Journal of Computational Science 44 (2020) 101130

Contents lists available at ScienceDirect

Journal of Computational Science

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

Cross-correlation dynamics and community structures of cryptocurrencies

Harshal Chaudhari, Martin Crane *

School of Computing, Dublin City University, Glasnevin, Dublin 9, Ireland

ARTICLE INFO

Article history:
Received 11 October 2019
Received in revised form 2 February 2020
Accepted 18 April 2020
Available online 12 March 2020

Keywords:
Cyptocurrency
Community structure
Correlation matrix
Random Matrix Theory

ABSTRACT

Cryptocurrencies have become a prominent investment tool recently with increasing interest in them and their relationships with stock and foreign exchange markets. We analyze here the cross-correlations of price changes of different cryptocurrencies using Random Matrix Theory and extract community structures by constructing minimum spanning trees, finding their eigenvalues contrast sharply with universal predictions of Random Matrix Theory. We reveal distinct transient community structures among different groupings of cryptocurrencies. By studying the cross-correlation dynamics of sub-communities we find evidence of collective behaviour. Furthermore, we compare eigenvalue changes and find prominent groupings following a community trend, useful for creating cryptocurrency portfolios.

1. Introduction

The exponential increase of Bitcoin and other cryptocurrencies have gained substantial attention in recent years. Bitcoin is a type of decentralized digital currency [1], where *Decentralized* means Bitcoin is peer-to-peer payment and is not regulated by any third party. Moreover, this cyptocurrency is independent of any other commodity market in the world.

Bitcoin's value has surged in the last few year and massive demands for the cryptocurrency have led it to reach an all-time high of USD 19,891 in December 2017.[1] Therefore, it is crucial to be able to forecast the value of Bitcoin to secure profitable investments. Recently, cryptocurrencies have become an investment tool where trading is carried out akin to trade in stock and foreign exchange market. Various trading platforms are available where you can buy and sell such assets [2].

Investors of cryptocurrencies commonly use traditional methods [3,4] in the stock market trading. For instance, the basic notion of buying the commodity when at a low price and selling at a high price is applied by the investors. Risks are evaluated ahead in time of investing and one such method commonly used for risk analysis is Market Technical Analysis (MTA). MTA recognizes the trend of the market given the historical market data. In such an analysis, candle graphs and market technical indicators [5] are used. However, such graphs are difficult and require experts to interpret. Furthermore, techniques like the Efficient Market Hypothesis (EMH) [6] are used to analyze market trends. However, the results of EMH would be inconsistent, as according to Fama [7,8] for analysis to work the prices should follow a random walk. Other such technical indicators such as Williams %R as a momentum indicator and Exponential Moving Average (EMA) can also be used as trend indicators [9]. EMA is the average movement of market price over some time, while %R represents the strength of the overbought or oversold market.

Financial markets can be represented as complex systems that have been vastly researched upon by physicists using methods used to describe complex systems [10–12]. Research into correlations between financial assets has been an area of interest for understanding the market as a complex system and also for developing investment portfolios [13,14]. One of the key questions that arises is whether the correlations in financial time series is a result of noise or genuine interactions [15,16] when market conditions are not always stationary and historical data are finite. Hence, it is a challenging task to quantify the interactions in financial systems.

The problem of quantifying correlations between different financial systems is studied by Plerou et al. [17]. They propose a method where the cross-correlation matrix is used to identify the correlation coefficients between different assets. The authors speculate that the corresponding matrix eigenvalues represent the collective behaviour of the market [17]. One can use these eigenvalues to analyze asset correlations, such as identifying the non-random properties of the system through deviations from universal predictions of the Random Matrix Theory (RMT) [16,18–20] and to reveal internal community structure (i.e. groups of cryptocurrencies that behave similarly).

* Corresponding author.
E-mail address: Martin.Crane@dcu.ie (M. Crane).

[1] https://www.oanda.com/currency/converter/.

https://doi.org/10.1016/j.jocs.2020.101130

Previously, such correlations have been a strong area of research [18,21] revealing collective market behaviour and correlations that spread throughout the entire system [17,16,19]. Gopikrishnan et al. [22] have discovered correlations being localized within various business sectors.

In recent years, the cryptocurrency market has created a new benchmark in the financial world. More than two thousand[2] cryptocurrencies are being traded in the market which relies on the same blockchain technology that is derived from Bitcoin or other similar currencies. Furthermore, cryptocurrencies are unique and extensively complex compared to other financial markets because of the unique nature of each cryptocurrency. Stosic et al. [23] have previously studied the interactions to investigate whether correlations in the market of cryptocurrencies exhibit similar properties to those of other financial markets. Ankenbrand and Bieri [24] examined the financial characteristics of cryptocurrency markets have concluded that currently no consensus exists on their uniqueness as a market or whether there exist similarities to other asset classes (e.g. stocks, bonds, commodities or foreign exchange). Reinforcing this, and investigating Cryptocurrency volatilities, Baur and Dimpfli [25] reported – for the 20 largest cryptocurrencies – a very different asymmetry as compared to equity markets: positive shocks increase volatility to a greater extent than negative shocks.

The proposed work updates and extends the research done by Stosic et al. [23]. We take the historical prices of 2 sets of cryptocurrencies with 119 in one and a smaller subset in the other over a total period of 300 days (April 18, 2017 to February 11, 2018) and analyze their cross-correlation dynamics and construct minimum spanning trees to identify distinct community structures and study their collective behaviour.

The paper covers the following: Section 2 covers the research done previously by various researchers. Section 3 describes the dataset we have used for this study. Section 4 describes the statistics of correlation matrix followed by Section 5 which depicts the methods that we have used to study the cross-correlation dynamics and construct a minimum spanning tree to reveal community structures. Lastly, Section 6 summarizes the findings of this study.

2. Literature review

Various articles and research articles and studies have been put forth in the past on forecasting market values. Guglielmo et al. [26] have examined 4 types of cryptocurrencies over the sample period of 2014–2017. They follow two approaches viz. Rescaled-Range R/S analysis and fractional integration. Time variation is particularly evident in the case of Litecoin, with the Hurst exponent dropping significantly from 0.7 in 2015 to 0.5 in 2017. Their findings indicate the market is persistent and its degree changes over time and conclude that the cryptocurrency market is inefficient.

Urquhart [4] has purposed another analysis of Bitcoin to study the market efficiency of Bitcoin and have used Hurst exponent analysis. The analysis shows when the full sample period is split into two sub-samples, the first sub-sample period rejects the null hypothesis of randomness and the R/S Hurst statistic indicates strong anti-persistence. However, when the second sub-sample period was studied, the Ljung-Box and Auto Variance Test (AVR) tests both fail to reject the null hypothesis, indicating absence of auto-correlation. This means that Bitcoin is inefficient, meaning investors cannot use previous information to forecast future values. The results obtained show that Bitcoin is efficient in the long run but currently the market is very volatile.

[2] https://coinmarketcap.com/coins/.

Bariviera [27] performed Hurst exponent analysis on Bitcoin computed using two alternative methods, Detrended Fluctuation Analysis (DFA) method for more robust results compared to the more commonly used R/S method. He found that the daily volatility of the market (as a proxy for the market risk) exhibits long memory over the time period studied. Rebane et al. [28] have used auto-regressive integrated moving average (ARIMA) and Seq2Seq Recurrent Neural Network(RNN) on Bitcoin and Altcoin to determine the best method to predict future values. The comparative analysis showed neural networks performed dramatically better than ARIMA as the cumulative errors were less for RNN. Furthermore, including social data from websites along with RNN reduced the error rate.

Karakoyun and Cibikdiken [29] have compared ARIMA and Long Short-Term Memory (LSTM) for forecasting values of Bitcoin and prices of the next 30 days were estimated. The results obtained were approx. 11.86% Mean Absolute Percentage Error (MAPE) with ARIMA against 1.40% MAPE with LSTM. The analysis used 1600 observations as training to predict the next 30 days. However, for LSTM only last 30 days are used to predict next days' prices as compared to the whole set of observations used in ARIMA.

Bakar and Rosbi [3] have also put forth a study of forecasting on Bitcoin using ARIMA and results indicate a non-stationary time series and obtain a MAPE of 5.36% over the period from January 2013 to October 2017. Kinderis et al. [30] have examined Bitcoin fluctuations using text mining of news articles and tweets to infer the relationship between these and cryptocurrency price direction using uses LSTM RNNs and a mix of hybrid models. Their modelling achieves higher accuracy in predicting the direction. However, their study indicates that sentiment analysis does not have an immediate effect on the cryptocurrency market.

Stosic et al. [23] have examined 119 different cryptocurrencies and analysed their cross-correlation matrix. Their findings indicate that the cross-correlation matrix of cryptocurrencies price changes exhibits non-trivial hierarchical structures and groupings in cryptocurrency pairs. For partial cross-correlation, anti-correlation was dominant in the matrix elements. Moreover, the findings revealed that most eigenvalues do not agree with universal predictions of the Random Matrix Theory (RMT), which is the exact opposite to the case of financial markets [19]. Later, the analysis of deviations from RMT revealed that the largest eigenvalue and its eigenvector represents the influence of the entire cryptocurrency market.

Dyhrberg and Haubo [31] explore the hedging capabilities of Bitcoin by applying Generalized Autoregressive Conditional Heteroskedasticity (GARCH) methodology. Their results indicate that Bitcoin has hedging capabilities against the Financial Times Stock Exchange (FTSE) and the American Dollar. They conclude that Bitcoin can be used alongside gold to minimize specific market risks. However, it should be emphasized that the correlations against the dollar are very small in value and indicative of short term capabilities.

Klein et al. [32] have also performed similar research into Bitcoin as a hedge. They initially analyse and compare the conditional variance properties of Bitcoin and gold. Next, they have implemented a Baba-Engle-Kraft-Kroner (BEKK-GARCH) model to estimate time-varying conditional correlations. Results obtained showed Bitcoin behaves exact opposite of gold and a positive correlation exists with downward markets. In conclusion, and in contrast to [31], Bitcoin and gold are fundamentally different which showed no evidence of Bitcoin having stable hedging capabilities.

