Politecnico di Milano

School of Industrial and Information Engineering
Master of Science in Mathematical Engineering

Log Periodic Power Law model for the detection of financial bubbles

Supervisor:
Prof. Daniele Marazzina

Author:
Daniele Bonanomi

Academic Year 2019 - 2020

Daniele Bonanomi
Log Periodic Power Law model for the detection of financial bubbles

Contents

List of Figures

Abstract

The purpose of the present work is to address the problem of financial bubbles, with the aim of developing a detection and forecasting procedure on historical time series.
Bubbles are events governed by irrationality and speculation and their bursts could heavily affect financial markets and the real economy too, with particularly dramatic consequences for the most unwitting investors. This work is an attempt to spread awareness about the risks related with irrational markets.
Specifically, we rely on the Log Periodic Power Law (LPPL) model, originally created by Johansen, Ledoit and Sornette in the late 1990s.
After an overview on the required theoretical background, the first part is dedicated to a deep review of the literature of the LPPL model, with particular focus on latest publications.
The second part is devoted to calibration methods with the purpose of implementing a coherent detection method able to find precursors of bubble crashes. In more detail, we present three different techniques: Ordinary Least Squares (OLS), Generalized Least Squares (GLS) and Maximum Likelihood Estimation (MLE), combined with the Genetic Algorithm, a robust optimization method able to deal with the high non linearity of the problem.
To conclude, we apply this methodology both to post-mortem and real-time case studies on Bitcoin's historical prices (BTC/USD). In particular, we consider the time series between December 2016 and January 2018, where Bitcoin's prices skyrocketed for the first time to almost 20'000 USD, and the most recent one between March 2020 and March 2021, where Bitcoin's price had a huge increase exceeding 60'000 USD.
Results obtained are in line with previous publications.

Keywords: Financial bubbles; Log Periodic Power Law model; Ordinary Least Squares; Generalized Least Squares; Maximum Likelihood Estimation; Genetic Algorithm; Bitcoin.

Sommario

Con questo elaborato si intende affrontare il problema delle bolle finanziarie, con l'obiettivo di sviluppare una procedura di identificazione e predizione su serie storiche. Le bolle finanziarie sono eventi governati da comportamenti irrazionali e speculativi e il loro scoppio può colpire duramente i mercati e l'economia reale, con conseguenze particolarmente drammatiche per gli investitori più inconsapevoli. Questo lavoro è un tentativo di diffondere maggiore consapevolezza sui rischi legati a mercati irrazionali. Nello specifico, adottiamo il "Log Periodic Power Law (LPPL) model", originariamente sviluppato da Johansen, Ledoit e Sornette alla fine degli anni Novanta.
Dopo una overview sulle nozioni teoriche preliminari, la prima parte è dedicata ad un'approfondita revisione della letteratura disponibile sul LPPL model, con particolare attenzione alle pubblicazioni più recenti.
La seconda parte tratta dei metodi di calibrazione, con l'obiettivo di implementare una procedura coerente in grado di individuare segnali di bolla precursori ai crolli. Più nel dettaglio, proponiamo tre diverse tecniche: Ordinary Least Squares (OLS), Generalized Least Squares (GLS) and Maximum Likelihood Estimation (MLE), combinate con Genetic Algorithm, un robusto metodo di ottimizzazione in grado di affrontare l'elevata non linearità del problema.
Per concludere, applichiamo questa metodologia sia ad analisi post-mortem che real-time, prendendo in considerazione la serie storica di Bitcoin (BTC/USD). In particolare, ci concentriamo sull'intervallo Dicembre 2016 - Gennaio 2018, dove Bitcoin ha quasi raggiunto i 20'000 USD, e il periodo più recente Marzo 2020 - Marzo 2021, dove Bitcoin ha avuto una notevole crescita oltrepassando i 60'000 USD.
I risultati ottenuti sono in linea con le pubblicazioni precedenti.

Parole chiave: Bolle finanziarie; Log Periodic Power Law model; Metodo dei minimi quadrati; Metodo dei minimi quadrati generalizzati; Metodo della massima verosomiglianza; Algoritmo Genetico; Bitcoin.

Acknowledgments

This paper is an important milestone for me, the end of my university studies. Therefore, I would like to thank all the people who accompanied me on this intense but wonderful journey.

First of all, I would like to thank my supervisor, Professor Daniele Marazzina, for his availability and precious advice.

The greatest of thanks goes to my parents. Thank you for supporting every choice of mine and for teaching me so much. Without your moral and financial support, this result would not have been possible.
To my brother Francesco, for always believing in me and encouraging me when I needed it. A thought also to my grandparents, my first supporters.

A special thanks also to Andrea, a lifelong friend, for always showing me reality with greater lightness.

A sincere thanks also to those people who have joined me since the very first day of University. Thank you Chiara L., Chiara M., Alessandro, Daniele and Davide for your precious advice and for sharing the happy, but especially the difficult, moments together.

Finally, I thank all my friends for helping me, with an advice or a laugh, to get closer to this goal.

Ringraziamenti

Questo elaborato è per me un traguardo importante, la conclusione del mio percorso universitario. Pertanto, desidero ringraziare tutte le persone che mi hanno accompagnato durante questo intenso ma meraviglioso viaggio.

In primis, vorrei ringraziare il mio relatore, il Professor Daniele Marazzina, per la disponibilità e i preziosi consigli.

Ai miei genitori va il più grande dei ringraziamenti. Grazie per aver supportato ogni mia scelta e per avermi insegnato tanto, da sempre. Senza il vostro sostegno morale ed economico, questo risultato non sarebbe stato possibile.
Grazie a mio fratello Francesco, per avere sempre creduto in me e avermi spronato quando ne avevo bisogno. Un pensiero anche ai miei nonni, i miei più accaniti sostenitori.

Un ringraziamento speciale inoltre ad Andrea, amico di una vita, per mostrarmi sempre la realtà con maggiore leggerezza.

Un sentito grazie inoltre a Chiara L., Chiara M., Alessandro, Davide e Luca per i vostri preziosi consigli e per aver condiviso i momenti felici, ma soprattutto quelli bui, tutti insieme.

Ringrazio infine tutti i miei amici per avermi aiutato, con un consiglio o una risata, ad avvicinarmi a questo traguardo.

1. Introduction

A financial bubble is a phenomenon where it is observed a consistent rise in the price of an asset, exceeding by far its fundamental value.
Because speculative demand, rather than intrinsic worth, fuels the inflated prices, the bubble eventually but inevitably pops, and massive sell-offs cause prices to decline, often quite dramatically.

The lack of a formal definition on the concept of bubble make its detection pretty challenging. Nonetheless, one can generally identify four stages:

1. *Displacement*: it occurs when investors get interested in a new paradigm, such as a disruptive technology or a new asset class. This may become the seed of a bubble.

2. *Boom*: prices rise slowly at first, but then more participants enter the market, fearing to miss a once in a lifetime opportunity. During this phase, the asset attracts widespread media coverage, increasing even more the number of investors into the fold.

3. *Euphoria*: during this phase prices skyrocket, completely disconnecting from the fundamental value of the underlying asset. New valuation metrics are invented to try to justify this relentless rise in prices. The irrational idea that there will always be buyers willing to pay more spreads everywhere.

4. *Panic*: even a relatively minor event can become the trigger of massive selloffs. The bubble has bursted and now investors just want to liquidate at any price. As supply overwhelms demand, asset prices may fall sharply.

