Detection of Chinese Stock Market Bubbles with LPPLS Confidence Indicator

Min Shu[1, 2, *], Wei Zhu[1, 2]

[1] Department of Applied Mathematics & Statistics, Stony Brook University, Stony Brook, NY, USA
[2] Center of Excellence in Wireless & Information Technology, Stony Brook University, Stony Brook, NY, USA

Abstract

This paper aims to present an advance bubble detection methodology based on LPPLS confidence indicator for the early causal identification of positive and negative bubbles in the Chinese stock market using the daily data on the Shanghai Shenzhen CSI 300 stock market index from January 2002 through April 2018. We account for the damping condition of LPPLS model in the search space and implement the stricter filter conditions for the qualification of the valid LPPLS fits by taking account of the maximum relative error, Lomb log-periodic test of the detrended residual, and unit-root tests of the logarithmic residual based on both the Phillips-Perron test and Dickey-Fuller test to improve the performance of LPPLS confidence indicator. Our analysis shows that the LPPLS detection strategy diagnoses the positive bubbles and negative bubbles corresponding to well-known historical events, implying the detection strategy based on the LPPLS confidence indicator has an outstanding performance to identify the bubbles in advance. We find that the probability density distribution of the estimated beginning time of bubbles appears to be skewed and the mass of the distribution is concentrated on the area where the bubbles start to have a super-exponential growth. This study presents that it is possible to detect the potential positive and negative bubbles and crashes ahead of time, which provides a prerequisite for limiting the bubble sizes and eventually minimizing the damage from the bubble crash.

Keywords: Financial bubble, Market crash, Log-periodic power law singularity, Chinese stock market, LPPLS confidence indicator, Bubble indicator

---

* Corresponding author: min.shu@stonybrook.edu

1. Introduction

In the modern society, financial bubbles and crashes are not rare phenomena and have great impact on the lives and livelihoods of most people all over the world. Approximately 100 financial crises worldwide have been observed in the past 30 years (Stiglitz, 2014). It is vital to identify bubbles in advance, limit their sizes, and eventually minimize the damage from the bubble crash. The causes of bubbles have been widely investigated and recent theories indicate bubbles of stock market can be generated because of (1) heterogeneous beliefs of investors together with short-time constraints, (2) positive feedback trading by noise traders, and (3) synchronization failures among rational traders (Yan, 2011).

In order to detect the presence of a bubble effectively, the Log Periodic Power Law Singularity (LPPLS) model (Johansen et al., 2000; Johansen et al., 1999; Sornette & Johansen, 2001) has been developed at the interface of financial economics, behavioral finance and statistical physics. In the LPPLS model based on the theory of rational expectation, the bubbles are believed to be characterized by faster-than-exponential (or super-exponential) growth of price leading to unsustainable growth ending with a finite crash-time t_c. The super-exponential growth of price of a bubble results from positive feedback mechanism in the valuation of assets created by imitation and herding behavior of noise traders and of boundedly rational agent results in price processes that exhibit a finite-time singularity at some future time (Yan, 2011). Because of the tension and competition between the value investors and the noise traders, the market price of an asset is deviated around the faster-than-exponential growth in the form of oscillations that are periodic in the logarithm of the time to t_c. Based on analyzing the price time series of an asset, the LPPLS model provided a flexible framework to detect financial bubbles. Over the past decade, the LPPLS model has been widely used to detect bubble and crashes in various markets, such as the real estate market in Las Vegas (Zhou & Sornette, 2008), the 2000-2003 real estate bubble in the UK (Zhou & Sornette, 2003), the USA real estate bubble (Zhou & Sornette, 2006), the 2006-2008 oil bubble (Sornette et al., 2009), the Chinese stock market bubbles in 2005–2007 and 2008–2009 (Jiang et al., 2010), and the Shanghai 2015 stock market bubble (Sornette et al., 2015).

In recent year, there is a growing research on the LPPLS model to aim at detecting bubbles. Yan et al. (2010) adapted LPPLS formula to model the negative bubbles, so that the market rebounds can be detected by implementing a pattern recognition. Brée et al. (2013) found that the LPPLS functions are intrinsically very hard to fit to time series by accounting for the sloppiness. Sornette et al. (2013) discussed the theoretical status and common calibration issues concerning LPPLS model. Filimonov and Sornette (2013) transformed the formulation of the LPPLS formula to reduce the number of nonlinear parameters in the function from four to three, which reduces complexity and improves the stability of the calibration. Geraskin and Fantazzini (2013) presents a detailed guide for modelling and identifying financial bubbles using LPPLS model. Lin et al. (2014) proposed a self-consistent model for explosive financial bubbles which combines a mean-reverting volatility process and a stochastic conditional return. Sornette et al. (2015) evaluated the performance of the real-time prediction of bubble crash in 2015 Shanghai stock market by constructing DS LPPLS Confidence indicator and DS LPPLS Trust indicator, and conducted the relevant post-mortem analysis on the effectiveness of LPPLS methodology. Zhang et al. (2016) adopted the quantile regression in LPPLS calibration and used a multi-scale

analysis to combine the many quantile regressions. The DS LPPLS confidence and trust indicators were also implemented to enrich the diagnostic of bubbles. Li (2017) investigated the critical times of three historical Chinese stock market bubbles and supported that the LPPLS performs well to predict the bubble crashes and the forecast gap is an alternative way for the market conversion warning. Demos and Sornette (2017) carried out systematic tests of the precision and reliability of determining the beginning and end time of a bubble, and found that the beginning of bubbles is much better constrained than their end. Filimonov et al. (2017) applied the modified profile likelihood inference method to calibrate the LPPLS model of the financial bubbles and obtained the interval estimation for the critical time. Demirer et al. (2019) applied the LPPLS confidence multi-scale indicators to evaluate the predictive power of market-based indicators and identified that the short selling and liquidity are two importance factors to contributing to the bubble indicators.

The financial system in China have evolved from Mao's single-bank system to Deng's four-bank system and is still dominated by its state-owned bank sector. China's stock markets opened in 1990, mainly as a platform for privatization of state-owned enterprises and the selected firms in the list strictly controlled by the government. Until 2005, only one-third of equity shares were tradable, and the total market capitalization was not over $1 trillion until 2006 (Carpenter & Whitelaw, 2017). However, due to a series of developments over the last decade, the China's economy has a stellar growth and China's GDP has more than tripled to over $11 trillion in 2016. The total market capitalization of China's stock market has grown more than five-fold to over $7 trillion until May 2017, and China's stock market becomes the world's second largest, which is attracting attention from mainstream research in financial economics.

With the rapid growth of China's economy, China's stock markets have experienced a roller coaster dynamics, with two large bubbles bursting respectively from May 2005 to October 2007, from November 2008 to August 2009, and from mid-2014 to June 2015 (Sornette et al., 2015). In mainland China, the organized stock market is composed of two stock exchanges: Shanghai stock exchange (SHSE) and Shenzhen stock exchange (SZSE). The one of the most important indexes for A-shares is the Shanghai Shenzhen CSI 300 index (CSI 300), which is a capitalization-weighted stock market index to replicate the performance of top 300 stocks traded in Shanghai and Shenzhen stock exchanges. The CSI 300 index has been calculated since April 8, 2005. The evolution of the price trajectories of the CSI 300 index is shown in Figure 1. In the Chinese stock bubble of 2007, the CSI 300 index soared 573.2% from 873 on December 1, 2015 to 5877.2 on October 16, 2007, and then the CSI 300 index suffered a more than 70% drop from the historical high during the period from October 2007 to October 2008. The 2015 Chinese Stock Market bubble crashed on June 12, 2015. The CSI 300 index has lost more than 42% from the peak on June 12, 2015 to the bottom on August 26, 2015.

In this study, we adopt the LPPLS methodology to detect the positive and negative bubbles in the Chinese stock market using the daily data on the CSI 300 stock market index from January 2002 through April 2018. This study is the first work in the literature that identifies the existence of bubbles in the Chinese stock market using the daily data of CSI 300 index with the advance bubble detection methodology of LPPLS confidence indicator. To improve the performance of LPPLS confidence indicator, the damping condition of LPPLS model is included in the search space and the stricter filter conditions for the qualification of the valid LPPLS fits are applied in

this study. This study also presents the additional results about the two "well-known" Chinese stock market bubbles: the Chinese stock bubble of 2007 and the 2015 Chinese Stock Market bubble to provides more detailed information of LPPLS methodology for detecting the bubbles and their termination.

