Physica A 507 (2018) 499–509

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Collective behavior of cryptocurrency price changes

Darko Stosic [a,*], Dusan Stosic [a], Teresa B. Ludermir [a], Tatijana Stosic [b]

[a] Centro de Informática, Universidade Federal de Pernambuco , Av. Luiz Freire s/n, 50670-901, Recife, PE, Brazil
[b] Departamento de Estatística e Informática, Universidade Federal Rural de Pernambuco, Rua Dom Manoel de Medeiros s/n, Dois Irmãos, 52171-900, Recife, PE, Brazil

HIGHLIGHTS

- We analyze cross correlations between price changes of different cryptocurrencies.
- They exhibit non-trivial groupings not present in the partial cross correlations.
- Most eigenvalues do not agree with the universal predictions of random matrix theory.
- We discover rather distinct community structures in their minimum spanning trees.
- Collective behaviors in cryptocurrency market differ from other financial markets.

ARTICLE INFO

Article history:
Received 7 February 2018
Received in revised form 11 April 2018

Keywords:
Correlation matrix
Partial correlation matrix
Random matrix theory
Minimum spanning tree
Cryptocurrencies
Bitcoin

ABSTRACT

Digital assets termed cryptocurrencies are correlated. We analyze cross correlations between price changes of different cryptocurrencies using methods of random matrix theory and minimum spanning trees. We find that the cross correlation matrix exhibits non-trivial hierarchical structures and groupings of cryptocurrency pairs, which are not present in the partial cross correlations. In sharp contrast to the predictions for other financial markets, we discover that most of the eigenvalues in the spectrum of the cross correlation matrix do not agree with the universal predictions of random matrix theory, but the few of the largest eigenvalues deviate as expected. The minimum spanning tree of cryptocurrency cross correlations reveals distinct community structures that are surprisingly stable. Collective behaviors that are present in the cryptocurrency market can be useful for the construction of portfolio of cryptocurrencies as well as for future research on the subject.

1. Introduction

The collective behavior of a complex system manifests in the correlations among its constituents. Financial markets are complex systems that have been extensively studied by physicists using concepts and methods developed to describe physical systems [1–5]. The study of correlations between different financial assets such as stocks and commodities is a topic of interest not only for scientific reasons of understanding the economy as a complex dynamical system, but also for practical reasons such as quantifying risk of investment portfolios [6–8]. Since market conditions are not always stationary and historical data is finite, an important question is whether the correlations between two financial time series are due only to omnipresent dynamical noise [9,10], or arise from genuine interactions [11]. The difficulty in answering this question lies in the fact that although financial assets should interact directly or indirectly with each other, the precise nature of their interactions remains unknown.

* Corresponding author.
E-mail addresses: ddstosic@bu.edu (D. Stosic), dbstosic@bu.edu (D. Stosic), tbl@cin.ufpe.br (T.B. Ludermir), tastosic@gmail.com (T. Stosic).

https://doi.org/10.1016/j.physa.2018.05.050
0378-4371/

In physical systems, one relates correlations to basic microscopic interactions, as in the case of a magnet with exchange interactions dictating the behavior of neighboring spins. However, the situation is quite different for financial systems as the underlying 'interactions' are not known [11,12]. The problem of quantifying correlations between different financial assets is similar to that of an N-body system [13]: one calculates the cross correlation matrix whose elements are the correlation coefficients between the observables of individual particles pairs. The corresponding eigenvalues (and eigenvectors) of the matrix then convey information about the collective behavior of the system. This leads to a plethora of methods that can be used for analyzing their correlations, ranging from identifying non-random properties of the system through deviations from universal predictions of random matrix theory [11,14–17] to searching for community structures in networks constructed from the correlations coefficients [18–23].

The correlations of financial markets have been extensively studied in literature [5,14,18,24–26] revealing important discoveries such as collective market behavior, or presence of correlations that pervade the entire system [11,12,15], and strong correlations localized between components within a given business sector [27]. Over the past few years digital assets termed cryptocurrencies have emerged creating new paradigms for economic transactions and alternative means of capital, with profound implications to central financial markets. These electronic cash systems introduce an entirely new technology, called the blockchain, to process and check transactions with no overseeing authority, but rather with all participants of the peer-to-peer network [28]. The market consists of over one thousand cryptocurrencies that share the same underlying blockchain technology, either as clones of Bitcoin or with significant technological innovations, and most of them are on isolated transaction networks [29]. Moreover, the cryptocurrency market is unique and increasingly complex compared to other financial markets. Part of the reason are the multiple usages and distinct nature of individual cryptocurrencies. They can interact with their own blockchain, or can be built on top of another blockchain to process transactions. Crowd funding centered around a cryptocurrency can be done in terms of initial coin offerings, as a source of capital for new companies. There are also cryptocurrencies with no value of proposition and possibly scams. Lastly, cryptocurrencies compete with each other for market capitalization, which further indicates the complex nature of their interactions. An important question that remains to be answered is whether correlations in the market of cryptocurrencies exhibit similar properties to those found for other financial markets.

The purpose of this work is to quantify the collective behavior among constituents in the cryptocurrency market. Here, we analyze cross correlations between price changes of 119 publicly traded cryptocurrencies in the time period from August 26, 2016 to January 18, 2018. After calculating the correlation matrix, we apply concepts and methods from random matrix theory and minimum spanning trees to investigate hierarchical structures that are indicative of collective behavior in the market of cryptocurrencies. The paper is organized as follows. Section 2 describes the data. Section 3 analyzes the matrix of cross correlations and partial cross correlations. Section 4 compares cryptocurrency correlations against universal properties of random matrix theory. Section 5 represents cross correlations of the cryptocurrency market as a minimum spanning tree and analyzes its topological properties. Section 6 draws the conclusions.

2. Data analyzed

To investigate collective behavior in the cryptocurrency market, we consider daily closing prices of the cryptocurrencies listed in https://coinmarketcap.com/. The website calculates prices by taking the volume weighted average of values reported for different exchange markets, where there will be missing data if the cryptocurrency was not exchanged on a given day. Therefore, to ensure simultaneity among the time series we remove days where the price for at least one cryptocurrency is missing. This can be interpreted as calculating correlations based on random measurements of cryptocurrency prices. We process the time series by selecting a set of cryptocurrencies with highest market capitalization (e.g., $N \geq 400$) and with a minimum number of trading days (e.g., $L \geq 500$). Our processing of the data leads to a total of $N = 119$ cryptocurrencies of $L = 200$ simultaneous days in the time period considered from August 26, 2016 to January 18, 2018 as listed in Table 1. The correlations can then be viewed as the statistical interactions between over one hundred cryptocurrencies based on roughly one year of random measurements. We achieved similar results for other combinations of cryptocurrencies and time series lengths.

