Volatility and returns connectedness in cryptocurrency markets: Insights from graph-based methods

An Pham Ngoc Nguyen [a,b,*], Tai Tan Mai [a,c], Marija Bezbradica [a,c], Martin Crane [a,c]

[a] School of Computing, Dublin City University, Collins Ave Ext, Whitehall, Dublin, D09 W6Y4, Ireland
[b] SFI Centre for Research Training in Artificial Intelligence, Dublin, D02 FX65, Ireland
[c] ADAPT Center for Digital Content Technology, Dublin, D02 PN40, Ireland

ABSTRACT

We employ graph-based methods to examine the connectedness between cryptocurrencies of different market caps over time. By applying denoising and detrending techniques inherited from Random Matrix Theory and the concept of the so-called *Market Component*, we are able to extract new insights from historical return and volatility time series. Notably, our analysis reveals that changes in volatility-based network structure can be used to identify major events that have, in turn, impacted the cryptocurrency market. Additionally, we find that these structures reflect investors' sentiments, including emotions like fear and greed. Using metrics such as *PageRank*, we discover that certain minor coins unexpectedly exert a disproportionate influence on the market, while the largest cryptocurrencies such as BTC and ETH seem less influential. We suggest that our findings have practical implications for investors in different ways: Firstly, helping them to avoid major market disruptions such as crashes, to safeguard their investments, and to capitalize on opportunities for high returns; Secondly, sharpening and optimizing the portfolios thanks to the understanding of cryptocurrencies' connectedness.

1. Introduction

Cryptocurrencies have become renowned for their sudden and unexpected price fluctuations [1]. While such price movements can be concerning for many investors — particularly inexperienced ones [2,3], observing common patterns across different cryptocurrencies can provide valuable insights to all investors. Specifically, by understanding the correlation in the volatility of prices between cryptocurrencies, investors can potentially gain high profits while avoiding significant losses [4].

Several methods exist to analyze the correlation between different cryptocurrencies and between cryptocurrencies and traditional assets (e.g. stocks, bonds, oil and fiat currencies). One of the first papers on this topic was conducted by Yermack [5], who tested the daily dependency between Bitcoin and national currencies (e.g., EUR, JPY, GBP) using *Pearson* correlation measurement. Since then, a variety of statistics-based correlation measures such as Wavelet Coherence, Vector Autoregressive (VAR), Granger causality and q-Dependent Detrended Correlation have been widely adopted on larger datasets, comprising different cryptocurrencies and asset classes. This diversification paints a more comprehensive picture of the dependencies present in the financial markets [6–9]. Notably, with the fast growth of Deep Learning, several methods such as Multilayer Perceptron (MLP) and Long Short Term Memory (LSTM) have been successfully applied to cryptocurrency-related correlation studies [10,11]. In summary, researchers found evidence of time-varying connectedness between cryptocurrencies themselves and between the cryptocurrency market and

traditional markets. They discovered that these connections tend to fluctuate, either increasing or decreasing frequently. However, one consistent finding is that different asset classes become more closely related to each other during the Covid-19 pandemic and economic crises. Additionally, the cryptocurrency market is becoming more mature over time, and its correlation with other asset classes gradually increases.

Recently, graph-based methods have seen wider applications in learning about co-movements and spillover effects in financial markets [12–14]. This is due to several factors: (1) their ability to observe the time-varying connectedness among multiple objects; (2) the diversity of metrics available in Graph Theory to analyze graph structures — such as *betweenness centrality*, *degree assortativity*, and *closeness centrality*. These metrics can be used to learn and understand the underlying characteristics of a graph; (3) Their ease of implementation in comparison with statistical and Neural Networks methods. This method creates a network of various objects where the distance between them is determined by their similarity.[1] The closer the objects are, the more alike they are. In the realm of financial markets, the measurement of similarity between two assets relies on their co-movement as reflected in their corresponding financial time series, such as prices, returns, and volatility. When two assets are completely similar, any upward or downward movement in one asset is mirrored simultaneously in the other. By examining the structure of the network and analyzing its features, we can observe the interactions between the objects.

In [13], we adopted the method described above to construct time-varying network structures using cryptocurrencies' return time series over a 2-year period from 2019 to 2021. This period covers different market conditions, including normal times and downturn times (e.g., the market crash in March 2020 due to the Covid-19 pandemic). We found that the network structure reflects investors' investment decisions. During normal times, investors arguably make their own investment decisions based on their personal market analysis and experience. On the other hand, during turbulent times, investors appear to trade only cryptocurrencies with high market capitalization, while smaller cryptocurrencies are mainly used for other purposes such as transaction fees, smart contract tokens or simply to run a digital platform.

One possible gap recognized in our previous work as well as in the existing literature is the limited use of volatility time series as the main data in cryptocurrencies-related experiments since very few works have mentioned this indicator. However, as proved in [4], return values cannot explain all phenomena in the cryptocurrency market. Instead, volatility information might also add more insights to the return ones. From this point of view, we contribute to the existing literature by using the volatility to observe the connectedness between different cryptocurrencies. In particular, we will use cryptocurrencies' volatility and return time series to construct time-varying networks with a finer window size, allowing us to capture not only normal and downturn times but also times when bull market conditions pertain. Our research questions are described as follows:

- RQ1: Is there evidence of a difference between returns-based network structures and volatility-based network structures? In other words, do volatility-based networks show different behaviors from returns-based networks?
- RQ2: If so, can we gain insights from these results in understanding the underlying structure (if any) in the cryptocurrency market? By this, we mean, what can we learn from these results and where to use them?

For our research questions, we use a 30-minute dataset sourced from the HitBTC exchange,[2] consisting of 34 closing price time series corresponding to the 34 most active cryptocurrencies. We rely on *Pearson* correlation and Minimum Spanning Tree (MST) to construct cryptocurrency networks. To observe the evolution of the network structure, we adopt 3 graph-based metrics, e.g. *betweenness centrality*, *degree assortativity* and *closeness centrality*. Additionally, we utilize the famous algorithm used by Google called *PageRank*[3] to examine the importance of each cryptocurrency in the network over time. Other features such as investors' sentiments and the number of transactions are also used to support and reinforce our findings.

Significantly, we remove the noise and trend in cryptocurrencies using Random Matrix Theory (RMT) and the Market Component concept. This technique has been widely used in other areas such as education [20] and stock markets [21,22] but is relatively new in the cryptocurrency market. We shall show later in this paper that the noise and trend removal scheme plays an important role in exploring the underlying characteristics of the cryptocurrency market, revealing several characteristics that become visible only after clearing the noise and trend.

The remainder of the article is organized as follows: Section 2 presents an overview of related works. Section 3 provides a description of the datasets. Section 4 discusses methodology, metrics and preprocessing procedures. Section 5 describes the experimental results followed by implications and hypothesis. Lastly, the conclusion of this study is given in Section 6.

2. Related works

2.1. Analysis of correlations in the cryptocurrency market and graph-based methods

The correlation between different financial assets has been the subject of extensive research over a considerable period of time, particularly in the realm of traditional markets (e.g. stocks, bonds, fiat currencies) [23–28]. This topic helps scholars and

---

[1] Several methods for measuring the similarity between two time series such as *Pearson* [15], *Spearman* [16], Dynamic time warping [17], Cosine similarity [18], Euclidean distance [19].

[2] https://hitbtc.com/.

[3] *PageRank* is an algorithm used by Google Search to rank web pages in their search engine results. It is named after both the term "web page" and co-founder Larry Page. The algorithm assigns a numerical weighting to each element of a hyperlinked set of documents, such as the World Wide Web, with the purpose of "measuring" its relative importance within the set.

traders understand the underlying interactions among various assets, incorporating various advantages. For instance, knowing the correlations between different assets can empower investors to construct well-diversified portfolios by strategically allocating assets with low correlations, potentially resulting in lower portfolio volatility and improving risk-adjusted returns [29,30]. Moreover, knowledge of asset correlations can aid in identifying opportunities for hedging against potential losses. Specifically, by understanding the correlations between different financial assets, investors can identify assets that have negative or low correlation with the rest in their portfolio and use them as hedges potentially to reduce their overall portfolio risk [31,32]. Additionally, by expanding correlation analysis to microeconomic variables such as interest rate, inflation rate and GDP growth rate, investors can be helped to anticipate changes in future asset prices [33].

Since the inception of the cryptocurrency market, several studies have also been conducted to support crypto investors in making trading decisions. Two main empirical findings from them are as follows: Firstly, the most renowned cryptocurrencies (Bitcoin and Ethereum) show the strongest correlation with other virtual currencies in this market and they are also seen to interact with each other. This correlation is primarily attributed to the impact of people's sentiment, widely recognized as a dominant factor influencing market dynamics [34]; Significantly, the correlation observed in the crypto market changes regularly with several consistent characteristics during a specific market condition. That is, different cryptocurrencies tend to show distinct behaviors during normal times, leading to a low correlation between them. However, they tend to display higher levels of correlation during downturn times, which is mirrored among conventional stocks [26]. This phenomenon appears to stem from changes in the investment decisions of cryptocurrency investors in response to varying market conditions [13].

Numerous studies have examined the relationship between cryptocurrencies and traditional asset classes such as stocks, commodities, and bonds. A common finding is that the cryptocurrency market is gradually aligning itself with traditional markets. For example, Drozdz et al. conducted an experiment to observe the statistical features of 70 cryptocurrencies that are actively traded nowadays, they cover different aspects of cryptocurrency time series such as returns, volatility and temporal multifractal correlations [14]. Their study revealed that the highest-capitalization as well as more matured cryptocurrencies BTC and ETH share common statistical properties with the traditional financial markets. By contrast, smaller cryptocurrencies and less matured ones, e.g. DOGE, FUN, XLM and ONT, demonstrated some variations in this regard. Notably, a study conducted by Yosra et al. [32] examined the return characteristics of cryptocurrencies in relation to traditional asset classes and found that adding Bitcoin to the portfolio can reduce the risk and increase the returns, thereby playing a significant role in optimizing the portfolio and mitigating liquidity risk.

The adoption of graph-based methods for analyzing the correlation between different cryptocurrencies or between cryptocurrencies and conventional assets has grown steadily in recent years. Graph-based approaches offer unique capabilities such as ease of Visualization, Clustering, Centrality and Pathfinding [25], enabling comprehensive exploration of correlation-related issues within the cryptocurrency market, not only at a specific time but also in different periods [13,35] thus potentially offering a way to observe the time-varying interactions within a system. Moreover, this method is relatively easy to implement and is not computationally intensive as is the case with existing statistical and regression methods [36–38].

Analyzing graph-based correlations adds more insights into the financial markets. For instance, the authors in [35] used Minimum Spanning Tree[4] to construct a graph of more than 100 cryptocurrencies. Based on the graph structure, they discovered the existence of different communities within the market, where each community consists of several cryptocurrencies that have similar return movements. Although these community structures do not persist over time, cross-correlation dynamics suggest a collective behavior exists among these communities. This result holds potential for portfolio diversification strategies, which has been explored and justified in [34].

Given the widespread use of graph theory for analyzing the correlations between different cryptocurrencies, we notice that most studies tend to focus on basic financial time series such as return time series and price time series. However, another essential aspect that has received limited attention is volatility. Recognizing this research gap, we aim to contribute to the cryptocurrency literature by conducting a time-varying correlation analysis in the cryptocurrency market using volatility time series in this study.

2.2. The cryptocurrency market during critical events

In [39], the authors utilized more than 100 cryptocurrencies to investigate the cryptocurrency market between 2015 and 2020. Based on a regression model called cross-sectional absolute dispersion (CSAD), they found herd effects during the intense period of the coronavirus outbreak. By contrast, crypto traders have not engaged in correlated investing since the end of March 2020, when the market started to recover. Instead, during the recovery time, they showed a more rational investment decision due to the restoration of the market and assets' prices. Moreover, from the experiments, the authors indicated that even though crypto traders tend to follow each other during both bull and bear market periods, their tendency to herd investing is stronger during the bullish regime while there are more rational investors during the bearish regime. On the same topic, Samuel in [40] examined the presence of herding for four different periods, e.g. pre-and during Covid-19, bear and bull markets between April 2019 and January 2021. He provided the same results as in [39] since his experiment revealed that investors make similar trading decisions for positive market returns and during the Covid-19 pandemic period.

On the other hand, [41] observed the cryptocurrency market during the pandemic by using the price returns-based network approach. In particular, they investigated the structure of the Minimum Spanning Tree (MST) constructed from 128 cryptocurrencies

---

[4] A technique to reduce the size of a graph (network) in order to identify the so-called *community structure* in the graph, in which each community comprises nodes that share similar characteristics and nodes in different communities have different characteristics.

Table 1
A list of 34 cryptocurrencies used in this study. Abbreviations are put in parentheses.

Cryptocurrencies					
Argur (REP)	Bitcoin SV (BSV)	Ethereum Classic (ETC)	MaidSafeCoin (MAID)	Ontology (ONT)	Tron (TRX)
Bancor (BNT)	Cardano (ADA)	FunToken (FUN)	Maker (MKR)	Ox (ZRX)	Verge (XVG)
Basic Attention Token (BAT)	Decentraland (MANA)	ICON (ICX)	Monero (XMR)	QTUM	Zcash (ZEC)
Bitcoin (BTC)	Dogecoin (DOGE)	IOST	Nem (XEM)	Ripple (XRP)	Zilliqa (ZIL)
Bitcoin Cash (BCH)	EOS	Lisk (LSK)	NEO	Stellar (XLM)	
Bitcoin Gold (BTG)	Ethereum (ETH)	Litecoin (LTC)	OMG Network (OMG)	Tezos (XTZ)	

before and during the pandemic to see whether this critical event causes the change of this structure. Indeed, the results showed that the network was distributed during the year 2019, which was before the onset of Covid-19. This network structure means that the cryptocurrencies had independent trends. Whereas, a centralized network appeared during the first half of 2020, in which cryptocurrencies formed a big community with a central node being USDT. In other words, the cryptocurrencies were closer to each other, sharing a common trend.

The existing studies confirmed the significant changes in the cryptocurrency market during the pandemic. However, to our knowledge, there is a scarcity in terms of the analysis of cryptocurrencies during the bull market 2021. Our study aims to focus not only on the Covid-19 event but also the bull market in 2021, observed via the time-varying correlation between various cryptocurrencies derived from a graph-based approach.

3. Dataset

3.1. Dataset description

For this study, we obtained tick-by-tick data of cryptocurrencies from the HitBTC exchange.[5] The dataset covers the period from 13/02/2019 until 06/04/2021. This timeframe encompasses two major events, which are the market crash triggered by the Covid-19 pandemic and the subsequent bull market starting in October 2020.

