entropy

MDPI

Article

Crash Diagnosis and Price Rebound Prediction in NYSE Composite Index Based on Visibility Graph and Time-Evolving Stock Correlation Network

Yuxuan Xiu [1,2], Guanying Wang [3] and Wai Kin Victor Chan [1,2,*]

[1] Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China; yuxuanxiu@gmail.com
[2] Tsinghua-Berkeley Shenzhen Institute, Tsinghua University, Shenzhen 518055, China
[3] College of Management and Economics, Tianjin University, Tianjin 300072, China; wangguanyingnk@163.com
* Correspondence: chanw@sz.tsinghua.edu.cn

Abstract: This study proposes a framework to diagnose stock market crashes and predict the subsequent price rebounds. Based on the observation of anomalous changes in stock correlation networks during market crashes, we extend the log-periodic power-law model with a metric that is proposed to measure network anomalies. To calculate this metric, we design a prediction-guided anomaly detection algorithm based on the extreme value theory. Finally, we proposed a hybrid indicator to predict price rebounds of the stock index by combining the network anomaly metric and the visibility graph-based log-periodic power-law model. Experiments are conducted based on the New York Stock Exchange Composite Index from 4 January 1991 to 7 May 2021. It is shown that our proposed method outperforms the benchmark log-periodic power-law model on detecting the 12 major crashes and predicting the subsequent price rebounds by reducing the false alarm rate. This study sheds light on combining stock network analysis and financial time series modeling and highlights that anomalous changes of a stock network can be important criteria for detecting crashes and predicting recoveries of the stock market.

Keywords: stock market; crash; rebound; log-periodic power law; visibility graph; stock correlation network; anomaly detection; extreme value theory

check for updates

Citation: Xiu, Y.; Wang, G.; Chan, W.K.V. Crash Diagnosis and Price Rebound Prediction in NYSE Composite Index Based on Visibility Graph and Time-Evolving Stock Correlation Network. *Entropy* **2021**, *23*, 1612. https://doi.org/10.3390/e23121612

Academic Editors: Zhi-Qiang Jiang and Wei-Xing Zhou

Received: 20 October 2021
Accepted: 28 November 2021
Published: 30 November 2021

1. Introduction

A stock market crash is one of the most significant systemic risks of the modern financial system, causing significant losses for investors. A recent example is the March 2020 stock market crash triggered by COVID-19 [1], during which the New York Stock Exchange (NYSE) Composite Index plunged roughly 35% within a month. Meanwhile, rebounds of the stock market after crashes usually signal the recovery of investors' confidence or the taking effect of bailout policies. Therefore, it is critical for investors and policy makers to detect stock market crashes and predict price rebounds.

In the past decades, different methods have been proposed to diagnose the stock market crashes and predict rebounds. One of the most representative methods is the log-periodic power-law (LPPL) model [2]. Originally, the LPPL model was proposed to predict the bursting point of financial bubbles. Yan et al. [3] adopt the LPPL model to study stock market crashes by considering them as the "mirror images" of financial bubbles, which are also known as "negative bubbles". The fundamental insight of modeling financial crashes with the LPPL model is to capture a particular pattern of the price time series, which can be described as "the faster-than-exponential decline accompanied by accelerating oscillations" [4]. The LPPL model and its extensions are successfully applied in diagnosing negative bubbles of many types of assets, such as crude oil [5] and cryptocurrency [6]. Despite previous achievements, all the existing LPPL-based models only focus on the price time series itself. However, it may not be sufficient to use only the price time series of the stock index when diagnosing stock market crashes. The price time series of the stock index can describe the overall fluctuation of the stock market, but it ignores the complex interactions among multiple assets.

Network analysis provides a novel tool for characterizing the complex interactions and co-movements in the stock market [7–9]. In fact, it has been discovered that stock market crashes and recoveries are accompanied by drastic changes in the topological structure of stock correlation networks, which can be captured by some network statistics such as assortativity [10], modularity [11], von Neumann entropy [12], the structural entropy [13,14] and the graph motif entropy [15]. Although the above literature qualitatively analyzes the dynamical changes of stock correlation networks during stock market crashes and rebounds, there is still a lack of work that applies such phenomenon to quantitatively diagnose stock market crashes and predict price rebounds.

This paper contributes to the literature by incorporating the analysis of a time-evolving stock correlation network into the LPPL model. The contribution of this paper has three folds. First, extending the LPPL model, we propose to characterize stock market crashes by two distinct characteristics: (1) faster-than-exponential decline of the stock index price; (2) abnormal changes in the market structure (i.e., the topology of the stock correlation network). Second, we design a prediction-guided anomaly detection method based on the extreme value theory (EVT) in order to define and detect "anomalies" for Characteristic (2). The intuition is that a properly trained predictor can forecast most of the "normal" situations well, with significant deviations when "abnormal" situations occur. The "normal" and "abnormal" deviations are distinguished based on EVT. Third, we propose a framework by combining Characteristic (1) and (2), where Characteristic (1) is captured by a visibility graph (VG)-based representation of the LPPL model proposed by Yan et al. [16]. A hybrid rebound indicator is calculated, which is a linear combination of Yan's VG-based indicator and the anomalies of the stock correlation network. Specifically, we normalize the anomalies to 0 to 1 based on EVT and treat them as confidence levels (i.e., weights) of the VG-based rebound indicator.

Experiments are conducted based on the data of the New York Stock Exchange (NYSE) Composite Index. We predict the subsequent price rebounds of the 12 major crashes in the U.S. stock market from 4 January 1991 to 7 May 2021. Experimental results demonstrate that our proposed prediction-guided anomaly detection algorithm is well capable of identifying abnormal changes in stock correlation networks during market crashes and recoveries. Furthermore, our proposed hybrid indicator outperforms Yan's VG-based indicator by reducing the false alarm rate. These findings imply that incorporating the analysis of the time-evolving stock correlation network into the modeling of the stock index time series is a promising direction for diagnosing and predicting financial markets.

The rest of this paper is organized as follows. Section 2 describes the data. Section 3 introduces our proposed framework for stock market crash diagnosis and rebound prediction. Section 4 presents the experimental results. Section 5 discusses the main findings, implications and limitations of this study. Section 6 concludes our work and provides future research directions.

2. Data Description and Labeling

2.1. Data Description

This paper uses the daily closing price of the NYSE Composite Index, which is collected from the Yahoo! financial database (http://finance.yahoo.com) (accessed on 25 November 2021). The time period is from 2 January 1986 to 7 May 2021, with the data from 2 January 1986 to 3 January 1994 as the training set and the rest of the data as the testing set. As shown in Table 1, there are 15 major crashes in the U.S. stock market, including 3 in the training set and 12 in the testing set. The list of stock market crashes in the U.S. before 2008 is provided in [15], while the crashes after 2008 are manually collected from Wikipedia (https://en.wikipedia.org/w/index.php?title=List_of_stock_market_crashes_and_bear_markets) (accessed on 25 November 2021). Furthermore, we select 199 stocks out of the 347 stocks in the

NYSE dataset [11], whose data are available for the entire time period from 1986 to 2021. Their price time series are used to construct the time-evolving stock correlation network.

Table 1. Major stock market crashes in the U.S. from January 1986 to May 2021.

Name	Date
Black Monday	19 October 1987
Friday the 13th mini-crash	13 October 1989
Early 1990s recession	3 July 1990
1997 Asian financial crisis	2 July 1997
Russian financial crisis	17 August 1998
Dot-com bubble	10 March 2000
September 11 attacks	11 September 2001
Stock market downturn of 2002	19 March 2002
Financial crisis of 2007–2008	31 October 2007
2009 Icelandic financial crisis	20 January 2009
European sovereign debt crisis	27 April 2010
August 2011 stock markets fall	1 August 2011
2015–2016 stock market selloff	18 August 2015
2018 cryptocurrency crash	20 September 2018
2020 stock market crash	24 February 2020

2.2. Labeling Rebounds of Price Time Series

The rebounds of the financial market are labeled based on the price time series of the stock index. Following the definition in the existing literature [16,17], we define the rebound as the time point at which a stock index turns from a downtrend to an uptrend after a stock market crash. This paper labels the trend of the stock index based on a recently proposed method [18]. The basic idea is that the price time series is considered to change from an upward trend to a downward trend when the price falls by more than w compared with the local peak. Similarly, when the price rises above w compared with the local trough, the price time series is considered to change from a downward trend to an upward trend. Here, w is a predetermined threshold for the proportion of price increases and decreases. The detailed procedure is illustrated in Appendix A.

For each stock market crash in Table 1, we detect the first turning point when the trend of the price time series changes from a downtrend to an uptrend. This turning point is defined as the time point of the price rebound after the corresponding stock market crashes. Here we choose $w = 0.15$, which is an empirical value proposed along with the trend labeling method [18]. In other words, once the price rises by more than 15% from the local trough after a financial crash, it is regarded as a switch from a downtrend to an uptrend, and the local trough is labeled as a rebound. The labeling results are further manually adjusted according to the historical records. Figure 1 shows the time series of the daily closing price of the NYSE Composite Index, as well as the labeled price rebounds and the corresponding stock market crashes.

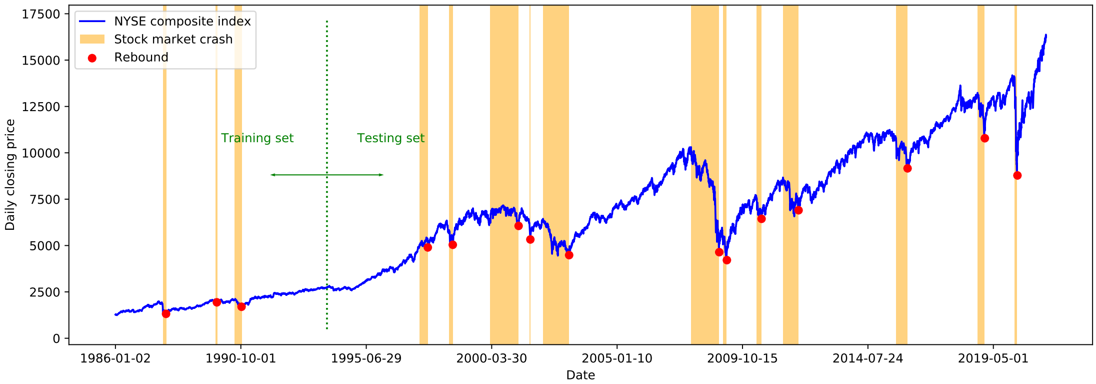


Figure 1. The NYSE Composite Index with labeled crashes and rebounds.

