Computational Economics (2021) 57:1149–1166
https://doi.org/10.1007/s10614-020-10007-3

Entropy of Graphs in Financial Markets

Chun-Xiao Nie[1] · **Fu-Tie Song**[2]

Accepted: 10 June 2020 / Published online: 2 July 2020

Abstract
This article analyzes the eigenvalues of financial graphs and discusses different types of graphs using random graph theory. We found that the energy-based Rényi index is an effective tool for studying the spectrum of financial graphs. The entropy of financial graphs is usually different from the theoretical predictions of random graph theory, which implies the existence of rich structures. This article also constructed some benchmark graphs for comparative analysis through the classic financial models. The calculations show that the geometric Brownian motion and the one factor model correspond to completely different entropy values based on eigenvalues, thus providing two extreme cases for characterizing real graph entropy. In particular, we find a high correlation between the degree-based Rényi index and the eigenvalue-based Rényi index based on real market data. This article shows the analysis of the structure and complexity of financial graphs from the perspective of graph entropy, thus providing a new way to analyze different types of financial graphs.

Keywords Rényi index · Financial graph · Correlation matrix · Eigenvalues

1 Introduction

In the past two decades, various methods have been developed to extract the structure of financial correlation matrices in the form of graphs or networks. The mainstream methods include *MST* (minimum spanning tree) algorithm, *PMFG* (planar maximally filtered graph) algorithm and threshold method (see Mantegna 1999; Tumminello et al. 2005; Yang and Yang 2008; Tse et al. 2010; Song et al. 2011). In

✉ Chun-Xiao Nie
niechunxiao2009@163.com; niechunxiao@mail.zjgsu.edu.cn

Fu-Tie Song
ftsong@ecust.edu.cn

[1] School of Statistics and Mathematics, Zhejiang Gongshang University, Hangzhou 310018, China

[2] Department of Finance, School of Business, East China University of Science and Technology, Shanghai 200237, China

previous studies, many researchers have focused on the topological structure of *MST* and *PMFG*, and found that power law distribution exists widely in financial graphs (Vandewalle et al. 2001; Wiliński et al. 2013, 2015).

In addition to using correlation coefficient matrices to construct financial graphs, some studies have used some different approaches to construct networks, such as partial correlation coefficients (Kenett et al. 2010a; Wang et al. 2018; Meng et al. 2014), Granger causality (Billio et al. 2012), and VAR models (Yi et al. 2018). In addition, some studies have analyzed financial risks from a network perspective, such as focusing on tail dependence (Hua et al. 2019; Gang-Jin and Chi 2016), or risk spillover effects (Hautsch et al. 2014; Gang-Jin et al. 2017). In particular, empirical studies have shown that analysis of the structure of financial networks not only helps measure risk, but also enhances portfolio performance (Zhao et al. 2016; Pozzi et al. 2013; Enguthaiwat 2018).

This article only analyzes financial graphs based on correlation matrices, which is the most widely studied category. In previous studies, the random matrix theory (*RMT*) provides the spectral characteristics of noise-driven correlation matrices and is therefore widely used to analyze financial correlation matrices. For example, some pioneering studies have revealed some differences between the empirical features exhibited by the correlation matrix of real world and the properties predicted by random matrix theory (Laloux et al. 1999; Plerou et al. 1999). The spectral analysis of the financial correlation matrix is important for analyzing dynamics in the market (Kumar and Deo 2012; Junior and Franca 2012; Conlon et al. 2009). Further, some researchers have combined the theory of random matrix with financial graphs to analyze the market (Jiang et al. 2014; Kenett et al. 2010b; Eom et al. 2009; Song et al. 2011; Dai et al. 2016). In summary, the correlation matrix can be analyzed from two perspectives of the network and the random matrix, and the latter has been widely used to construct a benchmark for evaluating the correlation coefficient matrix of the real market, and in particular, to analyze the eigenvalues of the correlation matrix. Corresponding to financial graphs, random graph theory also provides a benchmark for research. Therefore, similar to applying *RMT* to study correlation matrices, we use random graph theory to discuss the adjacency matrix of financial graphs, and focus on the heterogeneity of eigenvalues.

Although many researchers have studied the degree distribution of financial graphs in depth, the research on the distribution of eigenvalues of adjacency matrices has not been widely carried out. However, in complex network theory, a large number of theoretical and experimental work focuses on the distribution of eigenvalues of the adjacency matrix. For example, Farkas et al. (2001); Nadakuditi and Newman (2013) studied the eigenvalues of adjacency matrices of small world networks and scale-free networks. Since real networks often have rich structures, such as communities or modules, some researchers carefully studied the spectrum of networks with community or module structures (Chauhan et al. 2009; Zhang et al. 2014; Jalan et al. 2011). Sanjeev Chauhan et al. studied the spectrum of adjacency matrices of networks with community structures and discussed the application of spectral analysis in community detection (Chauhan et al. 2009). Xiao Zhang et al. analyzed the spectrum of networks with conditions of community structure and degree distribution (Zhang et al. 2014). In addition to undirected networks, studies have shown that

spectral analysis is important for analyzing the modular structure of directed networks (Jalan et al. 2011). Not only the community structure, the spectral analysis of the network can also be applied to the analysis of the microstructure of the network. For example, recently, Newman discussed the spectrum of the adjacency matrix of networks with short loops (Newman 2019). More details about the spectrum of the network can be found in the relevant review (Sarkar and Jalan 2018).

We mainly use the Rényi entropy of the graph to analyze the structure of the financial graph. The entropy proposed by Rényi is an important concept in information theory and fractal geometry (Rényi 1961; Jiang et al. 2019). Its special cases include Shannon entropy and Hartley entropy, as well as several other types of entropy. The Rényi entropy can be rewritten into a normalized form (Rényi index) and takes values in the interval [0, 1] (Eliazar 2011). Currently, the Rényi index has been used to analyze the topology of financial graphs (Nie et al. 2016; Nie and Song 2018). The researchers found that the Rényi index of financial graphs in the market is usually larger than the value based on Brownian motion, due to the heterogeneity of the degree distribution (Nie et al. 2016). Furthermore, previous research has shown that the Rényi index can be used in fractal analysis of financial graphs (Nie and Song 2018).

In order to study the heterogeneity of the eigenvalues of the adjacency matrix, we focus on the absolute value of the eigenvalues and its entropy. Researchers have defined generalized graph entropies based on eigenvalues, including Rényi entropy and Daròczy's entropy (Dehmer et al. 2015). Here, we still use the normalized Rényi entropy (Rényi index) to study the eigenvalues. One of the advantages of using the eigenvalues-based Rényi index is that the calculated values are normalized so that they can be used well in comparative analysis. Second, the Rényi index can be defined by the Lorenz curve, which can be intuitively expressed by the curve-based index (Eliazar and Sokolov 2012). Third, the energy is an important concept for studying graphs and has been extensively studied as a useful topic (Li et al. 2010). Following the existing research results, there is a direct relationship between the Rényi index based on the eigenvalue and the energy (Dehmer et al. 2015).