To summarize, previous studies done on the cryptocurrency market are limited to researchers focusing on selecting a few currencies based on market capital, volatility and crash in the market. Studies are often contradictory and the overalll lessons that can be taken regarding Bitcoin and other cryptocurrencies are regarding their volatility, differential behaviour vis à vis other

Table 1
List of 119 cryptocurrencies.

Cryptocurrency					
0x	Bytecoin-Bcn	Ethereum	Melon	Qtum	Tether
Adx-Net	Bytom	Ethereum-Classic	Metal	Quantum-Resistant-Ledger	Tierion
Aeternity	Centra	Ethos	Mobilego	Raiblocks	Triggers
Agoras-Tokens	Civic	Experience-Points	Monaco	Reddcoin	Tron
Aragon	Coindash	Factom	Monacoin	Ripple	Ubiq
Ardor	Counterparty	Funfair	Monero	Rlc	Vechain
Ark	Dash	Gamecredits	Nav-Coin	Santiment	Verge
Asch	Decentraland	Gas	Neblio	Siacoin	Veritaseum
Augur	Decred	Gnosis-Gno	Nem	Singulardtv	Vertcoin
Bancor	Dent	Golem-Network-Tokens	Neo	Skycoin	Viacoin
Basic-Attention-Token	Dentacoin	Gulden	Nexus	Smartcash	Wagerr
Binance-Coin	Digibyte	Gxshares	Nxt	Sonm	Walton
Bitbay	Digitalnote	Hshare	Omisego	Status	Waves
Bitcoin	Digixdao	Iconomi	Paccoin	Steem	Wings
Bitcoin-Cash	Dogecoin	Ion	Particl	Stellar	Xtrabytes
Bitcoindark	E-Coin	Iota	Peercoin	Storj	Zcash
Bitcore	Edgeless	Komodo	Pillar	Stratis	Zclassic
Bitshares	Einsteinium	Lisk	Pivx	Supernet-Unity	Zcoin
Blocknet	Emercoin	Loopring	Poet	Syscoin	Zencash
Byteball	Eos	Maidsafecoin	Populous	Tenx	

asset classes and consequent hedging properties. Also, despite the growing interest in dynamic correlations and its central role in the estimation of dynamic correlations, several important issues relating to this representation seem to have been ignored in the financial econometrics literature. Caporin et al. [33] mention these important issues include the absence of any derivation of dynamic conditional correlations and its mathematical properties, and a lack of any demonstration of the asymptotic properties of the estimated parameters. In fact, most published papers dealing with dynamic correlations simply do not discuss stationarity of the model, the regularity conditions, or the asymptotic properties of the estimators. In our study, we have considered a large set of cryptocurrencies to analyze their behaviour in the market. Thus, identifying key groupings of cryptocurrencies for investing and further research.

3. Data

To study evidence for their collective behaviour, we take the daily closing prices of the cryptocurriences from the source.[3] The data was collected from the website coinmarketcap.[4] We preprocessed our data into two sets of cryptocurrencies denoted by N. First set of $N = 150$ cryptocurrencies is listed in Table 1 and second set of $N = 59$ cryptocurrencies is listed in Table 2. For the first set, we take the data for 150 days (i.e. 150 observations, 1 observation per day) as the data was limited, which is denoted by $L = 150$. We then take two time period windows of $L = 150$ days from April 18, 2017 to February 11, 2018 for $N = 59$ to increase the Q factor. This was because data was only available on the 119 Cryptocurrencies in Table 1 for 200 days (i.e. continuous data) but considering only a subset of Cryptocurrencies (in Table 2) gives 300 days of continuous data during this period.

4. Statistics of cross-correlation matrix

In order to calculate the correlations, we first calculated the return of cryptocurrency $i = 1, \ldots, N$ over a time scale Δt

$$G_i(t) \equiv \ln S_i(t + \Delta t) - \ln S_i(t), \quad (1)$$

[3] https://data.world/pmohun/complete-historical-cryptocurrency-financial-data.
[4] https://www.coinmarketcap.com/.

Table 2
List of 59 cryptocurrencies (numbers refer to nodes shown in Fig. 7).

Cryptocurrency		
1. Agoras-tokens	21. Ethereum-classic	41. Reddcoin
2. Ardor	22. Experience-points	42. Siacoin
3. Augur	23. Factom	43. Singulardtv
4. Bitbay	24. Gamecredits	44. Steem
5. Bitcoin	25. Golem-network-tokens	45. Stellar
6. Bitcoindark	26. Gulden	46. Stratis
7. Bitshares	27. Iconomi	47. Supernet-unity
8. Blocknet	28. Ion	48. Syscoin
9. Byteball	29. Lisk	49. Tether
10. Bytecoin-bcn	30. Maidsafecoin	50. Triggers
11. Counterparty	31. Monacoin	51. Ubiq
12. Dash	32. Monero	52. Ripple
13. Decred	33. Nav-coin	53. Verge
14. Digibyte	34. Nem	54. Vertcoin
15. Digitalnote	35. Neo	55. Viacoin
16. Digixdao	36. Nexus	56. Waves
17. Dogecoin	37. Nxt	57. Zcash
18. Einsteinium	38. Paccoin	58. Zclassic
19. Emercoin	39. Peercoin	59. Zcoin
20. Ethereum	40. Pivx	

where $S_i(t)$ denotes the price of cryptocurrency i and we take $\Delta t = 1$ day. As different currencies have different values of volatility (standard deviation) we therefore take normalized returns:

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

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

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

where $C_{ij} = 1$ represents maximum correlation, $C_{ij} = -1$ represents maximum anti-correlation, and $C_{ij} = 0$ represents uncorrelated pairs of cryptocurrencies.

The cross-correlation matrix for the second set of $N = 59$ cryptocurrencies for the two time windows of $L = 150$ days each is shown in Fig. 1 and Fig. 2 respectively. We present the results from the second set of cryptocurrencies as during our initial experimentation on the first set of $N = 119$ cryptocurrencies exhibit odd results as mentioned further when we apply Kolmogorov-Smirnov test. It should be noted that there are many pairs with high positive correlation and few areas where the cryptocurrencies are uncorrelated. There

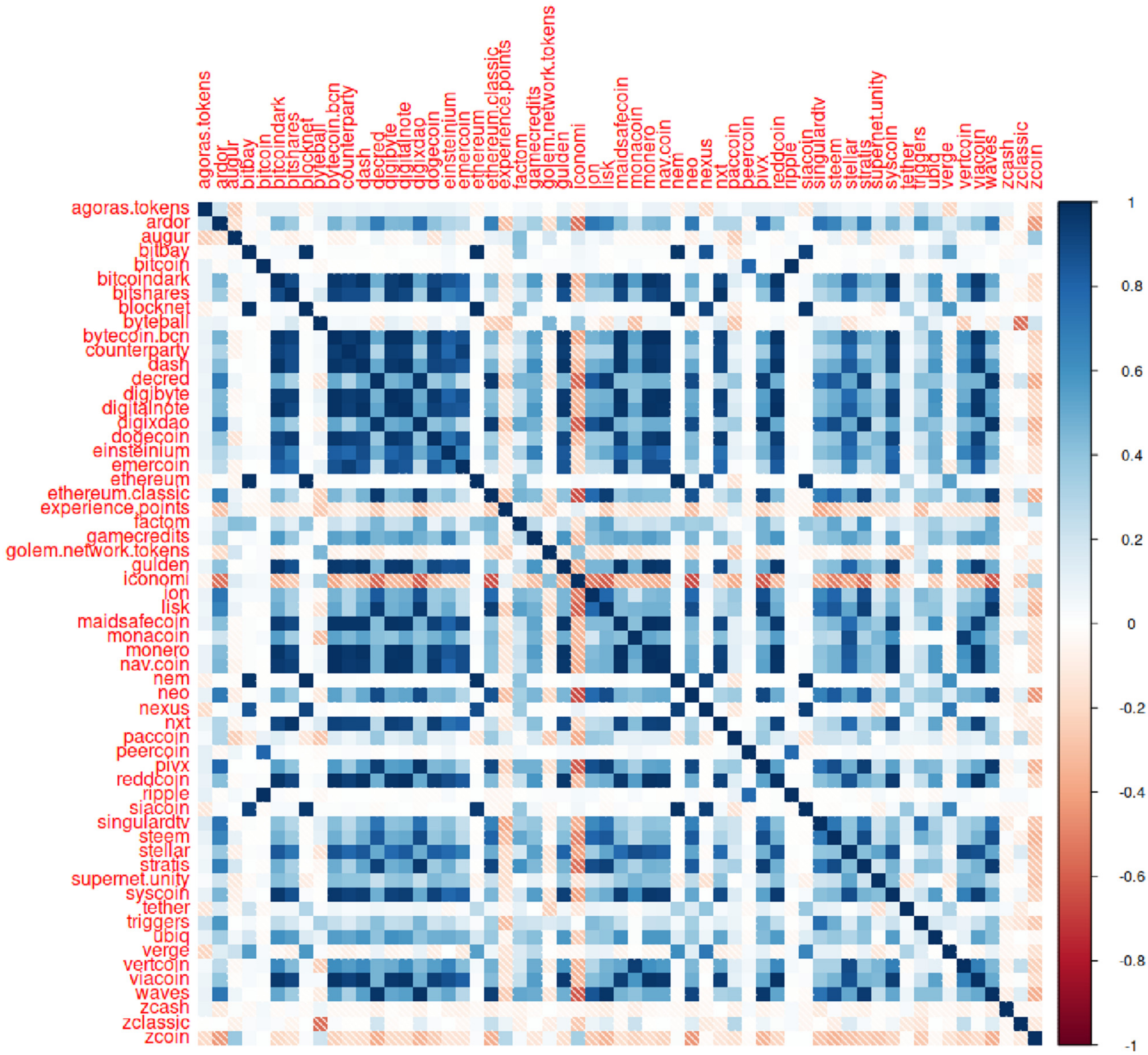

Fig. 1. Cross-correlation, C_{ij} of 59 Cryptocurrencies for the first window of 150 days.

are some cryptocurrencies which exhibit strong anti-correlation with the rest of the market such as *zcoin* and *iconomi*. This implies the existence of non-trivial groupings in the market as previously suggested by Stosic et al. [23].