3. Methodology

Our proposed framework predicts market rebounds by observing and quantifying a specific phenomenon, which is a faster-than-exponential decline in the stock index price time series accompanied by anomalous changes in the topology of the correlation network of constituent stocks. Therefore, our proposed framework consists of the following four steps: (1) quantifying faster-than-exponential decline in stock index price; (2) constructing time-evolving stock correlation network; (3) detecting anomalous topological changes of stock correlation networks; (4) calculating a rebound alarm index. In this section, we first introduce our proposed framework and then describe each of these four steps.

3.1. Our Proposed Framework

Our proposed framework is illustrated in detail in Figure 2. We first measure the faster-than-exponential decline of the price time series of the stock index based on Yan's VG-based LPPL model [16] (step *a*). The detailed procedures are described in Section 3.2.

The time series of the stock index can describe the overall fluctuation of the financial market, but it ignores the complex interactions between different assets. In our proposed framework, we exploit information of the market structure by investigating the time-evolving topology of the stock correlation network (step *b* and *c*). In step *b*, we first select the constituents of the stock index, then calculate the matrices of Pearson correlation coefficients based on the logarithmic return of the constituents through a sliding window, and finally extract the most important correlations in the matrix to form a stock correlation network. In step *c*, topological measurements are extracted from the time-evolving stock correlation network. Anomaly detection is performed on the extracted topological measurements to identify the emergence and disappearance of anomalous network topology. Sections 3.3–3.5 describe the implementation of step *b*, *c* and step *d*, respectively.

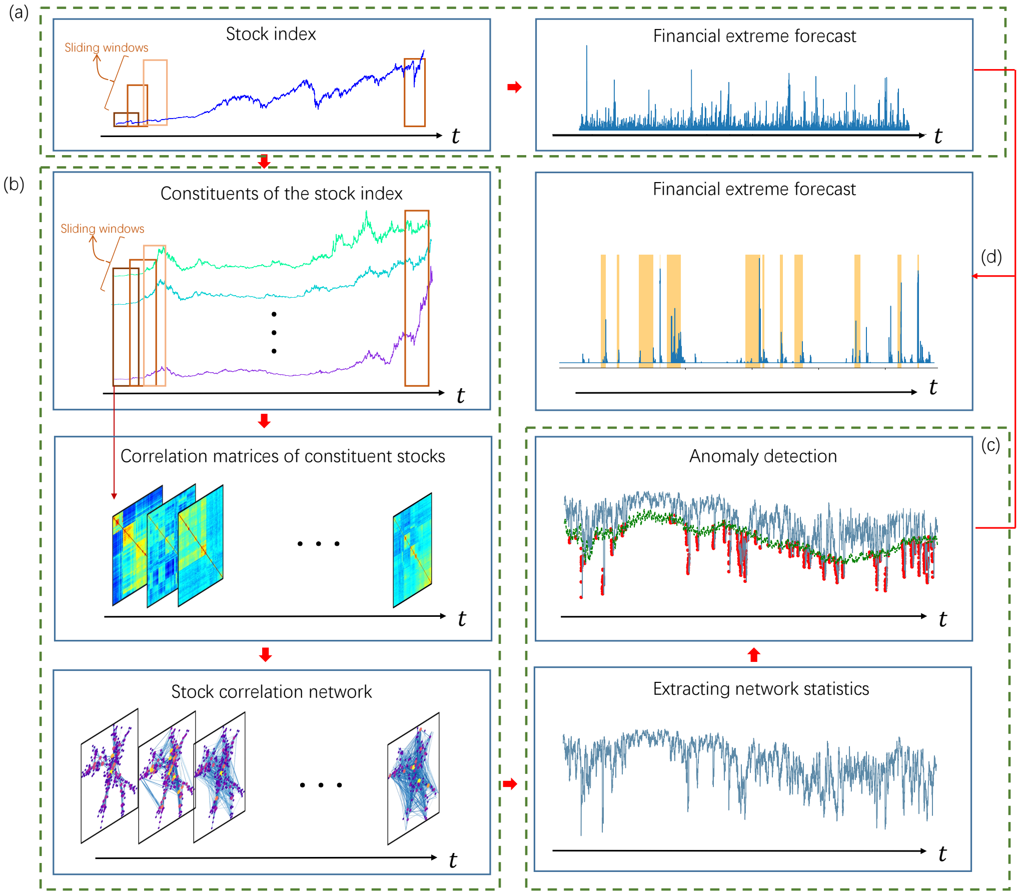

(a)
Stock index
Sliding windows
t
Financial extreme forecast
t
(b)
Constituents of the stock index
Sliding windows
t
Financial extreme forecast
t
(d)
Correlation matrices of constituent stocks
t
(c)
Anomaly detection
t
Stock correlation network
t
Extracting network statistics
t

Figure 2. Our proposed framework.

3.2. Quantifying Faster-Than-Exponential Decline in Stock Index Price

We adopt the method based on the visibility graph [16] to quantitatively describe the faster-than-exponential decline of the stock index prices. This subsection first briefly introduces the basic idea of the visibility graph and then describes the quantification of the faster-than-exponential decline of the stock index price based on the VG.

The visibility graph (VG) is first presented by Lacasa et al. [19] as an algorithm for converting time series into complex networks. Its basic idea is to map each time point in a discrete time series into a node in a network. The edges of the network are established based on the visibility criteria. The detailed procedure of constructing the visibility graph is introduced as follows.

For each pair of data points in the time series noted as (t_i, y_i) and $\left(t_j, y_j\right)$ where $i < j$, the VG algorithm first connects them with a straight line. If all the data points between i and j are below the straight line, (t_i, y_i) and $\left(t_j, y_j\right)$ are considered "visible" to each other, and an edge is created between them in the visibility graph. A more formal expression is: for a time series $[(t_1, y_1), (t_2, y_2), \cdots, (t_N, y_N)]$ of length N, the adjacency matrix W of its visibility graph is defined as

$$\omega_{ij} = \begin{cases} 1, & \text{if } y_k < y_i + \frac{t_k - t_i}{t_j - t_i}\left(y_j - y_i\right), \forall i < k < j \\ 0, & \text{otherwise} \end{cases}, \tag{1}$$

where ω_{ij} denotes the weight of the edge (i, j) in the visibility graph.

Yan et al. [16] first represent the log periodic power law (LPPL) model with the visibility graph. Originally, the LPPL model characterizes the financial bubbles by the faster-than-exponential growth of the stock market prices. Yan et al. construct the visibility graph based on the logarithmic prices of stock indices. Since the exponential growth is a straight line in the logarithmic-linear scale, faster-than-exponential growth is represented as a convex curve. Therefore, as the price time series goes up super-exponentially, the degree of the last node of its VG increases. Furthermore, if two nodes (t_i, y_i) and $\left(t_j, y_j\right)$ are "visible" to each other, the growth rate from t_i to t_j can be roughly considered as "faster-than-exponential".

Since our goal is to characterize the financial crashes, we adopt the absolute invisibility graph, which is exactly the opposite of the visibility graph. Its adjacency matrix $W = \left\{\omega_{ij}\right\}_{i,j=1,\cdots,N}$ is defined as

$$\omega_{ij} = \begin{cases} 1, & \text{if } y_k > y_i + \frac{t_k - t_i}{t_j - t_i}\left(y_j - y_i\right), \forall i < k < j \\ 0, & \text{otherwise} \end{cases}. \tag{2}$$

Similar to the visibility graph, if two nodes (t_i, y_i) and $\left(t_j, y_j\right)$ are "absolutely invisible" to each other, we can roughly conclude a faster-than-exponential decline from t_i to t_j. For each data point (t_i, y_i) in a time series, we obtain $T_{\mathrm{VG}} - 1$ historical data before t_i through a sliding window of length T_{VG} to construct the absolute invisibility graph. The magnitude of the faster-than-exponential decline at time point (t_i, y_i) is defined as

$$I_{\mathrm{VG}}(i) = \frac{1}{T_{\mathrm{VG}}} \sum_{j=i-T_{\mathrm{VG}}}^{i-1} \mathbf{1}_{[y_i < y_j]} \cdot \mathbf{1}_{\left[y_k > y_i + \frac{t_k - t_i}{t_j - t_i}\left(y_j - y_i\right)\right]}, \tag{3}$$

where $\mathbf{1}_{[\cdot]}$ is the indicator function.

3.3. Constructing the Time-Evolving Stock Correlation Network

This paper studies the time-evolving nature of the correlation network, which is formed by the constituents of a stock index. Our basic assumption is that the topology of the stock correlation network changes significantly during financial crashes, and that the recovery of the network topology implies the recovery of the financial market.

The most straightforward way of studying a time-evolving network is to look at snapshots of the network taken at different time points. By stacking the snapshots in temporal order, a time-evolving network is denoted as $\mathcal{G} = (G_1, \ldots, G_T)$. Each $G_t = \{\mathbf{N}_t, \mathbf{E}_t\}$ is a snapshot recorded as time t, where $\mathbf{N}_t$ is the set of nodes and $\mathbf{E}_t$ is the set of edges. Since this paper considers the time-evolving correlations among the same set of stocks, all the snapshots in $\mathcal{G}$ share the same set of nodes, that is, $\mathbf{N}_1 = \mathbf{N}_2 = \cdots = \mathbf{N}_T = \mathbf{N}$.

In this paper, each snapshot G^t is obtained from the correlation matrix C^t by a filtering method. The construction procedure is composed of four steps: (1) dividing time windows; (2) determining the constituents of the stock index; (3) calculating correlation matrices; (4) extracting single layer networks from the correlation matrices. The detailed procedure is illustrated as follows.

3.3.1. Dividing Time Windows

In terms of dividing time windows, three parameters need to be determined: (1) the length T_{tw} of the time window, (2) the step size Δ between two consecutive windows. The choice of time window length T_{tw} is a trade-off between over-smoothed and too noisy data [20]. In order to capture the dynamics of stock market correlations, we need to choose the smallest possible window length. On the other hand, the time window needs to be long enough to avoid the Epps effect [21]. Here, we choose a commonly used empirical value $T_{\text{tw}} = 25$ [15], meaning that each time window contains 25 trading days. To have a continuous tracking of the stock correlations, the step size is set as $\Delta = 1$, shifting the time window 1 day forward at each step.

3.3.2. Determining Constituents of the Stock Index

In this step, we determine the name list of the stocks to track, that is, the set of nodes $\mathbf{N}$. Notice that the constituents of the stock index are constantly adjusted over time. To ensure continuous and stable tracking of the market structure, we select the constituents of the stock index whose historical data are available during the entire experimental period. In this paper, we select $N = 199$ stocks from the NYSE dataset [11,22], whose daily closing prices are available from 2 January 1986 to 7 May 2021.