The consequences of a bubble can affect financial markets for years and spread to the real economy too. For example, the burst of the Dot-com bubble, inflated by excessive speculation on Internet related activities in the late 1990s, made the Nasdaq Composite Index loose more than 75% of its market capitalization and many companies go bankrupt.
In addition, the bursting of financial bubbles usually lead to weaker recoveries with respect to typical down swings in the business cycle. Considering again the case of the Dot-com bubble, the Nasdaq Composite Index managed to reach again its peak only in 2015.
Eventually, the role of credit dramatically affects how financial bubbles could hit economies. Indeed, when fueled by credit, bubble bursts are responsible of deeper recessions and slower recoveries (Jordà et al., 2015).

The present work focuses on the study and the detection of financial bubbles according to the Log Periodic Power Law (LPPL) model, developed by Johansen, Ledoit and Sornette from the late 1990s, (Sornette et al., 2001), (Sornette, 2003).
In a nutshell, the model states that the price of an asset is driven by a bubble behaviour if the following relation holds:

$$\ln[p_i] = \mathrm{LPPL}(t_i; A, B, C, t_c, \beta, \omega, \phi) + \epsilon_i,$$

where $\{p_1, ..., p_n\}$ are historical prices of the asset in correspondence of $\{t_1, ..., t_n\}$, $\{\epsilon_i\}_{i=1}^n$ are the residuals of the fitting procedure and $\mathrm{LPPL}(t_i; A, B, C, t_c, \beta, \omega, \phi)$ indicates the LPPL function, which corresponds to

$$\mathrm{LPPL}(t_i; A, B, C, t_c, \beta, \omega, \phi) = A + B(t_c - t)^\beta + C(t_c - t)^\beta \cos[\omega \ln(t_c - t) + \phi].$$

The parameters of the model, $\{A, B, C, t_c, \beta, \omega, \phi\}$, must be calibrated in order to optimally fit the asset prices on a fixed time window.
Moreover, if the estimated parameters satisfy specific constraints, the so called Sornette bounds, one can state that the asset is under a bubble regime and the critical time, t_c, represents the candidate crash time.
The calibration of the parameters of the model is not an easy task, because of the high non linearity of the LPPL function and the presence of many local minima where a linear optimization method could get trapped.
The original calibration procedure moves under the assumption that residuals are homoskedastic and uncorrelated, combining the Ordinary Least Square (OLS) method with Genetic Algorithm, a non linear optimization method able to attack the problem without any particular assumption on the shape of the function.

In the following decades, there has been a considerable body of research on this matter. We briefly recall the most relevant publications.
An extensive mathematical formulation of the model has been proposed by Fantazzini and Geraskin (Fantazzini et al., 2011).
A group of research, supervised by Didier Sornette himself, has been later established at ETH Zurich University, carrying on many real case studies and further theoretical development.
To this extent, we underline the institution of the Financial Crisis Observatory for the monitoring of financial bubbles on the main Equity and Commodity Indices.
Following the stream of publications, we recall the formulation of the Maximum Likelihood Estimation calibration procedure (Filomonov et al., 2016), which is able to provide confidence intervals, and not only punctual results, for the estimation of the parameters, providing a measure of the uncertainty related to the estimated critical time.
An attempt to relax the assumption on homoskedasticity and uncorrelation of residuals required in the original calibration procedure has then been made with the proposition of the Generalized Least Square calibration method (Scaringi, 2016), (Bianchetti et al., 2018). This variant opens up to correlated residuals following an autoregressive process and it has proved to provide more accurate forecasts.
Eventually, the Epsilon Drawdown/Drawup method and the Lagrange Regularisation

approach for the estimation of bubble start dates (Demos, Sornette, 2017), (Demos, Sornette et al., 2018) have provided a complete procedure for the post-mortem analysis of financial time series, where clearly there is no forecasting ambition, but only the desire to analyze the time series as a whole.

The rest of the thesis is organized as follows.
In Chapter 2, we provide a theoretical background, with basic notions on Times Series, Regression Theory and No Arbitrage Theory.
Chapter 3 is devoted to the mathematical formulation of the model, with particular focus on latest publications.
Chapter 4 deals with the calibration of the parameters of the model. In particular, we present the original calibration procedure, together with two more recent and reliable variants: the Maximum Likelihood Estimation (MLE) and the Generalized Least Squares (GLS) method.
In Chapter 5, we carry on practical applications on Bitcoin's historical prices, both via post-mortem and real-time analyses. Specifically, we have focused on Bitcoin's time series between December 2016 and January 2018, when Bitcoin prices skyrocketed from 1'000 USD to almost 20'000 USD, and on the recent period to establish whether the sharp increase in price observed throughout the last year was influenced by a bubble behaviour or not.
Eventually, Chapter 6 summarizes the main results obtained.

2. Theoretical Background

In this chapter, some relevant theoretical notions are presented with the idea of better understanding their application in the subsequent chapters.

First, we recall basic concepts on Time series following (Adhikari et al., 2013).
Then, we present some techniques of Regression Theory gradually increasing the degree of generalization.
Eventually, we recall a fundamental result from No Arbitrage Theory, that will be required in Chapter 3.

2.1. Time Series

A time series is a sequential set of data points measured over successive times and arranged in chronological order.
A time series is continuous if observations are measured at every instance of time, whereas it is discrete if observations are measured at discrete points.
In particular, we will deal with discrete time series, for example daily closing prices of a financial asset, mathematically represented as

$$x = \{x_1, x_2, ..., x_n\}.$$

A time series x is a non deterministic process, so we cannot predict with certainty what will occur in the future. For this reason, it is generally modeled as a stochastic process with x assumed to follow a certain probability model, which describes the joint distribution of the random variables x_i. Thus, observations $\{x_i\}_{i=1}^n$ are sample realizations of the stochastic process x that produced them.

Time series can present some relevant properties. One of these is certainly *stationarity*, which can be considered as a form of statistical equilibrium (Adhikari et al., 2013).

Definition 2.1. *A time serie x is strictly stationary if the probability distribution of the process is invariant under time translations, that is the joint probability of every pair of random variables $(x_t; x_{t-s})$ depends only on the temporal distance $t - s$.*

However for practical applications, the assumption of strong stationarity is not always needed and so a somewhat weaker form is accepted.

Definition 2.2. *A time serie x is weakly stationary of order k if the statistical moments of the process up to that order are time invariant.*

As a consequence, a stochastic process x is second order stationary if its mean and variance are time independent and its covariance depends only on s. Namely,

$$\begin{aligned} &\mathbb{E}[x_t] = \mu \qquad \forall t, \\ &\mathrm{Var}[x_t] = \sigma^2 \qquad \forall t, \\ &\mathrm{Cov}[x_t, x_{t-s}] = \sigma_s \qquad \forall t. \end{aligned}$$

An example of stationary process is the White Noise, $\mathcal{WN}(0, \sigma^2)$, i.e. a sequence of independent and identically distributed (i.i.d) random variables with zero mean and variance σ^2.
An example of non stationary process is instead the random walk

$$x_t = x_{t-1} + \epsilon_t, \tag{2.1}$$

with $\epsilon_t \sim \mathcal{WN}(0, \sigma^2)$.

In general, models for time series can have many forms representing different stochastic processes. We will focus in particular on Autoregressive processes.

Definition 2.3. *An Autoregressive process of order* p, $AR(p)$, *is written as*

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + ... + \phi_p x_{t-p} + \epsilon_t,$$

with $\{\phi_i\}_{i=1}^p \in \mathbb{R}$ *and* $\epsilon_t \sim \mathcal{WN}(0, \sigma^2)$.

It is straightforward to see how the process x at time t, x_t, is influenced by its previous values x_{t-1}, x_{t-2}, ..., x_{t-p}.

Within the class of Autoregressive processes, the easiest case is the $AR(1)$ process

$$x_t = \rho x_{t-1} + \epsilon_t, \qquad \rho \in \mathbb{R}.$$

For this specific class of processes, it can be proven that:

Proposition 2.1. *AR(1) process is weakly stationary* $\iff |\rho| < 1$.