The paper is organized as follows. Section 2 presents the technical descriptions of all the methods used in this study, including LPPLS model, LPPLS calibration, and LPPLS confidence indicator. The empirical analysis of the LPPLS confidence indicator application to the Chinese Stock Market are conducted in Section 3. Section 4 concludes this paper.
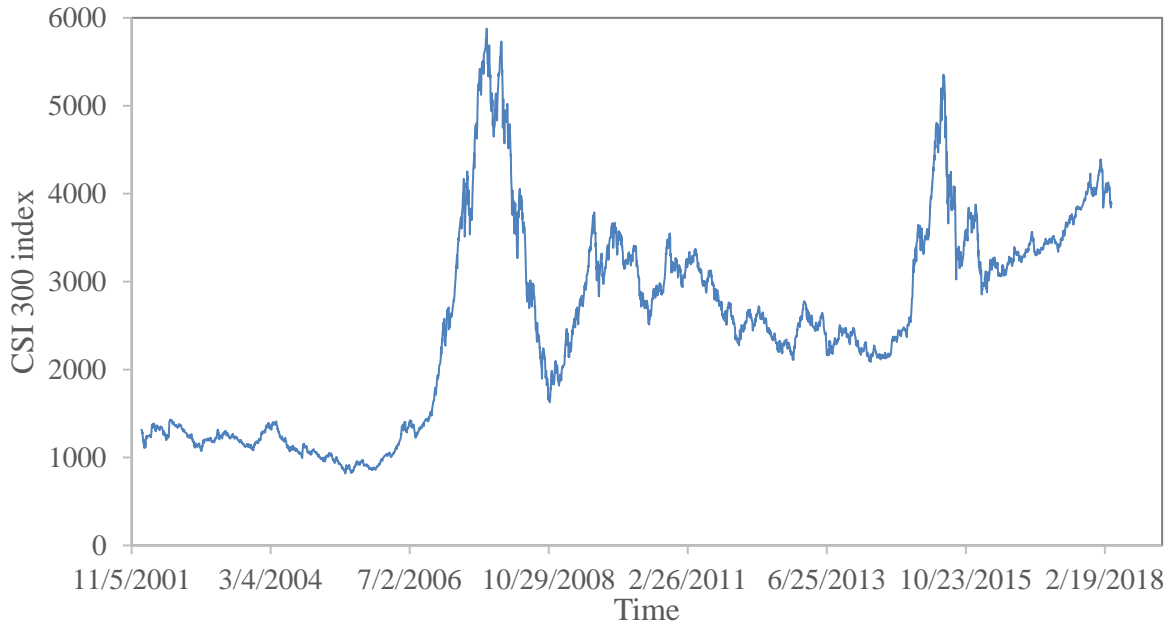

Figure 1. Evolution of the price trajectories of the CSI 300 index over the time interval of this analysis

2. Methodology

2.1 The Log-Periodic Power Law Singularity (LPPLS) Model

The LPPLS Model, originally called as the Johansen-Leoit-Sornette (JLS) model, is systematically explained by Johansen et al. (2000). In this section, the derivation of the LPPLS model is recalled based on the reference (Johansen et al., 2000). The LPPLS model is based on a risk neutral rational agent with rational expectations with ignoring the arbitrage, dividends, the interest rate, risk aversion, information asymmetry and the market clearing condition. The rise of the expected asset price must compensate for the expected risk, implying the asset price follows a martingale process, i.e. $E_t[p(t')] = p(t), \forall t' > t$, where $p(t)$ denotes the asset price at the time t and $E_t[\cdot]$ represents the conditional expectation given all previous data before and up to the time t. The occurrence of a crash or change can be modelled as a discontinuous jump process j with the value of 0 before the crash and 1 after the crash occurs at the critical time t_c. Due to the random nature of the crash occurrence, the t_c can be modeled by the cumulative distribution function $Q(t)$, the probability density function $q(t) = dQ/dt$, and a crash hazard rate $h(t) =$

4

$q(t)/[1 - Q(t)]$, which is the probability per unit of time of the crash taking place in the next instant conditional on the fact that it has not yet happened. Because $h(t)dt$ is the probability that the crash occurs between t and $t + dt$ given the crash has not yet happened, the expectation of dj can be determined as: $E_t[dj] = 1 \times h(t)dt + 0 \times (1 - h(t)dt) = h(t)dt$. For simplicity, it is assumed that the asset price falls during a crash at a fixed percentage $k \in (0,1)$. Then, the asset price dynamics before the crash occurring can be given by:

$$dp = \mu(t)p(t)dt - kp(t)dj + \sigma(t)p(t)dW \quad \Rightarrow$$

$$E_t[dp] = \mu(t)p(t)dt - kp(t)E_t(dj) + \sigma(t)p(t)E(dW) = \mu(t)p(t)dt - kp(t)h(t)dt \quad (1)$$

where $\mu(t)$ is the time-dependent return, $\sigma(t)$ is the volatility, dW is the infinitesimal increment of a standard Wiener process with zero mean and variance equal to dt. Under the assumption of the no arbitrage condition and rational expectations, the conditional expectation of the price dynamics $E_t[dp]$ is zero as the price process satisfies the martingale condition, so that $\mu(t)p(t)dt - kp(t)h(t)dt = 0$, yielding $\mu(t) = kh(t)$ that implies the return $\mu(t)$ is proportional to the risk of crash quantified by its crash hazard rate $h(t)$. Due to the existence of noise traders with herding behavior, the $h(t)$ playing a role of the driver of the bubble grows progressively on the no-arbitrage condition, leading to an instantaneous return $\mu(t)$ that grows together with $h(t)$ in order to remunerate investors who have willingness to invest in a risk asset (Sornette et al., 2015). Substituting the equality of the return $\mu(t)$ into Equation (1), the asset price dynamics, conditioned on the fact that no crash occurs, can be simplified as:

$$dp = kh(t)p(t)dt - kp(t) \times 0 + \sigma(t)p(t)dW = kh(t)p(t)dt + \sigma(t)p(t)dW \quad \Rightarrow$$
$$\frac{dp}{p(t)} = kh(t)dt + \sigma(t)dW \quad (2)$$

Its conditional expectation leads to $E_t[dp/p(t)] = kh(t)dt$ with the solution as follows:

$$E_t\left[ln\left[\frac{p(t)}{P(t_0)}\right]\right] = k \int_{t_0}^{t} h(t')dt' \quad (3)$$

To model the behavior the asset price before a crash, it is necessary to specify the key variable: the crash hazard rate $h(t)$, which quantifies the probability that a large number of agents will suppose the same sell position simultaneously resulting in the imbalance of financial market unless the asset price decrease substantially. In order to capture the imitative local micro-interactions, Johansen et al. (2000) proposed a model in which each agent i can have only two possible states s_i: "buy" ($s_i = +1$) or "sell" ($s_i = -1$). The state of agent i at a given point in time is given by the following Markov process:

$$s_i = \text{sign}\left(K \sum_{k \in N(i)} s_j + \sigma \varepsilon_i\right) \quad (4)$$

5

where sign($\cdot$) represents the sign function with the value of +1 (-1) for positive (negative) numbers, K is a positive constant qualifying the coupling strength between agents, $N(i)$ is the number of agents who influences agent i, s_j is the current state of agent s_j, σ is the tendency toward idiosyncratic behavior for all agents, ε_i is random draw from a standard normal distribution. The order K/σ determines the outcome if order in the network wins. When order wins, the agents will imitate their close neighbors, resulting in the spreading imitation in whole network, and eventually causing a crash. When a crash takes place, the K will approaches the critical value K_c, and all the agents will have the same state, either +1 or -1.

As also mentioned by Blanchard (1979), the higher the probability of a crash, the faster the price before the occurrence of crash should grow to satisfy the martingale condition, so that the investor induced to hold an asset with increasing risk of crash should be compensated by the higher return chance. At this point, Johansen et al. (2000) assumes that the behavior of the variable close to a critical point can be described by a power law, and the susceptibility of the critical system qualifying the degree of sensitivity of a system subjected to an external perturbation is expressed as $\chi \approx A(K_c - K)^{-\gamma}$ where A is a positive constant (=7/4 for the bi-dimensional Ising model) and γ is the positive critical exponent of the susceptibility. The susceptibility χ describes the chance that a large group of agents suddenly reach an agreement given the existent external influence in the network. In the 2-d Ising model, the interconnection of investors is only considered in an uniform way. However, in real modern financial market constituted of an ensemble of the investors which substantially differs in size ranging from individuals to gigantic professional funds, the interacting investors are organized inside a hierarchical network, where they locally influence each other at different levels. In order to appropriately represent the current structure of financial market, Johansen et al. (2000) proposed a Hierarchical Diamond Lattice (HDL) to model the rational imitation of the investors. The structure of HDL is created by starting with a pair of linked traders and then substituting each link with a new diamond with four links and two new nodes diagonally opposite each other. This operation is repeated until the stopping criterion is satisfied. After n iterations, there will be $\frac{2}{3}(2 + 4^n)$ traders and 4^n links among them. The HDL has the similar basic properties with the rational imitation model based on the bi-dimensional network. The only crucial difference is the that the critical exponent of the susceptibility γ can be a complex number in HDL. A version of HDL was solved Derrida et al. (1983) and the general solution is given by:

$$\chi \approx \text{Re}\left[A_0(K_c - K)^{-\gamma} + A_1(K_c - K)^{-\gamma+i\omega} + \cdots\right]$$
$$\approx A_0'(K_c - K)^{-\gamma} + A_1'(K_c - K)^{-\gamma}\cos[\omega \ln(K_c - K) + \phi] + \cdots \qquad (5)$$

where A_0', A_1', ω and ϕ are real numbers, and Re[$\cdot$] denotes the real part of a complex number. The oscillations correct the pure pow law singularity, accounting for the underlying approximate discrete scale invariance of the financial price dynamics (Sornette, 1998). The oscillations are called "log-periodic" because they are periodic in logarithm of the variable $(K_c - K)$ and the angular log-frequency is $\frac{\omega}{2\pi}$. When the oscillations reach the critical time, their frequency explodes, leading to the accelerating oscillations. Accounting for this mechanism, the crash hazard rate is assumed to behave in a similar way to the susceptibility in the neighborhood of the critical point. Therefore, the hazard rate has the following behavior:

$$h(t) \approx \alpha(t_c - t)^{m-1}(1 + \beta \cos[\omega \, ln(t_c - t) + \phi]) \qquad (6)$$

where $\alpha, \beta, \phi, m, \omega$ and t_c are parameters. This expression of the hazard rate shows that the risk of a crash per unit of time increase drastically when the interaction among investors increase before the occurrence of crash. Substituting the hazard rate in Equation (6) into the solution of the conditional expectation of the asset price in Equation (3), we get the evolution for the asset price before a crash, which is known as the Log Periodic Power Law Singularity (LPPLS) formula:

$$\text{LPPLS}(t) \equiv E_t[\ln p(t)] = A + B(t_c - t)^m\{1 + C \cos[\omega \, ln(t_c - t) + \phi]\} \quad (7)$$

where $A > 0$ is the expected value of the $\ln p(t_c)$ at the critical time t_c, $B = -k\alpha/m < 0$ for a positive bubble is the decrease in $\ln p(t)$ over the time unit if C is close to zero before a crash, $C = -k\alpha\beta/\sqrt{m^2 + \omega^2}$ is the proportional magnitude of the oscillations around the power law singular growth, $0 < m < 1$ is the exponent of the power law growth, ω is the angular log-frequency of the oscillation during a bubble, and $0 < \phi < 2\pi$ is a phase parameter. The Equation (7) is the fundamental equation of LPPLS formula describing the evolution of asset prices before a crash occurs and it has been proposed in different forms in several papers, e.g., Sornette (2003) and Lin et al. (2014).

Two common remarkable characteristics of the most speculative bubble are well documented in both developed and emerging stock markets, i.e., (1) a faster-than-exponential (or super-exponential) growth of the stock market, which ends when the bubble regime changes and (2) accelerating oscillations when approaching to the critical time of the bubble (Johansen & Sornette, 1999; Johansen et al., 1999; Sornette & Johansen, 2001). Both the significant features can be well captured by the LPPLS model in Equation (7). The feature of super-exponential growth of the bubble can be described by the power law singular component $A + B(t_c - t)^m$, which embodies the positive feedback mechanism of a bubble development. To ensure the super-exponential growth, it is required that $0 < m < 1$. The condition $m > 0$ makes sure that the price remains finite at the critical time t_c, while $m < 1$ expresses that a singularity exists. The positive bubble when the price of asset is arising is characterized by $B < 0$, while the negative bubble when the price is falling is featured by $B > 0$. $A > 0$ ensures the price of asset is positive. The asset price dynamics of antibubble can be obtained by replacing $t_c - t$ by $t - t_c$. The characteristic of accelerating oscillations of the bubble is captured by the component $C (t_c - t)^m \cos[\omega \, ln(t_c - t) + \phi]$, which represents the tension and competition between the value investors and the noise traders resulting in the deviation of the market price around the super-exponential growth in the form of oscillations that are periodic in the logarithm of the time to t_c. The term $C (t_c - t)^m$ describes the fact that the amplitude of the accelerating oscillation is falling to zero at the critical time t_c. The term $\omega \, ln(t_c - t)$ represents the local frequency of the log-periodic oscillations is accelerating to infinite at the critical time t_c. The parameter ϕ is related to the characteristic time unites for the oscillations. It should be noted that the critical time t_c is the most probable time for a change in regime at which the growth rate of the asset price changes. The regime change is often but not necessarily the time of a bubble crash. A change in regime refers to a change from super-exponential growth to an exponential or lower growth with the end of the accelerating oscillations.

2.2 LPPLS calibration

The original LPPLS formula in Equation (7) is consisted of the three linear parameters (A, B, C) and 4 nonlinear parameters (t_c, m, ω, ϕ). A common method of calibration for the LPPLS model in Equation (7) is the ordinary least squares method. The 3 linear parameters (A, B, C) are enslaved in the fitting algorithm to simplify the calibration and then estimated from the solved solutions of the 4 nonlinear parameters (t_c, m, ω, ϕ). However, the calibration of the LPPLS model by minimizing the nonlinear multivariate least squares functions is a non-trivial task because of the relatively large number of parameters and the strong nonlinear structure of the model and the multiple local minima can make the local optimization algorithms getting trapped. The solution for the global minimum may not be correctly discovered even using some metaheuristic methods such as taboo search (Cvijovic & Klinowski, 1995) or genetic algorithm (Jacobsson, 2009). In order to reduce the number of nonlinear parameters and gets rid of the interdependence between the angular log-frequency ω and the phase ϕ, Filimonov and Sornette (2013) proposed transforming the LPPLS formula to reduce the number of nonlinear parameters from 4 to 3 at the cost of increasing the number of linear parameters from 3 to 4. The transforming LPPLS formula is written as:

$$\text{LPPLS}(t) \equiv E_t[\ln p(t)] = A + B(t_c - t)^m + C_1(t_c - t)^m \cos[\omega \, ln(t_c - t)]$$
$$+ C_2(t_c - t)^m \sin[\omega \, ln(t_c - t)] \qquad (8)$$

where $C_1 = C cos\phi$ and $C_2 = C sin\phi$. The phase ϕ is contained by $C_1 \ and \ C_2$. The cost function in the least-squares method can be described as:

$$F(t_c, m, \omega, A, B, C_1, C_2) = \sum_{i=1}^{N} [\ln p(\tau_i) - A - B(t_c - \tau_i)^m - C_1(t_c - \tau_i)^m \cos(\omega \ln(t_c - \tau_i))$$
$$- C_2(t_c - \tau_i)^m \sin(\omega \ln(t_c - \tau_i))]^2 \qquad (9)$$

where $\tau_1 = t_1$ and $\tau_N = t_2$. Subordinating the 4 linear parameters $A, B, C_1 \ and \ C_2$ to the 3 nonlinear parameters t_c, m, ω, the nonlinear optimization problem is: $\{\hat{t}_c, \widehat{m}, \widehat{\omega}\} = arg \min_{t_c, m, \omega} F_1(t_c, m, \omega)$. The $F_1(t_c, m, \omega) = \min_{A, B, C_1, C_2} F_1(t_c, m, \omega, A, B, C_1, C_2)$. The linear parameters can be solved by:

$$
\begin{pmatrix}
N & \sum f_i & \sum g_i & \sum h_i \\
\sum f_i & \sum f^2_i & \sum f_i g_i & \sum f_i h_i \\
\sum g_i & \sum f_i g_i & \sum g^2_i & \sum h_i g_i \\
\sum h_i & \sum f_i h_i & \sum g_i h_i & \sum h^2_i
\end{pmatrix}
\begin{pmatrix}
\hat{A} \\
\hat{B} \\
\hat{C}_1 \\
\hat{C}_2
\end{pmatrix}
=
\begin{pmatrix}
\sum \ln p_i \\
\sum f_i \ln p_i \\
\sum g_i \ln p_i \\
\sum h_i \ln p_i
\end{pmatrix}
\qquad (10)
$$

where $f_i = (t_c - \tau_i)^m$, $g_i = (t_c - \tau_i)^m cos(\omega \ln(t_c - \tau_i))$, and $h_i = (t_c - \tau_i)^m sin(\omega \ln(t_c - \tau_i))$. The cost function of the transforming LPPLS model is characterized by good smooth properties, leading to the dramatical reduction of the complexity and tremendous improvement of stability in the fitting procedure, so that the metaheuristic methods are no longer necessary, and the fitting efficiency significantly increases. In this study, the covariance matrix adaptation

8

evolution strategy (CMA-ES) is adopted to search the best estimation of the three nonlinear parameters (t_c, m, ω) by minimizing the residuals (the sum of the squares of the differences) between the fitted LPPLS model and the observed price time series. The CMA-ES proposed by (Hansen et al., 1995) rates among the most successful evolutionary algorithms for real-valued single-objective optimization and is typically applied to difficult nonlinear non-convex black-box optimization problems in continuous domain and search space dimensions between three and a hundred. The parallel computing is applied to expedite the fitting process and remarkably reduce the computation time.

2.3 LPPLS confidence indicator

The LPPLS confidence indicator was introduced by Sornette et al. (2015) and is also one of key indicators in Financial Crisis Observatory (FCO) at ETH Zurich. The LPPLS confidence indicator is defined as the fraction of fitting windows in which the LPPLS calibrations satisfy the specified filter conditions. It is used to measure sensitivity of observed bubble pattern to the time interval between the end time and the start time in the fitting windows ($dt = t_2 - t_1$). A large value of the LPPLS confidence indicator indicates the more reliable of the LPPLS pattern. A small value of the indicator signals a possible fragility since the LPPLS pattern is presented in a few fitting windows.

A LPPLS confidence indicator for a specified data point t_2 (corresponding to a fictitious "present") can be obtained by the following five steps: (1) create the fitting time windows by shrinking in terms of t_1 moving toward the fixed endpoint t_2 with a step of dt_1, (2) determine the search space in the calibration procedure, (3) calibrate the LPPLS model for each fitting time window, (4) specify the filter conditions and summarize the number of fitting windows in which satisfy the specified filter condition, and (5) calculate the LPPLS confidence indicator from dividing the number of time windows satisfying the specified filter condition by the total number of the fitting windows.

