PAPER • OPEN ACCESS

Phase separation and scaling in correlation structures of financial markets

View the article online for updates and enhancements.

Journal of Physics: Complexity

PAPER

Phase separation and scaling in correlation structures of financial markets

Anirban Chakraborti[1,2,3,*] ⓘ, Hrishidev[4], Kiran Sharma[5] and Hirdesh K Pharasi[6]

1   School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi-110067, India
2   Centre for Complexity Economics, Applied Spirituality and Public Policy, Jindal School of Government and Public Policy, O P Jindal Global Universiity, Sonipat-131001, India
3   Centro Internacional de Ciencias, Cuernavaca-62210, México
4   Indian Institute of Science Education and Research, Pune-411008, India
5   Chemical & Biological Engineering, Northwestern University, Evanston, Illinois-60208, United States of America
6   Instituto de Ciencias Físicas, Universidad Nacional Autónoma de México, Cuernavaca-62210, México
*   Author to whom any correspondence should be addressed.

E-mail: anirban@jnu.ac.in

Abstract

Financial markets, being spectacular examples of complex systems, display rich correlation structures among price returns of different assets. The correlation structures change drastically, akin to critical phenomena in physics, as do the influential stocks (leaders) and sectors (communities), during market events like crashes. It is crucial to detect their signatures for timely intervention or prevention. Here we use *eigenvalue decomposition* and *eigen-entropy*, computed from *eigenvector centralities* of different stocks in the cross-correlation matrix, to extract information about the disorder in the market. We construct a 'phase space', where different market events (bubbles, crashes, etc) undergo *phase separation* and display *order–disorder* movements. An entropy functional exhibits *scaling* behavior. We propose a generic indicator that facilitates the continuous monitoring of the internal structure of the market—important for managing risk and stress-testing the financial system. Our methodology would help in understanding and foreseeing tipping points or fluctuation patterns in complex systems.

1. Introduction

Even before we could completely recover from the long-lasting effects of the global economic downturn in 2007–08 [1], we are threatened by another impending economic crisis that has been triggered by the coronavirus (COVID-19) pandemic. The last crisis had brought us both predicament and hope! Predicament, since the traditional theories in economics could not predict, not even warn, the near complete breakdown of the global financial system. Hope, since one began to 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 [2–4], which certainly merits more attention. Undoubtedly, the financial market serves as an ideal candidate for modeling a complex system [5, 6], which is generally composed of many constituents of diverse forms and nature 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 complex system through some simple mathematical equations, and new tools and interdisciplinary approaches are much needed. Historically, financial markets have often exhibited sharp and largely unpredictable drops at a systemic scale—'market crashes' [7]. Such rapid changes may be in some cases triggered by unforeseen stochastic events or *exogenous shocks* (e.g., coronavirus pandemic), or more often, they may be driven by certain underlying *endogenous processes* (e.g., housing bubble burst). New insights and concepts, such as systemic risk, tipping points, contagion and network resilience have surfaced in the financial literature, prompting people to better monitor the highly

interconnected macroeconomic and financial systems and, thus, anticipate future economic slowdowns or financial crises.

As a spectacular example of a complex system [8, 9], the financial market [10–13] displays rich correlation structures [14–19], among price returns of different assets, which have often been visualized as correlation-based networks [20–22] with the identification of dominant stocks as influential leaders and sectors as communities [23–25]. The correlation structures often change drastically, as do the leaders and communities in the market, especially during market events like crashes and bubbles [7]. Therefore, the continuous monitoring of the complex structures of the market correlations becomes very crucial and practical [18, 26, 27]. Recently, Pharasi *et al* [18, 19] used the tools of random matrix theory to determine market states and long-term precursors to crashes, and confirmed that during a market crash all the stocks behaved similarly such that the whole market acted like a single huge cluster or community. In contrast, during a bubble period, a particular sector got overpriced or over-performed, causing accentuation of disparities among the various sectors or communities. However, there are no existing formal definitions of market crashes or bubbles; in fact, a certain arbitrariness exists in declaring a market event as a crash or bubble. Hence, it is extremely difficult to detect the signatures of these events so that we can timely intervene or prevent them.