For an extended proof, see (Adhikari et al., 2013, pag. 20).

This result will be widely used in Chapter 4, as we will deal with $AR(1)$ processes requiring stationarity.

2.2. Regression Theory

The term *regression* refers to a wide range of statistical methods used to examine the relationship between an outcome variable y and one or more regressor variables $x = [x_1, ..., x_d]$. It is typically formulated as

$$y = f(x; \phi) + \epsilon, \tag{2.2}$$

where ϕ is a vector of unknown parameters and ϵ is an error term representing random noise.

Given n observations of y and x, the goal of Regression analysis is to estimate the vector of parameters $\hat{\phi}$ that most closely fits the data, i.e. find

$$\hat{y} = f(x; \hat{\phi}),$$

that better resembles y dynamics.

In particular, we will deal with Linear regression models, where f is a linear function and (2.2) is rewritten as

$$y = x^T \phi + \epsilon, \tag{2.3}$$

where x^T represents the transpose of vector x.

We will deal in more detail with the Ordinary Least Squares method, which is the most commonly used, and the Generalized Least Squares method, which is its generalization.

2.2.1. Ordinary Least Squares

Let us consider a linear mono-dimensional model defined as

$$y_i = \alpha + \beta x_i + \epsilon_i, \qquad i = 1, ..., n \tag{2.4}$$

with $\alpha, \beta \in \mathbb{R}$ and $\epsilon_i \sim \mathcal{WN}(0, \sigma^2)$.

The OLS method provides an estimate of the parameters, $\{\alpha, \beta\}$, in such a way that the sum of the squared residuals is minimized. Namely,

$$\{\hat{\alpha}, \hat{\beta}\} = \arg\min_{\alpha,\beta} S(\alpha, \beta), \tag{2.5}$$

where

$$\begin{aligned} S(\alpha, \beta) &= \sum_{i=1}^{n} (y_i - \hat{y}_i)^2, \\ &= \sum_{i=1}^{n} (y_i - \hat{\alpha} - \hat{\beta} x_i)^2, \\ &= \sum_{i=1}^{n} (e_i)^2. \end{aligned}$$

In particular, $\{e_i\}_{i=1}^n$, such that $e_i = y_i - \hat{y}_i \;\; i = 1, ..., n$, are called residuals and they describe the error in adapting the model to real observations.

Setting the partial derivatives of $S(\alpha, \beta)$ to be equal to zero,

$$\frac{\partial S(\alpha, \beta)}{\partial \alpha} = 0$$
$$\frac{\partial S(\alpha, \beta)}{\partial \beta} = 0,$$

we are able to deduce explicit formulas for $\{\hat{\alpha}, \hat{\beta}\}$ (Hubele et al., 2011):

$$\hat{\beta} = \frac{S_{xy}}{S_{xx}} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2},$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x},$$

where $\bar{x}$ indicates the sample mean of observations $\{x_i\}_{i=1}^{n}$.

Considering now the multi-dimensional case,

$$Y = X\beta + \epsilon, \tag{2.6}$$

where Y is a $n \times 1$ vector of outcome variables, X is a $n \times d$ matrix, with n observations of d independent variables, ϵ is a $n \times 1$ vector of disturbance variables and β is a $d \times 1$ vector of unknown parameters that require to be estimated.
In particular, the first column of X is an identity vector, $1 = [1, ..., 1]^T$, in order to address the model with a constant term.

Thanks to Gauss-Markov theorem, we can affirm the goodness of the OLS estimator (Cameron et al., 2005).

Theorem 2.1 (Gauss-Markov Theorem)**.** *Given Formula (2.6), which can be equivalently rewritten as*

$$Y_i = \sum_{j=1}^{d} \beta_j X_{ij} + \epsilon_i \qquad \forall i = 1, ..., n, \tag{2.7}$$

if X *has maximum rank and the disturbance variables* ϵ_i *are such that:*

- *they have zero mean, i.e.* $\mathbb{E}[\epsilon_i] = 0 \ \forall i$,
- *they are homoskedastic, i.e.* $\mathrm{Var}[\epsilon_i] = \sigma^2 \ \forall i$,
- *they are uncorrelated, i.e.* $\mathrm{Cov}[\epsilon_i, \epsilon_j] = 0 \ \forall i \neq j$,

then the Ordinary Least Squares Estimator,

$$\hat{\beta}_{OLS} = (X^T X)^{-1} X^T y, \tag{2.8}$$

is the Best Linear Unbiased Estimator (BLUE), meaning that

- $\hat{\beta}_{OLS}$ *is unbiased, i.e.* $\mathbb{E}[\hat{\beta}_{OLS}] = \beta$,
- $\hat{\beta}_{OLS}$ *is the most efficient estimator, i.e. it has the lowest variance compared to other unbiased estimators.*

In particular, the variance of $\hat{\beta}_{OLS}$ corresponds to

$$\text{Var}[\hat{\beta}_{OLS}] = \sigma^2(X^T X)^{-1}. \tag{2.9}$$

Moreover, note that the assumptions of homoskedasticity and zero correlation together imply the disturbances ϵ to be such that

$$\Omega = \text{Var}[\epsilon] = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix}.$$

Eventually, adding the gaussianity assumption on disturbances ϵ_i, the OLS estimate $\hat{\beta}$ is distributed in such a way that

$$\hat{\beta}_{OLS} \sim \mathcal{N}(\beta, \sigma^2(X^T X)^{-1}). \tag{2.10}$$

2.2.2. Generalized Least Squares

The Generalized Least Square (GLS) method is particularly useful when there is a certain degree of correlation between the residuals in a regression model, because in this case the OLS version could reveal inefficient and provide misleading results.

Let us consider again the model

$$Y = X\beta + \epsilon, \tag{2.11}$$

if we mantain the homoskedasticity assumption on disturbances, but we relax the hypothesis on zero correlations, we get that $\text{Var}[\epsilon] = \Omega \neq \sigma^2 I$.

Under the assumption that Ω is known, we can multiply (2.11) by $\Omega^{-1/2}$,

$$\Omega^{-1/2}Y = \Omega^{-1/2}X\beta + \Omega^{-1/2}\epsilon,$$

where $\Omega = \Omega^{1/2}\Omega^{1/2}$ is feasible thanks to Cholesky Decomposition, since Ω is a symmetric and definite positive matrix.
Note now that $\text{Var}[\Omega^{-1/2}\epsilon] = I$, hence the transformed residuals are homoskedastic and uncorrelated. For this reason, we can deduce the $\hat{\beta}_{OLS}$ estimator of $\Omega^{-1/2}Y$ on $\Omega^{-1/2}X$. This argument yields to the Generalized Least Squared (GLS) Estimator

$$\hat{\beta}_{GLS} = (X^T\Omega^{-1}X)^{-1}X^T\Omega^{-1}Y. \tag{2.12}$$

Moreover,

$$\mathrm{Var}[\hat{\beta}_{GLS}] = \sigma^2(X^T\Omega^{-1}X)^{-1}. \tag{2.13}$$

Thanks to Aitken Theorem, we can eventually state that $\hat{\beta}_{GLS}$ is the BLUE estimator of the model.

Theorem 2.2 (Aitken Theorem)**.** *Given Formula (2.11), where disturbances ϵ have some degree of correlation, i.e.* $\mathrm{Var}[\epsilon] = \Omega \neq \sigma^2 I$*, the* GLS *estimator for* β*,*

$$\hat{\beta}_{GLS} = (X^T\Omega^{-1}X)^{-1}X^T\Omega^{-1}Y,$$

is the BLUE estimator of the model, since it is unbiased and with lowest variance.

For an extended proof, see (Greene, 2002).