3.3.3. Calculating Correlation Matrices

For stock i in M^t, its logarithm return at time t is defined as

$$r_i(t) = \ln p_i(t) - \ln p_i(t-1), \tag{4}$$

where $p_i(t)$ is the adjusted closure price of stock i at time t. Then the Pearson correlation coefficient between stock i and stock j is calculated as

$$c_{ij}(t) = \frac{\mathbb{E}[\mathbf{r}_i(t) \odot \mathbf{r}_j(t)] - \mathbb{E}[\mathbf{r}_i(t)]\mathbb{E}[\mathbf{r}_j(t)]}{\sqrt{\left(\mathbb{E}[\mathbf{r}_i^2(t)] - (\mathbb{E}[\mathbf{r}_i(t)])^2\right)\left(\mathbb{E}\left[\mathbf{r}_j^2(t)\right] - (\mathbb{E}[\mathbf{r}_j(t)])^2\right)}}, \tag{5}$$

where $\mathbb{E}(\cdot)$ represents the sample mean, $\odot$ denotes element-wise multiplication of vectors and

$$\mathbf{r}_i(t) = [r_i(t - T_{\text{tw}} + 1), r_i(t - T_{\text{tw}} + 2), \cdots, r_i(t)] \tag{6}$$

is the logarithm return series of stock i within the time window. Thus a $N \times N$ correlation matrix $C(t)$ is obtained.

3.3.4. Extracting Single-Layer Networks from Correlation Matrices

The threshold-based method [23] is applied to extract the strong correlations and construct the network. For the correlation matrix $C(t)$, we choose a threshold $\rho(t)$ and only keep $c_{ij}(t) > \rho(t)$ as the edges of the network G^t. This paper sets $\rho(t)$ as the 85th percentile of all elements in $C(t)$. The connection criterion of network G^t is formally defined as

$$g_{ij}^t = \begin{cases} 1, & \text{if } c_{ij}(t) > \rho(t) \\ 0, & \text{otherwise} \end{cases}, \tag{7}$$

where g_{ij}^t denotes the weight of the edge (i, j) in the network G^t.

3.3.5. Calculating Singular Value Decomposition Entropy

We further measure the topology of each layer M^t by the singular value decomposition (SVD) entropy, which has been applied to the analysis of financial market networks [24–27]. The definition of the SVD entropy is introduced below. It is worth noting that the topological characteristic is not limited to the SVD entropy, any proper network statistic, such as the von Neumann entropy [28] or the graph motif entropy [15], is applicable to our proposed framework.

The SVD entropy is based on the singular value decomposition of the $N \times N$ adjacent matrix A of the network G,

$$A = U\Sigma V^T, \tag{8}$$

where Σ is a diagonal matrix of singular values,

$$\Sigma = \text{diag}(\sigma_1, \cdots, \sigma_N), \tag{9}$$

The SVD entropy is defined as

$$\text{Ent}_t = -\sum_i \bar{\sigma}_i \ln(\bar{\sigma}_i), \tag{10}$$

where $\bar{\sigma}_i$ is the normalized singular value defined as

$$\bar{\sigma}_i = \frac{\sigma_i}{\sum_j \sigma_j}. \tag{11}$$

By calculating the topological characteristics, we transform the time-evolving network into a time series, which is much easier to interpret. Based on the time series of the SVD entropy, we identify and measure anomalous changes in the topology of the stock temporal network in the following subsections.

3.4. Prediction-Guided Anomaly Detection Based on Extreme Value Theory

In this subsection, we detect the anomalous value of each of the topological indicators based on the extreme value theory (EVT) and time series prediction. Our intuition is that we first train advanced time series forecasting algorithms only based on “normal” values. We assume that such predictors are able to capture the “normal” dynamics properly, while being completely unaware of abnormal changes in the dynamics. Based on the commonly adopted normal distribution assumption of the forecasting error, the residuals between the predicted and true value should be small and normally distributed for “normal” data, while the residuals of “abnormal” data should be large and their distribution can be portrayed by EVT. We determine the threshold between the “normal” and “abnormal” residuals based on the training dataset, as well as the parameters of the distribution of the extreme values.

Subsequently, we make predictions and calculate the residual for each day in the testing dataset. If the residual exceeds the threshold, we treat it as an “abnormal” value and calculate its “anomaly score” based on the extreme value distribution. Inspired by [29], our designed method can be divided into an initialization step and an execution step, whose detailed procedures are described below.

Initialization Step

In the initialization step, we first generate the training set that only contains "normal" data. The detailed procedure is described in Appendix B. The training set is generated to train a predictor $F(\cdot)$. Without loss of generality, we assume that the predictor makes single-step predictions based on data from the previous d days as

$$I_t = F(I_{t-d}, \cdots, I_{t-1}) + \epsilon(t), \tag{12}$$

where $\epsilon(t)$ is a normally distributed error term.

To ensure that the predictor only learns the "normal" dynamics of the time series, we generate a training set that only contains the "normal" values. We first split time series $\mathbf{I}$ into different segments based on the date of the financial crashes. For example, the first segment is the "normal period" from 1 January 1986 to 15 October 1987, which is followed by the financial crash from 19 October 1987 to 4 December 1987. We then extract training samples from the "normal periods" using a sliding window of length $d+1$, where the data of the first d days are the inputs to the predictor, and the last datum is the expected output. Finally, we obtain the set of the training inputs $\mathbf{S}_x$ and its corresponding target output set $\mathbf{S}_y$ based on which the predictor $F(\cdot)$ is trained.

After training the predictor, for each day t in the first N days, we make the prediction based on the previous d days as

$$\hat{I}_t = F(I_{t-d}, \cdots, I_{t-1}). \tag{13}$$

Since we are looking for extremely small values, the residual is calculated as $X_t = \hat{I}_t - I_t$. Notice that here we consider both the "normal" periods and the financial crashes, so the set of the residuals should contain the extreme values corresponding to the financial crashes. Therefore, we analyze the tail distribution of the residual based on the EVT.

According to the EVT, the extreme values always follow the same type of distribution, regardless of the initial distribution of the data. It can be regarded as a theorem for the maximum values, which is similar to the central limit theorem for the mean values [30]. A mathematical formulation is provided by the Pickands–Balkema–de Haan theorem [31,32], which can be written as:

$$\bar{F}_t(x) = \mathbb{P}(X - \tau > x \mid X > \tau) \sim \left(1 + \frac{\xi x}{\sigma}\right)^{-\frac{1}{\xi}}. \tag{14}$$

This theorem shows that, for a random variable X_t, the excess over a sufficiently large threshold τ tends to follow a generalized Pareto distribution (GPD) with parameters ξ and σ [29].

A practical implication of EVT is that extreme and non-extreme events follow different distributions because they are often generated by different driving forces [33]. This is the theoretical basis for our use of EVT to identify abnormal changes in network topology. We argue that normal and abnormal residuals are caused by different driving forces, thus we aim to find the outliers that follow GPD. Based on the idea of EVT, we select the most appropriate threshold τ that allows GPD to fit the distribution of $X - \tau$ properly. This means that any value greater than τ can be regarded as an extreme value. Therefore, we consider a residual above the threshold τ as an anomalous value. The detailed procedure for obtaining the optimal value of τ is described in Appendix C.

3.5. Execution Step and Hybrid Rebound Indicator

The detailed procedure of the execution step is shown in Algorithm A4. It is designed to deal with the streaming data. For each day t, we calculate the residual $X_t =$

$F(I_{t-d}, \cdots, I_{t-1}) - I_t$. If the residual exceeds the threshold τ, we raise an alarm while calculating the corresponding alarm index by

$$I_{\text{Alm}}(t) = \text{Alm}(X_t) = 1 - \left[1 + \frac{\xi(X_t - \tau)}{\sigma}\right]^{-\frac{1}{\xi}}. \tag{15}$$

Notice that our alarm index is actually the CDF of the generalized Pareto distribution $F_{\xi,\sigma}(X_t - \tau)$.

Considering the development of the financial market, the internal dynamics of the topological changes of the stock correlation network may also be changing. Therefore, we need to constantly update the predictor based on the new data. Because the predictor is not expected to learn any information about the abnormal changes of the network, we only include the "normal" data into the training set. After every K new samples are added into the training set, the predictor is retrained to ensure that it keeps tracking the latest dynamics of the system.

We finally propose an indicator to characterize the phenomenon of "faster-than-exponential decline in the stock index price accompanied by anomalous changes in the market structure". The alarm index defined in Equation (15) has a value range of $(0, 1)$, and its magnitude indicates the extent to which we believe the network structure is anomalous. Therefore, it can be considered as a "confidence level", which is used as a mask to multiply the indicator defined in Equation (3). Considering that the anomalous changes in the stock correlation network are not necessarily perfectly synchronized with the plunge in the stock index, we make a moving average smoothing of $I_{\text{Alm}}(t)$ with sliding window length T_{Alm}. Notice that T_{Alm} should be a small integer. Based on the intuition that information from two weeks ago is hardly useful for forecasting, here we take $T_{\text{Alm}} < 10$. Since our proposed rebound indicator considers both the time series and the network, it is named as a hybrid indicator, whose formal definition is given as

$$I_{\text{Hybrid}}(t) = I_{\text{VG}}(t) \cdot \left(\frac{1}{T_{\text{Alm}}} \sum_{i=0}^{T_{\text{Alm}}-1} I_{\text{Alm}}(t-i)\right). \tag{16}$$

4. Experimental Results

In this section, we first qualitatively compare our proposed hybrid indicator with the baseline method [16] to give a general idea of how the analysis of the time-evolving stock correlation network helps to improve the prediction performance. We further quantitatively analyze the predictive power of our proposed framework based on the commonly adopted error diagram method. Experimental results demonstrate the effectiveness and robustness of our proposed method.

4.1. Qualitative Observation

In this subsection, we construct three indicators: (1) the VG-based indicator $I_{\text{VG}}(t)$; (2) the alarm index $I_{\text{Alm}}(t)$; (3) the hybrid indicator $I_{Hybrid}(t)$. The look-back scope of $I_{\text{VG}}(t)$ is set as $T_{\text{VG}} = 262$, which is the same as the original paper [16]. The prediction algorithm for calculating $I_{\text{Alm}}(t)$ is the Prophet forecasting model [34], whose open-source implementation is available at https://github.com/facebook/prophet (accessed on 25 November 2021). We use the data from 2 January 1986 to 3 January 1994 as the training set, and use the rest of the data as the testing set. The free parameter T_{Alm} is set as $T_{\text{Alm}} = 4$.

Figure 3 demonstrates the result of the prediction-guided anomaly detection procedure. The solid blue line in the figure indicates the SVD entropy of the stock correlation network for each day. The green dashed line indicates the alarm thresholds obtained based on the one-day ahead prediction, and each red dot indicates that an alarm is issued on that day. An alarm for the financial crisis will be raised on day t, if the SVD entropy on day t falls below the predicted alarm threshold on that day.