Similarly, looking at the cross-correlation matrix for the second time window, we can see that anti-correlation is more dominant among pairs of cryptocurrencies in the matrix. Interestingly, for *zcoin* and *iconomi* anti-correlation are less pronounced now, as compared to the previous correlation matrix which exhibited strong anti-correlation with the market. Also, certain pairs at the end now exhibit more anti-correlation with the overall market. Hence, providing additional support to the existence of non-trivial groupings.

5. Research methodology

This section describes the analysis of noise done using Random Matrix Theory and further illustrates the use of minimum spanning trees to discover community structure among cryptocurrencies.

5.1. Noise dynamics from Random Matrix Theory

To find more explicit correlation structures, we use statistics from Random Matrix Theory to extract genuine correlations between cryptocurrencies prices. Random Matrix Theory (RMT) was developed to explain the statistics of energy levels of complex quantum systems in nuclear physics [34] with RMT universal predictions represent overall average interactions [35,34]. Deviations from these predictions reveal non-random properties that are specific to the system and arise from the collective behaviour among cryptocurrencies.

RMT has been previously used in financial market analysis [18,16] and has been widely used to identify cross-correlations in stock markets [21]. To find the information of cross-correlations of the market, we take the cross-correlation matrix and compare it with a random matrix. The correlation matrix can be written as

$$C = \frac{1}{L} G G^T, \tag{4}$$

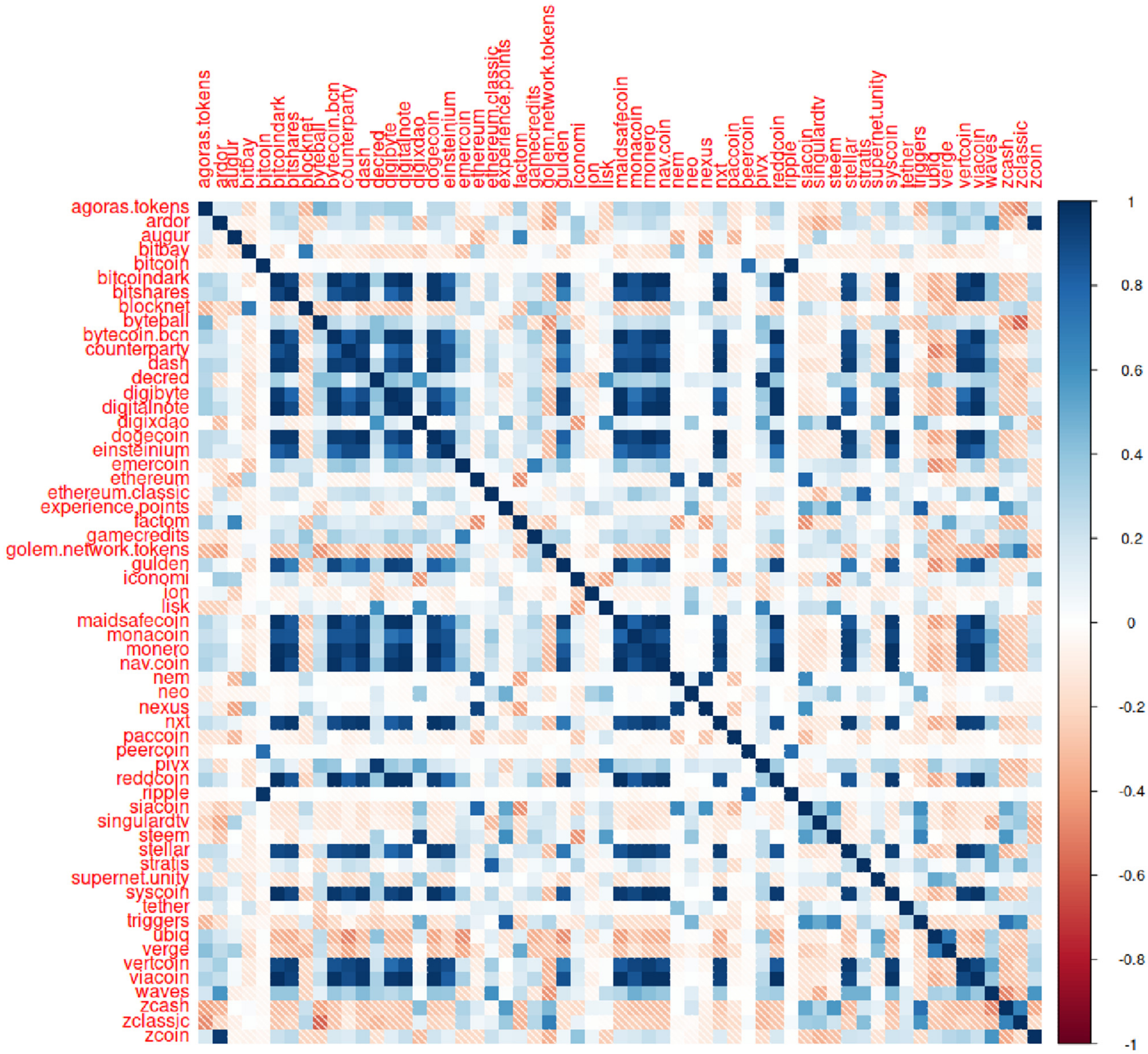

Fig. 2. Cross-correlation, C_{ij} of 59 Cryptocurrencies for the second window of 150 days.

where G is an $N \times L$ matrix with elements $g_{ik} \equiv g_i(k\Delta T)$ for $i = 1, \ldots, N$ and $k = 0, \ldots, L$. We then consider a random matrix such that

$$R = \frac{1}{L} A A^T, \tag{5}$$

where A is an N x L matrix with N time series of L random elements with zero mean and unit variance.

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 matrix R is given by

$$P_{RMT}(\lambda) = \frac{Q}{2\pi} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda}, \tag{6}$$

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

$$\lambda_\pm = 1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}}, \tag{7}$$

We study the statistics of C and compare with the universal properties of Random Matrix Theory [36]. We calculate the eigenvalue distribution and compare it with the matrix R generated from random time series. For $N = 59$ and $L = 150$ the eigenvalue bounds are $\lambda_- = 0.13$ and $\lambda_+ = 2.64$ and the Q[5] $= 2.54$. Fig. 3 shows the bulk of the eigenvalues of C for $\lambda_i \in \lambda_{bulk}$ falls within the bounds $\lambda_- < \lambda_{bulk} < \lambda_+$ for $P_{RMT}(\lambda)$. However, it can be seen that most eigenvalues deviate from the universal predictions of RMT as suggested earlier by Stosic et al. [23]. Similarly, we see there are deviating eigenvalues from RMT on the upper bound $\lambda_+ = 2.64$ for the largest few eigenvalues, which suggest genuine information exists between cross-correlation of cryptocurrencies. Further, we analyze the deviations of $P(\lambda)$ from $P_{RMT}(\lambda)$ by checking the statistics of the eigenvector components u_k^i, $k = 1, \ldots, N$. We do so by analyzing the distribution of these components, $\rho(u)$. RMT predicts that the components of the normalized eigenvectors of a random correla-

[5] Q is the defined as ratio of rows and columns of the matrix L/N.

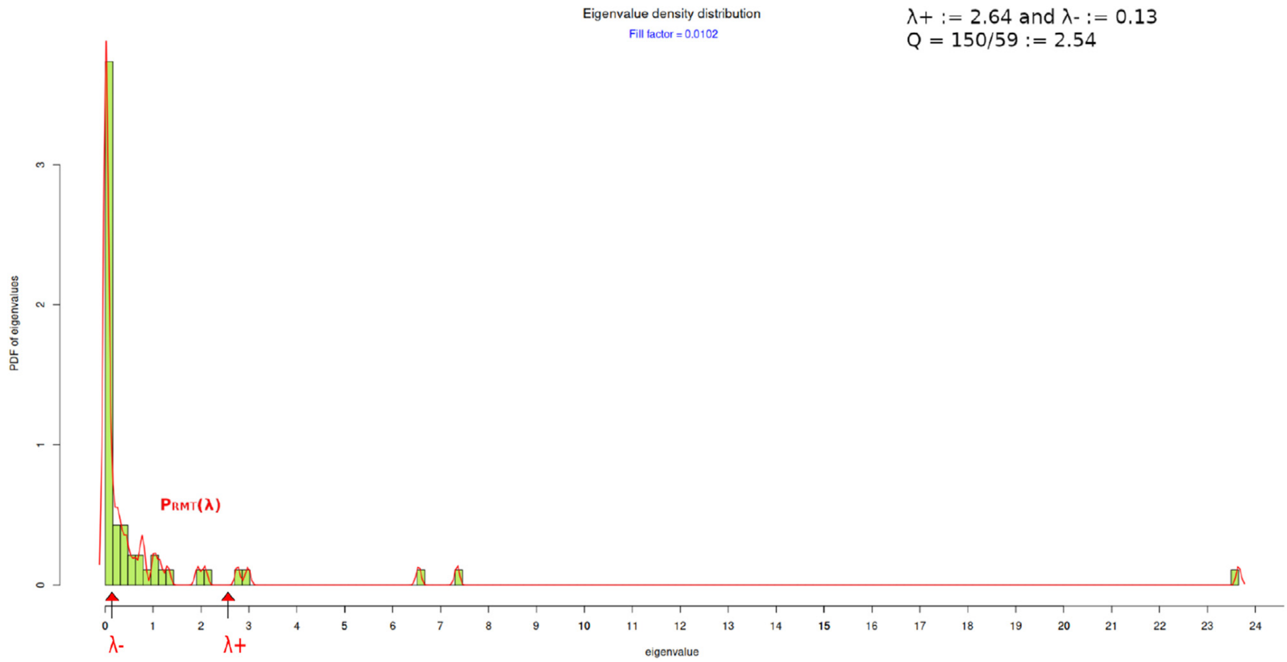

Fig. 3. Eigenvalue distribution $P(\lambda)$ for cross-correlation matrix C computed from $N = 59$ and $L = 150$.