Feasible Generalized Least Squares

To present the theoretical results above, we have assumed that Ω, the covariance matrix of the disturbances, was known, but actually most of the times this does not happen in practice, hence it is required to provide an estimated matrix, $\hat{\Omega}$.

Without loss of generality, we suppose that Ω depends over a parameter ρ, $\Omega = \Omega(\rho)$, so we can shift the consistency of Ω to ρ, meaning that $\hat{\Omega} = \Omega(\hat{\rho})$ and $\hat{\Omega}$ is consistent if $\lim \hat{\rho} = \rho$.
In this framework, the estimation procedure is called Feasible Generalized Least Squares (FGLS) and all the results presented above for the GLS case still hold. Consequently,

$$\hat{\beta}_{FGLS} = (X^T\hat{\Omega}^{-1}X)^{-1}X^T\hat{\Omega}^{-1}Y. \tag{2.14}$$

Before presenting explicitly the FGLS procedure for a specific case, we underline that $\hat{\beta}_{GLS}$ and $\hat{\beta}_{FGLS}$ are consistent only asymptotically (Cameron, 2005), hence for small samples consistency could not be satisfied and $\hat{\Omega}$ could be an unreliable estimate of Ω.

Now, we will consider in more detail FGLS procedure for a specific case, that will see further application in Chapter 4.
Recalling (2.11), suppose that disturbances are modelled by the following dynamics

$$Y_i = \sum_{j=1}^{d} \beta_j X_{ij} + \epsilon_i \qquad \forall i = 1, ..., n,$$

$$\epsilon_i = \rho\epsilon_{i-1} + u_i, \tag{2.15}$$

where $u_i \sim \mathcal{N}(0, \sigma^2)$ for $i = 1, ..., n$ and $\rho \in \mathbb{R}$, $|\rho| < 1$, is the parameter that we require to estimate.
In other words, ϵ_i are correlated following an autoregressive process of order 1, AR(1).

To get an estimation of $\hat{\beta}_{FGLS}$ we will follow Cochrane-Orcutt procedure (Scaringi, 2016):

1. Compute the covariance matrix Ω of the disturbances according to AR(1) process. This leads to

$$\Omega(\rho, \sigma^2) = \frac{\sigma^2}{1-\rho^2} \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{bmatrix} = \sigma^2 \Delta(\rho). \tag{2.16}$$

2. Apply Cholesky decomposition and deduce matrix Ψ, such that

$$\Delta(\rho)^{-1} = \Psi^T(\rho)\Psi(\rho). \tag{2.17}$$

   Specifically, one can prove that Ψ is a lower triangular matrix

$$\Psi(\rho) = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & \cdots & 0 & 0 \\ -\rho & 1 & 0 & \cdots & 0 \\ 0 & -\rho & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\rho & 1 \end{bmatrix}. \tag{2.18}$$

3. Consider the residuals $\hat{\epsilon}$ obtained via OLS estimation and deduce explicitely $\hat{\rho}$ via OLS method

$$\hat{\epsilon}_i = \hat{\rho}\hat{\epsilon}_{i-1} + u_i, \qquad i = 1, ..., n \tag{2.19}$$

4. Provide a first estimation of $\Psi(\rho)$ in (2.18), using $\hat{\rho}$ from (2.19).

5. To provide an estimation of $\hat{\beta}_{FGLS}$, multiply the model for $\Psi(\hat{\rho})$

$$\Psi(\hat{\rho})Y = \Psi(\hat{\rho})X\beta + \Psi(\hat{\rho})\epsilon, \tag{2.20}$$

   we can now assert that $\Psi(\hat{\rho})\epsilon = u \sim \mathcal{N}(0, \sigma^2)$, so we can move to the OLS framework.
   Namely, (2.20) corresponds to the following quasi-difference equation

$$y_i - \hat{\rho}y_{i-1} = (\alpha + \beta x_i + \epsilon_i) - \hat{\rho}(\alpha + \beta x_{i-1} + \epsilon_{i-1}), \qquad i = 2, ..., n \tag{2.21}$$

   which can be rewritten in a more compact way as

$$y_i^* = \alpha s^* + \beta x_i^* + \epsilon_i^*, \qquad i = 2, ..., n \tag{2.22}$$

   where

$$\begin{aligned} y_i^* &= y_i - \hat{\rho}y_{i-1} \\ s^* &= 1 - \hat{\rho} \\ x_i^* &= x_i - \hat{\rho}x_{i-1} \\ \epsilon_i^* &= \epsilon_i - \hat{\rho}\epsilon_{i-1} = u_i. \end{aligned}$$

Since disturbance, $\epsilon_i^* = u_i$, in (2.22) are uncorrelated, we are now able to minimize

$$S(\alpha, \beta) = \sum_{i=1}^{n} (y_i^* - \alpha s^* - \beta x_i^*)^2, \tag{2.23}$$

following the OLS formulation.

Eventually, note that the quasi-difference equation (2.21) holds starting from $i = 2$, because observations x_0 and y_0 do not exist.

2.3. No Arbitrage Theory

In this section, we will present a fundamental result of Asset Pricing theory, that will be required in Chapter 3.

We define *Arbitrage* as the practice of exploiting market inefficiences in order to make a profit out of nothing without taking any risk (Bjork, 2009). Namely,

Definition 2.4. *An arbitrage opportunity on a financial market is a self-financed portfolio* h *such that*

$$\begin{aligned} V^h(0) &= 0, \\ P(V^h(T) \geq 0) &= 1, \\ P(V^h(T) > 0) &> 0; \end{aligned}$$

where $V : t \in [0; T] \longrightarrow \mathbb{R}$ *is a function representing the value of a portfolio at time* t.

A typical example of arbitrage consists in the simultaneous buying and selling of a security in order to take advantage of price discrepancies within different markets.
Nonetheless, we will move under the assumption of *Market efficiency*, which requires markets to be free of arbitrage possibilities.
In other words, we will always assume that strategies that aims at making profits intrinsically hide some kind of risk.

Before presenting the First Fundamental Theorem of Asset Pricing, let us briefly recall some preliminary concepts on stochastic processes.

Definition 2.5. *A stochastic process is a mathematical object* $X = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in T}, (X_t)_{t\in T}, \mathbb{P})$, *where*

- $(\Omega, \mathcal{F}, \mathbb{P})$ *is a probability space;*
- $T \subset \mathbb{R}^+$ *is called Time Span;*
- $(\mathcal{F}_t)_{t\in T}$ *is a family of increasing* σ *- algebras,* $\mathcal{F}_s \leq \mathcal{F}_t \leq \mathcal{F}$ $\forall s \leq t$, *called Filtration;*
- $(X_t)_{t\in T}$ *is a family of random variables* $\forall t \in T$.

Specifically, we are interested on martingale processes, because of their property for which the conditional expectation of their future value, given the information accumulated up to now, correspond to their current value.

Definition 2.6. *Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ be a filtered Probability space and let $X = \{X_t\}_{t\geq 0}$ be a continuous stochastic process adapted to the Filtration $\mathcal{F}_t$ (or $\mathcal{F}_t$ - measurable $\forall t$), X is a martingale if*

1. $\mathbb{E}[X_t] < +\infty \quad \forall t$,
2. $\mathbb{E}[X_t|\mathcal{F}_s] = X_s \quad \forall s \leq t$.

We are now ready to enunciate the First Fundamental Theorem of Asset Pricing.