In this paper, we extract information about the *disorder* in the market using the *eigen-entropy* measure [28], computed from the *eigenvector centralities* (ranks) [25] of different stocks in the market, and show for the first time that different market events (correlation structures) undergo *phase separation* [29, 30] in a constructed 'phase space'. For the construction of the phase space we use transformed variables $|H - H_M|$ and $|H - H_{GR}|$, computed from the eigen-entropies $[H, H_M, H_{GR}]$ following the eigenvalue decomposition of the correlation matrices (C) into the market modes (C_M) and the composite group plus random modes (C_{GR}). We further show that all market events, characterized by the $[H, H_M, H_{GR}]$, are either 'business-as-usual' periods (located toward the interior of the phase space) or 'near-critical' events (located at the periphery). Thus, one can see movements in the *order–disorder* phases as market events evolve in the phase space, as observed in critical phenomena of physical systems [31–33]. For robustness, we chose two different financial markets—the US S & P-500 and Japanese Nikkei-225 over a 32 year period, and studied the evolution of the cross-correlation structures and their corresponding eigen-entropies. One of the entropy difference measures, $H - H_M$, displays *scaling* [34] behavior with respect to the mean market correlation μ. Further, a functional of the entropy difference measure, $-\ln(H - H_M)$, acts as a good gauge of the market fear (volatility index VIX) [35]. Analogous to the black-hole entropy that reveals about the internal structure of a black-hole, our methodology with eigen-entropy (measure of market disorder) would also reveal the nature of internal market structure. Further, the phase separation would help us to label the events as anomalies, bubbles, crashes, or other interesting type of events. It would also provide a few generic indicators that would facilitate the continuous monitoring of the internal structure of the correlation epochs [18, 19, 26, 27, 36–38]. We anticipate that this new methodology would help us to better understand the internal market dynamics and characterize the events in different phases as anomalies, bubbles, crashes, etc, which could help in better risk management and portfolio optimization [39]. This could also be easily adapted and broadly applied to the studies of other complex systems such as in brain science [28] or environment [40].
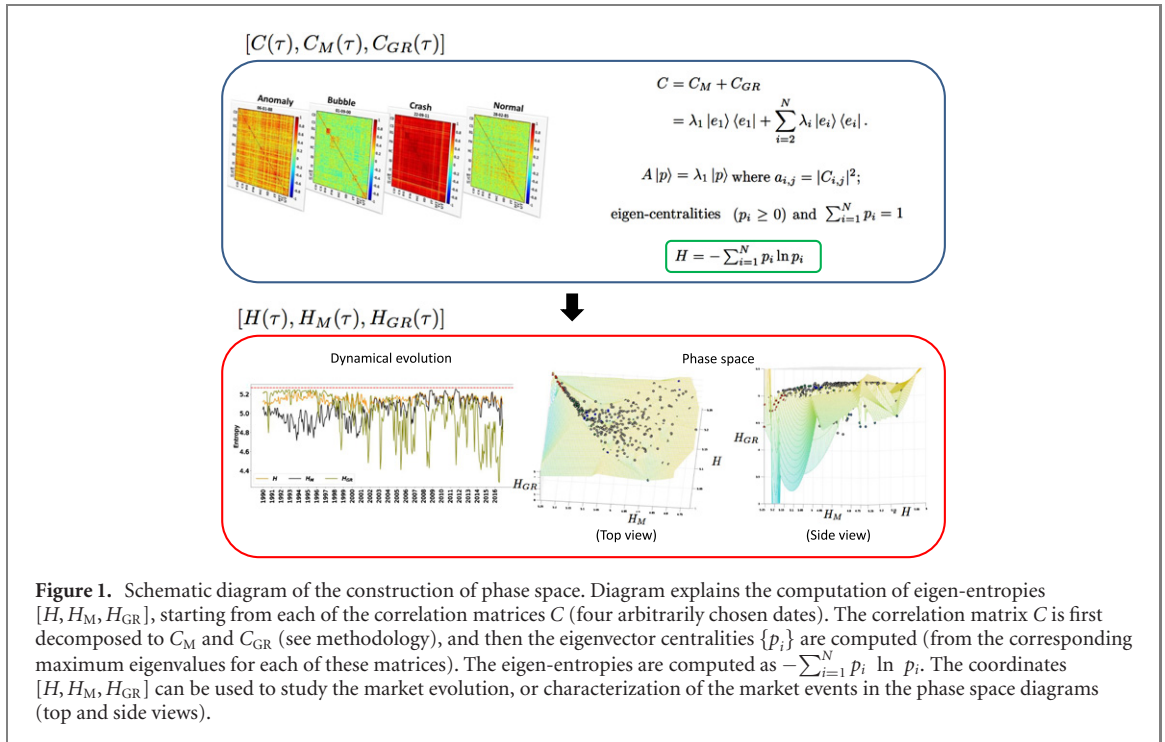
2. Data and methodology

2.1. Data description

2.1.1. Price returns
We have used the adjusted closure price time series for 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, from the Yahoo finance database (https://finance.yahoo.com/; accessed on 7th July, 2017). USA has data for $N = 194$ stocks and the period 02-01-1985 to 30-12-2016 ($T = 8068$ d). JPN has data for $N = 165$ stocks and the period 04-01-1985 to 30-12-2016 ($T = 7998$ d). Note that we have included only those stocks in our analyses, which are present in the data for the entire duration, and added zero return entries corresponding to the missing days. The list of stocks (along with the sectors) for the two markets and the sectoral abbreviations are given in the SI (https://stacks.iop.org/JPCOMPLEX/2/015002/mmedia) tables S1 and S2.

2.1.2. Volatility index
We have also used the daily closure volatility index (VIX) of the Chicago Board Options Exchange (CBOE) from Yahoo finance (https://finance.yahoo.com/; accessed on 13th October, 2019) for the period 02-01-1990 to 30-12-2016, for $T = 6805$ d. It acts as a popular measure of the expectation of volatility in the stock market implied by the S & P-500 index options. It is computed and displayed on a real-time basis by the CBOE, and acts as the 'fear index' or the 'fear gauge' [35].

Figure 1. Schematic diagram of the construction of phase space. Diagram explains the computation of eigen-entropies $[H, H_M, H_{GR}]$, starting from each of the correlation matrices C (four arbitrarily chosen dates). The correlation matrix C is first decomposed to C_M and C_{GR} (see methodology), and then the eigenvector centralities $\{p_i\}$ are computed (from the corresponding maximum eigenvalues for each of these matrices). The eigen-entropies are computed as $-\sum_{i=1}^{N} p_i \ln p_i$. The coordinates $[H, H_M, H_{GR}]$ can be used to study the market evolution, or characterization of the market events in the phase space diagrams (top and side views).

2.2. Methodology

2.2.1. Cross-correlation matrix

The returns series are constructed as $r_i(\tau) = \ln P_i(\tau) - \ln P_i(\tau - \Delta)$, where $P_i(\tau)$ is the adjusted closure price of stock i on day τ, and Δ is the shift in days. Instead of working with a long time series to determine the correlation matrix for N stocks, we work with a *short* time epoch of M days with a shift of Δ days. Then, the equal time Pearson correlation coefficients between stocks i and j are defined as $C_{ij}(\tau) = (\langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle)/\sigma_i \sigma_j$, where $\langle \ldots \rangle$ represents the expectation value computed over the time-epochs of size M and the day ending on τ, and σ_k represents standard deviation of the kth stock evaluated for the same time-epochs. We use $C(\tau)$ to denote the return correlation matrix for the time-epochs ending on day τ (see e.g., figure 1). Here, we show the results for $M = 40$ d with a shift of $\Delta = 20$ d (other choices of M and Δ in SI figure S1).

2.2.2. Eigenvector centrality

The correlation matrix C can be used to produce a correlation-based network [17, 26, 41]. For any given correlation-based network $G := (N, E)$ 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 $A|p\rangle = \lambda|p\rangle$. In general, there may exist many different eigenvalues λ for which a non-zero eigenvector solution $|p\rangle$ exists. We use the characteristic equation $|A - \lambda \mathbb{1}| = 0$ to compute the eigenvalues $\{\lambda_1, \ldots, \lambda_N\}$. However, the additional requirement that all the entries in the eigenvector be non-negative ($p_i \geqslant 0$) implies (by the Perron–Frobenius theorem) that only the *maximum* eigenvalue ($\lambda_{\max}$) results in the desired centrality measure. The ith component of the related eigenvector then gives the relative *eigenvector 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 A can be any matrix with real numbers representing the connection strengths. In order to enforce the Perron–Frobenius theorem, in the entire work, we use $a_{i,j} = |C_{i,j}|^2$, where $i, j = 1, \ldots, N$. In general, we can also consider higher powers n (any positive integer, as discussed and shown in SI figure S2).