To ensure the availability of data and minimize the percentage of missing values, we narrowed down our analysis to a final list of 34 cryptocurrencies that were most frequently traded. These cryptocurrencies are shown in Table 1. Regarding granularity, our previous study indicated that high-frequency data is more appropriate for analysis tasks since it contains sufficient information and can capture unexpected movements, especially with highly volatile time series like cryptocurrencies [13]. Eventually, we chose a granularity of 30 min as it is the highest frequency we can use while keeping the percentage of missing values low.

For each time series, we use its average value to replace missing data points. We acknowledge that there are other interpolation methods (e.g. nearest, linear, splines) that are worth considering. However, the average method is chosen for the sake of the mathematics involved. In particular, since our first pre-processing step is normalization [42], those average values will become zero as a result of the normalization formula. Thus, they will be eliminated in further calculation steps. In other words, the average method discriminates missing values in the time series and these values will not be used to do experiments. Consequently, our experimental results are derived solely from the available data. On the contrary, if other interpolation methods were used, they would result in non-zero interpolated values, which would be used in further calculation steps, potentially affecting the veracity of experimental results.

3.1.1. Ranking survey

We use the ranking[6] of each cryptocurrency to support our experiments. To do this, we collect monthly rankings from the Historical Snapshots page on Coinmarketcap from February 2019 to April 2021 and average them to obtain an overall ranking for each coin. This practice enables us to have a general assessment in terms of the size of a cryptocurrency during the considered period. A major coin tends to catch more attention of the general public and is utilized more often in cryptocurrency-related activities. This feature is displayed in Table 2

4. Methodology

4.1. Returns and volatility calculation

Given a price time series $x_i = \left(x_1^i, x_2^i, x_3^i, \ldots, x_T^i\right)$ of cryptocurrency i with a length of T, the log-return time series r_i is obtained using the formula $r_t^i = log(x_t^i/x_{t-1}^i)$, where r_t^i represents the return value at time t.

The volatility v_t at time t is calculated using the window of the last 48 data points from the return time series. The choice of this window size is explained by two reasons: Firstly, we use a 30-minute dataset, which results in 48 data points within a day (24 h). In

---

[5] https://hitbtc.com/.

[6] The ranking of a cryptocurrency is determined by its market capitalization (current price × number of tokens in circulation). The higher the market capitalization, the higher the corresponding rank assigned to the cryptocurrency.

Table 2
Average rankings of cryptocurrencies during the period between February 2019 and April 2021.

BTC	ETH	XRP	BCH	LTC	ADA	BSV	EOS	XLM
1	2	4	6	7	10	11	11	12
TRX	XMR	XTZ	NEO	XEM	ETC	MKR	DOGE	ZEC
15	15	18	22	25	28	30	32	32
ONT	BAT	OMG	QTUM	ZRX	BTG	ICX	LSK	REP
37	40	51	51	51	59	60	64	66
ZIL	IOST	XVG	MANA	MAID	BNT	FUN		
70	90	91	110	117	144	202		

other words, we capture the intraday movements of a cryptocurrency by employing a sliding window of size 48 with a sliding step of 30 min; Secondly, we tested the market efficiency[7] at each time window of 48 data points for each cryptocurrency, using two metrics: *permutation entropy*[8] and *statistical complexity*.[9] This testing procedure is proposed by [45]. Specifically, given an empirical time series in a corresponding time window, they first calculated permutation entropy and statistical complexity for the time series, then estimated the 95% random confidence interval for each measure by shuffling the time series randomly and calculating permutation entropy and statistical complexity for 30 independent realizations. The idea is that if two measures (e.g. permutation entropy and statistical complexity) derived from the empirical time series are within the 95% confidence intervals of randomly shuffled time series, then the empirical time series share common characteristics with random time series and thus no dependence patterns exist in the empirical time series since the values are just randomly distributed. In other words, the future value is independent of the previous values. Thus, the market efficiency presents in this time series. We found that the majority of cryptocurrencies in our study retain their efficiency for roughly 70 percent of the observed period. This percentage is relatively high for digital currencies, which suggests the equivalence between a coin's market value and the relevant news of that coin on the internet at that time, according to [45]. Therefore, the window size of 48 data points can capture and explain the movements of a cryptocurrency. The results of this experiment are shown in Table 3. We calculate the volatility in 2 ways to ensure the stability and transparency of experimental results.

i. Using the standard deviation of log-returns:

$$v_t^i = \sqrt{\frac{\sum_{k=0}^{47}(r_{t-k}^i - \mu)^2}{48}}$$

where μ is the mean of these log-returns.

ii. Using the moving average of squared log-returns, as suggested by Fernando et al. [46]:

$$v_t^i = \sqrt{\frac{\sum_{k=0}^{47} r_{t-k}^2}{48}}$$

4.2. Correlation matrix formation

Given a set of return or volatility time series z_i, $i = 1, 2, \ldots, 34$, corresponding to 34 cryptocurrencies mentioned in Section 3. We first calculate the similarity between each pair of cryptocurrencies C_{ij} as follows: $C_{ij} = \langle \hat{z}_i, \hat{z}_j \rangle$, with $\hat{z}_i = (z_i - \mu_i)/\sigma_i$, where i and j are cryptocurrencies, $\langle \cdot, \cdot \rangle$ represents the dot product, $\hat{z}_i$ is the normalized time series of z_i, μ_i is the mean value of z_i and σ_i is the standard deviation of z_i. Such a similarity measure is called *Pearson* correlation coefficient [15].

We acknowledge that this correlation calculation method has its limitations, such as being sensitive to outliers and not being able to capture non-linear relationships [47]. However, we have chosen to use it in our study for the following reasons:

i. We use returns and volatility to maintain the statistical nature of the time series. Other correlation metrics that address non-linear problem such as *Spearman* and *Kendall* show their own drawbacks. For instance, converting rational numbers

---

[7] The efficient market hypothesis states that the current prices of an asset reflect all available public market information relevant to that asset, making them fairly valued as they are present. Thus, no level of analysis or market timing strategy will yield opportunities for gaining excess returns. In other words, past price movements are not useful for predicting future prices, i.e. the future value does not depend on the historical values.

[8] Permutation entropy is an ordinal-based non-parametric complexity measure for studying the temporal dependence structure in a linear or non-linear time series. A low permutation entropy (close to 0) means that a future value can be predictable from the historical values. Whereas, a high permutation entropy ($\gg 0$) indicates that all values in the time series are independent and uniformly distributed. This measure is calculated based on the frequency of different ordinal patterns (permutations) in the data. Details of this complexity measure are described in [43].

[9] Statistical Complexity quantifies the structural complexity of dependence patterns in a time series. A higher statistical complexity means that the dependence patterns occur in a more complex fashion. In other words, the time series is more fluctuated and the future value is complicatedly relevant to historical values. This measure combines permutation entropy with a measure of disequilibrium, as described in [44].

Table 3

Testing the market efficiency for each cryptocurrency using 48 data points.

Crypto	Percentage		
ADA	70,00%	MANA	54,32%
BAT	55,00%	MKR	58,09%
BCH	70,37%	NEO	70,10%
BNT	63,80%	OMG	69,17%
BSV	70,32%	ONT	69,34%
BTC	69,37%	QTUM	69,77%
BTG	71,17%	REP	68,38%
DOGE	70,00%	TRX	70,21%
EOS	69,53%	XEM	70,80%
ETC	68,72%	XLM	70,63%
ETH	70,08%	XMR	69,34%
FUN	43,44%	XRP	70,89%
ICX	70,90%	XTZ	69,81%
IOST	67,61%	XVG	70,08%
LSK	70,08%	ZEC	71,43%
LTC	70,70%	ZIL	69,64%
MAID	69,81%	ZRX	69,47%

into integer rankings can result in the loss of important information from financial time series [48]. In addition, it has been demonstrated that rank correlation metrics (e.g. *Spearman* and *Kendall*) can also experience issues with non-linearity in certain situations [47].

ii. *Pearson* correlation is a widely accepted method in the literature for both cryptocurrency [41,49,50] and traditional asset markets [51–53]. This strongly reinforces our belief in the applicability of this method of correlation calculation for our problem.

iii. *Pearson* correlation is suitable for time series with repeated observations, as is the case in financial time series, unlike other methods that require independent observations [48].

The correlation matrix C is formed from similarity coefficients obtained above such that

$$C = \left(C_{ij}\right)_{1 \leq i \leq 34, 1 \leq j \leq 34}$$

A concern with this type of matrix is the reliability of the correlations, specifically whether the matrix accurately reflects true relationships between the time series being considered. To address this concern, the Random Matrix Theory provides a means of examination [13,35,54].

We have used RMT to test our correlation matrices and found that they all contain valuable information and are not random.

4.3. Denoising and detrending

4.3.1. Denoising

The noise effect was shown to appear in the stock market several decades ago [55]. This phenomenon is even more apparent in the cryptocurrency market. According to Thomas and Franziska [56], the average daily signal-to-noise ratio for the cryptocurrency market is only 36%, which is significantly lower than the 90% average daily signal-to-noise ratio for established US stock exchanges like NYSE and NASDAQ during the period from March 2017 to November 2017.

There are several factors contributing to the noise in the cryptocurrency market: Firstly, this market is vulnerable to *Pump and Dump* schemes, which have become a typical characteristic that distinguishes digital assets from traditional ones [57]; Secondly, although the purpose for inventing cryptocurrencies was to reduce the complexity of regulatory and trading operation that exists in traditional markets [58], a large number of transactions have been found to be linked to illegal purposes [58]. Eventually, more and more regulations have been enacted to control this young market [59]; Thirdly, transaction-related actions, in that investors can split their budget for one transaction into multiple transactions with smaller net worth to reduce transaction fees, which may result in unexpected price movements [60]; Additionally, other causes such as noise traders and arbitrageurs have been found to deliberately manipulate the movement of assets [55].

We use Random Matrix Theory [54] to distinguish the noise in cryptocurrencies. This theory states that the random correlation matrix C_R obtained from N randomly generated time series with zero mean, unit variance and length T (matching the size of the empirical correlation matrix C described earlier) has eigenvalues γ that asymptotically converges ($N \to +\infty$ and $T \to +\infty$ with $\frac{T}{N} > 1$) to the Marcenko–Pastur probability density function f [61]

$$f(\gamma) = \begin{cases} \frac{T}{N} \frac{\sqrt{(\lambda_+ - \gamma)(\gamma - \lambda_-)}}{2\pi\gamma} & \text{if } \gamma \in [\lambda_-, \lambda_+] \\ 0 & \text{if } \gamma \notin [\lambda_-, \lambda_+] \end{cases} \tag{1}$$
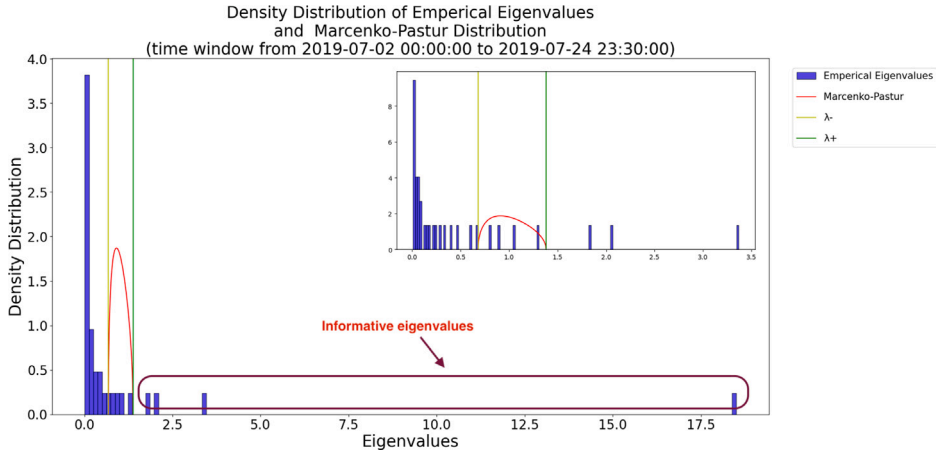
Fig. 1. Density distribution of eigenvalues (blue) obtained from our empirical correlation matrix (from 2019-07-02 00:00:00 to 2019-07-24 23:30:00) and the corresponding Marcenko–Pastur distribution (red). The yellow vertical line represents the minimum expected eigenvalue λ_- while the green vertical line represents the maximum expected eigenvalue λ_+ of the Marcenko–Pastur distribution. The top right sub-figure is obtained by removing the largest eigenvalue.

where $\lambda_- = (1 - \sqrt{\frac{N}{T}})^2$ is the minimum expected eigenvalue and $\lambda_+ = (1 + \sqrt{\frac{N}{T}})^2$ is the maximum expected eigenvalue of the Marcenko–Pastur distribution f.

Eventually, given eigenvalues λ obtained from the empirical correlation matrix C, $\lambda \in [\lambda_-, \lambda_+]$ are considered to exhibit random behavior as they belong to the random regime. Conversely, eigenvalues that fall outside of $[\lambda_-, \lambda_+]$ are expected to be informative signals. In practice, we only use informative eigenvalues that are greater than λ_+ since those that are less than λ_- carry negligible information [21]. Thus, we associate eigenvalues $\lambda \in [0, \lambda_+]$ with noise. Our objective in the denoising task is to eliminate these noisy eigenvalues from the empirical matrix C while keeping its trace unchanged. Consequently, the correlation between cryptocurrencies is calculated without the noise effect.

Fig. 1 illustrates an example of the density distribution of eigenvalues (shown in blue) obtained from our empirical correlation matrix (from 2019-07-02 00:00:00 to 2019-07-24 23:30:00) and the corresponding Marcenko–Pastur distribution (shown in red). In this case, the majority of eigenvalues fall below λ_+ (green vertical line), which are considered noise. There are a few eigenvalues that are greater than λ_+, which carry important information, however. We would like to keep these informative eigenvalues while processing the noisy ones to eliminate their impact from the correlation matrix. The sub-figure in the top right corner is obtained by removing the largest eigenvalue in order to help readers compare the Marcenko–Pastur distribution with empirical eigenvalues more easily.

For this, we employ the Eigenvector Clipping method [22]. One advantage of this approach is that it does not require any training parameters, making its outcome robust and more reliable. On the other hand, other denoising methods such as Linear shrinkage [62], Non-linear shrinkage [63] and Rotationally invariant, optimal shrinkage [64] require users to specify several parameters, leading to a quandary: how do we choose parameters? Furthermore, Eigenvector Clipping is straightforward to implement, with guaranteed efficiency as it keeps the information part, i.e., after the cleaning process, the trace of the correlation matrix remains unchanged [65]. This method has shown good performance in different studies and has been applied widely to different topics such as education, portfolio optimization and signal processing [20,66–68].