We organize this article as follows. After a brief description of the data set, we review the Rényi entropy and the graph energy in the method section. We analyze the relationship between normalized Rényi entropy (Rényi index) and energy. Then, we analyzed the value of the Rényi entropy of the random graph. In the results section, we analyzed the entropy values of the three types of financial graphs and compared them with the results generated by the financial benchmark model and the random graph model. At last, we dynamically analyze the relationship between the degree-based Rényi index and the eigenvalue-based Rényi index.

2 Data

The S&P500 index is a composite index that is commonly used to show the performance of the US market. The index is constructed by the stocks of the top 500 companies selected by market value. Here we used the daily closing price series of constituent stocks of the S&P500 index, which is from 2006/1/3 to 2018/8/15. There are a total of

3117 trading days. Since the data for some stocks was missing during the period considered, we used 432 stocks to construct the financial graphs. The data for all the stocks we use is extracted from Yahoo Finance (https://finance.yahoo.com/).

This study focuses on *PCC*-based graphs, where Pearson correlation coefficient (*PCC*) is generally considered to capture linear correlations. Recently, research on the stock market has revealed a relationship between *PCC* and non-linear measures (Hartman and Hlinka 2018). Before constructing a financial graph, the original price series needs to be pre-processed as a return series. To describe the method, we assign a label to each stock. If there are n stocks, then there are n labels $\{1, 2, \cdots, n\}$. We assume that the price time series of each stock k is $P_k = \{P_k(t)\}$, then its return series is $R_k = \{R_k(t)\}$ $(R_k(t) = log(P_k(t+1)) - log(P_k(t)))$, so that the Pearson correlation coefficient between stocks k and l is as shown in Eq. 1. Here the symbol $\langle R_l \rangle$ means calculating the mean of the series $\{R_l(t)\}$.

If the number of stocks is n, the financial graph generated based on the correlation matrix $\rho = [\rho(k, l)]$ has n nodes. To simplify the discussion, stock labels are used directly to mark network nodes. Each financial time series is mapped to a node through the graph model. The financial graph is a two-tuple $G(V, T)$, where $V = \{i | i = 1, \cdots, n\}$ and $T = [T(k, l)]$ are the node set and the adjacency matrix, respectively. The element $T(k, l) = 1$, if there is a link between k and l, otherwise it is 0, in addition, $T(k, k) = 0$.

$$\rho(k, l) = \frac{\langle R_k R_l \rangle - \langle R_k \rangle \langle R_l \rangle}{\sqrt{\langle R_k^2 - \langle R_k \rangle^2 \rangle \langle R_l^2 - \langle R_l \rangle^2 \rangle}} \tag{1}$$

Here, we use software Pajek to plot all the financial graphs (http://mrvar.fdv.uni-lj.si/pajek/).

3 Method

3.1 Classic Benchmark Financial Model

We use geometric Brownian motion (*GBM*) to simulate the price series of each selected component stock, where the mean (μ) and standard deviation (σ) of the return series are used as parameters for the simulated stock price series, as shown in Eq. 2 (Campbell et al. 1997). Here S_t represents the price process and W_t is the Wiener process. In this way, for each stock, we can generate a surrogate series, which can further calculate the correlation coefficient matrix and the correlation graph, where the correlation structure is eliminated. In addition, the simulation steps can be repeated multiple times to generate benchmarks for financial graphs.

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{2}$$

In addition to geometric Brownian motion, we also use the classic one-factor model (*OFM*) to generate simulated return series. The model includes only one factor, the market factor $R_m(t)$, as shown in Eq. 3 (Campbell et al. 1997). As in previous studies (Bonanno et al. 2003, 2004; Nie et al. 2016), we estimate the coefficients (α_k, β_k) of

Eq. 3 and the standard deviation s_k of the random terms, and then use Eq. 4 to generate a simulated series $\{R_k^s(t)\}$, where $\epsilon_N(0, s_k)$ is Gaussian noise. Finally, we apply model-based simulation series to generate correlation matrices and financial graphs.

$$R_k(t) = \alpha_k + \beta_k R_m(t) + \epsilon_k(t) \tag{3}$$

$$R_k^s(t) = \alpha_k + \beta_k R_m(t) + \epsilon_N(0, s_k) \tag{4}$$

3.2 *MST* and *PMFG*

The minimum spanning tree (*MST*) is an important concept of graph theory. At present, there are many algorithms for constructing *MST*, some of which are classical algorithms such as Kruskal algorithm and Prim algorithm (Kruskal 1956; Prim 1957). Here we use the Prim algorithm. *MST* is a connected graph that connects all nodes together, where the total edge weights are minimized. For n nodes, *MST* includes $n-1$ edges without cycles. Here, we use the classic method proposed by Mantegna to calculate the distance $D(k,l) = [2(1-\rho(k,l))]^{1/2}$ and generate the *MST* (Mantegna 1999). *PMFG* is a planar graph with 3-cliques and 4-cliques structures that do not exist in *MST*. Here, we only briefly review the construction method of *PMFG*, which includes the following main steps (Tumminello et al. 2005).

All upper triangular matrix elements in the correlation coefficient matrix are sorted in descending order to generate a set $L_{sort} = \{L_{sort}^l\}$, where each element L_{sort}^l corresponds to a pair of nodes and $L_{sort}^1 \geq L_{sort}^2 \geq \cdots L_{sort}^{n(n-1)/2}$. Then, we add the links to the n nodes in descending order of L_{sort}, that is, add links between the nodes corresponding to L_{sort}^l, and ensure that the graph after adding the links is a planar graph. Finally, a graph comprising $3(n-2)$ edges is constructed, where $3(n-2)$ is the maximum number of edges that a planar graph can include. In general, a graph with a genus of g includes at most $3(n-2+2g)$ edges (Tumminello et al. 2005).

3.3 Threshold Network

There are multiple ways to generate a threshold network, which relies on constructing an adjacency matrix from a correlation coefficient matrix or a distance matrix (Yang and Yang 2008; Tse et al. 2010). The main difference between different methods is the way to construct the threshold, such as directly using the correlation coefficient as the threshold (Tse et al. 2010; Boginski et al. 2005), or its absolute value as the threshold (Yang and Yang 2008). We use the method proposed in (Yang and Yang 2008) here, that is, the off-diagonal elements in the correlation coefficient matrix that are greater than or equal to a critical value p_{th} are converted to 1, and the other elements are converted to 0. The adjacency matrix $T_{p_{th}}$ of the threshold network $W(V, T_{p_{th}})$ is generated by the rules in Eq. 5.

$$\begin{aligned} T_{p_{th}}(k,l) &= 1, |\rho(k,l)| \geq p_{th} \\ T_{p_{th}}(k,l) &= 0, |\rho(k,l)| < p_{th} \end{aligned} \tag{5}$$

In this paper, in order to apply the random graph theory to analyze the threshold network, the threshold p_{th} needs to be converted into the proportion of links between nodes. In the calculation, we can calculate $p = \frac{\sum_{k,l} T_{p_{th}}(k,l)}{n(n-1)}$ directly from the adjacency matrix.

