



# Early Warning Signals of Financial Crises Using Persistent Homology and Critical Slowing Down: Evidence From Different Correlation Tests

Mohd Sabri Ismail\*, Mohd Salmi Md Noorani, Munira Ismail and Fatimah Abdul Razak

Department of Mathematical Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Malaysia

## OPEN ACCESS

### Edited by:

Mohd Hafiz Mohd,  
Universiti Sains Malaysia  
(USM), Malaysia

### Reviewed by:

Majid Khan,  
Universiti Sains Malaysia  
(USM), Malaysia  
Marian Gidea,  
Yeshiva University, United States

### \*Correspondence:

Mohd Sabri Ismail  
sabrilmohd92@gmail.com

### Specialty section:

This article was submitted to  
Mathematics of Computation and  
Data Science,  
a section of the journal  
Frontiers in Applied Mathematics and  
Statistics

Received: 10 May 2022

Accepted: 07 June 2022

Published: 30 June 2022

### Citation:

Ismail MS, Md Noorani MS, Ismail M  
and Abdul Razak F (2022) Early  
Warning Signals of Financial Crises  
Using Persistent Homology and  
Critical Slowing Down: Evidence From  
Different Correlation Tests.  
*Front. Appl. Math. Stat.* 8:940133.  
doi: 10.3389/fams.2022.940133

In this study, a new market representation from persistence homology, known as the  $L^1$ -norm time series, is used and applied independently with three critical slowing down indicators [autocorrelation function at lag 1, variance, and mean for power spectrum (MPS)] to examine two historical financial crises (Dotcom crash and Lehman Brothers bankruptcy) in the US market. The captured signal is the rising trend in the indicator time series, which can be determined by Kendall's tau correlation test. Furthermore, we examined Pearson's and Spearman's rho correlation tests as potential substitutes for Kendall's tau correlation. After that, we determined a correlation threshold and predicted the whole available date. The point of comparison between these correlation tests is to determine which test is significant and consistent in classifying the rising trend. The results of such a comparison will suggest the best test that can classify the observed rising trend and detect early warning signals (EWSs) of impending financial crises. Our outcome shows that the  $L^1$ -norm time series is more likely to increase before the two financial crises. Kendall's tau, Pearson's, and Spearman's rho correlation tests consistently indicate a significant rising trend in the MPS time series before the two financial crises. Based on the two evaluation scores (the probability of successful anticipation and probability of erroneous anticipation), by using the  $L^1$ -norm time series with MPS, our result in the whole prediction demonstrated that Spearman's rho correlation (46.15 and 53.85%) obtains the best score as compared to Kendall's tau (42.31 and 57.69%) and Pearson's (40 and 60%) correlations. Therefore, by using Spearman's rho correlation test,  $L^1$ -norm time series with MPS is shown to be a better way to detect EWSs of US financial crises.

**Keywords:** topological data analysis, persistent homology, critical transition, critical slowing down, correlation tests, early warning signal, financial crises, complex system

## INTRODUCTION

Understanding behaviors of financial crises (unexpected and huge declines in the stock market) are crucial to explaining the dynamics of the financial market. However, such financial events are very challenging to study. Among them, the challenges are understanding how the financial market behaves before financial crises and developing a method that is capable to detect early warning signals (EWSs) of the financial crises. The detection method will become more beneficial if it can

accurately distinguish the financial market into two classification periods, which are EWS periods (indicating a possible huge downtrend is coming in the market) and safe periods. Such a method helps investors and traders to develop an early precaution strategy and protect their investments from any losses due to a financial crisis.

In practice, the Chicago Board Options Exchange's CBOE Volatility Index or VIX index—a real-time market index representing the market's expectations for volatility over the coming 30 days—is always used by market participants to alert any upcoming financial crisis. The uptrend patterns in the VIX were observed before financial crises, unfortunately, low-level VIX also reported occurred before financial crises [1, 2]. This suggests that the VIX can provide EWSs; however, a reliable method to predict financial crises remains a challenge in this field. Therefore, extensive studies have been conducted by many researchers in an attempt to provide other rationales to explain why financial crises occurred and provide any other possible EWS. Some of those methods are bubble theory [3–5], financial stress indicator [6, 7], information-based measures [8, 9], financial network analysis [10–12], and graphical analysis [13].

However, this study is focusing on critical transition theory. This theory viewed the financial market as a complex system with episodes of critical transitions (abrupt shifts from a current stable state to another stable state when the system reaches critical points) [14, 15]. When approaching a critical point, this theory stated that a generic phenomenon happened known as critical slowing down (CSD) because of decreasing stability in the market and its recovery rate took longer to preserve the stability. At a critical point, the financial market loses its stability in the dynamic state and then suddenly causes a market movement into a financial crisis. CSD gives a rising trend in the time series of some indicators, such as autocorrelation function at lag 1 (AC1), variance (VAR), and mean for power spectrum (MPS) at low frequencies [16–18].

In the traditional method, the rising trend in the time series of CSD indicator that happens before the critical transition point is determined using Kendall's tau correlation test. By using Kendall's tau correlation test, previous studies have shown that the observed rising trend can provide EWSs before financial crises [19–23]. Despite the successful results obtained, there is a realization that some indicator like AC1 also tends to decline before financial crises. All of these lead to mixing results such as recorded in Guttal et al. [21] and Diks et al. [23]. However, as compared to critical transitions, Guttal et al. [21] argued that financial crises are more likely to follow stochastic transition, in which variability indicators (VAR and MPS) can signal early warnings for financial markets. This also points out that abrupt transitions in the financial market are hardly the same as a complex system in nature such as the earth's global climate and interaction between species in ecology. One of the reasons behind this is the financial market is involved with human behaviors that influence the market's movement [23].

Recently, using persistent homology (PH) (a robust method to compute topological features of financial data at different spatial scales [24]), Gidea and Katz [25] suggested a new market representative obtained from persistence landscapes

called  $L^1$ -norm time series. Since the persistence landscapes are robust under perturbations of the underlying data, the  $L^1$ -normtime series has the advantage to reflect the loss of stability in dynamic states of the original system. When the stock market becomes more volatile, loops in the relevant point clouds become much more pronounced and give more significant features within persistence landscapes [25]. At that time, the corresponding  $L^1$ -normvalues are more likely to jump up, and this behavior is believed to correspond with an undergoing CSD. Therefore, to test the presence of CSD, the CSD indicators are used with the  $L^1$ -normtime series, where these indicators should alert upcoming financial crises by showing their significant rising trend that can be classified by Kendall's tau correlation test. In their work, Gidea and Katz [25] showed that the  $L^1$ -normtime series grew substantially before the Dotcom crash and Lehman Brothers bankruptcy in the US market. Interestingly, by using CSD indicators and Kendall's tau correlation test on the  $L^1$ -normtime series, Gidea and Katz [25] confirmed that a significant strong rising trend happen in the MPS time series before those two financial crises. Therefore, the method is suggested as a new potential EWS. Later, the application of PH to financial data analysis has attracted more attention from researchers, especially in detecting EWSs. All the articles published in this area are discussed in Section Literature Review.

In this article, we also used PH and CSD indicators (AC1, VAR, and MPS) to detect EWSs of financial crises. Unlike Gidea and Katz [25], instead of using Kendall's tau correlation test to indicate the rising trend in the time series of the indicators, we also examined Pearson's and Spearman's rho correlation tests as a potential substitute for Kendall's tau correlation. The point of comparison between these correlation tests is to determine which test is significant and consistent in classifying the rising trend. The results of such a comparison will suggest the best test that can classify the observed rising trend and detect EWSs of impending financial crises. The remaining portion of this article is organized as follows. Section Literature Review briefly discusses our literature review, Section Persistent Homology introduces PH, Section Data Analysis elaborates on our data, Section Methods presents the applied methods, Section Result mentions our results and their corresponding discussions, and Section Conclusion wraps our conclusion.

## LITERATURE REVIEW

Topological information from financial data obtained using PH has been used to study financial problems. In recent times, this topic has attracted attention from many researchers around the globe. In the finance market, the most application currently investigated by PH is financial crises and their EWS detection tools. In a multivariate setting, by begin with examining chaotic time series with noise and a growing variance, Gidea and Katz [25] showed that PH can exhibit strong growth in its market representative time series prior to the critical point. This growth can be analyzed using CSD indicators (AC1, VAR, and MPS) and correlation tests (Kendall's tau) to indicate corresponding

CSD that happen before critical point. After that, Gidea and Katz [25] used the pipeline method to examine the Dotcom crash and Lehman Brothers bankruptcy in the US market as mentioned earlier.

After that, Ismail et al. [26] expanded the study of Gidea and Katz [25] by using PH and CSD indicators to examine financial crises in the US, Singapore, and Malaysia markets. Aspects of the method's robustness and prediction performance have been rigorously evaluated in Ismail et al. [26]. Meanwhile, Aromi et al. [27] substantiated that PH reflects changes in the underlying multivariate distribution and strong covariance could nullify the persistence of homologies. On the other hand, [28] also analyzed time-dependent correlation networks using PH to detect early signs of critical transitions in financial data.

In Guo et al. [29], an EWS based on PH is also built to detect the critical dates on the financial time series. Guo et al. [30] also used topological features of complex networks that were extracted using PH to find critical dates for financial crises. In addition, dynamics of financial market correlations based on topology and geometry are examined using PH in Yen et al. [31]. Moreover, Yen and Cheong [32] also tested PH to analyze Singapore and Taiwan markets. The extreme event called flash crash was also explored in Kim et al. [33] by applying PH and dynamic time series analysis.

Furthermore, anomalies detection in the dynamics of a market index also is studied with PH [34]. Katz and Biem [35] also showed that early signatures of growing market instabilities can be captured by PH. Besides, Gidea et al. [3] used PH and  $k$ -means clustering to detect critical transitions in the time series of cryptocurrencies. For Bitcoin [36], also uses PH and CSD indicators to detect such transitions and substantiated that PH can detect EWSs better than the detrending time-series approach (the most common approach used in previous studies to detect CSD indicators).