Figure 4 shows the price time series of the NYSE Composite Index as well as the three indicators. For the alarm index $I_{\text{Alm}}(t) \in (0, 1)$, a higher value indicates the higher confidence that the network structure on day t is anomalous. Thus, we can see that most of the financial crashes are accompanied by an anomalous change in the topology of the stock correlation network. However, an abnormal network topology does not necessarily imply the occurrence of a financial crisis. Similarly, a high $I_{\mathrm{Alm}}(t)$ usually implies that the stock index is close to a local trough, but there also exists a large number of false alarms. Our approach effectively suppresses the false alarms by considering both the faster-than-exponential decline in the stock index (i.e., I_{VG}) and the anomalous changes in the topology of the stock correlation network (i.e., I_{Alm}).

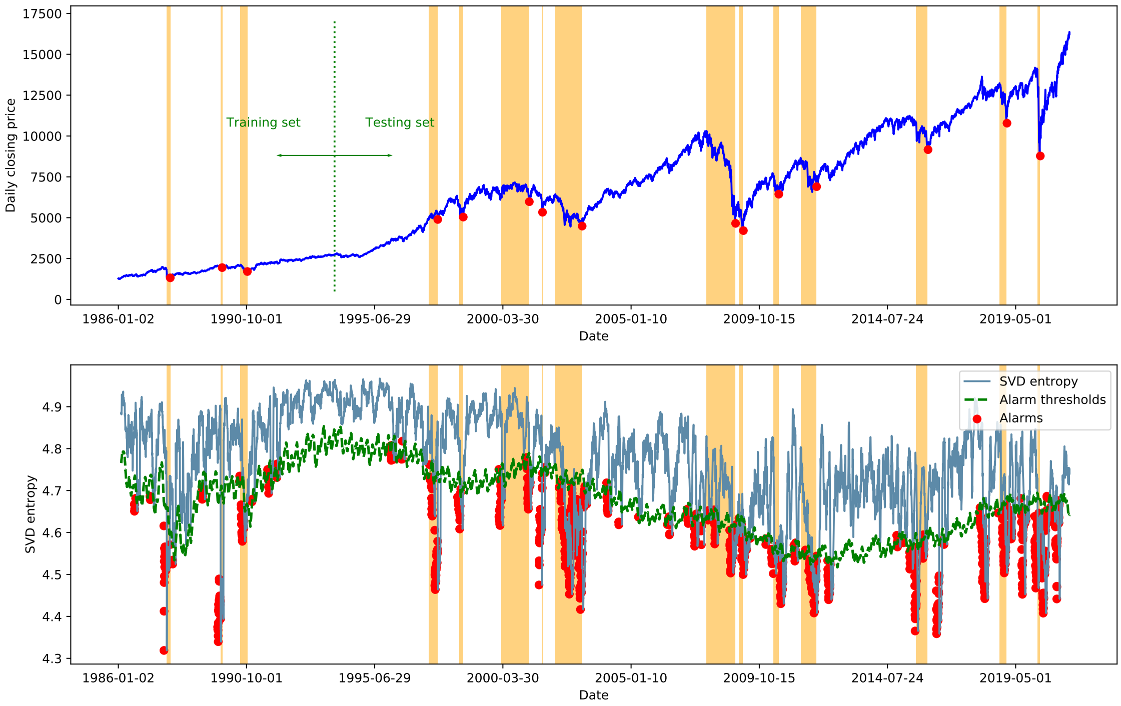


Figure 3. Result of the prediction-guided anomaly detection procedure.

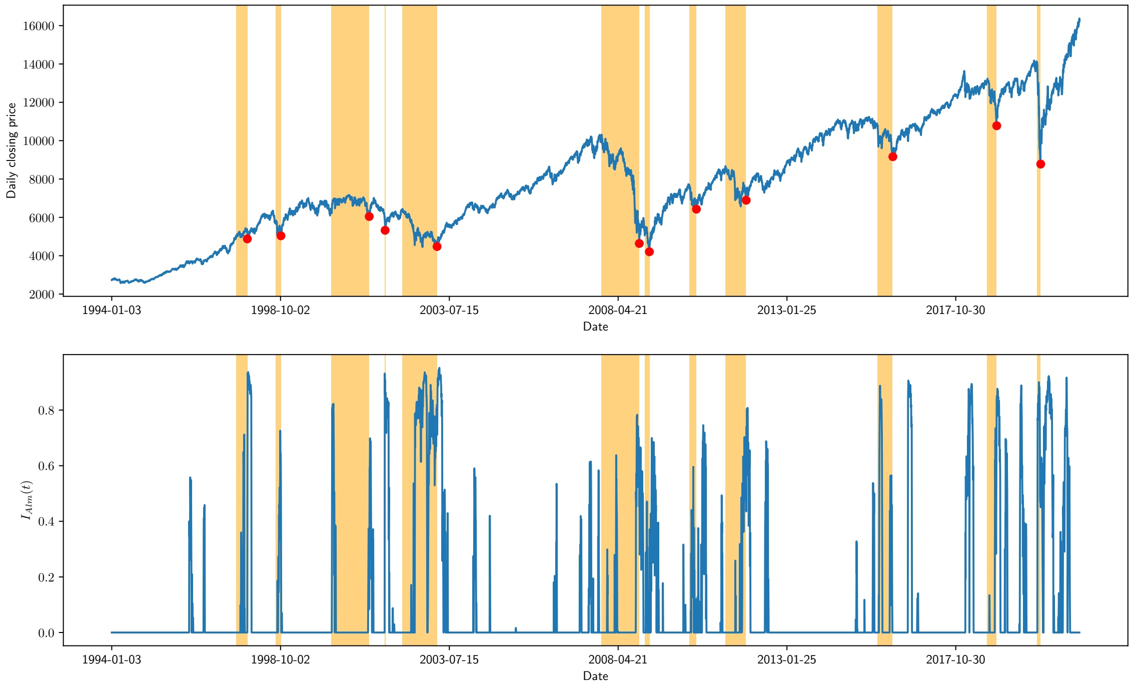


Figure 4. *Cont.*

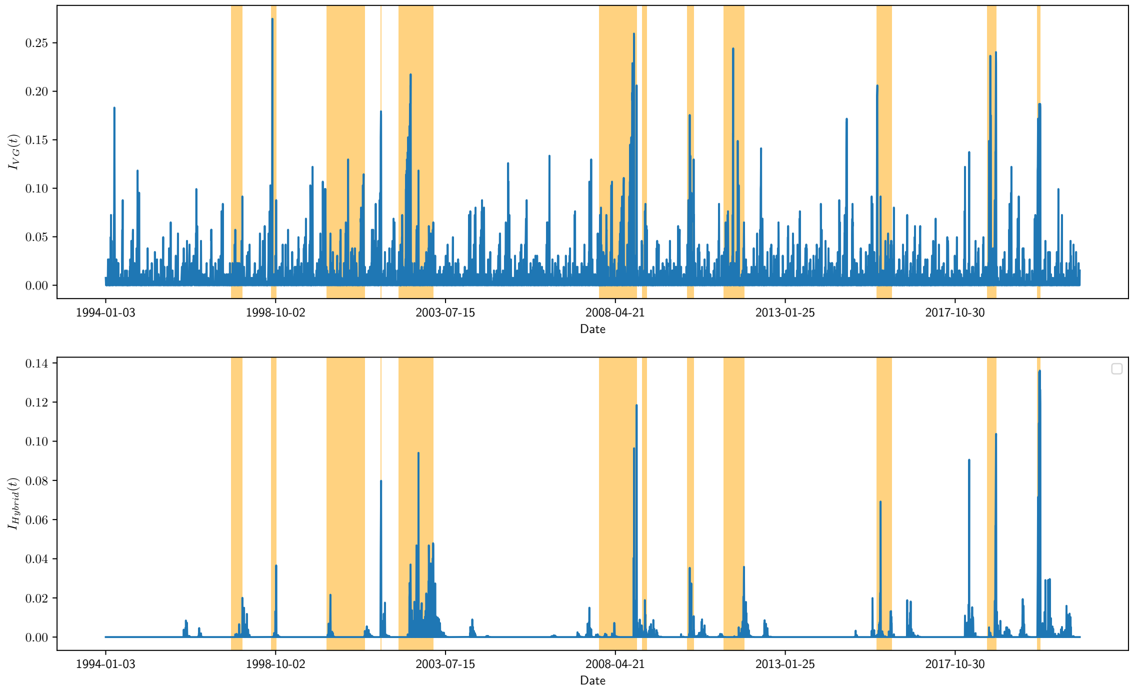


Figure 4. The NYSE Composite Index with three indicators.

4.2. Error Diagram

In this subsection, the predictive power of our proposed framework is evaluated based on the widely-adopted error diagram [4,16]. We also compare the short-term and long-term prediction performance of our proposed framework with the VG-based baseline method.

Error diagrams are often used to decide whether an indicator has predictive power for events that are difficult to predict, such as earthquakes [35–37] and financial extremes [3,16,38]. Just like the ROC (receiver operating characteristic) curve, the error diagram demonstrates the prediction performance of a certain indicator under different determination criteria (e.g., thresholds). Its x-axis is the "alarm ratio" R_{Alarm}, which is defined as the total number of alarms divided by the length of the testing period. The y-axis is the ratio of missed events R_{Miss}, which is defined as the number of missed events divided by the total number of events in the testing period. Therefore, the prediction performance of a random guess is represented by the straight line $y = 1 - x$ in the error diagram. Any prediction method that is better than a random guess follows a curve below the anti-diagonal $y = 1 - x$. Furthermore, the p-value of the hypothesis that the prediction indicator is better than a random guess is $p = A/A_{unit} = A$, where A is the area under a curve (AUC) of the error diagram and $A_{unit} = 1$ [16].

The error diagram of predicting financial rebounds of a stock index is created in the following way [4]:

1. Determine the forward-looking period a, that is, if a rebound occurs within a days after an alarm being raised, we consider the rebound has successfully predicted the alarm.
2. Count the number of rebounds in the testing set according to the definition in Section 2.
3. Sort the values of the rebound indicator time series in a decreasing order and save them in $\mathbf{I}_{sort}$. The largest value in $\mathbf{I}_{sort}$ is the first threshold.
4. We check the rebound indicator of each day during the testing period. If the rebound indicator on day t exceeds the predetermined threshold, an alarm is raised. If a rebound occurs during day t to day $t + a$, we consider the alarm successfully predicts the rebound.
5. We compare the successful predictions of the current threshold and the previous threshold. If there is no new successful prediction, the threshold is moved down to the next value in $\mathbf{I}_{sort}$.