3. Statistics of cross correlations and partial cross correlations

In order to quantify correlations, we calculate the price change (or "return") of cryptocurrency $i = 1, \ldots, N$ over a time scale Δt,

$$G_i(t) \equiv \ln S_i(t + \Delta t) - \ln S_i(t), \tag{1}$$

where $S_i(t)$ denotes the price of cryptocurrency i and we take $\Delta t = 1$ day. Different stocks have varying levels of volatility (or standard deviation), which makes them difficult to compare. Therefore, we consider a normalized return

$$g_i(t) \equiv \frac{G_i(t) - \langle G_i \rangle}{\sigma_i}, \tag{2}$$

Table 1
List of cryptocurrencies ranked by market capitalization.

$^{\text{rank}}$Cryptocurrency					
^{52}Aeon	^{50}CloakCoin	^{23}Factom	^{81}Mintcoin	^{110}PinkCoin	^{93}Steem Dollars
^{47}Agoras Tokens	^{42}Counterparty	^{68}FairCoin	^{27}MonaCoin	^{22}PIVX	^{15}Steem
^{109}ArtByte	^{62}Crown	^{59}Feathercoin	^{8}Monero	^{65}PotCoin	^{6}Stellar
^{99}AsiaCoin	^{7}Dash	^{91}FedoraCoin	^{86}MonetaryUnit	^{118}Primecoin	^{13}Stratis
^{19}Augur	^{21}Decred	^{92}FlorinCoin	^{51}Mooncoin	^{53}Pura	^{69}Synereo
^{80}Bean Cash	^{63}Diamond	^{100}FoldingCoin	^{79}Myriad	^{83}Radium	^{28}Syscoin
^{41}BitBay	^{20}DigiByte	^{101}Gambit	^{57}Namecoin	^{25}ReddCoin	^{11}Tether
^{67}bitCNY	^{32}DigitalNote	^{31}GameCredits	^{36}NavCoin	^{3}Ripple	^{39}Ubiq
^{1}Bitcoin	^{26}DigixDAO	^{102}GeoCoin	^{5}NEM	^{49}Rise	^{117}UFO Coin
^{34}BitcoinDark	^{73}Dimecoin	^{104}Global Curr. Resv.	^{87}NeosCoin	^{84}Rubycoin	^{106}Unobtanium
^{105}Bitcrystals	^{17}Dogecoin	^{77}GridCoin	^{88}Neutron	^{94}RussiaCoin	^{112}Vcash
^{108}BitSend	^{115}E-Dinar Coin	^{56}Groestlcoin	^{74}NewYorkCoin	^{58}Safe Exch. Coin	^{14}Verge
^{16}BitShares	^{114}EarthCoin	^{43}Gulden	^{24}Nexus	^{60}Salus	^{75}VeriCoin
^{95}bitUSD	^{44}Einsteinium	^{48}HempCoin	^{71}NuShares	^{55}Shift	^{37}Vertcoin
^{72}BlackCoin	^{33}Emercoin	^{66}Ion	^{30}Nxt	^{12}Siacoin	^{45}Viacoin
^{38}Blocknet	^{97}Energycoin	^{76}Jinn	^{111}OBITS	^{70}SIBcoin	^{54}Voxels
^{89}Boolberry	^{113}Espers	^{90}LEOcoin	^{82}OKCash	^{61}SolarCoin	^{18}Waves
^{46}Burst	^{9}Ethereum Classic	^{10}Lisk	^{85}Omni	^{116}Sphere	^{64}WhiteCoin
^{98}Circuits of Value	^{2}Ethereum	^{4}Litecoin	^{35}PACcoin	^{119}Sprouts	^{96}Xaurum
^{103}Clams	^{78}Expanse	^{29}MaidSafeCoin	^{40}Peercoin	^{107}Stealthcoin	

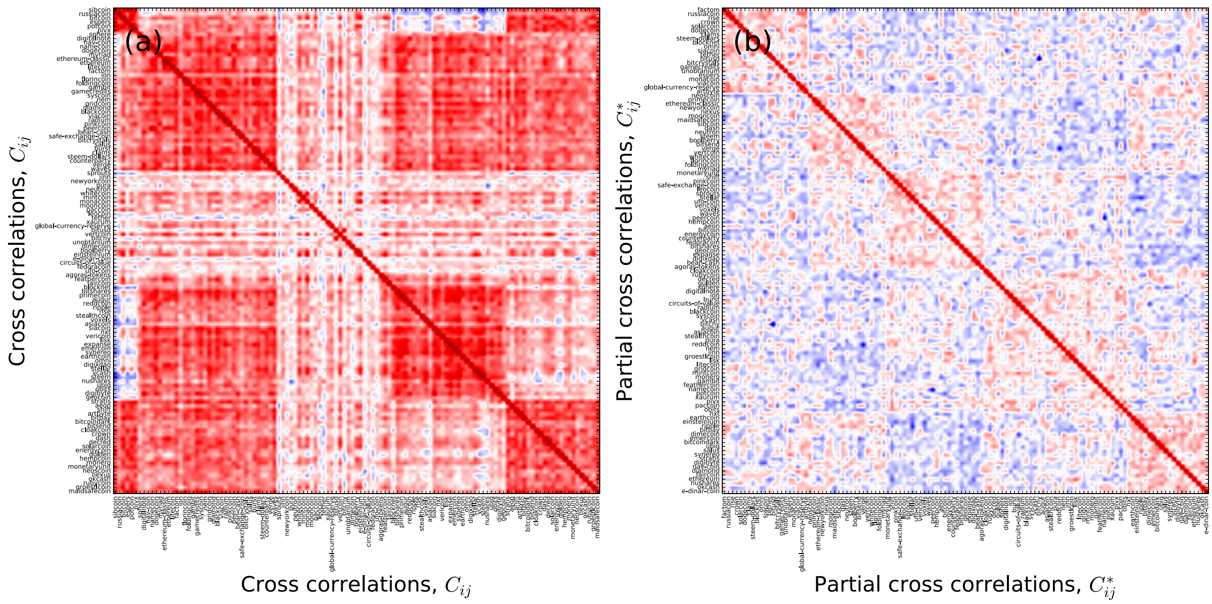

Fig. 1. Visual representation of (a) cross correlations matrix C and (b) partial cross correlations matrix C^* constructed from 119 cryptocurrencies over the two-year period 2016–2018. Colors represent the magnitude of the correlation coefficients ranging from strongly anti-correlated (blue) to uncorrelated (white) to strongly correlated (red). The elements in the matrix are interpolated using nearest neighbors. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

where $\sigma_i \equiv \sqrt{\langle G_i^2 \rangle - \langle G_i \rangle^2}$ is the standard deviation of G_i and $\langle \cdots \rangle$ denotes a time average over the period studied. We then compute the equal-time cross correlation matrix C with elements

$$C_{ij} \equiv \langle g_i(t) g_j(t) \rangle. \tag{3}$$

By construction the elements $C_{ij} = 1$ correspond to completely correlated, $C_{ij} = -1$ correspond to completely anti-correlated, and $C_{ij} = 0$ correspond to uncorrelated pairs of cryptocurrencies.