In this study, the length of the shrinking time windows $dt = t_2 - t_1$ is adopted to decrease from 750 trading days to 50 trading days in steps of 5 trading days. Thus, 141 fitting windows are obtained for each t_2. In order to minimize fitting problems and address the sloppiness of the model, we adopt the following search space:

$$m \in [0,1], \omega \in [1, 50], t_c \in \left[t_2, t_2 + \frac{t_2 - t_1}{3} \right], \frac{m|B|}{\omega \sqrt{C_1^2 + C_2^2}} \geq 1 \qquad (11)$$

The condition $t_c \in [t_2, t_2 + (t_2 - t_1)/3]$ ensure that the predicted critical time t_c should be after the endpoint t_2, and should not be too far away from the t_2 since the predictive capacity degrades far beyond t_2 (Jiang et al., 2010). The Damping parameter $m|B|/\left(\omega \sqrt{C_1^2 + C_2^2} \right) \geq 1$ under the condition that the crash hazard rate $h(t)$ is non-negative by definition (Bothmer & Meister, 2003). After calibrating the LPPLS models, the solutions should be filtered by the stricter conditions:

$$m \in [0.01, 0.99], \omega \in [2, 25], t_c \in \left[t_2, t_2 + \frac{t_2 - t_1}{5} \right], \frac{\omega}{2} \ln \left(\frac{t_c - t_1}{t_c - t_2} \right) \geq 2.5,$$

$$\max \left(\frac{|\hat{p}_t - p_t|}{p_t} \right) \leq 0.15, \; p_{lomb} \leq \alpha_{sign}, \ln(\hat{p}_t) - \ln(p_t) \sim AR(1) \qquad (12)$$

The filter conditions are derived from the empirical evidence gathered in investigations of previous bubbles (Jiang et al., 2010; Sornette et al., 2015) and are the stylized features of LPPLS model. The condition for the number of oscillations (half-periods) of the log-periodic component $(\omega/\pi)\ln[(t_c - t_1)/(t_c - t_2)] \geq 2.5$ is implemented to distinguish a genuine log-periodic signal from one that could be generated by noise (Huang et al., 2000). The condition of the maximum relative error $\max(|\hat{p}_t - p_t|/p_t) \leq 0.15$ ensure the fitted price of an asset $\hat{p}_t$ should be not too far from the actual asset price p_t. The condition $P_{lomb} \leq \alpha_{sig}$ ensures the logarithm-periodic oscillations in the fitting the logarithm of prices to the LPPLS model by applying the Lomb spectral analysis for the series of detrended residual $r(t) = (t_c - t)^{-m}(\ln[p(t)] - A - B(t_c - t)^m)$ (Sornette & Zhou, 2002). The probabilities that the maximum peak occurred by chance P_{lomb} is less than the specified significant level α_{sig}, indicating the existence the logarithm-periodic oscillations in the fitting LPPLS model. The $\ln(\hat{p}_t) - \ln(p_t) \sim AR(1)$ condition ensures that the LPPLS fitting residuals can be modeled by a mean-reversal Ornstein-Uhlenbeck (O-U) process when the logarithmic price in the bubble regime is attributed to a deterministic LPPLS component (Lin et al., 2014). Since the test for the O-U property of LPPLS fitting residuals can be translated into an AR(1) test for the corresponding residuals, both the Phillips-Perron unit-root test and Dickey-Fuller unit-root test are used to check the O-U property of LPPLS fitting residuals. The 10% significant level is adopted for the tests in this study. Only the calibrations satisfying filter conditions given in Equation (12) are considered valid and the others are discards.

3. Empirical analysis

In the following two subsections, this paper presents the detection analysis of the Chinese Stock Market bubble using the LPPLS confidence indicator described in Section 2, and the post-mortem analysis of Chinese stock market bubbles.

3.1 LPPLS bubble identification

In this study, we collected the daily data on the Shanghai Shenzhen CSI 300 stock market index from January 4, 2002 through April 2, 2018 for 3939 observations. These data come from the Bloomberg Financial Database. We adopted the length of the shrinking time windows $t_2 - t_1$ decreasing from 750 trading days to 50 trading days in steps of 5 trading days and the endpoint t_2 moving from March 1, 2005 through April 2, 2018 in steps of 5 trading days to generate 638 t_2. Since there are 141 fitting windows for each t_2, the total 89,958 fitting windows are generated in this study. The value of the LPPLS confidence indicator at a given time t_2 is causal since it is estimated only based on data prior to that time. The LPPLS confidence indicators for a series of varying t_2 provide useful insights into the time development of the bubble signal.

Both the positive and negative bubbles in Chinese stock market are detected in this study. The positive bubbles are associated with upwardly accelerating price increases, and are susceptible to regime changes that take the form of crashes or volatile sideway plateaus, while the negative bubbles are associated with downwardly accelerating price increases, and are susceptible to regime changes that take the form of rallies or volatile sideway plateaus. Figures 2-4 show the LPPLS confidence indicator for positive bubbles in red together with the CSI 300 index in blue from 3/1/2005 to 4/2/2018. Figures 5–7 presents the LPPLS confidence indicator for negative bubbles in red together with the CSI 300 index in blue from 3/1/2005 to 4/2/2018. These figures indicate the confidence level of the observed LPPLS bubble pattern. The LPPLS confidence indicator marks bubble by measuring the sensitivity of the bubble pattern to the selected starting time. When the LPPLS bubble pattern exists in more time windows for a given "present" time, the LPPLS confidence indicator has a higher value. The value of LPPLS confidence indicator can be up to one if the bubble pattern exists in most of the analyzed time windows and presents almost no sensitivity to the choice of the time windows. When the bubble pattern is only observed in a few of time windows, the LPPLS confidence indicator may have a value close to zero which indicates the over-fitting risk and needs careful consideration for the results.

As shown in Figure 1, a cluster of bubble patterns are detected from January 2007 to October 2007, indicating a positive bubble may have been born and developed over time since the value of LPPLS confidence indicator has a dramatic increase. Multiples peaks of the confidence indicator with large value are observed from May to October 2007 and the highest value of indicator is up to 0.27, representing the observed bubble signals are reliable and the regime change may occur in form of crash or volatile sideway plateaus, so that the growth rate of CSI 300 index would be changed from super-exponential growth to an exponential or lower growth. The diagnostic of the presence of possible bubble is confirmed by the Chinese stock bubble of 2007, in which the CSI 300 index reaches the historical peak of 5877.2 on October 16, 2007. The Chinese stock bubble of 2007 corresponded to an approximate 314.3% growth in just one year. After the bubble crashed, the CSI 300 index lost more than 70% from the historical high during the period from October 2007 to October 2008.

In Figure 2, a cluster of positive bubble signals are diagnosed from March to August in 2009 and the LPPLS confidence indicator reaches the peak of 0.10 on July 28, 2009. During the 2009 Chinese stock bubble, the CSI 300 index has risen more than 130% from November 4, 2008 to August 3, 2009. Following the bubble crash, the index fell by over 25% in August 31, 2009. Thus, the bubble patterns indeed detect the development and crash of the 2009 Chinese stock bubble.

From Figure 3, two clusters of positive bubble signals can be seen from November 7, 2014 to January 13, 2015, and from April 29, 2015 to June 11, 2015, separately. On June 11, 2015, the LPPLS confidence indicator reaches the top of 0.118, indicating the high risk of regime change. The first diagnostic of a "bubbly" CSI 300 index occurred November 2014 and persisted until January 2015, when a change of regime indeed occurred. Afterwards, the bubble pattern reappears and becomes stronger on April 2015 and persisted until the eventual burst of the 2015 Chinese Stock Market bubble. The CSI 300 index has suffered more than 42% drop from the peak on June 12, 2015 to the bottom on August 26, 2015. It is noted that some bubble signals

appear from October to November in 2016 and September in 2017, implying the potential bubble may be emerging and a significant change of regime may occur in the future.

Figure 4 shows two clusters of negative bubbles. The first one is from March 22, 2005 to August 2, 2005 with the peak of 0.085 occurs at July 12, 2005. This cluster captures the Chinese stock market negative bubbles in 2005. The CSI 300 index starts to fall from 1410.43 on April 9, 2004 to the historical lowest value 824.1 on July 11, 2005. The second cluster starts on April 8, 2008, and end on November 10, 2008 with the confidence indicator value 0.064. The CSI 300 index falls 4104 points from January 14, 2008 to November 4, 2008 (71.6% decline).

There are two main clusters of negative bubbles in Figure 5. The first cluster starts on September 14, 2011 and culminates on January 13, 2012 with the start of the rebound. In the negative bubble, the CSI 300 index has suffered more than 32% drop from April 15, 2011 to January 5, 2012. The second cluster is from July 30, 2012 to December 10, 2012. The LPPLS confidence indicator surges to 0.156 on December 10, 2012, followed by the rebound of the price. The CSI 300 index fell from 2717.8 on May 7, 2012 to 2108.9 on December 3, 2012 and then rebound to 2673.3 on February 28, 2013.