Other than that, PH also has been developed to improve portfolio investment strategies. Studied ten global indices and all their underlying assets, Goel et al. [37] showed that a new strategy based on PH leads to more robust portfolios. Baitinger and Flegel [38] also demonstrated that investment strategies relying on a PH-based turbulence detection outperform investment strategies based on other popular turbulence indices. In clustering and classification of financial time series, Majumdar and Laha [39] showed that PH outperforms other methods in this task.

Additionally, PH is applied with machine learning to predict the movement of financial data. Such a task has been done in Ismail et al. [40] by using PH and machine learning methods (logistic regression, neural network, support vector machine, and random forest) to predict the next-day direction of the Kuala Lumpur Composite Index (KLCI). Moreover, Baitinger and Flegel [41] also introduced PH to produce microstructural predictors, where these predictors are combined with machine learning and statistical factor extraction methods to predict asset returns.

## PERSISTENT HOMOLOGY

Persistent homology is a new quantitative method of topological data analysis to compute topology features (connected components, loops, voids, and others) that persistently emerge across multiple scales. Interestingly, PH is robust to small perturbations of input data, independent of dimensions and coordinates and provides a compact representation of the qualitative features of the input [42–44]. These characteristics make PH suitable to analyze complex, non-linear, noisy, and high-dimensional data like financial data [45].

Later, a summary of PH in the way that it is used in this article is provided. We noted that the concepts of PH stated here can also be found in most books and journals related to PH. To further explore the theories and other concepts of PH, we recommend Otter et al. [24] and Edelsbrunner and Harer [46] to the interested readers.

Input data analyzed by PH is called point cloud dataset (PCD), which can be denoted as  $X = \{x_i \in \mathbb{R}^d | i = 1, \dots, n\}$  for  $d \geq 2$ . Let  $X$  be a PCD, a Rips complex at a scale  $\varepsilon > 0$  (denoted as  $R(X, \varepsilon)$ ) can be constructed as follows:

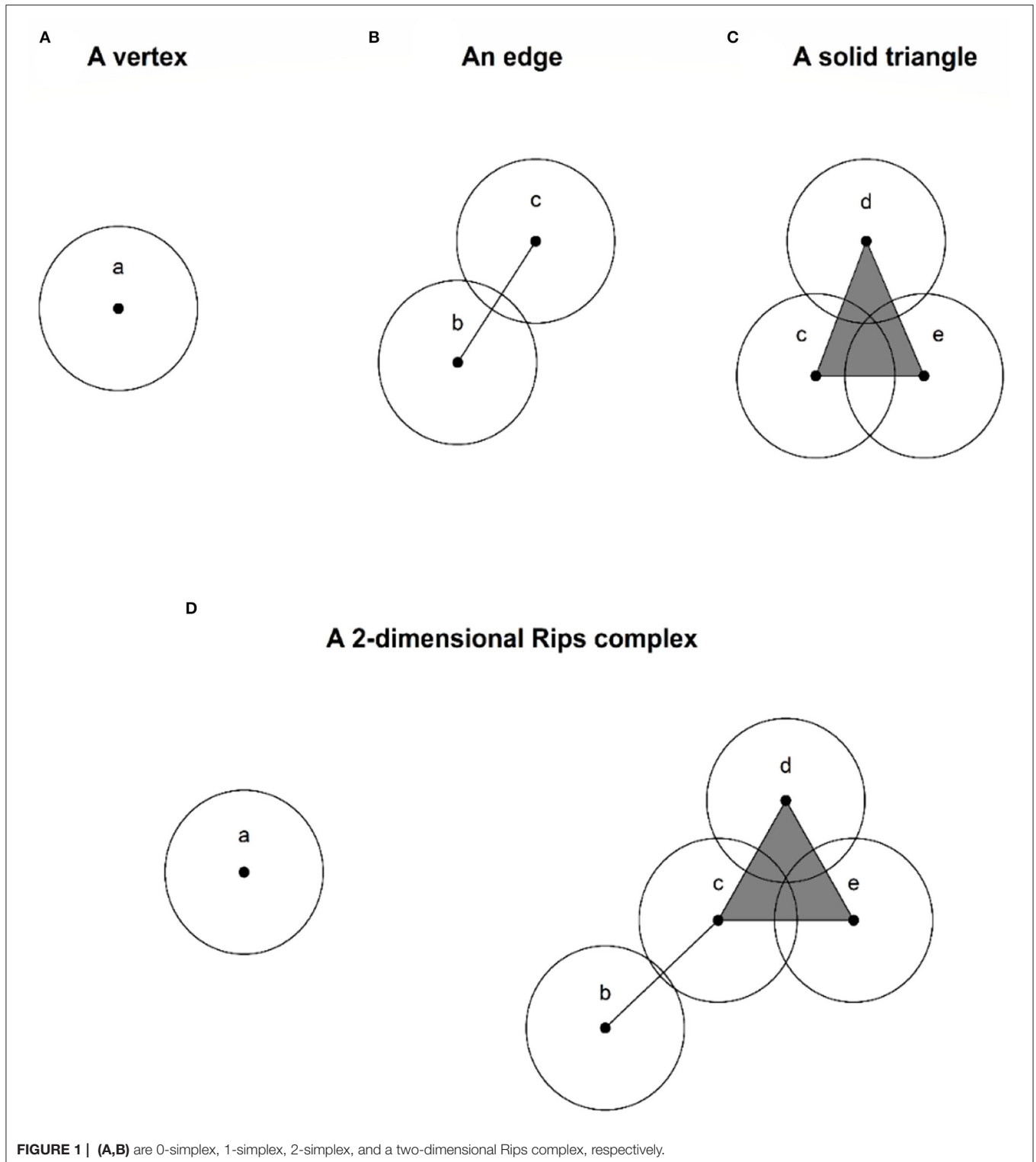
- For each dimensional  $k = 0, 1, \dots$ , a  $k$ -simplex of  $X$  points  $\{x_{i_1}, \dots, x_{i_{(k+1)}}\}$  belongs to a  $R(X, \varepsilon)$  if and only if for every pair  $\{x_{i_r}, x_{i_s}\}$ , we have  $|x_{i_r} - x_{i_s}| \leq \varepsilon$  for all  $x_{i_r}, x_{i_s} \in \{x_{i_1}, \dots, x_{i_{(k+1)}}\}$  [24].

Roughly speaking, a Rips complex is a combination of vertices (0-simplex), edges (1-simplex), solid triangles (2-simplex), and higher dimensional analog, joined according to the above-mentioned rule. **Figure 1** illustrates 0-simplex until 2-simplex and a two-dimensional Rips complex  $R(X, \varepsilon)$ , respectively. In **Figure 1**, the scale  $\varepsilon$  is the diameter of balls around each point in the different PCDs.

In a real-world application, finding a single  $\varepsilon$  is impractical since the real space behind PCD is mostly unknown. Therefore, PH provides a better way to interpret topological information regarding the shape behind PCD by varying the scale  $\varepsilon$ . Let  $\varepsilon \in \{\varepsilon_0, \dots, \varepsilon_m\}$  such that  $0 \leq \varepsilon_0 < \dots < \varepsilon_m$ , then filtration of Rips complexes is  $R(X, \varepsilon_0) \subset \dots \subset R(X, \varepsilon_m)$ . **Figure 2** shows a filtration containing six Rips complexes constructed of a PCD at multiple scales.

For  $i = 1, \dots, m$ , we can compute the  $k$ -homology of each Rips complex  $R(X, \varepsilon_i)$ , denoted as  $H_k(R(X, \varepsilon_i))$ . Roughly speaking,  $H_0(R(X, \varepsilon_i))$  is the free group generated by the connected components of  $R(X, \varepsilon_i)$ ,  $H_1(R(X, \varepsilon_i))$  is the free group generated by the loops in  $R(X, \varepsilon_i)$ ,  $H_2(R(X, \varepsilon_i))$  is the free group generated by the voids of  $R(X, \varepsilon_i)$ . The Betti numbers count the number of generators of such homology groups. This number of generators indicates the number of corresponding topological features that emerge at each scale.