Given eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_n$ and corresponding eigenvectors $v_1, v_2, \ldots, v_n$ of our empirical correlation matrix $\mathbf{C}$, we can identify $k \leq n$ such that $\lambda_k > \lambda_+$ and $\lambda_{k+1} \leq \lambda_+$. As a result, $\lambda_i, \forall i \leq k$ are noisy eigenvalues, according to the Random Matrix Theory. The Eigenvector Clipping replaces these noisy eigenvectors with their average value in which the trace of the correlation matrix $\mathbf{C}$ after denoising is similar to its origin. The denoised correlation matrix $\mathbf{C}_{denoised}$ is defined by [69]:

$$\mathbf{C}_{denoised} = \sum_{i=1}^{n} \lambda_i^* v_i v_i^{\mathsf{T}}, \lambda_i^* = \begin{cases} \frac{\lambda_{k+1} + \lambda_{k+2} + \cdots + \lambda_n}{n-k}, \forall i \geq k+1 \\ \lambda_i, \forall i \leq k \end{cases} \tag{2}$$

4.3.2. Detrending

As of 2023, there are nearly 23,000 cryptocurrencies in circulation in the market, as reported by one of the biggest websites for tracking digital coins Coinmarketcap. By contrast, only around 10 blockchain protocols are used to govern cryptocurrencies [70]. As a result, the same blockchain protocol is used by many cryptocurrencies [71]. Moreover, the cryptocurrency market is heavily manipulated by social media and breaking news, leading to 'herding behavior' among investors [72]. These factors contribute to the co-movement in prices across different cryptocurrencies, which is referred to as the *trend* effect in the market.

Since the underlying characteristics of each cryptocurrency are masked by the existence of trend, we should eliminate this effect from the cryptocurrencies in order to observe each individual cryptocurrency's movements. In other words, we examine the
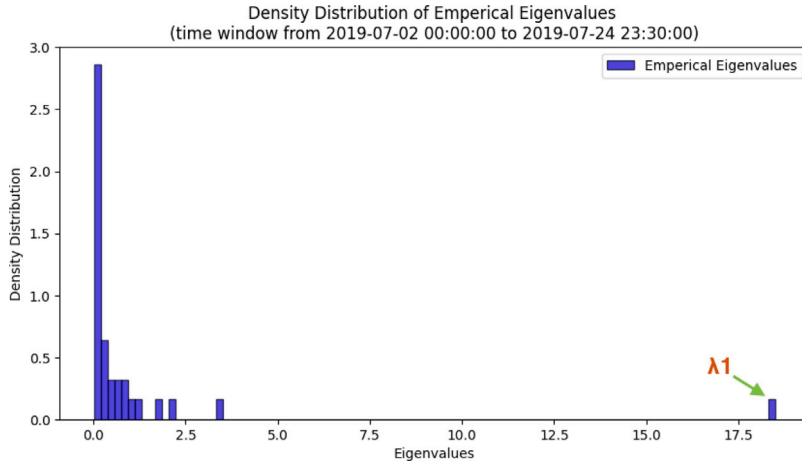
Density Distribution of Empirical Eigenvalues
(time window from 2019-07-02 00:00:00 to 2019-07-24 23:30:00)

Fig. 2. Market component (λ_1) in an empirical correlation matrix (from 2019-07-02 00:00:00 to 2019-07-24 23:30:00).

interaction among different cryptocurrencies without considering external factors such as shared protocols and mass media-driven price manipulation.

We adopt the concept of "market component" which is described in [21] to remove the trend effect. Such a component refers to the first eigenvalue (i.e. the largest eigenvalue) and corresponding eigenvector of the correlation matrix. It represents the common systematic risk that impacts most assets in the market and provides insights into the overall relationship between individual assets and the market as a whole. By removing this component, the correlation matrix can highlight the unique relationships between individual assets, providing insights into their specific dependencies and individual movements. More intuitively, according to [21], it is similar to removing a loud tone that prevents us from hearing other sounds. We call $\mathbf{C}_{filtered}$ the denoised and detrended correlation, which is obtained by subtracting the first eigenvalue and eigenvector from the denoised matrix $\mathbf{C}_{denoised}$.

$$\mathbf{C}_{filtered} = \mathbf{C}_{denoised} - \lambda_1 v_1 v_1^\mathsf{T} \tag{3}$$

Fig. 2 shows an example of the market component λ_1 in our empirical correlation matrix (from 2019-07-02 00:00:00 to 2019-07-24 23:30:00). As can be seen, there is a blue bar that stands significantly apart from the other blue bars, which represents the market component. For obvious reasons, the market component is always the largest eigenvalue of a correlation matrix.

Fig. 3 displays examples of the empirical correlation matrix (a) before denoising and detrending, (b) after denoising and (c) after denoising and detrending, respectively. The intensity of the red color (from white to dark red) represents the correlation between two cryptocurrencies, with darker red indicating higher correlation. The maximum value of the correlation is 1 which represents a complete similarity between two cryptocurrencies. By contrast, the minimum value of -1 represents a complete difference between two cryptocurrencies. Obviously, the correlation values in the diagonal line of each correlation matrix are all 1 because these values are derived from calculating the correlation between 2 identical cryptocurrencies. Fig. 3(a) shows that most cryptocurrencies are strongly correlated with each other since the dark red color is distributed across the matrix. This means that the majority of cryptocurrencies share a common trend, i.e. the price/volatility movements of different cryptocurrencies are similar. Additionally, we can infer that the noise actually exists in the cryptocurrencies since the denoised correlation matrix in Fig. 3(b) shows a brighter red color compared to the original correlation matrix. That is, the existence of noise results in the correlation between cryptocurrencies being higher than it should be. Regarding Fig. 3(c), this correlation matrix looks completely different from the first two cases since the low correlation is observed at a great number of cryptocurrency pairs. It is clear that the correlation between cryptocurrencies weakens significantly as a result of removing the trend effect. In general, the removal of noise and trend opens an opportunity to observe and explore underlying interactions between different cryptocurrencies. Hence, we can obtain novel findings that are hidden behind the noise and trend.

4.4. Cryptocurrency network construction and associated metrics

4.4.1. Network construction

Despite the usefulness of the correlation coefficient in understanding the relationships between cryptocurrencies, it fails to incorporate topological characteristics as it does not place connections in a metric space [73]. To overcome this limitation, the concept of *Distance Matrix* has been introduced as a substitute for the correlation matrix. Such a matrix, denoted as $\mathbf{D}$, is derived from the cleaned correlation matrix $\mathbf{C}$ such that each element d_{ij} of $\mathbf{D}$ is expressed as follows:

$$d_{ij} = \sqrt{2 \times \left(1 - c_{ij}\right)} \tag{4}$$

(a) Original correlation matrix

(b) Denoised correlation matrix

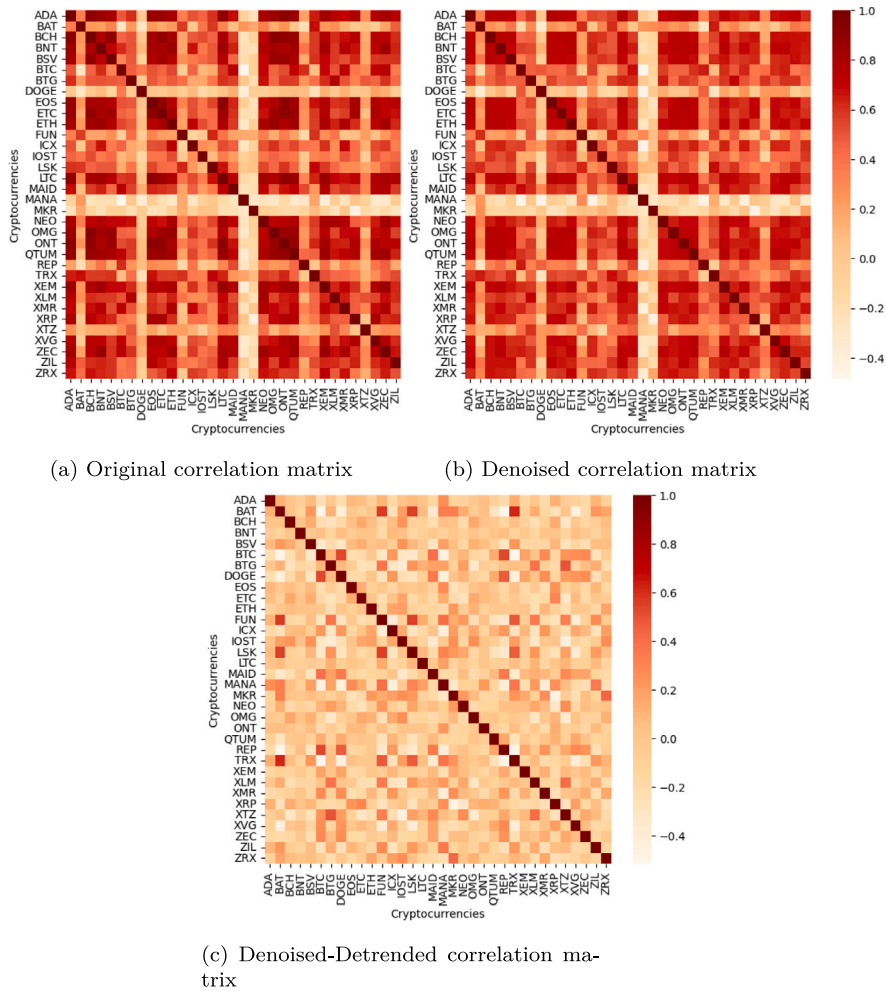(c) Denoised-Detrended correlation matrix

Fig. 3. Correlation matrix of 34 cryptocurrencies between 2019-07-02 00:00:00 and 2019-07-24 23:30:00. (a) Correlation matrix before denoising and detrending, (b) Correlation matrix after denoising, (c) Correlation matrix after denoising and detrending.

where d_{ij} ranges from zero to two with zero indicating complete dissimilarity while two represents complete similarity between 2 objects. By utilizing this approach, the resulting distance matrix satisfies the essential properties of a metric [73]: $d_{ij} \geq 0$, $d_{ij} = 0$ if $i = j$, and $d_{ij} = d_{ji}$. Consequently, this distance matrix enables the construction of a network (graph) of cryptocurrencies, where the proximity of nodes reflects their similarity and nodes with distinct behaviors are positioned further apart. The distance values between pairs of cryptocurrencies serve as the links (edges) within the network.

Nevertheless, this approach tends to create a dense network, i.e. each vertex is connected to every other vertices, thereby increasing complexity. To address this issue and focus on the most significant connections, the Minimum Spanning Tree (MST) [74] technique is employed. The MST is a specialized tree structure that connects all vertices with minimal total edge length. By selectively retaining only the $N-1$ most crucial edges, where N represents the number of nodes (cryptocurrencies), the MST effectively captures the underlying dynamics of network structures of a system [74]. This particular type of graph finds extensive applications across various domains [35,75,76], especially in studies of financial markets [14,74,77].

To derive the MST, the Kruskal algorithm is chosen due to its favorable performance and time complexity [78], which render it particularly suitable for relatively small networks like the one encompassing 34 cryptocurrencies in this study [79]. Furthermore, the Kruskal algorithm has demonstrated its applicability in numerous finance-related domains, thereby reinforcing its reliability and relevance [80,81].

4.4.2. Metrics for observing the evolution of a network

To track and evaluate the changes in cryptocurrency network structures over time, we use different graph-based metrics that can be categorized into 2 groups:

i. Category 1: Evaluation metrics for the entire network. These metrics allow us to observe the interactions among different cryptocurrencies over time and detect the changes in their interactions during different market conditions.

- Average Degree Assortativity: *Degree assortativity* measures the correlation between node degrees in a network. An assortative network has high-degree nodes connecting to other high-degree nodes, while a disassortative network has high-degree nodes connecting to low-degree nodes. The assortativity coefficient ranges from −1 (disassortative) to 1 (assortative). A smaller coefficient indicates a well-connected network, while a larger coefficient suggests a more scattered network [82].
- Average Betweenness Centrality: *Betweenness centrality* measures a node's importance in connecting others. It counts how many shortest paths pass through a node. A node with high betweenness centrality means that it lies on many shortest paths, indicating influence in information or resource flow. A high average betweenness centrality implies a scattered network with small communities, while a low average suggests a well-connected network forming a large community [7].
- Average Closeness Centrality: This measures the average distance from a node to all other nodes in the graph. A node with a high closeness centrality is one that is close to many other nodes. Thus, it has a higher influence on the network. This metric is used to identify the central node (the node that directly connects to many other nodes) in the network. For the first category, we average the closeness centrality of all nodes. Eventually, a bigger value suggests a well-connected network, while a smaller value suggests a more scattered network [25].

ii. Category 2: Evaluation metrics for each entity, which is used to observe the influence of each cryptocurrency on the whole network over time to identify which cryptocurrencies have the strongest influence on the market and gain an understanding of how cryptocurrencies function within the network.

- PageRank: An algorithm used to measure the importance of nodes in a graph based on the number of links pointing to it. This is the same algorithm used by Google to rank web pages in search results [83].

5. Experimental results and discussion

In this section, we conduct an analysis of network structures derived from two primary types of financial time series, e.g. returns and volatility, through different graph-based metrics, e.g. *degree assortativity*, *betweenness centrality*, *closeness centrality* and *PageRank*. This experiment has two main objectives: Firstly, we aim to gain insights into underlying traits of interactions among cryptocurrencies; Secondly, we seek to identify distinctive features between a returns-based crypto network and one based on volatility. We then relate these analyzed findings to different external factors such as investors' sentiment and transaction patterns.

5.1. The choice of optimum window size for network construction

We have presented the process of constructing a network of cryptocurrencies in Section 4. However, it is necessary to determine the optimum window size (length) of time series for this construction. In this regard, we propose two criteria for choosing a suitable window size for network construction: Firstly, the chosen window size should minimize the change in network metrics built on different window sizes. In other words, we choose a window size such that the characteristics of the networks are constant for small changes of window sizes; Secondly, it is also expected to minimize the discrepancy in network metrics (e.g. *degree assortativity*, *betweenness centrality* and *closeness centrality*) obtained by two different volatility measures.

For this experiment, we have selected a range of potential window sizes, consisting of 28 cases spanning between 48 and 1344 data points, which correspond to 1 day to 28 days (4 weeks). The rationale behind this selection is that the cryptocurrency market varies very quickly, thus a longer duration might fail to capture significant events happening to the market [84]. We use the linear correlation metric *Pearson* and the non-linear correlation metric *Spearman* to calculate the similarity between network structures in different cases via *betweenness centrality*, *degree assortativity* and *closeness centrality*.

