New Journal of Physics
The open access journal at the forefront of physics
Deutsche Physikalische Gesellschaft DPG IOP Institute of Physics

PAPER • OPEN ACCESS

Emerging spectra characterization of catastrophic instabilities in complex systems

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics
Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

CrossMark

OPEN ACCESS

RECEIVED
24 January 2020

REVISED
20 April 2020

ACCEPTED FOR PUBLICATION
6 May 2020

PUBLISHED
22 June 2020

PAPER

Emerging spectra characterization of catastrophic instabilities in complex systems

Anirban Chakraborti[1,2], Kiran Sharma[3], Hirdesh K Pharasi[4], K Shuvo Bakar[5], Sourish Das[6] and Thomas H Seligman[2,4]

1 School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi-110067, India
2 Centro Internacional de Ciencias, Cuernavaca-62210, México
3 Chemical & Biological Engineering, Northwestern University, Evanston, Illinois-60208, United States of America
4 Instituto de Ciencias Físicas, Universidad Nacional Autónoma de México, Cuernavaca-62210, México
5 Data61, CSIRO, Canberra-2601, Australia
6 Chennai Mathematical Institute, Chennai-603103, India

E-mail: anirban@jnu.ac.in

Keywords: market crash, random matrix theory, eigen spectra, environmental ozone, return cross-correlations, power mapping method

Supplementary material for this article is available online

Abstract

Random matrix theory has been widely applied in physics, and even beyond physics. Here, we apply such tools to study catastrophic events, which occur rarely but cause devastating effects. It is important to understand the complexity of the underlying dynamics and signatures of catastrophic events in complex systems, such as the financial market or the environment. We choose the USA S & P-500 and Japanese Nikkei-225 financial markets, as well as the environmental ozone system in the USA. We study the evolution of the cross-correlation matrices and their eigen spectra over different short time-intervals or 'epochs'. A slight non-linear distortion is applied to the correlation matrix computed for any epoch, leading to the *emerging spectrum* of eigenvalues, mainly around zero. The statistical properties of the emerging spectrum are intriguing—the smallest eigenvalues and the shape of the emerging spectrum (characterized by the spectral entropy) capture the system instability or criticality. Importantly, the smallest eigenvalue could also signal a precursor to a market catastrophe as well as a 'market bubble'. We demonstrate in two paradigms the capacity of the emerging spectrum to understand the nature of instability; this is a new and robust feature that can be broadly applied to other physical or complex systems.

1. Introduction

Critical transitions or sharp changes, which are usually unpredictable, are ubiquitous in complex systems found in nature and society—ranging from natural hazards, such as earthquakes, volcanic eruptions, hurricanes, lightning strikes, extreme weather conditions due to global warming [1–3], failure of physical and social structures due to large scale terror strikes [4, 5], market crashes and economic slowdowns, electric grid failures, traffic breakdowns, disease and epidemic spreads, etc [6–13]. Such extreme events often reveal underlying dynamical processes and thus provide ground for better scientific understanding of complex systems like stock markets, fractures or earthquakes [14, 15], climate, etc. The detailed evolution of complex systems may not have much significance; it may be more relevant to study certain phases of the evolution, like the rare critical events. A forecasting or prediction algorithm may be required for answering important questions related to rare events—predicting the occurrence of seismic events or temperature rise, forecasting crashes or bubbles, etc. Recently, scientists have therefore focused their attention on identifying generic indicators that may detect if a complex system is close to instabilities. Methods from self-organized criticality [16], networks [17], random matrix theory (RMT) [18, 19], etc have been used extensively to model and analyze such complex systems [20].

RMT has been applied in various problems [21, 22] such as in many-body physics [23], disordered systems and chaos [24], quantum chromodynamics [25], models for interacting fermions [26] and in other fields beyond physics, including financial markets [27–32]. Besides traditional methods, the RMT was proposed as a method for noise suppression in financial time series [33–35], which essentially proceeded as follows: the correlation matrices were diagonalized by an orthogonal transformation and then the bulk of the spectrum was eliminated; the inverse transformation reconstructed a singular correlation matrix with zero eigenvalues corresponding to the number of eigenvalues retained minus one. The power mapping introduced by Guhr and Kälber [36] was a less radical alternative; this was elaborated further in another paper [37]. The latter paper emphasized that the main advantage of the power mapping was not to omit all the information hidden in the statistics of the bulk. Using certain models, it was later demonstrated by Guhr and Wirtz [38], that the smallest eigenvalues, in particular, carry lot of information, making the highest eigenvalues and their eigenfunctions look like a rather rough approximation. Further applications of the power mapping were shown, e.g., in references [39, 40] as well as in reference [41], when focusing on data-sets of short time series in the financial sector. The interest in eigenvalues for sets of N time series of length T, for $N \gg T$, were mainly focused on the largest one (representing the market mode) or some of the larger ones (representing the market sectors), for the simple reason that there were only a few non-zero eigenvalues left. In reference [42], the authors discussed the use of the power mapping to break the linear dependence among the rows of the corresponding data matrices, leading to a large $N - T + 1$ dimensional subspace with eigenvalues in the ensuing correlation matrices. The power mapping would lift this degeneracy, even if the power was very close to unity. The resulting spectrum was well separated from the original one and was named as emerging spectrum. For very small deviations from the power unity, analytic results were given for Wishart matrices (white noise data) and numeric exploration of correlated Wishart ensembles were given, as well. Recently, in the context of market states [41], an intermediate view on power mapping for very short time series was adopted, where the significance of the largest eigenvalues were retained but the noise reduction was linked to the determination of optimal clustering choices. Following this idea, some extended studies have been recently done [19, 43–46].

Pharasi *et al* [19] had earlier studied the role of power mapping in noise suppression and detection of ‘market states’ using similarity measures between correlation matrices and corresponding multidimensional scaling maps. Pharasi *et al* [45] had also briefly reviewed the role of random matrix theory in the studying market dynamics, where they discussed the traditional Marcenko–Pastur distribution in WOE and CWOE, as well as the statistical properties of the emerging spectra arising in WOE and CWOE after application of power mapping.

Our overall aim is the detection of critical instabilities in complex systems using the emerging spectra. It must be mentioned that there is a closely related paper by Rinn *et al* [47], which explores such stabilities/instabilities in a dynamical way. It may also be mentioned that our approach to the analysis of complex dynamics could potentially be used in many different contexts such as in biological, ecological, and physiological systems [48–52].

In this paper, in the wider framework of RMT, we propose the usage of the shape of the ‘emerging spectrum’ (characterized by the spectral entropy [53]) as a simple yet generic indicator of critical periods, and the smallest eigenvalue as a signal for precursor to a market catastrophe and a market bubble. Importantly, the indicator can be bench-marked against the Wishart orthogonal ensemble (WOE) and it removes the arbitrariness of threshold, etc. We demonstrate its validity and robustness in two different complex systems.