3.4 Rényi Index

Based on the adjacency matrix T, the degree of node k is $d_k = \sum_l T(k,l)$, and the average degree is $d' = \frac{1}{n}\sum_k d_k$. Applying the definition of the Rényi index directly (Eliazar 2011), the degree-based Rényi index of the network can be defined as Eq. 6 (Nie et al. 2016), where α is a parameter. In this article α is a real number greater than zero. In addition, we use m to represent the number of links in the network, that is, the number of non-zero elements in the upper or lower triangular matrix of T.

$$\begin{aligned} R(\alpha) &= 1 - \left[\sum_{k=1}^{n}\left(\frac{d_k}{d'}\right)^{\alpha}\left(\frac{1}{n}\right)\right]^{\frac{1}{1-\alpha}}, \quad \alpha \neq 1 \\ R(1) &= 1 - \exp\left\{-\sum_{k=1}^{k}\left[\frac{d_k}{d'}\ln\left(\frac{d_k}{d'}\right)\right]\left(\frac{1}{1-n}\right)\right\}, \quad \alpha = 1 \end{aligned} \tag{6}$$

3.5 Graph Energy

The energy (ε_G) of the graph G is defined on the eigenvalues of the adjacency matrix (Eq. 7). Here, λ_m is the eigenvalue of T, and the symbol $|\lambda_m|$ means the absolute value of λ_m (Gutman 1978).

$$\varepsilon_G = \sum_m |\lambda_m| \tag{7}$$

The graph energy can also be expressed as an integral form by the Coulson integral formula (Eq. 8), where $i^2 = -1$, and $\phi(x)$ is the characteristic polynomial of the graph G (Li et al. 2010).

$$\varepsilon_G = \frac{1}{\pi}\int_{-\infty}^{+\infty}\left[n - \frac{ix\phi'(ix)}{\phi(x)}\right]dx \tag{8}$$

In particular, the energy of the Erdős-Rényi's random graph can be estimated. In an Erdős-Rényi random graph, the links between the nodes are independent and the probability is p. Then, the energy of almost every Erdős-Rényi graph $G_n(p)$ satisfies the relationship as Eq. 9 (Li et al. 2010). Here, $o(1)$ represents a high-order infinitesimal of 1.

$$\varepsilon_G = n^{\frac{3}{2}}\left(\frac{8}{3\pi}\sqrt{p(1-p)} + o(1)\right) \tag{9}$$

3.6 The Rényi Index Based on the Eigenvalues

Dehmer Matthias et al. defined the graph entropy based on Rényi entropy, where α is a parameter (Eq. 10) (Dehmer et al. 2015). In addition, if $\overline{\lambda_\alpha} = \sum_{i=1}^{n} |\lambda_i|^\alpha$ is defined, Eq. 11 can be derived (Dehmer et al. 2015).

$$I_\alpha = \frac{1}{1-\alpha} log \left[\sum_{i=1}^{n} \left(\frac{|\lambda_i|}{\sum_{j=1}^{n} |\lambda_j|} \right)^\alpha \right], \quad \alpha \neq 1 \tag{10}$$

$$I_\alpha = \frac{1}{1-\alpha} log \frac{\overline{\lambda_\alpha}}{\varepsilon_G}, \quad \alpha \neq 1 \tag{11}$$

There is a direct relationship between definition Eq. 10 and the Rényi index. Since the adjacency matrix has positive and negative eigenvalues, a suitable choice is to define the Rényi index on $\Lambda_+ = \{|\lambda_i| | i = 1, \cdots, n\}$. Applying the general Rényi index to the set Λ_+, we can define the Rényi index of the graph as Eq. 12, where $\lambda' = \frac{1}{n} \sum_k |\lambda_k|$.

$$\begin{aligned} R_\lambda(\alpha) &= 1 - \left[\sum_{i=1}^{n} (\frac{|\lambda_i|}{\lambda'})^\alpha \left(\frac{1}{n} \right) \right]^{\frac{1}{1-\alpha}}, \quad \alpha \neq 1 \\ R_\lambda(1) &= 1 - \exp \left\{ - \sum_{i=1}^{n} \left[\frac{|\lambda_i|}{\lambda'} \ln \left(\frac{|\lambda_i|}{\lambda'} \right) \right] \left(\frac{1}{1-n} \right) \right\}, \quad \alpha = 1 \end{aligned} \tag{12}$$

In this way, we can define the Rényi index of the graph in two ways. One is defined on the degree sequence of the graph (Eq. 6), and the other is defined on the eigenvalues of the adjacency matrix of the graph (Eq. 12). Therefore, here we call the former one is the degree-based Rényi index and the latter is the eigenvalues-based Rényi index. In this paper, we only choose $\alpha = 2$ to calculate the Rényi index, which makes it possible to apply the Erdös-Rényi $(E - R)$ random graph theory directly. In addition, changes in parameters do not affect the results of the comparative analysis.

It can be proved that the relationship between I_α (Eq. 10) and $R_\lambda(a)$ (Eq. 12) is as shown in Eq. 13. Therefore, the two definitions are equivalent when both m and n are determined. In our discussion, m and n of the financial graph are fixed values, where the m values of *MST* and *PMFG* are equal to $n - 1$ and $3(n - 2)$, respectively. Eq. 13 is a normalized value that can be used to compare different networks. In this paper, we mainly use the form expressed by Eq. 13 for calculation.

$$R_\lambda(\alpha) = 1 - \frac{exp(I_\alpha)}{n}, \quad \alpha \neq 1 \tag{13}$$

We apply a property to simplify the Rényi index when $\alpha = 2$. The sum of the squares of the eigenvalues is twice the number of edges (Eq. 14) (Bapat 2010). Based on the relationship Eq. 14, if we combine Eqs. 11 and 13, $R_\lambda(2)$ can be expressed as Eq. 15. Eqaution 15 means that when $\alpha = 2$, $R_\lambda(2)$ is essentially normalized graph

energy. This indicates that the evenness of the eigenvalues $\Lambda_+ = \{|\lambda_i|, i = 1 \cdots n\}$ corresponds to the energy of the graph. Since $\varepsilon_G \leq \sqrt{2mn}$ (McClelland 1971), $R_\lambda(2)$ takes a value in the interval [0, 1]. Therefore $R_\lambda(2)$ can be considered as a normalized graph entropy.

$$\sum_{i=1}^{n} |\lambda_i|^2 = 2m \tag{14}$$

$$R_\lambda(2) = 1 - \frac{\varepsilon_G^2}{2mn} \tag{15}$$

3.7 Entropy of a Random Graph

Since the energy of the random graph has been appropriately estimated (Eq. 9), we can estimate the entropy of the random graph ($R_\lambda(2)$). If the term $o(1)$ in Eq. 9 is ignored, there is an approximate relation $\varepsilon_G \approx n^{\frac{3}{2}}(\frac{8}{3\pi}\sqrt{p(1-p)})$. Combining the definition of $R_\lambda(2)$ (Eq. 15), the approximate relationship of Eq. 16 can be obtained. There is an approximate relationship $\frac{n^2}{n(n-1)} \approx 1$ when n is a large integer ($lim_{n\to\infty}\frac{n^2}{n(n-1)} = 1$). The limit form of Eqs. 16 is 17. Approximately, for a random graph with a ratio p, there is an approximate relationship Eq. 18. Equation 18 shows that there is a linear relationship between $R_\lambda(2)$ and p with a slope of $\frac{64}{9\pi^2}$ (≈ 0.7205) and the intercept term is $1 - \frac{64}{9\pi^2}$ (≈ 0.2795).