2.2.3. Eigenvalue decomposition of cross-correlation matrix

The correlation matrix C of size $N \times N$ will have N eigenvalues, say $\{\lambda_1, \ldots, \lambda_N\}$, which may be arranged in descending order of magnitude. Random matrix theory tells us that the asymptotic behavior of eigenvalues

of large random matrices (whose entries are independent identically distributed random variables) follow the Marčenko–Pastur distribution [42, 43]. By using the eigenvalue decomposition, we can thus filter the true correlations (coming from the signal) and the spurious correlations (coming from the random noise) [14, 15, 44, 45]. The maximum eigenvalue $\lambda_1 = \lambda_{\max}$ of the correlation matrix C, corresponds to a market mode C_M that reflects the aggregate dynamics of the market common across all stocks, and strongly correlated to the mean market correlation μ. The group modes C_G capture the sectoral behavior of the market, which are next few eigenvalues subsequent to the largest eigenvalue of the correlation matrix. Remaining eigenvalues capture the random modes C_R of the market. However, in practice it often becomes difficult to separate distinctly (without arbitrariness) the group modes and random modes. In order to avoid this, we choose to decompose the correlation matrix into the market mode C_M and the composite group plus random mode C_{GR}:

$$C = C_M + C_{GR}$$

$$= \lambda_1 |e_1\rangle\langle e_1| + \sum_{i=2}^{N} \lambda_i |e_i\rangle\langle e_i|. \tag{1}$$

2.2.4. Eigen-entropy
Following the tradition in information theory, we use the eigen-entropy $H = -\sum_{i=1}^{N} p_i \ln p_i$, since all the normalised eigenvector centralities are non-negative ($p_i \geqslant 0$) and $\sum_{i=1}^{N} p_i = 1$, by construction. The eigen-entropy may be described as kind of measure of disorder in the matrix A, where $a_{i,j} = |C_{i,j}|^2$; higher the eigen-entropy, higher is the disorder in the matrix; the highest being in the case of Wishart ensemble that are invariant under orthogonal transformations (WOE-Wishart Orthogonal Ensemble [46]), where $H \sim \ln N$. For empirical correlation matrices, the eigen-entropy will be bounded by these two limits $[0, \ln N]$.

Similarly, corresponding to C_M and C_{GR}, we can compute H_M and H_{GR}, respectively. Thus, from each cross-correlation matrix $C(\tau)$, we can use the eigenvalue decomposition to construct the set of matrices $[C(\tau), C_M(\tau), C_{GR}(\tau)]$, and then construct the set of phase space coordinates $[H(\tau), H_M(\tau), H_{GR}(\tau)]$, as illustrated in figure 1.

2.2.5. Market indicators
Traditionally, the market index returns r or the volatility index (VIX) are used to gauge the state of the market and detect market crises. It has been observed that the eigenvalues of the cross-correlation matrix as well as the mean market correlation μ can also act as indicators of the market state. In this paper, we propose that the eigen-entropy measures also provide important information about the market. Most importantly, the functional $-\ln(H - H_M)$ acts as an excellent market indicator (as discussed below).
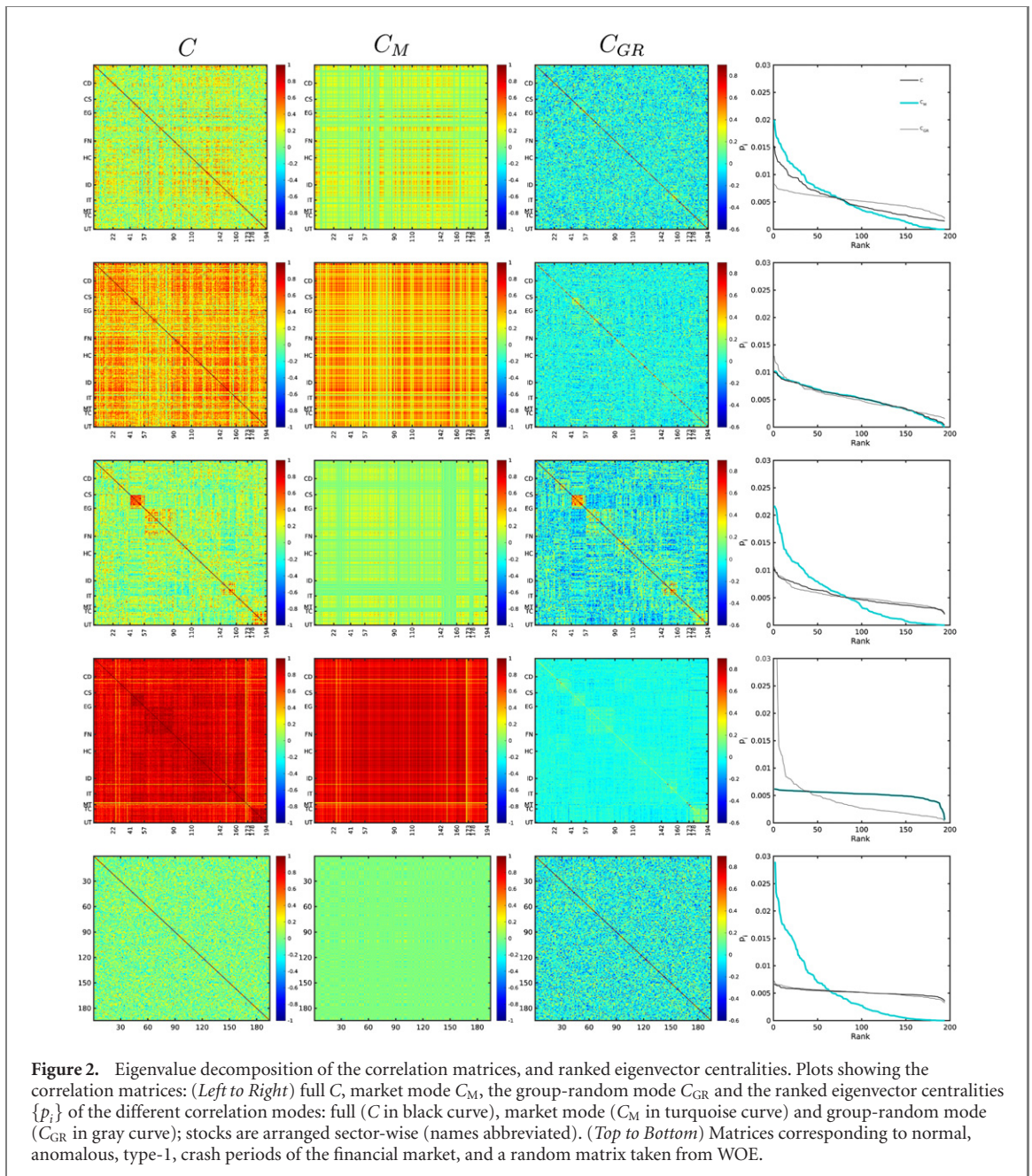
3. Results