As paradigm we first choose the stock market, which is a fascinating example of a complex system [54–56], where the coherent collective behavior of the economic agents and their repeated nonlinear interactions, often lead to rich structures of correlations and time-dependencies [27–29]. The movements in the market prices are often influenced by news or external shocks, which can result in the unforeseen and rapid drop in the prices of a large section of the stock market, labeled as a market crash! On the contrary, the widespread existence of bubbles in financial markets and extreme movements of price return series often result from the unstable relationship between macroeconomic fundamentals of the economy and the asset prices [57]. Since the societal impact of an extreme event like a market crash can be catastrophic [9, 58], the understanding of such events [59], the assessment of the associated risks [60], and possible prediction of these events have drawn attention from all quarters: governments, industry participants and academia.

Second, we choose to study environmental pollution [61, 62], which in the last few decades has reached alarming levels! One major component is air pollution, where harmful or excessive quantities of substances including gases and particles are introduced into earth’s atmosphere [63, 64]. According to the intergovernmental panel on climate change, top three most important air pollutants are carbon dioxide (CO_2), methane (CH_4), and *tropospheric* ozone (O_3) [65, 66]. The tropospheric O_3 has been radically worsening the pollution and changing the climatic behavior [64, 67]. The ozone pollution in the earth's environment is an example of a complex system—a product of complex interlinked networks of relationships between humans interactions and non-living objects surrounding them, as well as the cycles controlling the flows of chemical elements or compounds that support life and regulate climate. Polluting actions of humans often result in positive feedback loops that are known to destabilize an environmental system and drive it toward an extreme or critical state. We focus on the study of catastrophic instabilities in the environmental ozone pollution vis-à-vis the financial market.

2. Methodology and results

Our method relies on the time evolution of the cross-correlation matrices for N time-series and the eigen spectra over different time-epochs (of size M), as traditionally analyzed in RMT or in the analysis of adaptive complex systems like financial markets [19, 41, 45, 68–70]. The market returns series are constructed as $r_i(\tau) = \ln P_i(\tau) - \ln P_i(\tau - 1)$, where $P_i(\tau)$ is the adjusted closure price of stock i on day τ. Then the equal time Pearson correlation coefficients between stocks i and j is 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 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 $\boldsymbol{C}(\tau)$ to denote the return correlation matrix for the time-epochs ending on day τ.

For this type of analysis, one assumes stationarity of the underlying time series. As this assumption manifestly fails for longer time series, it is often useful to break the long time series of length T, into n time-epochs of size M (such that $T/M = n$). The assumption of stationarity improves for the shorter time-epochs used. However, if there are N return time series such that $N > M$, this implies an analysis of highly singular correlation matrices with $N - M + 1$ zero eigenvalues, which lead to poor statistics. This problem can be avoided by using the non-linear 'power mapping', which was introduced to reduce noise [36, 37, 42]. It breaks the degeneracy of the zero eigenvalues also named as zero modes [71]. Thus, we apply a small non-linear distortion to the coefficients of the cross-correlation matrix: $C_{ij} \to (\mathrm{sign}\, C_{ij})|C_{ij}|^{1+\epsilon}$, where $\epsilon = 0.01$. This breaks the degeneracy of zero-eigenvalues, giving rise to the *emerging spectrum* near zero [42, 71] and the smallest eigenvalues are often negative! We had earlier used the power mapping to study the properties of markets and market states [19, 45]. Here, we would like to focus on the evolution of the emerging spectra.

2.1. Financial market

For empirical data, we have used the adjusted closure price time series from the Yahoo finance database [72], for two countries: United States of America (USA) S & P-500 index (for the period 02-01-1985 to 30-12-2016 consisting of $T = 8068$ trading days, and number of stocks $N = 194$) and Japan (JPN) Nikkei-225 index (for the period 04-01-1985 to 30-12-2016 consisting of $T = 7998$ trading days and number of stocks $N = 165$). We have included the stocks that are present in the index for the entire duration. Note that we have $T = 7897$ trading days data for the Nikkei-225 index and $T = 7998$ trading days data for 165 stocks; so we add zero entries corresponding to the missing index data for the entire time-series (without affecting the results or conclusions). In figure 1(a), four correlation structures with $M = 200$ days (non-singular matrices) are shown. Evidently, the correlation structure (along with the mean market correlation $\mu(\tau)$) varies over time—the market has a highly correlated structure (with high mean correlation) during the critical or crash epoch of 200 days ending on 22-09-2011, and an interesting structure mixed with correlations and anti-correlations (with low mean) during a relatively calm epoch of 200 days ending 28-02-1985. At times, one can see that there are strong correlations within certain stocks/sectors and anti-correlations with respect to other stocks/sectors (06-01-1988 and 01-09-2000); at certain times, *all* the stocks/sectors are correlated, with the mean market correlation being very high (critical periods). Figure 1(b) shows a cross-correlation matrix computed for *short* epoch of $M = 20$ days ending on 08-07-1985. It has both correlation and anti-correlation present in the correlation pattern and shows non-critical (normal) behavior of market. Figure 1(c) shows the eigen spectrum of the correlation matrix of non-critical (normal) period, evaluated for the *short* time series of returns for the epoch of $M = 20$ days ending on 08-07-1985, with the maximum eigenvalue $\lambda_{\max} = 29.59$ (not plotted). Inset of figure 1(c) shows the emerging spectrum generated using power mapping ($\epsilon = 0.01$) is a deformed semi-circle, with the smallest eigenvalue $\lambda_{\min} = -0.011$. Figure 1(d) shows a critical (crash) period correlation matrix, evaluated for the *short* time-epoch of $M = 20$ days ending on 15-09-2008. Its eigenvalue spectrum is shown in figure 1(e) with the maximum eigenvalue $\lambda_{\max} = 94.92$ (not plotted). Inset of figure 1(e) shows the emerging spectrum using power mapping ($\epsilon = 0.01$), which is heavy-tailed

