILD. CO., LTD.	Capital Goods	CG
6	S-7012	KAWASAKI HEAVY IND., LTD.	Capital Goods	CG
7	S-9202	ANA HOLDINGS INC.	Capital Goods	CG
8	S-1801	TAISEI CORP.	Capital Goods	CG
9	S-1802	OBAYASHI CORP.	Capital Goods	CG
10	S-1803	SHIMIZU CORP.	Capital Goods	CG
11	S-1808	HASEKO CORP.	Capital Goods	CG
12	S-1812	KAJIMA CORP.	Capital Goods	CG
13	S-1925	DAIWA HOUSE IND. CO., LTD.	Capital Goods	CG
14	S-1928	SEKISUI HOUSE, LTD.	Capital Goods	CG
15	S-1963	JGC CORP.	Capital Goods	CG
16	S-5631	THE JAPAN STEEL WORKS, LTD.	Capital Goods	CG
17	S-6103	OKUMA CORP.	Capital Goods	CG
18	S-6113	AMADA HOLDINGS CO., LTD.	Capital Goods	CG
19	S-6301	KOMATSU LTD.	Capital Goods	CG
20	S-6302	SUMITOMO HEAVY IND., LTD.	Capital Goods	CG
21	S-6305	HITACHI CONST. MACH. CO., LTD.	Capital Goods	CG
22	S-6326	KUBOTA CORP.	Capital Goods	CG
23	S-6361	EBARA CORP.	Capital Goods	CG
24	S-6366	CHIYODA CORP.	Capital Goods	CG
25	S-6367	DAIKIN INDUSTRIES, LTD.	Capital Goods	CG
26	S-6471	NSK LTD.	Capital Goods	CG
27	S-6472	NTN CORP.	Capital Goods	CG
28	S-6473	JTEKT CORP.	Capital Goods	CG
29	S-7004	HITACHI ZOSEN CORP.	Capital Goods	CG
30	S-7011	MITSUBISHI HEAVY IND., LTD.	Capital Goods	CG
31	S-7013	IHI CORP.	Capital Goods	CG
32	S-7911	TOPPAN PRINTING CO., LTD.	Capital Goods	CG
33	S-7912	DAI NIPPON PRINTING CO., LTD.	Capital Goods	CG
34	S-7951	YAMAHA CORP.	Capital Goods	CG
35	S-1332	NIPPON SUISAN KAISHA, LTD.	Consumer Goods	CN
36	S-2002	NISSHIN SEIFUN GROUP INC.	Consumer Goods	CN
37	S-2282	NH FOODS LTD.	Consumer Goods	CN
38	S-2501	SAPPORO HOLDINGS LTD.	Consumer Goods	CN
39	S-2502	ASAHI GROUP HOLDINGS, LTD.	Consumer Goods	CN

40	S-2503	KIRIN HOLDINGS CO., LTD.	Consumer Goods	CN
41	S-2531	TAKARA HOLDINGS INC.	Consumer Goods	CN
42	S-2801	KIKKOMAN CORP.	Consumer Goods	CN
43	S-2802	AJINOMOTO CO., INC.	Consumer Goods	CN
44	S-2871	NICHIREI CORP.	Consumer Goods	CN
45	S-8233	TAKASHIMAYA CO., LTD.	Consumer Goods	CN
46	S-8252	MARUI GROUP CO., LTD.	Consumer Goods	CN
47	S-8267	AEON CO., LTD.	Consumer Goods	CN
48	S-9602	TOHO CO., LTD	Consumer Goods	CN
49	S-9681	TOKYO DOME CORP.	Consumer Goods	CN
50	S-9735	SECOM CO., LTD.	Consumer Goods	CN
51	S-8331	THE CHIBA BANK, LTD.	Financials	FN
52	S-8355	THE SHIZUOKA BANK, LTD.	Financials	FN
53	S-8253	CREDIT SAISON CO., LTD.	Financials	FN
54	S-8601	DAIWA SECURITIES GROUP INC.	Financials	FN
55	S-8604	NOMURA HOLDINGS, INC.	Financials	FN
56	S-3405	KURARAY CO., LTD.	Materials	MT
57	S-3407	ASAHI KASEI CORP.	Materials	MT
58	S-4004	SHOWA DENKO K.K.	Materials	MT
59	S-4005	SUMITOMO CHEMICAL CO., LTD.	Materials	MT
60	S-4021	NISSAN CHEMICAL IND., LTD.	Materials	MT
61	S-4042	TOSOH CORP.	Materials	MT
62	S-4043	TOKUYAMA CORP.	Materials	MT
63	S-4061	DENKA CO., LTD.	Materials	MT
64	S-4063	SHIN-ETSU CHEMICAL CO., LTD.	Materials	MT
65	S-4183	mitsui chemicals, inc.	Materials	MT
66	S-4208	UBE INDUSTRIES, LTD.	Materials	MT
67	S-4272	NIPPON KAYAKU CO., LTD.	Materials	MT
68	S-4452	KAO CORP.	Materials	MT
69	S-4901	FUJIFILM HOLDINGS CORP.	Materials	MT
70	S-4911	SHISEIDO CO., LTD.	Materials	MT
71	S-6988	NITTO DENKO CORP.	Materials	MT
72	S-5002	SHOWA SHELL SEKIYU K.K.	Materials	MT
73	S-5201	ASAHI GLASS CO., LTD.	Materials	MT
74	S-5202	NIPPON SHEET GLASS CO., LTD.	Materials	MT
75	S-5214	NIPPON ELECTRIC GLASS CO., LTD.	Materials	MT