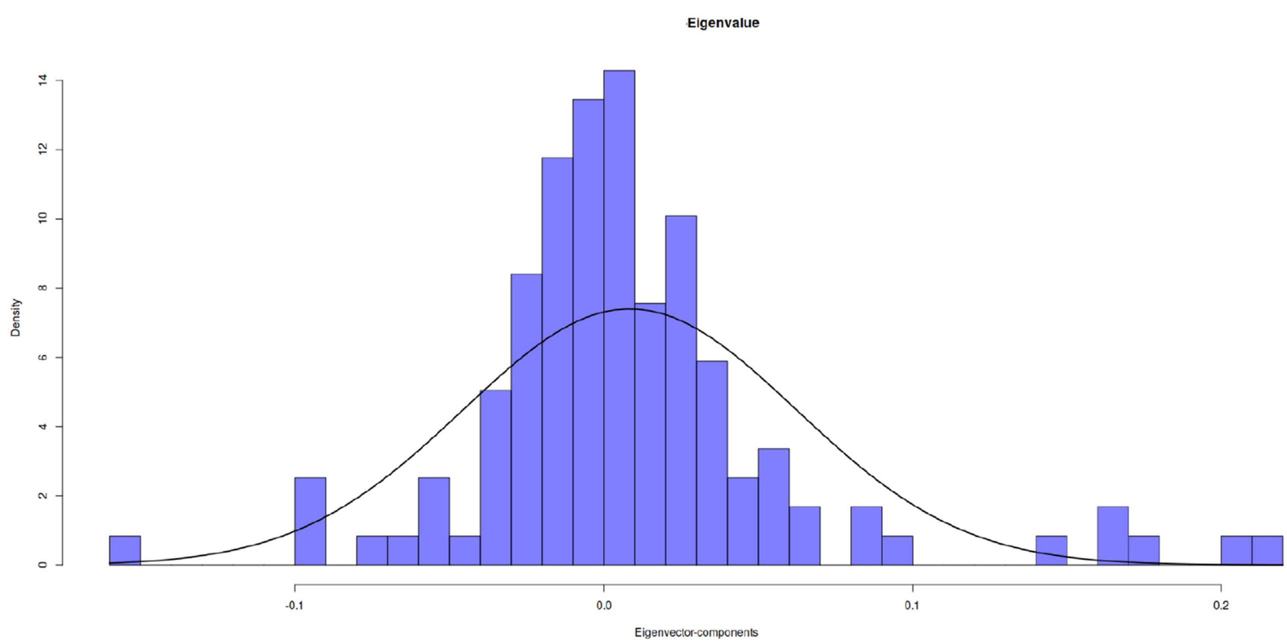

Fig. 4. Distribution $\rho(u)$ of eigenvector components for a eigenvalue in the bulk $\lambda_- < \lambda < \lambda_+$.

tion matrix R are distributed according to Gaussian with zero mean and unit variance [36]:

$$\rho_{RMT}(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \quad (8)$$

Fig. 4 shows $\rho(u)$ for an eigenvector u_i from the bulk shows good agreement with predictions from RMT. However, eigenvectors corresponding to eigenvalues outside the bulk (or $\lambda_i > \lambda_+$ deviate from the $\rho_{RMT}(u)$. From Fig. 5 we can see that for an eigenvalue outside the bulk the distribution is nearly uniform, which is similar to that suggested by Stosic et al. [23] in their analysis of eigenvector distribution.

We also further use Kolmogorov–Smirnov (KS) test to check whether the eigenvector components of deviating eigenvalues from the predictions of RMT come from the same distribution as those from the bulk. Surprisingly, for the initial cross-correlation matrix C with $N = 119$ and $L = 150$, we find that KS test fails to reject the null hypothesis. This could have something to do with the Q factor. As we had limited data the only way to increase the Q value was for the number of cryptocurrencies to decrease. Therefore, we consider a smaller subset of cryptocurrencies for further analysis.

Thus, we conducted KS tests with the $N = 59$ and $L = 150$ for the two time windows of $L = 150$ days each and we find that KS test is rejected. This suggests genuine information exists in the eigenvector components of eigenvalues outside the bulk.

As we move further from the upper bound λ_+ of RMT, the deviations of $\rho(u)$ are more significant. We therefore further quantify the components that participate in each eigenvector, which shows

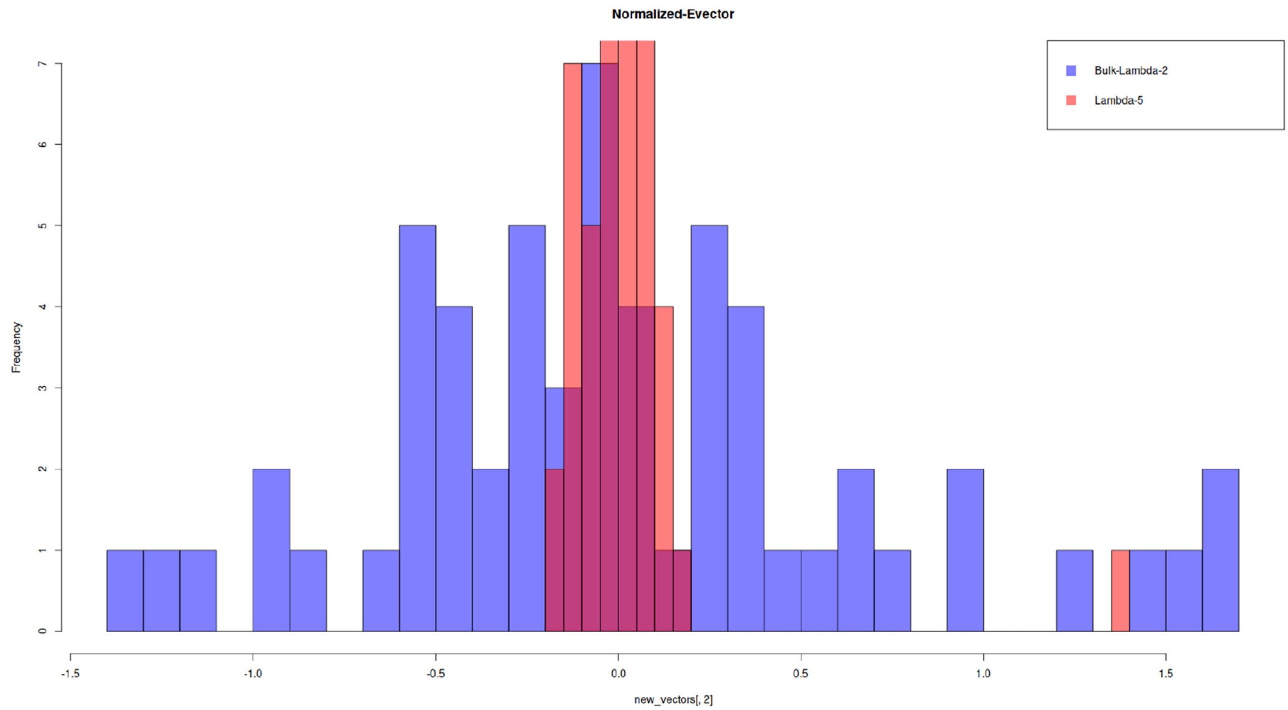

Fig. 5. Distribution $\rho(u)$ of eigenvector components for two eigenvalues, one from the bulk $\lambda_- < \lambda < \lambda_+$ (blue) and one from $\lambda_i > \lambda_+$ (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

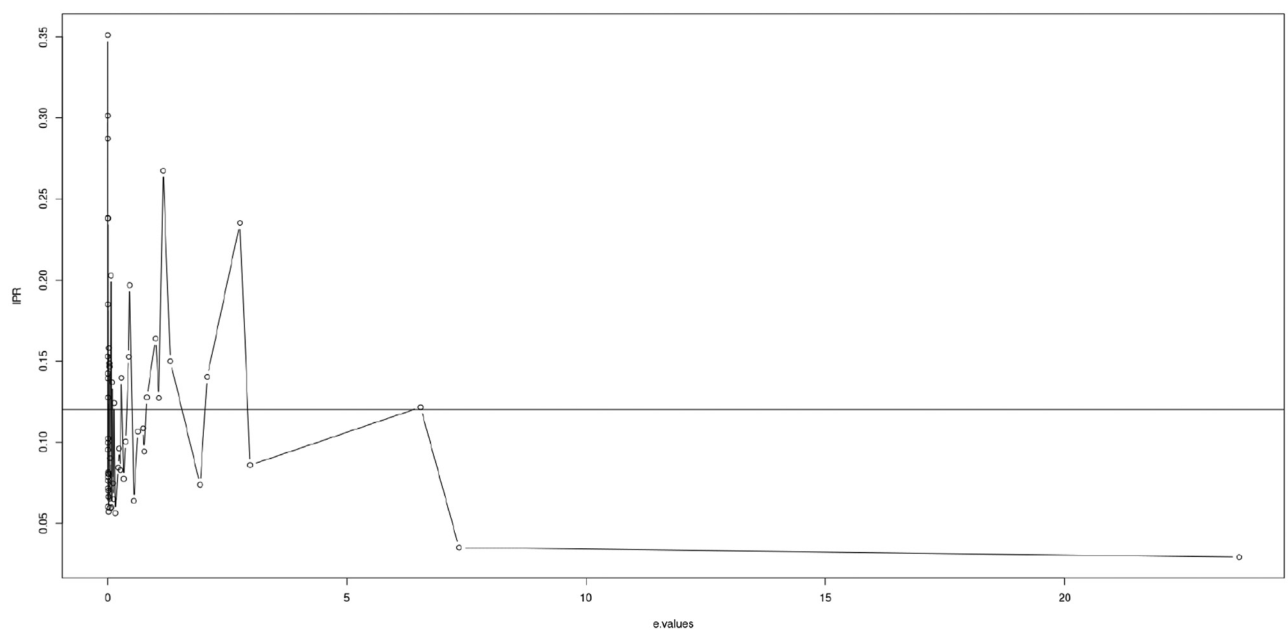

Fig. 6. Inverse participation ratio (IPR) as a function of eigenvalue λ for the cross correlation matrix C.

the degree of deviation of the distribution of eigenvectors from RMT [17]. We use the Inverse Participation Ratio[6] (IPR) [36,37] to quantify the following, as defined by:

$$I^i = \sum_{k=1}^{N} [u_k^i]^4, \tag{9}$$

where u_k^i, $k = 1, \ldots, N$ are the components of eigenvector u^i.

