Advances in Complex Systems
Vol. 25, No. 4 (2022) 2250006 (23 pages)

DOI: 10.1142/S0219525922500060

THE BACKBONE OF THE FINANCIAL INTERACTION NETWORK USING A MAXIMUM ENTROPY DISTRIBUTION

MAURICIO A. VALLE

Facultad de Economía y Negocios,
Universidad Finis Terrae, Santiago, Chile
mvalle@uft.cl

FELIPE URBINA

Vicerrectoría de Investigación,
Universidad Mayor, Santiago, Chile
furbinaparada@gmail.com

Received 27 February 2022
Revised 21 July 2022
Accepted 22 July 2022
Published 19 October 2022

We modeled the stocks of the financial system as a set of many interacting like spins derived from binary daily returns. From the empirical observation of these returns, we used a Boltzmann machine to infer a distribution of states equivalent to a maximum entropy distribution. This model describes the interaction couplings between each stock pair in the system, which can be considered a complete network with $N(N-1)/2$ couplings. We then engage in a coupling removal process to find a critical graph that can describe the observed states of the system with the minimum number of edges. We interpret the critical graph as the backbone of the system, and it allows us to evaluate the importance of markets in their relation to others in the system. We also found that the structure of this critical graph is highly variable over time and appears to be dependent on the level of entropy of the system.

Keywords: Pairwise Ising model; financial network; transition; Boltzmann learning; Boltzmann machine; maximum entropy principle.

1. Introduction

Using the idea that the stocks markets could be represented as spins magnets, we show that it is possible to characterize the critical structure of a financial market by identifying a phase shift in the recovery of system states. The view of stocks as random magnets provides a paradigm to study the interactions between various components of the financial market. In fact, since the work of [1], it has been possible to show that the dynamics of the correlations between stock price changes of different companies is satisfactorily explained by assuming a behavior similar to that of spin glasses. The main idea under this paradigm is that the mutual influences between each system component (the stocks in this case) can represent the observed macroscopic structures (market behavior). There are no prior assumptions regarding rules that each component must follow in the relationship with another in this process. Examples of this type of energy-based model have already been studied in finance [2–5] where couplings, fields and the energy of the system describe the behavior of the system. From the machine learning perspective, the Boltzmann machines [6, 7] offer an interesting possibility to the problem of inference of the parameters of energy-based distributions, which is equivalent to an inverse Ising inference process under the maximum entropy principle [8–11]. The inference process of these interactions is based on finding a maximum entropy model describing the state space of the complex system [12]. The first step is to find a model that satisfactorily recovers the correlations between all the stock markets. The model, as indicated above, is described by the couplings and field energies. In this sense, we can think of the system as a fully connected graph of interactions or a couplings network. Using this coupling network, it is possible to analyze what modifications can be made to this model and observe the correlations' recoverability between spins. To be more specific, we find the simplest coupling network so that it recovers the observed correlations of the financial system in a given period. We call this network the critical graph. For this purpose, we use as a measure of comparison between the observed distribution of p_{obs} states and the distribution achieved by the simplest model, with the Kullback–Leibler divergence (KL). Thus, we can build phase diagrams to study the extent to which the system can maintain the empirical moments of the distribution before they get lost.

This work aims to establish a methodology based on discovering maximum entropy principle (MEP) that allows us to analyze the interactions existing between the components of a system. From these interactions, we undertake a simplification process of the model that can reproduce the system's behavior, discarding all those interactions that do not play an essential role in this behavior. The final result is a simplified model of the original that can still correctly describe the system's statistical properties. This simplified model is called the critical network because if we try to simplify the model even further, it drastically loses the ability to reproduce its properties. The results can provide descriptive but valuable information regarding which system components have more influence than others.

Section 2 presents the inference process used to find the MEP-based distribution of a system in different periods and the critical graph problem. In Sec. 3, we introduce the results of the inference process and analysis of the critical graphs in Sec. 3, and finally in Sec. 5 the conclusions are presented.

2. Model and Problem Formulation

2.1. *The maximum entropy model*

The state of the system with N assets can be represented by the state vector $\mathrm{s} = (s_1, \ldots, s_N)$ similar to the Ising model, where $s_i = \pm 1$ represents whether the return of asset i is positive or negative orientation. The return of any asset i can be calculated as the difference in natural logarithm of the asset's value between time t and $t-1$ as $r_i(t) = \log(v_i(t)) - \log(v_i(t-1))$, thus, the orientation of the spin i is $s_i = +1$ if $r_i \geq 0$, and $s_i = -1$ otherwise.

The idea is to find a model that fits the observed state distribution $p_{\text{obs}}(\text{s})$. This is not an easy task since, due to the high dimensionality of the space of possible states, it would be impossible in practice to have a sample with sufficient data that includes all of that space. Another approach is to fit a parametric model with several parameters much smaller than the dimensionality of the state space.

An alternative model is the maximum entropy model, in which we consider the pairwise connections between assets in the system using the Gibbs distribution as

$$p_{\text{Ising}}(\text{s}) \sim \left\{ -\sum_i h_i s_i - \frac{1}{2} \sum_{i<j} J_{ij} s_i s_j \right\}, \tag{1}$$

where h_i represents the external fields, and J_{ij} coupling parameters. These two sets of parameters must be such that the resulting distribution $p_{\text{Ising}}(\mathbf{s})$ has the same observed mean orientation of the spins and the same correlation structure,

$$\langle s_i \rangle_{\text{Ising}} = \langle s_i \rangle_{\text{obs}}, \tag{2a}$$

$$\langle s_i s_j \rangle_{\text{Ising}} = \langle s_i s_j \rangle_{\text{obs}}, \tag{2b}$$

where the symbols $\langle * \rangle$ indicate the average over the quantities. To find the parameters $\{\mathbf{J}, \mathbf{h}\}$ we can use a potentially exact approximation with Boltzmann Learning [13]. The Boltzmann Learning is an iterative algorithm in which the couplings and fields are adjusted from an initial guess. Under these parameters, the distribution of the system states is calculated and compared with the observed distribution. According to this difference, the parameters are adjusted in such a way that Eqs. (2a) and (2b) are fulfilled. The adjustments of the parameters in each iteration are made according to

$$\delta h_i = \nu(\langle s_i \rangle_{\text{obs}} - \langle s_i \rangle_{\text{Ising}}), \tag{3a}$$

$$\delta J_{ij} = \nu\big(\langle s_i s_j \rangle_{\text{obs}} - \langle s_i s_j \rangle_{\text{Ising}}\big). \tag{3b}$$

The ν is the learning constant, decreasing as the algorithm converges. Under a sufficient number of iterations and with a sufficient number of Monte Carlo steps to find the distribution of the states, it is assured, in principle, that the process converges to the exact values of the parameters; however, this procedure suggests that it is a slow algorithm [14].

2.2. *Critical graph problem formulation*

The Boltzmann machine is a complete signed network, in which the weight of each of the links between nodes (spins) is the value of the couplings J_{ij}. The problem is to find a network (not necessarily connected), which is a subset of the original network that manages to recover the original information of the distribution of states of the system.

The problem can be defined as follows: The Boltzmann machine is a complete undirected graph $G = (V, J)$ where V is the set of nodes representing the spins of the system (in our case, the stock market). The cardinality of V is $|V| = N$. The set J contains the edges between each of the spins in V and is equivalent to the weight matrix of G: the inferred coupling values. Self-loops are not allowed, so $J_{ii} = 0$, $\forall i$. For each period, we have a graph G from which we wish to extract a graph $\tilde{G} = (V, \tilde{J})$ where $\tilde{J} \subset J$, i.e. $\tilde{G}$ maintains the same nodes of the original Boltzmann machine, but with a reduced number of connections. The question of interest is what is the graph $\tilde{G}$, call it the *critical graph*, that preserves the information of $p_{\mathrm{obs}}(\mathbf{s})$ before there is a drastic loss of this information? To do this, it is necessary to estimate the distribution $p_{\tilde{G}}(\mathbf{s})$, which corresponds to the distribution estimated using the network $\tilde{G}$. To evaluate the information loss, we compute the Kullback–Leiber (D_{KL}) divergence, between the observed distribution $p_{\mathrm{obs}}(\mathbf{s})$ and the distribution inferred from $\tilde{G}$, i.e. $D_{\mathrm{KL}}(p_{\mathrm{obs}}||p_{\tilde{G}})$. If this divergence is null, then the distribution from $\tilde{G}$ fully recovers all the information of the $p_{\mathrm{obs}}(\mathbf{s})$, otherwise, the graph $\tilde{G}$ is not sufficient to recover the information. Evidently, since $\tilde{G}$ is a subset of the original Boltzmann network, then it must be satisfied that $D_{\mathrm{KL}}(p_{\mathrm{obs}}||p_{\tilde{G}}) > D_{\mathrm{KL}}(p_{\mathrm{obs}}||p_{\mathrm{Ising}}(\mathbf{s}))$.