$$\begin{aligned} R_\lambda(2) \approx & 1 - n^3 \frac{64}{9\pi^2} \cdot \frac{p(1-p)}{2mn} \\ R_\lambda(2) \approx & 1 - \frac{64n^2}{9\pi^2} \cdot \frac{1}{2m} \cdot \frac{2m}{n(n-1)} \cdot (1-p) \\ R_\lambda(2) \approx & 1 - \frac{64n^2}{9\pi^2} \cdot \frac{1}{n(n-1)} \cdot (1-p) \end{aligned} \tag{16}$$

$$lim_{n\to\infty} R_\lambda(2) = 1 - \frac{64}{9\pi^2} \cdot (1-p) \tag{17}$$

$$R_\lambda(2) \approx 1 - \frac{64}{9\pi^2} \cdot (1-p) \tag{18}$$

4 Results

4.1 The Rényi Entropy of Financial *MST* and *PMFG*

In this section, we analyze the differences between *MST* and *PMFG* in real markets and model-based financial graphs. First, we select the price data of the constituent

stocks of S&P500 to calculate *MST* (Fig. 1a) and *PMFG* (Fig. 1b). Second, we use two benchmark models (*GBM*, *OFM*) to simulate the return series and generate *MSTs* (Fig. 1c, e) and *PMFGs* (Fig. 1d, f)) for comparison.

Below, we separately analyze the distribution of the eigenvalues of the six financial graphs in Fig. 1. First, we generate the frequency histogram of Λ_+ for each graph as shown in Fig. 2. Comparing Fig. 2a with c, we find that the distribution of the latter is flatter and the maximum is smaller than the former. Similarly, the maximum eigenvalue in Fig. 2b is greater than the value in Fig. 2d. Furthermore, the difference between the average values of Λ_+ is small, as shown in the second and third columns of Table 1.

Then we analyze the Rényi index of the different graphs. According to previous research, we have known that the degree-based Rényi index of *MST* and *PMFG* based on real data is usually larger than the value based on geometric Brownian motion (Nie et al. 2016; Nie and Song 2018). Consistent with previous research, we found that the degree-based Rényi index of Fig. 1a is greater than that of Fig. 1c, and the analysis for Fig. 1b and d is similar.

Similar to the comparative analysis of $R(2)$, we find that the $R_\lambda(2)$ value of the real data is also significantly larger than the calculated value based on *GBM*, as shown in the fifth column of Table 1. From Eq. 14, if the values of m and n of the two graphs are equal, we already know that the smaller $R_\lambda(2)$ value corresponds to the larger ε_G. Therefore, the energy values of Fig. 1a and b are smaller than those of Fig. 1c and d, respectively.

Previous studies have shown that simulation data based on the one factor model can generate *MST* with super hub nodes (Bonanno et al. 2004; Nie et al. 2016). Figure 1e illustrates one such type of *MST*, which includes three hub nodes, resulting in a high $R(2)$ value. In particular, here we find that $R_\lambda(2)$ corresponding to Fig. 1e is also close to 1, which implies a high level of heterogeneity. Similarly, we find that the *PMFG* (Fig. 1f) based on the one factor model also shows an $R(2)$ value close to 1, and $R_\lambda(2) = 0.6175$, which indicates that the level of heterogeneity of its set Λ_+ is lower than that of network Fig. 1e. Figure 2f shows the set Λ_+ of Fig. 1f, where the maximum is significantly larger than the maximum in Fig. 2e (Table 1). This means that heterogeneity does not only depend on the maximum eigenvalue.

We can also calculate the value of $R_\lambda(2)$ directly from column 6 of Table 1 (Eq. 14). For example, we calculated $R_\lambda(2) = 0.4458$ based on the energy value of the network expressed in Fig. 1a ($R_\lambda(2) = 1 - 454.3052^2/(432 \times 431)$), which is consistent with the value calculated from the adjacency matrix in Table 1.

We can find that the heterogeneity of the eigenvalues of financial graphs in the real market is between the results based on geometric Brownian motion and the one factor model. That is, the model provides two extreme cases that make it possible to compare the spectra of different financial graphs.

4.2 Comparative Analysis of Financial Graphs and Random Graphs

In the previous section we have found that the energy and entropy of financial graphs in the market are significantly different from the calculated values provided