[6] Inverse Participation Ratio essentially can be interpreted as the reciprocal of the number of eigenvector components that contribute significantly.

Fig. 6 shows that the average value of I^i is 0.12, suggesting that almost all cryptocurrencies contribute to the eigenvectors, which was also evident in the research done by Stosic et al. [23]. They also suggested that lack of deviations from I at the end of the eigenvalue spectrum implies that eigenvectors are not localized.

5.2. Community structures from minimum spanning trees

Financial markets have been previously modelled by translating their correlations into networks by using a distance matrix D defined as (Tables 3 and 4)

Table 3
Distinct communities groups by colour for first time window of 150 days of 59 cryptocurrencies.

Groups by colour							
Red	Lgreen	Purple	Blue	Orange	Green	Black	Brown
Agoras-tokens	Bitshares	Bitbay	Decred	Byteball	Ardor	Bytecoin-bcn	Monacoin
Augur	Digibyte	Blocknet	Digixdao	Golem-network-tokens	Bitcoindark	Counterparty	Stellar
Bitcoin	Dogecoin	Ethereum	Ethereum-classic	Iconomi	Dash	Emercoin	Supernet-unity
Peercoin	Nav-coin	Factom	Experience-points	Neo	Digitalnote	Gulden	Syscoin
Ripple	Nxt	Ion	Lisk	Paccoin	Einsteinium	Maidsafecoin	Vertcoin
		Nem	Singulardtv	Zclassic	Gamecredits	Monero	
		Nexus	Triggers	Zcoin	Pivx	Ubiq	
		Siacoin	Waves		Reddcoin		
		Steem	Zcash		Viacoin		
		Stratis					
		Tether					
		Verge					

Table 4
Distinct communities groups by colour for second time window of 150 days of 59 cryptocurrencies.

Groups by colour							
Red	Lgreen	Purple	Blue	Orange	Green	Black	Brown
Bitshares	Ethereum	Augur	Bitcoindark	Bitcoin	Ardor	Agoras-Tokens	Decred
Dogecoin	Nem	Bitbay	Bytecoin-bcn	Experience-points	Dash	Byteball	Digixdao
Einsteinium	Nexus	Blocknet	Digibyte	Ion	Ethereum-classic	Golem-network-tokens	Lisk
Monero	Siacoin	Counterparty	Digitalnote	Neo	Monacoin	Iconomi	Pivx
Nxt	Tether	Emercoin	Gulden	Peercoin	Stellar	Singulardtv	
Syscoin		Factom	Maidsafecoin	Ripple	Stratis	Steem	
		Gamecredits	Nav-coin		Vertcoin	Triggers	
		Paccoin	Reddcoin		Waves	Zcash	
		Supernet-unity	Viacoin		Zcoin	Zclassic	
		Ubiq					
		Verge					

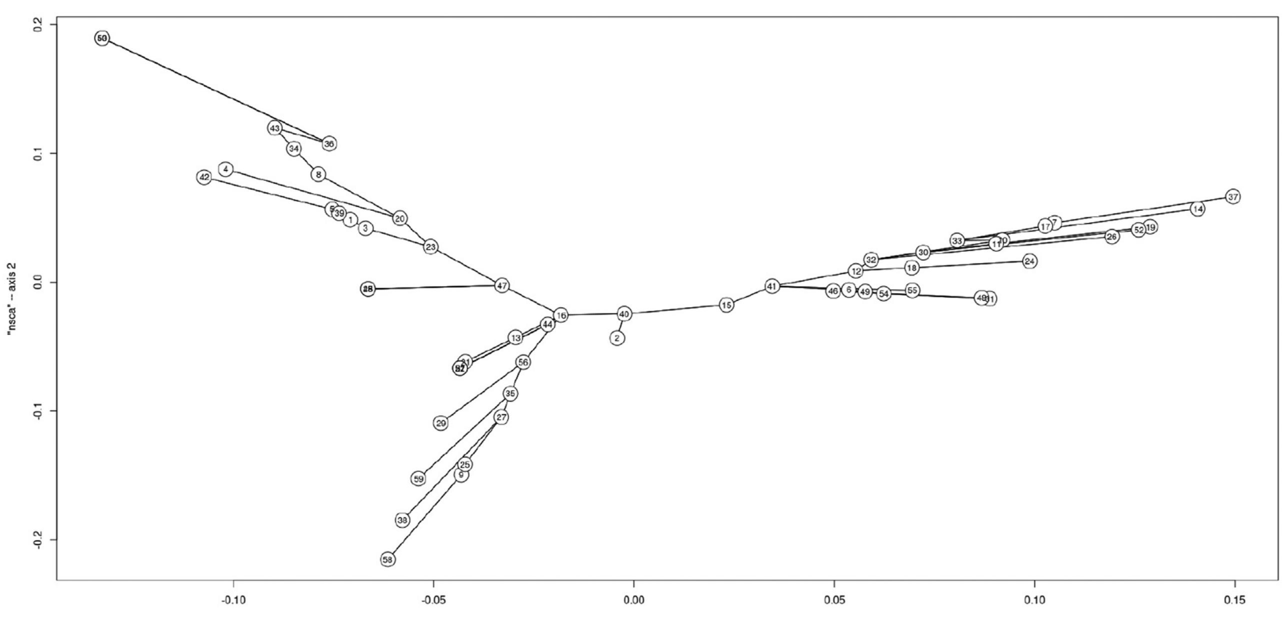

Fig. 7. Minimum spanning tree from cross-correlations of the cryptocurrencies. Each node number corresponds to the order of cryptocurrencies listed in Table 2.

$$D_{ij} = \sqrt{2(1 - C_{ij})}, \tag{10}$$

Network-based models for studying correlations in financial time series have been proposed by, e.g. [38]. In these methods the distance matrix D are defined as adjacency matrix, $A = D$. The Minimum Spanning Tree (MST) is the most-used network in the analysis of financial time series. MSTs are types of network that connect all nodes without having any loops. Therefore these have N nodes, and $N-1$ edges to connect them. The minimum refers to the fact that the sum of all edges is minimum for all spanning tress defined on the distance matrix [38].

We next construct the MST using a smaller subset of cryptocurrencies (59) to analyze their collective behaviour in the market. Fig. 7 shows the MST network for the correlation matrix C with $N = 59$ and $L = 150$. We can see that the network is split into vari-

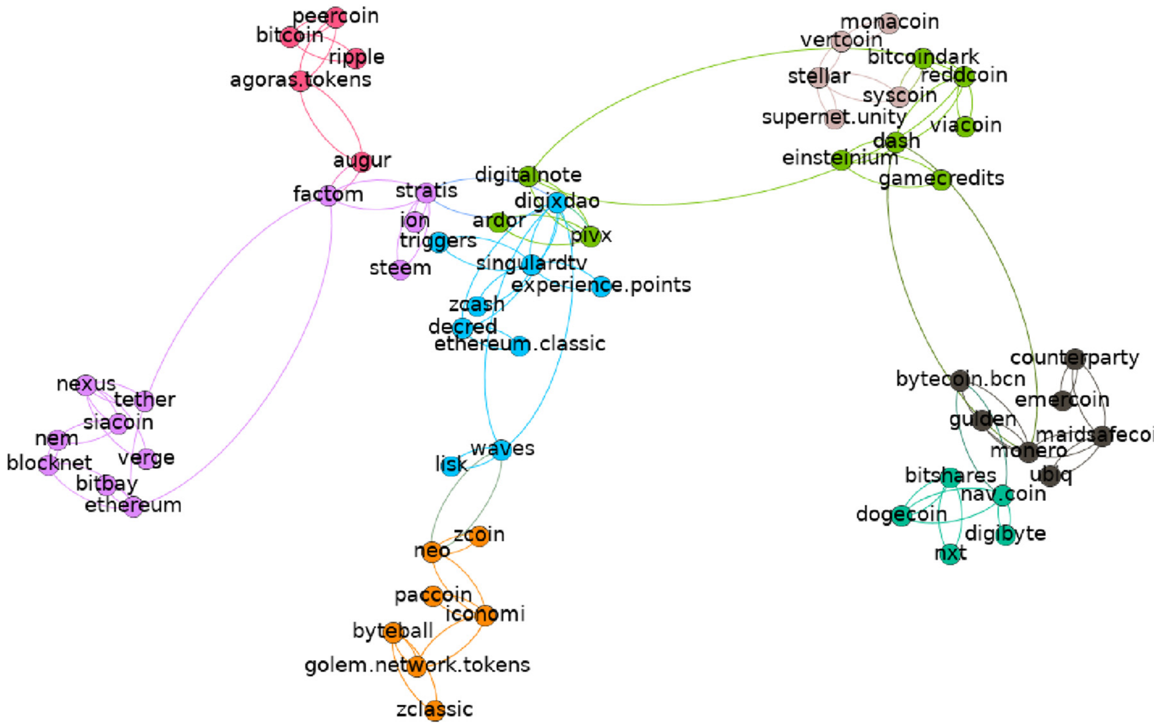

Fig. 8. Minimum spanning tree with distinct communities for 59 cryptocurrencies for first time window of 150 days. Each colour represent a unique community. (The distances between nodes are not actual representations from the distance matrix.) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ous groups and densely connected. However, this network is not enough to quantify the structures among them. Therefore, we further move to community detection methods such as developed by Blondel et al. [39]. Their algorithm is similar to other community detection algorithms such as Girvan–Newman [40]. Although, it uses heuristics that are based on modularity optimization which is more robust and requires less computing.