From our experiments, we gained valuable findings that assisted us in choosing a favorable network window size. In particular, by testing the first criterion, we found that the characteristics of the network structures do not change greatly when the window size is large. This suggests that the network structures can remain relatively stable if we use relatively large window sizes. Remarkably, we observed a plateau point starting at a window size of 816 (corresponding to 17 days) and the similarity slowly increases from here until the last case (corresponding to 28 days). Thus, based on the first criterion, we exclude window sizes that are less than 816. An example of this experimental result is shown in Fig. 4, where we use *Pearson* and *Spearman* to calculate the similarity between *betweenness centrality* values of different network structures constructed using different window sizes.

Regarding the second criterion, we also found a rather similar result with the one obtained from the first criterion. Specifically, the network structures with the window size of 816 or more show negligible changes, indicating a high similarity between network structures with large window sizes. The only difference is that *Pearson* and *Spearman* values drop slightly after the window size of 1104 (corresponding to 23 days), before resuming a steady increase from the window size of 1248 (corresponding to 26 days). An example of this experiment is shown in Fig. 5, where we use *betweenness centrality* as a feature to measure the similarity between two network structures derived from two different volatility formulas.

In summary, we chose the window size of 1104 (23 days) to construct a network. This window size guarantees the consistency in results irrespective of the volatility formulas we use, thus helping to minimize the change in a network structure when using a
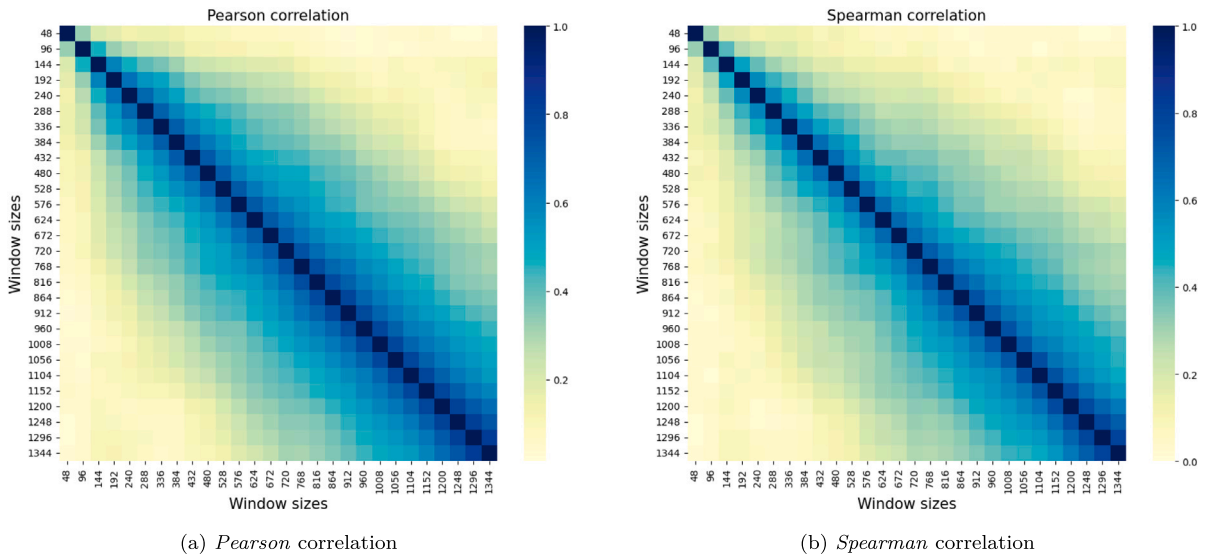
(a) *Pearson* correlation (b) *Spearman* correlation

Fig. 4. The similarity of each pair of time-varying network structures using different window sizes, observed by *betweenness centrality*. (a) Measured by *Pearson* algorithm, (b) Measured by *Spearman* algorithm.
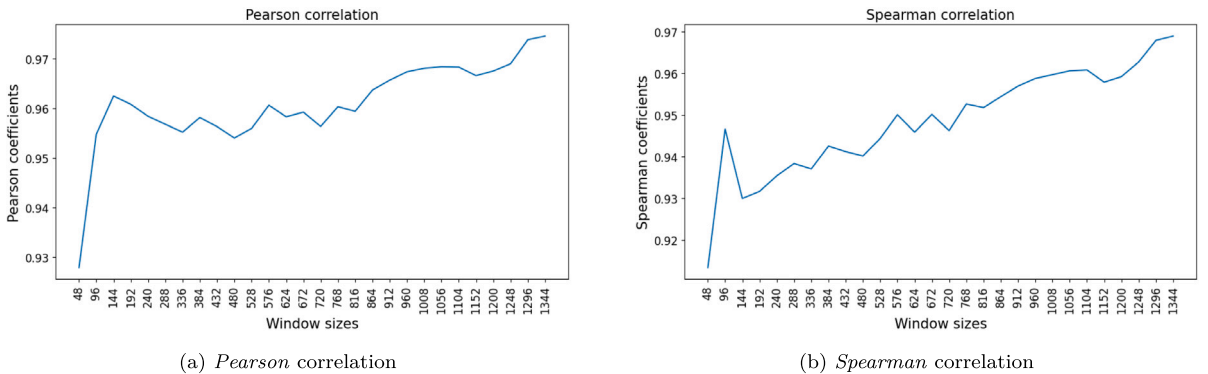

(a) *Pearson* correlation (b) *Spearman* correlation

Fig. 5. The similarity between two time-varying network structures derived from 2 corresponding volatility formulas, observed by *betweenness centrality*. (a) Measured by *Pearson* algorithm, (b) Measured by *Spearman* algorithm.

different window size and especially, large enough to satisfy the requirement of Random Matrix Theory (as described in Section 4.2) and minimize the time and space complexity.

Additionally, we note that two volatility formulas show negligible differences in terms of graph-based experimental results when using this window size. Thus, without loss of generality, we only use the more commonly used standard deviation-based volatility to conduct the remaining experiments.

5.2. Analysis of cross-interactions in cryptocurrencies through time-varying networks

5.2.1. Returns-based time-varying networks

We first observe the evolution of network structures obtained by return time series, which is the most frequently used financial feature in the existing literature. We implement this experiment in two cases, before and after the noise and trend removal, respectively.

Figs. 6(a) and 6(b) depict the changes in network structure every 30 min, expressed by three different metrics before and after noise and trend removal, respectively. The red line indicates the original result, the blue line indicates the centered moving average over a 14-day period, the green line indicates the average value and the black line indicates the border of each period.

As expected, these graph-based metrics show significant fluctuations over the considered period, both before and after denoising and detrending. This means that the interactions between different cryptocurrencies change over time, regardless of whether noise and trend are present or not. However, it can be seen that the changes in the network structure with the effect of noise and trend present seem to be rather random. By contrast, we see noticeable patterns in the cleaned network structure. Specifically, the values
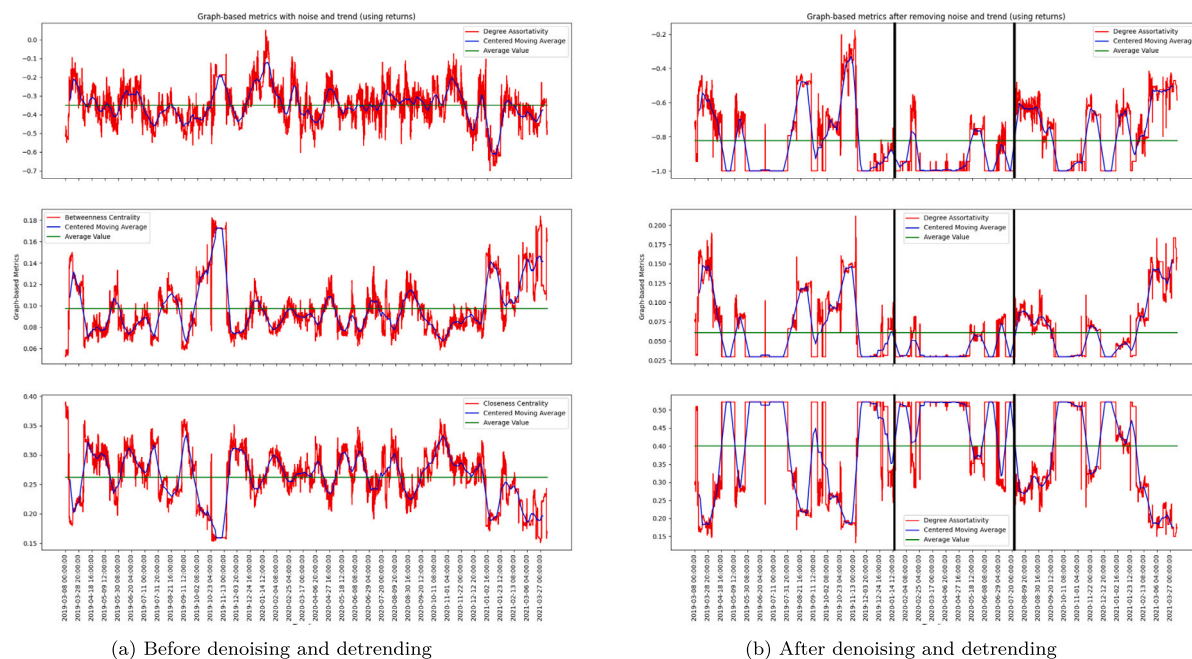
(a) Before denoising and detrending   (b) After denoising and detrending

Fig. 6. Evolution of the returns-based network structure, measured by *degree assortativity*, *betweenness centrality* and *closeness centrality*, respectively. (a) Before denoising and detrending, (b) After denoising and detrending. The red line indicates the original result, the blue line indicates the centered moving average over the 14-day period, the green line indicates the mean of the time series, and the black line indicates the boundary of each period.

Table 4
The percentage of 3 SAX segments in each period. "a" stands for the first segment, corresponding to the first 50 percentiles of values of graph-based metrics, "b" stands for the second segment, ranging from the 51st percentile to the 80th percentile, "c" stands for the third segment, covering values that are greater than the 80th percentile.

	SAX type	Betweenness	Degree assortativity	Closeness
08/03/2019–15/01/2020	a	47.6	47.7	47.6
	b	22.3	30.8	22.4
	c	30.1	21.5	30.0
16/01/2020–23/07/2020	a	75.3	76.8	75.3
	b	24.2	21.2	24.4
	c	0.50	2.00	0.30
24/07/2020–06/04/2021	a	34.3	33.4	33.9
	b	43.8	35.1	43.9
	c	21.9	31.5	22.2

of graph-based metrics are relatively small during the period between January and July 2020, which means that the network tends to be compressed during this time, i.e. nodes are close to each other and form one large homogeneous group. On the other hand, large values are dominant during the first period from February 2019 until the end of the same year, implying a scattered network where nodes are more distant from each other, forming multiple smaller groups. The last period, after July 2020, reveals a mixture of characteristics observed in the first two periods.

For the purpose of verifying this phenomenon, we use the Symbolic Aggregate Approximation (SAX) [85] to compare the differences between the three mentioned periods relative to each other. This technique is designed to find similar patterns of behavior within a single time series even when the exact magnitudes of values differ due to factors like trend and seasonality [86]. In other words, SAX can potentially find patterns in what classic time series decomposition would consider noise, as these patterns do not occur regularly and are not part of the trend or seasonal behavior.

The SAX representation for the cleaned graph-based metrics is shown in Table 4. To implement SAX, we begin by splitting the time series' values into different segments based on the structure of the network. In particular, our experiments have revealed that graph-based metric values within the first 50 percentiles represent a compressed network (i.e. nodes are close to each other and form a big group). Values surpassing the 80th percentile represents a scattered network (i.e. node are far from each other and spread out). On the other hand, values falling between the 51st and 80th percentile show a mixed network, displaying characteristics of both compression and scattering. Next, we partition the time series into three periods, including 08/03/2019–15/01/2020, 16/01/2020–23/07/2020 and 24/07/2020–06/04/2021. This division is based on the observation that the graph-based metric time series between
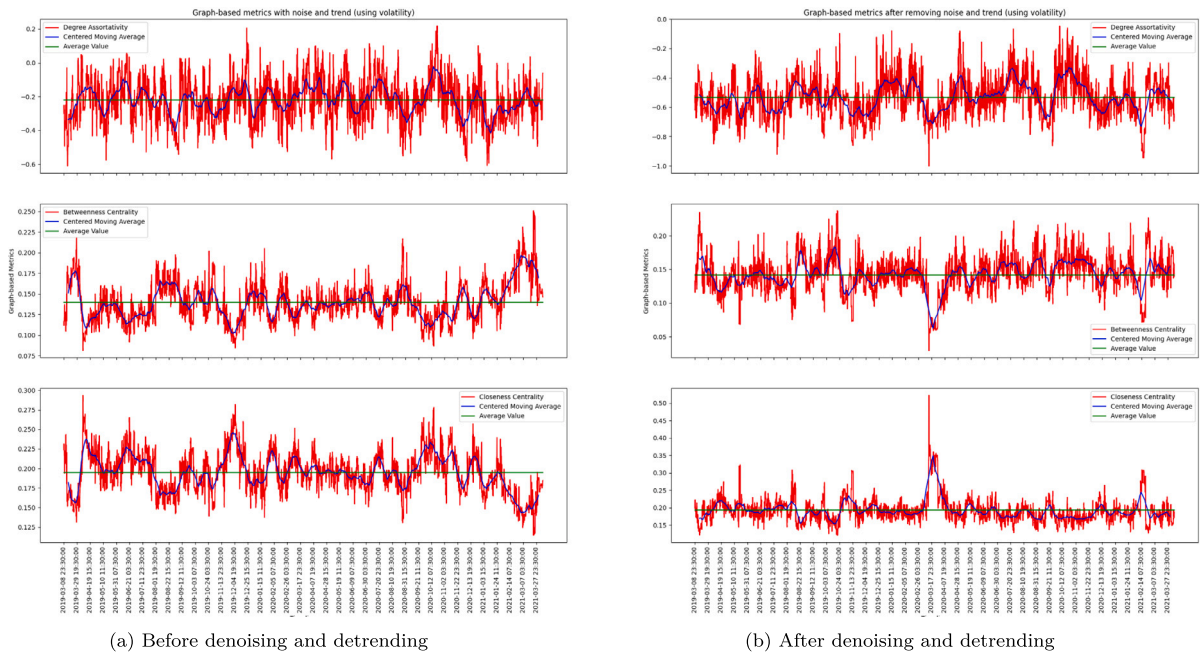
(a) Before denoising and detrending   (b) After denoising and detrending

Fig. 7. Evolution of the volatility-based network structure, measured by *degree assortativity*, *betweenness centrality* and *closeness centrality*. (a) Before denoising and detrending, (b) After denoising and detrending. The red line indicates the original result, the blue line indicates the centered moving average over the 14-day period, the green line indicates the mean of the time series.