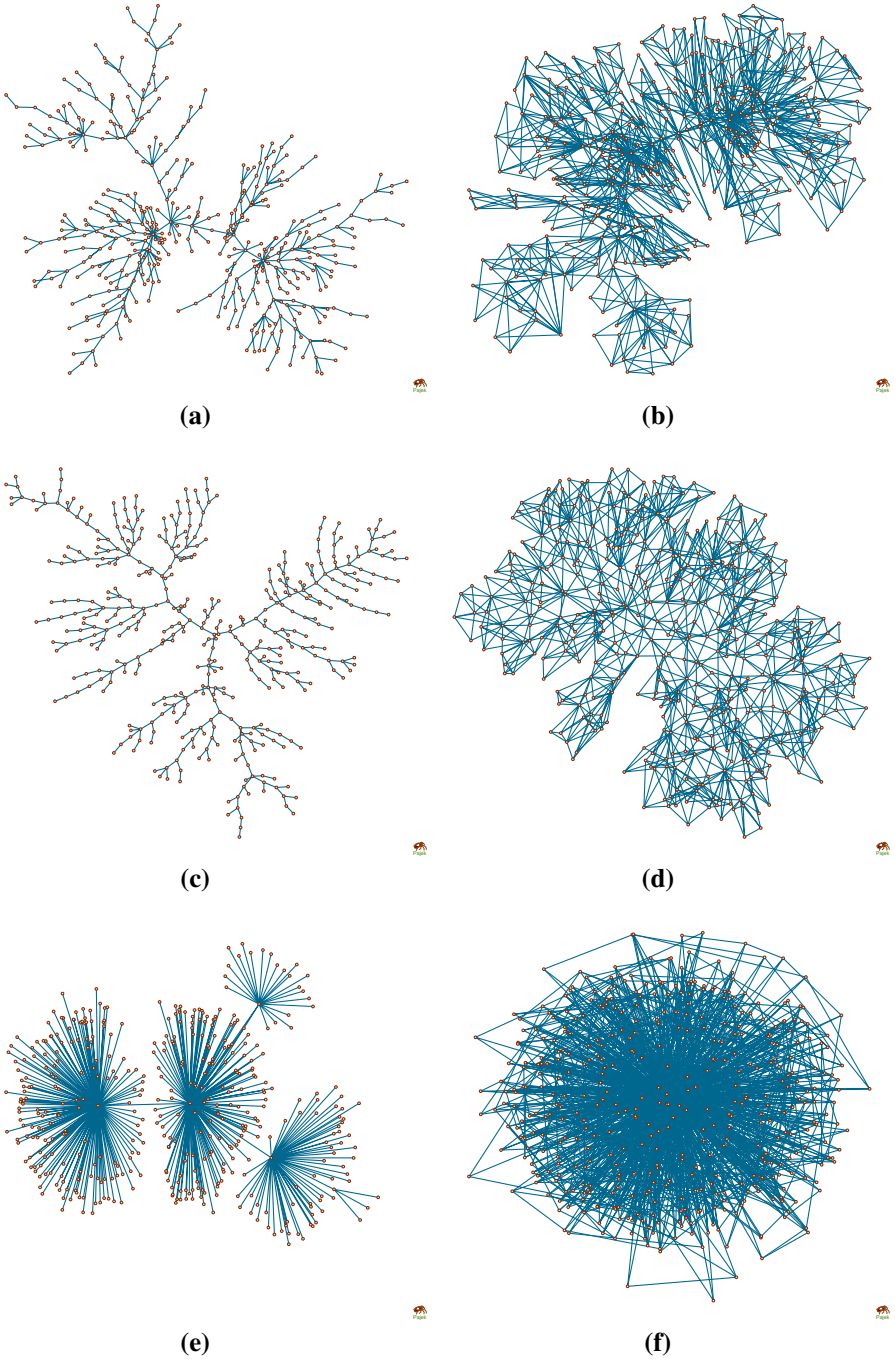

Fig. 1 The six sub-graphs show the financial graph of the real market and the financial graph generated by the models. **a**, **b** show the *MST* and *PMFG* of the real market. **c**, **d** show financial graphs of simulated data generated by geometric Brownian motion. **e**, **f** show financial graphs of simulated data generated by one factor model

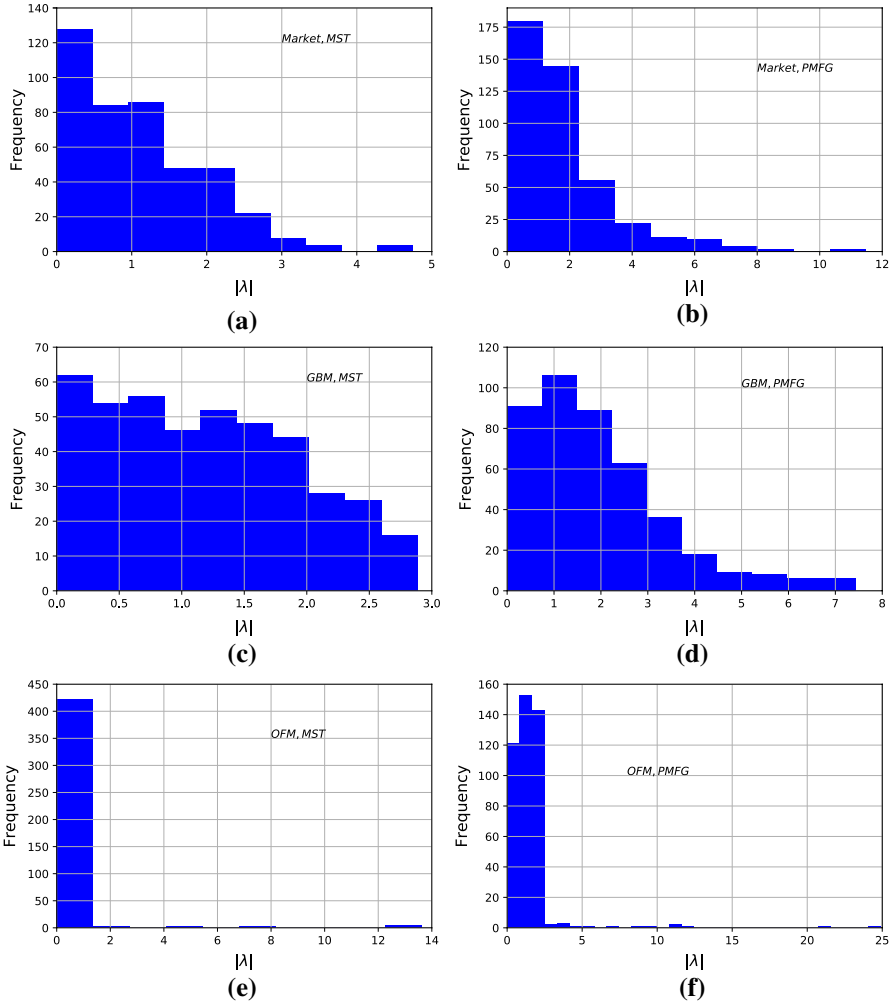

Fig. 2 The frequency histograms of the eigenvalues (Λ_+) of the adjacency matrices, wherein the six subgraphs correspond to the six graphs in Fig. 1, respectively

Table 1 The Rényi index and energy values of the six subgraphs in Fig. 1

Figure	$max(\Lambda_+)$	$mean(\Lambda_+)$	$R(2)$	$R_\lambda(2)$	ε_G
Figure 1a	4.7491	1.0516	0.4766	0.4458	454.3052
Figure 1b	11.4792	1.8188	0.4682	0.4461	785.7172
Figure 1c	2.8818	1.1752	0.2196	0.3078	507.6896
Figure 1d	7.4456	1.9508	0.1988	0.3628	842.7469
Figure 1e	13.6417	0.1920	0.9732	0.9815	82.9339
Figure 1f	24.8739	1.5114	0.9296	0.6175	652.9034

by the two benchmark models. In this section, we use the $E-R$ random graph as the benchmark.

The connection densities of *MST* and *PMFG* with 432 nodes are 0.0046 and 0.0139, respectively. We use $p = 0.0046$ and $p = 0.0139$ as parameters to generate two groups of graphs, each of which includes 1000 random graphs, and calculate the $R_\lambda(2)$ value of each graph, as shown in Fig. 3.

We find that the $R_\lambda(2)$ values of the random graphs corresponding to MST are significantly different from those estimated by Eqs. 1 and 2. However, the difference between the calculated value and the theoretical value of the random graph with the connection density of *PMFG* is small.

From the frequency histogram, it can be found that the $R_\lambda(2)$ values of the random graph are significantly different from the predicted values of Eq. 18, where the theoretical predicted values are 0.2828 $(1-\frac{64}{9\pi^2}\cdot(1-0.0046))$ and 0.2895 $(1-\frac{64}{9\pi^2}\cdot(1-0.0139))$. In addition, we find that the $R_\lambda(2)$ value of Fig. 1a is greater than $0.4451(ave+3\times std)$, which is statistically different from the entropy value of the random graph. Similarly, the $R_\lambda(2)$ value of Fig. 1b is greater than $0.3215(ave+3\times std)$.

In summary, we find that there is a large error between the predicted entropy value based on the random graph theory and the simulated value when the p value is small. The calculations in the next section show that the predicted values of Eq. 18 agree well with the calculated values with larger p values.