3.1. Eigenvalue decomposition and eigenvector centrality
Figure 2 shows the eigenvalue decompositions of the correlation matrices, for: (*Top to Bottom*) normal, anomalous, type-1 event, crash, and WOE. We have denoted the different matrices as: full correlation C, market mode C_M, group-random mode C_{GR}, and displayed the results in figure 2 (*Left to Right*). The last column shows the results for the ranked eigenvector 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 gray curve). Evidently, the internal structure of the cross-correlation matrix changes a lot with time [18, 19, 26, 27, 36–38] and causes the change in the importance/hierarchy of the stocks (leaders) and block structures (communities). This further changes the eigen-entropies $[H(\tau), H_M(\tau), H_{GR}(\tau)]$ that are used to create a phase space where each frame is represented by a point. As time evolves, different parts of the phase space are occupied and this allows us to identify certain phases (restricted to some regions) and characterize the market events as crashes, etc (see supplementary videos 1, 2). 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; in the interesting type-1 period (classified due to the position of the point in certain region of phase space), the curves corresponding to the full and the group-random modes coincide while there is a strict hierarchy in the eigenvector 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 eigenvector centralities of the group-random mode; and for the WOE (without internal structure), once again the curves corresponding to the full and the group-random modes coincide while there is a strict hierarchy in the eigenvector centralities of the market mode.

3.2. Eigen-entropy and phases
As seen in figure 1, the events in phase space $[H, H_M, H_{GR}]$ appear to be scattered on a complicated manifold (created by MATLAB's surface interpolation). However, there appears to be different phases. We thus

Figure 2. Eigenvalue decomposition of the correlation matrices, and ranked eigenvector centralities. Plots showing the correlation matrices: (*Left to Right*) full C, market mode C_M, the group-random mode C_{GR} and the ranked eigenvector 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 gray curve); stocks are arranged sector-wise (names abbreviated). (*Top to Bottom*) Matrices corresponding to normal, anomalous, type-1, crash periods of the financial market, and a random matrix taken from WOE.

plotted the events, as shown in figure 3, in another phase space with transformed variables $[H - H_M, H_M - H_{GR}, H - H_{GR}]$; the points are seen to lie on a flat surface (plane), with different phases (but not segregated). The time-evolution of the transformed variables $[H - H_M, H_M - H_{GR}, H - H_{GR}]$ show interesting dynamics. The characterized events (figures 3(a) and (c)) are indicated as vertical lines in the time-evolution plots (figures 3(b) and (d)). For frame-wise evolution, see SI supplementary videos 1, 2. We found that many anomalies occurred just around the major crashes and intriguing patterns (termed as interesting events of type-1 and type-2, belonging to two distinct regions in the phase space) appeared. 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 'Dot-com bubble' that appear in the $H - H_{GR} \simeq 0$ axis are termed as interesting events of type-1. The events which lie far away from the origin and both the axes, are termed as interesting events of type-2, which include frames with exogenous shocks (like Hurricane Katrina, etc). The events lying close to the origin are like anomalies happening right before or right after major crashes.