Figure 6 presents two clusters of negative bubbles from March 28, 2014 to July 17, 2014 and from August 21, 2015 to September 8, 2015. As shown in Figure 1, the CSI 300 index has a valley from March to July 2014, and then rises dramatically in the next one year to form the well-known 2015 Chinese Stock Market bubble. The 2015 Chinese Stock Market bubble dropped to the bottom on September 8, 2015 and then the regime changes in form of rebound.
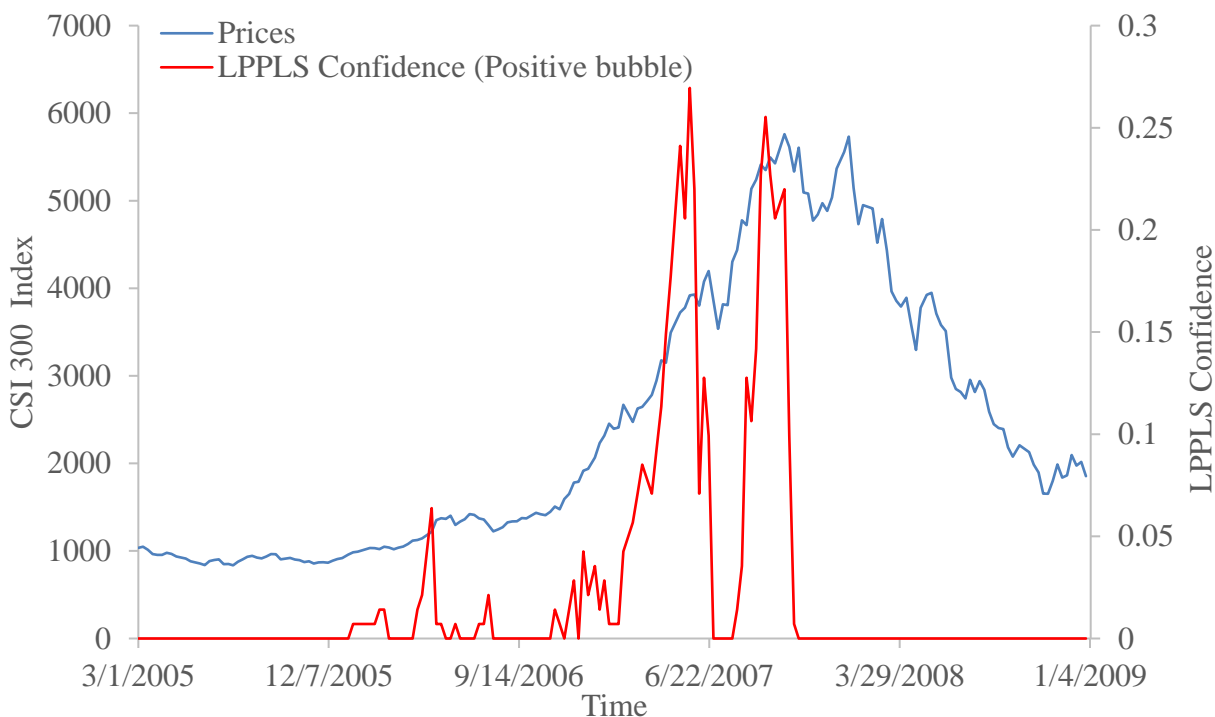

Figure 1. LPPLS confidence indicator for positive bubbles in red (right scale) together with the CSI 300 index in blue (left scale) from 3/1/2005 to 12/29/2008
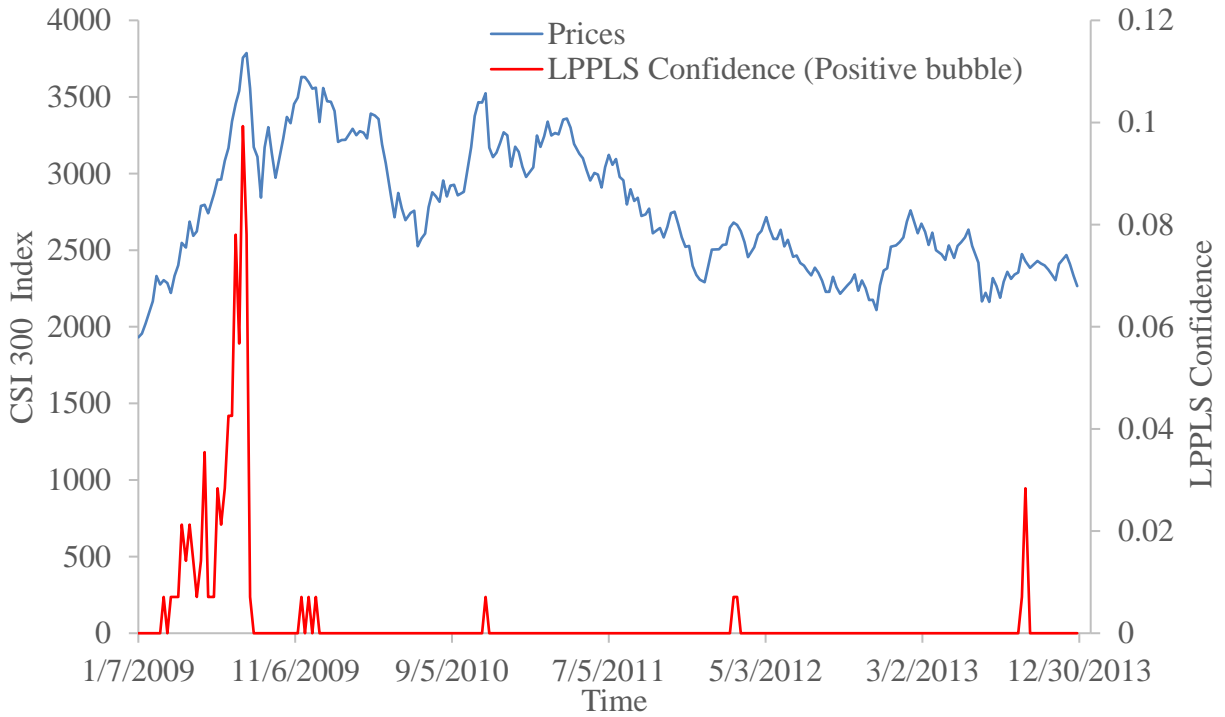
Figure 2. LPPLS confidence indicator for positive bubbles in red (right scale) together with the CSI 300 index in blue (left scale) from 1/7/2009 to 12/26/2013
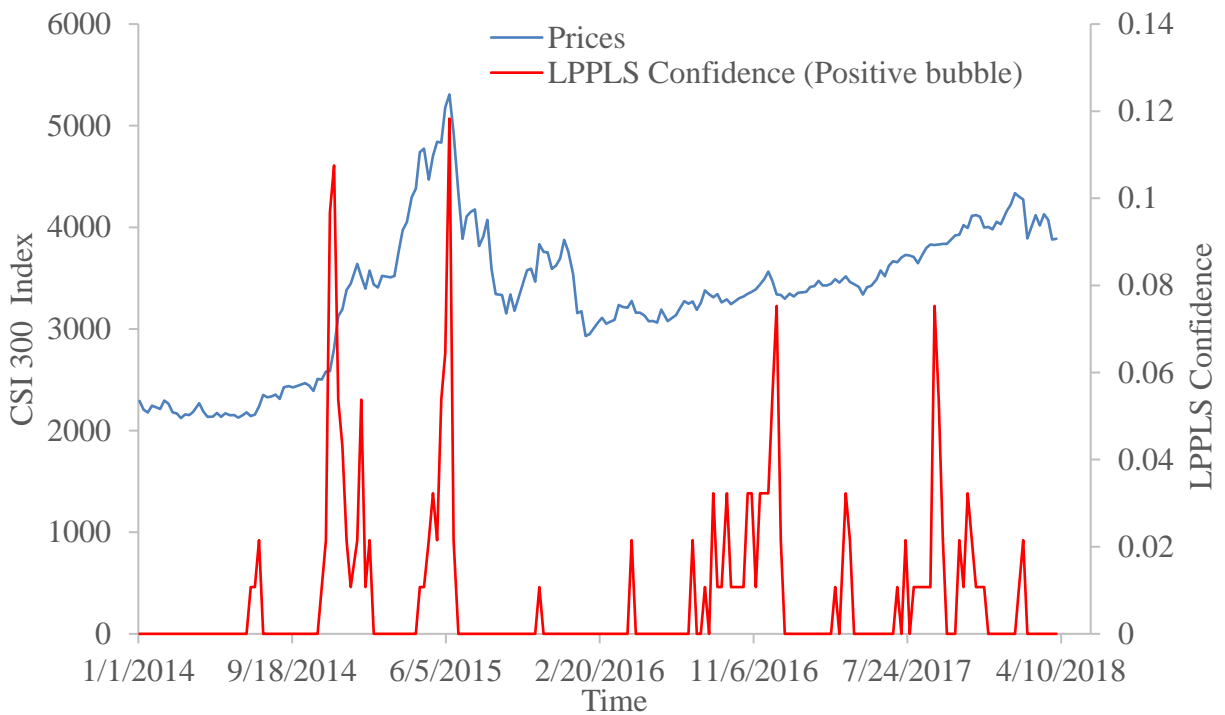