76	S-5232	SUMITOMO OSAKA CEMENT CO., LTD.	Materials	MT
77	S-5233	TAIHEIYO CEMENT CORP.	Materials	MT
78	S-5301	TOKAI CARBON CO., LTD.	Materials	MT
79	S-5332	TOTO LTD.	Materials	MT
80	S-5333	NGK INSULATORS, LTD.	Materials	MT
81	S-5706	mitsui Mining & Smelting Co.	Materials	MT
82	S-5707	TOHO ZINC CO., LTD.	Materials	MT
83	S-5711	MITSUBISHI MATERIALS CORP.	Materials	MT
84	S-5713	SUMITOMO METAL MINING CO., LTD.	Materials	MT
85	S-5714	DOWA HOLDINGS CO., LTD.	Materials	MT
86	S-5715	FURUKAWA CO., LTD.	Materials	MT
87	S-5801	FURUKAWA ELECTRIC CO., LTD.	Materials	MT
88	S-5802	SUMITOMO ELECTRIC IND., LTD.	Materials	MT
89	S-5803	FUJIKURA LTD.	Materials	MT
90	S-5901	TOYO SEIKAN GROUP HOLDINGS, LTD.	Materials	MT
91	S-3865	HOKUETSU KISHU PAPER CO., LTD.	Materials	MT
92	S-3861	OJI HOLDINGS CORP.	Materials	MT
93	S-5101	THE YOKOHAMA RUBBER CO., LTD.	Materials	MT
94	S-5108	BRIDGESTONE CORP.	Materials	MT
95	S-5401	NIPPON STEEL & SUMITOMO METAL CORP.	Materials	MT
96	S-5406	KOBE STEEL, LTD.	Materials	MT
97	S-5541	PACIFIC METALS CO., LTD.	Materials	MT
98	S-3101	TOYOBO CO., LTD.	Materials	MT
99	S-3103	UNITIKA, LTD.	Materials	MT
100	S-3401	TEIJIN LTD.	Materials	MT
101	S-3402	TORAY INDUSTRIES, INC.	Materials	MT
102	S-8001	ITOCHU CORP.	Materials	MT
103	S-8002	MARUBENI CORP.	Materials	MT
104	S-8015	TOYOTA TSUSHO CORP.	Materials	MT
105	S-8031	MITSUI & CO., LTD.	Materials	MT
106	S-8053	SUMITOMO CORP.	Materials	MT
107	S-8058	MITSUBISHI CORP.	Materials	MT