If a pair of cryptocurrencies i and j is correlated with other elements in the market, then the correlation C_{ij} between them may introduce spurious information. For example, this would be typical of cryptocurrencies that trade on another blockchain. The misleading information can be avoided by considering the partial correlations, which quantify the true correlation between any two variables by measuring the degree of association between them, and removing the effects of all other variables. In order to quantify the partial correlations [30], we first invert the cross correlation matrix, $C' = C^{-1}$.

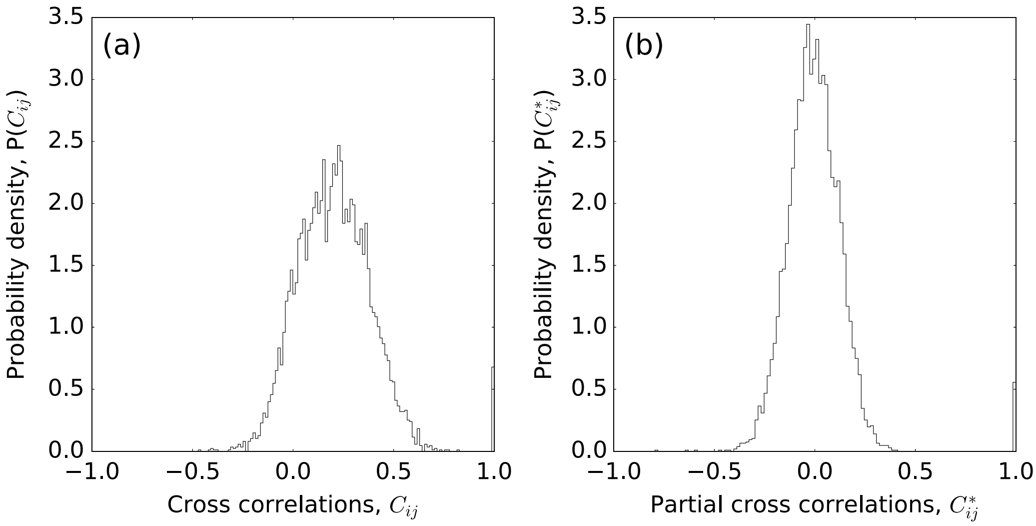

Fig. 2. Probability density function of (a) cross correlations coefficients C_{ij} and (b) partial cross correlations coefficients C^*_{ij}.

We then compute the partial cross correlation matrix C^* with elements

$$C^*_{ij} \equiv -\frac{C'_{ij}}{C'_{ii}C'_{jj}}, \tag{4}$$

where C'_{ij} is an element of C' and C^*_{ij} is the correlation between cryptocurrencies i and j when all other influences from the elements in C are removed.

The rows and columns (or cryptocurrency pairs) of the cross correlation matrix C and partial cross correlation matrix C^* are grouped using a spectral block clustering procedure [31]. Fig. 1(a) reveals that a complex arrangement of positive cross correlations ($C_{ij} > 0$) exists between the different cryptocurrencies. The matrix consists of blocks of strongly correlated cryptocurrencies, mostly along the early columns of C, but also contains uncorrelated features ($C_{ij} \approx 0$) in unexpected regions, such as in the final rows of C. Moreover, strong anti-correlations ($C_{ij} < 0$) appear between very specific cryptocurrency pairs. Lastly, a set of cryptocurrencies, amounting to one-fifth of the market including Tether which is pegged to the Dollar, appears to be mostly uncorrelated with the rest of the market. This implies the presence of non-trivial hierarchical structures and groupings in the correlations. Fig. 1(b) on the other hand reveals a very distinct situation. Here most of the hierarchical structures are lost, and anti-correlations appear to dominate the elements of C^*. A possible interpretation is that cryptocurrency pairs are significantly influenced by others, indicating a form of collective behavior in the market of cryptocurrencies.

We next investigate the probability density functions $P(C_{ij})$ and $P(C^*_{ij})$ of the elements of C and C^*, respectively. Fig. 2 shows that both cross correlations and partial cross correlations are normally distributed. On average, $P(C_{ij})$ is centered at $\langle C_{ij}\rangle > 0$ meaning that positive correlations are dominant among cryptocurrency pairs. $P(C^*_{ij})$ indicates that partial interactions between cryptocurrency prices are mostly anti-correlated or uncorrelated. These results confirm our previous findings from the visual representations of C and C^*.

4. Deviations from universal properties of random matrix theory

By construction the elements of C are supposed to express correlations that exist in the system, but in practice their meaning is not clear because of non-stationarity and of finite lengths involved in their calculation. For these reasons empirical cross correlations contain random contributions (or "noise") that is difficult to distinguish from genuine interactions. The question then arises how can we extract from C the cross correlations that are significant. The problem of interpreting the correlations between cryptocurrency prices is reminiscent of the difficulties in interpreting the spectra of complex nuclei in the 1950s [32–34]. Random matrix theory (RMT) was developed in this context in order to explain the statistics of energy levels of complex quantum systems in nuclear physics [32]. RMT predictions represent an average over all possible interactions [32,33]. Deviations from their universal predictions reveal non-random properties that are specific to the system and arise from the presence of collective behavior [35,36]. Since RMT was introduced in financial markets [5,11,14], the method has been used extensively to investigate the statistical properties of cross-correlations in different financial markets, including stock markets [17,25,37,38], the housing market [39], and the foreign exchange market [26]. Time-lag generalizations of RMT were recently explored [40–42] due to the importance of the lagged correlations. The idea was to

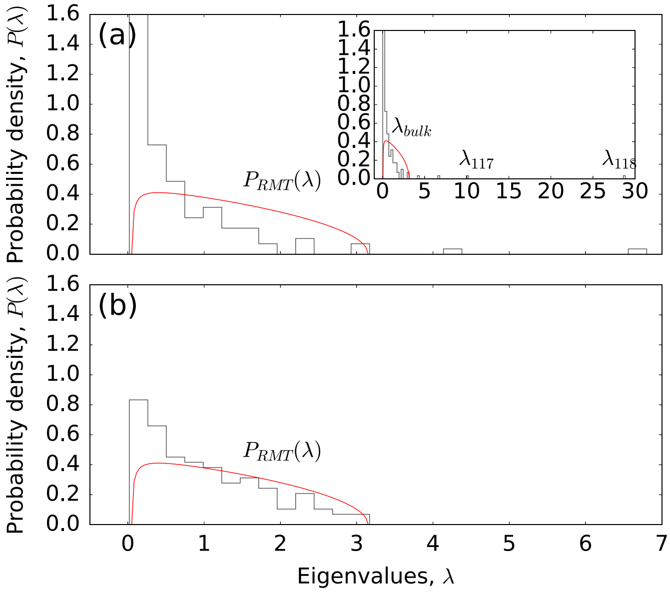

Fig. 3. (a) Eigenvalue distribution $P(\lambda)$ for the cross correlation matrix C. The inset shows the largest deviating eigenvalues $\lambda > \lambda_+$. (b) Eigenvalue distribution for the random correlation matrix R computed from $N = 119$ uncorrelated time series with length $L = 200$. The solid curve shows the RMT prediction $P_{RMT}(\lambda)$ from (7).

analyze the correlations at different instances in time, where both short-range and long-range cross correlations have been reported for a wide range of complex systems [16,24].