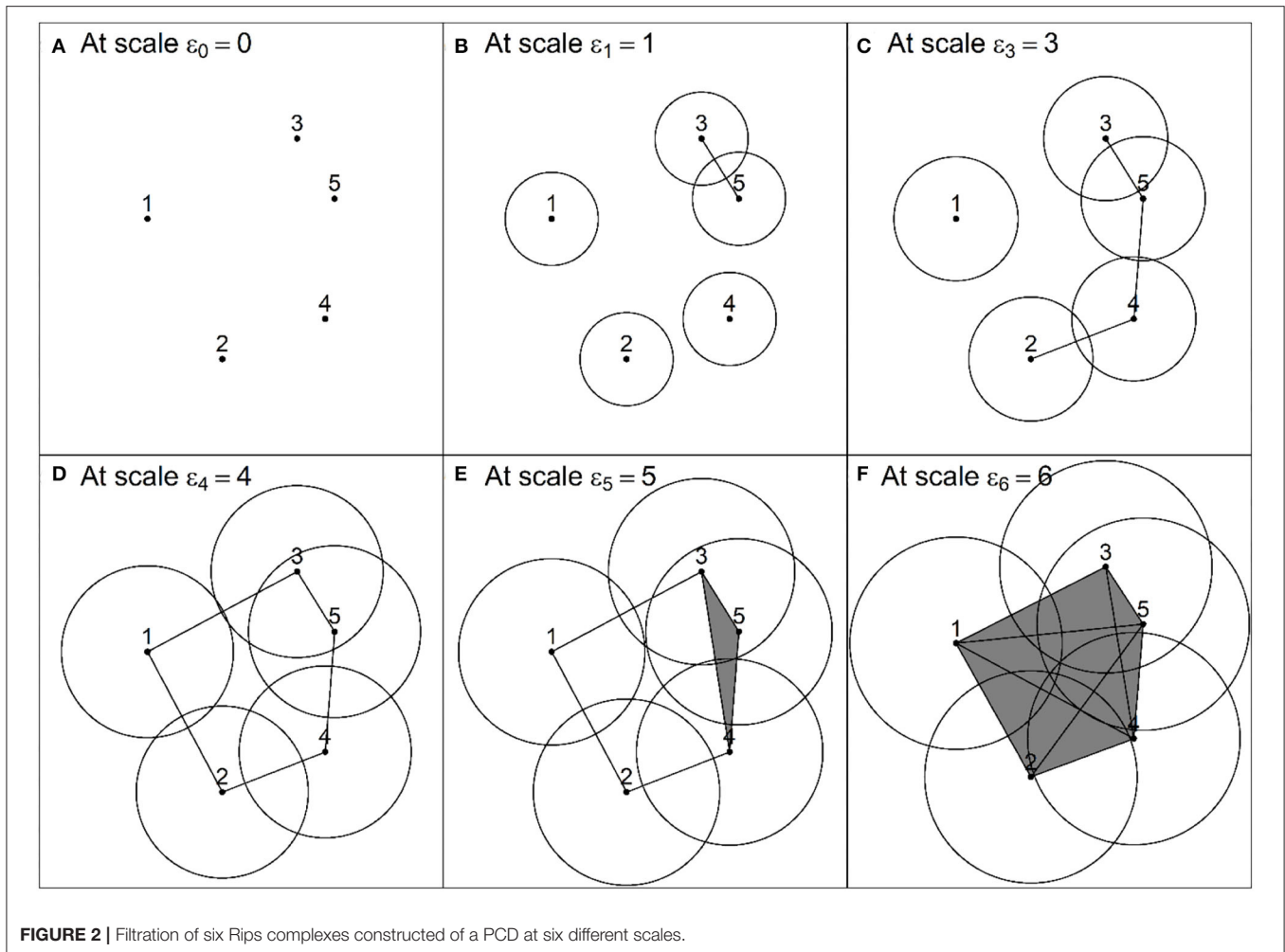
In addition, information regarding the lifespans of topological features also can be obtained by using PH. In the PH community, most believe that a topological feature that has a longer lifespan (persists for a bigger range of scales) can be viewed as a more



significant one, whereas a feature that has a shorter lifespan (persists for a smaller range) can be viewed as a less significant, or a noisy feature. Nevertheless, the theoretical justification for this is unclear and may be dependent on the problem at

hand [45]. However, this study used all the obtained lifespan in our computation.

Let  $\{(\varepsilon_b, \varepsilon_d) \mid i = 1, \dots, n\}$  be a collection of the lifespans of topological features such that  $\varepsilon_b < \varepsilon_d$  and  $\varepsilon_b, \varepsilon_d \in \{\varepsilon_0, \dots, \varepsilon_m\}$ .



**FIGURE 2** | Filtration of six Rips complexes constructed of a PCD at six different scales.

The simplest tool to present this information is a persistent barcode, which is a collection of  $n$  half-closed intervals  $[\varepsilon_b, \varepsilon_d)$  representing the lifespans of topological features, see **Figure 3A**. Other tools are a persistent diagram and a persistent landscape. A persistent diagram is a finite collection of  $n$  birth–death points  $(\varepsilon_b, \varepsilon_d) \in \mathbb{R}^2$  that lie along or above a diagonal line. If there are redundant birth–death points, then the points will be represented as a single point but multiplied by its size to correspond to the frequency of this point. **Figure 3B** provides an example of a persistent diagram.

On the other hand, the persistence landscape  $\lambda$  is quite a recent tool introduced by Bubenik and Dłotko [47] and Bubenik [48]. To define the persistence landscape, we transformed each lifespan into a piecewise linear function  $f_{(\varepsilon_b, \varepsilon_d)_i} : \mathbb{R} \rightarrow [0, \infty)$ , which is defined below:

$$f_{(\varepsilon_b, \varepsilon_d)_i}(x) = \begin{cases} x - \varepsilon_b & \text{if } x \in (\varepsilon_b, (\varepsilon_b + \varepsilon_d)/2] \\ -x + \varepsilon_d & \text{if } x \in ((\varepsilon_b + \varepsilon_d)/2, \varepsilon_d) \\ 0 & \text{if } x \notin ((\varepsilon_b + \varepsilon_d)) \end{cases},$$

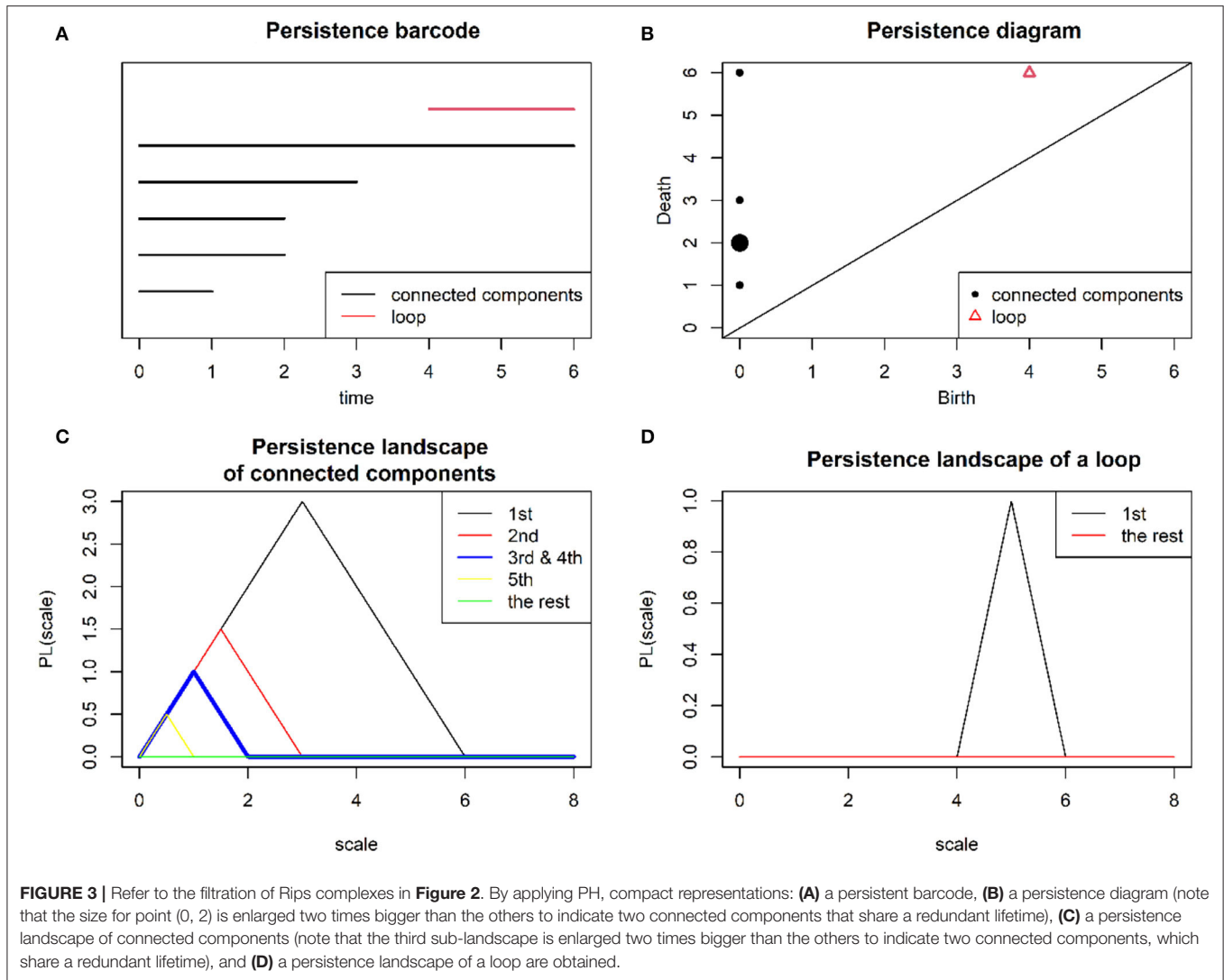
for  $i = 1, \dots, n$ . (1)

As a result, we obtained a set of  $n$  piecewise linear functions  $\{f_{(\varepsilon_b, \varepsilon_d)_i}(x) | i = 1, \dots, n\}$ . Later the layer in the persistence landscape was obtained for each  $k \in \mathbb{Z}$ , which is a function  $\lambda_k(x) : \mathbb{R} \rightarrow [0, \infty)$  that is defined as follows  $\lambda_k(x) = k - \max(\{f_{(\varepsilon_b, \varepsilon_d)_i}(x) | i = 1, \dots, n\})$ , where  $k - \max$  denotes the  $k$ th largest value of the piecewise linear functions. If the  $k$ th largest value does not exist anymore for  $k = l$ , then  $\lambda_k(x) = 0$  for all remaining  $k \geq l$ . The persistence landscape  $\lambda$  is the infinite sequence of  $\lambda_k(x)$ , which can be denoted as  $\{\lambda_1(x), \lambda_2(x), \dots\}$ . **Figures 3C, D** presents the persistence landscapes of connected components and a loop, respectively.

In this study, we used persistent landscapes because this representation can be summarized into a summary point, that is a  $L^p$ -norm. This norm is a function  $\|\cdot\|_p : \lambda \rightarrow \mathbb{R}$ , which is defined as follows:

$$\|\lambda\|_p = \left[ \sum_{k=1}^{\infty} \int |\lambda_k(x)|^p dx \right]^{\frac{1}{p}} \text{ for } 1 \leq p < \infty, \tag{2}$$

$$\text{and } \|\lambda\|_{\infty} = \sup_{k,x} |\lambda_k(x)| \text{ for } p = \infty. \tag{3}$$



Furthermore, statistical properties of these norm values also can be analyzed. Therefore, the CSD indicators are computed from these norm values, and they are used to detect EWSs of impending financial crises in this study.