To compute the KL divergences, we used a Metropolis-Hasting process, keeping the information of the fields of each spin unchanged. That is, for each distinct set of $\tilde{G}$ graphs, the external h_i fields of the period are kept constant in the simulation process. Now, it is worth asking how we find the graph $\tilde{G}$. We must remember that each coupling J_{ij} can be understood as a ferromagnetic or antiferromagnetic link between two $i - j$ spins and, according to the theory of spins glass, its value comes from a Gaussian distribution. In this sense, the probability of finding a very high coupling magnitude (with positive or negative signs) is improbable. On the other hand, such high magnitude couplings are theoretically the most influential on the system's behavior. This is because the system's energy is dependent on the couplings. In turn, on the probability, so high magnitude couplings will have a greater influence in driving the system towards a particular state. Thus, the simplest procedure to evaluate the ability of the graph to retain information is to order all the couplings from smallest to largest in terms of their magnitude. Then iteratively eliminate those of smaller ones. This is equivalent to discarding links from the system that are more likely but less influential.

The process of eliminating links from smallest to largest stops until $\tilde{G}$ is a wholly disconnected graph. This last graph is equivalent to an independent model with no interactions in which only the fields influence the state of each spin. At this point, we have all KL divergences, from the original graph model of the Boltzmann machine to the independent model. We select the graph as critical as the one just before there is an increase of KL distance.

3. Results

3.1. *Data*

We analyze the stock market by selecting indices of six European countries: The FTSE index for the London stock market (UK), CAC for France, DAX for Germany, IBEX for Spain, FTSEMIB for Milan, and AEX for Amsterdam. For each country index, we took daily prices from August 7, 2002, to April 22, 2021. This period covers approximately 4876 market trading days. This extended analysis period covers the Subprime crisis between 2007 and 2010, and recently the COVID19 outbreak beginning in December 2019.

To study the dynamics of the pairwise distribution's parameters over time, we split the sample into 40-time segments of 250 days, equivalent to one year of market trading days, with an overlap of 125 days between time segments. In this way, we carried out 40 inference processes with the Boltzmann Learning.

For the learning process over each time segment, we use 250 days of returns from each country index and a learning rate of 0.95 with a decay of 0.0004, and 25,000 iterations to compute the approximations Eq. (3). In each of these iterations we use 1500 Monte Carlo sampling steps to compute $\langle s_i \rangle_{\text{Ising}}$ and $\langle s_i s_j \rangle_{\text{Ising}}$.

3.2. *Consistency*

As a measure of consistency between the parameters of the observed $p_{\text{obs}}(\mathbf{s})$ state distribution and the maximum entropy distribution $p_{\text{Ising}}(\mathbf{s})$, we compare the mean of the orientation of the spins $\langle s_i \rangle_{\text{Ising}}$ with the observed ones and similarly, the covariances $C_{ij} = \langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle$. The recovered parameters of each in the 40 time segments are calculated based on a Metropolis-Hasting dynamics with 1500 iterations repeated 1000 times each one. This is equivalent to providing a measure of compliance with Eqs. 3(a) and 3(b). Figure 1 shows a scatterplot of the recovered versus observed pooled parameters from all time segments.

For the mean orientation of the spins, we find that the mean Root Mean Squared Error (RMSE) was 0.04339 and a standard deviation of 0.03159. The maximum value of RMSE was 0.1449, corresponding to the sample of July 15, 2015, to June 29, 2016. Except for this value, all other RMSEs are less than 0.010. The mean of Pearson correlations between retrieved and observed orientations were 0.9131 and a standard deviation of 0.01010. The minimum correlation value was 0.4604 in the sample of the 25 from February 20, 2013, to February 05, 2014. The sample of September 14, 2011, to August 29, 2012, present a correlation of 0.6087. All the others have a value higher than 0.90.

For the covariance, we find that mean of RMSE was 0.0269 with a standard deviation of 0.0103. The worst RMSE is on the sample of September 19, 2010, to September 14, 2011, with a value of 0.0543. The mean of Pearson correlations between observed covariances and retrieved ones was 0.9772, with a standard deviation of 0.0137. The lower correlation was on the sample May 30, 2018, to May 15, 2019.

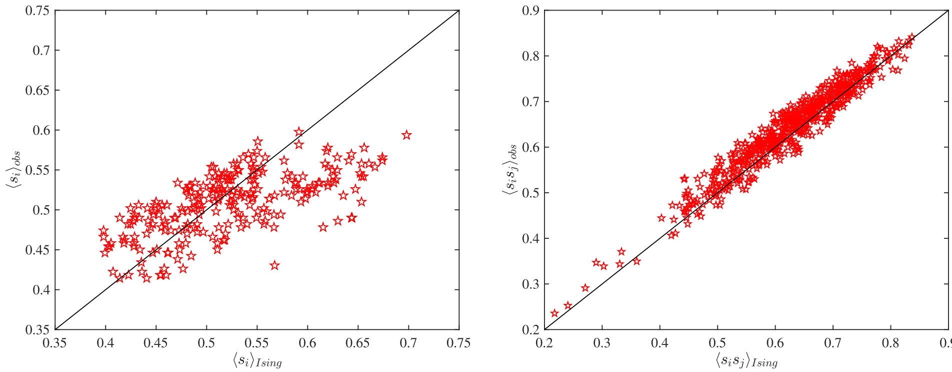


Fig. 1. Left plot — Scatterplot of the mean orientations of spins $\langle s_i \rangle$ between observed distributions and Ising distribution. Right plot — Scatterplot of the covariances $\langle s_i s_j \rangle$ between observed distributions $p_{\text{obs}}(\text{s})$ and Ising distribution $p_{\text{Ising}}(\text{s})$.

In addition, we conduct several simulations to investigate whether the pairwise model corresponding to the MEP is able to describe better the distribution of observed system states versus an independent model in which there are no couplings. The orientations of the spins (market returns) are driven only as a function of the fields, i.e. there is no mutual influence between the assets. For each of the 40 time segments, we perform 100 simulations to compute the KL divergence between the observed distribution of states $p_{\text{obs}}(\mathbf{s})$, and the maximum entropy distribution $p_{\text{Ising}}(\mathbf{s})$, i.e. $D_{\text{KL}}(p_{\text{obs}}||p_{\text{Ising}})$. In parallel, we compute $D_{\text{KL}}(p_{\text{obs}}||p_1)$ where p_1 is the distribution resulting from an independent model in which spins are independent of each other and there is no interaction between them [15]. As a summary of the analysis, we can state that for all time segments, $D_{\text{KL}}(p_{\text{obs}}|p_{\text{Ising}}) < D_{\text{KL}}(p_{\text{obs}}|p_1)$, i.e. the pairwise model description is better than the independent model. The overall mean of $D_{\text{KL}}(p_{\text{obs}}|p_1)$ was 2.09 ± 0.043, while that of $D_{\text{KL}}(p_{\text{obs}}||p_{\text{Ising}})$ was 1.89 ± 0.038. These results indicate that the pairwise model does a better job than an independent model in describing the behavior of the system.

3.3. *Estimated parameters of maximum entropy distribution* p_{Ising}

This section shows the results of the parameters inferred with Boltzmann learning over the 40 samples. Since we do not impose machine learning restrictions and have $N = 6$ country market indices, the system will have 15 interaction couplings between each pair of i, j spins. Figure 2 presents the distribution of all couplings inferred by the Boltzmann machine, as well as the fields in each of the 40 time segments. There are $15 \times 40 = 600$ couplings and $6 \times 40 = 240$ fields. It is interesting to note that similar to what was found by [3–5], the distribution of couplings closely resembles a Gaussian distribution. Throughout all the periods the system is predominantly

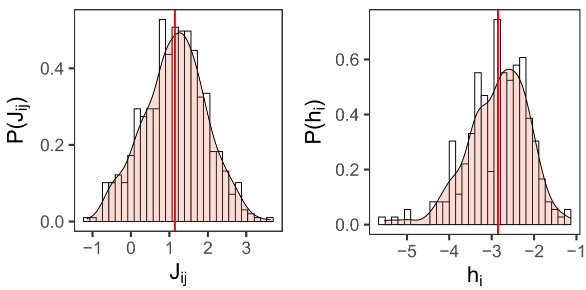


Fig. 2. (Color online) Histograms and densities of couplings J_{ij} (left) and fields h_i (right) for all time segments. Red lines indicate the mean. Red lines indicate $\langle J_{ij} \rangle$ and $\langle h_i \rangle$ for couplings and filed, respectively.

ferromagnetic ($\langle J_{ij} \rangle = 1.143$) and external fields with antiferromagnetic influence ($\langle h_i \rangle = -2.853$). From the point of view of the Spin Glasses theory [16], we can consider that the bonds across spins are quenched, i.e. they are fixed, at least for the time that the period lasts. This is an important assumption, which allows us later to carry out the threshold analysis and determine the critical couplings of the system.