Theorem 2.3 (First Fundamental Theorem of Asset Pricing)**.** *Consider a market $\mathcal{M}$ constituted by a risk-free asset $B = \{B_t\}_{t\geq 0}$ and by a risky asset $S = \{S_t\}_{t\geq 0}$, modelled as a continuous stochastic process.*
The following statements are equivalent:

1. *S does not allow arbitrage opportunities;*
2. *There exists a probability measure $\mathbb{Q}$ on the measurable space $(\Omega, \mathcal{F})$, which is equivalent to $\mathbb{P}$,*

$$\mathbb{Q} \sim \mathbb{P},$$

meaning that $\mathbb{P}(A) = 0 \iff \mathbb{Q}(A) = 0, \forall A \in \mathcal{F}$.
Moreover, S is a martingale process under $\mathbb{Q}$, that is

$$\mathbb{E}^{\mathbb{Q}}\left[\frac{S_t}{B_t}|\mathcal{F}_s\right] = S_s, \qquad \forall s < t.$$

3. The Log Periodic Power Law Model

Since the model originally created by Johansen, Ledoit and Sornette in the late 1990s, there has been a considerable body of research adopting an approach which relies on agent-based models and using critical points with log-periodic corrections to identify financial bubbles.
Such methodology is called Johansen-Ledoit-Sornette model or Log-Periodic Power Law (LPPL) model.

3.1. A brief historical overview

The idea that certain phenomena subject to critical conditions could be described with a log-periodic behaviour first came in 1991.
The original application concerned aerospace flights during the European missions of Ariane IV and V rockets. Daniel Sornette realized that during the propulsion phase the materials covering the rocket was characterized by specific critical oscillations. The power laws and log-periodic patterns discovered in this context were found to perform quite well.
In 1995, Sornette and Sammis extended the field of application to earthquakes, since the model seemed to forecast quite reliably the critical rupture in heterogeneous materials.
Approximately in the same period as the extension to earthquakes was proposed, Feigenbaum, Freund and Sornette independently suggested that another interesting field of application could have been large financial crashes.
Since then, a theoretical framework has been developed and many practical applications have been made.

3.2. Financial crashes dynamics and predictability

Market drawdowns are undoubtedly feared financial events, but it is important to underline that they are intrinsically part of the process.
Indeed, bull markets in the last decades always had one or more corrections. This is a natural part of the market cycle that wise investors welcome as a pullback for the market to consolidate before going toward higher highs.
Nonetheless, a huge effort has been spent trying to explain the dynamics which might trigger market crashes, which are more sudden and violent than simple corrections, at least in case of financial bubbles when the crash can be somehow expected and justified by a previous irrational behaviour of the market.

According to LPPL model, a crash occurs because the market has entered an unstable phase and any disturbance might be the trigger of the instability.
According to this view, a crash has fundamentally an endogenous origin and exogenous shocks only work as triggering factors. As a consequence, the origin of the crashes is built progressively by the market as a self-organizing process.

Sornette presented a series of observations showing how financial markets constitute one among many other systems exhibiting a complex organization and dynamics with similar behaviour (Sornette, 2003).
In this scenario, the complex system approach, which consists in the analysis of the system as a whole, is preferred to the analytical approach, which relies on the decomposition in sub parts and on the detailed study of each component in order to determine the functioning of the whole system.

Because of their complexity, these systems often cannot be provided with a mathematical formulation or even an analytical description, but they can be explored only by means of numerical simulations. For these reasons, they are considered to be generally unpredictable.
However, they are also characterized by the occurrence of large scale collective behaviors resulting from repeated interactions among its constituents.
It turns out that most complex systems in natural and social sciences exhibit rare and sudden transitions, ranging from large natural catastrophes, such as earthquakes or volcanic eruptions, to the failure of engineering structures or crashes in financial markets.
It is essential to realize that the long-term behavior of these complex systems is often controlled in large part by these rare catastrophic events and that they result from a progressive and more global cooperative process occurring over the whole system by repetitive interactions.

Considering more closely the reality of financial markets crashes, it is convenient to distinguish between Black Swans and Dragon Kings (Sornette, 2009).
While drawdowns called Black Swans are considered to be unpredictable, Dragon Kings are outliers, meaning that they are rare events with significantly elevated value.
To better understand the difference between Black Swans and Dragon Kings, Sornette decided to consider financial crashes as drawdowns defined as follows.
Let $X = (X_t)_{t>0}$ be a random process with $X_0 = 0$, the drawdown at time T, D(T), is

$$D(T) := \max\Big\{0, \max_{t\in(0,T)} X(t) - X(T)\Big\}.$$

A drawdown can be then viewed as the measure of the decline from a historical peak.

The analysis of drawdowns of Nasdaq Composite Index brought the presence of dragon kings, as showed in Figure 3.1.
Sornette found out that the majority of the Dragon Kings were actually preceded by a consistent asset price increase (Sornette, 2009). This result drove to the idea that these

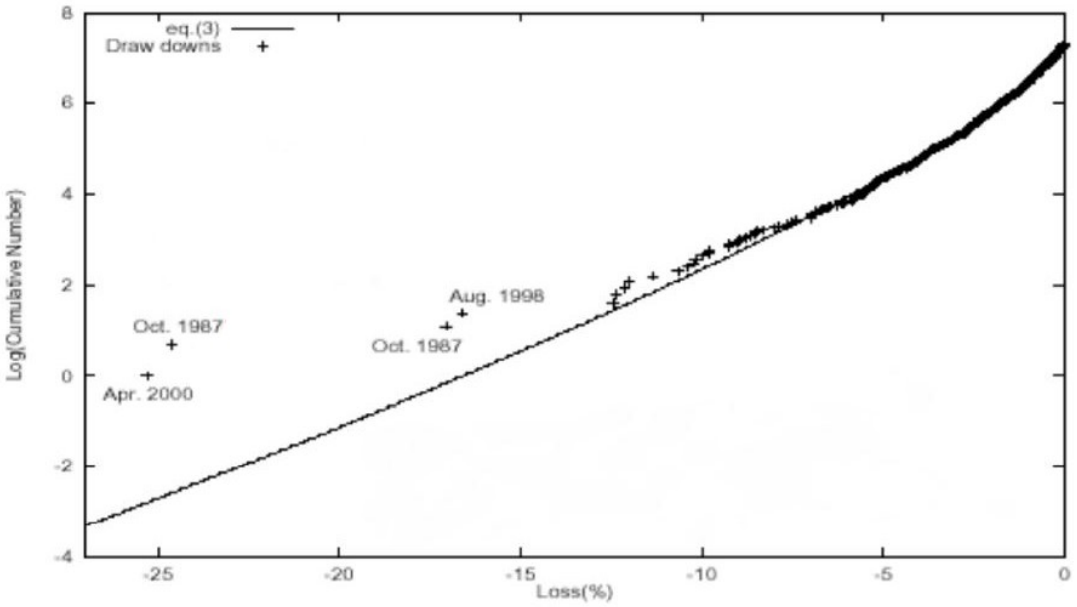

Figure 3.1: Distribution of Nasdaq Composite Index drawdowns from (Sornette, 2009).

crashes would have been the inevitable consequence to periods of unsustainable growth in price, due to a self-fulfilling enthusiasm which produced irrational behaviours.

3.3. The Original Formulation

Now we will deal with the theoretical presentation of the model as originally formulated by Johansen, Ledoit and Sornette (Johansen et al., 2001) and reported by many other successive works (Fantazzini et al., 2011), (Bingcun et al., 2018).

Let us consider an ideal market where dividends, interest rates, liquidity constraints and risk aversion are ignored.
One can reasonably assume that the fundamental value of an asset is $p(t) = 0$, hence any positive value of $p(t)$ represents a bubble and similarly any negative value of $p(t)$ indicates that the asset is undervalued.
As a consequence, $p(t)$ can be seen as the price in excess from the fundamental value of an asset.

In this framework, there exist only two types of agents: rational players, identical in their preferences and characteristics, and a group of irrational players with herding behaviour.
The tendency of irrational agents to imitate a particular behaviour in the market, for

example a long position on the same asset, is the main burst to the development of a financial bubble. This behaviour will continue until a critical value, where a large portion of investors will then assume the same short position, thus causing a crash.
Nonetheless, since the crash is not a certain deterministic outcome of the bubble it remains rational for investors to remain in the market since they are compensated by a higher return. Indeed, there is a finite probability of attaining the end of the bubble without crash, but instead with a smooth correction.