We use the community detection algorithm to reveal distinct communities as shown in Figs. 8 and 9 for the two time windows from the cross-correlation matrix of 59 cryptocurrencies respectively. These distinct communities represent (we speculate) collective behaviour in the market. This would seem to be in contrast to the concept of some major cryptocurrencies influencing the whole market, as advanced earlier by Stosic et al. [23]. We discover eight distinct communities in both the periods studied. Also, we find that the community structures change over time and exhibit few similarities over the period studied. Unlike Stosic et al. [23], we conclude that, over the two periods, the cryptocurrencies in the communities are not persistent.

We suspect this could be to do with the changes led by exogenous effects studied previously by Tan et al. [41]. They use the structural change model to study the effects of exogenous variables on the market. The authors use 2 different types of cryptocurrency indices which shows that the changes occur frequently for the price series followed by changes in return series, although this is not consistently the case. The use of 2 cryptocurrency indices may not be useful to capture the whole market's movement due to its fast-changing nature. This is because the cryptocurrencies constantly change their ranks based on the market capitalization. The authors advise financial researchers to take into account these instabilities of parameters that may exist in all aspects of the cryptocurrency analysis and modelling process due to the frequent existence of change points affected by underlying internal or external factors. Our research attempts to capture the underlying interactions within the market and we need more data of various cryptocurrency indices to capture these exogenous effects on ranks and market capitalization.

Further work by Sovbetov [42] studies the factors influencing prices of most common five CC such Bitcoin, Ethereum, Litecoin, Dash and Monero. First, they find market beta, trading volume, and volatility appear to be significant determinant for all five cryptocurrencies both in short and long run. Second, the trends led by attractiveness have impact on price, but only in long run. They also examine relation with S&P500 index and discover it seems to have weak positive long run impact on Bitcoin, Ethereum and Litecoin. Similarly, Poyser [43] defines three types of cryptocurrency influence factors organized into internal and external factors: Supply and demand being the major internal factor whereas attractiveness, legalization, and some macro-finance factors being the external factors.

This leads to several questions about these groupings such as why do they group? Do they group because of similar underlying framework? or Do they group because trading volumes are similar? remain a subject of future work. We need more data to confirm whether similar cryptocurrencies in communities persists over a long period.

Furthermore, we study and compare the cross-correlation dynamics of similar community groupings from the two periods. We take the cross-correlation matrices C_1 for first time window with $L = 150$ and C_2 for second time window with $L = 150$. Next, we compare their correlation matrices by selecting communities with similar colour groupings from Figs. 8 and 9 . By visual inspection, we find that most communities exhibit a transition from positive to negative correlation. Particularly, in Figs. 10 and 11 negative correlation is dominant in their second period. Interestingly, we find that for the community in Fig. 12 the positive correlations are dominant in the next period, suggesting that a community specific behaviour exists in the market. Our visual findings are also con-

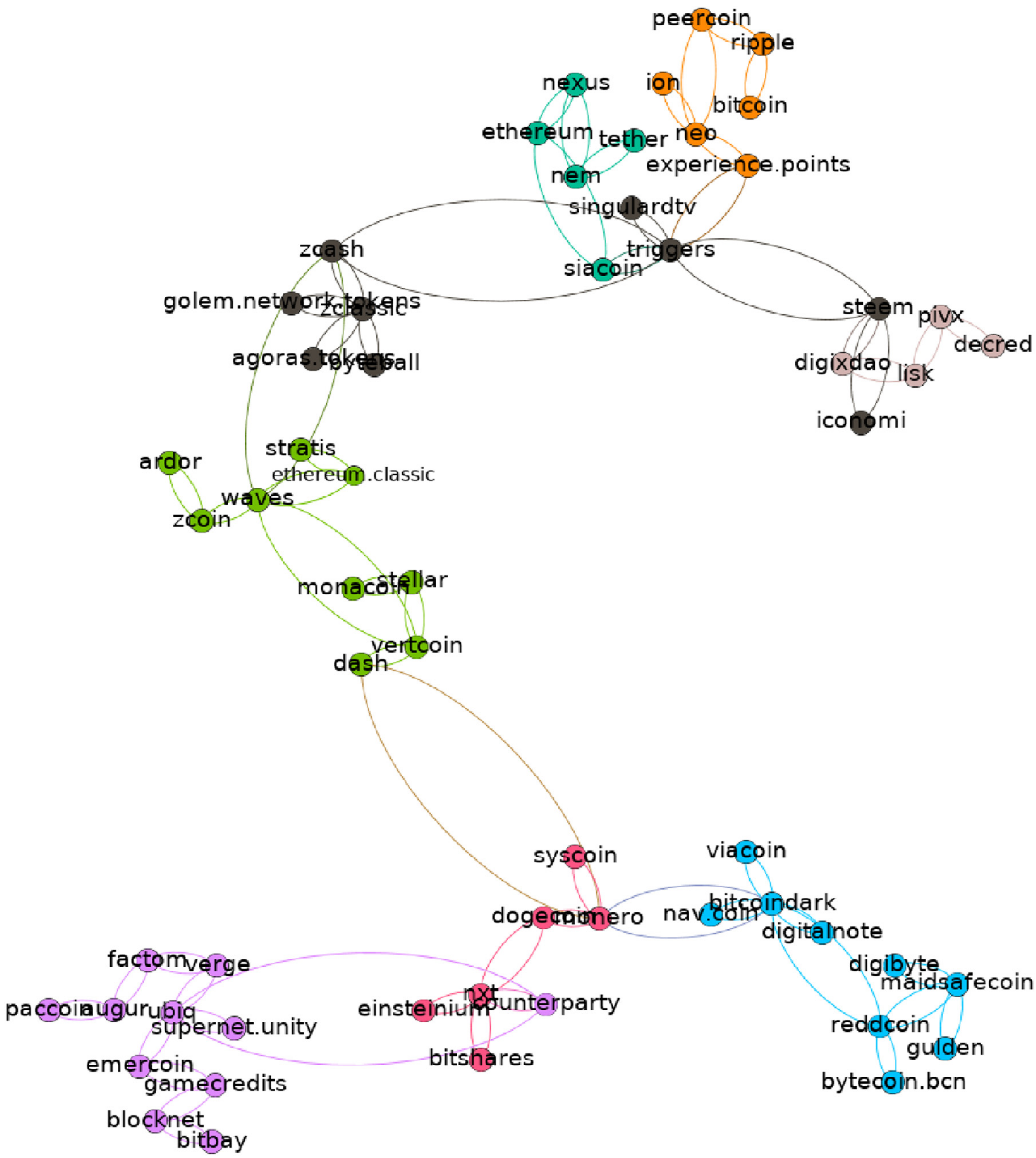

Fig. 9. Minimum spanning tree with distinct communities for 59 cryptocurrencies for second time window of 150 days. Each colour represent a unique community. (The distances between nodes are not actual representations from the distance matrix.) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

firmed by examining the changes in the largest eigenvalues from both matrices C_1 and C_2.

5.3. *Possible implications of the results*

These communities can have certain economic implications as using these community groupings can be useful for constructing diverse investment portfolios. This was previously suggeseted by Mensi et al. [44] who study Bitcoin and five major cryptocurrenies (Dash, Ethereum, Litecoin, Monero and Ripple) and examine their portfolio risk implications. They consider different portfolio strategies and examine the implications of diversification. The results show that use of mixed portfolio strategy provides better diversification and risk reductions for portfolio managers and investors. Also, the findings show such risk minimization is time-horizon dependent which suggests investors need to be wary of changes in each community grouping for different time periods. Although we would suggest that implications of these changes remain to be confirmed as part of future work.

Moreover, Zhang et al. [45] consider a value-weighted cryptocurrency index and compare the cross-correlations with the DJIA index. They find persistent cross-correlation exists between the two. They use 9 different cryptocurrenies in their analysis and conclude that all the cryptocurrenies studied are efficient and that further work can be done to study their changing degree of efficiency and investigate hedging properties. Zhang et al. [46] have studied some stylized facts of 8 different cryptocurrenies which represents almost 70% cryptocurrency market capitalization