It is of interest to observe the mean of the couplings $\langle J_{ij} \rangle$ to find the ferromagnetism level. Figure 3(Left) shows the dynamics of the couplings means over the different periods. The same graph shows the average of the correlations between spins. Although correlations and couplings are not the same, there is some proportionality between the two observables, which is expected [4, 5]. A positive coupling $J_{ij} \geq 0$ implies a tendency for an alignment in the same direction of the spins, so in this case, the pairwise connection would be expected to be positive $s_i s_j \geq 0$. On the contrary, it is expected a negative correlation between spins with antiferromagnetic couplings $J_{ij} < 0$. It is worth noting that the highest mean values of couplings are

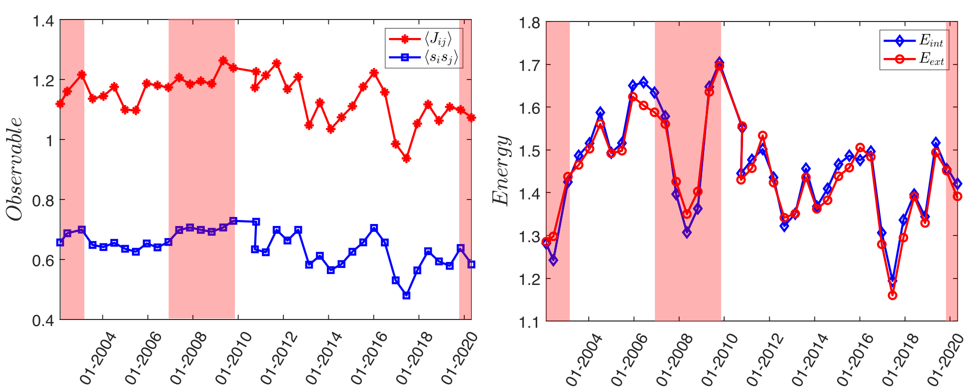


Fig. 3. (Left) The mean of estimated couplings $\langle J_{ij} \rangle$ and pairwise connections between spins $\langle s_i s_j \rangle$ for each time segment. Ljung-Box tests were performed on both data series with non-overlapping time segments. In all cases, the null hypothesis is not rejected, and the data are not correlated, so they can be considered a random process. (Right) Internal h_{int} and external h_{ext} energies by periods.

found in the period from 2008 to 2010, the period in which the subprime crisis is found, as well as the highest level of correlation between spins [2]. This behavior is also related to the synchronization phenomenon in which the correlation of returns between assets tends to increase in periods of financial turmoil [17, 18].

To compare coupling and field estimation, we carried out an estimation using a Naive Mean-Field approximation (see Appendix A) which offers a competitive performance relative to parameter inference with the Boltzmann machine. Figure 4 shows the distribution of KL distances (D_{KL}) between the observed distribution $p_{\mathrm{obs}}(\mathbf{s})$ and the distribution simulated using the inferred Mean-Field parameter set $\{\mathbf{J}_{\mathrm{MF}}, \mathbf{h}_{\mathrm{MF}}\}$ and the inferred parameter set with the Boltzmann machine $\{\mathbf{J}_{\mathrm{Boltz}}, \mathbf{h}_{\mathrm{Boltz}}\}$. Similar to the results of [14], in small sets of neurons (spins) in the network, alternative methods such as Mean-Field seem to do a good job in the process of describing the state distribution of the system.

The energy of the system or Hamiltonian is defined as [19]

$$\mathcal{H}(\mathbf{s}) = -\sum_i h_i s_i - \frac{1}{2}\sum_{i<j} J_{ij} s_i s_j. \tag{4}$$

According to Eq. (1) this energy defines the state probability of the system. The Hamiltonian can be decomposed into two parts: one due to the contribution of the system itself through interactions, and the other due to the external magnetic field, which represents external influences that *bias* the system states in one direction or another. Similarly to [3] for the internal energy due to interactions we can write it as $h_{\mathrm{int}} = \frac{1}{2}\sum_j J_{ij}\langle s_i\rangle$, while that due to external fields as $h_{\mathrm{ext}} = h_i\langle s_i\rangle$.

Figure 3(right) shows a comparison of the energy of interactions and fields over the 40 periods. In most periods, we can see that the system's energy due to

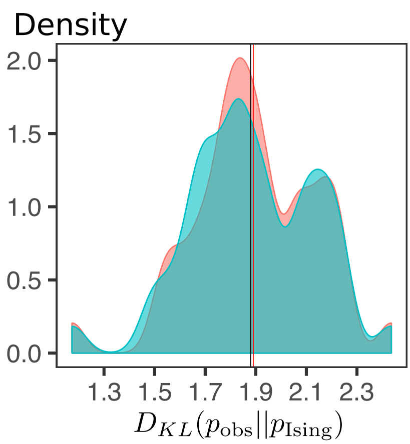
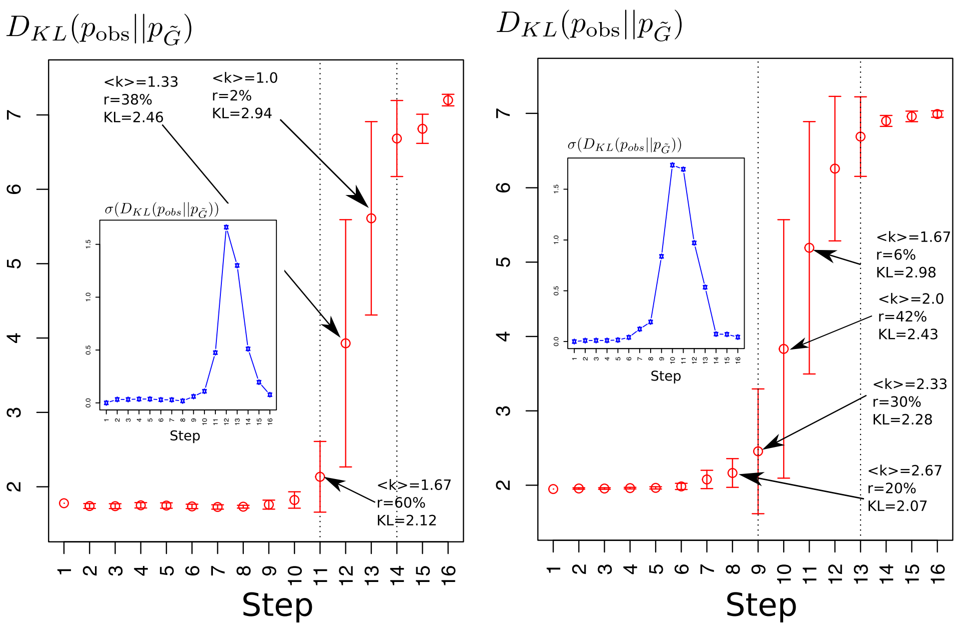


Fig. 4. (Color online) Densities of KL divergences achieved from 30 simulations of each time segment under Mean-field estimation (green curve) and Boltzmann Machine (red curve). Means are 1.89 an 1.88 for Boltzmann (red line) and Mean-Field (black line) parameters, respectively.

interactions and external agents is quite similar. However, only in 11 periods is it true that $h_{\text{ext}} > h_{\text{int}}$, i.e. external agents are predominant over the system states. Interestingly, this occurs during the periods coinciding with the Subprime crisis. This result is in line with what was found by [4] in that the energy due to external influences exceeds that of the system's interactions in times of financial turmoil. However, as a cautionary note, it should be noted that from a statistical point of view, the difference between these influences is not statistically significant, as found in a series of bootstrap hypothesis testings carried out in each period.