Since our aim is to extract information about cross correlations of the cryptocurrency market, we can compare C with a random cross correlation matrix [43]. The correlation matrix can be written as

$$C = \frac{1}{L} G G^T, \tag{5}$$

where G is an $N \times L$ matrix with elements $g_{ik} \equiv g_i(k\Delta t)$ for $i = 1, \dots, N$ and $k = 0, \dots, L$. We can therefore consider a random correlation matrix of the form

$$R = \frac{1}{L} A A^T, \tag{6}$$

where A is an $N \times L$ matrix containing N time series of L random elements with zero mean and unit variance. Statistical properties of random matrices such as R are established in literature [43]. For $N \to \infty$ and $L \to \infty$ such that $Q = L/N$, the probability density function $P_{RMT}(\lambda)$ of eigenvalues λ of the random correlation matrix R is given by

$$P_{RMT}(\lambda) = \frac{Q}{2\pi} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda}, \tag{7}$$

for λ within the bounds $\lambda_- \le \lambda_i \le \lambda_+$, where λ_- and λ_+ represent the minimum and maximum eigenvalues given by

$$\lambda_\pm = 1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}}. \tag{8}$$

$P_{RMT}(\lambda)$ has a rapidly decaying edge instead of an abrupt cutoff for finite L and N [44].

In analogy to the analysis of correlations in physical systems, we study the statistical properties of C and compare them against the universal properties of random matrix theory [35,36]. We first diagonalize the cross correlations matrix C and rank-order its eigenvalues λ_k such that $\lambda_{k+1} > \lambda_k$ with eigenvectors u^k. Next, we calculate the eigenvalue distribution and compare it with analytical results for a cross correlation matrix R generated from finite uncorrelated time series [43]. For $N = 119$ and $L = 200$ the theoretical eigenvalues bounds are $\lambda_- = 0.05$ and $\lambda_+ = 3.14$ and the value of Q is equal to $L/N = 1.68$. Fig. 3(a) shows that the “bulk” of the eigenvalues of C $\lambda_i \in \lambda_{bulk}$ falls within the bounds $\lambda_- < \lambda_{bulk} < \lambda_+$ for $P_{RMT}(\lambda)$. However, their probabilities deviate significantly from theoretical predictions with most of the eigenvalues located below the lower bound $\lambda_- = 0.05$. This indicates there is a clear dominance for small eigenvalues. Moreover, the fact that λ_{bulk} does not necessary follow RMT suggests that the cryptocurrency market exhibits markedly different correlations from stocks markets [12,45] and currencies [46]. We also find deviations from RMT on the upper bound $\lambda_+ = 3.15$ for the largest few eigenvalues. The largest eigenvalue is particularly interesting, which is approximately 9 times larger than the

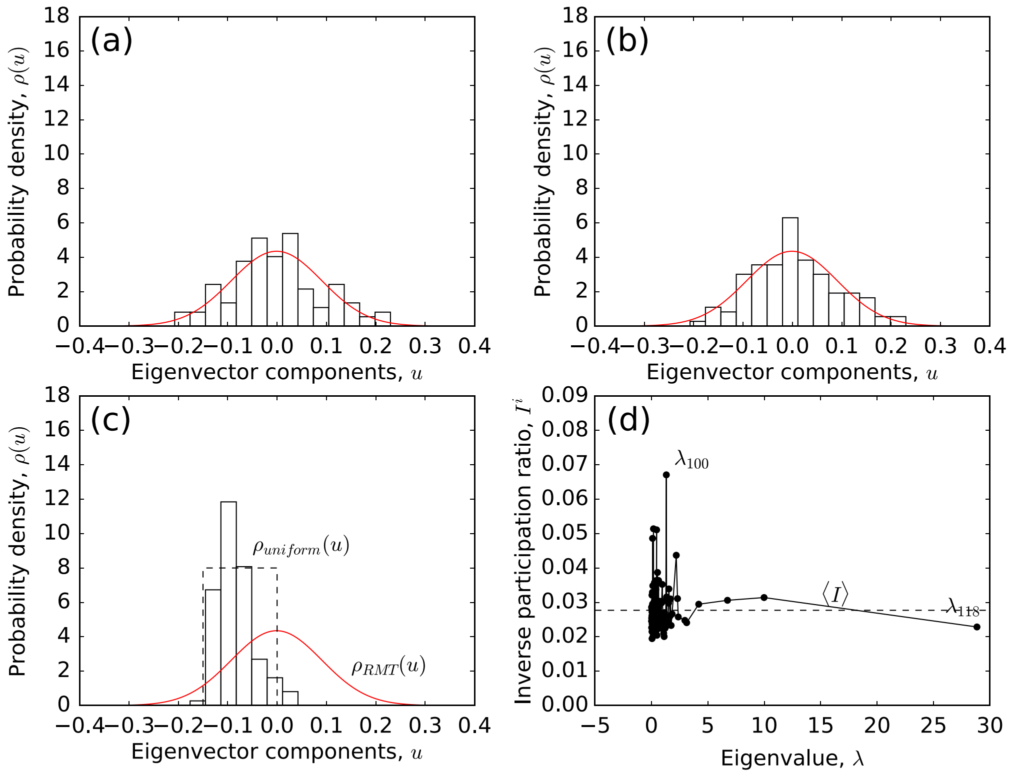

Fig. 4. Distribution $\rho(u)$ of eigenvector components for (a) an eigenvalue in the bulk $\lambda_- < \lambda < \lambda_+$, (b) the second largest eigenvalue λ_{117}, (c) the largest eigenvalue λ_{118}. The solid curve corresponds to the RMT prediction $\rho_{RMT}(u)$ in (9), and dashed curve is that of a uniform distribution. (d) Inverse participation ratio (IPR) as a function of eigenvalue λ for the cross correlation matrix C.

value predicted for a random correlation matrix, suggesting genuine information about the correlations between different cryptocurrencies.

Since (7) is strictly valid only for $L \to \infty$ and $N \to \infty$, we must test that the deviations for the largest few eigenvalues are not an effect of finite values of L and N. Fig. 3(b) shows that the eigenvalues of a random correlation matrix constructed from $N = 119$ mutually uncorrelated time series each of the same length $L = 200$ are in good agreement with $P_{RMT}(\lambda)$, which means that the deviating eigenvalues are not a result of finite effects. Moreover, we find that similar eigenvalue distributions $P(\lambda)$ are found by varying the number of random measurements (L) and the number of chosen cryptocurrencies (N).