### SAMPLE DATA

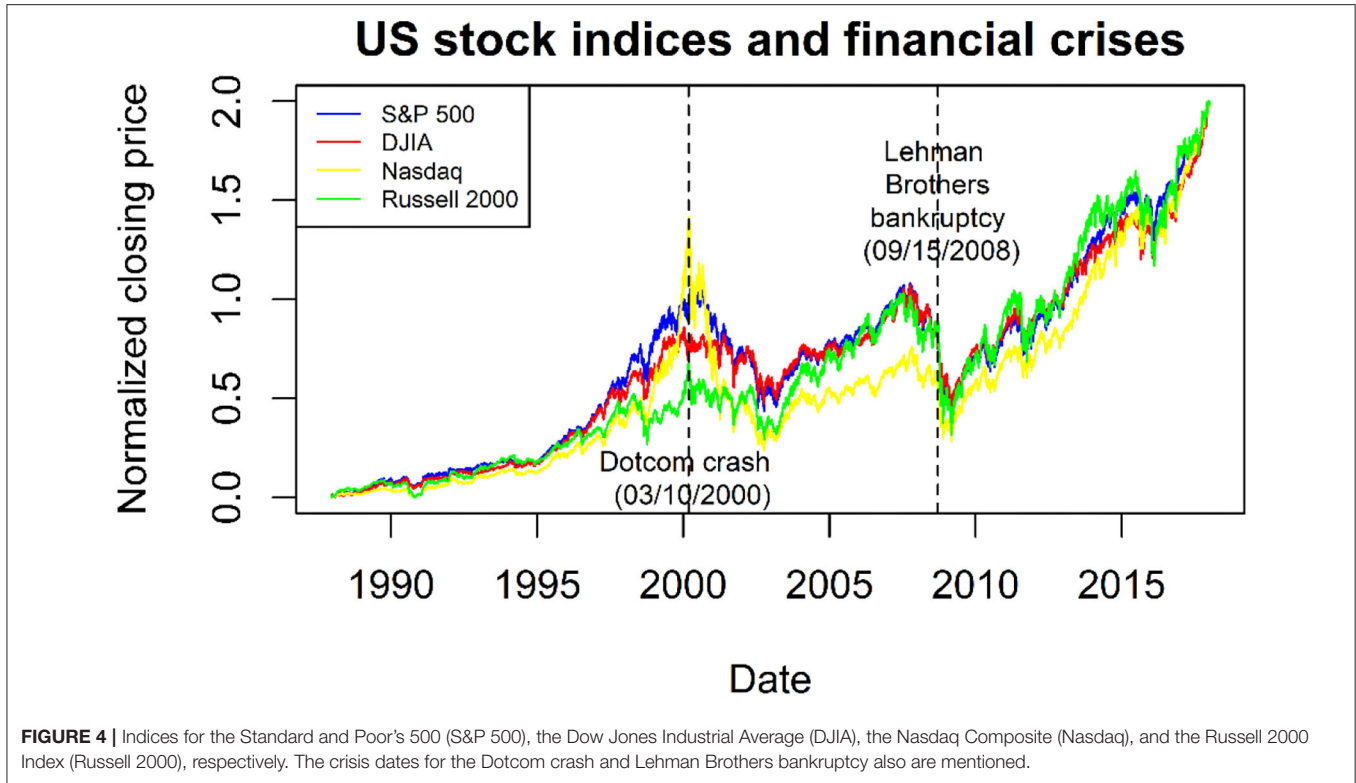
For this study, four main stock indices of the US market were collected, which span from 22/12/1987 until 29/12/2017. The indices were the Standard and Poor's 500 (S&P 500), the Dow Jones Industrial Average (DJIA), the Nasdaq Composite (Nasdaq), and the Russell 2000 Index (Russell 2000), which were derived from the Yahoo Finance. All these indices are shown in **Figure 4**. In **Figure 4**, for illustration purposes, the indices are normalized using the max/min normalization using the formula  $p_{norm} = 2 \times ((p_t - \min) / (\max - \min))$ , where  $p_{t,\max}$  and  $\min$  are the closing price at date  $t$ , the maximum closing price, and the minimum closing price of an index, respectively. Such normalization does not involve our method

as discussed in Section Methods. Furthermore, in **Figure 4**, the crisis dates for the Dotcom crash and Lehman Brothers bankruptcy are mentioned.

### METHODS

#### Pre-crisis Dataset

In this study, two pre-crisis datasets were generated. The two datasets contained closing prices of the four indices (S&P 500, DJIA, Nasdaq, and Russell 2000) with a length of 1000 before the crisis date of the Dotcom crash and Lehman Brothers bankruptcy, respectively. The objective of creating the two pre-crisis datasets was to study the financial market's behaviors before the financial crises, where the obtained information based on PH and CSD would be proceeded to form a EWS detection. Later, we predicted and compared the performance for the case of using different correlation tests: Kendall's tau, Pearson's, and Spearman's rho.



### Persistent Homology

For  $i = 1, \dots, 4$ , 1000 closing prices of an index  $i$  in the pre-crisis were transformed to daily log returns using the formula  $r_t^i = \ln(p_t^i/p_{t-1}^i)$  for  $t = 1, \dots, 1000$ , where  $p_t^i$  is the closing price at date  $t$ . As a result, for each pre-crisis dataset, we obtained a transformed pre-crisis dataset, denoted as  $X = \{r_t \in \mathbb{R}^4 | t = 1, \dots, 1000\}$ , where  $r_t = (r_t^1, \dots, r_t^4)$  is a point of the  $X$  at the date  $t$ .  $X$  can be illustrated in matrix form as below:

$$X = \begin{bmatrix} r_1^1 & r_1^2 & \dots & r_1^4 \\ r_2^1 & r_2^2 & \dots & r_2^4 \\ \vdots & \vdots & \ddots & \vdots \\ r_{1000}^1 & r_{1000}^2 & \dots & r_{1000}^4 \end{bmatrix} \tag{4}$$

Furthermore, the daily sliding window of length 50 approach was applied to segment the  $X$  to obtain PCDs. For financial data analysis, it had been demonstrated that a length of 50 was enough to extract topological information through PH as reported in the previous literature reviews [3, 25, 26, 35]. Consequently, each respective PCD from  $X$  at the date  $t$  also can be illustrated in matrix form as follows:

$$X(t) = \begin{bmatrix} r_{((t-50)+1)}^1 & r_{((t-50)+1)}^2 & \dots & r_{((t-50)+1)}^4 \\ r_{((t-50)+2)}^1 & r_{((t-50)+2)}^2 & \dots & r_{((t-50)+2)}^4 \\ \vdots & \vdots & \ddots & \vdots \\ r_{((t-50)+50)}^1 & r_{((t-50)+50)}^2 & \dots & r_{((t-50)+50)}^4 \end{bmatrix}, \tag{5}$$

for  $t = 1, \dots, 1000$ .

For each  $t$ , as briefed in Section Persistent Homology, PH of Rips filtration that built on PCD  $X(t)$ , the corresponding persistence landscape at the date  $t$  and the corresponding  $L^1$ -norm value at the date  $t$  were computed, accordingly. By doing so, we obtained  $L^1$ -normtime series, denoted by  $Y = \{\|\lambda\|_{1,t} | t = 50, \dots, 1000\}$ , where  $\|\lambda\|_{1,t}$  is a  $L^1$ -norm value at the date  $t$ .

### Critical Slowing Down Indicators

Consequently, the  $L^1$ -normtime series  $Y = \{\|\lambda\|_{1,t} | t = 50, \dots, 1000\}$  was segmented by using the daily sliding window of the length 50 to obtain sequences of 500  $L^1$ -norm values. The length 50 was selected based on the literature mentioned in Guttal et al. [21], Diks et al. [23], and Ismail et al. [26], which uses half of the length of the pre-crisis dataset. Each sequence of 500  $L^1$ -norm values at the date  $t$  can be denoted as below:

$$Y(t) = \{\|\lambda\|_{1,((t-50)+1)}, \dots, \|\lambda\|_{1,((t-50)+50)}\}, \tag{6}$$

for  $t = 549, \dots, 1000$ .

For each  $t \in \{549, \dots, 1000\}$ , we computed a value based on the CSD indicators: AC1, VAR, and MPS at low frequencies as accordingly defined below:

- The AC1 value at trading  $t$  is  $ac1_t = \rho_{1,t} / \text{var}_t$ , where  $\rho_{1,t} = \frac{1}{500-1} \sum_{j=((t-50)+1)}^{((t-50)+50)-1} (\|\lambda\|_{1,j} - \overline{\|\lambda\|_{1,j}}) (\|\lambda\|_{1,(j+1)} - \overline{\|\lambda\|_{1,j}})$ ,  $\overline{\|\lambda\|_{1,j}}$  is the mean of  $Y(t)$  and  $\text{var}_t$  is the VAR of  $Y(t)$  as defined in the point below.

- The VAR value at the date  $t$  is  $\text{var}_t = \frac{1}{500-1} \sum_{j=((t-500)+1)}^{((t-500)+500)-1} (\|\lambda\|_{1,j} - \overline{\|\lambda\|_{1,j}})^2$ , where  $\overline{\|\lambda\|_{1,j}} = \frac{1}{500} \sum_{j=((t-500)+1)}^{((t-500)+500)} \|\lambda\|_{1,j}$  the mean of  $Y(t)$ .
- Given  $Y(t) = \{\|\lambda\|_{1,((t-500)+1)}, \dots, \|\lambda\|_{1,((t-500)+500)}\}$ , we defined its discrete Fourier transformation as  $F_{k,t} = \sum_{j=((t-500)+1)}^{((t-500)+500)} \|\lambda\|_{1,j} \cdot e^{-2\pi i k j / 500}$ , where  $k \in \{1, \dots, 500\}$ . The power spectrum is  $PS_{k,t} = |F_{k,t}|^2$  for  $k \in \{1, \dots, 500\}$ . Then, the MPS value at the date  $t$ , denoted as  $m\text{ps}_t$  is the mean of all  $PS_{k,t}$  for  $k \in \{2, \dots, \frac{1}{8} \times 500\}$ .