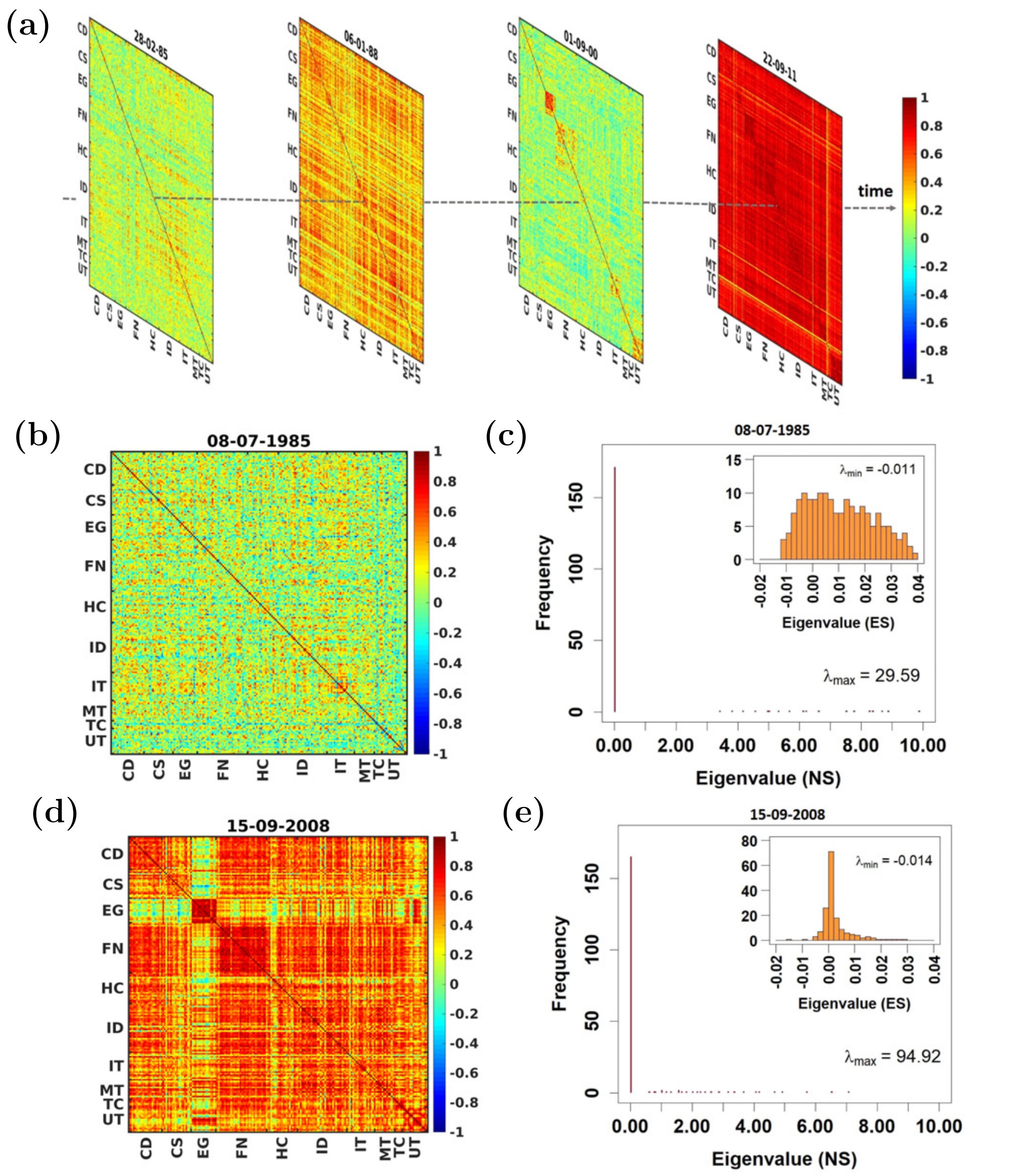

Figure 1. USA market correlations and eigen spectra. (a) Schematic diagram showing four sample correlation matrices for different time-epochs of $M = 200$ days ending on 28-02-1985, 06-01-1988, 01-09-2000, and 22-09-2011. (b) Non-critical (normal) period cross-correlation matrix computed for *short* time series of $M = 20$ days ending on 08-07-1985. (c) Eigen spectrum corresponding to (b). Inset: emerging spectrum using power mapping ($\epsilon = 0.01$) is deformed semi-circular. (d) Critical (crash) period cross-correlation matrix computed for *short* time series of $M = 20$ days ending on 15-09-2008. (e) Eigen spectrum corresponding to (d). Inset: emerging spectrum using power mapping ($\epsilon = 0.01$) is heavy-tailed Lorentzian-like. The sectoral abbreviations are as follows: **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.

Lorentzian-like, with the smallest eigenvalue $\lambda_{\min} = -0.014$. Note, the insets of figure 1(c) and (e) elucidate that the emerging spectra are considerably different—slightly deformed semi-circular/Lorentzian-like as the market is normal/critical.

Figures 2(a) and (b) show the evolution of the emerging spectra for the USA and JPN markets, respectively.

Figures 3(a) and (b) show the statistical time series analyses: (i) market returns $r(\tau)$, (ii) mean market correlation $\mu(\tau)$, (iii) smallest eigenvalue of the emerging spectrum $\lambda_{\min}$, (iv) t-value of the t-test, which tests if lag-1 smallest eigenvalue $\lambda_{\min}(\tau - 1)$ has statistically significant effect with mean market correlation $\mu(\tau)$, and (v) spectral entropy S_{ES}. The mean of the correlation coefficients and the smallest eigenvalue in the emerging spectra are correlated ($\sim$0.6) to a large extent. Notably, the smallest eigenvalue behaves differently (sharply rising *or* falling) at the same time when the mean market correlation is very high (crash). It is evident that for the USA, from 2001 onward, the financial market has become more turbulent and in case of JPN, from 1990 onward.