The key variable to model the price behaviour before a crash is the hazard rate $h(t)$, which is the probability per unit of time that the crash will take place, given that it has not yet occurred.
Financially speaking, the hazard rate measures the probability that a great number of investors will take simultaneously a sell position, causing a substantial decrease in the price of the asset.

3.3.1. Microscopic Modelling

The formulation proposed by Johansen, Ledoit and Sornette assumes that the group of irrational players is connected into a network. Each agent i, $i = 1, ..., I$, can only have two possible states: "long" ($s_i = +1$) or "short" ($s_i = -1$).
For "long" position we mean that investor i has a positive view about the growth of the asset and he owns it into his portfolio or is going to buy it. Consequently, the same investor i has a short position whenever he has a negative opinion on the future growth of the asset.
Moreover, each player is directly connected with $N(i)$ agents with a direct influence on each one of them.

In this framework, the state of agent i is determined as:

$$s_i = \begin{cases} +1 & \text{if} \quad K \sum_{k \in N(i)} s_k + \sigma \epsilon_i > 0, \\ -1 & \text{if} \quad K \sum_{k \in N(i)} s_k + \sigma \epsilon_i \leq 0, \end{cases} \tag{3.1}$$

where $K > 0$ and ϵ_i is i.i.d. $\mathcal{N}(0; 1)$.
While K governs the tendency of imitating the behaviour of other investors $k = 1, ..., N(i)$, σ quantifies the idiosyncratic behaviour.
It is clear that the consequence of an increase in K is a better ordered system, while the reverse is true when σ increases.
It is reasonable to affirm that there exists a critical point K_c, which is the separation between these two different regimes: when $K < K_c$ the disorder prevails and there is not a significative predominance of one state. When $K \longrightarrow K_c$, groups of agents with same positions start to form and the market becomes extremely sensitive to small global disturbances, like external news.
Finally, for $K > K_c$, the tendency to imitation is so strong that one position prevails

among the community of players.

For sake of completeness, it can be interesting to include in the modelization also a generic term $G \in \mathbb{R}$, called global influence, measuring external influence on each single agent.
This quantity summarizes external contributions able to affect most or all the network, like global news or events.
Adding G to (3.1), one gets:

$$s_i = \begin{cases} +1 & \text{if} \quad \left(K \sum_{k \in N(i)} s_k + \sigma \epsilon_i\right) + G > 0, \\ -1 & \text{if} \quad \left(K \sum_{k \in N(i)} s_k + \sigma \epsilon_i\right) + G \leq 0. \end{cases} \tag{3.2}$$

Defining the average state of the market as

$$M = \frac{\sum s_i}{I},$$

for $G = 0$ $\mathbb{E}[M] = 0$, since ϵ_i is such that $\mathbb{E}[\epsilon_i] = 0$. Analogously, for $G > 0$ one will get $\mathbb{E}[M] > 0$, and for $G < 0$ $\mathbb{E}[M] < 0$.
This means that if the global influence favors the "buy" state ($G > 0$) the network is pushed towards that specific state and viceversa.

At this point, it is convenient to introduce the susceptibility of the system χ defined as

$$\chi = \frac{d\mathbb{E}[M]}{dG}\Big|_{G=0}.$$

χ is a measure of the sensitivity of M to a small change in G. In other words, it can be seen as a measure of the impact on a player if a directly connected agent is forced to a specific state.

3.3.2. Price Dynamics and Derivation of the Model

We required the rational agent to be risk neutral and with rational expectations. Under these assumptions, the asset price $p(t)$ will follow a martingale process

$$\mathbb{E}[p(t)|\mathcal{F}_s] = p(s), \qquad \forall t > s.$$

Previous bubble crashes on financial markets, for example the bursts of the dot-com bubble in 2000 or the Bitcoin bubble in 2018, showed that there is a not zero probability for the crash to happen, hence we can model such crash, happening at time t_c, as a jump process j, equal to 0 before it and equal to 1 after.
Since t_c is unknown, it can be modelled by a stochastic variable having density $q(t)$ and cumulative distribution $Q(t)$.
At this point, we will report an important result (Fantazzini et al., 2011) which allows us to express the hazard rate $h(t)$ as a function of $q(t)$ and $Q(t)$.

Theorem 3.1. *Given a jump process* j *with density* $q(t)$ *and cumulative distribution* $Q(t)$*, the hazard rate* $h(t)$*, probability per unit of time that the crash will take place given that it has not yet occurred, is defined as:*

$$h(t) = \frac{q(t)}{1 - Q(t)}.$$

Proof. Just apply the definition of hazard rate:

$$\frac{P(t_c < t + h | t_c > t)}{h} = \frac{P(t < t_c < t + h)}{hP(t_c > t)} = \frac{1}{h} \frac{\int_t^{t+h} q(x)dx}{1 - Q(t)} \longrightarrow \frac{q(t)}{1 - Q(t)} = h(t).$$

□

Let us consider the asset price dynamics as given by:

$$dp = \mu(t)p(t)dt - kp(t)dj, \tag{3.3}$$

where we are assuming that during a crash the price falls of a fixed percentage $k \in (0; 1)$.

The no arbitrage martingale condition together with the assumption of rational expectations require that $\mathbb{E}[dp] = 0$, so

$$\begin{aligned}\mathbb{E}[dp] &= \mu(t)p(t)dt - kp(t)\mathbb{E}[dj] \\ &= \mu(t)p(t)dt - kp(t)[P(dj = 0)(dj = 0) + P(dj = 1)(dj = 1)] \\ &= \mu(t)p(t)dt - kp(t)h(t)dt,\end{aligned}$$

because $P(dj = 1) = h(t)dt$.
One can easily deduce the following:

$$\mu(t) = kh(t). \tag{3.4}$$

Substituting it in the price dynamics (3.3) before the crash, i.e. when $j = 0$,

$$\begin{aligned} dp &= \mu(t)p(t)dt \\ &= kh(t)p(t)dt \\ d(\ln p(t)) &= kh(t)dt \\ \ln \frac{p(t)}{p_0} &= k \int_{t_0}^{t} h(s)ds \end{aligned}$$

$$\ln p(t) = p_0 \cdot k \int_{t_0}^{t} h(s)ds. \tag{3.5}$$

Hence it is clear that to capture the behaviour of $p(t)$, the hazard rate $h(t)$ needs to be specified.

In order to do that, it is necessary to better clarify the nature of the network where agents are connected.

Bidimensional Ising Model
The first result was obtained for the bidimensional Ising model, which considers players connected in an uniform way as shown in Figure 3.2. Every node is an agent and every line is a connection with another investor.

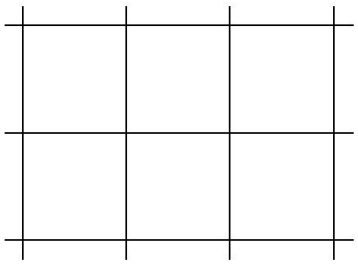

Figure 3.2: Representation of the two dimensional Ising model.

In this case, the following theorem holds (Johansen et al., 2008):

Theorem 3.2. *A system of variables close to a critical point can be described by a power law and the susceptibility χ of the system diverges as:*

$$\chi \approx A(K - K_c)^{-\gamma}.$$

Here K_c is the critical value of K defined in the previous section, A is a positive constant and $\gamma > 0$ is the critical exponent of the susceptibility.
In particular, for the two-dimensional Ising model $\gamma = \frac{7}{4}$.