6. If new predictions are made based on a threshold, we count the missed rebounds and calculate the ratio of missed events as

$$R_{\text{Miss}} = \frac{\text{Number of missed rebounds}}{\text{Total number of rebounds}} \tag{17}$$

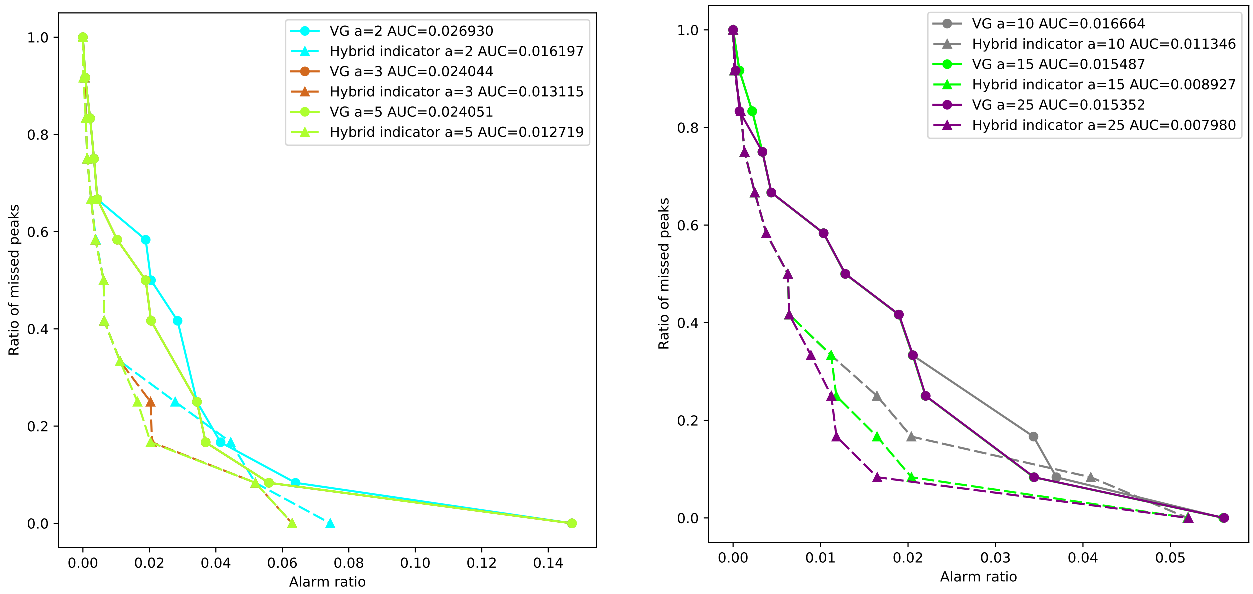
The alarm ratio is calculated as

$$R_{\text{Alarm}} = \frac{\text{Number of alarms}}{T - a}, \tag{18}$$

where T is the length of the testing period. We further plot $(R_{\text{Alarm}}, R_{\text{Miss}})$ in the error diagram.

7. Steps 4 to 6 are continuously repeated until all the rebounds are successfully predicted.

Figure 5 compares the short-term and long-term prediction performance of our proposed method (i.e., the hybrid indicator) with the VG-based indicator. We use the same color to represent the same forward-looking period a. Solid lines with circular icons correspond to the VG-based method, while dashed lines with triangular icons represent the results for the hybrid indicator. It can be observed that our proposed hybrid indicator outperforms the benchmark VG-based indicator for any forward-looking period.



(**a**) (**b**)

Figure 5. Error diagrams. (**a**) Short-term prediction of VG-based and hybrid indicator. (**b**) Long-term prediction of VG-based and hybrid indicator.

Table 2 shows the p-values of the predictions for all the forward-looking periods $a = 2, 3, 5, 10, 15, 25$. It can be observed that, for each forward-looking period a, the p-value of the VG-based indicator is smaller than 0.03. This indicates that the VG-based indicator is superior to a random guess on a significance level 3%. In other words, the predictability of the VG-based indicator is significant on a significance level 3%, which matches the result in the previous paper [16]. Similarly, our proposed hybrid indicator is superior to a random guess at a 2% significance level. We can conclude that the predictability of our proposed hybrid indicator is significant at a significance level of 2%. We can further observe that the p-value of our proposed indicator is smaller than the p-value of the VG-based indicator for each a, indicating that our proposed method constantly outperforms the benchmark method.

Table 2. The p-values of the predictions for all the forward-looking periods.

Forward-Looking Period a	VG-Based Indicator	Our Proposed Hybrid Indicator
2	0.026930	0.016197
3	0.024044	0.013115
5	0.024051	0.012719
10	0.016664	0.011346
15	0.015487	0.008927
25	0.015352	0.007980

We finally test the robustness of our proposed method. We choose T_{Alm} from $\{2, 3, \cdots, 10\}$ and calculate the nine corresponding p-values for each forward-looking period a. The box plot in Figure 6 shows the statistics of the p-values for different T_{Alm}. In this figure, each box represents the nine p-values for a forward-looking period a. The green dots are the mean, while the orange solid lines are medians. The lower and upper edges of each box are the first and third quartiles, which are denoted as Q_1 and Q_3, respectively. The top and bottom boundaries are $Q_3 + 1.5\text{IQR}$ and $Q_1 - 1.5\text{IQR}$, where $\text{IQR} = Q_3 - Q_1$ is the interquartile range. Points out of the upper and lower boundaries are outliers. For the convenience of comparison, we label the p-value of VG-based indicator with the blue dashed line for each a. It can be observed that the p-values of our method are constantly smaller than the VG-based benchmark, indicating that the performance of our proposed framework is consistently better. We can conclude that our proposed framework is robust to the choice of the free parameter T_{Alm}.

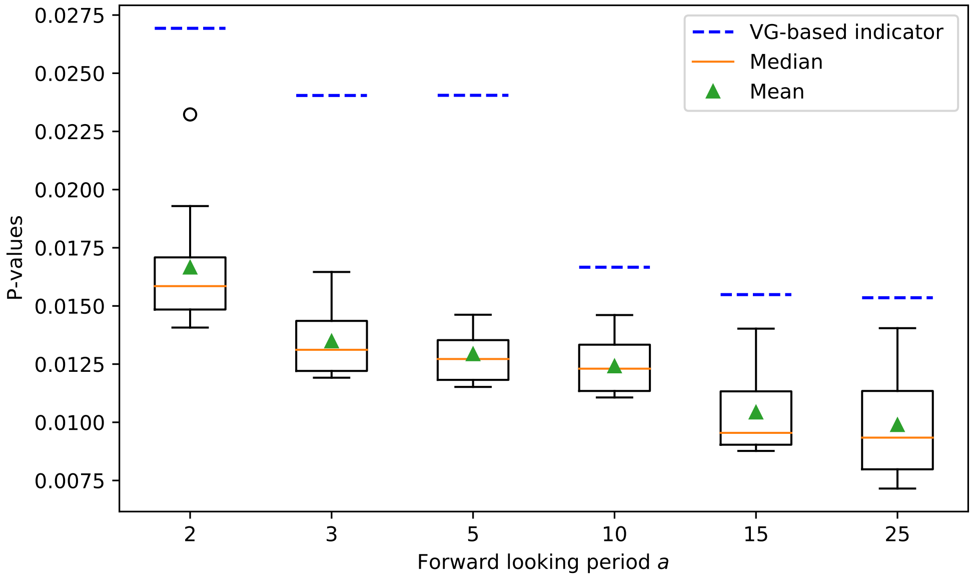


Figure 6. Statistics of the p-values for different parameters.

5. Discussion

The findings of this study indicate that incorporating the analysis of a time-evolving network can improve the performance of traditional time series models. Qualitative observations show the rapid decline of SVD entropy during financial crashes and suggest that the network-based indicator helps with reducing the false alarm rate. Quantitative comparisons further confirm that the predictive power is improved by combining the network-based indicator.

According to the qualitative observations, the SVD entropy of the stock correlation network of the NYSE Composite Index decreases significantly during stock market crashes. This is consistent with the existing studies of various stock indices based on different definitions of network entropy [13–15,39]. For example, Zhang et al. [15] have observed decreases in the von Neumann entropy and the graph motif entropy of the NYSE network during financial crashes. Furthermore, it is also observed that the structural entropy declines rapidly during crisis periods in the stock correlation network of FTSE100 and NIKKEI225 [13]. Since "entropy" can be interpreted as "diversity", such observations suggest that the "structural diversity" of the stock market commonly declines during financial crashes, and the recovery of structural diversity reflects the recovery of the market. We may further interpret this econophysics phenomenon from the perspective of behavioral finance. The structural diversity of the stock market may represent investors' perceptions in the heterogeneity of different firms. During financial crashes, spillover effects and investors' herding behavior may erase the differences between "good" and "bad" firms, and thus can eliminate the structural diversity of the stock market.

In addition, it can be observed that the network-based indicator helps with reducing the false alarm rate of the time series-based indicator. This finding can shed light on a better understanding of the stock market crashes. Traditionally, research on stock market crashes focuses on modeling the price time series of stock indices [16,17,40–43]. In particular, the benchmark indicator of this study (i.e., the VG-based LPPL indicator) [16] captures the faster-than-exponential decline in stock index price. However, our experimental results show that detection and prediction based on the time series alone can contain a number of false alarms. In other words, a rapid decline in stock index prices does not necessarily mean a financial crash. An intuitive explanation is that the plunge of highly weighted sectors or companies may drive down the stock index, but as long as the plunge is not propagated through the stock network, it will not trigger the crash of the whole market. By combining network-based and time series-based indicators, this study contributes to the LPPL-related literature and highlights that anomalous topological changes of the stock correlation network can be important criteria for confirming the occurrence and predicting the recovery of stock market crashes.

Quantitative comparisons confirm that the incorporation of SVD entropy improves the predictive power of the time series-based model. This finding agrees with the existing literature on the predictive power of SVD entropy for stock market dynamics [24–27]. For a number of representative stock indices such as the Dow Jones Industrial Average [24], the Shenzhen Component Index [25] and the Shanghai Component Index [27], it has been found that the SVD entropy of the constituent stock network has a predictive ability to the dynamics of the stock index. However, the existing literature mainly focuses on examining the predictive power of SVD entropy using the Granger causality test, without developing methods or models to practically leverage such predictive ability. This study contributes to this research gap by proposing a framework that jointly considers the SVD entropy and the stock index price. It is demonstrated that the predictive power of the SVD entropy can be practically applied.