16/01/2020 and 23/07/2020 indicates a distinct pattern compared to the other periods. Lastly, we calculate the percentage of occurrence for each segment in each of the three periods.

Table 4 shows the percentage of each segment (e.g. a,b and c) in each period. Notably, the second period clearly displays a different behavior compared to the others. Specifically, the cryptocurrency network structure tends to be compressed during the second period (16/01/2020–23/07/2020) with more than 75 percent of the time displaying an "a" SAX type and nearly 100 percent of the time showing "a" and "b" SAX types. Conversely, the first and third periods are similar to some extent. The notable difference is that there is an increase of "b" SAX type and a decrease of "a" SAX type between the first and third periods. This result is in line with the findings in our previous study [13] and also other existing studies with different datasets and methodologies [87,88], where significant changes in the network structure of cryptocurrencies during the Covid-19 pandemic from January to July 2020 were discovered. In the meanwhile, the post-pandemic period demonstrates a combination of network structure characteristics observed in both the pre-pandemic and peak pandemic seasons.

In summary, the return time series reveals a change between normal times and downturn times in terms of the correlation between cryptocurrencies. However, from the graph-based metrics, as shown in Fig. 6, no other significant patterns can be discerned by this financial feature.

5.2.2. Volatility-based time-varying networks

As mentioned earlier, volatility also plays an essential role in exploring and explaining the underlying characteristics of financial assets [4]. In this regard, we expect that the volatility information can add more insights into understanding the correlation among different cryptocurrencies. Thus, in this section, we analyze the time-varying network structures that are constructed by volatility time series.

Figs. 7(a) and 7(b) illustrate the results of *degree assortativity*, *betweenness centrality* and *closeness centrality* of the volatility-based cryptocurrency network over time, before and after denoising and detrending, respectively. Similar to the returns-based results, we observe changes in the correlation between cryptocurrencies over time in both pre-and post-denoise and detrend scenarios. Again, the graph-based metrics, when affected by noise and trend, still fail to reveal any meaningful patterns as they appear to fluctuate randomly. This is particularly evident when looking at the *degree assortativity* in Fig. 7(a). However, after denoising and detrending, we observe abnormal spikes in the results, as shown in Fig. 7(b). Notably, two common spikes are present across all three metrics, including March 2020 and February 2021. These spikes indicate a compressed network, where its nodes are close to each other and form a big group. In other words, our examined cryptocurrencies share the same movement in terms of volatility.

The first common spike was between 13/03/2020 and 27/03/2020, which was equivalent to the worst time in the first wave of the Covid-19 pandemic, in the sense that the global economy experienced a crash due to the regulations and health protection measures enacted by almost all countries across the globe to defend against the pandemic [89]. During this time, public attention towards the pandemic increased dramatically, which was revealed by the increased volume of searches for COVID-19-related terms,
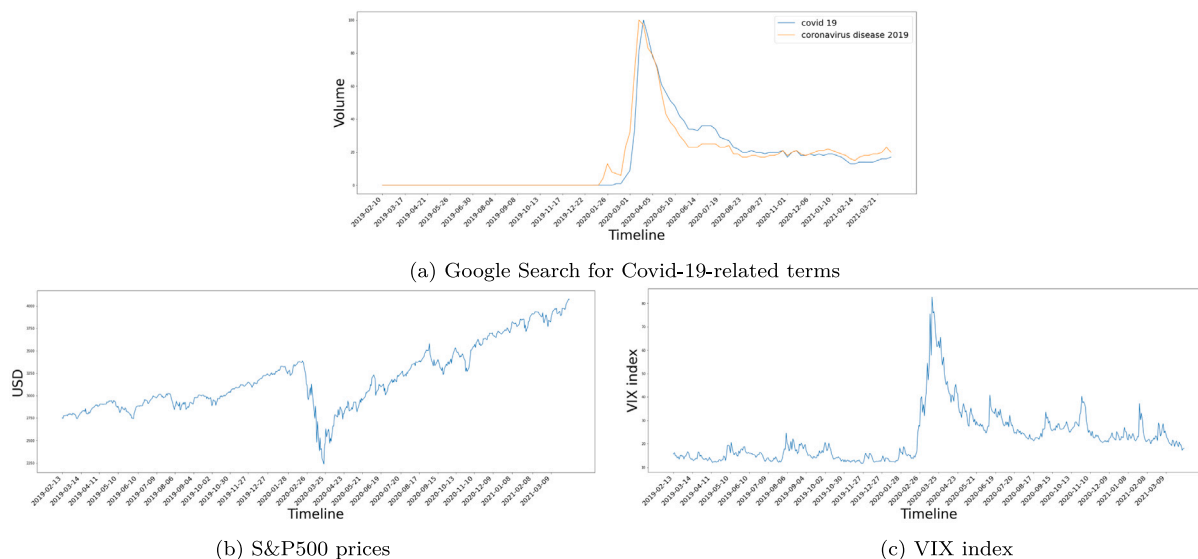
(a) Google Search for Covid-19-related terms

(b) S&P500 prices

(c) VIX index

Fig. 8. Public attention (a), stock volatility index (VIX) (b) and S&P500 index (c), from 13/02/2019 until 4/06/2021.

i.e. "covid 19" and "coronavirus disease 2019" (Fig. 8(a)). Regarding the stock market, one of the largest stock market indices ($S\&P$500) experienced one of the biggest downturns ever seen with a more than 30 percent loss (Fig. 8(b)), while the stock volatility index VIX increased by more than 500% in a matter of 2 months (Fig. 8(c)). The cryptocurrency market also reacted to this global crisis since the majority of digital coins started to lose their values in mid-February 2020 and reached the lowest values on 13/03/2020 as shown on Coinmarketcap, which was also the date graph-based metrics had the highest peak, i.e. the cryptocurrency network was most compressed or cryptocurrencies were closest to each other. The financial markets quickly started to recover thereafter while graph-based metrics also gradually went back to their pre-pandemic values.

The second spike occurred between 12/02/2021 and 20/02/2021, coinciding with a bullish period in the cryptocurrency market as shown on Coinmarketcap. This was the first time the cryptocurrency market experienced a significant surge in prices, not only with major coins such as BTC and ETH but also with minor coins like BAT and MKR. However, the prices of most cryptocurrencies started to decline again after 19/02/2021. Interestingly, it can be seen that the network metrics fell back to their values prior to the spike period after 20/02/2021, as shown in Fig. 7(b).

These spikes in graph-based metrics can be attributed to herding behavior [90], which is mainly driven by the collective trading decisions of naïve (irrational) traders [39,40]. In particular, the cryptocurrency market is shown to be ruled by such investors (irrational investors), they tend to mimic other investment decisions and are strongly influenced by the market sentiment, especially during market crashes and bullish periods [90]. That is, they tend to sell off their shares when the market experiences a downturn or bad news circulates on the internet in order to avoid the loss. By contrast, they are more likely to purchase new shares during bullish phases or when positive news about cryptocurrencies spreads among the general public due to the fear of missing out on increasing the cryptocurrency price. As a result, similar trading actions occur across different cryptocurrencies of all sizes, causing a consistent impact on the volatility of different coins, which in turn, leads to an increase in the correlation within the cryptocurrency market with respect to the volatility. This collective correlation has been discussed in a study by James and Menzies [4], where the authors calculated the proportional contribution of one cryptocurrency to the total volatility of 52 major cryptocurrencies throughout the period between April 2019 and June 2021. They found that the proportion of the market's total volatility is more evenly distributed among all the cryptocurrencies during the Covid-19 market crisis and the bull market of 2021. There was less deviation between different volatilities of individual cryptocurrencies in these two periods, i.e. everything was similarly volatile together.

To reinforce our interpretation of the aforementioned spikes in graph-based metrics, we examine the correlation between these two spikes and the Fear and Greed index[10] of investors in the cryptocurrency market. Without loss of generalization, we display the correlation between the *betweenness centrality* and the Fear and Greed index in Fig. 9. In accordance with our expectations, the spike in March 2020 corresponds to the time people are extremely pessimistic about the market due to the significant decrease in prices. Regarding the spike in February 2021, it corresponds to the time people are most optimistic about cryptocurrencies, boosted by the continuous growth of this market. This result clearly points out the relationship between the volatility network structure and people's sentiments. That is, the cryptocurrency network constructed by volatility time series is strongly compressed (i.e. the

---

[10]   Assessing the emotions and sentiments of people towards the cryptocurrency market. The index ranges from 0 to 100, which is classified into 5 levels in total, including extreme fearful (0–25), fearful (26–46), neutral (47–54), greedy (55–75) and extreme greedy (76–100), website: https://alternative.me/crypto/fear-and-greed-index/.
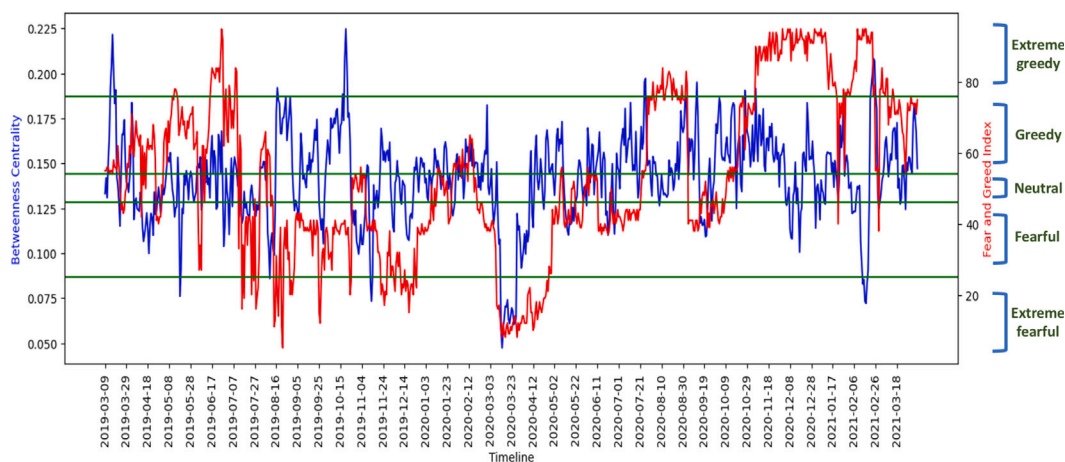
Fig. 9. The movement of *betweenness centrality* and *Fear and Greed* index from 09/03/2019 until 06/04/2021. The red line refers to the *betweenness centrality*, the blue line refers to the sentiment index, and each green line corresponds to a boundary for a type of sentiment, including extreme fearful, fearful, neutral, greedy and extreme greedy.

correlation between different cryptocurrencies increases, and the network forms a big group) when investors overreact to movements in the market, which is a consequence of a herding phenomenon, either buying (positive sentiment) or selling (negative sentiment) behavior.

We also notice that the magnitude of the first spike in March 2020 is much larger than that of the second spike in February 2021. This is not a trivial phenomenon but rather, it carries a significant implication regarding the influence of people's sentiment on the cryptocurrency market. Specifically, crypto investors, being naïve, tend to overreact and trade irrationally when confronted with negative news. They engage in herding activities by imitating others' actions in an attempt to avoid substantial losses. This causes an increase in volatility and influences a wide range of cryptocurrencies. On the other hand, the majority of investors demonstrate greater wisdom and make more rational investment decisions when the market sentiment is positive, thereby reducing herding behavior. These findings contradict several studies conducted before 2020 such as [91,92], which argued that positive returns exerted a stronger influence on volatility than negative returns. The rise in volatility can be interpreted as being due to the dominance of uninformed traders, this time reacting to positive news. However, our result is in line with a recent study utilizing a dataset covering 3 continuous years from 2019 until 2021 [93]. More importantly, the stronger influence of negative news compared to positive news on investors has been well-established in traditional markets and has stood as an enduring fact [94]. This implies that the cryptocurrency market is becoming progressively more mature and more similar to traditional ones. This is aligned with [14], where the authors showed that major and famous cryptocurrencies are becoming more and more similar to the stock market.

We acknowledge that there are other spikes in graph-based metrics (Fig. 7(b)). However, they are only consistent in one or two graph-based metrics and the duration for these spikes is much shorter, within one or two days. For instance, three spikes are in *betweenness centrality* and *closeness centrality* but not in *degree assortativity* including May, August and November 2019. We suspect that this might stem from illicit trading activities which we cannot track or collect information on, such as money laundering, terrorist financing and corruption [95]. In addition, another potential reason stems from social media-related activities such as market manipulation or pump-and-dump schemes, where a portion of crypto traders try to create herding to take advantage of the associated upward price movement. This possibly explains the reason why people changed their sentiment regularly during the year 2019, as shown in Fig. 9. This phenomenon will be the preserve of future work since it is outside the scope of this study.

5.3. Information flow in the cryptocurrency network: Which cryptocurrency acts as the information transmission center?

In this section, we aim to analyze the influence of each cryptocurrency on a network comprising 34 diverse cryptocurrencies. The influence of a cryptocurrency on this network is determined by the extent to which the other cryptocurrencies' values change when the value of that particular cryptocurrency fluctuates. Therefore, a cryptocurrency with a greater influence is associated with a larger number of interconnected cryptocurrencies. For this purpose, we use an algorithm launched by Google which is originally used to assess the importance of web pages based on their links, called *PageRank* [83]. It assigns a numerical score between 0 and 1 to each node, considering factors such as the number and quality of incoming links. Nodes with higher scores are considered more influential, and thus transfer their information to a wider range of coins in the network.

We calculate the *PageRank* value for each cryptocurrency at each timestamp based on the corresponding network structure, as observed via both returns and volatility. To assess the influence of each cryptocurrency in the network, we define two metrics inherited from *PageRank* that quantify the magnitude of influence for each cryptocurrency:

Table 5

Total Accumulation (TA) for each cryptocurrency with the effect of noise and trend, observed via volatility (a) and returns (b), respectively. Each table has 4 columns: Crypto represents cryptocurrency symbols; TA represents Total Accumulation values; Rank represents the ranking of cryptocurrencies; Percentage represents the ratio of one cryptocurrency's Total Accumulation value to the sum of all Total Accumulation values. The results are ordered from high to low.