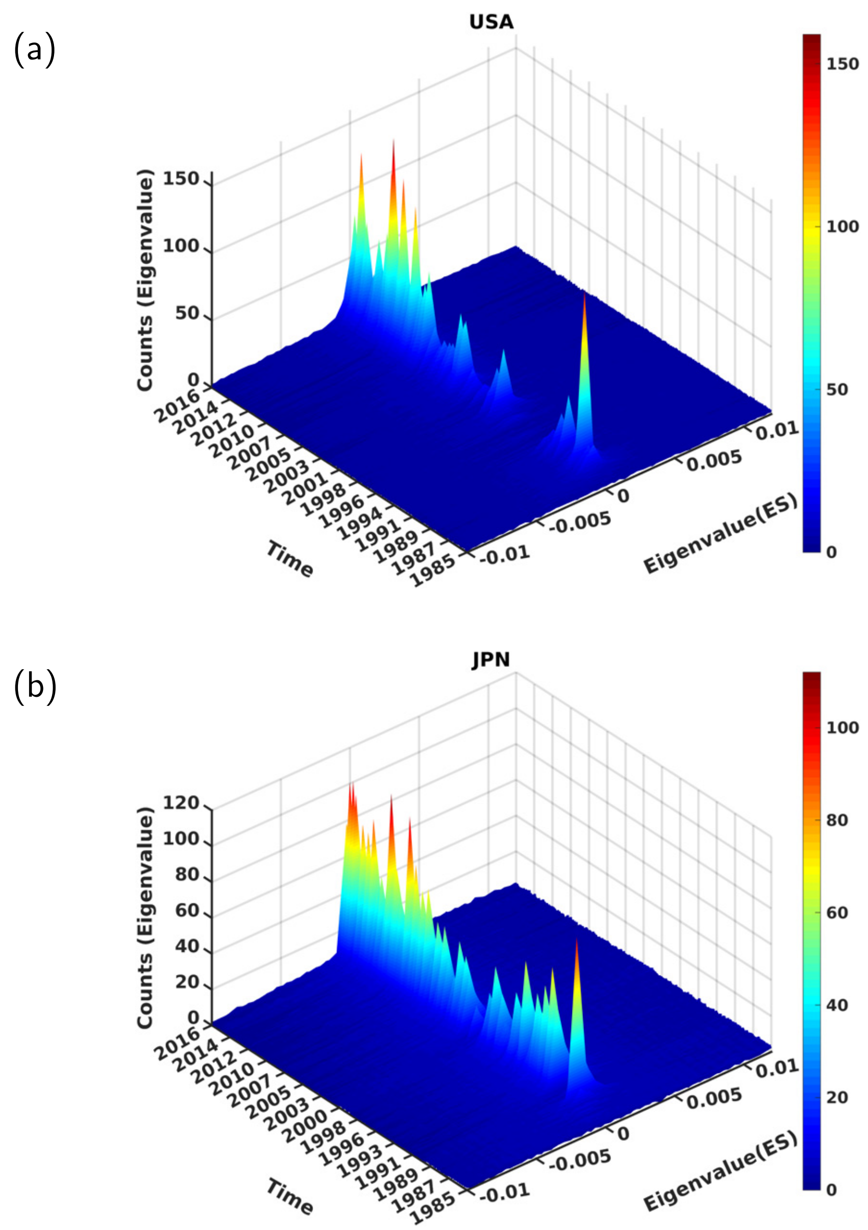

Figure 2. Evolution of emerging eigen spectra for identification of catastrophic instabilities. (a) **USA** and (b) **JPN**.

2.2.1. Linear regression model for market correlation on lagged smallest eigenvalue

In order to test the statistical significance of the correlation between last day's (lag-1) smallest eigenvalue $\lambda_{\min}(\tau-1)$ and the current day's mean market correlation $\mu(\tau)$, we consider a linear regression model for μ (mean market correlation):

$$\mu(\tau) = \beta_0 + \beta_1\lambda_{\min}(\tau-1) + \beta_2\lambda_{\min}(\tau-2) + \ldots$$
$$+ \ \beta_p\lambda_{\min}(\tau-p) + \epsilon(\tau),$$

where the β's are the coefficients to be estimated, $\epsilon \sim N(0,\sigma^2)$, $\tau = 0,1,2,\ldots,T$ is the white noise, and $\lambda_{\min}(\tau-p)$'s are the lag-p smallest eigenvalues. The choice of p is arbitrary to some extent. We tried with values $p = 1,2,\ldots,5$. For $p = 1$ and 2, we observed that almost all t-values are overwhelmingly large, especially because we considered a moving window approach. In the cases of $p = 1$ and $p = 2$, due to significant overlap over the windows, the large t-values resulted in too many false positives. On the other hand, when we chose $p = 3$ and above, the overlap between the windows reduced. As a result, the t-values become moderate as compared to $p = 1$ and 2 (see figure S6 in the supplementary information: http://stacks.iop.org/NJP/22/063043/mmedia). We also observed that the predictive powers for $p = 3,4$ and 5 remained similar. Hence, we chose $p = 3$ as a parsimonious model choice.

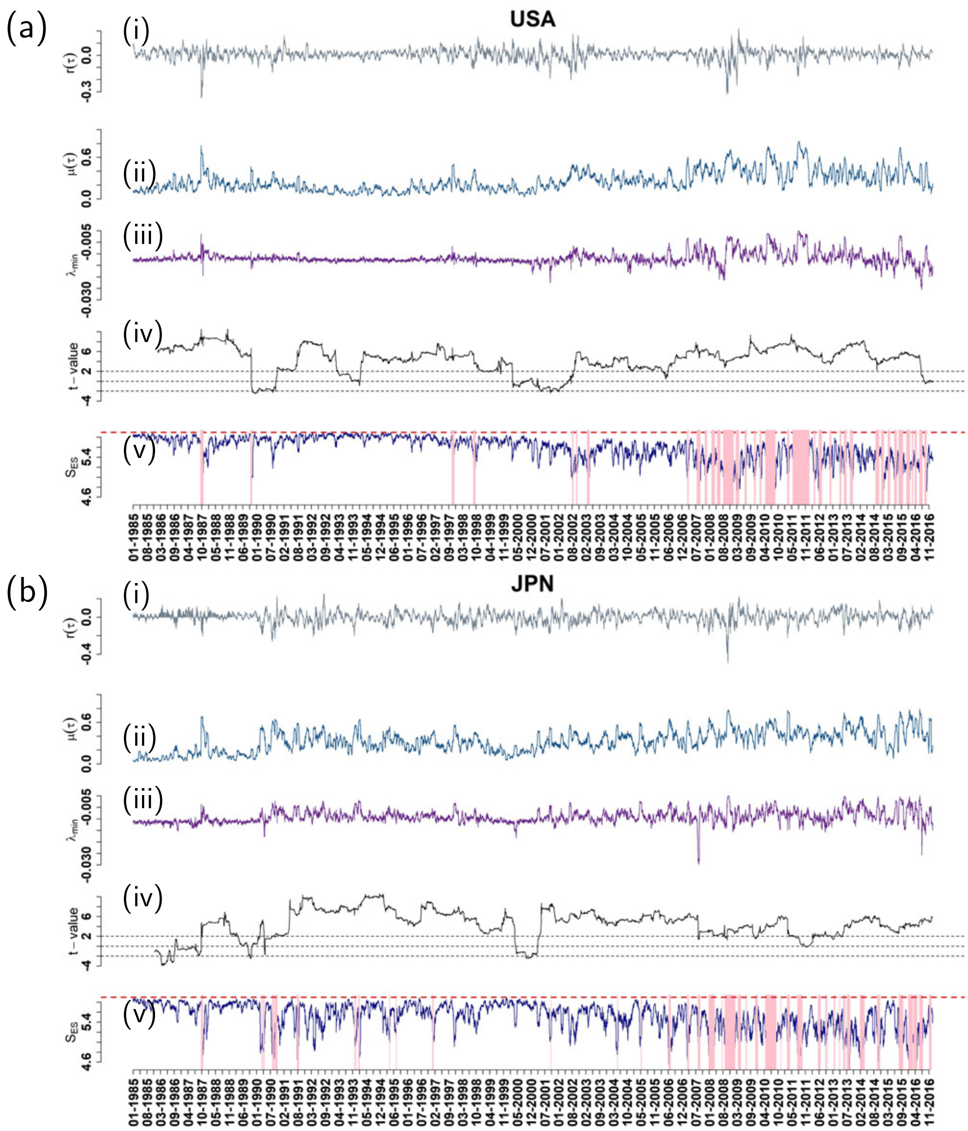

Figure 3. Evolution of the statistical properties of the financial return series for identification of crashes. (a) **USA** and (b) **JPN**: (i) market returns $r(\tau)$, (ii) mean market correlation $\mu(\tau)$, (iii) smallest eigenvalue of the emerging spectrum $\lambda_{\min}$, (iv) t-value of the t-test, which tests if lag-1 smallest eigenvalue $\lambda_{\min}(\tau - 1)$ has statistically significant effect over mean market correlation $\mu(\tau)$, and (v) spectral entropy S_{ES}. The spectral entropy S_{ES} acts as the indicator and reflects the shape of the emerging spectra: high, when the shape resembles the semi-circular Wigner distribution (similar to WOE in red dash) and low, when the shape resembles Lorentzian-like. The S_{ES} is undefined (indicated by pink vertical line) for the epoch that is heavy-tailed Lorentzian-like, which corresponds to a crash (catastrophic instability).