108	S-4151	KYOWA HAKKO KIRIN CO., LTD.	Pharmaceuticals	PH
109	S-4503	ASTELLAS PHARMA INC.	Pharmaceuticals	PH
110	S-4506	SUMITOMO DAINIPPON PHARMA CO., LTD.	Pharmaceuticals	PH
111	S-4507	SHIONOGI & CO., LTD.	Pharmaceuticals	PH
112	S-4519	CHUGAI PHARMACEUTICAL CO., LTD.	Pharmaceuticals	PH
113	S-4523	EISAI CO., LTD.	Pharmaceuticals	PH
114	S-7201	NISSAN MOTOR CO., LTD.	Information Technology	IT
115	S-7202	ISUZU MOTORS LTD.	Information Technology	IT
116	S-7205	HINO MOTORS, LTD.	Information Technology	IT
117	S-7261	MAZDA MOTOR CORP.	Information Technology	IT
118	S-7267	HONDA MOTOR CO., LTD.	Information Technology	IT
119	S-7270	SUBARU CORP.	Information Technology	IT
120	S-7272	YAMAHA MOTOR CO., LTD.	Information Technology	IT
121	S-3105	NISSHINBO HOLDINGS INC.	Information Technology	IT
122	S-6479	MINEBEA MITSUMI INC.	Information Technology	IT
123	S-6501	HITACHI, LTD.	Information Technology	IT
124	S-6502	TOSHIBA CORP.	Information Technology	IT
125	S-6503	mitsubishi electric corp.	Information Technology	IT
126	S-6504	FUJI ELECTRIC CO., LTD.	Information Technology	IT
127	S-6506	YASKAWA ELECTRIC CORP.	Information Technology	IT
128	S-6508	MEIDENSHA CORP.	Information Technology	IT
129	S-6701	NEC CORP.	Information Technology	IT
130	S-6702	FUJITSU LTD.	Information Technology	IT
131	S-6703	OKI ELECTRIC IND. CO., LTD.	Information Technology	IT
132	S-6752	PANASONIC CORP.	Information Technology	IT
133	S-6758	SONY CORP.	Information Technology	IT
134	S-6762	TDK CORP.	Information Technology	IT
135	S-6770	ALPS ELECTRIC CO., LTD.	Information Technology	IT
136	S-6773	PIONEER CORP.	Information Technology	IT
137	S-6841	YOKOGAWA ELECTRIC CORP.	Information Technology	IT
138	S-6902	DENSO CORP.	Information Technology	IT
139	S-6952	CASIO COMPUTER CO., LTD.	Information Technology	IT
140	S-6954	FANUC CORP.	Information Technology	IT
141	S-6971	KYOCERA CORP.	Information Technology	IT
142	S-6976	TAIYO YUDEN CO., LTD.	Information Technology	IT
143	S-7752	RICOH CO., LTD.	Information Technology	IT
144	S-8035	TOKYO ELECTRON LTD.	Information Technology	IT
145	S-4543	TERUMO CORP.	Information Technology	IT
146	S-4902	KONICA MINOLTA, INC.	Information Technology	IT
147	S-7731	NIKON CORP.	Information Technology	IT
148	S-7733	OLYMPUS CORP.	Information Technology	IT
149	S-7762	CITIZEN WATCH CO., LTD.	Information Technology	IT
150	S-9501	TOKYO ELECTRIC POWER COMPANY HOLDINGS, I	Transportation & Utilities	TU

151	S-9502	CHUBU ELECTRIC POWER CO., INC.	Transportation & Utilities	TU
152	S-9503	THE KANSAI ELECTRIC POWER CO., INC.	Transportation & Utilities	TU
153	S-9531	TOKYO GAS CO., LTD.	Transportation & Utilities	TU
154	S-9532	OSAKA GAS CO., LTD.	Transportation & Utilities	TU
155	S-9062	NIPPON EXPRESS CO., LTD.	Transportation & Utilities	TU
156	S-9064	YAMATO HOLDINGS CO., LTD.	Transportation & Utilities	TU
157	S-9101	NIPPON YUSEN K.K.	Transportation & Utilities	TU
158	S-9104	NETSUKAISEN, LTD.	Transportation & Utilities	TU
159	S-9107	KAWASAKI KISEN KAISHA, LTD.	Transportation & Utilities	TU
160	S-9001	TOBU RAILWAY CO., LTD.	Transportation & Utilities	TU
161	S-9005	TOKYU CORP.	Transportation & Utilities	TU
162	S-9007	ODAKYU ELECTRIC RAILWAY CO., LTD.	Transportation & Utilities	TU
163	S-9008	KEIO CORP.	Transportation & Utilities	TU
164	S-9009	KEISEI ELECTRIC RAILWAY CO., LTD.	Transportation & Utilities	TU
165	S-9301	NETSUKAISEN, LTD.	Transportation & Utilities	TU