As a result, we obtained three sets of time series based on AC1, VAR, and MPS values, which can be denoted as  $AC1 = \{ac1_t | t = 549, \dots, 1000\}$ ,  $VAR = \{\text{var}_t | t = 549, \dots, 1000\}$ , and  $MPS = \{m\text{ps}_t | t = 549, \dots, 1000\}$  accordingly.

### Correlation Tests

In the study, Kendall's tau correlation was used to determine a rising trend in the indicator's time series (AC1, VAR, or MPS). Nonetheless, Pearson's and Spearman's rho correlations were included in addition to Kendall's tau correlation as alternative measures to detect the trend. In brief, Kendall's tau, Pearson's, and Spearman's rho correlations computed the strength of concordance dependency, the range of linear relationship, and the degree of the association, respectively.

Daily sliding window of length 250 was used to attain sequences containing 125 indicator values. In addition, the length of 250 was chosen as half data of the previous sliding window with length of 500. This length was considered sufficient to capture rising trend in the indicators as shown in Guttal et al. [21] and Ismail et al. [26]. Therefore, three tests based on Kendall's tau, Pearson's, and Spearman's rho correlations from each sequence containing 250 indicator values were computed. Considering AC1 as an example, each sequence containing 250 AC1 values at the date  $t$  can be denoted as  $AC1(t) = \{ac1_t | t = ((t-250)+1), \dots, ((t-250)+250)\}$ , for  $t = 798, \dots, 1000$ . For each  $t \in \{798, \dots, 1000\}$ , Kendall's tau, Pearson's, and Spearman's rho correlations from  $AC1(t)$  are computed as follow:

- $\tau_{AC1,t} = \frac{(C-D)/Z}{\sqrt{\frac{2(2n-1)}{n(n-1)}}}$ , where  $C$  is the number of concordant pairs between  $AC1(t)$  and  $\{((t-250)+1), \dots, ((t-250)+250)\}$ ,  $D$  is the number of discordant pairs between  $[[\text{Mathtype-mtefl-eqn-136.mtf}]]$  and  $\{((t-250)+1), \dots, ((t-250)+250)\}$ .  $Z = (250 \times (250 - 1))/2$  is the total number of different possible pair combinations.
- $r_{AC1,t} = \frac{\left(250 \sum_{l=((t-250)+1)}^{((t-250)+250)} ac1_l \cdot l - \left(\sum_{l=((t-250)+1)}^{((t-250)+250)} ac1_l\right) \left(\sum_{l=((t-250)+1)}^{((t-250)+250)} l\right)\right)}{\sqrt{\left(\sum_{l=((t-250)+1)}^{((t-250)+250)} l\right)^2}}$

$$\left( \sqrt{250 \left( \sum_{l=((t-250)+1)}^{((t-250)+250)} (ac1_l)^2 \right) - \left( \sum_{l=((t-250)+1)}^{((t-250)+250)} ac1_l \right)^2} \right) / \sqrt{250 \left( \sum_{l=((t-250)+1)}^{((t-250)+250)} (l)^2 \right) - \left( \sum_{l=((t-250)+1)}^{((t-250)+250)} l \right)^2}$$

- $p_{AC1,t} = 1 - \frac{6 \sum_i d_i^2}{250(250^2-1)}$ , where  $d_i$  is the difference between the ranks of the corresponding variables of  $AC1(t)$ , and  $\{((t-250)+1), \dots, ((t-250)+250)\}$ .

As a result, three sets of time series based on Kendall's tau, Pearson's, and Spearman's rho correlation values computed from AC1 were obtained, which can be denoted as  $\tau_{AC1} = \{\tau_{AC1,t} | t = 798, \dots, 1000\}$ ,  $r_{AC1} = \{r_{AC1,t} | t = 798, \dots, 1000\}$ , and  $p_{AC1} = \{p_{AC1,t} | t = 798, \dots, 1000\}$ , respectively. For, VAR and MPS, the same three sets of time series based on these three correlation values also can be obtained using the above formula.

Note that if any correlation (Kendall's tau, Pearson's, or Spearman's rho) provides a positive real value at the date  $t$ , we conclude that a rising trend happened in the 250 indicator values from the date  $(t-250+1)$  to  $t$ . In critical transition theory, it was expected that these indicators' time series: AC1, VAR, and MPS are increasing before a financial crisis. In addition, the rising trend in the indicators will give rise to a positive real number of the correlation ahead of a financial crisis.

### Significant and Skewness Tests

Furthermore, we performed a significance test in this study to examine whether the rising trend in the last 125 indicator values at the date from 750 to 999, which was summarized by the correlation value at the date  $t = 999$  (1 day before the financial crisis), was statistically significant at level 5%. To conduct the test, we computed the  $p$ -value for each of the correlation values at the date  $t = 999$ . The  $p$ -value obtained is a doubled probability of getting the statistical value or a value with even greater evidence against  $H_0$ .

Let us take AC1 as an example; here,  $H_0$  represents no monotonic trend (either rising or declining) that happens in the time series  $AC1(999)$ . A significance test at level 5% translates to a requirement  $p$ -value less than 0.005 to interpret any monotonic trend (rising or decline) in the time series  $AC1(999)$  as significant or otherwise. The same significant test was also applied to VAR and MPS time series, that is,  $VAR(999)$  and  $MPS(999)$ , respectively. This significance test is vital to verify whether PH via  $L^1$ -norm time series associated with the used indicator (AC1, VAR, or MPS) can provide a reliable EWS at least 1 day before the observed financial crisis.

In addition, since the Pearson correlation is usually applied to time series whose probability distribution is symmetric, such as the normal or Gaussian probability distribution, it is better to consider other association correlation measures (Kendall's tau or Spearman's rho) when the probability distribution for the considered time series is asymmetric. Therefore, we also employed the skewness test to observe whether the probability distribution of the last 125 indicator values at the date from 750



to 999 is asymmetric. As a result, we also computed the skewness value for each of the Pearson correlation at the date  $t = 999$ .

If the skewness value equals zero, the observed series is perfectly symmetrical. But a skewness of exactly zero is quite unlikely for real-world data, so we interpreted the skewness value as follows:

- If skewness is between  $-1/2$  and  $1/2$ , the distribution can be called approximately symmetric. For this case, the Pearson correlation is applicable to detect EWSs of impending financial crises.
- Otherwise, the distribution is called skewed, and the Pearson correlation is not considered.

## Threshold

In this study, by using the test (Kendall's tau, Pearson's, or Spearman's rho correlation), if there exist two significant correlation values, which indicate the significant rising trend in the observed indicator (AC1, VAR, or MPS) 1 day before the two corresponding observed financial crises (Dotcom crash and Lehman Brothers bankruptcy) in the US, a threshold was determined by choosing the smaller value of those two values. Let us assume that  $T_1$  and  $T_2$  are these two thresholds for the US market, then  $T = \min\{T_1, T_2\}$  is the threshold covering all available dates. The minimum value was chosen because it corresponded to the most extreme cases, which provided us with the longest period of significant rising trends in the correlation time series (AC1, VAR, and MPS) before the Dotcom crash and Lehman Brothers bankruptcy.

After computing the threshold, all available dates were covered, and the likelihood of an indicative period of significant rising trends, which lay above the threshold, was determined. The latter was made following the range of time at which the indicative trend could be observed. Finally, all the recorded events (periods with significant rising trends and breakpoints or without that signal) were classified either as EWSs, false alarms (FAs), false negatives (FNs), or true negatives (TNs) as described in the later section.

## Classification

In our practice, we classified any recorded event (a period with significant rising trends and breakpoints or without that signal) as a EWS, an FA, an FN, or a TN and is as follows:

- If there is a nearest financial crisis that happened within a period of continuous significant rising trends with breakpoints, which lies above the threshold, we consider the event in this period as an EWS.
- If there is no nearest financial crisis that happened within a period of continuous significant rising trends with breakpoints, which lies above the threshold, the event on this period is considered as an FA.
- If there is a financial crisis that happened within a period without the observed signal, which lies below the threshold, we consider the event in this period as an FN.
- If there is no nearest financial crisis within a period without the observed signal, which lies below the threshold, the event in this period is considered as a TN.

**TABLE 1** | The classification matrix used in this study.

	Does there exist a nearest financial crisis within the duration?	
	Exist	None
Has the correlation duration exceeded the threshold?	Yes EWS	FA
	No FN	TN

## Evaluation Measures

For evaluation measures, the classification matrix used in our study is as given in **Table 1**. To evaluate the method's performance, we used two classification scores (in percentage), namely probability of successful anticipation (PSA) and probability of erroneous anticipation (PEA). These two scores are defined as follows: let A, B, C, and D be the total number of EWS, FA, FN, and TN, respectively, then