So, the null hypothesis and the alternate hypothesis can be stated mathematically as: $H_0 : \beta_1 = 0 \,\text{vs}\, H_A : \beta_1 \neq 0$.

The t-value for estimated $\hat{\beta}_1$ is calculated as

$$t = \frac{\hat{\beta}_1 - 0}{\text{se}(\hat{\beta}_1)},$$

where se is the standard error in statistics. If $|t - \text{value}| > 2$, we can say that the last day's smallest eigenvalue ($\lambda_{\min}(\tau - 1)$) has statistically significant effect over today's mean correlation ($\mu(\tau)$). The t-value itself signifies the strength of the signal and shown in figures 3(a) and (b). It seems that for most of the time, the correlation is statistically significant at 2σ levels or higher. The only periods when the $\lambda_{\min}(\tau - 1)$ fails to detect with high significance, are the broad periods 1990-91, 2000-02, etc, which act like bubbles/anomalies—the 'Dot-com bubble', 'Housing bubble', etc. Thus, smallest eigenvalue of the emerging spectrum can be effectively used for the characterization of market crashes and as a signal for market bubbles.

We run the kernel density estimate to evaluate the normalized probability distribution function of the eigenvalues around zero (emerging spectrum), and compute the probabilities $\{p_i\}$. We define the *entropy of*

Table 1. List of major crashes and their characterization.

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	Bond market crisis	1994	JPN, USA
4	Dot com bubble	1994–2000	USA, JPN
5	Asian financial crisis	02-07-1997	JPN, USA
6	Russia devalues ruble	07-08-1998	USA
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	IPO facebook debut	18-05-2012	USA, JPN
15	Flash freeze	22-08-2013	USA, JPN
16	Treasury freeze	15-10-2014	USA, JPN
17	Chinese black monday	24-08-2015	USA
18	Brexit	20-06-2016	USA, JPN

emerging spectrum $S_{ES} = -\sum_i p_i \ln p_i$, and use it to characterize the shape of the this spectrum. The shape of the emerging spectra is quantified by the S_{ES}. When the system changes from normal to critical periods, the shape (distribution) of the emerging spectra changes from distorted Wigner semi-circle to heavy-tailed Lorentzian-like shape; the value of S_{ES} changes from low (normal) to high (critical). We take the mean S_{ES} for the WOE of equivalent size N, as benchmark. The spectral entropy S_{ES} reflects the shape of the emerging spectra: high, when the shape resembles semi-circular Wigner distribution (red dash) and low, when the shape is Lorentzian-like. The S_{ES} is undefined (indicated by pink vertical line) for the epoch that is heavy-tailed Lorentzian, which corresponds to a crash (catastrophic instability). The list of the major crashes is given in table 1. These features could shed light on detection of market anomalies/instabilities and designing appropriate market strategies, etc.

2.2. Environmental ozone pollution

We use the data of ozone pollution [73] with daily 8 h maximum ground level ozone pollution from 1 January 2012 to 31 December 2017 ($T = 2192$ days) monitored from 115 locations in the six states of the USA distributed as follow: **IL**–Illinois (24), **IN**–Indiana (01), **NJ**–New Jersey (16), **NY**–New York (26), **OH**–Ohio (03), and **PA**–Pennsylvania (45).

The inter-spatial interpolation process using Bayesian modeling and computational details are given in SI. A few salient points about the computational part are: (i) the spatial model is fitted after doing a Markov chain Monte Carlo calculations (number of MCMC steps = 5,000, with first 1,000 as transient (burn-in) period). (ii) The spatial process is then extracted from the model for all 115 locations. (iii) The median of the spatial process, calculated from the $(5000 - 1000 =)$ 4000 MCMC samples of the spatial process, is reported. (iv) The median spatial process is then interpolated/kriged to 200×200 block grid points, i.e., 40,000 grid points, as plotted in figures 4(a) and (b).

Figures 4(a)–(f) show the interpolated spatial processes, correlation matrices and emerging spectra of daily maximum 8 h average ozone concentration data. The plots of the interpolated spatial processes in 200×200 grid points over six states are shown in figure 4(a) for January 2014 and figure 4(b) for September 2012, respectively. The interpolated values of spatial process for January are seen in figure 4(a) to have lower mean ($\sim 44.90 \pm 1.03$, as shown in the inset), with a very steep gradient of the contour lines that implies there is less spatio-temporal correlation in the time-series for 115 stations; however, for September (see figure 4(b)), the mean is much higher ($\sim 62.96 \pm 0.61$, as shown in inset) with the gradient of the contour lines changing very little that implies high spatio-temporal correlation across stations over extended regions. Figures 4(c) and (d) show the plots of 30 days epoch correlation-matrices from data of 115 ozone stations, confirming the low and high mean correlations in January and September, respectively. Figures 4(e) and (f) show the plots of eigen spectra and emerging spectra (insets) generated by the power mapping with $M = 30$ days and $\epsilon = 0.01$, for January and September, respectively. Figure 4(e) shows the eigen spectrum of the correlation matrix of non-critical (normal) period (January), with the maximum eigenvalue $\lambda_{max} = 17.494$ (not plotted). Inset shows the emerging spectrum generated using power mapping ($\epsilon = 0.01$) is deformed semi-circular, with the smallest eigenvalue $\lambda_{min} = -0.004$. Figure 4(f) shows the eigen spectrum of the correlation matrix of a critical period (September) correlation matrix, with