Bibliography

- [1] Venkateswararao Vemuri. *Modeling of Complex Systems: An Introduction*. Academic Press, New York, 1978.
- [2] Murray Gell-Mann. What is complexity? *Complexity*, 1:16–19, 1995.
- [3] Rosario Nunzio Mantegna and Harry Eugene Stanley. *An introduction to econophysics: correlations and complexity in finance*. Cambridge University Press, Cambridge, 2007.
- [4] Jean-Philippe Bouchaud and Marc Potters. *Theory of Financial Risk and Derivative Pricing: from Statistical Physics to Risk Management*. Cambridge University Press, 2003.
- [5] Sitabhra Sinha, Arnab Chatterjee, Anirban Chakraborti, and Bikas K Chakrabarti. *Econophysics: an introduction*. John Wiley & Sons, 2010.
- [6] Anirban Chakraborti, Ioane Muni Toke, Marco Patriarca, and Frédéric Abergel. Econophysics review: I. empirical facts. *Quantitative Finance*, 11(7):991–1012, 2011.
- [7] Anirban Chakraborti, Damien Challet, Arnab Chatterjee, Matteo Marsili, Yi-Cheng Zhang, and Bikas K Chakrabarti. Statistical mechanics of competitive resource allocation using agent-based models. *Physics Reports*, 552:1–25, 2015.
- [8] Mark Buchanan. *Ubiquity: Why Catastrophes Happen*. Three Rivers Press, New York, 2000.
- [9] Didier Sornette. *Why Stock Markets Crash: Critical Events in Complex Financial Systems*. Princeton University Press, 2004.
- [10] H Eugene Stanley. *Phase transitions and critical phenomena*. Clarendon Press, Oxford, 1971.
- [11] James Sethna. *Statistical mechanics: entropy, order parameters, and complexity*, volume 14. Oxford University Press, 2006.
- [12] Andrew Ross Sorkin. *Too Big to Fail: The Inside Story of How Wall Street and Washington Fought to Save the Financial System—and Themselves*. Viking, New York, 2009.

- [13] Daron Acemoglu, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. Systemic risk and stability in financial networks. *American Economic Review*, 105(2):564–608, 2015.
- [14] Kiran Sharma, Balagopal Gopalakrishnan, Anindya S Chakrabarti, and Anirban Chakraborti. Financial fluctuations anchored to economic fundamentals: A mesoscopic network approach. *Scientific Reports*, 7(1):8055, 2017.
- [15] Kiran Sharma, Anindya S Chakrabarti, and Anirban Chakraborti. Multi-layered network structure: Relationship between financial and macroeconomic dynamics. In *New Perspectives and Challenges in Econophysics and Sociophysics*, pages 117–131. Springer, 2019.
- [16] Yahoo finance database. <https://finance.yahoo.co.jp/>, 2017. Accessed on 7th July, 2017, using the R open source programming language and software environment for statistical computing and graphics.
- [17] Martin Sewell. Characterization of financial time series. *Rn*, 11(01):01, 2011.
- [18] Rama Cont. Empirical properties of asset returns: stylized facts and statistical issues. 2001.
- [19] Eugene F Fama. Mandelbrot and the stable paretian hypothesis. *The journal of business*, 36(4):420–429, 1963.
- [20] Benoit B Mandelbrot. The variation of certain speculative prices. In *Fractals and scaling in finance*, pages 371–418. Springer, 1997.
- [21] Sarat Chandra Nayak, Bijan Bihari Misra, and Himansu Sekhar Behera. Efficient financial time series prediction with evolutionary virtual data position exploration. *Neural Computing and Applications*, 31(2):1053–1074, 2019.
- [22] Stephen J Taylor. *Modelling financial time series*. world scientific, 2008.
- [23] Francis Galton. Regression towards mediocrity in hereditary stature. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 15:246–263, 1886.
- [24] Stephen M Stigler. Francis galton’s account of the invention of correlation. *Statistical Science*, pages 73–79, 1989.
- [25] Anirban Chakraborti, Hrishidev, Hirdesh K. Pharasi, and Kiran Sharma. Phase separation and universal scaling in markets: Fear and fragility. *arXiv preprint arXiv:1910.06242*, 2020.
- [26] Giovanni Bonanno, Guido Caldarelli, Fabrizio Lillo, and Rosario N. Mantegna. Topology of correlation-based minimal spanning trees in real and model markets. *Phys. Rev. E*, 68:046130, 2003.