It should be acknowledged that this work has several limitations. First, this study transforms the time-evolving stock correlation network into the time series of a network statistic (i.e., the SVD entropy) and therefore ignores the finer topology in the network. In particular, our proposed framework is not able to recognize the potential micro- and meso-level structural changes. This limitation also hinders a more in-depth observation of the topological changes during and after financial crashes. As a consequence, a number of critical research questions are left unanswered. For example, it is still unclear if similar micro- and meso-level patterns can be observed among multiple financial crashes throughout history. If so, developing models and methods based on such patterns to detect crashes and predict rebounds can be another interesting research topic. Another limitation of this study is that an in-depth analysis of the mechanism is not performed. For example, this study focuses on the diagnosis of major financial crashes. However, we have also observed rapid declines in stock indices accompanied by anomalous network changes during some periods that are not regarded as major financial crashes. Furthermore, the topological changes in the stock correlation network have become increasingly drastic from 1968 to the present. The causes of these phenomena are still unclear.

This study has several practical implications for professionals in a variety of fields. For both institutional and individual investors, our proposed rebound indicator provides a referential signal for market timing strategies. It can be observed that weak signals start to appear when a stock market crash occurs. As the stock market crash progresses, the rebound indicator gradually gets stronger and clusters around certain dates. When the rebound indicator reaches the peak and starts to decline, a change of regime is more likely to occur. Therefore, in the aftermath of stock market crashes, investors can start applying long strategies after observing the clustering and peaking of the rebound indicator. For policymakers, this study provides useful information that can help with the detection and management of market risks. When abnormal changes in the stock network are detected, policymakers can be alerted to conduct further in-depth research to determine whether a systemic risk exists. Furthermore, this study recommends introducing policies to stabilize key firms and sectors once a systemic risk occurs, since the contagion of risks may destroy the market structure and trigger herding behavior of investors. From a broader perspective, this study also has practical implications for machine learning researchers and algorithm developers. Our findings highlight the importance of incorporating stock network analysis into traditional time series models. Novel algorithms can be developed along this direction to better diagnose and forecast extreme financial events.

6. Conclusions

This study proposes a framework to diagnose stock market crashes and predict the subsequent price rebounds by jointly modeling plunges in the stock index price and abnormal changes in the stock correlation network. Experiments based on the NYSE Composite Index show that our proposed framework outperforms the benchmark VG-based LPPL model. In line with the existing literature, we observe rapid declines of the stock network's SVD entropy during market crashes. It suggests that the elimination of structural diversity in the stock market can be an important characteristic of financial crashes. In addition, it is observed that we reduce the false alarm rate of the LPPL-based indicator by incorporating the network anomaly-based indicator. This finding can shed light on bridging the gap between stock network analysis and financial time series modeling, which are often considered as two relatively independent research directions. From the perspective of market participants (e.g., policymakers and investors), this study can provide referential signals for risk management and market timing strategies.

The main limitation of this paper is the conversion of the time-evolving network into a time series, thus losing the micro- and meso-level topological patterns. Future research can benefit from using representation learning methods (e.g., graph neural network and tensor decomposition) that can directly process dynamic networks. Techniques for interpretable machine learning can also be applied to identify specific topological patterns in the course of market crashes and recoveries. Moreover, future research can be conducted to reveal the mechanism of abnormal topological changes in stock correlation networks during market crashes. Finally, practitioners can integrate our proposed rebound indicator into their trading strategies to control risk and make profits in the aftermath of stock market crashes.

Author Contributions: Conceptualization, Y.X.; methodology, Y.X.; software, Y.X.; validation, Y.X., G.W. and W.K.V.C.; formal analysis, Y.X.; investigation, Y.X.; resources, W.K.V.C.; data curation, Y.X.; writing—original draft preparation, Y.X.; writing—review and editing, G.W. and W.K.V.C.; visualization, Y.X.; supervision, W.K.V.C.; project administration, W.K.V.C.; funding acquisition, W.K.V.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research is funded by the Shenzhen Science and Technology Innovation Commission (Grant No. JCYJ20210324135011030, WDZC20200818121348001 and KCXFZ202002011010487), the National Natural Science Foundation of China (Grant No. 71971127 and 72171164), Guangdong Pearl River Plan (2019QN01X890), and the Hylink Digital Solutions Co., Ltd. (120500002).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: This study uses publicly available datasets, which can be found in http://finance.yahoo.com and https://en.wikipedia.org/w/index.php?title=List_of_stock_market_crashes_and_bear_markets (accessed on 25 November 2021).

Acknowledgments: We gratefully acknowledge Peiming Li from School of Finance, Central University of Finance and Economics for his technical support in data acquisition and processing. Likewise, we are grateful to Xinyue Ren, Wenxiu Zhang, Weida Wang and Xinglu Liu from Tsinghua-Berkeley Shenzhen Institute, Tsinghua University for proofreading the manuscript and sharing their thoughts with us. Furthermore, we would like to thank the four anonymous reviewers for their constructive and insightful comments and the editor for his help in reviewing and processing the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A. Trend Labeling Algorithm

For completeness, we illustrate the detailed algorithm for labeling the trend of the stock index price time series, which is proposed by Wu et al. [18]. The detailed procedure is introduced in Algorithm A1. The input of this algorithm is the time series of daily prices $P = [p_1, p_2, p_3, \cdots, p_T]$, as well as the predetermined proportion threshold $w \in (0,1)$. The output of the algorithm is the corresponding label vector $Y = [\text{ label }_1, \text{ label }_2, \text{ label }_3, \ldots, \text{ label }_T]$, where

$$\text{label}_t = \begin{cases} 1, & \text{if a rebound occurs on day } t. \\ -1, & \text{if day } t \text{ is a turning point from uptrend to downtrend.} \\ 0, & \text{otherwise} \end{cases} \tag{A1}$$

Before executing the labeling algorithm, the local peak and the local trough are initialized as $\text{LP} = p_1$ and $\text{LT} = p_1$. The trend label is initialized as $trend = 0$, where $trend = 1$ indicates an upward trend and $trend = -1$ indicates a downward trend. The label vector Y is initialized as $Y = [0, 0, \cdots, 0]$. The index of the first financial extreme t is initialized as $t = 0$.

Algorithm A1 Method for labeling rebounds

```
Input:
    $P = [p_1, p_2, p_3, \cdots, p_T]$;
    $w = 15$
Output:
    $Y = [\text{ label }_1, \text{ label }_2, \text{ label }_3, \ldots, \text{ label }_T]$.
for $i = 1$ to $T$ do
    if $P[i] > \text{LP}(1 + w)$ then
        Labeling the first peak: $\text{LP} = P[i]$, $trend = 1$, $t = i$, $Y[i] = 1$.
        break
    end if
    if $P[i] < \text{LT}(1 - w)$ then
        Labeling the first trough: $\text{LT} = P[i]$, $trend = -1$, $t = i$, $Y[i] = 1$.
        break
    end if
end for
for $i = t$ to $T$ do
    if $trend == 1$ then
        if $P[i] > \text{LP}$ then
            Updating the local peak: $\text{LP} = P[i]$
        end if
        if $P[i] < \text{LP}(1 - w)$ then
            Labeling the turning point: $\text{LT} = P[i]$, $Y[i] = -1$, $trend = -1$.
        end if
    end if
    if $trend == -1$ then
        if $P[i] < \text{LT}$ then
            Updating the local trough: $\text{LT} = P[i]$
        end if
        if $P[i] > \text{LT}(1 + w)$ then
            Labeling the turning point: $\text{LP} = P[i]$, $Y[i] = 1$, $trend = 1$
        end if
    end if
end for
return $Y$;
```

Appendix B. Algorithm for the Generation of a Training Set

This appendix describes the algorithm for the generation of training set, whose detailed procedure is described in Algorithm A2. The input of this algorithm is the first N data points of the indicator time series $\mathbf{I} = [I_1, I_2, \cdots, I_N]$ with labeled financial crashes $\mathbf{L} = [l_1, l_2, , \cdots, l_T]$, where

$$l_t = \begin{cases} 1, & \text{if day } t \text{ is during a financial crash.} \\ 0, & \text{otherwise} \end{cases} \tag{A2}$$

Algorithm A2 Generating training set

```
Input:
    First $N$ indicators $\mathbf{I} = [I_1, I_2, \cdots, I_N]$;
    Corresponding event labels $\mathbf{L} = [l_1, l_2, , \cdots, l_N]$;
    The sliding window length $d$.
Output:
    Training input set $\mathbf{S}_x$ and target output set $\mathbf{S}_y$.
$i, j = 0$, $\mathbf{W} \leftarrow []$
while $i < N$ do
    while $j < N - 1$ and $l_j == l_{j+1}$ do
        $j = j + 1$
        if $l_i == 0$ then
            $W_i = \mathbf{I}[i : j]$, $\mathbf{W} \leftarrow \mathbf{W} \cup W_i$
        end if
        $j = j + 1, i = j$
    end while
end while
$\mathbf{S}_x, \mathbf{S}_y \leftarrow []$
for $i = 1$ to length($\mathbf{W}$) do
    for $j = d + 1$ to length($W_i$) - 1 do
        $\mathbf{S}_x \leftarrow \mathbf{S}_x \cup W_i[j - d : j]$
        $\mathbf{S}_y \leftarrow \mathbf{S}_y \cup W_i[j + 1]$
    end for
end for
return $\mathbf{S}_x, \mathbf{S}_y$
```

Appendix C. Obtaining Optimal Threshold Based on Extreme Value Theory

In this appendix, we describe in detail the method of obtaining the optimal threshold for residuals, which is used to distinguish normal residuals from abnormal residuals. The training data that we are using are the SVD entropy of the NYSE financial network from 10 February 1986 to 3 January 1994. The residuals are calculated as

$$X_t = F(I_{t-d}, \cdots, I_{t-1}) - I(t). \tag{A3}$$

In this case, we use the Prophet forecasting model [34] as the predictor $F(\cdot)$, whose open-source implementation is available at https://github.com/facebook/prophet (accessed on 25 November 2021).

The general procedure for obtaining τ is summarized in Algorithm A3. Such algorithm mainly contains three key steps: (1) generating possible candidates, (2) estimating parameters for GPD; (3) selecting the optimal threshold based on the Kolmogorov–Smirnov statistic. The detailed procedure is described as follows.