(a) Volatility								(b) Returns							
Crypto	TA	Rank	Percentage	Crypto	TA	Rank	Percentage	Crypto	TA	Rank	Percentage	Crypto	TA	Rank	Percentage
ETH	1564	2	4,29	BAT	1008	40	2,76	BTC	3414	1	9,35	BAT	822	40	2,25
BCH	1387	6	3,8	DOGE	998	32	2,74	ETH	3090	2	8,46	MKR	797	30	2,18
BTC	1375	1	3,77	OMG	996	51	2,73	LTC	1692	7	4,63	ZIL	791	70	2,17
XLM	1361	12	3,73	XMR	991	15	2,72	NEO	1495	22	4,09	REP	785	66	2,15
NEO	1335	22	3,66	BSV	982	11	2,69	BCH	1473	6	4,03	XTZ	783	18	2,14
LTC	1324	7	3,63	ICX	976	60	2,68	EOS	1443	11	3,95	XVG	783	91	2,14
EOS	1304	11	3,58	BNT	971	144	2,66	ONT	1206	37	3,30	MANA	780	110	2,13
ADA	1265	10	3,47	XVG	965	91	2,64	ADA	1161	10	3,18	ZEC	774	32	2,12
ONT	1211	37	3,32	XEM	964	25	2,64	QTUM	1095	51	3,00	XMR	772	15	2,11
QTUM	1137	51	3,12	XTZ	952	18	2,61	TRX	1054	15	2,89	LSK	756	64	2,07
TRX	1070	15	2,93	MANA	925	110	2,53	XRP	1010	4	2,77	BSV	754	11	2,06
ZEC	1058	32	2,90	REP	908	66	2,49	XLM	985	12	2,70	MAID	754	117	2,06
ETC	1050	28	2,88	BTG	901	59	2,47	BNT	912	144	2,50	DOGE	753	32	2,06
ZRX	1044	51	2,86	MKR	894	30	2,45	ZRX	895	51	2,45	BTG	747	59	2,05
XRP	1037	4	2,84	FUN	860	202	2,36	ETC	879	28	2,41	XEM	733	25	2,01
LSK	1032	64	2,83	MAID	823	117	2,26	ICX	856	60	2,34	FUN	732	202	2,00
ZIL	1010	70	2,77	IOST	803	90	2,20	OMG	848	51	2,32	IOST	703	90	1,93

- *Total Accumulation* (TA): calculates the sum of *PageRank* values for each cryptocurrency across the period.

$$\textbf{Total Accumulation}_i = \sum_{k=1}^{T} pr_k^i \tag{5}$$

- *Dominance Score* (DS): counts the number of times that a cryptocurrency has the highest *PageRank* value.

$$\textbf{Dominance Score}_i = \sum_{k=1}^{T} hpr_k^i$$

With

$$hpr_k^i = \begin{cases} 1, & \text{if } pr_k^i \text{ is the highest value at time } k \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

where i represents the cryptocurrency i, T represents the length of the period and pr_k^i represents the *PageRank* value of the cryptocurrency i at time k.

Tables 5 to 8 display the results of this experiment before and after denoising and detrending. Our experiment demonstrates that returns and volatility yield similar outcomes. In particular, with the influence of noise and trend, the prominent cryptocurrencies such as BTC and ETH appear to have a greater influence on the market than the rest, as evidenced by their higher *PageRank* values throughout the observed period. This is a common result obtained by various studies whose experiments are based on original data with the presence of noise and trend [72,96,97]. By contrast, this characteristic disappears after removing the noise and trend contributions. Specifically, smaller and less famous cryptocurrencies tend to have a larger impact on the market. In contrast, BTC and ETH show negligible impact on the market. We note that this phenomenon remains consistent across different window sizes of the network construction.

The significant influence of minor cryptocurrencies on the market can be attributed to 3 key factors: the prevalence of Pump & Dump schemes, the transaction activity of crypto investors and the treatment of investors on minor cryptocurrencies (e.g. how minor cryptocurrencies are used by investors).

In relation to Pump & Dump, it has become a familiar characteristic of the cryptocurrency market, which is organized regularly by a wide range of investors. A study [98] observed 355 such events on 2 major exchanges, Binance and Yobit. The researchers collected data from 2 different sources, namely Telegram and PumpAnalysis.com during a 4-year period from 2018 to 2021, to analyze the characteristics of coins that are most likely to be targeted by manipulators for a Pump & Dump scheme. Their findings revealed that manipulators target relatively illiquid coins (low-ranking coins). This means that individuals will be more attracted to pumps of low-ranking coins. These schemes were found to impact not only the pumped cryptocurrency but also other cryptocurrencies. In particular, a study conducted by Balcilar and Ozdemir [99] investigated the risk spillover effect caused by the Pump & Dump among various cryptocurrencies, using the frequency connectedness approach. They found evidence of the risk spillover among different cryptocurrencies during pump events. Therefore, a Pump & Dump scheme often originates from a low-ranked cryptocurrency, resulting in synchronized fluctuations between that coin and other coins.

In order to investigate the connection between investor transaction activity and the significant influence of small cryptocurrencies, we calculate the correlation between our examined cryptocurrencies with respect to the number of transactions recorded every

Table 6
Total Accumulation (TA) for each cryptocurrency after removing noise and trend, observed via volatility (a) and returns (b), respectively. Each table has 4 columns: Crypto represents cryptocurrency symbols; TA represents Total Accumulation values; Rank represents the ranking of cryptocurrencies; Percentage represents the ratio of one cryptocurrency's Total Accumulation value to the sum of all Total Accumulation values. The values are ordered from high to low.

(a) Volatility								(b) Returns							
Crypto	TA	Rank	Percentage	Crypto	TA	Rank	Percentage	Crypto	TA	Rank	Percentage	Crypto	TA	Rank	Percentage
BAT	1818	40	4,98	BTC	981	1	2,69	MKR	3354	30	9,18	BTC	787	1	2,15
MKR	1627	30	4,46	BSV	975	11	2,67	FUN	3152	202	8,63	BCH	771	6	2,11
FUN	1460	202	4,00	OMG	952	51	2,61	MANA	2318	110	6,35	ZIL	768	70	2,10
MANA	1405	110	3,85	ADA	940	10	2,58	BAT	2130	40	5,83	QTUM	766	51	2,10
XTZ	1309	18	3,59	XEM	937	25	2,57	DOGE	1550	32	4,24	REP	740	66	2,03
ZIL	1288	70	3,53	ETC	931	28	2,55	ZRX	1309	51	3,58	ICX	717	60	1,96
IOST	1260	90	3,45	ZEC	919	32	2,52	BNT	1243	144	3,40	ADA	714	10	1,95
XVG	1206	91	3,30	NEO	907	22	2,49	OMG	1207	51	3,31	XRP	711	4	1,95
MAID	1190	117	3,26	ETH	904	2	2,48	MAID	1177	117	3,22	LTC	709	7	1,94
BNT	1182	144	3,24	LTC	903	7	2,47	XTZ	1054	18	2,89	ETH	707	2	1,93
BTG	1079	59	2,96	EOS	880	11	2,41	NEO	923	22	2,53	XEM	692	25	1,89
REP	1077	66	2,95	TRX	872	15	2,39	IOST	896	90	2,45	TRX	684	15	1,87
DOGE	1074	32	1,94	XMR	863	15	2,37	ONT	877	37	2,40	ETC	676	28	1,85
ZRX	1073	51	2,94	ONT	850	37	2,33	LSK	864	64	2,36	XLM	670	12	1,83
XLM	1069	12	2,93	XRP	837	4	2,30	BTG	856	59	2,34	BSV	653	11	1,79
LSK	1046	64	2,87	QTUM	813	51	2,23	XVG	798	91	2,19	XMR	641	15	1,75
ICX	1042	60	2,86	BCH	811	6	2,22	EOS	794	11	2,17	ZEC	619	32	1,70

Table 7
Dominance Score (DS) for each cryptocurrency with the effect of noise and trend, observed via volatility (a) and returns (b), respectively. Each table has 4 columns: Crypto represents cryptocurrency symbols; DS represents Dominance Score values; Rank represents the ranking of cryptocurrencies; Percentage represents the ratio of one cryptocurrency's Dominance Score value to the sum of all Dominance Score values. The values are ordered from high to low.

(a) Volatility								(b) Returns							
Crypto	TA	Rank	Percentage	Crypto	TA	Rank	Percentage	Crypto	TA	Rank	Percentage	Crypto	TA	Rank	Percentage
ETH	4244	2	11,63	ZRX	579	51	1,59	BTC	14204	1	39,89	XLM	5	12	0,01
LTC	3544	7	9,71	DOGE	511	32	1,40	ETH	11481	2	31,43	BNT	4	144	0,01
BCH	3347	6	9,17	MAID	490	117	1,34	LTC	3618	7	9,90	MAID	2	117	0,01
NEO	3094	22	8,48	ICX	475	60	1,30	NEO	1697	22	4,65	BSV	0	11	0,00
BTC	2913	1	7,98	LSK	442	64	1,21	EOS	1356	11	3,71	BTG	0	59	0,00
XLM	2563	12	7,03	BAT	319	40	0,87	QTUM	601	51	1,65	DOGE	0	32	0,00
EOS	1588	11	4,35	OMG	313	51	0,86	TRX	600	15	1,64	FUN	0	202	0,00
QTUM	1499	51	4,11	MANA	248	110	0,68	ADA	521	10	1,43	IOST	0	90	0,00
ONT	1372	37	3,76	XTZ	238	18	0,65	XRP	517	4	1,42	LSK	0	64	0,00
ZIL	1327	70	3,64	XMR	208	15	0,57	ONT	474	37	1,30	MKR	0	30	0,00
ADA	1326	10	3,63	BTG	190	59	0,57	ICX	434	60	1,19	OMG	0	51	0,00
ETC	977	28	2,68	BSV	190	11	0,52	ZRX	278	51	0,76	REP	0	66	0,00
TRX	963	15	2,64	BNT	177	144	0,49	ETC	259	28	0,71	XEM	0	25	0,00
XEM	947	25	2,60	REP	154	66	0,42	BCH	214	6	0,59	XMR	0	15	0,00
XRP	916	4	2,51	MKR	48	30	0,13	BAT	188	40	0,51	XTZ	0	18	0,00
ZEC	675	32	1,85	IOST	6	90	0,02	MANA	66	110	0,18	XVG	0	91	0,00
XVG	598	91	1,64	FUN	0	202	0,00	ZEC	9	32	0,02	ZIL	0	70	0,00

30 min throughout the period, the result is illustrated in Fig. 10. We collect the number of transactions within every 30 min for each cryptocurrency throughout our data period, between 13/02/2019 and 06/04/2021, we then use the non-linear *Spearman* correlation metric to calculate the similarity between each pair of cryptocurrencies. Additionally, we also include the ranking information in this figure to facilitate a comparison between high-ranking and low-ranking cryptocurrencies. Specifically, we classify cryptocurrencies ranked 21[11] or lower as low-ranking, which are colored green in the figure, while the rest are considered high-ranking, which are colored red in the figure.

Our analysis has revealed that low-ranking cryptocurrencies tend to have similar transaction activity patterns. That is, an increase or decrease in the number of transactions of one coin is reflected in trading activity in other coins, resulting in a similarity among their volatility movements. However, the biggest coins such as Bitcoin, Ethereum and Litecoin tend to behave independently from the rest since they show nearly zero correlation with other coins, showing that they have their unique trading patterns. In the cryptocurrency market, only a few assets are well-known and have a large market capitalization, while the majority of coins are unpopular and do not attract much public attention. Therefore, in the absence of noise and trend, we should see real

---

[11] Based on our analysis, the number of circulating cryptocurrencies exceeded 5000 as of April 2021. However, only 20 cryptocurrencies had an average market capitalization (throughout the duration of our dataset) of more than 1.5 billion USD. This is the reason for classifying the top 20 cryptocurrencies with the highest market capitalization as high-ranking cryptocurrencies.

Table 8
Dominance Score (DS) for each cryptocurrency after removing noise and trend, observed via volatility (a) and returns (b), respectively. Each table has 4 columns: Crypto represents cryptocurrency symbols; DS represents Dominance Score values; Rank represents the ranking of cryptocurrencies; Percentage represents the ratio of one cryptocurrency's Dominance Score value to the sum of all Dominance Score values. The values are ordered from high to low.

| (a) Volatility								(b) Returns							
Crypto	TA	Rank	Percentage	Crypto	TA	Rank	Percentage	Crypto	TA	Rank	Percentage	Crypto	TA	Rank	Percentage
BAT	4842	40	13,27	BNT	777	144	2,13	MKR	6328	30	17,32	ETH	429	2	1,17
MKR	3358	30	9,20	ETC	716	28	1,96	FUN	5876	202	16,09	LSK	388	64	1,06
IOST	2026	90	5,55	BTC	691	1	1,89	MANA	3782	110	10,35	BTC	346	1	0,95
ADA	1839	10	5,04	XEM	648	25	1,78	BAT	3489	40	9,55	LTC	187	7	0,51
XTZ	1696	18	4,65	OMG	647	51	1,77	ZRX	2747	51	7,52	XEM	64	25	0,18
ZIL	1612	70	4,42	LSK	612	64	1,68	DOGE	2053	32	5,62	XLM	52	12	0,14
DOGE	1601	32	4,39	ONT	557	37	1,53	OMG	1924	51	5,27	ICX	48	60	0,13
MANA	1573	110	4,31	BCH	552	6	1,51	MAID	1864	117	5,10	REP	23	66	0,06
XVG	1471	91	4,03	NEO	541	22	1,48	BTG	1479	59	4,05	TRX	13	15	0,04
REP	1265	66	3,47	ICX	515	60	1,41	NEO	1105	22	3,03	ETC	4	28	0,01
FUN	1231	202	3,37	XMR	488	15	1,34	XTZ	993	18	2,72	ZIL	3	70	0,01
MAID	1165	117	3,19	BSV	458	11	1,26	EOS	683	11	1,87	XRP	2	4	0,01
ETH	1080	2	2,96	TRX	299	15	0,82	QTUM	664	51	1,82	ZEC	2	32	0,01
BTG	1002	59	2,75	EOS	168	11	0,46	IOST	619	90	1,69	BCH	1	6	0,00
XLM	993	12	2,72	XRP	143	4	0,39	BNT	479	144	1,31	ADA	0	10	0,00
ZEC	879	32	2,41	QTUM	111	51	0,30	ONT	448	37	1,23	BSV	0	11	0,00
ZRX	835	51	2,29	LTC	90	7	0,25	XVG	433	91	1,19	XMR	0	15	0,00
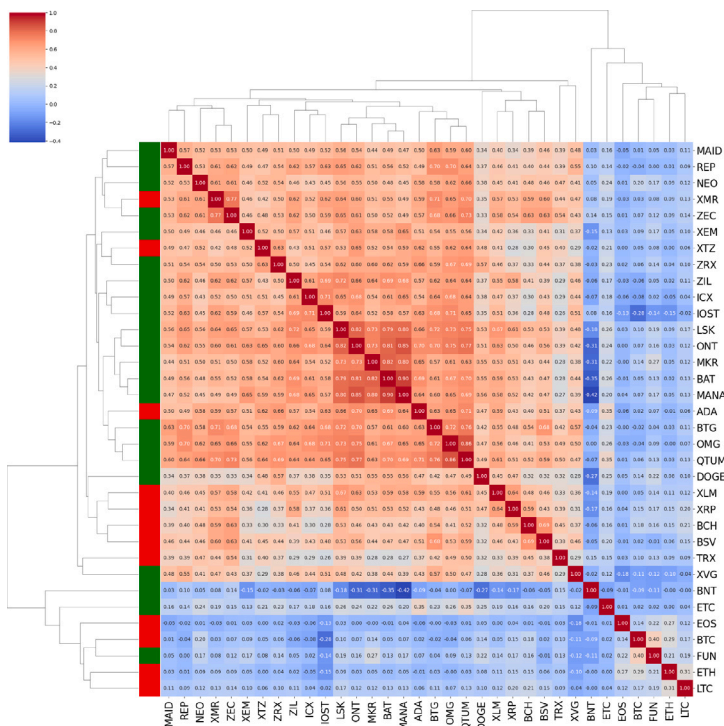
Fig. 10. The correlation between 34 cryptocurrencies in terms of the number of transactions from 13/02/2019 until 06/04/2021 (on a 30-minute basis). Cryptocurrencies which are ranked 21 or lower are considered low-ranking (marked as green), while the rest are considered high-ranking (marked as red).

connections between different cryptocurrencies. That is, when the volatility of a small coin changes, it tends to affect the majority of cryptocurrencies. Conversely, changes in the volatility of a major coin are usually isolated due to its distinct trading pattern.