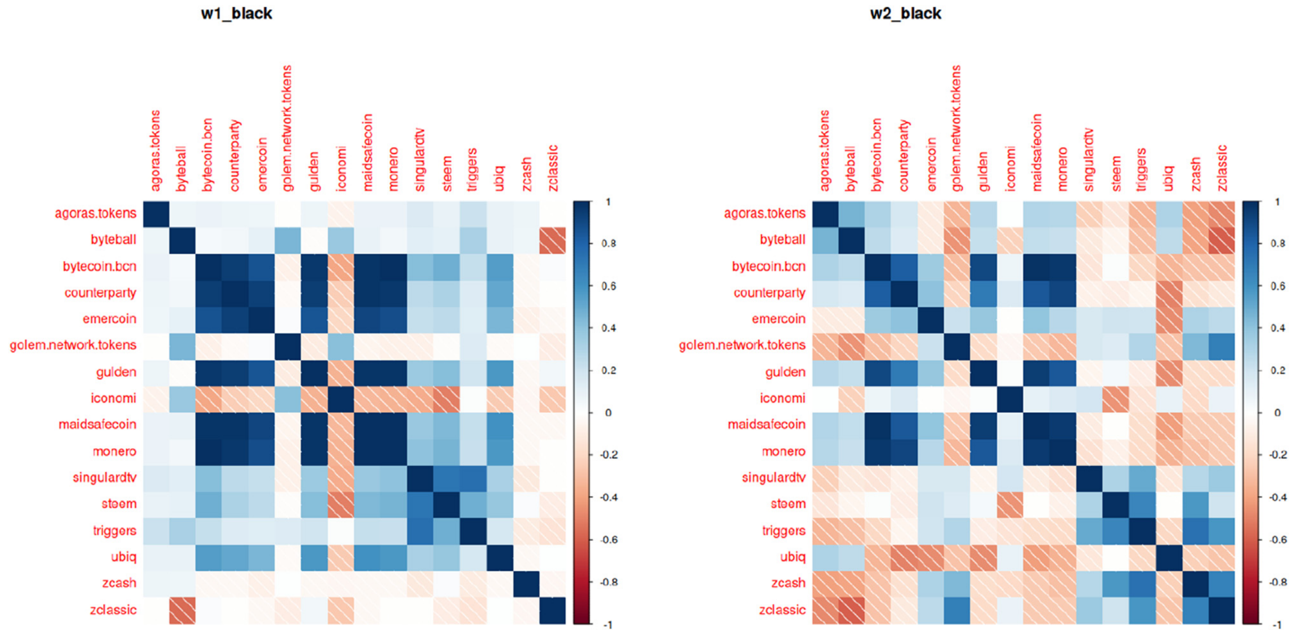

Fig. 10. cross-correlation matrices C_1 and C_2 for communities in 'Black'.

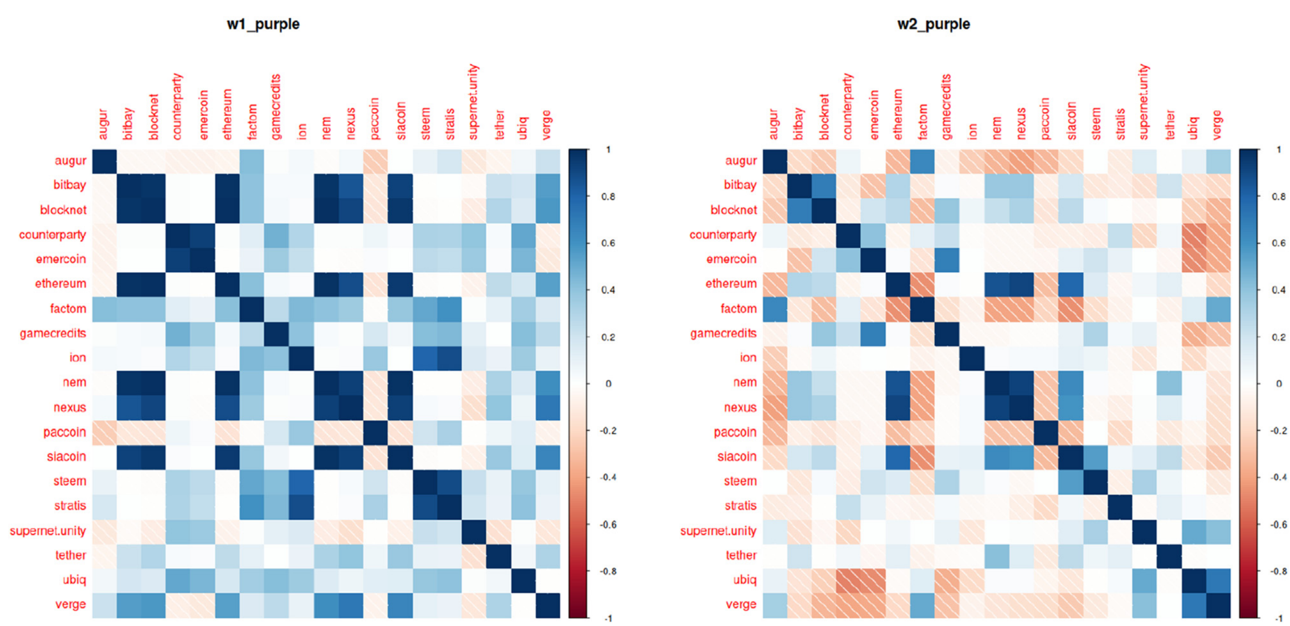

Fig. 11. cross-correlation matrices C_1 and C_2 for communities in 'Purple'.

there exists. Their findings reveal heavy tails for all the returns of cryptocurrenies, an absence of autocorrelations, returns of cryptocurrencies displaying strong volatility clustering, leverage effects and long-range dependence for the return of cryptocurrencies. The authors suggest investors in cryptocurrencies should take into account these stylized facts. However, the authors must confess an inability to satisfactorily explain the difference between returns and volatility in long range dependence in the case of cryptocurrencies and whether this has to do with the number of determinants of these for this particular asset class is an open question.

6. Conclusion

In summary, we study the collective behaviour for the cryptocurrency market using correlations of 119 and 59 cryptocurrencies. We find and verify that cross-correlations matrix have non-trivial structures and groupings among cryptocurrency pairs. Also, we discover for different numbers of cryptocurrencies and time periods the eigenvalue spectrum does not always agree with universal predictions of Random Matrix Theory.

We analyze the eigenvector components to validate their influence on the market. Furthermore, we construct minimum spanning trees and discover distinct community structures. Although, these communities structures do not persist over time but cross-correlation dynamics suggests a collective behaviour exists among these communities. Lastly, we conclude our analysis of community groupings can be useful to construct cryptocurrency investment portfolios.

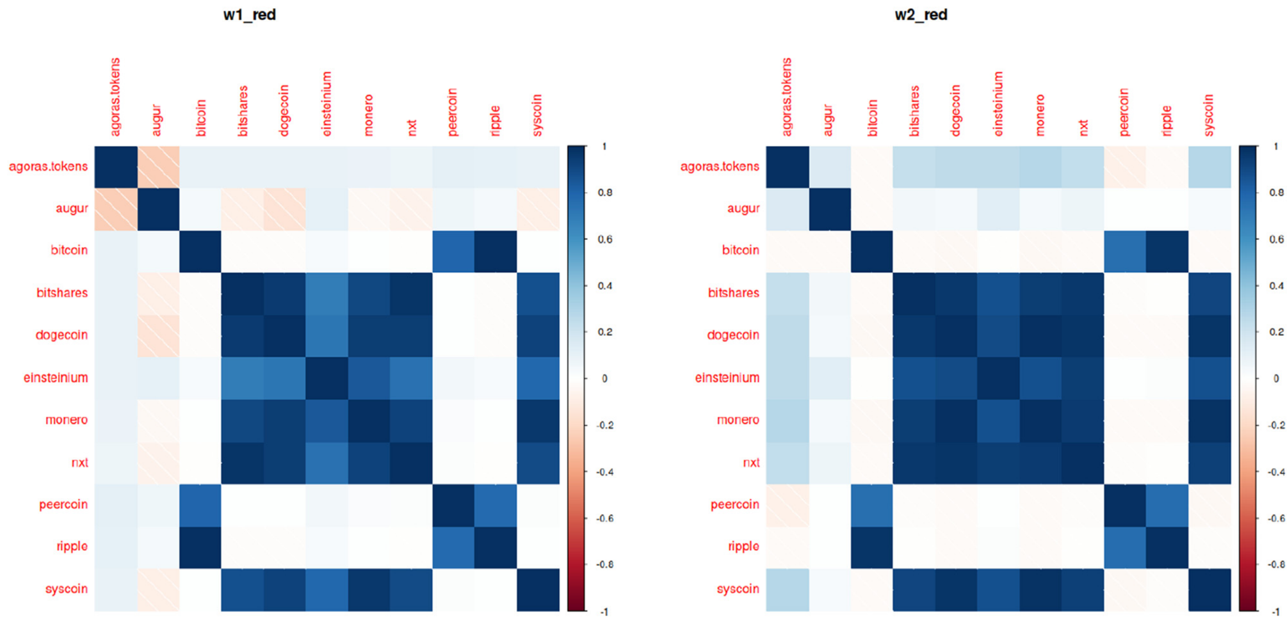

Fig. 12. cross-correlation matrices C_1 and C_2 for communities in 'Red'.

Authors' contribution

Harshal Chaudhari: conceptualization, software, formal analysis, investigation, data curation, writing – original draft, visualization. Martin Crane: methodology, validation, resources, writing – review & editing, supervision, project administration.

Conflict of interest

None declared.

Acknowledgements

The authors are indebted to two anonymous referees the suggestions of whom enabled us improve the paper substantially. This research was (partially) conducted with the financial support of Science Foundation Ireland (SFI) under Grant Agreement No. 17/SP/5447 at the ADAPT SFI Research Centre at DCU. The ADAPT SFI Centre for Digital Media Technology is funded by Science Foundation Ireland through the SFI Research Centres Programme and is co-funded under the European Regional Development Fund (ERDF) through Grant no. 13/RC/2106.