We do not know the dynamics that drive K, but it is reasonable to assume that it evolves smoothly so that we can use a Taylor expansion around the critical point.
Let us define t_c as the first time such that $K(t_c) = K_c$, then before the critical time t_c we have that:

$$K_c - K(t) \approx C \cdot (t_c - t).$$

Using this approximation, we posit that the hazard rate $h(t)$ behaves as the susceptibility χ close to the critical point (Johansen et al., 2008):

$$h(t) \approx B \cdot (t_c - t)^{-\alpha}, \tag{3.6}$$

with $B > 0$ and $\alpha \in (0; 1)$.

It is crucial to remember that t_c is not the exact time of the crash, but only the most probable, and that there exists the possibility for the crash not to happen at all.
This is fundamental for the theory presented, otherwise rational agents would anticipate the crash.

Moreover, this model looses its reliability just after the critical time as the system enters in a new unpredictable phase.

Although we would able to find a closed formula for the evolution of $\ln p(t)$, the bidimensional Ising model is not a realistic description of financial markets, since they are actually formed by very heterogeneous groups of investors differing significantly in size and available connections, going from individual investors to large asset management funds.

Hierarchical Diamond Lattice
A more appropriate representation is given by the Hierarchical Diamond Lattice reported in Figure 3.3, which can be described in the following way:
First, take two players linked to each other. Secondly, substitute this link with four new links forming a diamond with two new traders added to the network. Then, continue in the same way substituting each connection with four new ones and adding two new traders at every substitution.
As a result, after n iterations there will be $N = \frac{2}{3}(2+4^n)$ agents and $L = 4^n$ links among them. Nonetheless, note that the last added agents only have two links, while the initial players will have 2^n neighbors.

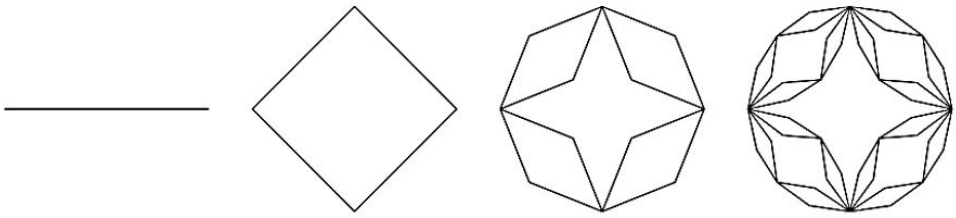

Figure 3.3: Stages of the evolution of a Hierarchical Diamond Lattice.

For this network, Theorem 3.2 still holds, but the critical exponent γ is now a complex number (Johansen et al., 2008).
In this case, the susceptibility is approximated as:

$$\begin{aligned}\chi &\approx \mathrm{Re}[A_0(K_c - K)^{-\gamma} + A_1(K_c - K)^{-\gamma+i\omega} + ...] \\ &\approx A_0'(K_c - K)^{-\gamma} + A_1'(K_c - K)^{-\gamma}\cos[\omega \ln(K_c - K) + \psi] + ...,\end{aligned}$$

with A_0, A_1, ω and $\psi \in \mathbb{R}$.
It is interesting to note that now the power law is corrected with log-periodic oscillations which accelerate while approaching the critical time.

Following the same steps as in the previous case, we can deduce the hazard rate of a crash:

$$h(t) \approx B_0(t_c - t)^{-\alpha} + B_1(t_c - t)^{-\alpha}cos[\omega \ln(t_c - t) + \psi']. \tag{3.7}$$

Note that tha hazard rate explodes near the critical date but now it also displays log-periodic oscillations.
Applying (3.7) to (3.5), we eventually get this expression for the evolution of the asset price before the crash

$$\ln[p(t)] \approx \ln[p(t_c)] - \frac{\kappa}{\beta}\left\{B_0(t_c - t)^{\beta} + B_1(t_c - t)^{\beta}cos[\omega \ln(t_c - t) + \phi]\right\}, \tag{3.8}$$

rewritten as

$$\ln[p(t)] \approx A + B(t_c - t)^{\beta} + C(t_c - t)^{\beta}cos[\omega \ln(t_c - t) + \phi]. \tag{3.9}$$

This is the final formulation of the Log Periodic Power Law (LPPL).
The model requires seven parameters: A, B, C, ω, β, ϕ and t_c, which is the most relevant one since it expresses the critical time of the crash.

4. Model Calibration

The Log-Periodic Power Law (LPPL) model has been developed for the detection of financial bubbles and crashes.
As shown in the previous chapter, the main result of the model is that the logarithm of the price of an asset is driven by a bubble trend if it follows a super exponential behaviour represented by the so called LPPL function:

$$\ln[p(t)] = A + B(t_c - t)^{\beta} + C(t_c - t)^{\beta}\cos[\omega \ln(t_c - t) + \phi]. \tag{4.1}$$

The seven parameters, $\{A, B, C, t_c, \beta, \omega, \phi\}$ must be calibrated to optimally fit the asset prices on a fixed time window.
If the estimated parameters then satisfy certain constraints, the so called *Sornette bounds* which we will discuss soon, it is possible to state that the asset behaviour is regulated by a bubble regime and the critical time represents the candidate crash time.

The LPPL model theory claims that during a bubble the following relation holds

$$\ln[p_i] = \mathrm{LPPL}(t_i; A, B, C, t_c, \beta, \omega, \phi) + \epsilon_i, \tag{4.2}$$

where $\{t_1, ..., t_n\}$ and $\{p_1, ..., p_n\}$ are the historical dates and prices of the time serie considered, while ϵ_i, for $i = 1, ..., n$, are the residuals of the fitting procedure.

The calibration of the model can be seen as an optimization problem which consists in finding the best set of LPPL parameters, $\{\hat{A}, \hat{B}, \hat{C}, \hat{t}_c, \hat{\beta}, \hat{\omega}, \hat{\phi}\}$, which minimizes a cost function, as shown in Figure 4.1.
It is important to underline that this is a non trivial problem because we are dealing with a non linear minimization problem with seven different parameters.

In this chapter, we describe three different methods for the optimization problem: Ordinary Least Squares (OLS), Generalized Least Squares (GLS) and Maximum Likelihood Estimation (MLE).
These three alternative methods are consistent with each other and they gave coherent results in the practical applications where they were applied.

The Ordinary Least Squares (OLS) is the mostly used in LPPL model framework, since it belongs to the original formulation of the model. The residuals ϵ_i in (4.2) are modelled as a White Noise process, $\mathcal{WN}(0, \sigma^2)$.

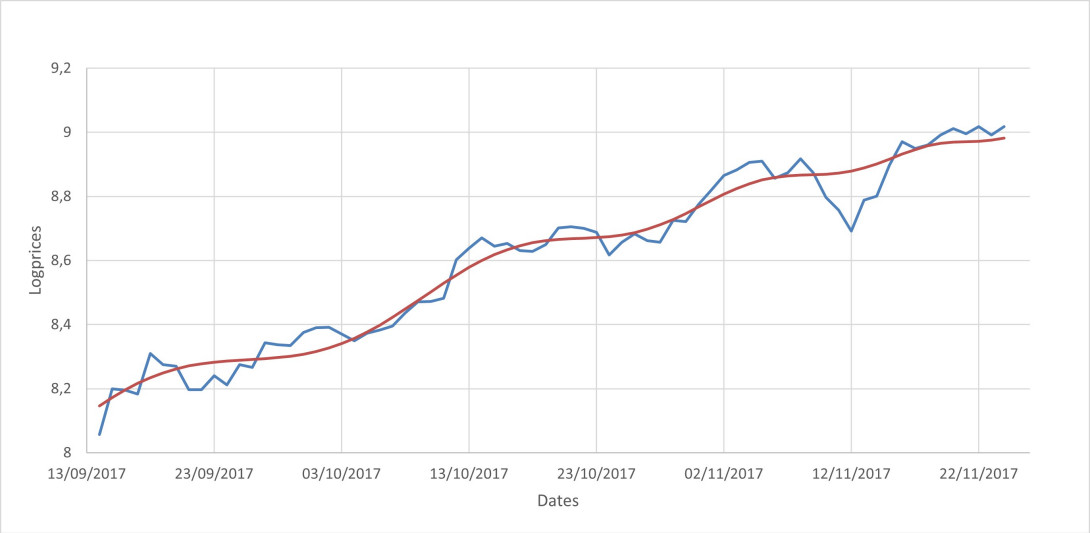

Figure 4.1: Example of calibrated LPPL function to historical logprices.