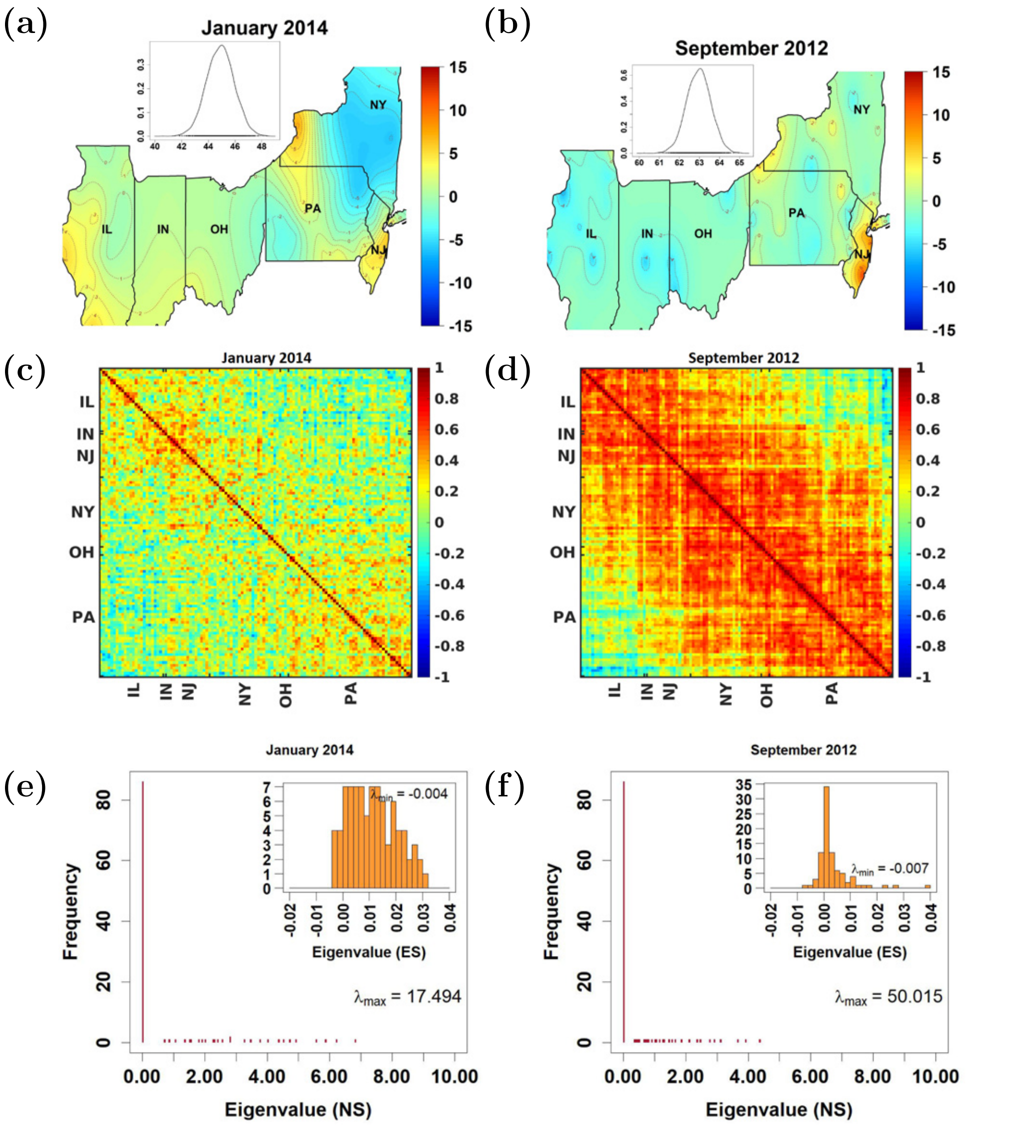

Figure 4. Daily maximum 8 h average ozone concentration data: interpolated spatial processes, and correlation matrices in six states of the USA. Plots of the interpolated spatial process in 200 × 200 grid points over six states, in: (a) January 2014 and (b) September 2012, respectively. In January(a), the interpolated values of spatial process are seen to have lower mean ($\sim$44.90 ± 1.03, in inset), with the gradient of the contour lines very steep that implies that there is less spatio-temporal correlation in the time-series for 115 stations; however, for September (b), the mean is much higher ($\sim$62.96 ± 0.61, in inset) with the gradient of the contour lines changing very little that implies high spatio-temporal correlation across stations over extended regions. (c) And (d) plots of short time-epoch cross-correlation matrices from data of 115 ozone stations, confirming the low and high mean correlations in January and September, respectively. (e) And (f) plots of eigen spectra and emerging spectra generated by the power mapping with $M = 30$ days and $\epsilon = 0.01$, for January and September, respectively. The map and latitude-longitude details are given in SI.

the maximum eigenvalue $\lambda_{\max} = 50.015$ (not plotted). Inset shows the emerging spectrum using power mapping ($\epsilon = 0.01$), which is heavy-tailed Lorentzian-like, with the smallest eigenvalue $\lambda_{\min} = -0.007$.

Figure 5(a) shows the evolution of the emerging spectra from distorted semi-circle (non-critical period) to Lorenztian-like shape (critical period) with time. Figure 5(b) displays the variation of spectral entropy S_{ES}, which reflects the shape of the emerging spectra: high (e.g., January 2014), when the shape resembles Wigner's semi-circle distribution (red dash) and low (e.g., September 2012), when the shape is Lorentzian-like. However, in ozone system, the mean correlation is never too high ($\leqslant 0.35$) and the Lorentzian-like peaks are low with tails less fat (compared to financial markets in figure 3), indicating the fact that the system never approaches critical state!

It is interesting to note that the smallest eigenvalue $\lambda_{\min}$ of the emerging spectra for the environmental ozone neither shows rapid fluctuations nor large changes as compared to market crashes in the financial market. Also, note that the relative changes in the maximum eigenvalue $\lambda_{\max}$ is similar for both the cases, whereas for smallest eigenvalues $\lambda_{\min}$ it is much bigger during the market crash.

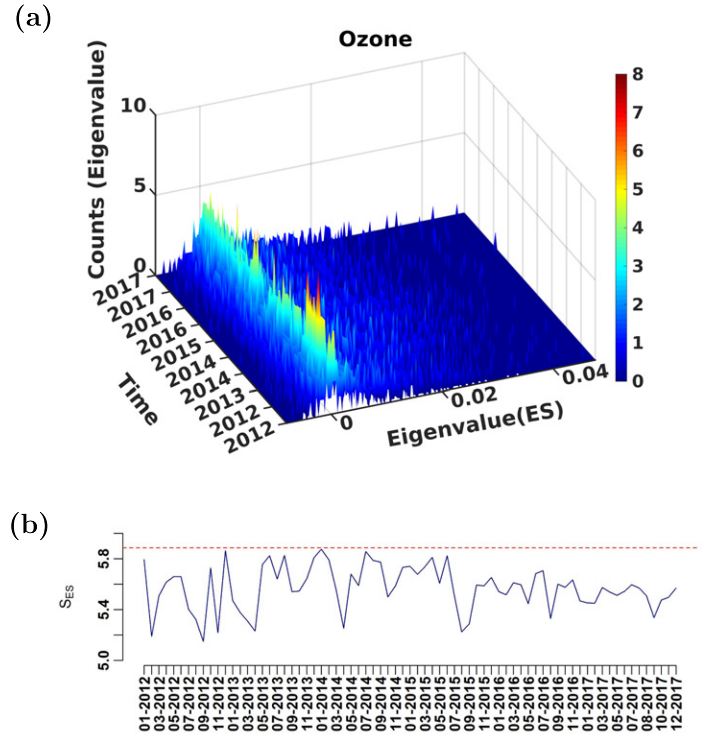

Figure 5. Evolution of emerging spectra and spectral entropy in six states of the USA. (a) Evolution of the emerging spectra from distorted semi-circular (non-critical period) to Lorenztian-like (critical period) with time. (b) The spectral entropy S_{ES} reflects the shape of the emerging spectra: high (e.g., January 2014), when the shape resembles semi-circular Wigner distribution (similar to WOE in red dash) and low (e.g., September 2012), when the shape is Lorentzian-like.