Overall, there is one main finding that has been uncovered from the results. In particular, the significant impact of the largest coins in the network when considering the noise and trend effect highlights the crucial role of people's sentiment in the cryptocurrency market. Individuals primarily focus on the most prominent coins like BTC, ETH, and LTC, using them as a basis to evaluate the entire market. Consequently, they react when the prices of these coins change and anticipate similar changes in other assets [100,101]. This results in similar trading decisions, ultimately leading to parallel volatility movements between different cryptocurrencies. However, without the effects of noise and trend, the dominance of large cryptocurrencies tends to diminish, giving way to the rising influence of minor and less famous coins in the network referred to above. This observation highlights the true

nature of the cryptocurrency market. That is, small cryptocurrencies are frequently targeted for a Pump & Dump, leading to an increase in their volatility and a spillover effect that impacts a broad spectrum of coins [98]. Additionally, the transaction activity appears to exhibit similarities among the majority of small cryptocurrencies, while major cryptocurrencies demonstrate their own distinctive trading patterns (see Fig. 10). The similarity in transaction activity among small cryptocurrencies might arise from the lead–lag phenomenon[12] between different cryptocurrencies. That is, small cryptocurrencies often follow the price movements of major cryptocurrencies like Bitcoin (BTC) and Ethereum (ETH) with a slight delay, usually within a few minutes. As a result, it is feasible to generate a profit by purchasing small cryptocurrencies whenever the price of Bitcoin (or ETH) increases, waiting a few minutes until the price of the small coins increases as well, and then selling them [102]. Another possible reason for the similarity in transaction activity is that cryptocurrencies with a low market capitalization are often used as safe-havens in a portfolio, especially during turbulent economic and market conditions [103].

Furthermore, we notice another common result from Tables 5 to 8. That is, the cryptocurrencies with the highest influence are BAT, MKR and FUN. This seems to stem from their increasing popularity and use cases. Specifically, these cryptocurrencies gain higher and higher usage demand and are associated with various cryptocurrencies in the market. In particular:

- **BAT**[13]: Since the beginning, BAT has been shown to have a robust business model and real-world use cases [104]. This cryptocurrency has experienced a continuous surge in usage, partly thanks to the increase in usage of the Brave browser. To date, Brave has 55 million monthly active users, 16 million daily active users and millions of verified creators accepting BAT. Bat is considered one of the most successful alt-coins to date. Moreover, it is bridged across Ethereum and Solana blockchains and offers utility to both ecosystems [105,106].
- **MKR**[14]: MKR is created to govern the MakerDAO platform. Specifically, MakerDAO offers its community of MKR holders the right to vote on risk management and business logic. Thus, this coin plays an important role in the development of the platform [107]. Moreover, MKR is also utilized for paying various fees associated with generating DAI- one of the most famous stablecoins at the time of writing [107]. Initially, DAI served primarily as a means of lending and borrowing crypto assets on MakerDAO. Over time, its utility expanded to a wide range of applications on the Ethereum blockchain such as the creation of smart contracts. Thanks to the success of the Ethereum blockchain, DAI maintains its popularity and is relevant to various cryptocurrencies [107,108]. Consequently, MKR experiences increased usage and is associated with numerous cryptocurrencies, as a result of the DAI's flourishing.
- **FUN**[15]: FUN is the only token accepted for making as well as receiving payment from players and game creators, respectively, on FunFair. Thus, players have to exchange other cryptocurrencies or fiat currencies into FUN to be able to play on FunFair. This online gaming platform has experienced a significant increase in different aspects, including business scale, daily active users, and net worth [109]. Especially, since the outbreak of the pandemic, online gaming platforms have attracted a record amount of users as a result of their being forced to stay at home due to the lockdown policy [110]. Consequently, the demand for using FUN keeps increasing over time.

6. Conclusions

In this study, we constructed a time-varying network of 34 cryptocurrencies with a mix of low and high rankings. Our main contributions to the existing literature arise from 2 aspects: Firstly, the use not only of the return time series but also the volatility time series to observe the time-varying correlations between different cryptocurrencies; Secondly, the use of a noise and trend removal scheme, which is not widely considered in the cryptocurrency market at the moment. To our knowledge, this is one of the first studies using volatility time series to calculate the correlation coefficients between different cryptocurrencies. Our expectation is that volatility information can provide new insights distinct from return information. Thus, this draws a more complete picture of cryptocurrency correlations. Besides this, removing noise and trend from cryptocurrencies is expected to help us find out underlying characteristics of the cryptocurrency market that are invisible with the existence of noise and trend.

We found that the use of volatility provides different signals compared to the return one. In particular, the returns-based correlation network reacts to the change in market condition between January and July 2020 (e.g. during the outbreak of the Covid-19 pandemic and the global economic crisis) by changing its structure and this change lasts until the market recovers to its previous condition. In the meantime, the volatility-based network can point out the most changing period within a critical event, such as the worst period of the market crash in March 2020 and the most bullish period in 2021. From these findings, investors can adjust their portfolios to benefit from a particular market condition. Moreover, they can anticipate the movements in the market, especially major events such as the economic crisis and health crisis by using the volatility-based network, thereby preparing and proposing a reasonable investment strategy to adapt to those movements.

---

[12] A lead–lag effect describes the situation where one (leading) variable is cross-correlated with the values of another (lagging) variable at later times.

[13] A cryptocurrency invented for revolutionizing the advertising industry on the internet. This coin is used on the Brave web browser as a reward for Brave users when they read an ads. This browser has a unique mechanism to block ads and only recommends relevant ads based on users' preferences (https://basicattentiontoken.org/).

[14] A main token on a peer-to-peer decentralized protocol called MakerDAO. This platform serves as an online bank for cryptocurrencies, facilitating borrowing, lending and savings of various cryptocurrencies (https://makerdao.com/en/).

[15] A token on the decentralized, cryptocurrency-based casino gaming platform called FunFair (https://funfair.games/), which is the number one online gaming platform in the world.

We also used the *PageRank* method to analyze the influence of each cryptocurrency on the collection of 34 coins collected from the HitBTC exchange and found contradictory results between pre-and post-removal of the noise and trend. Specifically, low-ranking cryptocurrencies have the greater influence on the collection after denoising and detrending while the most influential coins in the pre-denoised and detrended collection are high-ranking cryptocurrencies (e.g. BTC, ETH and BCH). More interestingly, we found that the most influential cryptocurrencies after denoising and detrending are BAT, MKR and FUN. This is explained by their increasing popularity and use cases. From these results, investors should be aware of low-ranking cryptocurrencies and not underestimate them. Moreover, knowing the correlation between different cryptocurrencies helps investors adjust their portfolios to reduce the risk while maximizing future returns. For instance, they can avoid investing in cryptocurrencies with strong correlations and diversify their portfolios by choosing cryptocurrencies with low correlations.

This work can be expanded in the future in different ways: Firstly, one can take the correlation of different cryptocurrencies into consideration to create portfolio optimization models. For example, the models can avoid choosing strongly correlated cryptocurrencies while spending a portion of their budget on safe-haven cryptocurrencies; Secondly, researchers can expand on this study's concepts to other financial asset classes such as stocks, commodities as well as a combination of different asset classes.

CRediT authorship contribution statement

An Pham Ngoc Nguyen: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing – original draft, Visualization. **Tai Tan Mai:** Methodology, Validation, Writing – review & editing, Supervision. **Marija Bezbradica:** Conceptualization, Methodology, Validation, Data curation, Writing – review & editing, Supervision. **Martin Crane:** Conceptualization, Methodology, Validation, Data curation, Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgments

References

[1] G. Giudici, A. Milne, D. Vinogradov, Cryptocurrencies: market analysis and perspectives, J. Ind. Bus. Econ. 47 (2020) 1–18.
[2] Y. Liu, A. Tsyvinski, Risks and returns of cryptocurrency, Rev. Financ. Stud. 34 (6) (2021) 2689–2727.
[3] D. Fantazzini, S. Zimin, A multivariate approach for the simultaneous modeling of market risk and credit risk for cryptocurrencies, J. Ind. Bus. Econ. 47 (2020) 19–69.
[4] N. James, M. Menzies, Collective correlations, dynamics, and behavioural inconsistencies of the cryptocurrency market over time, Nonlinear Dynam. 107 (4) (2022) 4001–4017.
[5] D. Yermack, Is Bitcoin a real currency? An economic appraisal, in: Handbook of Digital Currency, Elsevier, 2015, pp. 31–43.
[6] A. Bejaoui, W. Frikha, A. Jeribi, A.F. Bariviera, Connectedness between emerging stock markets, gold, cryptocurrencies, DeFi and NFT: Some new evidence from wavelet analysis, Physica A 619 (2023) 128720.
[7] M. Balcilar, H. Ozdemir, B. Agan, Effects of COVID-19 on cryptocurrency and emerging market connectedness: Empirical evidence from quantile, frequency, and lasso networks, Physica A 604 (2022) 127885.
[8] I. Omri, Directional predictability and volatility spillover effect from stock market indexes to Bitcoin: evidence from developed and emerging markets, J. Risk Financ. (ahead-of-print) (2023).
[9] M. Wątorek, J. Kwapień, S. Drożdż, Cryptocurrencies are becoming part of the world global financial market, Entropy 25 (2) (2023) 377.
[10] S. Hyun, J. Lee, J.-M. Kim, C. Jun, What coins lead in the cryptocurrency market: using Copula and neural networks models, J. Risk Final. Manag. 12 (3) (2019) 132.
[11] M. Maiti, Y. Vyklyuk, D. Vuković, Cryptocurrencies chaotic co-movement forecasting with neural networks, Int. Technol. Lett. 3 (3) (2020) e157.
[12] P. Pagnottoni, Superhighways and roads of multivariate time series shock transmission: Application to cryptocurrency, carbon emission and energy prices, Physica A 615 (2023) 128581.
[13] A.P.N. Nguyen, T.T. Mai, M. Bezbradica, M. Crane, The cryptocurrency market in transition before and after covid-19: An opportunity for investors? Entropy 24 (9) (2022) 1317.
[14] S. Drożdż, J. Kwapień, M. Wątorek, What is mature and what is still emerging in the cryptocurrency market? Entropy 25 (5) (2023) 772.
[15] I. Cohen, Y. Huang, J. Chen, J. Benesty, J. Benesty, J. Chen, Y. Huang, I. Cohen, Pearson correlation coefficient, in: Noise Reduction in Speech Processing, Springer, 2009, pp. 1–4.
[16] J.H. Zar, Spearman rank correlation, in: Encyclopedia of Biostatistics, Vol. 7, Wiley Online Library, 2005.