4.3 Rényi Entropy of the Financial Threshold Network

The structure of the threshold network is related to the threshold, so that we only focus on the relationship between the entropy and the threshold. Here, we specify a series of p_{th} values and calculate the network sequence $W_{p_{th}}$, and further study the relationship between $R_\lambda(2)$ and p_{th}. Figure 4a shows a network with $p_{th} = 0.1003$, which includes 9339 edges, and a network with a larger p_{th} value includes more links. Figure 4b shows a threshold network with the same parameter in the real market.

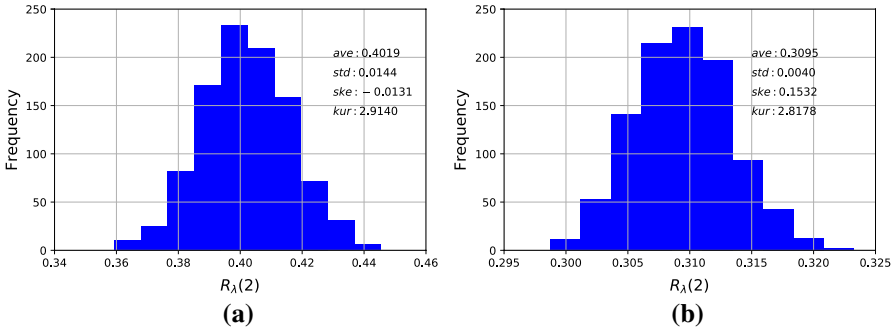

Fig. 3 **a** and **b** Show the $R_\lambda(2)$ values of 1000 random graphs, respectively, where **a** and **b** correspond to $p = 0.0046$ and $p = 0.0139$, respectively

We find differences between the structures of the two threshold networks. The network shown in Fig. 4a is a fully connected network. The threshold network in the real market includes a large component, which includes 338 nodes, as shown in Fig. 4b. Here we ignore the isolated nodes and small subgraphs. The $R_\lambda(2)$ value of the network Fig. 4a is equal to 0.3363, while the $R_\lambda(2)$ value of the threshold network in the real market is equal to 0.8356, which implies that the threshold network of the former is significantly different from the latter.

Next, we calculate the relationship between p_{th} value and Rényi index, as shown in Fig. 5. For comparison, we show the relationship between $R_\lambda(2)$ and p_{th} based on *GBM* in Fig. 5, where Rényi index and p_{th} are linear in the interval [0.05, 0.95]. When the threshold value p_{th} is in the interval [0.01, 0.1], the value of $R_\lambda(2)$ does not change significantly. If we extract the interval where $p_{th} \in [0.1, 0.8]$, we can find that Rényi and p_{th} can be well fitted by the line, where the equation is $R_\lambda(2) = 0.7001 * p_{th} + 0.2531$. The sum of the slope and intercept terms is 0.9531, which is close to the predicted value based on the random graph theory (Eq. 18). Comparing the solid line with the dotted line, we find that the $R_\lambda(2)$ value of the threshold network in the real market is different from the reference values based on *GBM*.

4.4 A Comparative Analysis of the Degree-Based Rényi Index and the Eigenvalues-Based Rényi Index

In the previous section, we have found that the spectrum of the financial graph of the real market is significantly different from the spectrum of the graph used for the benchmark, suggesting that the non-trivial structure present in the former leads to this difference. We next study the relationship between the degree-based Rényi index and the eigenvalues-based Rényi index. The former is an index describing the

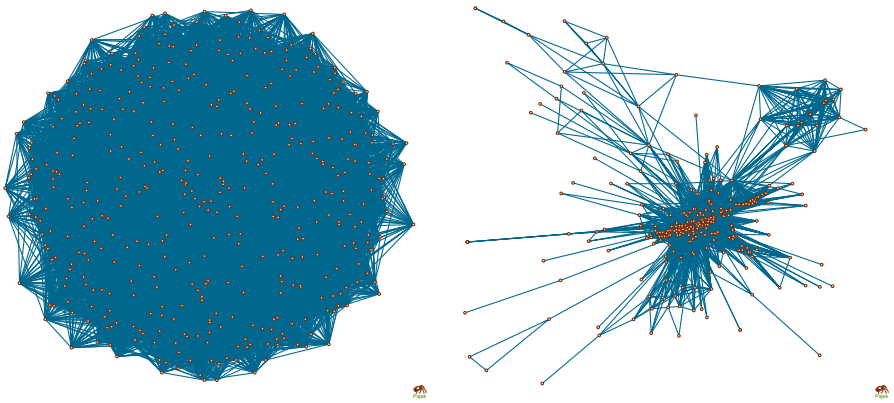

Fig. 4 **a** Shows the threshold network generated by geometric Brownian motion. **b** Shows the largest connected component in the threshold network of S&P500 constituents

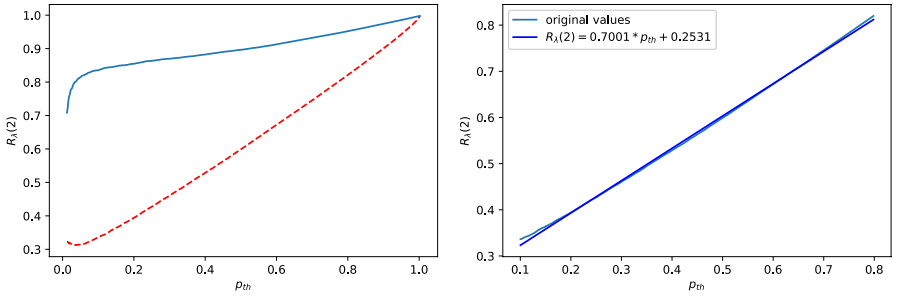

Fig. 5 The abscissa of the figure represents p_{th}, and the ordinate represents the $R_\lambda(2)$ value. The solid line in **a** shows the relationship between the $R_\lambda(2)$ value and p_{th} of the threshold network based on the S&P500 constituent stocks, and the dotted line corresponds to the relationship based on geometric Brownian motion. **b** Shows a part of the dotted line in **a**, where the p_{th} value belongs to the interval [0.1, 0.8]

topological structure, while the latter is an index defined on the eigenvalues. The results shown in Table 1 have suggested a correlation between $R(2)$ and $R_\lambda(2)$.

In this section, we examine the relationship between the two concepts in detail. We slide the time window to calculate the graph sequence so that multiple $R(2)$ and $R_\lambda(2)$ values can be calculated for comparative analysis. We use data from stocks in S&P500 index and specify a window length of 126 days. In addition, the sliding window is 5 days, so that a total of 611 *MSTs* or *PMFGs* are calculated. Then, we calculate the degree-based Rényi index and the eigenvalues-based Rényi index for each *MST* separately. A comparison of the two time series is shown in Fig. 6. Similarly, the Rényi index of *PMFG* can be analyzed, and the calculation result is shown in Fig. 7. Intuitively, a high correlation between the two pairs of time series can be observed. In fact, the Pearson correlation coefficients in the two figures are 0.8745 (Fig. 6) and 0.9381 (Fig. 7), respectively.