References

[1] J. Yli-Huumo, D. Ko, S. Choi, S. Park, K. Smolander, Where is current research on blockchain technology? – a systematic review, PLOS ONE 11 (10) (2016), http://dx.doi.org/10.1371/journal.pone.0163477.
[2] Binance Cryptocurrency Exhance. https://www.binance.com/en (Accessed 30 June 2010.
[3] N.A. Bakar, S. Rosbi, Autoregressive integrated moving average (ARIMA) model for forecasting cryptocurrency exchange rate in high volatility environment: a new insight of bitcoin transaction, Int. J. Adv. Eng. Res. Sci. 4 (11) (2017) 130–137, http://dx.doi.org/10.22161/ijaers.4.11.20.
[4] A. Urquhart, The inefficiency of Bitcoin, Econ. Lett. 148 (2016) 80–82, http://dx.doi.org/10.1016/j.econlet.2016.09.019.
[5] S. Nison, Japanese Candlestick Charting Techniques, Simon and Schuster, 1991.
[6] M.P. Taylor, H. Allen, The use of technical analysis in the foreign exchange market, J. Int. Money Finance 11 (3) (1992) 304–314.
[7] E.F. Fama, Efficient capital markets: a review of theory and empirical work, J. Finance 25 (2) (1970) 383–417, http://dx.doi.org/10.2307/2325486.
[8] B.G. Malkiel, Efficient Market Hypothesis, Palgrave Macmillan UK, London, 1989, pp. 127–134, http://dx.doi.org/10.1007/978-1-349-20213-3_13.
[9] H. Ni, H. Yin, Exchange rate prediction using hybrid neural networks and trading indicators, Neurocomputing 72 (13–15) (2009) 2815–2823, http://dx.doi.org/10.1016/j.neucom.2008.09.023.
[10] I. Kertesz, I. Kondor, Econophysics: An Emerging Science, Kluwer, 1999 https://books.google.ie/books?id=rm53PgAACAAJ.
[11] R. Mantegna, H. Stanley, An Introduction to Econophysics: Correlations and Complexity in Finance, vol. 53, 2000, http://dx.doi.org/10.1063/1.1341926.
[12] J. Kwapien, S. Drozdz, Physical approach to complex systems, Phys. Rep. V 515 (2012) 115–226, http://dx.doi.org/10.1016/j.physrep.2012.01.007.
[13] E. Elton, M. Gruber, S. Brown, W. Goetzmann, Modern Portfolio Theory and Investment Analysis, John Wiley & Sons, 2007.
[14] J.-P. Bouchaud, M. Potters, Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management.
[15] O. Rosso, H. Larrondo, M. Martin, A. Plastino, M. Fuentes, Distinguishing noise from chaos, Phys. Rev. Lett. 99 (2007) 154102, http://dx.doi.org/10.1103/PhysRevLett.99.154102.
[16] V. Plerou, P. Gopikrishnan, B. Rosenow, L. Amaral, H. Stanley, Universal and nonuniversal properties of cross correlations in financial time series, Phys. Rev. Lett. 83 (1999) 1471–1474, http://dx.doi.org/10.1103/PhysRevLett.83.1471.
[17] V. Plerou, P. Gopikrishnan, B. Rosenow, L.A.N. Amaral, T. Guhr, H.E. Stanley, A random matrix approach to cross-correlations in financial data, Phys. Rev. E 65 (6) (2002), http://dx.doi.org/10.1103/PhysRevE.65.066126.
[18] L. Laloux, P. Cizeau, J.-P. Bouchaud, M. Potters, Noise dressing of financial correlation matrices, Phys. Rev. Lett. 83 (1999) 1467, http://dx.doi.org/10.1103/PhysRevLett.83.1467.
[19] V. Plerou, P. Gopikrishnan, B. Rosenow, L. Amaral, H. Stanley, Collective behavior of stock price movements – a random matrix theory approach, Physica A 299 (2001) 175–180, http://dx.doi.org/10.1016/S0378-4371(01)00293-X.
[20] B. Podobnik, D. Wang, D. Horvatic, I. Grosse, H. Stanley, Time-Lag Cross-Correlations in Collective Phenomena, 2010, http://dx.doi.org/10.1209/0295-5075/90/68001.
[21] J. Shen, B. Zheng, Cross-correlation in financial dynamics, EPL (Europhys. Lett.) 86 (4) (2009) 48005, http://dx.doi.org/10.1209/0295-5075/86/48005.
[22] P. Gopikrishnan, B. Rosenow, V. Plerou, H.E. Stanley, Identifying Business Sectors from Stock Price Fluctuations, 2000 arXiv:cond-mat/0011145.
[23] D. Stosic, D. Stosic, T.B. Ludermir, T. Stosic, Collective behavior of cryptocurrency price changes, Physica A: Stat. Mech. Appl. 507 (2018) 499–509, http://dx.doi.org/10.1016/j.physa.2018.05.050.
[24] T. Ankenbrand, D. Bieri, Assessment of cryptocurrencies as an asset class by their characteristics, Invest. Manag. Finan. Innov. 15 (2018) 169–181, http://dx.doi.org/10.21511/imfi.15(3).2018.14.
[25] D.G. Baur, T. Dimpfl, Asymmetric volatility in cryptocurrencies, Econ. Lett. 173 (C) (2018) 148–151.
[26] G.M. Caporale, L. Gil-Alana, A. Plastun, Persistence in the cryptocurrency market, Res. Int. Business Finance 46 (2018) 141–148, http://dx.doi.org/10.1016/j.ribaf.2018.01.002.
[27] A.F. Bariviera, The inefficiency of Bitcoin revisited: a dynamic approach, Econ. Lett. 161 (2017) 1–4, http://dx.doi.org/10.1016/j.econlet.2017.09.013.
[28] J. Rebane, I. Karlsson, S. Denic, P. Papapetrou, Seq2seq RNNs and ARIMA Models for Cryptocurrency Prediction: A Comparative Study, KDD Data Science in Fintech Workshop, 2018, pp. 5.

[29] E. Karakoyun, A. Cibikdiken, Comparison of ARIMA time series model and LSTM deep learning algorithm for Bitcoin price forecasting, International Academic Conference in Prague 2018 (2018) 171–179.
[30] M. Kinderis, M. Bezbradica, M. Crane, Bitcoin Currency Fluctuation, in: Proceedings of the 3rd International Conference on Complexity, Future Information Systems and Risk, SCITEPRESS – Science and Technology Publications, 2018, pp. 31–41, http://dx.doi.org/10.5220/0006794000310041.
[31] A.H. Dyhrberg, Hedging capabilities of Bitcoin. Is it the virtual gold? Finance Res. Lett. 16 (2016) 139–144, http://dx.doi.org/10.1016/j.frl.2015.10.025.
[32] T. Klein, H. Pham Thu, T. Walther, Bitcoin is not the New Gold – a comparison of volatility, correlation, and portfolio performance, Int. Rev. Finan. Anal. 59 (2018) 105–116, http://dx.doi.org/10.1016/j.irfa.2018.07.010.
[33] M. Caporin, M. McAleer, Ten things you should know about the dynamic conditional correlation representation, Econometrics 1 (1) (2013) 115–126, http://dx.doi.org/10.3390/econometrics1010115.
[34] A.S.W. Jagdish Mehra, The Collected Works of Eugene Paul Wigner: Part A: The Scientific Papers, 1st ed., Springer-Verlag Berlin Heidelberg, 1993.
[35] M.L. Mehta, F.J. Dyson, Statistical theory of the energy levels of complex systems. V, J. Math. Phys. 4 (5) (1963) 713–719, http://dx.doi.org/10.1063/1.1704009.
[36] T. Guhr, A. Mueller-Groeling, H.A. Weidenmueller, Random matrix theories in quantum physics: common concepts, Phys. Rep. 299 (4-6) (1998) 189–425, http://dx.doi.org/10.1016/S0370-1573(97)00088-4.
[37] Y.V. Fyodorov, A. Mirlin, Analytical derivation of the scaling law for the inverse participation ratio in quasi-one-dimensional disordered systems, Phys. Rev. Lett. 69 (1992) 1832-L.
[38] R. Mantegna, Hierarchical Structure in Financial Markets, Quantitative Finance Papers 11. doi:10.1007/s100510050929.
[39] V. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks, J. Stat. Mech. Theory Exp. (2008), http://dx.doi.org/10.1088/1742-5468/2008/10/P10008.
[40] M. Girvan, M.E.J. Newman, Community structure in social and biological networks, Proc. Natl. Acad. Sci. U.S.A. 99 (2001) 7821–7826.
[41] C.Y. Tan, Y.B. Koh, K.H. Ng, K.H. Ng, Structural Change Analysis of Active Cryptocurrency Market, 2019 arXiv:1909.10679.
[42] Y. Sovbetov, Factors influencing cryptocurrency prices: evidence from bitcoin, ethereum, dash, litcoin, and monero, J. Econ. Finan. Anal. 2 (2) (2018) 1–27, http://dx.doi.org/10.1991/jefa.v2i2.a16.
[43] O. Poyser, Exploring the Determinants of Bitcoin's Price: An Application of Bayesian Structural Time Series, 2017 arXiv:1706.01437.
[44] W. Mensi, M.U. Rehman, K.H. Al-Yahyaee, I.M.W. Al-Jarrah, S.H. Kang, Time frequency analysis of the commonalities between Bitcoin and major Cryptocurrencies: Portfolio risk management implications, N. Am. J. Econ. Finance 48 (2019) 283–294, http://dx.doi.org/10.1016/j.najef.2019.02.013.
[45] W. Zhang, P. Wang, X. Li, D. Shen, The inefficiency of cryptocurrency and its cross-correlation with Dow Jones Industrial Average, Physica A: Stat. Mech. Appl. 510 (2018) 658–670, http://dx.doi.org/10.1016/j.physa.2018.07.032.
[46] W. Zhang, P. Wang, X. Li, D. Shen, Some stylized facts of the cryptocurrency market, Appl. Econ. 50 (55) (2018) 5950–5965, http://dx.doi.org/10.1080/00036846.2018.1488076.

Harshal Chaudhari received his M.Sc. in Computing majoring in Data Analytics from the School of Computing at Dublin City University, Ireland. His main research interests include time series analysis, financial data analysis and forecasting. Prior to this he obtained his B.Eng. in Computer Engineering from the University of Pune, India. He is also involved in the open source community and passionate about applying machine learning algorithms to interdisciplinary fields of statistics, data analytics, applied mathematics and high performance computing.

Martin Crane received his B.A. B.A.I. (Mech. Eng.) degrees from Trinity College Dublin in 1989 and a Ph.D. in Numerical Analysis from the same institution in 1993. He has been carrying out research at DCU School of Computing since 1999 and Associate Professor since 2007. My areas of Research include many fields of Complex Systems in Quantitative Finance and Systems Biology. He is the author of more than 80 journal papers