[17] E. Keogh, C.A. Ratanamahatana, Exact indexing of dynamic time warping, Knowl. Inf. Syst. 7 (2005) 358–386.
[18] V. Marella, B. Upreti, J. Merikivi, V.K. Tuunainen, Understanding the creation of trust in cryptocurrencies: the case of Bitcoin, Electron. Mark. 30 (2) (2020) 259–271.
[19] S.A. Manavi, G. Jafari, S. Rouhani, M. Ausloos, Demystifying the belief in cryptocurrencies decentralized aspects. A study of cryptocurrencies time cross-correlations with common currencies, commodities and financial indices, Physica A 556 (2020) 124759.
[20] T.T. Mai, M. Bezbradica, M. Crane, Learning behaviours data in programming education: Community analysis and outcome prediction with cleaned data, Future Gener. Comput. Syst. 127 (2022) 42–55.
[21] M.M.L. de Prado, Machine Learning for Asset Managers, Cambridge University Press, 2020.
[22] J.-P. Bouchaud, M. Potters, Financial applications of random matrix theory: a short review, 2009, arXiv preprint arXiv:0910.1205.
[23] T.C. Silva, P.V.B. Wilhelm, B.M. Tabak, The effect of interconnectivity on stock returns during the Global Financial Crisis, North Am. J. Econ. Finance 67 (2023) 101940.
[24] S. Shafiq, S.S. Qureshi, M. Akbar, Dynamic relationship of volatility of returns across different markets: evidence from selected next 11 countries, J. Econ. Adm. Sci. (2023) no. ahead-of-print.
[25] T. Lan, L. Shao, H. Zhang, C. Yuan, The impact of pandemic on dynamic volatility spillover network of international stock markets, Empir. Econ. (2023) 1–30.
[26] B. Korkusuz, D.G. McMillan, D. Kambouroudis, Complex network analysis of volatility spillovers between global financial indicators and G20 stock markets, Empir. Econ. 64 (4) (2023) 1517–1537.
[27] K.-H. Choi, R.P. McIver, S. Ferraro, L. Xu, S.H. Kang, Dynamic volatility spillover and network connectedness across ASX sector markets, J. Econ. Finance (2021) 1–15.
[28] S.J.H. Shahzad, E. Bouri, L. Kristoufek, T. Saeed, Impact of the COVID-19 outbreak on the US equity sectors: Evidence from quantile return spillovers, Final. Innov. 7 (1) (2021) 1–23.
[29] G.B. Koumou, Diversification and portfolio theory: a review, Financial Mark. Portfolio Manag. 34 (3) (2020) 267–312.
[30] D.K. Nguyen, N. Topaloglou, T. Walther, Asset classes and portfolio diversification: Evidence from a stochastic spanning approach, 2020, Available at SSRN 3721907.
[31] B. Abuzayed, N. Al-Fayoumi, Diversification and hedging strategies of green bonds in financial asset portfolios during the COVID-19 pandemic, Appl. Econ. (2022) 1–11.
[32] Y. Ghabri, K. Guesmi, A. Zantour, Bitcoin and liquidity risk diversification, Finance Res. Lett. 40 (2021) 101679.
[33] H.M. Reyad, M.A. Zariyawati, T.S. Ong, H. Muhamad, The impact of macroeconomic risk factors, the adoption of financial derivatives on working capital management, and firm performance, Sustainability 14 (21) (2022) 14447.
[34] J. Gavin, M. Crane, Community detection in cryptocurrencies with potential applications to portfolio diversification, 2021, arXiv preprint arXiv:2108.09763.
[35] H. Chaudhari, M. Crane, Cross-correlation dynamics and community structures of cryptocurrencies, J. Comput. Sci. 44 (2020) 101130.
[36] C. Chen, L. Liu, N. Zhao, Fear sentiment, uncertainty, and bitcoin price dynamics: The case of COVID-19, Emerg. Mark. Finance Trade 56 (10) (2020) 2298–2309.
[37] M. Jose, G. Silvio, R. Heather, S. Adel, C. Martin, Correlation of Worldwide Markets' Entropies, Centro, 2006.
[38] M. Kumar, G.S. Priya, P. Gadipudi, I. Agarwal, S. Sumalatha, Exploring patterns and correlations between cryptocurrencies and forecasting crypto prices using influential tweets, in: Machine Learning, Image Processing, Network Security and Data Sciences: 4th International Conference, MIND 2022, Virtual Event, January 19–20, 2023, Proceedings, Part II, Springer, 2023, pp. 308–321.
[39] G. Rubbaniy, S. Polyzos, S.K.A. Rizvi, A. Tessema, COVID-19, Lockdowns and herding towards a cryptocurrency market-specific implied volatility index, Econom. Lett. 207 (2021) 110017.
[40] S.A. Gyamerah, COVID-19 pandemic and herding behaviour in cryptocurrency market, Appl. Finance Lett. 10 (2021) 58–66.
[41] S. Drożdż, J. Kwapień, P. Oświęcimka, T. Stanisz, M. Wątorek, Complexity in economic and social systems: Cryptocurrency market at around COVID-19, Entropy 22 (9) (2020) 1043.
[42] S. Nayak, B.B. Misra, H.S. Behera, Impact of data normalization on stock index forecasting, Int. J. Comput. Inf. Syst. Ind. Manage. Appl. 6 (2014) (2014) 257–269.
[43] C. Bandt, B. Pompe, Permutation entropy: a natural complexity measure for time series, Phys. Rev. Lett. 88 (17) (2002) 174102.
[44] R. Lopez-Ruiz, H.L. Mancini, X. Calbet, A statistical measure of complexity, Phys. Lett. A 209 (5–6) (1995) 321–326.
[45] H.Y. Sigaki, M. Perc, H.V. Ribeiro, Clustering patterns in efficiency and the coming-of-age of the cryptocurrency market, Sci. Rep. 9 (1) (2019) 1440.
[46] F. Pérez-Cruz, J.A. Afonso-Rodriguez, J. Giner, Estimating GARCH models using support vector machines, Quant. Finance 3 (3) (2003) 163.
[47] D. Tjøstheim, H. Otneim, B. Støve, Statistical dependence: Beyond Pearson's ρ, Stat. Sci. 37 (1) (2022) 90–109.
[48] S. Lieberson, Limitations in the application of non-parametric coefficients of correlation, Am. Sociol. Rev. 29 (5) (1964) 744–746.
[49] C.-X. Nie, Analysis of critical events in the correlation dynamics of cryptocurrency market, Physica A 586 (2022) 126462.
[50] H.R. Atiya, H.N. Nawaf, Prediction of link weight of bitcoin network by leveraging the community structure, in: IOP Conference Series: Materials Science and Engineering, Vol. 928, IOP Publishing, 2020, 032080.
[51] C.-X. Nie, A network-based method for detecting critical events of correlation dynamics in financial markets, Europhys. Lett. 131 (5) (2020) 50001.
[52] J.-P. Onnela, A. Chakraborti, K. Kaski, J. Kertesz, Dynamic asset trees and Black Monday, Physica A 324 (1–2) (2003) 247–252.
[53] L. Zhao, W. Li, X. Cai, Structure and dynamics of stock market in times of crisis, Phys. Lett. A 380 (5–6) (2016) 654–666.
[54] A. Edelman, N.R. Rao, Random matrix theory, Acta Numer. 14 (2005) 233–297.
[55] J.B. De Long, A. Shleifer, L.H. Summers, R.J. Waldmann, Noise trader risk in financial markets, J. Political Econ. 98 (4) (1990) 703–738.
[56] T. Dimpfl, F.J. Peter, Nothing but noise? Price discovery across cryptocurrency exchanges, J. Financial Mark. 54 (2021) 100584.
[57] T. Li, D. Shin, B. Wang, Cryptocurrency pump-and-dump schemes, 2021, Available at SSRN 3267041.
[58] K. Ciupa, Cryptocurrencies: opportunities, risks and challenges for anti-corruption compliance systems, in: OECD Conference on Global Anti-Corruption and Integrity Forum. Paris, March, 2019, pp. 20–21.
[59] R. Auer, S. Claessens, Regulating cryptocurrencies: assessing market reactions, in: BIS Quarterly Review September, 2018.
[60] A. Sarr, T. Lybek, Measuring liquidity in financial markets, 2002.
[61] V.A. Marchenko, L.A. Pastur, Distribution of eigenvalues for some sets of random matrices, Mat. Sb. 114 (4) (1967) 507–536.
[62] Z. Burda, A. Jarosz, Cleaning large-dimensional covariance matrices for correlated samples, Phys. Rev. E 105 (3) (2022) 034136.
[63] O. Ledoit, M. Wolf, Nonlinear shrinkage estimation of large-dimensional covariance matrices, Ann. Statist. 40 (2) (2012) 1024–1060.
[64] J. Bun, A. Knowles, An optimal rotational invariant estimator for general covariance matrices: The outliers, 2018, Preprint.
[65] R. Markowitz, Cleaning correlation matrices, 2016.
[66] T. Conlon, H.J. Ruskin, M. Crane, Random matrix theory and fund of funds portfolio optimisation, Physica A 382 (2) (2007) 565–576.
[67] L. Yang, M.R. McKay, R. Couillet, High-dimensional MVDR beamforming: Optimized solutions based on spiked random matrix models, IEEE Trans. Signal Process. 66 (7) (2018) 1933–1947.
[68] T.T. Mai, M. Crane, M. Bezbradica, Students' learning behaviour in programming education analysis: Insights from entropy and community detection, Entropy 25 (8) (2023) 1225.

[69] T. Heimo, K. Kaski, J. Saramäki, Maximal spanning trees, asset graphs and random matrix denoising in the analysis of dynamics of financial networks, Physica A 388 (2–3) (2009) 145–156.
[70] B. Council, 10 Blockchain protocols to know in 2023, 2023, https://www.blockchain-council.org/blockchain/10-blockchain-protocols/.
[71] S. Zhang, J.-H. Lee, Analysis of the main consensus protocols of blockchain, ICT Express 6 (2) (2020) 93–97.
[72] P.V.J. da Gama Silva, M.C. Klotzle, A.C.F. Pinto, L.L. Gomes, Herding behavior and contagion in the cryptocurrency market, J. Behav. Exp. Finance 22 (2019) 41–50.
[73] M. O'Searcoid, Metric Spaces, Springer Science & Business Media, 2006.
[74] C. Huang, X. Zhao, R. Su, X. Yang, X. Yang, Dynamic network topology and market performance: A case of the Chinese stock market, Int. J. Finance Econ. 27 (2) (2022) 1962–1978.
[75] B. Ghosh, M. Motagh, M.H. Haghighi, M.S. Vassileva, T.R. Walter, S. Maghsudi, Automatic detection of volcanic unrest using blind source separation with a minimum spanning tree based stability analysis, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 14 (2021) 7771–7787.
[76] A. Denkowska, S. Wanat, Linkages and systemic risk in the European insurance sector. New evidence based on Minimum Spanning Trees, Risk Manag. 24 (2) (2022) 123–136.
[77] Q. Nguyen, N. Nguyen, L. Nguyen, Dynamic topology and allometric scaling behavior on the Vietnamese stock market, Physica A 514 (2019) 235–243.
[78] J.B. Kruskal, On the shortest spanning subtree of a graph and the traveling salesman problem, Proc. Am. Math. Soc. 7 (1) (1956) 48–50.
[79] F. Huang, P. Gao, Y. Wang, Comparison of prim and Kruskal on Shanghai and Shenzhen 300 index hierarchical structure tree, in: 2009 International Conference on Web Information Systems and Mining, IEEE, 2009, pp. 237–241.
[80] R.N. Mantegna, Hierarchical structure in financial markets, Eur. Phys. J. B 11 (1999) 193–197.
[81] B. Zhao, W. Yang, J. Wen, W. Zhang, The Financial Market in China under the COVID-19, Emerg. Mark. Finance Trade 58 (13) (2022) 3726–3738.
[82] R. Khanna, I. Guler, Degree assortativity in collaboration networks and invention performance, Strateg. Manag. J. 43 (7) (2022) 1402–1430.
[83] S. Ilbeigipour, B. Teimourpour, A social network analysis approach to evaluate the relationship between the mobility network metrics and the COVID-19 outbreak, Health Serv. Insights 16 (2023) 11786329231173816.
[84] R. Grinis, A. Kislitsyn, I. Drozdov, K. Shulga, Are cryptocurrency markets running behind the fed? A significant shift in crypto markets microstructure, in: A Significant Shift in Crypto Markets Microstructure (October 20, 2021), 2021, Available at SSRN: https://ssrn.com/abstract=3951812.
[85] V. Jha, P. Tripathi, Semantic modelling of multivariate time-series data in cognitive IoT, in: 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Vol. 1, IEEE, 2023, pp. 943–948.
[86] W. Song, Z. Wang, Y. Ye, M. Fan, Empirical studies on symbolic aggregation approximation under statistical perspectives for knowledge discovery in time series, in: 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), IEEE, 2015, pp. 1040–1046.
[87] P. Katsiampa, L. Yarovaya, D. Zięba, High-frequency connectedness between Bitcoin and other top-traded crypto assets during the COVID-19 crisis, J. Int. Final. Mark., Inst. Money 79 (2022) 101578.
[88] A. Kumar, N. Iqbal, S.K. Mitra, L. Kristoufek, E. Bouri, Connectedness among major cryptocurrencies in standard times and during the COVID-19 outbreak, J. Int. Final. Mark., Inst. Money 77 (2022) 101523.
[89] W. Huang, P. Molyneux, S. Ongena, R. Xie, The new challenges of global banking and finance, Eur. J. Finance (2023) 1–7.
[90] J. Almeida, T.C. Gonçalves, A systematic literature review of investor behavior in the cryptocurrency markets, J. Behav. Exp. Finance (2023) 100785.
[91] E. Bouri, G. Azzi, A.H. Dyhrberg, On the return-volatility relationship in the Bitcoin market around the price crash of 2013, Economics 11 (1) (2017).
[92] N.B. Cheikh, Y.B. Zaied, J. Chevallier, Asymmetric volatility in cryptocurrency markets: New evidence from smooth transition GARCH models, Finance Res. Lett. 35 (2020) 101293.
[93] S. Kakinaka, K. Umeno, Asymmetric volatility dynamics in cryptocurrency markets on multi-time scales, Res. Int. Bus. Finance 62 (2022) 101754.
[94] D. Easley, N.M. Kiefer, M. O'hara, J.B. Paperman, Liquidity, information, and infrequently traded stocks, J. Finance 51 (4) (1996) 1405–1436.
[95] J. Collins, Crypto, crime and control, 2022.
[96] J.Y. Song, W. Chang, J.W. Song, Cluster analysis on the structure of the cryptocurrency market via Bitcoin–Ethereum filtering, Physica A 527 (2019) 121339.
[97] J. Kwapień, M. Wątorek, S. Drożdż, Cryptocurrency market consolidation in 2020–2021, Entropy 23 (12) (2021) 1674.
[98] A. Dhawan, T.J. Putniņš, A new wolf in town? Pump-and-dump manipulation in cryptocurrency markets, Rev. Finance 27 (3) (2023) 935–975.
[99] M. Balcilar, H. Ozdemir, On the risk spillover from bitcoin to altcoins: The fear of missing out and pump-and-dump scheme effects, J. Risk Final. Manag. 16 (1) (2023) 41.
[100] P.D. DeVries, An analysis of cryptocurrency, bitcoin, and the future, Int. J. Bus. Manag. Commer. 1 (2) (2016) 1–9.
[101] H. Bukhary, K. Shibano, G. Mogi, in: S.Y. Yurish (Ed.), High-Frequency Spillover Analysis of Cryptocurrency in the Exchange Listings, 2022, p. 61.
[102] B. Anderson, A tick-by-tick level measurement of the lead-lag duration between cryptocurrencies: The case of Bitcoin versus Cardano, Invest. Manag. Final. Innov. 20 (1) (2023) 174–183.
[103] N. Kyriazis, S. Papadamou, P. Tzeremes, S. Corbet, The differential influence of social media sentiment on cryptocurrency returns and volatility during COVID-19, Q. Rev. Econ. Finance 89 (2023) 307–317.
[104] A. Bylund, 3 cryptocurrencies to buy in any bear market, 2023, https://www.fool.com/investing/2023/01/30/3-cryptocurrencies-to-buy-in-any-bear-market/.
[105] J. Caroline, Brave browser surpasses 25 million monthly active users, 2021, https://zycrypto.com/brave-browser-surpasses-25-million-monthly-active-users/.
[106] Y. Khatri, Privacy-focused brave browser added 1M new users in March, 2020, https://finance.yahoo.com/news/privacy-focused-brave-browser-added-190400408.html.
[107] Kriptomat, What is cryptocurrency maker (MKR) and how does it work?, 2023, https://kriptomat.io/cryptocurrencies/maker/what-is-maker/.
[108] M. Young, How dai is dominating the world of NFTs and digital art, 2020, https://beincrypto.com/how-dai-is-dominating-the-world-of-nfts-and-digital-art/.
[109] M. Pechman, FunFair (FUN) gains 183% as the blockchain-based gaming industry grows, 2021, https://cointelegraph.com/news/funfair-fun-gains-183-as-the-blockchain-based-gaming-industry-grows.
[110] E. Howcroft, Virtual real estate plot sells for record $2.4 million, 2021, https://www.reuters.com/markets/currencies/virtual-real-estate-plot-sells-record-24-million-2021-11-23/.