The deviations of $P(\lambda)$ from $P_{RMT}(\lambda)$ suggests that these deviations should also be displayed in the statistics of the eigenvector components u^i_k, $k = 1, \ldots, N$. Therefore we can analyze the distribution of eigenvector components, $\rho(u)$. RMT predicts that the components of the normalized eigenvectors of a random correlation matrix R are distributed according to a Gaussian with zero mean and unit variance [36]

$$\rho_{RMT}(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}. \tag{9}$$

Fig. 4 shows that $\rho(u)$ for a typical eigenvector u^i from the bulk shows good agreement with the RMT prediction, while analysis on the other eigenvectors belonging to eigenvalues within λ_{bulk} yields similar results. On the other hand, eigenvectors with eigenvalues outside the bulk (or $\lambda_i > \lambda_+$) deviate significantly from $\rho_{RMT}(u)$. In particular, Fig. 4(c) shows that the eigenvector corresponding to the largest eigenvalue λ_{118} is nearly uniformly distributed, meaning that all cryptocurrencies contribute almost equally. This suggests that this eigenvector represents a collective mode in which all cryptocurrencies participate, and the magnitude of the largest eigenvalue seems to reflect the degree of collective behavior [15]. In addition, we find that almost all components of u^{118} have the same sign, causing $\rho(u)$ to shift to one side, which suggests that the significant participants of eigenvector u^{118} have a common component that affects all of them with the same bias [12].

The deviations of $\rho(u)$ from the RMT prediction of $\rho_{RMT}(u)$ are more pronounced as the separation from the RMT upper bound λ_+ increases. We therefore quantify the number of components that participate significantly in each eigenvector, which reflects the degree of deviation of the distribution of eigenvectors from RMT [12]. We use the notion of the inverse

Table 2
Properties of community structures in the MST of C.

Community	Hub	# of Nodes	Market Cap.
#1	Bitcoin	27	190690385600
#2	Ethereum	17	124413330300
#3	Synereo	23	66434438900
#4	GridCoin	18	51270539400
#5	MaidSafeCoin	34	20272480800
MST	–	119	453081175000

participation ratio (IPR) [11,36,47] defined as

$$I^i = \sum_{k=1}^{N} [u_k^i]^4, \tag{10}$$

where u_k^i, $k = 1, \ldots, N$ are the components of eigenvector u^i. The power term is chosen such that I^i can be interpreted as the reciprocal of the number of eigenvector components that contribute significantly: $I^i = 1/N$ for a vector with identical components $u_k^i = 1/\sqrt{N}$ and $I^i = 1$ for a vector with one component $u_k^i = 1$ and the remainder zero [11]. Fig. 4(d) shows that the average value of I^i is $\langle I \rangle \approx 0.03 \approx 1/N$, indicating that the eigenvectors are extended [47], or almost all cryptocurrencies contribute to them. In contrast to financial markets [11,12,45,46], fluctuations around this average value are significant, showing that there are varying numbers of cryptocurrencies contributing to these eigenvectors. Particularly, I^i values of λ_{100} are approximately four to five times larger than the average $\langle I \rangle$. The lack of deviations from $\langle I \rangle$ at the edges of the eigenvalue spectrum implies that the eigenvectors are not localized [48] and that C does not resemble a random matrix band [47] as suggested for stock markets [11,12].

5. Hierarchical structures in minimum spanning trees

Financial markets are well represented as complex systems. On the side of modeling financial markets using concepts and methods developed for complex networks, one can consider different assets as nodes in a network, and their correlations as the connections between them. Such treatment allows for the detection of hierarchical structures (or collective behavior) in a given financial market [18]. To represent the cryptocurrency market as a complex network, there is the need to quantify a distance between the different cryptocurrencies that are being traded. Each element of the correlation matrix must be associated to a metric distance

$$D_{ij} = \sqrt{2(1 - C_{ij})}, \tag{11}$$

where D denotes the distance matrix. Correlations of the cryptocurrency market are then translated into a network by defining D as the adjacency matrix, $A = D$. Several studies have proposed network based models for studying correlations of financial time series, such as planar maximally-filtered graph (PMFG) [19], correlation threshold [20], and other relevant financial networks [21–23] (see Ref. [49] for a more detailed review). Among the most popular networks is the Minimal Spanning Tree (MST). Spanning trees are particular types of networks which connect all the nodes without forming any loop. Therefore if the number of node is N, one has $N - 1$ edges to connect them. The minimum means that the sum of all edges is minimal among all spanning trees defined on the distance matrix. Ref. [18] outlines the method for constructing the MST out of correlations in financial time series.

From the MST construction, we can analyze collective behavior between correlations in the cryptocurrency market. Fig. 5 reveals that the corresponding network divides naturally into groups of nodes with dense connections internally and sparser connections between them. This leads to the existence of community structures [50], which can be found using community detection methods such as Girvan–Newman [50]. The algorithm detects communities by progressively removing edges from the original network, leaving behind only the communities. Existence of communities implies that different collective behavior exists between cryptocurrencies, which is in sharp contrast to the popular and perhaps intuitive belief that Bitcoin exerts a global influence on the entire market because of its dominant position. Particularly, we find that the MST of cryptocurrencies consists of five communities. Since similar communities are found for different random measurements (or time periods L) and choices of cryptocurrencies (or set N), we can conclude that they represent genuine information about the cryptocurrency market.

The communities found in the MST of cryptocurrencies can be further characterized by their properties. The community hubs (or node with the largest number of links within a community) are formed by Bitcoin, Ethereum, Synereo, GridCoin and MaidSafeCoin as listed in Table 2. These hubs are either formed or clustered by cryptocurrencies with large market capitalization. Table 2 reveals that Bitcoin and MaidSafeCoin communities are by far the largest with respect to the number of cryptocurrencies. The fact that the Bitcoin community is extremely close to the Ethereum community suggests that important cryptocurrencies with significant innovations to the blockchain are grouped together. We also find that Bitcoin and Ethereum communities retain most of the market capitalization. On the other hand, the Synereo community consists