3. Discussions and summary

In summary, our study of the statistical properties of the emerging spectra illustrates for the first time that the shape of the emerging spectrum (captured by the spectral entropy) reflects the instability in a complex system. Also, the smallest eigenvalues of the emerging spectrum also contain significant information about the system's correlation structure. We used the emerging spectral entropy S_{ES}, to quantify the shape of the emerging spectrum. When the system changed from normal to critical periods, the shape (distribution) of the emerging spectra changed from distorted semi-circular to heavy-tailed Lorentzian-like; the value of S_{ES} changed from high (normal) to low (critical). In each case, we took the mean S_{ES} for WOE of equivalent size N, as benchmark. As expected, the spectral entropy S_{ES} never goes above this benchmark. For the financial market, we found that the S_{ES} was undefined (indicated by pink vertical line) for the epoch which corresponds to a crash. For ozone data S_{ES} is always defined. We found that we could retrieve all the major crashes as given in table 1. We also found a few extra instabilities or crashes (the false-positives), for which we could not get any information from internet sources (see table S4 in the supplementary information). In the financial market context, a striking and far reaching result found was that in certain instabilities/crashes the smallest eigenvalue of the emerging spectrum was positively correlated with the largest eigenvalue (and thus with the mean market correlation) rather than just a trivial anti-correlation. We further ran a linear regression model for the mean market cross-correlation $\mu(\tau)$ as function of the time-lagged smallest eigenvalue $\lambda_{\min}(\tau-1)$, and found that the two variables have statistically significant correlation except during bubbles/anomalies, implying that an indicator function exists for bubbles which could signal a precursor to market catastrophe. It may be noted that we are interested in the time evolution of the statistical significance (measured by t-value) of the *correlation* between the variables, μ and $\lambda_{\min}(\tau-1)$, and not the predicted values *per* se. The proposed regression model can be used for short term forecasting of market instability in a situation like the ongoing COVID-19. We would have to develop a thorough back-testing strategy (or out-of-sample testing strategy) for such forecasting models. However, these are beyond the scope of the present paper.

We also observed a lead-lag effect of the crashes across the globe in terms of the behavior of $\lambda_{\min}$ in various markets (to be reported elsewhere). These features could shed light on detection of market anomalies/instabilities and designing appropriate market strategies, etc. In ozone system, the mean correlation is never too high ($\leqslant 0.35$) (see SI) and the Lorentzian-like peaks are low with less fat tails, which allow the S_{ES} to be well defined (compared to financial markets in figure 3), indicating the fact that the system never approaches critical state. But the emerging spectral entropy does have high value (e.g., January 2014), when the shape resembles Wigner's semi-circle distribution (red dash) and low value (e.g., September 2012), when the shape is Lorentzian-like. So, it can indicate periods which have high variability (less mean correlation) and low variability (less mean correlation), in a computationally less intensive way, as compared to Bayesian spatial interpolation methods using MCMC.

In our proposed method, the only two parameters that can be varied are epoch size M and distortion parameter ϵ, which we have studied in details computationally with empirical and surrogate data in references [19, 45]. We found that the S_{ES} is not very sensitive to small values of the distortion parameter ϵ (see figure S7 in the supplementary information). Therefore, our method to identify and characterize critical periods or instabilities is very simple, computationally cheap, and general (without any arbitrary threshold). Further, we may mention that power mapping reduces the noise and spurious signals quite a bit (for large ϵ, as confirmed and reported in [19]). Using different choices of parameters (like window size, shift of windows, distortion parameter) we have observed that the fluctuations still manifest, and so in a sense they are independent of the arbitrariness in the choice of parameters. Using other techniques (see [19, 74]) we have confirmed that the financial markets actually have become very volatile (with increasing number of sizable fluctuations), especially after the year 2002 in USA and 1990 in Japan. As mentioned earlier, in the history of financial crashes, not all cases of sizable fluctuations are recorded as market crashes, even though these correlation matrices bear similar signatures of crashes.

In conclusion, we note that the behavior of the spatial correlations of ozone pollutant have interesting and important properties that may help the policy makers on providing risk assessment decisions. On the other hand, it does not display the catastrophic instabilities as seen in financial markets. The new quantities discussed here allow detection and further characterization in simple numerical terms that was not previously available. Thus, these results are certainly of deep significance for the understanding of critical behavior in complex systems and risk management, but beyond that open a new window to the exploration of other complex systems that display catastrophic instabilities.

Acknowledgments

The authors are grateful to R Chatterjee for taking part in initial stages; AS Chakrabarti, BK Chakrabarti, S Chakravarty, F Leyvraz and S Sadhukhan for their critical inputs. SD is partially supported by the Infosys grant to CMI. HKP is grateful for financial support provided by UNAM-DGAPA and CONACYT Proyecto Fronteras 952. THS acknowledges the support grant by CONACyT through Project FRONTERAS 201. AC, KS and THS acknowledge support from the project UNAM-DGAPA-PAPIIT AG100819. KSB acknowledges support from the Australian Academy of Science AISRF-EMCR Fellowship (2018-19) and the UbiSENSE project Data61, CSIRO Australia.

References

[1] Coumou D and Rahmstorf S 2012 *Nat. Clim. Change* **2** 491
[2] Dale V H *et al* 2001 *BioScience* **51** 723–34
[3] Cai W *et al* 2014 *Nat. Clim. Change* **4** 111
[4] Husain S S, Sharma K, Kukreti V and Chakraborti A 2020 *Phys.* A **540** 123113
[5] Siman-Tov M, Bodas M and Peleg K 2016 *Soc. Sci. Q.* **97** 75–85
[6] Scheffer M 2009 *Critical Transitions in Nature and Society Princeton Studies in Complexity* (Princeton, NJ: Princeton University Press)
[7] Scheffer M *et al* 2012 *Science* **338** 344–8
[8] May R M, Levin S A and Sugihara G 2008 *Nature* **451** 893–5
[9] Sornette D 2004 *Why Stock Markets Crash: Critical Events in Complex Financial Systems* (Princeton, NJ: Princeton University Press)
[10] Cardarelli R, Elekdag S and Lall S 2011 *J. Financ. Stabil.* **7** 78–97
[11] Weiss R A and McMichael A J 2004 *Nat. Med.* **10** S70
[12] Wang W, Tang M, Stanley H E and Braunstein L A 2017 *Rep. Prog. Phys.* **80** 036603
[13] Kleinberg J 2007 *Nature* **449** 287
[14] Kawamura H, Hatano T, Kato N, Biswas S and Chakrabarti B K 2012 *Rev. Mod. Phys.* **84** 839
[15] Flores J, Novaro O and Seligman T 1987 *Nature* **326** 783
[16] Bak P 1996 *How Nature Works: The Science of Self-Organized Criticality* (New York: Copernicus Press)
[17] Albert R and Barabási A L 2002 *Rev. Mod. Phys.* **74** 47
[18] Mehta M L 2004 *Random Matrices* (New York: Academic)
[19] Pharasi H K, Sharma K, Chatterjee R, Chakraborti A, Leyvraz F and Seligman T H 2018 *New J. Phys.* **20** 103041
[20] Goldenfeld N and Kadanoff L P 1999 *Science* **284** 87–9
[21] Forrester P, Snaith N and Verbaarschot J 2003 *J. Phys. A: Math. Gen.* **36** R1
[22] Brody T A, Flores J, French J B, Mello P, Pandey A and Wong S S 1981 *Rev. Mod. Phys.* **53** 385
[23] Beenakker C W 1997 *Rev. Mod. Phys.* **69** 731