3.3. Phase separation, order–disorder movements and scaling
The above interesting features led us to try the transformed variables $|H - H_M|$ and $|H - H_{GR}|$ as independent coordinates of phase space. Very interestingly, as evident from figures 4(a) and (b), the event frames show clear phase separation—anomalies (green region), crashes (red region), normal (gray), type-1 (light
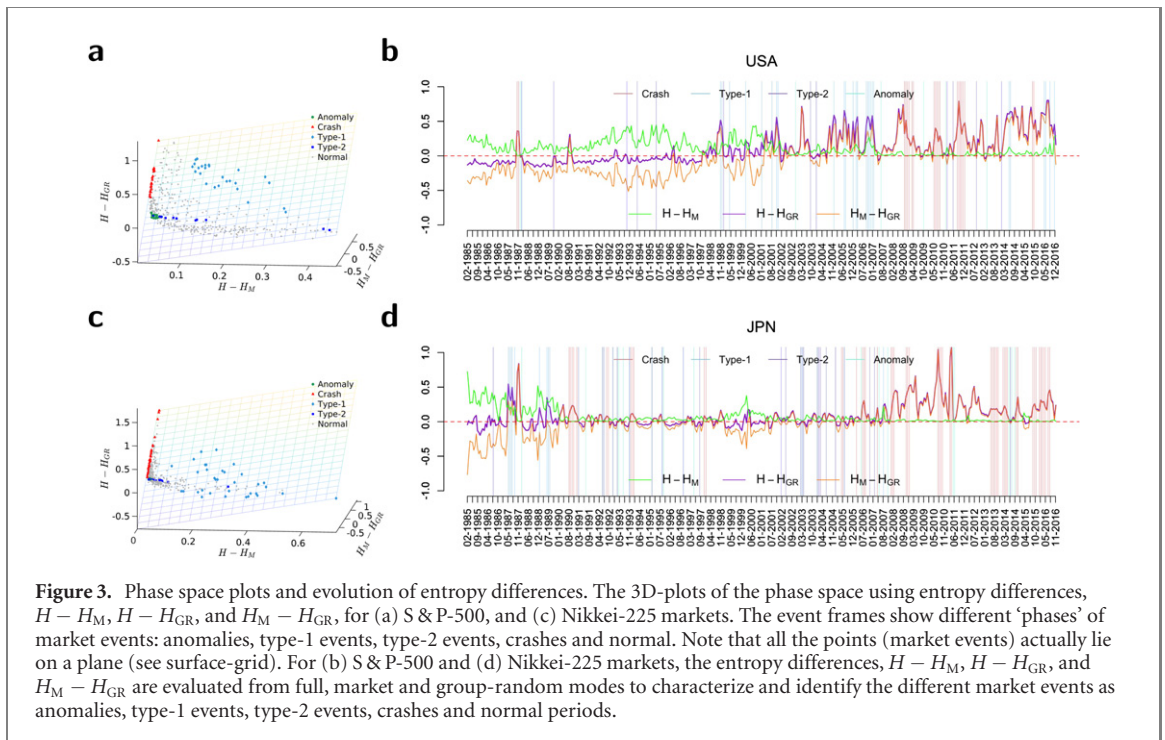
Figure 3. Phase space plots and evolution of entropy differences. The 3D-plots of the phase space using entropy differences, $H - H_M$, $H - H_{GR}$, and $H_M - H_{GR}$, for (a) S & P-500, and (c) Nikkei-225 markets. The event frames show different 'phases' of market events: anomalies, type-1 events, type-2 events, crashes and normal. Note that all the points (market events) actually lie on a plane (see surface-grid). For (b) S & P-500 and (d) Nikkei-225 markets, the entropy differences, $H - H_M$, $H - H_{GR}$, and $H_M - H_{GR}$ are evaluated from full, market and group-random modes to characterize and identify the different market events as anomalies, type-1 events, type-2 events, crashes and normal periods.
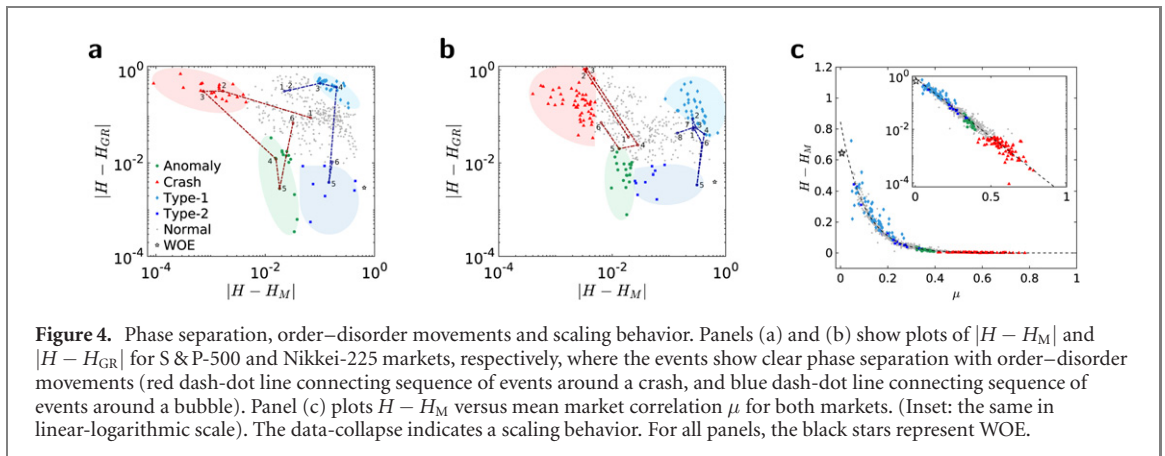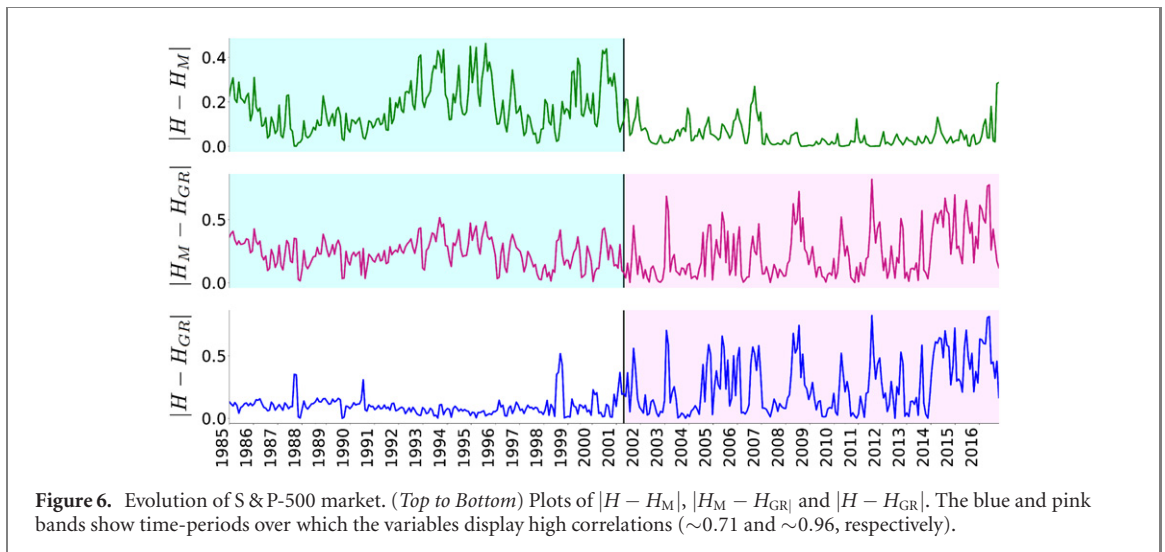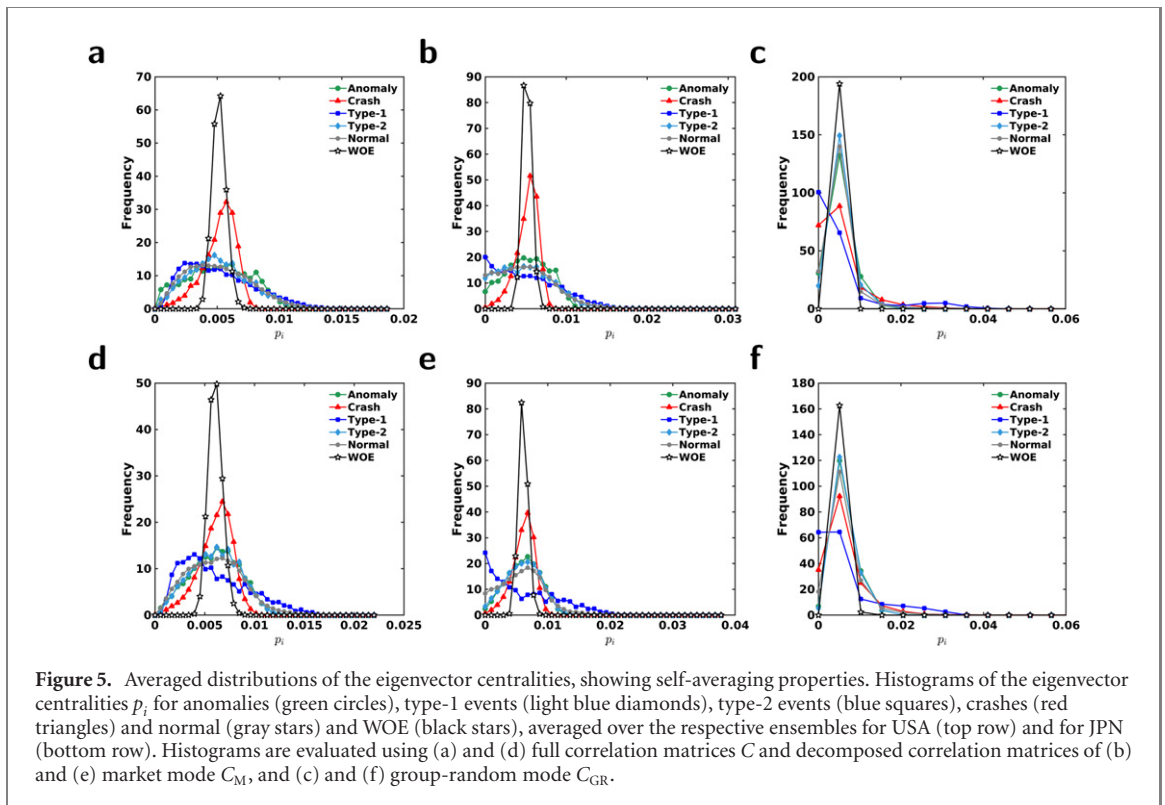