3.4. *Unfolding graphs transition*

As an illustrative example, we choose two time segments. For each one, we have a network $G(V, J)$ described in Sec. 2.2. An iterative process of removing edges from smallest to largest is started. For each iteration, a Metropolis-Hasting simulation process is carried out to evaluate the KL divergence between the observed state distribution and the one recovered from the network $\tilde{G} = (V, \tilde{J})$. This process is repeated 50 times for the same time segment. Figure 5(a) shows the KL divergence



(a) Sample from Feb 5, 2014 to Jan 21, 2015 (b) Sample from Nov 6, 2019 to Oct 23, 2020

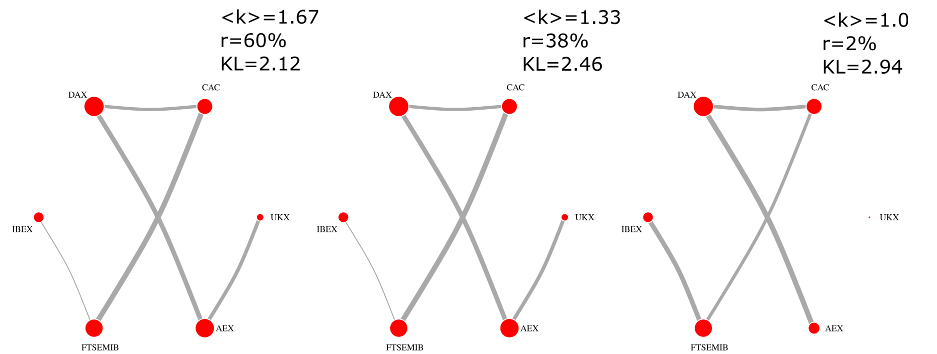
Fig. 5. The KL path from the original Boltzmann machine graph to an independent model of a graph with no edges. Dotted lines indicate the zone of unstable conditions. Arrows indicate some points of the transition zone. Also, the mean degree $\langle k \rangle$ of the critical graph, the fraction of times the graph r occurs, and the average of divergence are shown. Subgraphs (a) and (b) indicate the corresponding time segment. The insets represent the standard deviation of the KL divergence in every step.

$D_{\mathrm{KL}}(p_{\mathrm{obs}}||p_{\tilde{G}})$ achieved versus the removing-coupling iteration for time-segment number 12 corresponding to dates between February 5, 2014 and January 21, 2015. Some comments can be made from this example: First, the KL divergence remains virtually constant as the graph's edges are removed. This is expected because the removing process starts with smaller couplings, i.e. interactions with little influence between the assets. However, it is interesting to note that the KL divergence remains stable even after removing between 9 and 11 edges, representing between 60 and 73.3% of all the edges of the complete network (remember that the complete network has 15 edges). Second, there is a point at which the following removal of edges produces a sudden increase of KL divergence. This phenomenon occurs in step 11, and is exacerbated in steps 12 and 13, in which an increase in the standard deviation of the KL divergence is also observed. This behavior suggests a phase change in which the graph $\tilde{G} = (V, \tilde{J})$ loses its full ability to recover the state distribution of the system. After that, the following edge removal does nothing but keep increasing KL divergence. This behavior indicates that it is possible to disentangle a network $\tilde{G} = (V, \tilde{J})$ with a minimum number of interactions that manage to recover the original distribution (or at least in part) before the graph loses the ability to recover the observed distribution. Call this the critical graph.

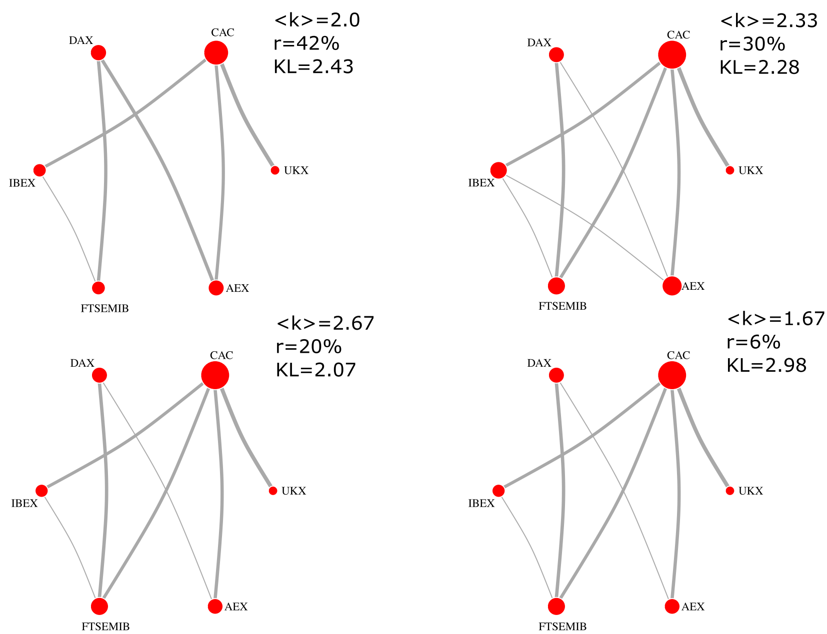
For the time segment under analysis, we identify step 11 as the moment of a critical graph: At this point, the KL Divergence and its standard deviation increase dramatically. Figure 6(a) shows three graphs. The first one corresponds to the critical graph at step 11, but also two more graphs at steps 12 and 13 in the transition zone, where there is a high level of variability in the KL divergence. Looking at these graphs, some patterns can be identified. For example, in all cases, certain connections survive. For example, IBEX is always connected to FTSEMIB, as well as the CAC-DAX-AEX trio. This result suggests that the critical graph are part of the backbone of the system. In other words, it is the minimum possible network from which the orientations and correlations between spins can be recovered in an acceptable manner. In this case, the critical network is characterized by a mean degree of 1.67, compared to the full graph mean degree of $N - 1 = 5$.

Figure 5(b) shows the transition in a more recent period (from November 6, 2019 to October 23, 2020). It suggests that the beginning of the transition occurs at step 9. At this point, there is an evident increase in KL divergence and its variability. The graphs associated at different steps of edge removal on the transition zone can be found in Fig. 6(b). For this case, the graph structure is completely different. The level of connectivity of the network is higher than in the previous example. The mean degree for the critical graph is 2.33. These results suggest that for this period, the pairs UKX-CAC, DAX-AEX, DAX-FTSMIB, the CAC-IBEX-AEX and CAC-IBEX-FTSMIB triangles and fundamental structures in the system's behavior.

It should be noted that the transition does not occur abruptly but rather corresponds to a gradual process. Consequently, it is not possible to determine a single critical network. The point at which the removal of edges stops represents a trade-off between the KL divergence that the researcher is willing to increase, versus the



(a) Sample from Feb 5, to Jan 21, 2015



(b) Sample from Nov 6, 2019 to Oct 23, 2020

Fig. 6. The critical graphs $\tilde{G} = (V, \tilde{J})$ achieved before the transition to the destruction of the coupling network in two different time segments. The first one in (a) from January 5, 2014 to January 21, 2015, and the second one in (b) from November 6, 2019 to October 23, 2020, covering COVID19 crisis. The mean degree $\langle k \rangle$, the fraction of times the network r occurs, and the average of divergence $D_{\mathrm{KL}}(p_{\mathrm{obs}}||p_{\tilde{G}})$ are shown for each graph.

simplicity of the desired graph in terms of the number of edges. The point where the standard deviation is maximum is a good indicator of the transition; however, a critical graph is not desirable just at this point because the KL divergence is already too high; therefore, a graph before that transition, where the KL divergence remains low, and a sign of increased variability is observed, would be advisable. A more objective and practical choice to stop the edge removal is to determine a threshold choosing a percentile of the empirical distribution of the simulations of the KL divergences. Specifically, we suggest observing the empirical distribution of KL divergences differences $\Delta D_{\mathrm{KL}} = D^{t}_{\mathrm{KL}} - D^{t-1}_{\mathrm{KL}}$ between step t and $t-1$ and then determining a threshold based on a percentile of this distribution. In our case, we have taken the 70th percentile. Note that the higher the percentile, the higher the KL divergence level to find the critical network and the higher the number of edges. On the other hand, with a lower percentile, we would be stricter, and we would have a critical network with low KL divergence but with more edges.