[24] Efetov K 1999 *Supersymmetry in Disorder and Chaos* (Cambridge: Cambridge University Press)
[25] Verbaarschot J and Wettig T 2000 *Annu. Rev. Nucl. Part. Sci.* **50** 343–410
[26] Guhr T, Müller-Groeling A and Weidenmüller H A 1998 *Phys. Rep.* **299** 189–425
[27] Bouchaud J P and Potters M 2003 *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management* (Cambridge: Cambridge University Press)
[28] Mantegna R N and Stanley H E 2007 *An Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge: Cambridge University Press)
[29] Sinha S, Chatterjee A, Chakraborti A and Chakrabarti B K 2010 *Econophysics: an Introduction* (New York: Wiley)
[30] Chakraborti A, Muni Toke I, Patriarca M and Abergel F 2011 *Quant. Finance* **11** 991–1012
[31] Chakraborti A, Muni Toke I, Patriarca M and Abergel F 2011 *Quant. Finance* **11** 1013–41
[32] Chakraborti A, Challet D, Chatterjee A, Marsili M, Zhang Y C and Chakrabarti B K 2015 *Phys. Rep.* **552** 1–25
[33] Plerou V, Gopikrishnan P, Rosenow B, Amaral L A N and Stanley H E 1999 *Phys. Rev. Lett.* **83** 1471
[34] Laloux L, Cizeau P, Potters M and Bouchaud J P 2000 *Int. J. Theor. Appl. Finance* **3** 391–7
[35] Gopikrishnan P, Rosenow B, Plerou V and Stanley H E 2001 *Phys. Rev.* E **64** 035106
[36] Guhr T and Kälber B 2003 *J. Phys. A: Math. Gen.* **36** 3009
[37] Schäfer R, Nilsson N F and Guhr T 2010 *Quant. Finance* **10** 107–19
[38] Wirtz T and Guhr T 2014 *J. Phys. A: Math. Theor.* **47** 075004
[39] Chetalova D, Schmitt T A, Schäfer R and Guhr T 2015 *Int. J. Theor. Appl. Finance* **18** 1550012
[40] Schmitt T A, Chetalova D, Schäfer R and Guhr T 2013 *Europhys. Lett.* **103** 58003
[41] Münnix M C, Shimada T, Schäfer R, Leyvraz F, Seligman T H, Guhr T and Stanley H E 2012 *Sci. Rep.* **2** 644
[42] Vinayak, Schäfer R and Seligman T H 2013 *Phys. Rev.* E **88** 032115
[43] Ramos M M M 2018 Caracterización estadśtica de mercados europeos *Master's Thesis* México Universidad Nacional Autónoma de México
[44] González E S O 2018 Mapeo de Guhr–Kaelber aplicado a matrices de correlaciń singulares de dos mercados financieros *Master's Thesis* México Universidad Nacional Autónoma de México
[45] Pharasi H K, Sharma K, Chakraborti A and Seligman T H 2019 *Complex Market Dynamics in the Light of Random Matrix Theory* (Berlin: Springer) pp 13–34
[46] Pharasi H K, Seligman E and Seligman T H 2020 arXiv:2003.07058
[47] Rinn P, Stepanov Y, Peinke J, Guhr T and Schäfer R 2015 *Europhys. Lett.* **110** 68003
[48] Lombardi F, Gómez-Extremera M, Bernaola-Galván P, Vetrivelan R, Saper C B, Scammell T E and Ivanov P C 2020 *J. Neurosci.* **40** 171–90
[49] Lo C C, Amaral L N, Havlin S, Ivanov P C, Penzel T, Peter J H and Stanley H E 2002 *Europhys. Lett.* **57** 625
[50] Lo C C, Chou T, Penzel T, Scammell T E, Strecker R E, Stanley H E and Ivanov P C 2004 *Proc. Natl. Acad. Sci.* **101** 17545–8
[51] Lo C C, Bartsch R P and Ivanov P C 2013 *Europhys. Lett.* **102** 10008
[52] Wang S W, Bitbol A F and Wingreen N S 2019 *PLoS Comput. Biol.* **15** e1007010
[53] De Domenico M and Biamonte J 2016 *Phys. Rev.* X **6** 041062
[54] Vemuri V 1978 *Modeling of Complex Systems: An Introduction* (New York: Academic)
[55] Gell-Mann M 1995 *Complexity* **1** 16–9
[56] Bar-Yam Y 2002 *Encyclopedia of Life Support Systems (EOLSS)* (Oxford: EOLSS Publishers)
[57] Shiller R J 1981 *Am. Econ. Rev.* **71** 421–36
[58] Buchanan M 2000 *Ubiquity: Why Catastrophes Happen* (New York: Three Rivers Press)
[59] Sorkin A R 2009 *Too Big to Fail: The inside Story of How Wall Street and Washington Fought to Save the FinancialSystem–and Themselves* (New York: Viking)
[60] Acemoglu D, Ozdaglar A and Tahbaz-Salehi A 2015 *Am. Econ. Rev.* **105** 564–608
[61] Hill M K 2010 *Understanding Environmental Pollution* (Cambridge: Cambridge University Press)
[62] Gilbert R O 1987 *Statistical Methods for Environmental Pollution Monitoring* (New York: Wiley)
[63] Brunekreef B and Holgate S T 2002 *Lancet* **360** 1233–42
[64] Godish T, Davis W T and Fu J S 2014 *Air Quality* (Boca Raton, FL: CRC Press)
[65] Nakicenovic N, Alcamo J, Grubler A, Riahi K, Roehrl R, Rogner H H and Victor N 2000 *Special Report on Emissions Scenarios (SRES), a Special Report of Working Group III of the Intergovernmental Panel on Climate Change* (Cambridge: Cambridge University Press)
[66] Stocker T 2014 *Climate Change 2013: The Physical Science Basis: Working Group I Contribution to the Fifth assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge: Cambridge University Press)
[67] Murazaki K and Hess P 2006 *J. Geophys. Res. Atmos.* **111** D05301
[68] Onnela J P, Chakraborti A, Kaski K, Kertesz J and Kanto A 2003 *Phys. Rev.* E **68** 056110
[69] Kulkarni V and Deo N 2007 *Eur. Phys. J.* B **60** 101–9
[70] Kumar S and Deo N 2012 *Phys. Rev.* E **86** 026101
[71] Damgaard P H, Splittorff K and Verbaarschot J J M 2010 *Phys. Rev. Lett.* **105** 162002
[72] 2017 Yahoo Finance Database (https://finance.yahoo.co.jp/)
[73] 2019 Usepa (https://www.epa.gov/outdoor-air-quality-data/)
[74] Chakraborti A, Hrishidev, Sharma K and Pharasi H K 2019 arXiv:1910.06242