$$PSA = ((A + D)/(A + B + C + D)) \times 100\% \quad (7)$$

$$PEA = ((B + C)/(A + B + C + D)) \times 100\% \quad (8)$$

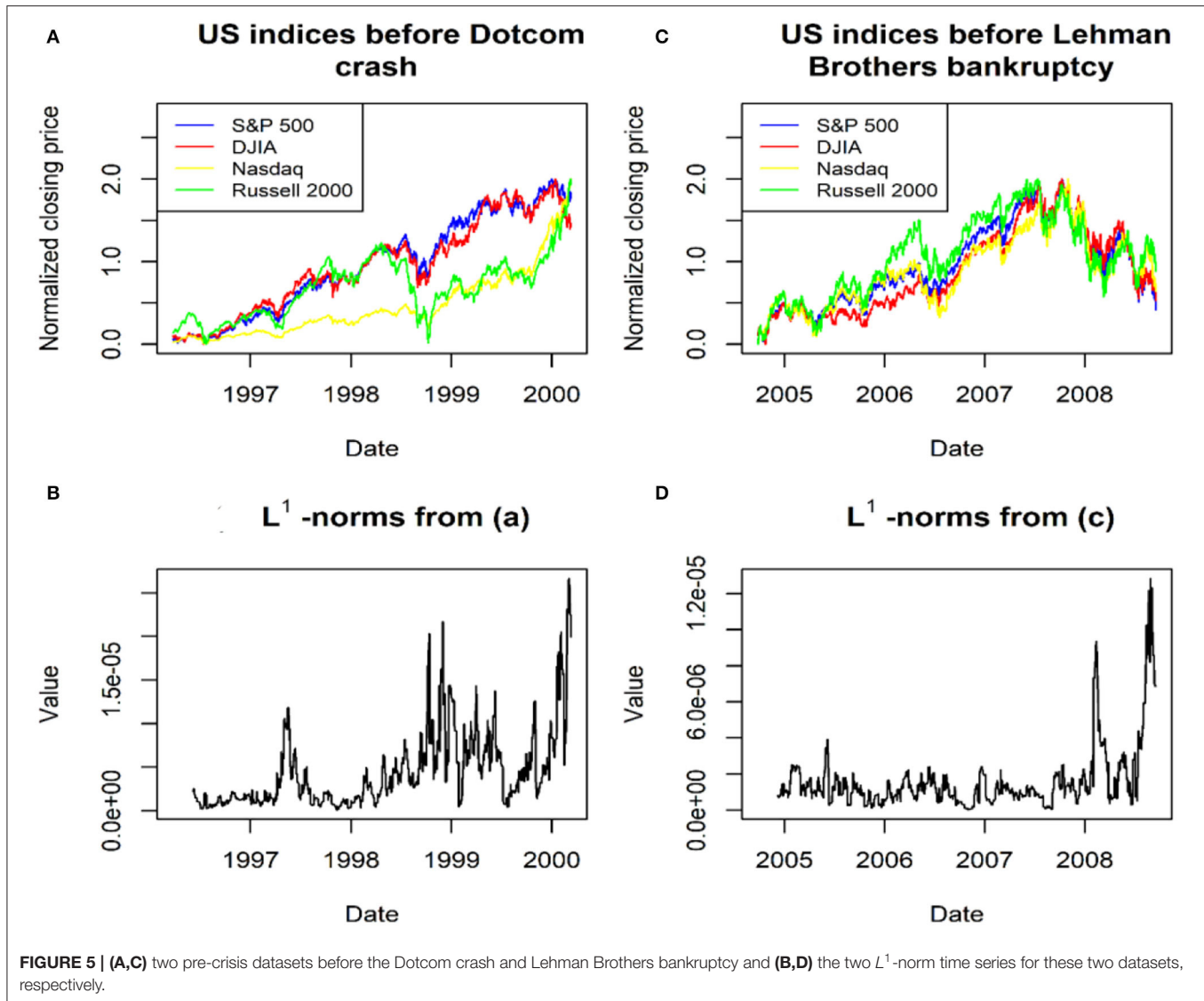
To evaluate, a method that scores the highest PSA and the lowest PEA is considered the best EWS detection tool.

## RESULT

Two pre-crisis datasets were analyzed, whereby each of those datasets had a length of 1,000 days before a financial crisis. In the first row of **Figure 5**, the two pre-crisis datasets before the Dotcom crash and Lehman Brothers bankruptcy are presented, respectively. For all indices in each pre-crisis dataset, we computed all its corresponding daily log-return time series and then combined all the computed daily log returns to build a high-dimensional time series. Furthermore, the daily sliding window of length 50 was applied to obtain PCDs. Then, we applied PH on each PCD to obtain a corresponding  $L^1$ -norm value. As a result, we acquired a  $L^1$ -norm time series. All obtained  $L^1$ -norm time series for every pre-crisis dataset is illustrated in the second row of **Figure 5**.

**Figure 5** demonstrates that the  $L^1$ -norm time series exhibits strong growth toward a primary peak prior to the Dotcom crash and Lehman Brothers bankruptcy in the US market. These results are consistent with Gidea and Katz [25] and Ismail et al. [26]. As stock indices become increasingly volatile when moving closer to a financial crisis, more peaks appear in its corresponding persistent landscape. The latter gives growth in the  $L^1$ -norm values prior to the Dotcom crash and Lehman Brothers bankruptcy [26].

Furthermore, the daily sliding window with the length of 500 was applied to each  $L^1$ -norm time series as shown in **Figure 5**, and the corresponding time series of three CSD indicators, including AC1, VAR, and MPS at low frequencies were obtained. Furthermore, for each computed CSD indicator's time series, we applied the daily sliding window with a length of 250 and



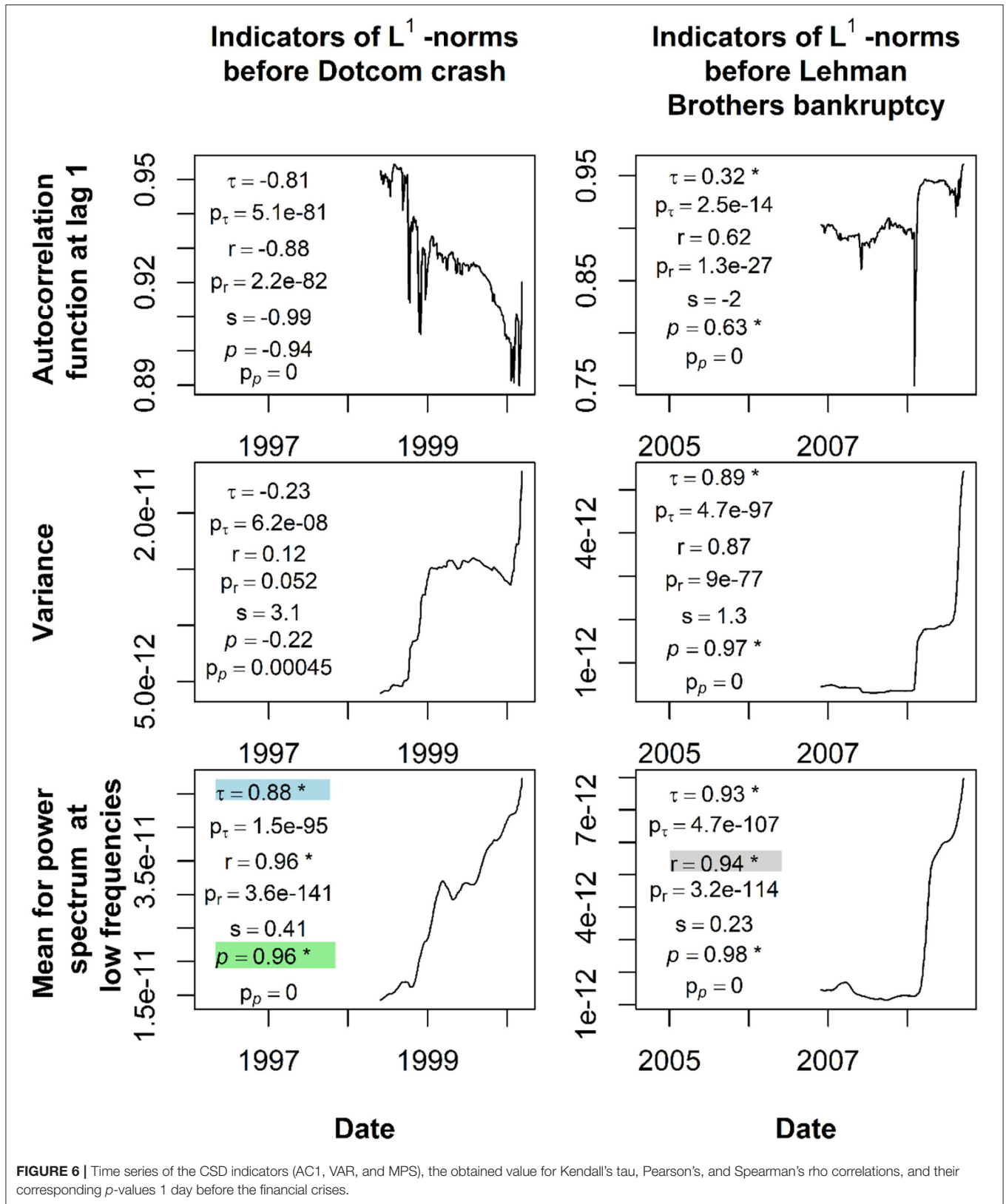
**FIGURE 5 | (A,C)** two pre-crisis datasets before the Dotcom crash and Lehman Brothers bankruptcy and **(B,D)** the two  $L^1$ -norm time series for these two datasets, respectively.