3.5. *Critical graphs changes across time*

Critical networks $\tilde{G}$ are not fixed over time. The interactions and external fields are dynamic and change as new continental and global events unfold. Thus, some interactions between pairs of markets that would be considered the *backbone* of the network are no longer a part of it, and now some other bonds perform the function of concentrating most of the market's information. Appendix B shows the evolution of critical networks over the 40-time segments. Since these graphs represent a backbone structure of the system, it is interesting to recognize: (1) which nodes (markets) are the most connected to other nodes via couplings J_{ij}, and (2) changes on the topology of the critical graph. In the first case, we measure the strength st_i of each node, where $st_i = \sum_{\forall k \neq i} J_{ik}$. A node (market) with higher strength than another indicates more importance because it is tied with higher interaction strength with the rest of the others. As the couplings describe the intensity in which two market agents interact with each other, a market with greater strength has greater potential to influence and be influenced by other markets than one with less strength. In the second case, we can evaluate the robustness of the critical network topology by calculating: (a) the number of edges that are in common of two successive periods [20] at time t and $t-1$ as $S(t) = |\tilde{J}(t) \cap \tilde{J}(t-1)|$, where $\tilde{J}(t)$ is the set of couplings of the critical network $\tilde{G}$ at time segment t, (b) calculating the number of edges that disappear from one period to the next as $D(t) = |\tilde{J}(t-1) \notin \tilde{J}(t)|$ and, (c) calculating the number of new edges born at time t as $B(t) = |\tilde{J}(t) \notin \tilde{J}(t-1)|$. If we divide these quantities by the number of edges at time t, i.e. $|\tilde{J}(t)|$, we have the instantaneous rates of survivals, deaths and birth of edges, and they measure the stability of couplings in critical networks. Note that $S(t) + D(t) = |\tilde{J}(t)|$. When $D(t) > B(t)$ or when $D(t) < B(t)$, then the network is left with fewer bonds that make the skeleton of the network simpler, while in the second case, the bonds between nodes increase, increasing also the level of critical interdependencies between markets. When $D(t) = B(t)$ there is also a reconfiguration of the critical graph; however, the number of bonds at $t-1$ and t remains unchanged, i.e. the number of critical interdependencies does not vary.

The upper part of Fig. 7 shows the strength st_i of each country market in each of the periods. Despite the variability of the strength of each one, it can be seen that the

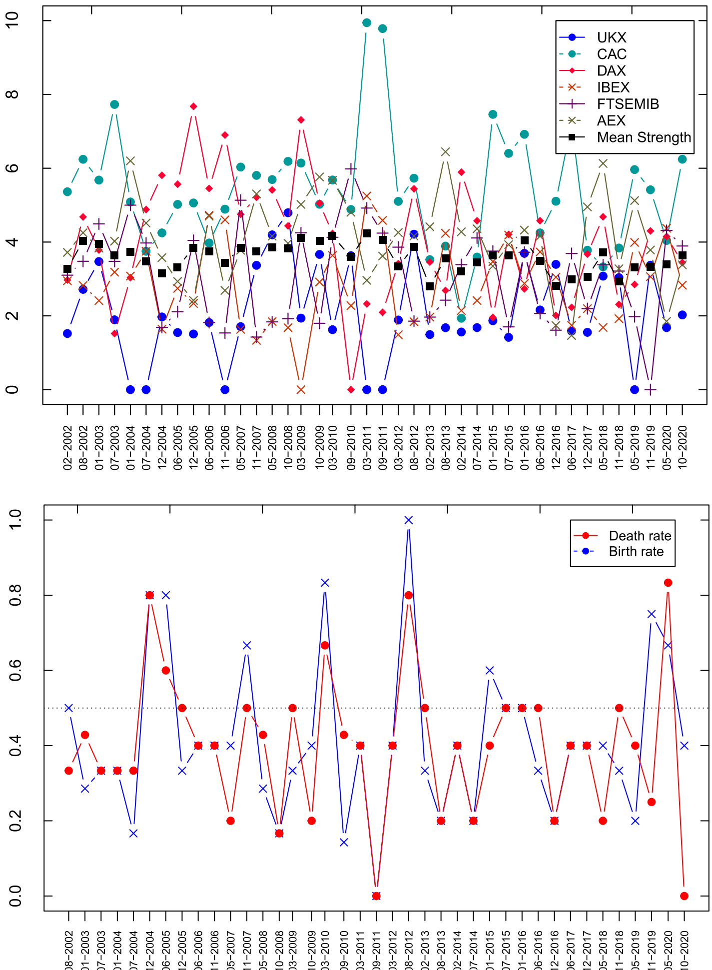


Fig. 7. (Color online) (Upper) Strength st_i of each market by time segment. Black line represents the mean strength $\langle st_i \rangle$. A one-way ANOVA was carried out to test if there are differences between the strength values of each market. The results indicate a significant difference at 1%. Then a *posthoc* Turkey test reveals that all the values are significantly different at 5%, except for the IBEX-UKX, AEX-DAX, and FTSEMIB-IBEX pairs. (Lower) The $D(t)$ and $B(t)$ rates in red and blue colors respectively. The dotted black line indicates a rate of 0.5, indicating that half of the edges in period t die or were born.

French market (CAC) tends to have a high level of importance in relative terms. It tends to be permanently tied to other markets. The same seems true for the German market (DAX), at least in the first 20 time segments. On the other hand, the English (UKX) and Milan (FTSEMIB) markets tend to have the lowest levels of this measure. To compare these results, a similar result is found, but using correlation networks of returns (see Appendix C).

The lower part of Fig. 7 shows the $D(t)$ and $B(t)$ ratios. It is interesting to note that the critical networks exhibit drastic topological changes at specific periods.

However, the most drastic changes do not necessarily correspond to the periods of financial turmoils.

At the time segment 08-2012, the death rate was 0.8, indicating that 4 of 5 edges are no longer present from one period to the next. The birth rate was one, indicating that there are five new edges. This is a drastic reconfiguration of the critical network (see Figure in Appendix B). Also, there are other periods in which a significant reconfiguration of the critical network is observed (12-2004, 08-2005, 03-2010, and 05-2020). In all but the last, the birth rate is higher than the death rate, suggesting an increase in the number of critical network edges, leaving more critical ties between markets. In almost half of the periods, however, the birth rate is greater than or equal to the death rate, suggesting a tendency for some stability of the critical networks to keep the level of ties constant in the system.

3.6. *Financial irregularities*

It is known that changes in the distribution of the net orientation of the spins are a clue that the system is reorganizing itself [3]. With most of the couplings being positive, the system may be in order, favoring most market indices at -1 or $+1$ values. The distribution of the net orientation, in this case, is of low entropy. On the contrary, we see a uniform distribution of the net orientation and consequently high entropy in periods of higher disorder. The first case is associated with unlikely events attributable to financial crises [5], and also to periods of financial overconfidence and bull markets [21].

One way to characterize the order-disorder level of the system is to estimate the mean-field approximation of the entropy S_{MF} [22] over each time segment as

$$S_{\mathrm{MF}} = -\sum_{i=1}^{N} \frac{1+m_i}{2}\ln\frac{1+m_i}{2} + \frac{1-m_i}{2}\ln\frac{1-m_i}{2}, \tag{5}$$

where $m_i = \langle s_i \rangle$ is the mean orientation of market i, and N is total number of markets, in our case the six European markets.

Figure 8 upper shows the empirical relationship between the mean orientation of all spins $m = 1/N \sum_{n}^{N} \langle s_i \rangle$ in each of the 40 time segments versus the mean-field entropy S_{MF}. This scatterplot highlights the characteristic quadratic curve of entropy as a function of mean orientation. Its shape comes from Eq. (5). We recognize that there are two extremes of low entropy given by low and high mean net orientations. The maximum entropy is given when the average of the orientations is equal to zero, i.e. when the market movement is disordered, and the spins that represent them are aligned half downwards and half upwards.

We identify precisely three-time segments in the zone of low entropy and a low negative mean orientation, representing a bear market, and these points are precisely the periods of three well-recognized financial crises, which are marked in a red circle in Fig. 8 (lower). For example, in a bearish market, all stock prices drop, as happened in the 2008 Subprime crisis, in the 2002 crisis after the dot-com bubble, and, in the

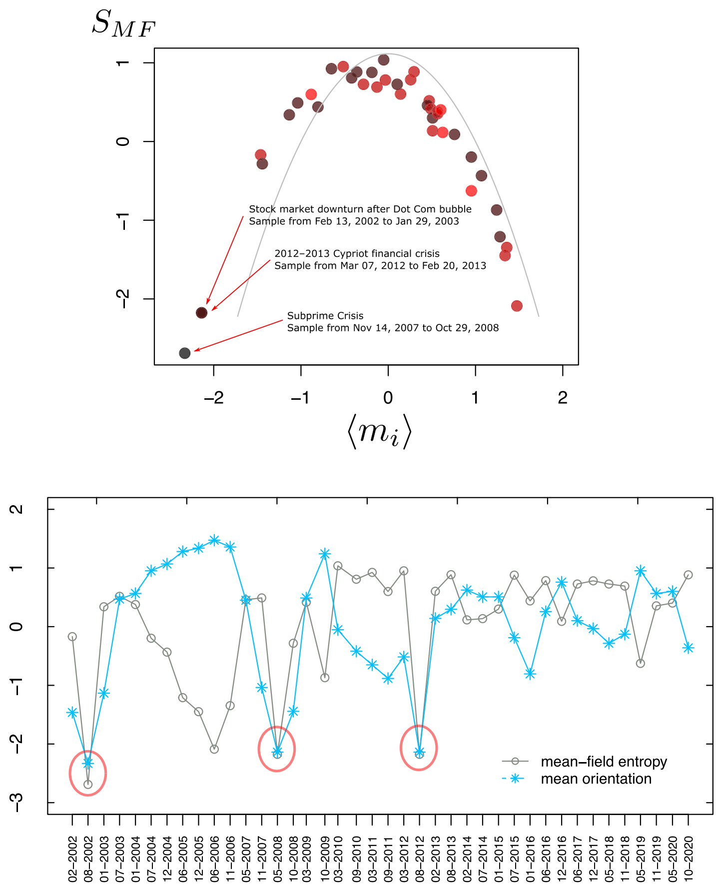


Fig. 8. (Upper) The relationship between mean-field S_{MF} entropy and the mean orientations of all market indices. All values are normalized. The color of the points indicates the mean degree of the critical graphs. Black color indicates low mean degree and red indicates high mean degree. The gray curve indicates the theoretical relationship. (Lower) Mean-field entropies S_{MF} and mean orientations of all market indices $m = 1/N \sum_n^N \langle s_i \rangle$ across each time segment. Circles indicate three financial crisis.