Figure 3. LPPLS confidence indicator for positive bubbles in red (right scale) together with the CSI 300 index in blue (left scale) from 1/3/2014 to 4/2/2018
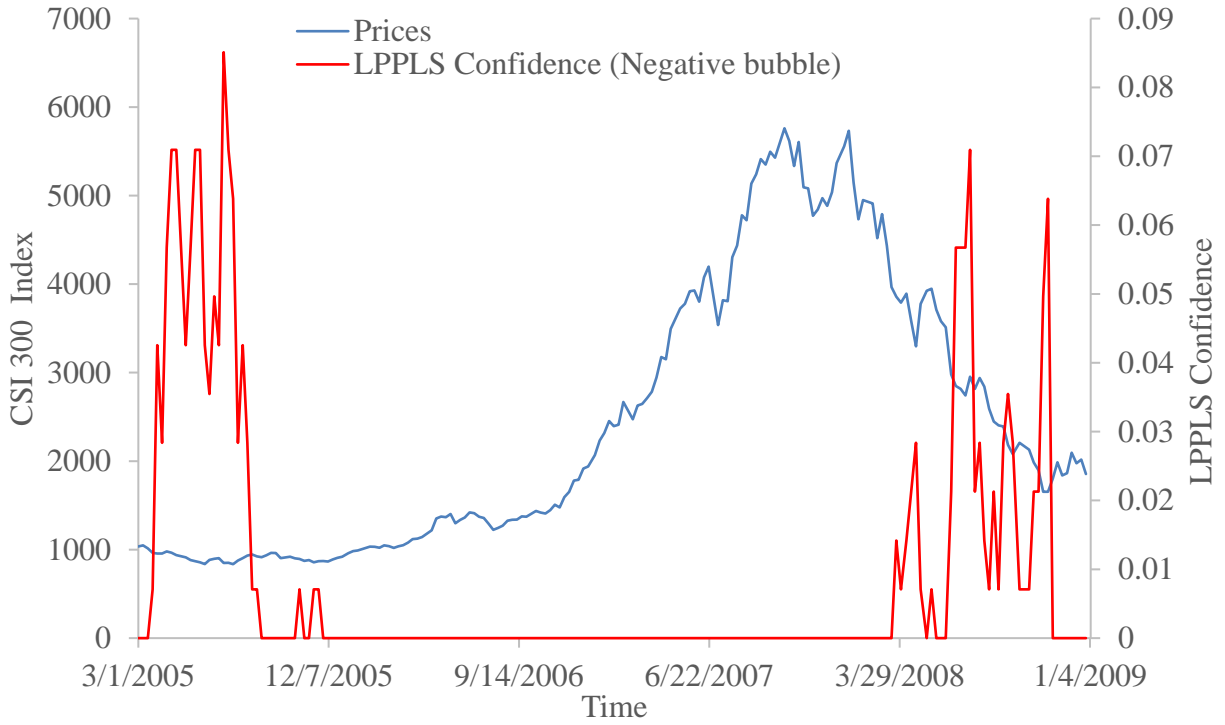
Figure 4. LPPLS confidence indicator for negative bubbles in red (right scale) together with the CSI 300 index in blue (left scale) from 3/1/2005 to 12/29/2008
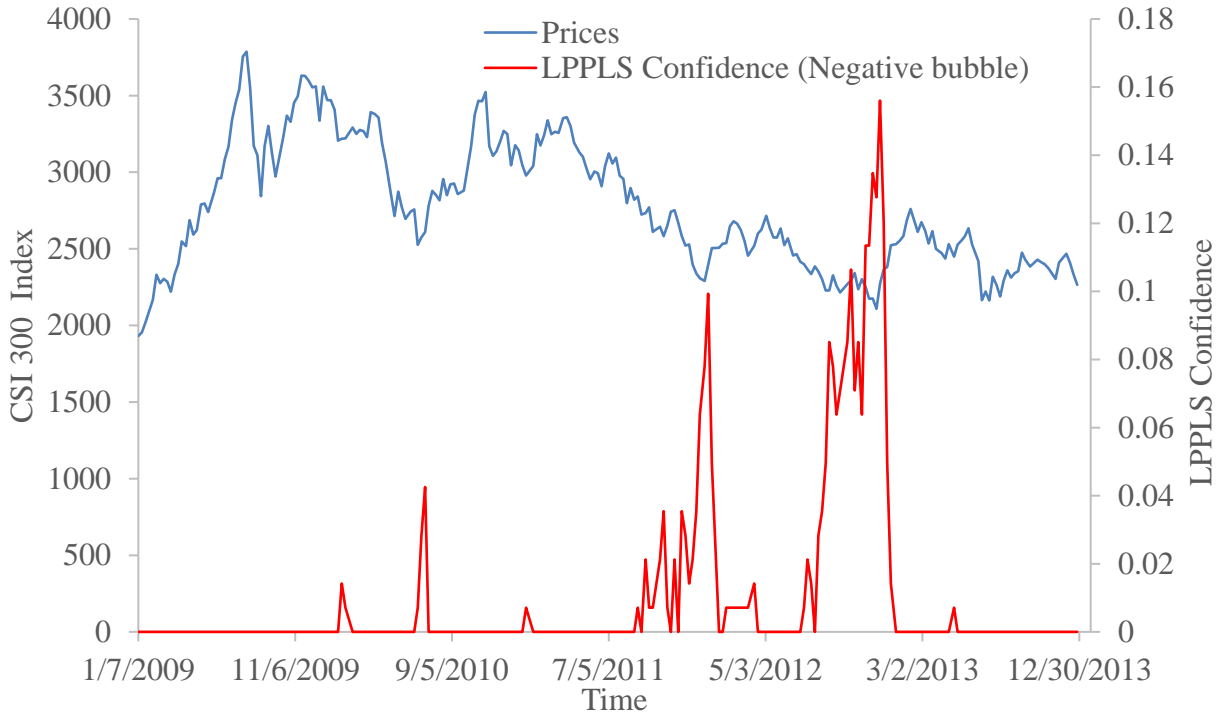

Figure 5. LPPLS confidence indicator for negative bubbles in red (right scale) together with the CSI 300 index in blue (left scale) from 1/7/2009 to 12/26/2013
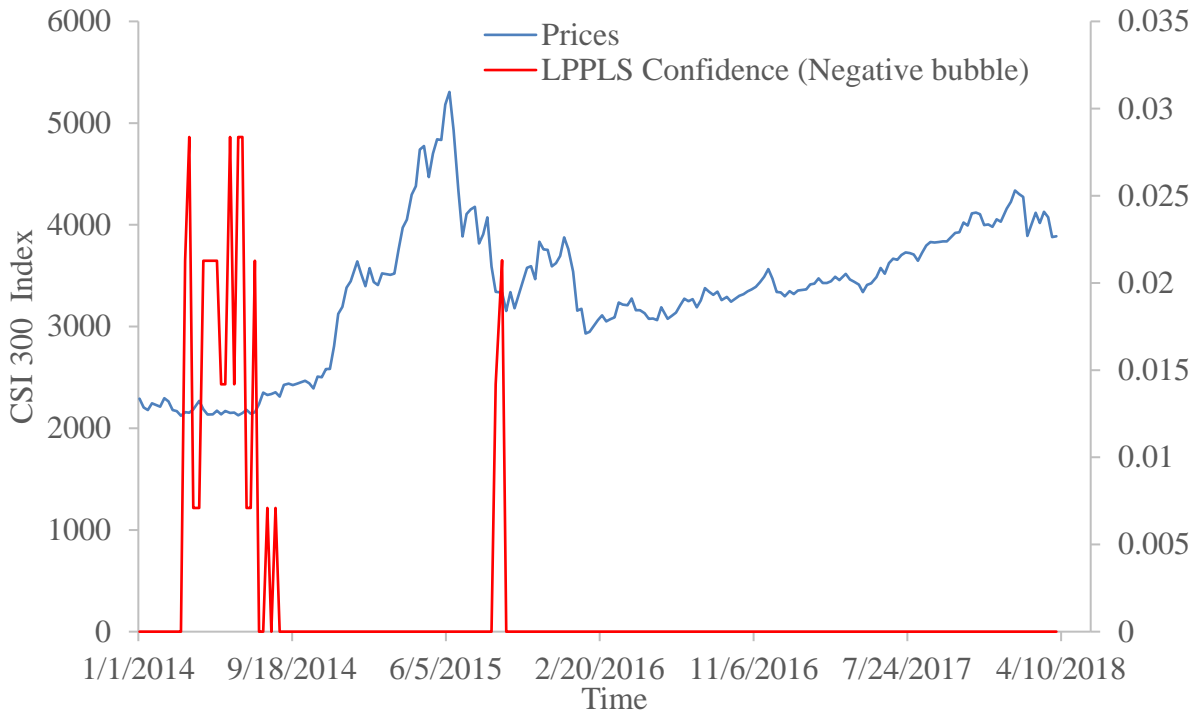
14

Figure 6. LPPLS confidence indicator for negative bubbles in red (right scale) together with the CSI 300 index in blue (left scale) from 1/3/2014 to 4/2/2018

3.2 Post-mortem analysis for the bubbles

This section presents the additional results about the two "well-known" Chinese stock market bubbles: the Chinese stock bubble of 2007 and the 2015 Chinese Stock Market bubble. This analysis provides more detailed information of LPPLS methodology for detecting the bubbles and their termination.