Figure 4. Phase separation, order–disorder movements and scaling behavior. Panels (a) and (b) show plots of $|H - H_M|$ and $|H - H_{GR}|$ for S & P-500 and Nikkei-225 markets, respectively, where the events show clear phase separation with order–disorder movements (red dash-dot line connecting sequence of events around a crash, and blue dash-dot line connecting sequence of events around a bubble). Panel (c) plots $H - H_M$ versus mean market correlation μ for both markets. (Inset: the same in linear-logarithmic scale). The data-collapse indicates a scaling behavior. For all panels, the black stars represent WOE.

blue region) and type-2 (deep blue region), for both S & P-500 and Nikkei-225 markets. The order–disorder movements—normal (at the central region) to near-critical phases (at the peripheral regions) are intriguing.

We have also studied in detail the sequence of six frames to follow the order–disorder movements (SI figures S4 and S5) in cases of major crashes and bubbles (SI table S3). The similar nature of the order–disorder movements in all the major crashes and Dot-com bubbles, nine events in USA and eleven events in JPN, certainly indicate robustness of the method. Moreover, we found that $(H - H_M) \sim \alpha \exp(-\beta\mu)$, where α and β are constants (see figure 4(c) for USA and JPN). We found that the best-fit line yields $\alpha \simeq 0.85 \pm 0.03$ and $\beta \simeq 10.22 \pm 0.25$; adjusted $R^2 = 0.95$. Interestingly, the market event frames segregate into different portions, interspersed by the normal events. This data-collapse on a single curve indicates a scaling behavior [34], which implies that the co-movements in price returns for different financial assets and varying across countries are governed by the same statistical law—certainly non-trivial and striking behavior! This suggests that markets have an inherent structure that remains pretty invariant—it has an average structure with fluctuations (dispersion). The dispersion around the average behavior is slightly more in JPN than USA.

The phase properties are found to be pretty robust, though the phase boundaries are not very sharp (and may depend on the parameters like window choice, shift, etc; SI figure S1and S2). In fact, once we characterize the epochs (event frames) into different 'phases', we can actually create different ensembles of anomalies, type-1 events, type-2 events, crashes and normal events. All frames in a certain phase have very similar properties (hierarchies in ranks of stocks) and can be averaged over to represent a certain phase (figure 5). For each type of event, we find that eigenvector centralities have distinct ranges of values and the sorted eigenvector centrality

Figure 5. Averaged distributions of the eigenvector centralities, showing self-averaging properties. Histograms of the eigenvector centralities p_i for anomalies (green circles), type-1 events (light blue diamonds), type-2 events (blue squares), crashes (red triangles) and normal (gray stars) and WOE (black stars), 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-random mode C_{GR}.

Figure 6. Evolution of S & P-500 market. (*Top to Bottom*) Plots of $|H - H_M|$, $|H_M - H_{GR}|$ and $|H - H_{GR}|$. The blue and pink bands show time-periods over which the variables display high correlations (~0.71 and ~0.96, respectively).

curves have interesting features (hierarchies) in the eigenmodes. The eigen-entropies actually quantify these features appropriately. For the S & P-500 and Nikkei-225 markets, we compute the histograms of the eigenvector centralities p_i. Figure 5 shows the histograms (for S & P-500 (*Top*) and Nikkei-225 (*Bottom*)) for all the characterized events (anomalies, crashes, etc), averaged over the respective ensembles, for the full/decomposed matrices. For comparison, we also plot the results for the WOE (black stars). This helps us understand what actually happens in the market, during these different types of events (characterized as phases) and what type of hierarchies exist within the stocks's eigenvector centralities. This would shed new light into the understanding of formation of type-1 events, their development and crashes, etc. It is interesting to note that the properties remain similar across different markets (USA and JPN) and across various periods of time.