Cypriot financial crisis, indicated in the figure with the arrows and circles. In this case, all markets showed a tendency to be 'aligned' in a bearish direction. Interestingly, it appears that under these conditions, the critical graphs have a low level of connectivity, as measured by the mean degree (black dots).

4. Another Example with more Markets

The system examined in the previous sections was only with a few European markets. This section offers an analysis of a system with a larger number of markets. In particular, we analyze two interesting periods that allow us to compare the transition in times of low and high volatility in stock market returns. We take this time eleven markets, adding to the previous ones, the markets of Russia (MOEX), Switzerland (SMI), Sweden (OMX), Standard & Poor's of USA (SPC), and Honk-Kong (HSI). The first period we call the pre-Covid19 (pre-c19) outbreak between April 13, 2019 and December 18, 2020, encompassing a period of general growth and low volatility. The second, the COVID19 (c19) outbreak between December 19, 2019 and August 20, 2020. In this case, the period includes the time of significant negative returns in all markets and then the months after March of gradual recovery. This period covers times of high volatility in market returns.

In the system analyzed previously with $N = 6$ markets, the total number of possible couplings was 15, and the transition graph is observed to be staggered. However, with $N = 11$, the number of coupling is now 55. This indicates that the number of parameters to be inferred by the Boltzmann machine increases substantially with the size of the system (for each additional node, $N + 1$ new parameters must be inferred among couplings and fields), making the inference process slower and consuming more computational resource.

Figure 9 (left) presents two transition diagrams: one for the pre-c19 outbreak period and the other for the c19 outbreak period. Both transitions appear to occur at quite different times in the edge removal process. In pre-c19, the increase in KL divergence occurs earlier relative to c19 period. As before, we took the 70th percentile of the KL divergence differences as a threshold to determine the critical network,

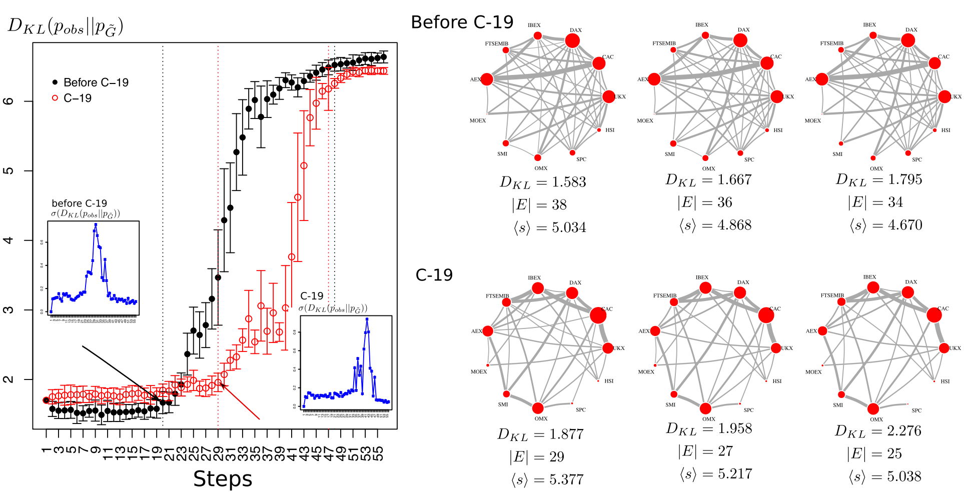


Fig. 9. (Color online) Phase plot for KL distance in the period pre-Covid19 outbreak (black) and COVID19 outbreak (red). Bars show ±1 standard deviation of the mean KL distance calculated over 30 Metropolis-Hasting simulations at each step. Arrows indicate the critical networks' point; Steps 20 and 29 for pre-c19 and c19 outbreak periods, respectively. The right side shows the critical networks at steps 18, 20, and 22 and steps 27, 29, and 31 for pre-c19 and c19 outbreak periods, respectively. Also indicated are the KL Distance D_{KL}, the number of couplings $|E|$, and the mean strength $\langle s \rangle$ of each graph. The insets represent the standard deviation of the KL divergence in every step.

which indicates the point just before where the KL divergence starts to increase. In pre-c19 outbreak, we identify this change at step 20, while in the c19 outbreak, the increase appears to be much later, at step 29. This is interesting because it indicates that the critical network has fewer edges under c19, in this case, at times of higher financial distress and crisis, than in regular periods as pre-c19.

In either case, the increases from one step to the next are not as drastic, so it is worth looking at the critical graphs in the neighboring steps. This is presented in Fig. 9 right-hand side, where it can be seen which edges are removed and remain active for the step selected as the beginning of the transition. The critical networks in both periods are different. In the pre-c19 period, 36 couplings support the system, out of 55 in total, i.e. 65.45%, while in c19, there are only 27, i.e. 49.1%. Interestingly, the average strength of the network in c19 is higher than in pre-c19 (5.22 versus 4.87), indicating a greater ferromagnetic behavior between markets. The critical network is sustained with fewer couplings in market turbulence, and the mutual influence between markets is greater than in pre-c19. At the market level, it is observed that CAC, DAX, and AEX have a high strength between other markets in both periods, while, for example, the influence of SPC and HSI decreases in c19. It is also observed that while in pre-c19, the AEX-CAC, CAC-FTSEMIB, IBEX-DAX, among others, couplings were very strong, while in c19, their intensity decreases.

5. Conclusion

The use of the Boltzmann machine to find the maximum entropy distributions turns out to be helpful to represent the stock markets, using only the interactions/couplings between them and their fields. We have seen that this representation allows us to find the most important bonds that drive the market through an iterative process of eliminating couplings, leaving only those able to replicate the description of the data in terms of the original distribution of states of the system. The result of this process is a subgraph of the complete network of couplings (Boltzmann machine), which has the minimum number of edges that allow recovering the system's observed distribution satisfactorily. We call this subgraph the critical graph.

In line with what has been found in other studies [4, 3], the interaction strengths are time-dependent, so the critical networks are also changing, and consequently, the set of interaction strengths that hold the system is not constant. However, when calculating the strength of the couplings that connect to each node, it is observed that there is a tendency for certain markets to have a higher level of strength than others. For example, of the six European markets, it is observed that consistently in most of the periods under analysis, the French and German markets tend to dominate. In contrast, the Spanish and UK markets tend to have a lower level of strength.

An interesting result is that the changes in critical networks, measured as the rate of edges dying and being born from one overlapped period to another, are not constant. It is observed that drastic changes in the topology of the critical networks are approximately two years after the end of the two crises, i.e. in periods of growth and not necessarily in crashes. On the other hand, it is observed that every time that from one period to another, the level of critical connectivity of the market increases (the rate of births of edges is higher than the rate of deaths), in the following period, precisely the opposite occurs, i.e. the level of critical interdependence decreases (the rate of births is lower than the rate of deaths). In other circumstances, both rates are equal. This suggests that the system somehow reconfigures itself to maintain a certain level of equilibrium in the mutual interdependence between markets.

We see an increase in the level of ferromagnetism of the system at times of crisis in agreement with the findings of [4]. We observe variations of the interaction strengths and fields but not necessarily drastic changes in the strength of the nodes. In other words, if Germany and France are influential markets on the continent, in the face of a crisis, it does not mean that they will now cease to be so. This raises some connection with the problem of systemic risk in the financial system, in which the market seems to be much more sensitive to the failure of one particular market than another. In this case, for example, the effect that a failure in the French or German market could have on the rest of the European market is greater than that of the Spanish or UK markets. Something similar occurs in stock markets at the national level in which companies with a higher degree in the correlation networks tend to have a greater potential to produce a harmful effect on the rest of the market in the event of a failure or adverse event of the company [3, 23].