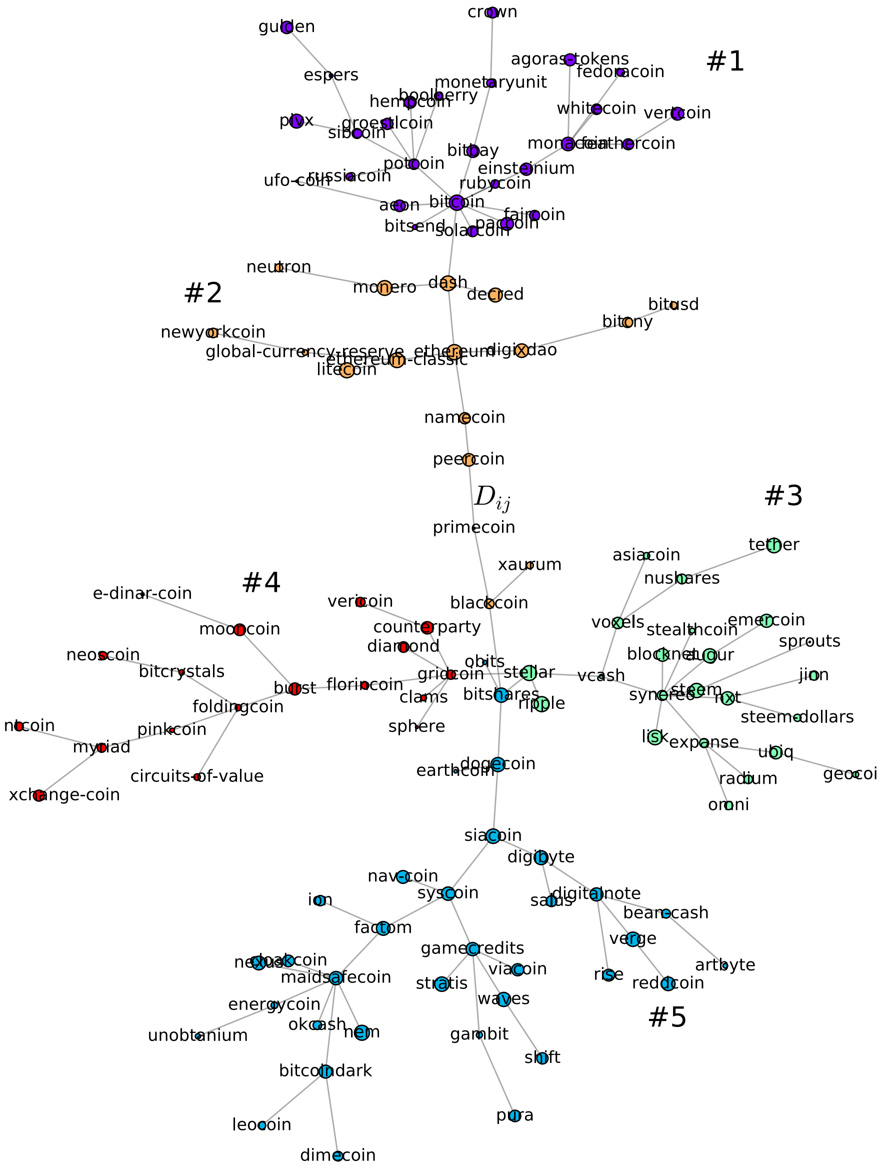

Fig. 5. Minimum spanning tree of cross correlations of the cryptocurrency market. Nodes correspond to individual cryptocurrencies and connections to elements of the distance matrix *D*. The node sizes are proportional to the rank of the associated cryptocurrency by market capitalization. Colors represent distinct community structures. Numbers correspond to the community labels in Table 2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of cryptocurrencies with strong anti-correlations (such as Tether, Ripple and Stellar) found in Fig. 1. Interestingly, the MaidSafeCoin community is composed of smaller cryptocurrencies by market capitalization, and therefore has the least influence in the market, while the GridCoin community holds all of those in-between. Results in Table 2 further indicate that the communities have very different properties than the average properties of the MST. One useful application involves the management of investment portfolios, where diversification (or lower risk) of the portfolio can be guaranteed by selecting cryptocurrencies from distinct communities.

The properties of individual nodes in the MST can also provide useful information regarding the cryptocurrencies. We therefore analyze the degree d_i and the influence strength [51], $S_i = \sum_j C_{ij}$, for each node in the spanning tree. Table 3 shows that the largest values of d_i and S_i are found for the community hubs, which means they exert greatest influence in their respective communities. Another useful measure is the betweenness centrality [52], $B_i = \sum_{i \neq j \neq k} = \sigma^i_{jk}/\sigma_{jk}$, where σ^i_{jk} is the number of shortest paths between nodes (or cryptocurrencies) j and k that run through node i, and σ_{jk} is the total number of shortest paths between j and k. We find that cryptocurrencies along the center of the MST have the highest B_i, which quantifies the importance of a node when positioned between other nodes in the network. Lastly we consider the closeness

Table 3
Network properties of individual nodes in the MST of C.

Node	d_i	Node	S_i	Node	B_i	Node	C_i
Bitcoin	0.085	Bitcoin	5.375	BitShares	0.672	BitShares	0.174
Synereo	0.068	Synereo	4.937	Stellar	0.512	Blackcoin	0.167
MaidSafeCoin	0.059	MaidSafeCoin	3.671	Ethereum	0.474	Stellar	0.165
Gridcoin	0.051	Potcoin	3.474	Blackcoin	0.473	Dogecoin	0.161
Potcoin	0.051	Gridcoin	3.306	Primecoin	0.457	Primecoin	0.159
Monacoin	0.042	GameCredits	2.910	Peercoin	0.452	Peercoin	0.151
GameCredits	0.042	Monacoin	2.761	Namecoin	0.446	VCash	0.149
Ethereum	0.034	Stellar	2.487	Siacoin	0.398	Siacoin	0.149
Dash	0.034	Ethereum	2.447	Dogecoin	0.395	Obits	0.149
Stellar	0.034	Expanse	2.346	Dash	0.394	Gridcoin	0.148