One could also simulate (to be reported elsewhere) various correlation structures from a correlated WOE with the mean correlation as tuning parameter. The non-trivial inherent market structure (sectors or communities) plays a crucial role in the observed scaling behavior. We also observed from the evolution of the variables $|H - H_M|$, $|H_M - H_{GR}|$ and $|H - H_{GR}|$ (figure 6), and other market indicators (SI figures S6 and S7) that the market behavior has changed radically after 2002 (USA) and 1990 (JPN) corroborating the findings of our earlier work [18].

Figure 7. Evolution of S & P-500 market indicators. (*Top to Bottom*) Plots of index returns r, volatility index (VIX), mean market correlation μ and the functional $-\ln(H - H_{\mathrm{M}})$, for 1990–2016. Light blue bands are representative critical periods (see SI table S3).
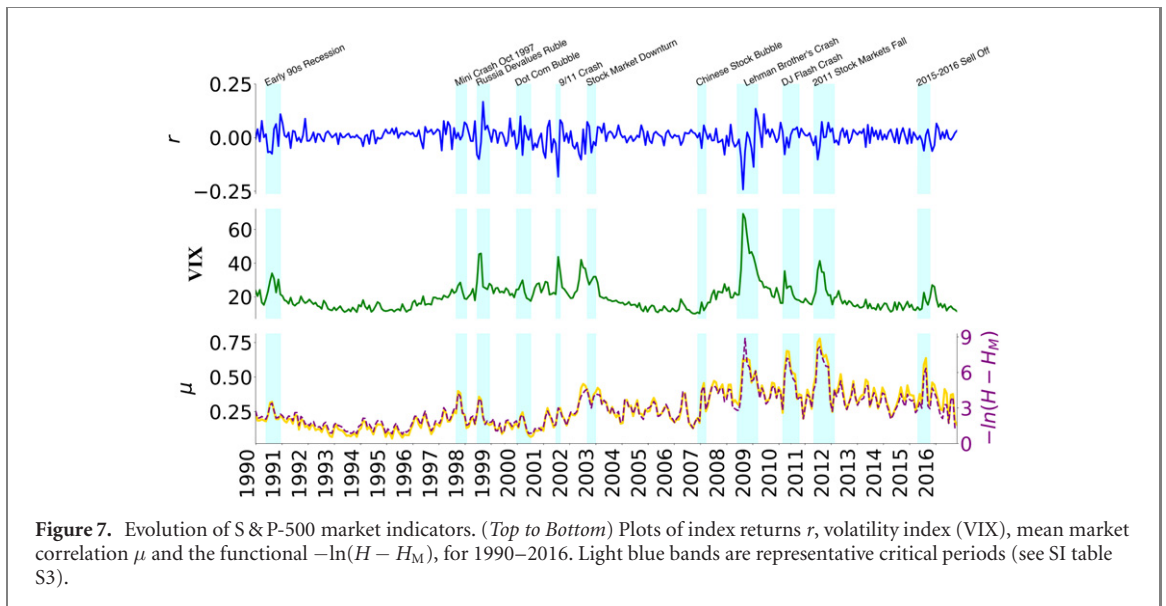
3.4. Generic market indicator

Finally, the functional $-\ln(H - H_{\mathrm{M}})$ is found to act as a good gauge of the market characteristic (μ) and market fear (VIX) (figure 7). There exist significant and non-trivial correlations between these variables, and the other market indicators (SI figure S8). Hence, this functional $-\ln(H - H_{\mathrm{M}})$ can serve as a very good generic indicator.

4. Summary and discussions

We summarize the findings of our paper:

(a) We have proposed a methodology to extract information about the *disorder* in the market using the eigen-entropy measures $[H, H_{\mathrm{M}}, H_{\mathrm{GR}}]$, computed from the distributional properties of the eigenvector centralities $\{p_i\}$ of correlation matrices (and eigenmodes) of asset returns. We emphasize here that our eigen-entropy measure has a few advantages-uniquely determined, non-arbitrary, computational cheap (low complexity), when compared to existing methods, e.g., structural entropy [26]. Note that the structural entropy (or any other network-based entropy measures) is very sensitive to the community structure and construction of the network. An algorithm [24] involves identifying the group mode from the correlation matrix, which may be hard and non-arbitrary (the boundary determined by the eigenvalues of the correlation matrix is not sharp).

(b) We have shown that the different market events (corresponding to different correlation structures) like crashes, bubbles, etc undergo phase separation in a constructed 'phase space'. For the demonstration of phase separation in a phase space, we used transformed variables $|H - H_{\mathrm{M}}|$ and $|H - H_{\mathrm{GR}}|$, computed from the eigen-entropies $[H, H_{\mathrm{M}}, H_{\mathrm{GR}}]$ following the eigenvalue decomposition of the correlation matrices (C) into the market modes (C_{M}) and the composite group plus random modes (C_{GR}). We reiterate that this type of phase separation behavior has never been recorded for financial markets; it is very distinct from the two-phase behavior in financial markets reported earlier by Plerou *et al* [47].

(c) We also showed that all market events, characterized by the $[H, H_{\mathrm{M}}, H_{\mathrm{GR}}]$, are either 'business-as-usual' periods (located toward the interior of the phase space) or 'near-critical' events (located at the periphery). We demonstrated movements in the *order–disorder* phases, for all the major crashes and bubbles (nine events in USA and eleven events in JPN). Our results are pretty robust, as we found similar features in two different financial markets—the US S & P-500 and Japanese Nikkei-225, and over a very long span of time (32 years).

(d) One of the entropy difference measures, $H - H_{\mathrm{M}}$, displayed scaling behavior (data collapse) with respect to the mean market correlation μ. The data collapse certainly 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 scaling 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. In addition, this may lead to a foundation for understanding scaling in a broader context, and providing us with altogether new concepts not anticipated previously.