While it is interesting to analyze the most important bonds between stock markets, our analysis must invariably face the problem of inference of coupling and field parameters. This is not a trivial problem, and the solution that is possible to find through the Boltzmann machine, although accurate, takes a lot of time and computational resources because it must iterate a sufficient number of times to ensure that convergence has been achieved. In addition, we must consider that to find the critical network, it is necessary to carry out several metropolis hasting simulations for each iteration of the edge-removal process. Consequently, the process is slow, and it is convenient to use other approximate and recent inference methods [4, 10].

Finally, an interesting aspect is that the transition point at which the gradual and significant increase of the KL distance starts to occur seems to depend on the level of entropy on the system. Under crisis periods associated with lower entropy, the transition occurs later under higher entropy. This means that the critical network in crisis periods requires fewer bonds to represent the system. This behavior may be related to the fact that in times of higher volatility of returns, the co-movement of the markets is in order, the correlation of returns increases. In this sense, fewer couplings are needed to represent the system states. We have observed that the curves on the period pre-c19 outbreak and on the period c19 outbreak developed a significant distance between them. In this context, it is interesting to wonder if this gap can represent a critical distance between stable and unstable markets or towards the market's instability. Given the evidence of our results, the direct relationship between critical network connectivity and system entropy is a conjecture. We have observed that as the mean orientation of the market is lower during a financial crisis, our system is in a ferromagnetic state then, and analogous to the Ising model, the entropy value is also low. This behavior has been reported in different phenomena [24, 25]. However, a study with a more significant number of nodes and more samples would be necessary to investigate this possible phenomenon. A vein of future research has to see if these changes in the number of couplings to represent the system has any correspondence with the early detection of crisis signals. The idea is to test the possibility that financial crises are instances of critical transitions similar to what occurs in other complex systems [26], and add cross-sectional time-lagged dependences to estimate temporal dependency networks [27].

In conclusion, a look at a complex system such as the stock market system using critical networks is an exciting alternative to reveal the interactions that sustain the system and eventually offer another view to analyze the financial stability and contagion problem in future studies.

Appendix A. Mean Field Approximation

This appendix considers the Naive mean-field method [28, 29], where the values of coupling parameters and the external fields have been calculated through of the next equations:

$$J_{ij} = -\left(C^{-1}\right)_{ij}, \tag{A.1}$$

$$h_i = \tanh^{-1}(m_i) - \sum_j J_{ij} m_j. \tag{A.2}$$

The matrix C corresponds to the covariance matrix, and its elements C_{ij} are $C_{ij} = \langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle$. The parameter m_i is the magnetization, defined as the average of the spins orientations $m_i = \langle s_i \rangle$.

This method has been used by other studies by showing a very well approximation on the reconstruction of the parameters such as the magnetic field and the coupling constants [4, 30]. We call the set of parameters estimated with the mean-field approximation $\{\mathbf{J}_{\mathrm{MF}}, \mathbf{h}_{\mathrm{MF}}\}$.

Appendix B. Critical Graphs Across Time Segments

In this appendix we provide additional information concerning the critical networks $G(\tilde{V}, \tilde{J})$ found for each of the 40 time segments of 250 days each. For this we have identified the transition phase of each period, identifying the positive and sudden change in KL. In some cases, this change is rather gradual, but it is always possible to find the same S-shape of the transition graphs, which allows us to find the point where the deviation of KL from one step to another becomes significantly larger.

Figure B.1 illustrates how the critical networks change over time. A closer look at the graphs shows how the bonds are changing. It is very rare for critical networks to be disconnected (only 9 out of 40), indicating the nature of close dependence between stocks in the market.

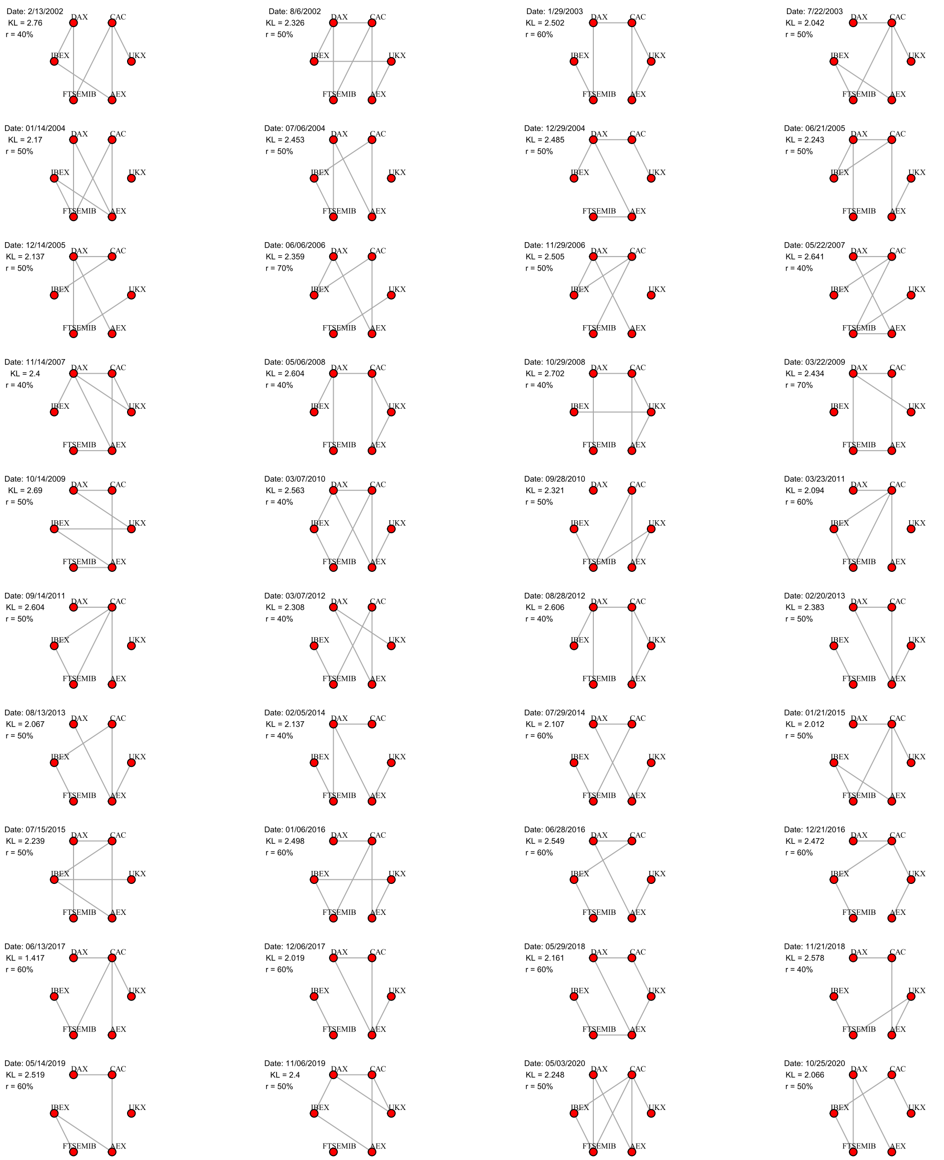


Fig. B.1. Critical graph across time segments for the European stock markets with seven countries. The indicated date in each graph is the middle date of the time segment. KL is the Kullback–Leibler distance $D_{\mathrm{KL}}(p_{\mathrm{obs}}||p_{\tilde{G}})$. The parameter r is the fraction of times the graph appears as critical in the 50 simulations carried out for each time segment.

Appendix C. Strength st_i Computed from the PMFG Graphs

This appendix offers an alternative result in the calculation of market strengths as a way to compare the results found using the critical networks. For the comparison task, we used the Planar Maximally Filtered Graph (PMFG), which allows filtering information from complex data through the extraction of a subgraph with the most representative links of the original graph, controlling the genus of the graph [31]. It applies to correlation networks, such as, for example, networks formed from the correlations of returns between pairs of stocks [32]. In these cases, the complete network of $N(N-1)/2$ edges is reduced to only $3N-6$ edges.

In this case, the procedure, in brief, is as follows: The Pearson correlation matrix between the returns of the market indices is calculated. The correlation matrix is transformed into a distance metric. When two markets have a high level of correlation, the distance is small, while otherwise, the distance is large at a low level of correlation. The distance matrix is equivalent to the weighted matrix of a complete network, which serves as input to find the PMFG (see [31] for a detailed description).