Figure 7 shows the probability density distribution of predicted t_c's as well as estimated beginning t_1's for the Chinese stock bubble of 2007. This is obtained by scanning over 141 t_1's from a maximum 750 trading days to a minimum of 50 trading days in steps of 5 trading days prior to each the end t_2 in which the fitting windows passed the filter conditions in Equation (12) are collected, and repeating this procedure for different t_2 in steps of 5 trading days, and eventually generating the probability density distribution by statistical analysis on the chosen fitting windows. The ranges of t_1 and t_2 for the 220 selected fitting windows are from August 30, 2005 to June 28, 2007 and from August 23, 2007 to October 11, 2007, respectively. The optimal values for the bubble starting date t_1 are represented by the probability density distribution $pdf(t_1)$ in green. It can be seen that the $pdf(t_1)$ is concentrated in the time interval where the CSI 300 index starts to super-exponentially accelerate. This allow us to determine the beginning of the Chinese stock bubble of 2007 as early as August 30, 2005. The forecasted critical time t_c depicted by the probability density distribution $pdf(t_c)$ in red present a strong probability measure at the time of the crash.

15

As shown in Figure 7, the 20%/80% and 5%/95% quantile range of values of the crash dates t_c for the Chinese stock bubble of 2007 are from September 21, 2007 to October 22, 2007, and from August 30, 2007 to November 12, 2007, respectively. The observed market peak date for the CSI 300 index is October 16, 2007, which lies in the quantile ranges of the predicted crash dates t_c fitted based on data before the actual stock market crash. Figure 8 also presents three typical fitting examples are corresponding to t_1= 18 October 2005 and t_2= 13 September 2007, t_1= 20 November 2006 and t_2= 11 October 2007, and t_1= 19 April 2007 and t_2= 20 September 2007, which represent the different time scale windows, respectively.
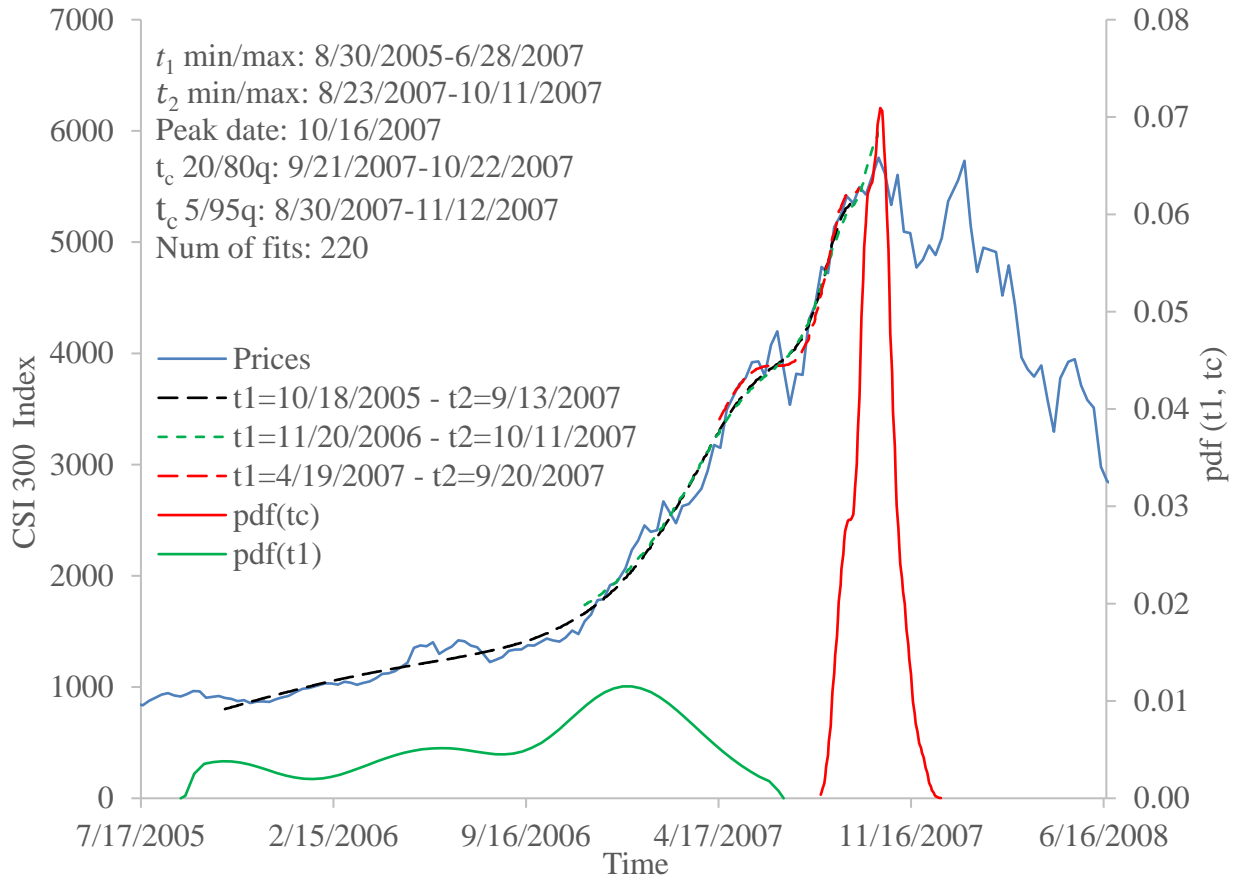

Figure 7. The probability density distributions $pdf(t_1, t_c)$ for the Chinese stock bubble of 2007 (right scale) together with the CSI 300 index in blue (left scale) from 7/17/2005 to 6/16/2008

The probability density distribution of predicted t_c's as well as estimated beginning t_1's for the 2015 Chinese Stock Market bubble is shown in Figure 8. The ranges of t_1 and t_2 for the 48 selected fitting windows are from February 7, 2015 to March 17, 2015 and from April 22, 2015 to June 11, 2015, respectively. The probability density distribution $pdf(t_1)$ in green presents the optimal values for the bubble starting date t_1. It is observed that the $pdf(t_1)$ is negative skew and the mass of the distribution is concentrated on the right of the figure where the CSI 300 index has a super-exponentially growth. From Figure 8, we can see that 2015 Chinese Stock Market bubble began as early as February 7, 2015. The probability density distribution $pdf(t_c)$ in red presents the predicted critical time t_c.

Figure 8 shows that the 20%/80% and 5%/95% quantile range of values of the crash dates t_c for the 2015 Chinese Stock Market bubble are from June 10, 2015 to July 22, 2015, and from May 27, 2015 to August 6, 2015, respectively. The observed market peak date for the CSI 300 index (June 12, 2015) lies in the quantile ranges of the predicted crash dates t_c fitted based on data before the actual stock market crash. Figure 8 also illustrates three typical fitting examples are corresponding to $t_1 = 14$ April 2014 and $t_2 = 14$ May 2015, $t_1 = 21$ November 2014 and $t_2 = 14$ May 2015, and $t_1 = 20$ January 2015 and $t_2 = 4$ June 2015 to represent the long, median and short time scale windows, respectively.
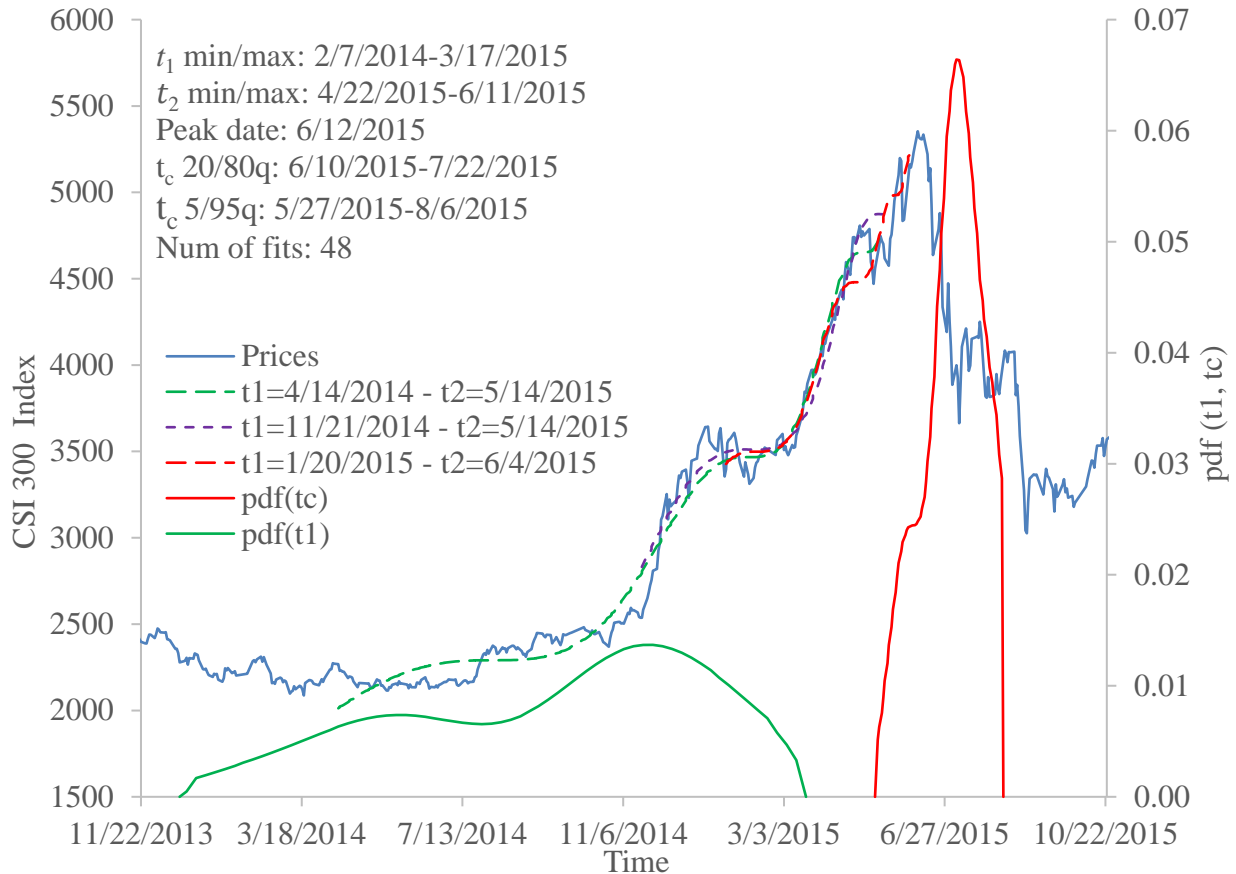