(e) We also showed that the functional of the entropy difference measure, $-\ln(H - H_{\mathrm{M}})$, acted as a good gauge of the market fear (volatility index VIX). This methodology may be generalized and used in other complex systems (to be reported elsewhere) to understand and foresee tipping points and fluctuation patterns. Our proposed methodology may further help us to understand the market events and their dynamics, as well as find the time-ordering and appearances of the bubbles (formations or bursts) and crashes, separated by normal periods.

Acknowledgments

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

Anirban Chakraborti ⬤ https://orcid.org/0000-0002-6235-0204

References

[1]   Altman R C 2009 *Foreign Aff.* **88** 1
[2]   Sharma K, Gopalakrishnan B, Chakrabarti A S and Chakraborti A 2017 *Sci. Rep.* **7** 8055
[3]   Sharma K, Chakrabarti A S and Chakraborti A 2019 Multi-layered network structure: relationship between financial and macroeconomic dynamics *New Perspectives and Challenges in Econophysics and Sociophysics* (Berlin: Springer) pp 117–31
[4]   Chakrabarti A S, Chakraborti A, Upmanyu S 2020 Macroeconomic and financial networks: review of some recent developments in parametric and non-parametric approaches *Proc. of the Int Con. of Game Theory and Networks-ICGTN 2019* ed *et al*
[5]   Boccara N 2010 *Modeling Complex Systems* (Berlin: Springer)
[6]   Foote R 2007 *Science* **318** 410–2
[7]   Sornette D 2004 *Why Stock Markets Crash: Critical Events in Complex Financial Systems* (Princeton, NJ: Princeton University Press)
[8]   Goldenfeld N and Kadanoff L P 1999 *Science* **284** 87–9
[9]   Arthur W B 1999 *Science* **284** 107–9
[10] Mantegna R N and Stanley H E 2007 *An Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge: Cambridge University Press)
[11] Bouchaud J P and Potters M 2003 *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management* (Cambridge: Cambridge University Press)
[12] Sinha S, Chatterjee A, Chakraborti A and Chakrabarti B K 2010 *Econophysics: An Introduction* (New York: Wiley)
[13] Chakraborti A, Toke I M, Patriarca M and Abergel F 2011 *Quant. Finance* **11** 991–1012
[14] Laloux L, Cizeau P, Bouchaud J-P and Potters M 1999 *Phys. Rev. Lett.* **83** 1467–70
[15] Plerou V, Gopikrishnan P, Rosenow B, Nunes Amaral L A and Stanley H E 1999 *Phys. Rev. Lett.* **83** 1471–4
[16] Kullmann L, Kertész J and Kaski K 2002 *Phys. Rev.* E **66** 026125
[17] Heimo T, Tibély G, Saramäki J, Kaski K and Kertész J 2008 *Phys.* A **387** 5930–45
[18] Pharasi H K, Sharma K, Chatterjee R, Chakraborti A, Leyvraz F and Seligman T H 2018 *New J. Phys.* **20** 103041
[19] Pharasi H K, Sharma K, Chakraborti A and Seligman T H 2019 Complex market dynamics in the light of random matrix theory *New Perspectives and Challenges in Econophysics and Sociophysics* (Berlin: Springer) pp 13–34
[20] Bonanno G, Caldarelli G, Lillo F and Mantegna R N 2003 *Phys. Rev.* E **68** 046130
[21] Onnela J P, Chakraborti A, Kaski K, Kertesz J and Kanto A 2003 *Phys. Rev.* E **68** 056110
[22] Tumminello M, Lillo F and Mantegna R N 2010 *J. Econ. Behav. Organ.* **75** 40–58
[23] Fortunato S 2010 *Phys. Rep.* **486** 75–174
[24] MacMahon M and Garlaschelli D 2015 *Phys. Rev.* X **5** 021006
[25] Barabási A L 2016 *Network Science* (Cambridge: Cambridge University Press)
[26] Almog A and Shmueli E 2019 *Sci. Rep.* **9** 10832
[27] Kuyyamudi C, Chakrabarti A S and Sinha S 2019 *Phys. Rev.* E **99** 052306
[28] Fan Y, Zeng L L, Shen H, Qin J, Li F and Hu D 2017 *Entropy* **19** 471
[29] Mazurin O V and Porai-Koshits E 1984 *Phase Separation in Glass* (Amsterdam: Elsevier)
[30] Poole P H, Sciortino F, Essmann U and Stanley H E 1992 *Nature* **360** 324
[31] Stanley H E 1971 *Phase Transitions and Critical Phenomena* (Oxford: Clarendon)
[32] Sethna J 2006 *Statistical Mechanics: Entropy, Order Parameters, and Complexity* vol 14 (Oxford: Oxford University Press)
[33] Goldenfeld N 2018 *Lectures on Phase Transitions and the Renormalization Group* (Boca Raton, FL: CRC Press)
[34] Stanley H E 1999 *Rev. Mod. Phys.* **71** S358
[35] 2020 *Volatility index (VIX)* https://en.wikipedia.org/wiki/VIX (accessed 7 July 2019)
[36] Drożdż S, Grümmer F, Górski A, Ruf F and Speth J 2000 *Phys.* A **287** 440–9

[37] Drożdż S, Grümmer F, Ruf F and Speth J 2001 *Phys.* A **294** 226–34
[38] Sandhu R S, Georgiou T T and Tannenbaum A R 2016 *Sci. Adv.* **2** e1501495
[39] Markowitz H 1952 *J. Finance* **7** 77–91
[40] Chakraborti A, Sharma K, Pharasi H K, Bakar K S, Das S and Seligman T H 2020 *New J. Phys.* **22** 063043
[41] Kukreti V, Pharasi H K, Gupta P and Kumar S 2020 *Front. Phys.* **8** 323
[42] Marčenko V A and Pastur L A 1967 *Math. USSR-Sb.* **1** 457
[43] Wishart J 1928 *Biometrika* **20A** 32–52
[44] Kwapień J, Drożdż S, Oświe P *et al* 2006 *Phys.* A **359** 589–606
[45] Kwapień J and Drożdż S 2012 *Phys. Rep.* **515** 115–226
[46] Vinayak and Seligman T H 2014 Time series, correlation matrices and random matrix models *AIP Conf. Proc.* **1575** 196–217
[47] Plerou V, Gopikrishnan P and Stanley H E 2003 *Nature* **421** 130