Figure C.1. shows the strength of the nodes of each PMFG over time. Compared with Fig. 7, it can be noted some consistency with both methods in the sense that the importance of the markets measured by strength is maintained, at least in the case of the French market (CAC), which continues to reveal its importance in this six-node system. At the same time, the United Kingdom (UKX) is the one that presents the lowest level of importance. Similarly, as seen in Fig. 7, the Spanish market (IBEX)

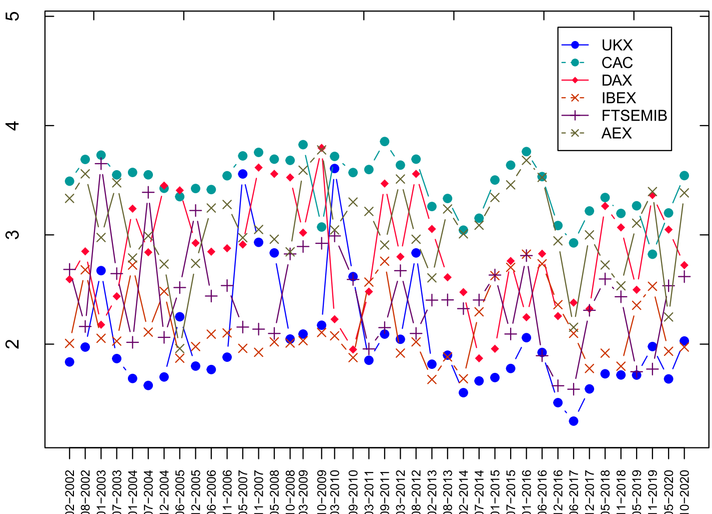


Fig. C.1. Strengths of each market st_i across time segments computed from the PMFG graphs. A one-way ANOVA was carried out to test if there are differences between the strength values of each market. The results indicate a significant difference at 1%. Then a *posthoc* Turkey test reveals that all the values are significantly different at 5%, except for the IBEX-UKX and AEX-DAX pairs.

also shows less importance. In contrast, the German market (DAX) seems vital in all periods but always slightly below the French market.

It should be clarified that the strength in the PMFG is simply a sum of the distances (proxy of linear correlations) that node i has to each of the other nodes with which it is connected. In this sense, a node with a higher strength than another node means that it is more connected to other nodes (higher degree) or less connected but at a greater distance. On the other hand, the strength of the nodes in the critical network is a direct measure of importance because it is calculated based on the couplings J_{ij} that node i has to all the others with which it is connected. Because of the way the coupling network and, therefore, the critical network is constructed (based on couplings and fields, and not correlations), the measurement with our method better describes the influence that a node has on affected other components of the system.

References

[1] Rosenow, B., Gopikrishnan, P., Plerou, V. and Eugene Stanley, H., Random magnets and correlations of stock price fluctuations, *Physica A* **314**(1–4) (2002) 762–767.

[2] Bury, T., A statistical physics perspective on criticality in financial markets, *J. Stat. Mech. Theory Exp.* **2013**(11) (2013) P11004.

[3] Bury, T., Statistical pairwise interaction model of stock market, *Eur. Phys. J. B* **86**(3) (2013) 1–7.

[4] Borysov, S. S., Roudi, Y. and Balatsky, A. V., U.S. stock market interaction network as learned by the Boltzmann machine, *Eur. Phys. J. B* **88**(12) (2015) 321.

[5] Valle, M. A., Lavín, J. F. and Magner, N. S., Equity market description under high and low volatility regimes using maximum entropy pairwise distribution, *Entropy* **23**(10) (2021) 1307.

[6] Sherrington, D. and Kirkpatrick, S., Solvable model of a spin-glass, *Phys. Rev. Lett.* **35** (1975) 1792–1796.

[7] LeCun, Y., Chopra, S., Ranzato, M. and Huang, F.-J., Energy-based models in document recognition and computer vision, in *Ninth Int. Conf. Document Analysis and Recognition (ICDAR 2007)*, Vol. 1 (IEEE, 2007), pp. 337–341.

[8] Jaynes, E. T., Information theory and statistical mechanics, *Phys. Rev.* **106** (1957) 620–630.

[9] Jaynes, E. T., Information theory and statistical mechanics. ii, *Phys. Rev.* **108** (1957) 171–190.

[10] Zeng, H.-L., Lemoy, R. and Alava, M., Financial interaction networks inferred from traded volumes, *J. Stat. Mech., Theory Exp.* **2014**(7) (2014) P07008.

[11] Chau Nguyen, H., Zecchina, R. and Berg, J., Inverse statistical problems: From the inverse Ising problem to data science, *Adv. Phys.* **66**(3) (2017) 197–261.

[12] Schneidman, E., Berry, M. J., Segev, R. and Bialek, W., Weak pairwise correlations imply strongly correlated network states in a neural population, *Nature* **440**(7087) (2006) 1007–1012.

[13] Ackley, D. H., Hinton, G. E. and Sejnowski, T. J., A learning algorithm for Boltzmann machines, *Cogn. Sci.* **9**(1) (1985) 147–169.

[14] Roudi, Y., Tyrcha, J. and Hertz, J., Ising model for neural data: Model quality and approximate methods for extracting functional connectivity, *Phys. Rev. E* **75**(5) (2009) 051915.

[15] Roudi, Y., Aurell, E. and Hertz, J. A., Statistical physics of pairwise probability models, *Front. Comput. Neurosci.* **3** (2009) 22.

[16] Dotsenko, V., *An Introduction to the Theory of Spin Glasses and Neural Networks*, Vol. 54 (World Scientific, 1995).

[17] Magner, N., Lavin, J. F., Valle, M. and Hardy, N., The predictive power of stock market's expectations volatility: A financial synchronization phenomenon, *PLos One* **16**(5) (2021) e0250846.

[18] Yang, C.-X., Wu, H.-F., Zhang, Y.-C., Xia, B.-Y. and Itoh, M., Phase synchronization detection of financial market crises, *Mod. Phys. Lett. B* **25**(4) (2011) 243–254.

[19] Sornette, D., Physics and financial economics (1776–2014): Puzzles, Ising and agent-based models, *Rep. Prog. Phys.* **77**(6) (2014) 062001.

[20] Onnela, J.-P., Chakraborti, A., Kaski, K., Kertesz, J. and Kanto, A., Dynamics of market correlations: Taxonomy and portfolio analysis, *Phys. Rev. E* **68**(5) (2003) 056110.

[21] Welch, I., Herding among security analysts, *J. Financ. Econ.* **58**(3) (2000) 369–396.

[22] Binder, K. and Peter Young, A., Spin glasses: Experimental facts, theoretical concepts, and open questions, *Rev. Mod. Phys.* **58**(4) (1986) 801.

[23] Gao, Y.-C., Wei, Z.-W. and Wang, B.-H., Dynamic evolution of financial network and its relation to economic crises, *Int. J. Mod. Phys. C* **24**(2) (2013) 1350005.

[24] Brandani, G. B., Schor, M., MacPhee, C. E., Grubmüller, H., Zachariae, U. and Marenduzzo, D., Quantifying disorder through conditional entropy: An application to fluid mixing, *PLoS One* **8**(6) (2013) e65617.

[25] Vijayaraghavan, V. S., James, R. G. and Crutchfield, J. P., Anatomy of a spin: The information-theoretic structure of classical spin systems, *Entropy* **19**(5) (2017) 214.

[26] Diks, C., Hommes, C. and Wang, J., Critical slowing down as an early warning signal for financial crises? *Empir. Econ.* **57**(4) (2019) 1201–1228.

[27] Begušić, S., Kostanjčar, Z., Kovač, D., Eugene Stanley, H. and Podobnik, B., Information feedback in temporal networks as a predictor of market crashes, *Complexity* **2018** (2018).

[28] Peterson, C. and Hartman, E., Explorations of the mean field theory learning algorithm, *Neural Netw.* **2**(6) (1989) 475–494.

[29] Tanaka, T., Mean-field theory of Boltzmann machine learning, *Phys. Rev. E* **58** (1998) 2302–2310.

[30] Roudi, Y., Aurell, E. and Hertz, J., Statistical physics of pairwise probability models, *Front. Comput. Neurosci.* **3** (2009) 22.

[31] Tumminello, M., Aste, T., Di Matteo, T. and Mantegna, R. N., A tool for filtering information in complex systems, *Proc. Natl. Acad. Sci.* **102** (2005) 10421–10426.

[32] Tumminello, M., Lillo, F. and Mantegna, R. N., Correlation, hierarchies, and networks in financial markets, *J. Econ. Behav. Organ.* **75** (2010) 40–58.