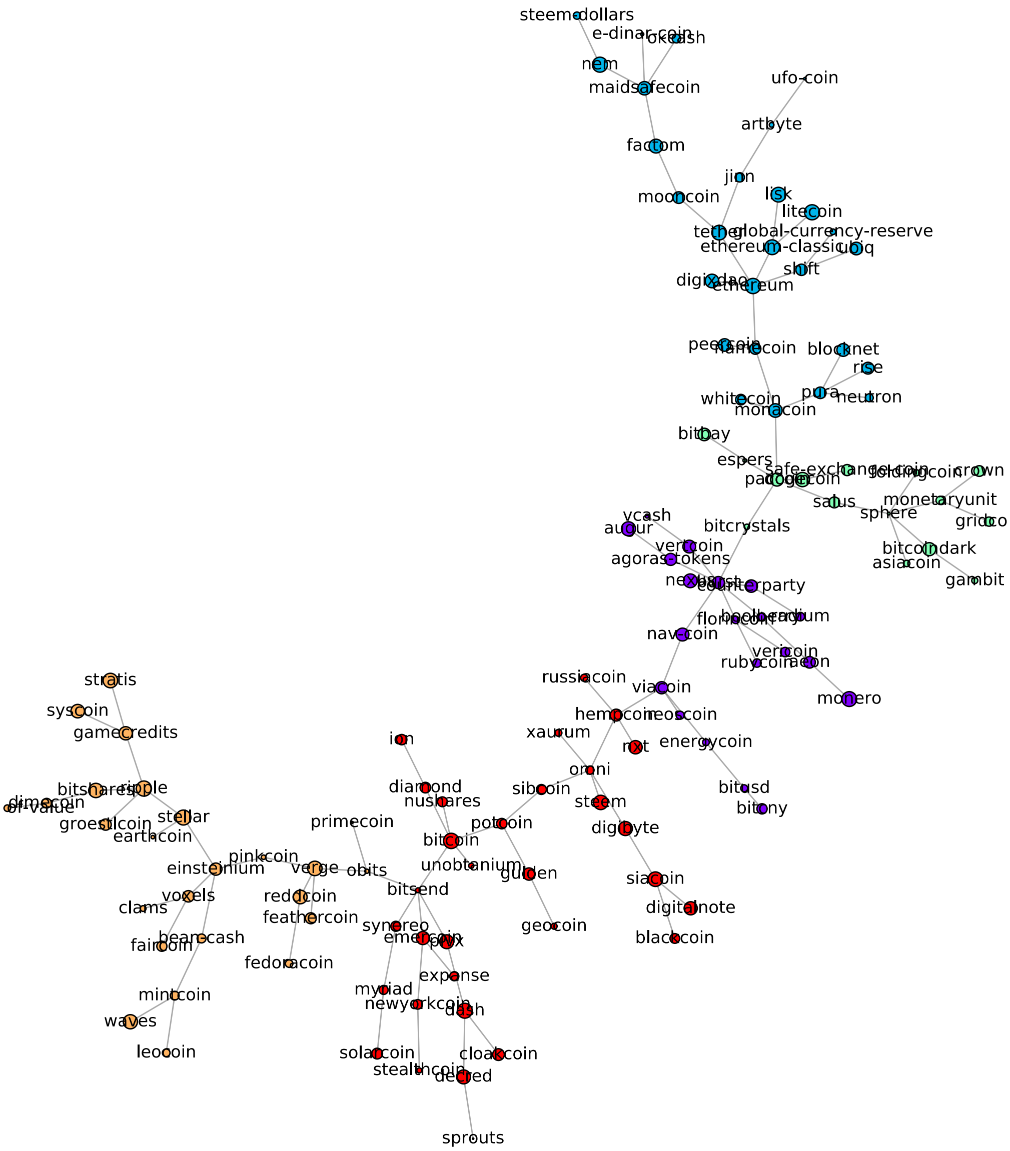

Fig. 6. Minimum spanning tree of partial cross correlations of the cryptocurrency market. Nodes correspond to individual cryptocurrencies and connections to elements of the distance matrix D^*. The node sizes are proportional to the rank of the respective cryptocurrency by market capitalization. Colors represent distinct community structures. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

centrality [53], $C_i = (N-1)/\sum_j l_{ij}$, where l_{ij} is the shortest path length from node i to j. Here the central cryptocurrencies are again the closest to other nodes in the network. Interestingly, Bitcoin only holds the largest value for S_i, but not for B_i and C_i. This implies that Bitcoin exerts a localized influence in its community, but does not drive the behavior of the wider cryptocurrency market.

The minimum spanning tree of the partial correlations can also be considered. Here we calculate a new distance matrix with elements $D^*_{ij} = \sqrt{2(1 - C^*_{ij})}$ from the partial correlation matrix C^*, and construct the MST from the adjacency matrix $A = D^*$, following the recipe in Ref. [18]. Fig. 6 reveals that hierarchical structures in the MST differ from those in Fig. 5. The communities have closer connections between them and they are composed of different cryptocurrencies. For example, the Bitcoin and Ethereum communities are widely apart, while they were neighboring communities for the MST in Fig. 5. These results indicate that the correlations between cryptocurrency pairs are strongly influenced by other elements of the market of cryptocurrencies.

6. Conclusions

To summarize, we quantify the collective behavior of the market of cryptocurrencies through the correlations of 119 of their constituents. We find that the cross correlation matrix of cryptocurrency price changes exhibits non-trivial hierarchical structures and groupings of cryptocurrency pairs. For partial cross correlations most of these structures are lost and anti-correlations seem to dominate the matrix elements. Surprisingly, we discover that most eigenvalues in the spectrum of the cross correlation matrix do not agree with the universal predictions of random matrix theory, which is in sharp contrast to the predictions for other financial markets. We further analyze the deviations from RMT and find that the largest eigenvalue and its corresponding eigenvector represent the influence of the entire market on all cryptocurrencies. Lastly, we represent the correlations of the cryptocurrency market as a complex network and discover distinct community structures in its minimum spanning tree. Our results reveal the presence of multiple collective behaviors in the market of cryptocurrencies, which can be useful when constructing cryptocurrency investment portfolios.

Acknowledgments

This work is supported by Brazilian agencies CAPES and CNPq, Brazil (grants No. 442668/2014-7 and No. 140840/2016-8).

References

[1] I. Kertesz, I. Kondor, Econophysics: An Emerging Science, Kluwer, 1999.
[2] R. Mantegna, H. Stanley, Introduction to Econophysics: Correlations and Complexity in Finance, Cambridge University Press, 1999.
[3] R.N. Mantegna, H.E. Stanley, Scaling behaviour in the dynamics of an economic index, Nature 376 (1995) 46–49.
[4] L.A.N. Amaral, S.V. Buldyrev, S. Havlin, M.A. Salinger, H.E. Stanley, Power law scaling for a system of interacting units with complex internal structure, Phys. Rev. Lett. 80 (1998) 1385–1388.
[5] J. Kwapień, S. Drożdż, Physical approach to complex systems, Phys. Rep. 515 (3) (2012) 115–226.
[6] E. Elton, M. Gruber, Modern Portfolio Theory and Investment Analysis, Wiley, 1995.
[7] J.P. Bouchaud, M. Potters, Theory of Financial Risk, Cambridge University Press, 2000.
[8] J. Campbell, A.W. Lo, A.C. MacKinlay, The Econometrics of Financial Markets, Princeton University Press, 1997.
[9] H. Wold, A Study in the Analysis of Stationary Time Series, Almqvist and Wiksell, 1938.
[10] O.A. Rosso, H.A. Larrondo, M.T. Martin, A. Plastino, M.A. Fuentes, Distinguishing noise from chaos, Phys. Rev. Lett. 99 (2007) 154102.
[11] V. Plerou, P. Gopikrishnan, B. Rosenow, L.A. Nune. Amaral, H.E. Stanley, Universal and nonuniversal properties of cross correlations in financial time series, Phys. Rev. Lett. 83 (1999) 1471–1474.
[12] V. Plerou, P. Gopikrishnan, B. Rosenow, L.A.N. Amaral, T. Guhr, H.E. Stanley, Random matrix approach to cross correlations in financial data, Phys. Rev. E 65 (2002) 066126.
[13] V. Plerou, P. Gopikrishnan, B. Rosenow, L. Amaral, H. Stanley, A random matrix theory approach to financial cross-correlations, Physica A 287 (3) (2000) 374–382.
[14] L. Laloux, P. Cizeau, J.-P. Bouchaud, M. Potters, Noise dressing of financial correlation matrices, Phys. Rev. Lett. 83 (1999) 1467–1470.
[15] V. Plerou, P. Gopikrishnan, B. Rosenow, L. Amaral, H. Stanley, Collective behavior of stock price movements—a random matrix theory approach, Physica A 299 (1) (2001) 175–180.
[16] B. Podobnik, D. Wang, D. Horvatic, I. Grosse, H.E. Stanley, Time-lag cross-correlations in collective phenomena, Europhys. Lett. 90 (6) (2010) 68001.
[17] G.-J. Wang, C. Xie, S. Chen, J.-J. Yang, M.-Y. Yang, Random matrix theory analysis of cross-correlations in the us stock market: Evidence from pearson's correlation coefficient and detrended cross-correlation coefficient, Physica A 392 (17) (2013) 3715–3730.
[18] R.N. Mantegna, Hierarchical structure in financial markets, Eur. Phys. J. B 11 (1) (1999) 193–197.
[19] M. Tumminello, T. Aste, T. Di Matteo, R.N. Mantegna, A tool for filtering information in complex systems, Proc. Natl. Acad. Sci. 102 (30) (2005) 10421–10426.
[20] V. Boginski, S. Butenko, P.M. Pardalos, Statistical analysis of financial networks, Comput. Statist. Data Anal. 48 (2) (2005) 431–443.
[21] G.-J. Wang, C. Xie, Tail dependence structure of the foreign exchange market: A network view, Expert Syst. Appl. 46 (2016) 164–179.
[22] G.-J. Wang, C. Xie, K. He, H.E. Stanley, Extreme risk spillover network: application to financial institutions, Quant. Finance 17 (9) (2017) 1417–1433.
[23] G.-J. Wang, C. Xie, S. Chen, Multiscale correlation networks analysis of the us stock market: a wavelet analysis, J. Econ. Interact. Coord. (3) (2017) 561–594.
[24] D. Wang, B. Podobnik, D. Horvatić, H.E. Stanley, Quantifying and modeling long-range cross correlations in multiple time series with applications to world stock indices, Phys. Rev. E 83 (2011) 046121.
[25] J. Shen, B. Zheng, Cross-correlation in financial dynamics, Europhys. Lett. 86 (4) (2009) 48005.
[26] S. Drozdz, A. Górski, J. Kwapién, World currency exchange rate cross-correlations, Eur. Phys. J. B 58 (2007) 499–502.
[27] P. Gopikrishnan, B. Rosenow, V. Plerou, H.E. Stanley, Identifying business sectors from stock price fluctuations, arXiv:cond-mat/0011145.