Appendix C.1. Generating Possible Candidates

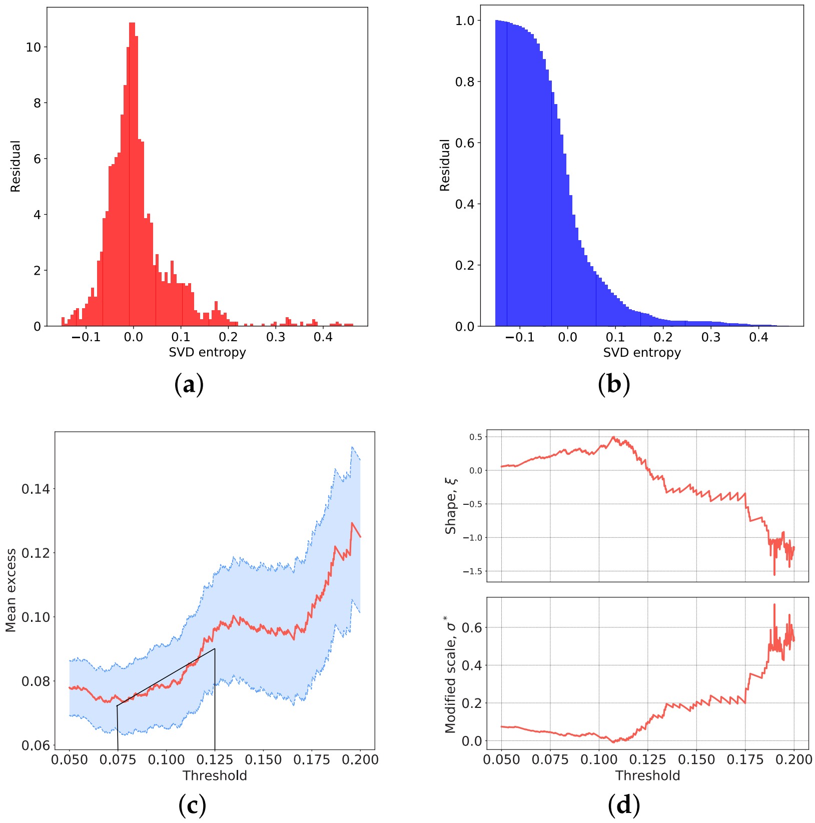
The possible candidates for the threshold are selected based on a hybrid approach combining EVT and observation of the data. The two most common methods for selecting thresholds in EVT applications are the mean residual life (MRL) plot and the parameter stability plot [44,45]. The MRL plot shows the relationship between the *mean excess* (i.e., the average of $\mathbf{Y}_\tau$) and τ. Its theoretical foundation is the conclusion proved by [46], that is, if the threshold τ_0 enables $X - \tau_0$ to be the GPD excess, for any higher threshold $\tau > \tau_0$, the average value of $X - \tau$ (i.e., the *mean excess*) is $E(X - \tau \mid X > \tau) = (\sigma + \xi\tau)/(1 - \xi)$, which is linear to τ. Therefore, the MRL plot is approximately linear within a suitable range of τ. Meanwhile, too low a threshold τ would violate the assumption of extreme value theory and thus lead to a tail distribution that does not obey the GPD. Too high a threshold will lead to too few excesses, which brings a large uncertainty to the parameter estimation. Therefore, the MRL plot is not a straight line for a too high or too low threshold τ.

Algorithm A3 Finding the threshold τ based on EVT

```
Input:
    The residuals for the first $N$ days $\mathbf{X} = [X_1, X_2, X_3, \cdots, X_N]$.
Output:
    $\tau^*$, $\xi(\tau^*)$, and $\sigma(\tau^*)$.
Generating possible candidates for the threshold $\boldsymbol{\tau} = [\tau_1, \tau_2, \tau_3, \cdots]$.
for $\tau$ in $\boldsymbol{\tau}$ do
    $\mathbf{Y}_\tau \leftarrow \{X_i - \tau \mid X_i > \tau\}$;
    Estimating parameters $\hat{\xi}(\tau)$ and $\hat{\sigma}(\tau)$ based on $\mathbf{Y}_\tau$;
    Calculating the Kolmogorov–Smirnov statistic KS$(\tau)$ between GPD$\big(\hat{\xi}(\tau), \hat{\sigma}(\tau)\big)$ and $\mathbf{Y}_\tau$;
end for
return $\tau^*$ with the smallest KS$(\tau)$, as well as two parameters $\xi(\tau^*)$ and $\sigma(\tau^*)$ of the GPD.
```

Generally, the MRL plot can provide a rough range of the proper threshold. However, the MRL plot in practical applications may not be ideally straight. In fact, it is widely recognized by researchers that the interpretation of practical MRL plots can be challenging [33]. As a supplementary method, the parameter stability plot involves plotting the estimated parameters $\hat{\sigma}(\tau)$ and $\hat{\xi}(\tau)$ of the GPD model against τ. Within the proper range of τ, $\hat{\sigma}(\tau)$ and $\hat{\xi}(\tau)$ should be near-constant, so the parameter stability plot should be an approximately horizontal straight line.

In addition, observing the probability density function (PDF) and complementary cumulative distribution function (CCDF) of the data also gives us an intuitive understanding of the appropriate threshold values. Therefore, we plotted these four figures (i.e., PDF, CCDF, the MRL plot and the parameter stability plot) in Figure A1.



(**a**) (**b**) (**c**) (**d**)

Figure A1. Selecting the threshold for the NYSE Composite Index. (**a**) Empirical PDF. (**b**) Empirical CCDF. (**c**) MRL Plot. (**d**) Parameter Stability Plot.

In this figure, (a), (b), (c) and (d) are the PDF, CCDF, the MRL plot and the parameter stability plot of the residual. From the PDF, it can be observed that the major part of this distribution approximates a normal distribution, which is consistent with our assumptions on the residual $\epsilon(t)$. Furthermore, the distribution has a heavy tail where the outliers are located. The CCDF is also obviously heavy-tailed, which supports our hypothesis based on EVT. Combining the approximately linear part of the MRL plot and the near-constant part of the parameter stability plot, we determined the search interval for the threshold τ as $(0.075, 0.125)$. It is worth noting that we actually relax the search interval. The possible candidates for the threshold are the data points that lie within the search interval, that is, $\mathbf{T} = \{X_t \mid 0.075 < X_t < 0.125\}$.

Appendix C.2. Estimating Parameters for GPD

For each possible candidate $\tau \in \mathbf{T}$, we assume that the threshold excesses $\mathbf{Y}_\tau = \{X_i - \tau \mid X_i > \tau\}$ follow the GPD. The parameters $\hat{\xi}(\tau)$ and $\hat{\sigma}(\tau)$ are estimated based on the maximum likelihood estimation (MLE), which is considered as a more efficient and robust approach compared with the method of moments (MOM) method and the probability weighted moments (PWM) method [47].

We follow the parameter estimation procedure illustrated by [29], which is based on Grimshaw's trick [48]. The MLE method aims at maximizing the log-likelihood function of GPD:

$$\ell(\xi,\sigma) = \log \mathcal{L}(\xi,\sigma) = -N_t \log \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{N_t} \log\left(1 + \frac{\xi}{\sigma} Y_i\right) \tag{A4}$$

where $Y_i \in \mathbf{Y}_\tau$ are the threshold excesses. The global optimal solutions ξ^* and σ^* that maximize $\ell(\xi,\sigma)$ must satisfy

$$\nabla \ell(\xi^*, \sigma^*) = 0. \tag{A5}$$

Grimshaw [48] has demonstrated that the solutions of Equation (A5) satisfy the following conditions

$$\begin{cases} u(x)v(x) = 1 \\ u(x) = \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{1}{1+xY_i} \\ v(x) = 1 + \frac{1}{N_t} \sum_{i=1}^{N_t} \log(1 + xY_i) \end{cases}, \tag{A6}$$

where $x = \xi^*/\sigma^*$. Solutions of Equation (A6) can be obtained by performing numerical root searches in the interval $(x_{\min}, x_{\max})$. The lower bound of the interval is $x_{\min} = -1/\max\{\mathbf{Y}_\tau\}$ and the upper bound is

$$x_{\max} = 2\frac{\overline{\mathbf{Y}}_\tau - \min\{\mathbf{Y}_\tau\}}{(\min\{\mathbf{Y}_\tau\})^2}, \tag{A7}$$

where $\overline{\mathbf{Y}}_\tau$ is the mean of $\mathbf{Y}_\tau$. For each solution x of Equation (A6), the corresponding ξ and σ can be calculated as,

$$\begin{cases} \xi = v(x) - 1 \\ \sigma = \xi/x \end{cases} \tag{A8}$$

For all solutions of Equation (A6), we compute their corresponding ξ and σ, which are substituted into Equation (A4) to compute $\ell(\xi,\sigma)$. We select the ξ and σ that maximize $\ell(\xi,\sigma)$ as the estimated parameters of the GPD.

Appendix C.3. Threshold Selection Based on the Kolmogorov–Smirnov Statistic

In the previous step, we estimated the parameters $\hat{\xi}(\tau)$ and $\hat{\sigma}(\tau)$ of GPD for each possible threshold $\tau \in \mathbf{T}$. Here we calculate the Kolmogorov–Smirnov (KS) statistic between the empirical distribution function (EDF) of $\mathbf{Y}_\tau$ and the cumulative distribution function (CDF) of GPD with parameters $\hat{\xi}(\tau)$ and $\hat{\sigma}(\tau)$. Figure A2 plots the KS statistics against the threshold τ for the residual. We choose the optimal threshold τ^* such that the corresponding KS statistic is minimized, which means that EDF and CDF are the closest and the best fit is obtained. As is shown in Figure A2, the optimal threshold is $\tau^* = 0.1164$, with $\xi(\tau^*) = 0.3004$ and $\sigma(\tau^*) = 0.0276$.

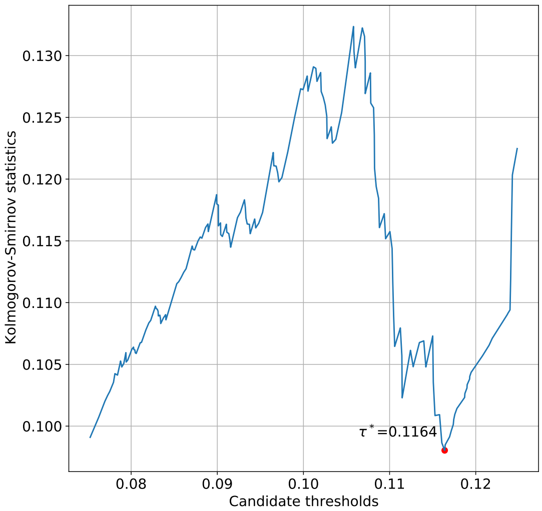


Figure A2. Threshold selection based on the Kolmogorov–Smirnov statistic.

Appendix D. Execution Step

This appendix presents the detailed procedure of the execution step. With the input of streaming data, we keep updating the training set, which is used to continuously fine-tune the predictor. It is worth noting that we distinguish between "normal" and "abnormal" data based on EVT, and only include "normal" data into the updated training set. In this way, the predictor is capable of tracking the "normal" changes in the dynamics of the time series.