The high correlation between the entropy defined by two different approaches implies a deeper connection between the two concepts.

5 Discussion and Conclusions

5.1 Discussion

Similar to applying the random matrix theory to analyze the correlation matrix, here we use the random graph theory to analyze the correlation-based graph (network). Financial graphs show patterns that are different from those predicted by random graph theory, thereby suggesting that non-trivial structures can be analyzed using spectral theory.

At present, complex network theory has provided some insights based on the spectrum of adjacency matrix, such as analysis of community structure (Farkas et al. 2001; Nadakuditi and Newman 2013; Chauhan et al. 2009; Zhang et al. 2014; Jalan

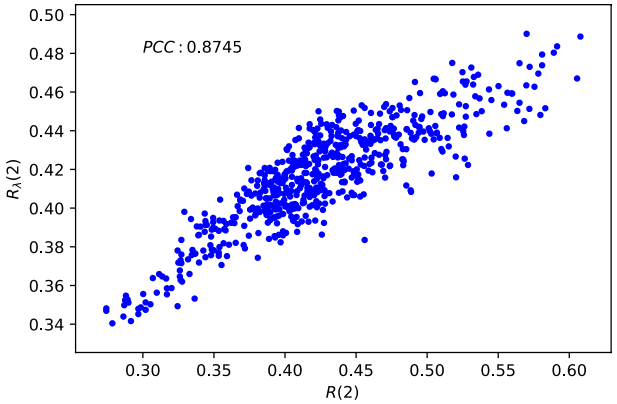

Fig. 6 The figure shows the relationship between $R_\lambda(2)$ and $R(2)$ of the *MST*

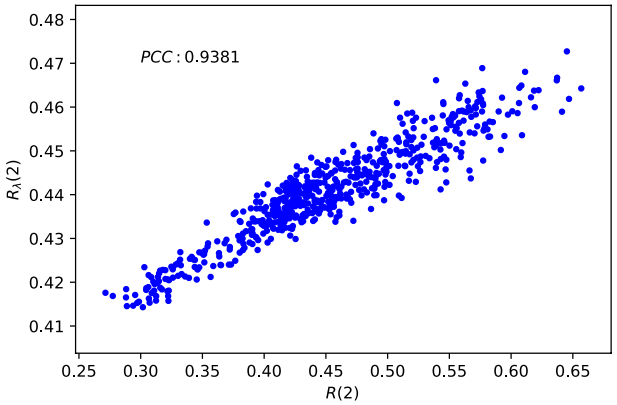

Fig. 7 The figure shows the relationship between $R_\lambda(2)$ and $R(2)$ of the *PMFG*

et al. 2011). In general, based on the spectrum of the network, we can analyze more properties related to the network structure. In this way, we can apply network theory to analyze financial networks in a broader sense, not just networks constructed from correlation matrices. In particular, for financial networks constructed from time series, the surrogate time series generated by classic financial models provide useful benchmarks, such as geometric Brownian motion. In addition, previous studies have shown that the analysis of the structure of financial graphs helps to discuss some classic topics (Zhao et al. 2016; Pozzi et al. 2013; Enguthaiwat 2018). A related topic is to discuss risk management and asset investment based on structural indicators of spectrum theory.

In this article, we report that there is a high correlation between Rényi indices defined by two different approaches. However, this is only a statistical correlation and does not mean that the more general cases involving non-planar graphs have similar conclusions. The spectrum of the adjacency matrix contains a lot of

information related to the network structure, and the Rényi index only describes one of its global characteristics. A detailed analysis of the eigenvalues of the financial graph may reveal the structure of the financial network more deeply.

5.2 Conclusions

We study the entropy of financial graphs and its relationship with topological structure, and find that there is a high correlation between the eigenvalue-based Rényi index and the degree-based Rényi index. We use random graphs as a benchmark model and find that the entropy of financial graphs in the real market is clearly different from theoretical predictions. In addition, calculations show that the geometric Brownian motion and the one factor model provide two extreme cases that allow comparative analysis of financial graphs in the market. The extreme cases provided by the models allow us to observe the deviation level of a financial graph from the model graph.

In summary, our research shows that there is a correlation between the eigenvalue distribution of the adjacency matrix and the topological structure. This also shows that the entropy based on eigenvalues can be used as an a useful indicator to characterize the properties of financial graphs. In addition, our research provides a new perspective that is different from the entropy of studying a single time series, and focuses on the entropy of the correlation structure, which helps to analyze the complexity of the graph structure in the market.

References

Bapat, R. B. (2010). *Graphs and matrices* (p. 77). New York: Springer.

Billio, M., Getmansky, M., Lo, A. W., & Pelizzon, L. (2012). Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *Journal of Financial Economics*, *104*(3), 535–559.

Boginski, V., Butenko, S., & Pardalos, P. M. (2005). Statistical analysis of financial networks. *Computational Statistics & Data Analysis*, *48*(2), 431–443.

Bonanno, G., Caldarelli, G., Lillo, F., & Mantegna, R. N. (2003). Topology of correlation-based minimal spanning trees in real and model markets. *Physical Review E*, *68*(4), 046130.

Bonanno, G., Caldarelli, G., Lillo, F., Miccichè, S., Vandewalle, N., & Mantegna, R. N. (2004). Networks of equities in financial markets. *The European Physical Journal B*, *38*, 363–371.

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The econometrics of financial markets* (pp. 219–252). Princeton: Princeton University Press.

Chauhan, S., Girvan, M., & Ott, E. (2009). Spectral properties of networks with community structure. *Physical Review E*, *80*, 056114.

Conlon, T., Ruskin, H. J., & Crane, M. (2009). Cross-correlation dynamics in financial time series. *Physica A: Statistical Mechanics and its Applications*, *388*, 705–714.

Dai, Y. H., Xie, W. J., Jiang, Z. Q., Jiang, G. J., & Zhou, W. X. (2016). Correlation structure and principal components in the global crude oil market. *Empirical Economics*, *51*(4), 1501–1519.

Dehmer, M., Li, X., & Shi, Y. (2015). Connections between generalized graph entropies and graph energy. *Complexity*, *21*, 35–41.

Eliazar, I. (2011). Randomness, evenness, and rényi's index. *Physica A: Statistical Mechanics and its Applications*, *390*(11), 1982–1990.

Eliazar, I. I., & Sokolov, I. M. (2012). Measuring statistical evenness : A panoramic overview. *Physica A: Statistical Mechanics and its Applications*, *391*(4), 1323–1353.

Enguthaiwat, H. (2018). Stock market return predictability: Does network topology matter? *Review of Quantitative Finance and Accounting*, *51*(2), 433–460.

Eom, C., Oh, G., Jung, W. S., Jeong, H., & Kim, S. (2009). Topological properties of stock networks based on minimal spanning tree and random matrix theory in financial time series. *Physica A*, *388*, 900–906.