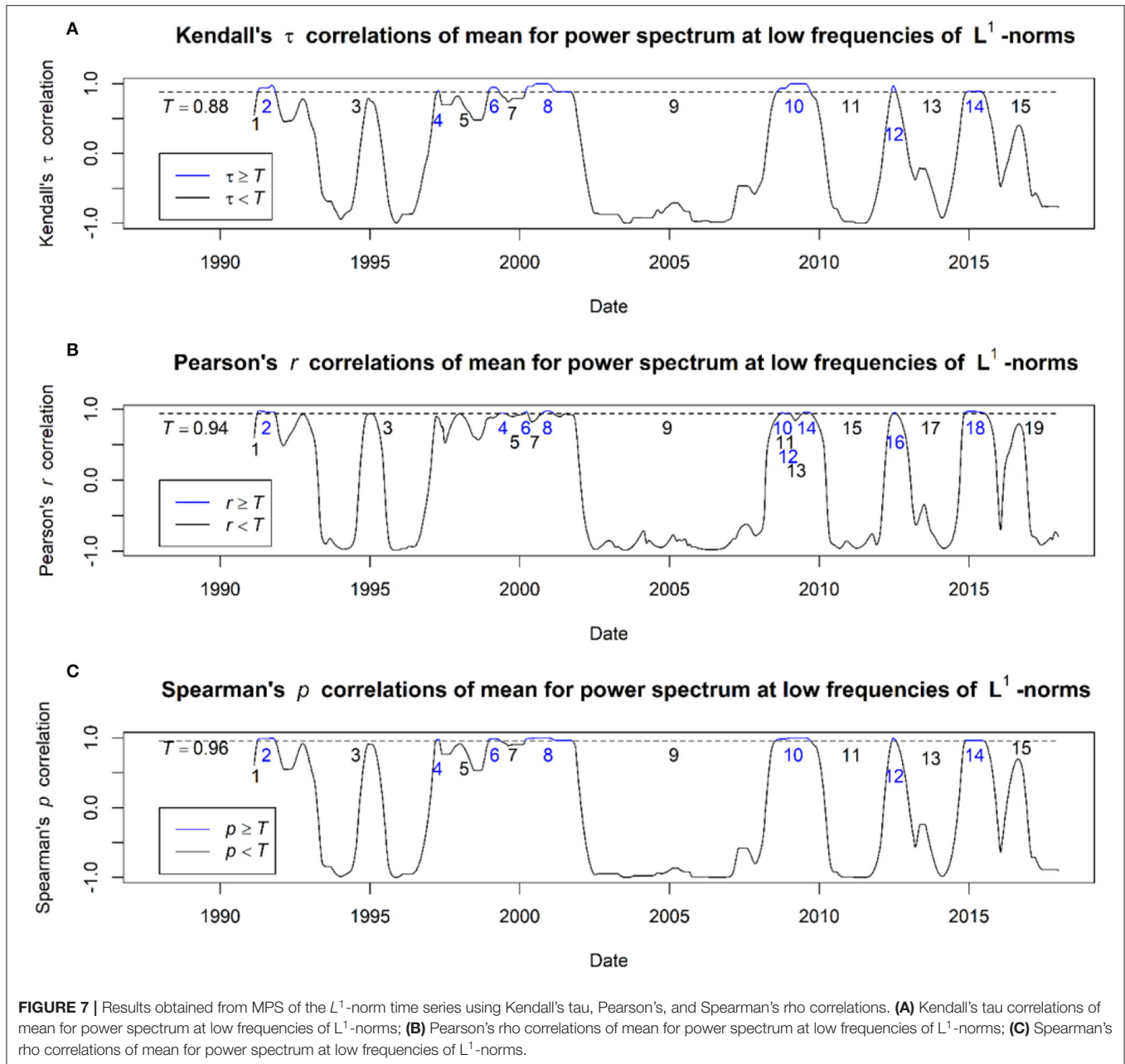
computed the time series of three tests: Kendall's tau, Pearson's, and Spearman's rho correlations. In addition, we also employed a significant test to compute  $p$ -values for each correlation test 1 day before the Dotcom crash and Lehman Brothers bankruptcy. **Figure 6** shows the time series of these CSD indicators (AC1, VAR, and MPS) and presented the obtained value for Kendall's tau, Pearson's, and Spearman's rho correlations before the financial crises. Moreover, in **Figure 6**, corresponding  $p$ -values for Kendall's tau, Pearson's, and Spearman's rho correlations and skewness measure (for the Pearson correlation only) are also mentioned. Kendall's tau correlation, Pearson's correlation, Spearman's rho correlation,  $p$ -values for Kendall's tau correlation,  $p$ -values for Pearson's correlation,  $p$ -values for Spearman's rho correlation, and skewness values in **Figure 6** are symbolized with  $\tau$ ,  $r$ ,  $\rho$ ,  $p_\tau$ ,  $p_r$ ,  $p_\rho$ , and  $s$ , respectively.

Based on the results presented in **Figure 6**, for any CSD indicator of the  $L^1$ -norm time series, which consistently exhibits a significant rising trend (indicated by the correlation values,

the corresponding  $p$ -values, and skewness) before the observed financial crises, we considered such method as a EWS detection tool. It is clearly shown that only the MPS time series of the  $L^1$ -norm time series fulfill this condition, and its significant rising trend can be indicated by Kendall tau, Pearson, and Spearman rho correlation. These because Kendall tau and Spearman rho provide positive values and their corresponding  $p$ -values also less than 0.05. For Pearson, this test obtains a positive value, its corresponding  $p$ -values is less than 0.05, and its skewness values lie in between  $-1/2$  and  $1/2$  indicating that the data has an approximately symmetrical distribution. For the others, they did not achieve the above-mentioned characteristics, therefore not considered as potential EWS detection tools.

In **Figure 6**, the highlighted values from different correlation tests (Kendall's tau, Pearson's, and Spearman's rho) are the chosen threshold to predict the whole available date. These values are 0.88, 0.94, and 0.96 for Kendall's tau, Pearson's, and Spearman's rho correlations, respectively. Therefore, we have





three predictions performed using three different correlations, and each correlation is obtained from the MPS time series of the  $L^1$ -norm time series. For each prediction, all individual events (periods of continuous significant rising trends or otherwise) are recorded, and these events are then classified whether they are EWSs, FAs, FNs, or TNs.

Results obtained from MPS of the  $L^1$ -norm time series using Kendall's tau, Pearson's, and Spearman's rho correlations, respectively, are illustrated in **Figure 7**. All periods of continuous significant rising trends in **Figure 7** are indicated by blue numbers and lines, which lie above the threshold. In contrast, correlation values that lie below the threshold are indicated

with black numbers and lines. Details on these recorded events and their corresponding classification as shown in **Figure 7** for Kendall's tau, Pearson's, and Spearman's rho correlations are provided in tables of **Supplementary Material File**.

By using MPS of the  $L^1$ -norm time series, our results reported that Spearman's rho correlation test obtained the total number of EWS, FA, FN, TN, PSA, and PEA of 4, 7, 7, 8, 46.15, and 53.85%, respectively. Furthermore, Kendall's tau correlation test provided the total number of EWS, FA, FN, TN, PSA, and PEA of 3, 8, 7, 8, 42.31, and 57.69%, respectively. In addition, the Pearson correlation test achieved the total number of EWS, FA, FN, TN, PSA, and PEA of 2, 9, 9, 10, 40, and 60%, respectively.

**TABLE 2** | A summary for all obtained results.

Evaluation measure	Kendall's tau correlation test	Pearson's correlation test	Spearman's rho correlation test
EWS	3	2	<b>4</b>
FA	8	9	<b>7</b>
FN	<b>7</b>	9	<b>7</b>
TN	<b>8</b>	10	<b>8</b>
PSA (%)	42.31	40	<b>46.15</b>
PEA (%)	57.69	60	<b>53.85</b>

All these results obtained are summarized in **Table 2**, where all the bolded values represent the obtained highest scores for the corresponding evaluation measure.

From **Table 2**, based on PSA and PEA, Spearman's rho and Kendall's tau correlations are showed to obtain a better result than the Pearson correlation. This shows that non-parametric rank correlation, which computes statistical associations based on the ranks of the data like Spearman's rho and Kendall's tau correlations is better than the Pearson correlation, which measures the degree of the linear relationship between related variables. However, Spearman's rho correlation, which measures the degree of association between two variables is reported and has a better result as compared to Kendall's tau correlation, which measures the strength of dependence between two variables. In addition, our result in **Table 2** clearly shows that Spearman's rho correlation outperforms other correlation tests (Kendall's tau and Pearson's correlations) with the highest score of every evaluation measure. Therefore, by using MPS of the  $L^1$ -norm time series, our results concluded that Spearman's rho correlation can detect EWSs better than Kendall's tau and Pearson's correlations.

## CONCLUSION

In this study, PH and CSD were proposed to detect EWSs of major financial crashes in the US market. Preliminarily, two financial crises: Dotcom Crash and Lehman Brothers Bankruptcy that happened in the US market were examined. By using PH,  $L^1$ -norm time series was obtained for each financial crisis and used to compute the CSD indicators: AC1, VAR, and MPS.

By using three different correlation tests, Kendall's tau, Pearson's, and Spearman's rho, the rising trend in these indicators is observed prior to the financial crises. Furthermore, this study applied significance and skewness tests to determine whether the rising trends in the indicators (AC1, VAR, or MPS) are statistically significant. This test aims to conclude that the rising trend does not happen by chance. Subsequently, a threshold is

## REFERENCES

1. D.S. Bates. Jumps and stochastic volatility: exchange rate processes implicit in deutsche mark options. *Rev Financ Stud.* (2015) 9:69–107. doi: 10.1093/rfs/9.1.69

determined to predict the whole date, and then the classification performance of our method is evaluated by using PSA and PEA.

Our result shows that the  $L^1$ -norm time series exhibits a strong growth before Dotcom Crash and Lehman Brothers Bankruptcy. This portrays that the  $L^1$ -norm time series has the potential to be a representative to detect EWSs of major financial crashes in the US market. Moreover, MPS from the  $L^1$ -norm time series is significantly rising before these two financial crises. It has also been demonstrated that all correlation tests, Kendall's tau, Pearson's, and Spearman's rho, can indicate the observed significant rising trend.

Overall, based on PSA and PEA, our results revealed that Spearman's rho correlation predicts the US market better than Kendall's tau and Pearson's correlations. Therefore, this study demonstrates that PH *via* its  $L^1$ -norm time series with MPS and Spearman's rho correlation offers a new potential EWS detection tool for financial crises in the US market. For future studies, we plan to examine and refine this method of PH to provide more reliable EWSs for upcoming financial crises.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: <https://finance.yahoo.com>

## AUTHOR CONTRIBUTIONS

MSI: conceptualization, methodology, software, validation, writing—original draft, and writing—review and editing. MM: conceptualization, supervision, validation, writing—review and editing, and funding acquisition. MI: supervision, data curation, validation, writing—review and editing, funding acquisition. FA: supervision, validation, writing—review and editing. All authors contributed to the article and approved the submitted version.

## ACKNOWLEDGMENTS

We would like to express our gratitude to the Universiti Kebangsaan Malaysia and the Ministry of Higher Education Malaysia for their financial support *via* the provision of two grants with numbers: FRGS/1/2019/STG06/UKM/01/3 and GUP-2020-032.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fams.2022.940133/full#supplementary-material>

2. Gopalakrishnan EA, Sharma Y, John T, Dutta PS, Sujith RI. Early warning signals for critical transitions in a thermoacoustic system. *Sci Rep.* (2016) 6:35310. doi: 10.1038/srep35310
3. Gidea M, Goldsmith D, Katz Y, Roldan P, Shmalo Y. Topological recognition of critical transitions in time series of cryptocurrencies.