[28] S. Nakamoto, Bitcoin: A peer-to-peer electronic cash system, 2008, http://bitcoin.org/bitcoin.pdf.
[29] A. ElBahrawy, L. Alessandretti, A. Kandler, R. Pastor-Satorras, A. Baronchelli, Evolutionary dynamics of the cryptocurrency market, Royal Society Open Science 4(11).
[30] G.-J. Wang, C. Xie, H.E. Stanley, Correlation structure and evolution of world stock markets: Evidence from pearson and partial correlation-based networks, Comput. Econ.
[31] I.S. Dhillon, Co-clustering documents and words using bipartite spectral graph partitioning, in: Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '01, ACM, New York, NY, USA, 2001, pp. 269–274.
[32] E.P. Wigner, On a class of analytic functions from the quantum theory of collisions, Ann. of Math. 53 (1) (1951) 36–67.
[33] F.J. Dyson, Statistical theory of the energy levels of complex systems. i, J. Math. Phys. 3 (1) (1962) 140–156.
[34] M.L. Mehta, F.J. Dyson, Statistical theory of the energy levels of complex systems, J. Math. Phys. 4 (5) (1963) 713–719.
[35] T.A. Brody, J. Flores, J.B. French, P.A. Mello, A. Pandey, S.S.M. Wong, Random-matrix physics: spectrum and strength fluctuations, Rev. Modern Phys. 53 (1981) 385–479.
[36] T. Guhr, A. Müller-Groeling, H.A. Weidenmüller, Random-matrix theories in quantum physics: common concepts, Phys. Rep. 299 (4) (1998) 189–425.
[37] D.-M. Song, M. Tumminello, W.-X. Zhou, R.N. Mantegna, Evolution of worldwide stock markets, correlation structure, and correlation-based graphs, Phys. Rev. E 84 (2011) 026108.
[38] G. Oh, C. Eom, F. Wang, W.-S. Jung, H.E. Stanley, S. Kim, Statistical properties of cross-correlation in the korean stock market, Eur. Phys. J. B 79 (1) (2011) 55–60.
[39] H. Meng, W.-J. Xie, Z.-Q. Jiang, B. Podobnik, W.-X. Zhou, H.E. Stanley, Systemic risk and spatiotemporal dynamics of the us housing market, Sci. Rep. 4 (2014) 3655.
[40] M. Potters, J.-P. Bouchaud, L. Laloux, Financial applications of random matrix theory: Old laces and new pieces, Phys. Pol. B 36 (2005) 2767.
[41] J.-P. Bouchaud, L. Laloux, M.A. Miceli, M. Potters, Large dimension forecasting models and random singular value spectra, Eur. Phys. J. B 55 (2) (2007) 201–207.
[42] T. Nakamura, M. Small, Correlation structures in short-term variabilities of stock indices and exchange rates, Physica A 383 (1) (2007) 96–101.
[43] A.M. Sengupta, P.P. Mitra, Distributions of singular values for some random matrices, Phys. Rev. E 60 (1999) 3389–3392.
[44] M.J. Bowick, E. Brézin, Universal scaling of the tail of the density of eigenvalues in random matrix models, Phys. Lett. B 268 (1) (1991) 21–28.
[45] M.E. Alaoui, Random matrix theory and portfolio optimization in moroccan stock exchange, Physica A 433 (2015) 92–99.
[46] J. Daly, M. Crane, H. Ruskin, Random matrix theory filters and currency portfolio optimisation, J. Phys. Conf. Ser. 221 (1) (2010) 012003.
[47] Y.V. Fyodorov, A.D. Mirlin, Analytical derivation of the scaling law for the inverse participation ratio in quasi-one-dimensional disordered systems, Phys. Rev. Lett. 69 (1992) 1093–1096.
[48] Y.V. Fyodorov, A.D. Mirlin, Level-to-level fluctuations of the inverse participation ratio in finite quasi 1d disordered systems, Phys. Rev. Lett. 71 (1993) 412–415.
[49] G. Marti, F. Nielsen, M. Bińkowski, P. Donnat, A review of two decades of correlations, hierarchies, networks and clustering in financial markets. arXiv:1703.00485.
[50] M. Girvan, M.E.J. Newman, Community structure in social and biological networks, Proc. Natl. Acad. Sci. 99 (12) (2002) 7821–7826.
[51] H.-J. Kim, Y. Lee, B. Kahng, I.-M. Kim, Weighted scale-free network in financial correlations, J. Phys. Soc. Japan 71 (9) (2002) 2133–2136.
[52] L.C. Freeman, A set of measures of centrality based on betweenness, Sociometry 40 (1) (1977) 35–41.
[53] L.C. Freeman, Centrality in social networks conceptual clarification, Social Networks 1 (3) (1978) 215–239.