Farkas, I. J., Derényi, I., Barabási, A. L., & Vicsek, T. (2001). Spectra of "real-world" graphs: Beyond the semicircle law. *Physical Review E*, *64*, 026704.

Gang-Jin, W., & Chi, X. (2016). Tail dependence structure of the foreign exchange market: A network view. *Expert Systems with Applications*, *46*, 164–179.

Gang-Jin, W., Chi, X., Kaijian, H., & Stanley, H. E. (2017). Extreme risk spillover network: Application to financial institutions. *Quantitative Finance*, *17*(9), 1417–1433.

Gutman, I. (1978). The energy of a graph. *Ber Math Stat Sekt Forschungsz Graz*, *103*, 1–22.

Hartman, D., & Hlinka, J. (2018). Nonlinearity in stock networks. Chaos: An Interdisciplinary Journal of Nonlinear Science 28(8):083127.

Hautsch, N., Schaumburg, J., & Schienle, M. (2014). Financial network systemic risk contributions. *Review of Finance*, *19*(2), 685–738.

Hua, W. F., Xin, Y., & Wei-Xing, Z. (2019). Tail dependence networks of global stock markets. *International Journal of Finance & Economics*, *24*(1), 558–567.

Jalan, S., Zhu, G., & Li, B. (2011). Spectral properties of directed random networks with modular structure. *Physical Review E*, *84*, 046107.

Jiang, X. F., Chen, T. T., & Zheng, B. (2014). Structure of local interactions in complex financial dynamics. *Scientific Reports*, *4*, 05321.

Jiang, Z. Q., Xie, W. J., Zhou, W. X., & Sornette, D. (2019). Multifractal analysis of financial markets: A review. *Reports on Progress in Physics*, *82*(12), 125901.

Junior, L. S., & Franca, I. D. P. (2012). Correlation of financial markets in times of crisis. *Physica A*, *391*, 187–208.

Kenett, D. Y., Tumminello, M., Madi, A., Gur-Gershgoren, G., Mantegna, R. N., & Ben-Jacob, E. (2010a). Dominating clasp of the financial sector revealed by partial correlation analysis of the stock market. *PloS One*, *5*(12), e15032.

Kenett, D. Y., Tumminello, M., Madi, A., Gur-Gershgoren, G., Mantegna, R. N., & Ben-Jacob, E. (2010b). Dominating clasp of the financial sector revealed by partial correlation analysis of the stock market. *PLoS ONE*, *5*(12), e15032.

Kruskal, J. B. (1956). On the shortest spanning subtree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical society*, *7*(1), 48–50.

Kumar, S., & Deo, N. (2012). Correlation and network analysis of global financial indices. *Physical Review E*, *86*, 026101.

Laloux, L., Cizeau, P., Bouchaud, J. P., & Potters, M. (1999). Noise dressing of financial correlation matrices. *Physical Revie Letter*, *83*(7), 1467–1470.

Li, X., Shi, Y., & Gutman, I. (2010). *Graph Energy*. Berlin: Springer.

Mantegna, R. N. (1999). Hierarchical structure in financial markets. *The European Physical Journal B*, *11*, 193–197.

McClelland, B. J. (1971). Properties of the latent roots of a matrix: The estimation of π - electron energies. *The Journal of Chemical Physics*, *54*, 640–643.

Meng, H., Xie, W. J., Jiang, Z. Q., Podobnik, B., Zhou, W. X., & Stanley, H. E. (2014). Systemic risk and spatiotemporal dynamics of the US housing market. *Scientific Reports*, *4*, 3655.

Nadakuditi, R. R., & Newman, M. E. J. (2013). Spectra of random graphs with arbitrary expected degrees. *Physical Review E*, *87*, 012803.

Newman, M. E. J. (2019). spectra of networks containing short loops. *Physical Review E*, *100*, 012314.

Nie, C. X., & Song, F. T. (2018). Relationship between entropy and dimension of financial correlation-based network. *Entropy*, *20*(3), 177.

Nie, C. X., Song, F. T., & Li, S. P. (2016). Rényi indices of financial minimum spanning trees. *Physica A: Statistical Mechanics and its Applications*, *444*, 883–889.

Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, L. A. N., & Stanley, H. E. (1999). Universal and nonuniversal properties of cross correlations in financial time series. *Physical Revie Letter*, *83*(7), 1471–1474.

Pozzi, F., Matteo, T. D., & Aste, T. (2013). Spread of risk across financial markets: Better to invest in the peripheries. *Scientific Reports*, *3*(1), 1665–1665.

Prim, R. C. (1957). Shortest connection networks and some generalizations. *Bell System Technical Journal*, *36*(6), 1389–1401.

Rényi, A. (1961). On measures of entropy and information. In *Proceedings of the fourth berkeley symposium on mathematical statistics and probability Vol. 1: Contributions to the Theory of Statistics* (Vol. 1, pp. 547–561).

Sarkar, C., Jalan, S. (2018). Spectra of networks. arXiv preprint arXiv:1810.01254

Song, D. M., Tumminello, M., Zhou, W. X., & Mantegna, R. N. (2011). Evolution of worldwide stock markets, correlation structure, and correlation-based graphs. *Physial Review E*, *84*(2), 026108. https://doi.org/10.1103/PhysRevE.84.026108.

Tse, C. K., Liu, J., & Lau, F. C. M. (2010). A network perspective of the stock market. *Journal of Empirical Finance*, *17*(4), 659–667.

Tumminello, M., Aste, T., Matteo, T. D., & Mantegna, R. N. (2005). A tool for filtering information in complex systems. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 10421–10426.

Vandewalle, N., Brisbois, F., & Tordoir, X. (2001). Non-random topology of stock markets. *Quantitative Finance*, *1*(3), 372–374.

Wang, G. J., Xie, C., & Stanley, H. E. (2018). Correlation structure and evolution of world stock markets: Evidence from pearson and partial correlation-based networks. *Computational Economics*, *51*(3), 607–635.

Wiliński, M., Sienkiewicz, A., Gubiec, T., Kutner, R., & Struzik, Z. R. (2013). Structural and topological phase transitions on the german stock exchange. *Physica A*, *392*(23), 5963–5973.

Wiliński, M., Szewczak, B., Gubiec, T., Kutner, R., & Struzik, Z. (2015). Temporal condensation and dynamic λ-transition within the complex network: An application to real-life market evolution. *The European Physical Journal B*, *88*(2), 1.

Yang, Y., & Yang, H. (2008). Complex network-based time series analysis. *Physica A: Statistical Mechanics and its Applications*, *387*(5–6), 1381–1386.

Yi, S., Xu, Z., & Wang, G. J. (2018). Volatility connectedness in the cryptocurrency market: Is bitcoin a dominant cryptocurrency? *International Review of Financial Analysis*, *60*, 98–114.

Zhang, X., Nadakuditi, R. R., & Newman, M. E. J. (2014). Spectra of random graphs with community structure and arbitrary degrees. *Physical review E*, *89*, 042816.

Zhao, L., Li, W., & Cai, X. (2016). Structure and dynamics of stock market in times of crisis. *Physics Letters A*, *380*(5–6), 654–666.