- [27] J.-P. Onnela, K. Kaski, and J. Kertész. Clustering and information in correlation based financial networks. *The European Physical Journal B*, 38(2):353–362, 2004.
- [28] M. Tumminello, T. Di Matteo, T. Aste, and R. N. Mantegna. Correlation based networks of equity returns sampled at different time horizons. *The European Physical Journal B*, 55(2):209–217, 2007.
- [29] Michele Tumminello, Fabrizio Lillo, and Rosario N. Mantegna. Correlation, hierarchies, and networks in financial markets. *Journal of Economic Behavior & Organization*, 75(1):40–58, 2010. Transdisciplinary Perspectives on Economic Complexity.
- [30] Mark E.J. Newman, Albert-László Barabási, and Duncan J. Watts. *The structure and dynamics of networks*. Princeton University Press, Princeton, 2006.
- [31] Albert-László Barabási. *Network science*. Cambridge University Press, Cambridge, 2016.
- [32] G. Bonanno, G. Caldarelli, F. Lillo, S. Micciché, N. Vandewalle, and R. N. Mantegna. Networks of equities in financial markets. *The European Physical Journal B*, 38(2):363–371, 2004.
- [33] Mark E.J. Newman. *Networks: an introduction*. Oxford University Press, Oxford, 2010.
- [34] Eugene P Wigner. On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, pages 325–327, 1958.
- [35] Eugene P Wigner. Random matrices in physics. *SIAM review*, 9(1):1–23, 1967.
- [36] Vasiliki Plerou, Parameswaran Gopikrishnan, Bernd Rosenow, Luís A Nunes Amaral, and H Eugene Stanley. Universal and nonuniversal properties of cross correlations in financial time series. *Physical review letters*, 83(7):1471, 1999.
- [37] Laurent Laloux, Pierre Cizeau, Jean-Philippe Bouchaud, and Marc Potters. Noise dressing of financial correlation matrices. *Physical review letters*, 83(7):1467, 1999.
- [38] Mel MacMahon and Diego Garlaschelli. Community detection for correlation matrices. *Phys. Rev. X*, 5:021006, 2015.
- [39] Raj Kumar Pan and Sitabhra Sinha. Collective behavior of stock price movements in an emerging market. *Physical Review E*, 76(4):046116, 2007.
- [40] Manan Vyas, T Guhr, and TH Seligman. Multivariate analysis of short time series in terms of ensembles of correlation matrices. *Scientific reports*, 8(1):1–12, 2018.
- [41] John Wishart. The generalised product moment distribution in samples from a normal multivariate population. *Biometrika*, pages 32–52, 1928.

- [42] Vladimir A Marčenko and Leonid Andreevich Pastur. Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1(4):457, 1967.
- [43] Rosario N Mantegna and H Eugene Stanley. *Introduction to econophysics: correlations and complexity in finance*. Cambridge university press, 1999.
- [44] Otakar Borvka. O jistém problému minimálním. 1926.
- [45] Jukka-Pekka Onnela, Anirban Chakraborti, Kimmo Kaski, Janos Kertesz, and Antti Kanto. Asset trees and asset graphs in financial markets. *Physica Scripta*, 2003(T106):48, 2003.
- [46] Robert Clay Prim. Shortest connection networks and some generalizations. *The Bell System Technical Journal*, 36(6):1389–1401, 1957.
- [47] J.-P. Onnela, Anirban Chakraborti, Kimmo Kaski, and J. Kertiész. Dynamic asset trees and portfolio analysis. *The European Physical Journal B - Condensed Matter and Complex Systems*, 30:285–288, 2002.
- [48] James Ladyman, James Lambert, and Karoline Wiesner. What is a complex system? *European Journal for Philosophy of Science*, 3(1):33–67, 2013.
- [49] Filipi N Silva, Cesar H Comin, Thomas K Peron, Francisco A Rodrigues, Cheng Ye, Richard C Wilson, Edwin Hancock, and Luciano da F Costa. Modular dynamics of financial market networks. *arXiv preprint arXiv:1501.05040*, 2015.
- [50] John Von Neumann. *Mathematical foundations of quantum mechanics: New edition*. Princeton university press, 2018.
- [51] Filippo Passerini and Simone Severini. The von neumann entropy of networks. *arXiv preprint arXiv:0812.2597*, 2008.
- [52] Lin Han, Francisco Escolano, Edwin R Hancock, and Richard C Wilson. Graph characterizations from von neumann entropy. *Pattern Recognition Letters*, 33(15):1958–1967, 2012.
- [53] Cheng Ye, Richard C. Wilson, César H. Comin, Luciano da F. Costa, and Edwin R. Hancock. Approximate von neumann entropy for directed graphs. *Phys. Rev. E*, 89:052804, 2014.
- [54] Jianjia Wang, Chenyue Lin, and Yilei Wang. Thermodynamic entropy in quantum statistics for stock market networks. *Complexity*, 2019, 2019.
- [55] Jianjia Wang, Richard C Wilson, and Edwin R Hancock. Spin statistics, partition functions and network entropy. *Journal of Complex Networks*, 5(6):858–883, 2017.
- [56] Satyendra Nath Bose. Plancks gesetz und lichtquantenhypothese. 1924.