Figure 8. The probability density distributions $pdf(t_1, t_c)$ for the 2015 Chinese Stock Market bubble (right scale) together with the CSI 300 index in blue (left scale) from 11/22/2013 to 10/22/2015

4. Conclusions

In this study, we apply the LPPLS methodology to carry out the early causal identification of positive and negative bubbles in the Chinese stock market using the daily data on the Shanghai Shenzhen CSI 300 stock market index from January 2002 through April 2018. In order to improve the performance of LPPLS confidence indicator, we adjust the search space to account for the damping condition of LPPLS model and implement the stricter filter conditions for the qualification of the valid LPPLS fits by taking account of the maximum relative error, Lomb log-periodic test of the detrended residual, and unit-root tests of the logarithmic residual based on

both the Phillips-Perron test and Dickey-Fuller test. This study is the first of its kinds that identifies the existence of bubbles in the Chinese stock market using the daily data of CSI 300 index with the advance bubble detection methodology of LPPLS confidence indicator. Our analysis shows that the LPPLS detection strategy diagnoses three periods of positive bubbles and four periods of negative bubbles in the period from March 1, 2005 to April 2, 2018. The bubble periods detected by our methodology correspond to well-known historical events, implying the detection strategy based on the LPPLS confidence indicator has an outstanding performance to identify the potential positive and negative bubbles in advance.

This study implements the post-mortem analysis for the two "well-known" Chinese stock market bubbles: the Chinese stock bubble of 2007 and the 2015 Chinese Stock Market bubble. The probability density distribution and quantile range of predicted critical time t_c's as well as estimated beginning t_1's provide a strong probability measure about the time of the crash and the beginning estimation of the bubbles. It is observed that the probability density distribution of the estimated beginning time of bubbles appears to be skewed and the mass of the distribution is concentrated on the area where the bubbles start to have a super-exponentially growth.

It can also be found that the regime shifts and changes are not a rare phenomenon and may occur more frequently in the future. This study shows that it is possible to detect the potential positive and negative bubbles and crashes ahead of time, which provides a prerequisite for limiting the bubble sizes and eventually minimizing the damage from the bubble crash.

References

Blanchard, O. J. (1979). Speculative bubbles, crashes and rational expectations. *Economics letters, 3*(4), 387-389.

Bothmer, H.-C. G. v., & Meister, C. (2003). Predicting critical crashes? A new restriction for the free variables. *Physica A: Statistical Mechanics and its Applications, 320*, 539-547.

Brée, D. S., Challet, D., & Peirano, P. P. (2013). Prediction accuracy and sloppiness of log-periodic functions. *Quantitative Finance, 13*(2), 275-280. doi:10.1080/14697688.2011.607467

Carpenter, J. N., & Whitelaw, R. F. (2017). The Development of China's Stock Market and Stakes for the Global Economy. *Annual Review of Financial Economics, 9*, 233-257.

Demirer, R., Demos, G., Gupta, R., & Sornette, D. (2019). On the predictability of stock market bubbles: evidence from LPPLS confidence multi-scale indicators. *Quantitative Finance, 19*(5), 843-858.

Demos, G., & Sornette, D. (2017). Birth or burst of financial bubbles: which one is easier to diagnose? *Quantitative Finance, 17*(5), 657-675.

Derrida, B., De Seze, L., & Itzykson, C. (1983). Fractal structure of zeros in hierarchical models. *Journal of Statistical Physics, 33*(3), 559-569.

Filimonov, V., Demos, G., & Sornette, D. (2017). Modified profile likelihood inference and interval forecast of the burst of financial bubbles. *Quantitative Finance, 17*(8), 1167-1186.

Filimonov, V., & Sornette, D. (2013). A stable and robust calibration scheme of the log-periodic power law model. *Physica A: Statistical Mechanics and its Applications, 392*(17), 3698-3707. doi:10.1016/j.physa.2013.04.012

Geraskin, P., & Fantazzini, D. (2013). Everything you always wanted to know about log-periodic power laws for bubble modeling but were afraid to ask. *The European Journal of Finance, 19*(5), 366-391.

Hansen, N., Ostermeier, A., & Gawelczyk, A. (1995). *On the Adaptation of Arbitrary Normal Mutation Distributions in Evolution Strategies: The Generating Set Adaptation.* Paper presented at the the Sixth International Conference on Genetic Algorithms, San Francisco, CA.

Huang, Y., Johansen, A., Lee, M., Saleur, H., & Sornette, D. (2000). Artifactual log-periodicity in finite size data: Relevance for earthquake aftershocks. *Journal of Geophysical Research: Solid Earth, 105*(B11), 25451-25471.

Jacobsson, E. (2009). *How to predict crashes in financial markets with the Log-Periodic Power Law.* (Master's Thesis), Stockholm University.

Jiang, Z.-Q., Zhou, W.-X., Sornette, D., Woodard, R., Bastiaensen, K., & Cauwels, P. (2010). Bubble diagnosis and prediction of the 2005–2007 and 2008–2009 Chinese stock market bubbles. *Journal of Economic Behavior & Organization, 74*(3), 149-162.

Johansen, A., Ledoit, O., & Sornette, D. (2000). Crashes as critical points. *International Journal of Theoretical and Applied Finance, 3*(02), 219-255.

Johansen, A., & Sornette, D. (1999). Critical crashes. *Journal of Risk, 12*(1), 91-94.

Johansen, A., Sornette, D., & Ledoit, O. (1999). Predicting financial crashes using discrete scale invariance. *arXiv preprint cond-mat/9903321.*

Li, C. (2017). Log-periodic view on critical dates of the Chinese stock market bubbles. *Physica A: Statistical Mechanics and its Applications, 465*, 305-311.

Lin, L., Ren, R. E., & Sornette, D. (2014). The volatility-confined LPPL model: A consistent model of 'explosive' financial bubbles with mean-reverting residuals. *International Review of Financial Analysis, 33*, 210-225. doi:10.1016/j.irfa.2014.02.012

Sornette, D. (1998). Discrete-scale invariance and complex dimensions. *Physics Reports, 297*(5), 239-270.

Sornette, D. (2003). Critical market crashes. *Physics Reports, 378*(1), 1-98. doi:10.1016/s0370-1573(02)00634-8

Sornette, D., Demos, G., Zhang, Q., Cauwels, P., Filimonov, V., & Zhang, Q. (2015). Real-time prediction and post-mortem analysis of the Shanghai 2015 stock market bubble and crash. *Journal of Investment Strategies, 4*(4), 77–95.

Sornette, D., & Johansen, A. (2001). Significance of log-periodic precursors to financial crashes. *Quantitative Finance, 1*(4), 452-471.

Sornette, D., Woodard, R., Yan, W., & Zhou, W.-X. (2013). Clarifications to questions and criticisms on the Johansen–Ledoit–Sornette financial bubble model. *Physica A: Statistical Mechanics and its Applications, 392*(19), 4417-4428.

Sornette, D., Woodard, R., & Zhou, W.-X. (2009). The 2006–2008 oil bubble: Evidence of speculation, and prediction. *Physica A: Statistical Mechanics and its Applications, 388*(8), 1571-1576.

Sornette, D., & Zhou, W. X. (2002). The US 2000-2002 market descent: How much longer and deeper? *Quantitative Finance, 2*(6), 468-481.

Stiglitz, J. E. (2014). *The lessons of the North Atlantic crisis for economic theory and policy*: Cambridge, MA: MIT Press.

Yan, W. (2011). *Identification and Forecasts of Financial Bubbles.* (Ph.D. Thesis), ETH Zurich.

Yan, W., Woodard, R., & Sornette, D. (2010). Diagnosis and prediction of tipping points in financial markets: Crashes and rebounds. *Physics Procedia, 3*(5), 1641-1657.

Zhang, Q., Zhang, Q., & Sornette, D. (2016). Early warning signals of financial crises with multi-scale quantile regressions of Log-Periodic Power Law Singularities. *PloS one, 11*(11), e0165819.

Zhou, W., & Sornette, D. (2003). 2000–2003 real estate bubble in the UK but not in the USA. *Physica A: Statistical Mechanics and its Applications, 329*(1), 249-263.

Zhou, W., & Sornette, D. (2006). Is there a real-estate bubble in the US? *Physica A: Statistical Mechanics and its Applications, 361*(1), 297-308.

Zhou, W., & Sornette, D. (2008). Analysis of the real estate market in Las Vegas: Bubble, seasonal patterns, and prediction of the CSW indices. *Physica A: Statistical Mechanics and its Applications, 387*(1), 243-260.