Algorithm A4 Execution step

```
The set for new training samples $W \leftarrow []$.
$c = 0$
for $t > N$ do
    $\hat{I}_t = F(I_{t-d}, \cdots, I_{t-1})$
    $X_t = \hat{I}_t - I_t$
    if $X_t > \tau$ then
        Raise an alarm with the alarm index $F_{\xi,\sigma}(X_t - \tau)$.
        Empty new training sample set: $W \leftarrow []$
    else
        Append new training sample set: $W \leftarrow W \cup I_t$.
        if length$(W) > d$ then
            $\mathbf{S}_x \leftarrow \mathbf{S}_x[1:] \cup W[1:d]$
            $\mathbf{S}_y \leftarrow \mathbf{S}_y[1:] \cup W[d+1]$
            $W \leftarrow []$
            $c \leftarrow c + 1$
        end if
    end if
    if mod$(c, K) == 0$ then
        Train the predictor $F(\cdot)$ by the updated training set $\mathbf{S}_x$ and $\mathbf{S}_y$.
    end if
end for
```

References

1. Mazur, M.; Dang, M.; Vega, M. COVID-19 and the March 2020 stock market crash: Evidence from S&P1500. *Financ. Res. Lett.* **2021**, *38*, 101690.
2. Sornette, D.; Johansen, A.; Bouchaud, J.P. Stock market crashes precursors and replicas. *J. Phys. I* **1996**, *6*, 167–175. [CrossRef]
3. Yan, W.; Woodard, R.; Sornette, D. Diagnosis and prediction of tipping points in financial markets: Crashes and rebounds. *Phys. Procedia* **2010**, *3*, 1641–1657. [CrossRef]
4. Yan, W.; Woodard, R.; Sornette, D. Diagnosis and prediction of rebounds in financial markets. *Phys. A Stat. Mech. Its Appl.* **2012**, *391*, 1361–1380. [CrossRef]
5. Fantazzini, D. The oil price crash in 2014/15: Was there a negative financial bubble? *Energy Policy* **2016**, *96*, 383–396. [CrossRef]
6. Fry, J.; Cheah, E.T. Negative bubbles and shocks in cryptocurrency markets. *Int. Rev. Financ. Anal.* **2016**, *47*, 343–352. [CrossRef]
7. Stavroglou, S.K.; Pantelous, A.A.; Stanley, H.E.; Zuev, K.M. Hidden interactions in financial markets. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 10646–10651. [CrossRef]
8. Bardoscia, M.; Barucca, P.; Battiston, S.; Caccioli, F.; Cimini, G.; Garlaschelli, D.; Saracco, F.; Squartini, T.; Caldarelli, G. The physics of financial networks. *Nat. Rev. Phys.* **2021**, *3*, 490–507. [CrossRef]
9. Shen, Y.Y.; Jiang, Z.Q.; Ma, J.C.; Wang, G.J.; Zhou, W.X. Sector connectedness in the Chinese stock markets. *Empir. Econ.* **2021**, *1*, 1–28. [CrossRef]
10. Heiberger, R.H. Stock network stability in times of crisis. *Phys. Stat. Mech. Its Appl.* **2014**, *393*, 376–381. [CrossRef]
11. Silva, F.N.; Comin, C.H.; Peron, T.K.D.; Rodrigues, F.A.; Ye, C.; Wilson, R.C.; Hancock, E.; Costa, L.d.F. Modular dynamics of financial market networks. *arXiv* **2015**, arXiv:1501.05040.
12. Anand, K.; Bianconi, G. Entropy measures for networks: Toward an information theory of complex topologies. *Phys. Rev. E* **2009**, *80*, 045102. [CrossRef]
13. Almog, A.; Shmueli, E. Structural entropy: Monitoring correlation-based networks over time with application to financial markets. *Sci. Rep.* **2019**, *9*, 10832. [CrossRef]
14. Shi, Y.; Zheng, Y.; Guo, K.; Jin, Z.; Huang, Z. The evolution characteristics of systemic risk in China's stock market based on a dynamic complex network. *Entropy* **2020**, *22*, 614. [CrossRef]
15. Zhang, Z.; Chen, D.; Bai, L.; Wang, J.; Hancock, E.R. Graph motif entropy for understanding time-evolving networks. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *1*, 1–15. [CrossRef]
16. Yan, W.; van Tuyll van Serooskerken, E. Forecasting financial extremes: A network degree measure of super-exponential growth. *PLoS ONE* **2015**, *10*, e0128908. [CrossRef]
17. Chen, D.R.; Liu, C.; Zhang, Y.C.; Zhang, Z.K. Predicting financial extremes based on weighted visual graph of major stock indices. *Complexity* **2019**, *2019*, 5320686. [CrossRef]
18. Wu, D.; Wang, X.; Su, J.; Tang, B.; Wu, S. A labeling method for financial time series prediction based on trends. *Entropy* **2020**, *22*, 1162. [CrossRef]
19. Lacasa, L.; Luque, B.; Ballesteros, F.; Luque, J.; Nuño, J.C. From time series to complex networks: The visibility graph. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 4972–4975. [CrossRef]
20. Onnela, J.P.; Chakraborti, A.; Kaski, K.; Kertész, J.; Kanto, A. Dynamics of market correlations: Taxonomy and portfolio analysis. *Phys. Rev. E* **2003**, *68*, 056110. [CrossRef]
21. Epps, T.W. Comovements in stock prices in the very short run. *J. Am. Stat. Assoc.* **1979**, *74*, 291–298.
22. Wang, J.; Lin, C.; Wang, Y. Thermodynamic entropy in quantum statistics for stock market networks. *Complexity* **2019**, *2019*, e1817248. [CrossRef]
23. Tse, C.K.; Liu, J.; Lau, F.C.M. A network perspective of the stock market. *J. Empir. Financ.* **2010**, *17*, 659–667. [CrossRef]
24. Caraiani, P. The predictive power of singular value decomposition entropy for stock market dynamics. *Phys. A Stat. Mech. Its Appl.* **2014**, *393*, 571–578. [CrossRef]
25. Gu, R.; Xiong, W.; Li, X. Does the singular value decomposition entropy have predictive power for stock market?—Evidence from the Shenzhen stock market. *Phys. A Stat. Mech. Its Appl.* **2015**, *439*, 103–113. [CrossRef]
26. Gu, R.; Shao, Y. How long the singular value decomposed entropy predicts the stock market?—Evidence from the Dow Jones industrial average index. *Phys. A Stat. Mech. Its Appl.* **2016**, *453*, 150–161. [CrossRef]
27. Jiang, J.; Gu, R. Using Rényi parameter to improve the predictive power of singular value decomposition entropy on stock market. *Phys. A Stat. Mech. Its Appl.* **2016**, *448*, 254–264. [CrossRef]
28. Minello, G.; Rossi, L.; Torsello, A. On the von Neumann entropy of graphs. *J. Complex Netw.* **2019**, *7*, 491–514. [CrossRef]
29. Siffer, A.; Fouque, P.A.; Termier, A.; Largouet, C. Anomaly detection in streams with extreme value theory. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, 13–17 August 2017; pp. 1067–1075.
30. Bień-Barkowska, K. Looking at extremes without going to extremes: A new self-exciting probability model for extreme losses in financial markets. *Entropy* **2020**, *22*, 789. [CrossRef]
31. Balkema, A.A.; De Haan, L. Residual life time at great age. *Ann. Probab.* **1974**, *2*, 792–804. [CrossRef]
32. Pickands, J., III. Statistical inference using extreme order statistics. *Ann. Stat.* **1975**, *3*, 119–131.
33. Scarrott, C.; MacDonald, A. A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT–Stat. J.* **2012**, *10*, 33–60.

34. Taylor, S.J.; Letham, B. Forecasting at scale. *Am. Stat.* **2018**, *72*, 37–45. [CrossRef]
35. Molchan, G.M. Earthquake prediction as a decision-making problem. *Pure Appl. Geophys.* **1997**, *149*, 233–247. [CrossRef]
36. Han, P.; Zhuang, J.; Hattori, K.; Chen, C.H.; Febriani, F.; Chen, H.; Yoshino, C.; Yoshida, S. Assessing the potential earthquake precursory information in ULF magnetic data recorded in Kanto, Japan during 2000–2010: Distance and magnitude dependences. *Entropy* **2020**, *22*, 859. [CrossRef] [PubMed]
37. Wang, R.; Chang, Y.; Miao, M.; Zeng, Z.; Chen, H.; Shi, H.; Li, D.; Liu, L.; Su, Y.; Han, P. Assessing earthquake forecast performance based on b value in Yunnan Province, China. *Entropy* **2021**, *23*, 730. [CrossRef] [PubMed]
38. Sornette, D.; Zhou, W.X. Predictability of large future changes in major financial indices. *Int. J. Forecast.* **2006**, *22*, 153–168. [CrossRef]
39. Shi, Y.; Zheng, Y.; Guo, K.; Ren, X. Relationship between herd behavior and Chinese stock market fluctuations during a bullish period based on complex networks. *Int. J. Inf. Technol. Decis. Mak.* **2021**, *1*, 1–17. [CrossRef]
40. Zhang, Q.; Zhang, Q.; Sornette, D. Early warning signals of financial crises with multi-scale quantile regressions of log-periodic power law singularities. *PLoS ONE* **2016**, *11*, e0165819. [CrossRef]
41. Jiang, Z.Q.; Canabarro, A.; Podobnik, B.; Stanley, H.E.; Zhou, W.X. Early warning of large volatilities based on recurrence interval analysis in Chinese stock markets. *Quant. Financ.* **2016**, *16*, 1713–1724. [CrossRef]
42. Jiang, Z.Q.; Wang, G.J.; Canabarro, A.; Podobnik, B.; Xie, C.; Stanley, H.E.; Zhou, W.X. Short term prediction of extreme returns based on the recurrence interval analysis. *Quant. Financ.* **2018**, *18*, 353–370. [CrossRef]
43. Ghosh, B.; Kenourgios, D.; Francis, A.; Bhattacharyya, S. How well the log periodic power law works in an emerging stock market? *Appl. Econ. Lett.* **2021**, *28*, 1174–1180. [CrossRef]
44. Cirillo, P.; Taleb, N.N. Tail risk of contagious diseases. *Nat. Phys.* **2020**, *16*, 606–613. [CrossRef]
45. Mehrnia, N.; Coleri, S. Non-stationary wireless channel modeling approach based on extreme value theory for ultra-reliable communications. *IEEE Trans. Veh. Technol.* **2021**, *70*, 8264–8268. [CrossRef]
46. Davison, A.C.; Smith, R.L. Models for exceedances over high thresholds. *J. R. Stat. Soc. Ser. B (Methodol.)* **1990**, *52*, 393–442. [CrossRef]
47. Beirlant, J.; Goegebeur, Y.; Segers, J.; Teugels, J.L. *Statistics of Extremes: Theory and Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2006.
48. Grimshaw, S.D. Computing maximum likelihood estimates for the generalized Pareto distribution. *Technometrics* **1993**, *35*, 185–191. [CrossRef]