- [57] Enrico Fermi. Sulla quantizzazione del gas perfetto monoatomico. *Rendiconti Lincei*, 145, 1926.
- [58] Ginestra Bianconi and Albert-László Barabási. Bose-einstein condensation in complex networks. *Physical review letters*, 86(24):5632, 2001.
- [59] SHEN Yi. Fermi—dirac statistics of complex networks shen yi, zhu di-ling, liu wei-ming (£).
- [60] Assaf Almog and Erez Shmueli. Structural entropy: Monitoring correlation-based networks over time with application to financial markets. *Scientific reports*, 9(1):10832, 2019.
- [61] Hirdesh K Pharasi, Kiran Sharma, Rakesh Chatterjee, Anirban Chakraborti, Francois Leyvraz, and Thomas H Seligman. Identifying long-term precursors of financial market crashes using correlation patterns. *New Journal of Physics*, 20(10):103041, 2018.
- [62] Hirdesh K. Pharasi, Kiran Sharma, Anirban Chakraborti, and Thomas H. Seligman. *Complex Market Dynamics in the Light of Random Matrix Theory*, pages 13–34. Springer International Publishing, Cham, 2019.
- [63] Chandrashekar Kuyyamudi, Anindya S. Chakrabarti, and Sitabhra Sinha. Emergence of frustration signals systemic risk. *Phys. Rev. E*, 99:052306, 2019.
- [64] Yiming Fan, Ling-Li Zeng, Hui Shen, Jian Qin, Fuquan Li, and Dewen Hu. Lifespan development of the human brain revealed by large-scale network eigen-entropy. *Entropy*, 19(9):471, 2017.
- [65] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47, 2002.
- [66] Oleg V Mazurin and EA Porai-Koshits. *Phase separation in glass*. Elsevier, 1984.
- [67] Douglas R. Lloyd, Kevin E. Kinzer, and H.S. Tseng. Microporous membrane formation via thermally induced phase separation. i. solid-liquid phase separation. *Journal of Membrane Science*, 52(3):239 – 261, 1990. Selected papers presented at the Third Ravello Symposium on Advanced Membrane Science and Technology.
- [68] Peter H Poole, Francesco Sciortino, Ulrich Essmann, and H Eugene Stanley. Phase behaviour of metastable water. *Nature*, 360(6402):324, 1992.
- [69] P. van de Witte, P.J. Dijkstra, J.W.A. van den Berg, and J. Feijen. Phase separation processes in polymer solutions in relation to membrane formation. *Journal of Membrane Science*, 117(1):1 – 31, 1996.
- [70] Eduard Leonovich Nagaev et al. *Colossal magnetoresistance and phase separation in magnetic semiconductors*. World Scientific, 2002.

- [71] Vasiliki Plerou, Parameswaran Gopikrishnan, and H. Eugene Stanley. Two-phase behaviour of financial markets. *Nature*, 421:130, 2003.
- [72] H Eugene Stanley. Scaling, universality, and renormalization: Three pillars of modern critical phenomena. *Reviews of modern physics*, 71(2):S358, 1999.
- [73] Michael R. Norman. The challenge of unconventional superconductivity. *Science*, 332(6026):196–200, 2011.
- [74] List of stock market crashes and bear markets. https://en.wikipedia.org/wiki/List_of_stock_market_crashes_and_bear_markets, 2019. Accessed on 7th July, 2019.
- [75] Bull markets. <https://bullmarkets.co/u-s-stock-market-in-1996/>, 2019. Accessed on 7th July, 2019.
- [76] United states housing bubble. https://en.wikipedia.org/wiki/United_States_housing_bubble, 2019. Accessed on 7th July, 2019.
- [77] A short history of stock market crashes. <https://www.cnbc.com/2016/08/24/a-short-history-of-stock-market-crashes.html>, 2019. Accessed on 7th July, 2019.
- [78] Stock market selloff. https://en.wikipedia.org/wiki/2015-16_stock_market_selloff, 2019. Accessed on 7th July, 2019.