- Physica A Stat Mech Appl.* (2020) 548:123843. doi: 10.1016/j.physa.2019.123843
4. Virtanen T, Tölö E, Virén M, Taipalus K. Can bubble theory foresee banking crises? *J Financial Stab.* (2018) 36:66–81. doi: 10.1016/j.jfs.2018.02.008
  5. Sornette D, Cauwels P. Financial bubbles: mechanisms and diagnostics. *Swiss Finance Institute Research Paper.* (2014). doi: 10.2139/ssrn.2423790
  6. H.J. Edison. Do indicators of financial crises work? An evaluation of an early warning system. *Int J Finance Econ.* (2003) 8:11–53. doi: 10.1002/ijfe.197
  7. Hubrich K, Tetlow RJ. Financial stress and economic dynamics: The transmission of crises. *J Monet Econ.* (2015) 70:100–15. doi: 10.1016/j.jmoneco.2014.09.005
  8. Quax R, Kandhai D, Sloot P. Information dissipation as an early-warning signal for the Lehman Brothers collapse in financial time series. *Sci Rep.* (2013) 3:1–7. doi: 10.1038/srep01898
  9. Gatfaoui H, Peretti PDe. Flickering in information spreading precedes critical transitions in financial markets. *Sci Rep.* (2019) 9:1–11. doi: 10.1038/s41598-019-42223-9
  10. Squartini T, Van Lelyveld I, Garlaschelli D. Early-warning signals of topological collapse in interbank networks. *Sci Rep.* (2013) 3:1–9. doi: 10.1038/srep03357
  11. Saracco F, Clemente RDi, Gabrielli A, Squartini T. Detecting early signs of the 2007–2008 crisis in the world trade. *Sci Rep.* (2016) 6:1–11. doi: 10.1038/srep30286
  12. Almog A, Shmueli E. Structural entropy: monitoring correlation-based networks over time with application to financial markets. *Sci Rep.* (2019) 9:1–13. doi: 10.1038/s41598-019-47210-8
  13. Flood MD, Lemieux VL, Varga M, Wong BW. The application of visual analytics to financial stability monitoring. *J Financial Stab.* (2016) 27:180–97. doi: 10.1016/j.jfs.2016.01.006
  14. Battiston S, Farmer JD, Flache A, Garlaschelli D, Haldane AG, Heesterbeek H, et al. Complexity theory and financial regulation. *Science.* (2016) 351:818–9. doi: 10.1126/science.aad0299
  15. D. Sornette, *Why Stock Markets Crash: Critical Events in Complex Financial Systems.* Princeton: University Press (2017). doi: 10.23943/princeton/9780691175959.001.0001
  16. Van Nes EH, Scheffer M. Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift. *Am Nat.* (2007) 169:738–47. doi: 10.1086/516845
  17. Dakos V, Carpenter SR, Brock WA, Ellison AM, Guttal V, Ives AR, et al. Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLoS ONE.* (2012) 7:e41010. doi: 10.1371/journal.pone.0041010
  18. Veraart AJ, Faassen EJ, Dakos V, van Nes EH, Lürling M, Scheffer M. Recovery rates reflect distance to a tipping point in a living system. *Nature.* (2012) 481:357–9. doi: 10.1038/nature10723
  19. Tan JPL, Cheong SSA. Critical slowing down associated with regime shifts in the US housing market. *Eur Phys J B.* (2014) 87:1–10. doi: 10.1140/epjb/e2014-41038-1
  20. Tan J, Cheong SA. The regime shift associated with the 2004–2008 US housing market bubble. *PLoS ONE.* (2016) 11:e0162140. doi: 10.1371/journal.pone.0162140
  21. Guttal V, Raghavendra S, Goel N, Hoarau Q. Lack of critical slowing down suggests that financial meltdowns are not critical transitions, yet rising variability could signal systemic risk. *PLoS ONE.* (2016) 11:e0144198. doi: 10.1371/journal.pone.0144198
  22. Wen H, Ciamarra MB, Cheong SA. How one might miss early warning signals of critical transitions in time series data: A systematic study of two major currency pairs. *PLoS ONE.* (2018) 13:e0191439. doi: 10.1371/journal.pone.0191439
  23. Diks C, Hommes C, Wang J. Critical slowing down as an early warning signal for financial crises? *Empir Econ.* (2019) 57:1201–28. doi: 10.1007/s00181-018-1527-3
  24. Otter N, Porter MA, Tillmann U, Grindrod P, Harrington HA, A. roadmap for the computation of persistent homology. *EPJ Data Science.* (2017) 6:1–38. doi: 10.1140/epjds/s13688-017-0109-5
  25. Gidea M, Katz Y. Topological data analysis of financial time series: Landscapes of crashes. *Phys A Stat Mech Appl.* (2018) 491:820–34. doi: 10.1016/j.physa.2017.09.028
  26. Ismail MS, Noorani MSM, Ismail M, Razak FA, Alias MA. Early warning signals of financial crises using persistent homology. *Phys A Stat Mech Appl.* (2022) 586:126459. doi: 10.1016/j.physa.2021.126459
  27. Aromi LL, Katz YA, Vives J. Topological features of multivariate distributions: Dependency on the covariance matrix. *Commun Nonlinear Sci Numer Simul.* (2021) 103:105996. doi: 10.1016/j.cnsns.2021.105996
  28. M. Gidea, Topological Data Analysis of Critical Transitions in Financial Networks. In: Shmueli E, Barzel B, Puzis R, (Eds.), *3rd International Winter School and Conference on Network Science.* Cham: Springer International Publishing (2017). pp. 47–59. doi: 10.1007/978-3-319-55471-6\_5
  29. Guo H, Xia S, An Q, Zhang X, Sun W, Zhao X. Empirical study of financial crises based on topological data analysis. *Phys A Stat Mech Appl.* (2020) 558:124956. doi: 10.1016/j.physa.2020.124956
  30. Guo H, Zhao X, Yu H, Zhang X. Analysis of global stock markets' connections with emphasis on the impact of COVID-19. *Phys A Stat Mech Appl.* (2021) 569:125774. doi: 10.1016/j.physa.2021.125774
  31. Yen PT, Xia K, Cheong SA. Understanding changes in the topology and geometry of financial market correlations during a market crash. *Entropy.* (2021) 23:1211. doi: 10.3390/e23091211
  32. Yen PT, Cheong SA. Using Topological Data Analysis (TDA) and persistent homology to analyze the stock markets in Singapore and Taiwan. *Front Physics.* (2021) 9:20. doi: 10.3389/fphy.2021.572216
  33. Kim W, Kim Y-J, Lee G, Kook W. Investigation of flash crash via topological data analysis. *Topol Appl.* (2021) 301:107523. doi: 10.1016/j.topol.2020.107523
  34. Nguyen NKK, Bui M. Detecting anomalies in the dynamics of a market index with topological data analysis. *Int J Syst Innovation.* (2021) 6:37–50.
  35. Katz YA, Biem A. Time-resolved topological data analysis of market instabilities. *Phys A Stat Mech Appl.* (2021) 571:125816. doi: 10.1016/j.physa.2021.125816
  36. Ismail MS, Hussain SI, Noorani MSM. Detecting early warning signals of major financial crashes in bitcoin using persistent homology. *IEEE Access.* (2020) 8:202042–57. doi: 10.1109/ACCESS.2020.3036370
  37. Goel A, Pasricha P, Mehra A. Topological data analysis in investment decisions. *Expert Syst Appl.* (2020) 147:113222. doi: 10.1016/j.eswa.2020.113222
  38. Baitinger E, Flegel S. The better turbulence index? Forecasting adverse financial markets regimes with persistent homology. *Financ Mark Portf Manag.* (2021) 35:277–308. doi: 10.1007/s11408-020-00377-x
  39. Majumdar S, Laha AK. Clustering and classification of time series using topological data analysis with applications to finance. *Expert Syst Appl.* (2020) 162:113868. doi: 10.1016/j.eswa.2020.113868
  40. Ismail MS, Md Noorani MS, Ismail M, Abdul Razak F, Alias MA. Predicting next day direction of stock price movement using machine learning methods with persistent homology: Evidence from Kuala Lumpur Stock Exchange. *Appl Soft Comput.* (2020) 93:106422. doi: 10.1016/j.asoc.2020.106422
  41. E. Baitinger, and S. Flegel, *New Concepts in Financial Forecasting: Network-Based Information, Topological Data Analysis and their Combination.* (2021). doi: 10.2139/ssrn.3962148
  42. V. Robins, Towards computing homology from finite approximations. *J Adv Stud Topol.* (1999) 24: 503–32. Available online at: <http://topology.nipissingu.ca/tp/reprints/v24/tp24222.pdf>
  43. Edelsbrunner H, Letscher D, Zomorodian A. Topological persistence and simplification. In: *Proceedings 41st Annual Symposium on Foundations of Computer Science.* New York, NY: IEEE (2000). p. 454–63.
  44. Zomorodian A, Carlsson G. Computing persistent homology. *Discrete Comput Geom.* (2005) 33:249–74. doi: 10.1007/s00454-004-1146-y
  45. G. Carlsson. Topology and data. *Bull New Ser Am Math Soc.* (2009) 46:255–308. doi: 10.1090/S0273-0979-09-01249-X
  46. Edelsbrunner H, Harer JL, *Computational Topology: An Introduction.* American Mathematical Society (2022).
  47. Bubenik P, Dlotko P, A. persistence landscapes toolbox for topological statistics. *J Symb Comput.* (2017) 78:91–114. doi: 10.1016/j.jsc.2016.03.009

48. Bubenik P. Statistical topological data analysis using persistence landscapes. *J Mach Learn Res.* (2015) 16:77–102. doi: 10.48550/arXiv.1207.6437

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor MM declared a past collaboration with the authors.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of

the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

*Copyright © 2022 Ismail, Md Noorani, Ismail and Abdul Razak. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*