The Generalized Least Squares (GLS) is a generalization which better applies to the problem. Now the residuals are governed by an autoregressive process AR(1). Namely,

$$\epsilon_i = \rho\epsilon_{i-1} + u_i, \tag{4.3}$$

where $|\rho| < 1$ and $u_i \sim \mathcal{WN}(0, \sigma^2)$.

Finally, the Maximum Likelihood Estimation (MLE) is the most recent version proposed in (Filomonov et al., 2016). It is able to provide interval estimations for the parameters, meaning that it can return a more realistic interval forecast where the burst of the bubble is more probable to happen.

Each method requires to be combined with a nonlinear optimization method. We decide to use the Genetic Algorithm because of its ability to solve the problem without any particular assumption on the shape of the function.
Let us underline the fact that, once the optimal solution is obtained, it is necessary to check that every parameter falls inside the *Sornette Bounds*, otherwise the critical time t_c cannot be considered a valid bubble signal.

Sornette Bounds
Sornette Bounds are a list of constraints which require to be tested in order to deduce whether an optimal set of parameters is a valid bubble signal or not. They come from both mathematical considerations and empirical results on previous bubbles. We present the adopted ranges in Table 4.1, from (Filimonov et al., 2016) and (Scaringi, 2016).

Item	Filtering Condition
A	$A > 0$
B	$B < 0$
C	$\lvert C\rvert < 1, C \neq 0$
t_c	$t_2 < t_c < t_2 + 0.375(t_2 - t_1)$
β	$0 < \beta < 1$
ω	$6 < \omega < 13$
ϕ	$0 \leq \phi \leq 2\pi$
D	$D = \frac{\beta\lvert B\rvert}{\omega C} \geq 0.8$

Now we briefly discuss them in relation to their meaning:

- The parameter A is required to be strictly positive because it is equal to the logprice of the asset in correspondence of the critical time t_c.
- We impose that the amplitude of power law acceleration B to be strictly negative, in order to assure an explosive and accelerating behaviour of the hazard rate.
- The condition on C, $|C| < 1$ and $C \neq 0$, is required in order to control the amplitude of the oscillations.
- The constraint on t_c prevents the predicted critical time to be too distant from the time window and for this reason unreliable.
- The limit $0 < \beta < 1$ is necessary to have an accelerating hazard rate up to t_c but a finite and less than 1 integral up to t for all $t \leq t_c$.
- Empirical studies on previous bubbles (Filimonov et al., 2017) suggest to contain the parameter ω between 6 and 13.
- Finally, the damping factor D is imposed to be bigger than 0.8.

The theoretical assumption that the crash occurs in one immediate negative jump actually requires D to be strictly bigger than 1. However, since this hypothesis is counterfactual and crashes generally have a duration of weeks or even months, the constraint on the damping factor can be relaxed.
It can be interesting to underline that it is not excluded a priori the possibility of negative bubbles, which revert all the arguments discussed above. In this case, we have a general pessimistic sentiment which produces negative feedback and the consequent falling of the price.
All the theoretical foundations remain unchanged, except for the sign of the percentage of price crash which becomes positive according to the possible rebound of the price.
The Sornette Bounds remain unchanged except for the parameter B and the damping factor D which have opposite signs in order to capture the different and opposite price behaviour.

Time Windows
When the model is used with the scope of detecting bubbles, it is important to calibrate it on time windows of different length.
The reason behind this choice is motivatd by the fact that only a subsample of the whole historical serie could actually present signs of log-periodic behaviour. Moreover, as underlined above, the calibration process is a nontrivial task which is strongly influenced by data.
However, it is strongly suggested to consider time windows of at least one month in order to keep an acceptable level of quality in the data inspected.

Before going further, it can be useful to specify the proper "nature" of the historical dates considered. Two options are available: business dates, which do not consider weekends and holidays, and calendar dates, where t continuously increases.
While the price dynamics considered are invariant to this choice, previous works were carried on using calendar dates. This seems to be in line with the formulation of the model where flows of information continuously transmit in the network and so influencing irrational agents, even when trading operations are not allowed.

4.1. Original Calibration Procedure

Estimating LPPL model in general has never been easy, due to the presence of many local minima of the cost function where the minimization algorithm could get trapped.

To find the optimal set of parameters, it is required to minimize the distances between the historical logprices and the LPPL function (Fantazzini et al., 2011). This is done by requiring the minimization of the sum of squared errors (SSE), which corresponds to

$$SSE = \sum_{i=1}^{n} (\ln p_i - LPPL(t_i; A, B, C, t_c, \beta, \omega, \phi))^2. \tag{4.4}$$

Let us recall now the well known LPPL formula:

$$LPPL(t_i; A, B, C, t_c, \beta, \omega, \phi) = A + B(t_c - t_i)^\beta + C(t_c - t_i)^\beta \cos[\omega \ln(t_c - t_i) + \phi]. \tag{4.5}$$

The original calibration technique requires to reduce the number of free parameters by slaving the three linear parameters, {A, B, C}, and compute them from the estimated nonlinear parameters.
More specifically, (4.5) can be rewritten in the following compact way:

$$LPPL(t_i; A, B, C, t_c, \beta, \omega, \phi) = A + Bf_i + Cg_i, \tag{4.6}$$

where $y_i = \ln p_i$, $f_i = (t_c - t_i)^\beta$ and $g_i = (t_c - t)^\beta \cos[\omega \ln(t_c - t) + \phi]$.
It is straightforward to see that the linear parameters A, B and C can be obtained

analytically using Ordinary Least Squares for fixed values of the nonlinear parameters, $\{t_c, \beta, \omega, \phi\}$.
The minimization problem is thus transformed into

$$\{\hat{t}_c, \hat{\beta}, \hat{\omega}, \hat{\phi}\} = \arg \min_{t_c,\beta,\omega,\phi} F_1(t_c, \beta, \omega, \phi), \tag{4.7}$$

where the cost function $F_1(t_c, \beta, \omega, \phi)$ is given by

$$F_1(t_c, \beta, \omega, \phi) = \min_{A,B,C} SSE(t_c, \beta, \omega, \phi, A, B, C). \tag{4.8}$$

Now we have only four parameters to estimate but, because of the non linearity of the function involved, it is required to use a robust method that is able to avoid local minima and converge efficiently to the optimal solution, i.e. optimal set of parameters. To this scope, it is frequently adopted the Genetic Algorithm.

Both the Ordinary Least Squares and the Genetic Algorithm will be discussed in detail below, but first let us present a convenient reformulation of the model which has proved to be very useful in the calibration procedure.

Reformulation of the model
An alternative formulation to the classic LPPL Formula consists in applying a variable change that considerably simplifies the calibration of the parameters (Filimonov et al., 2016)

$$\begin{cases} C_1 = C\cos(\phi) \\ C_2 = C\sin(\phi) \end{cases}, \tag{4.9}$$

in such a way that (4.5) becomes

$$\ln[p(t)] \approx A + B(t_c - t)^\beta + C_1(t_c - t)^\beta \cos[\omega \ln(t_c - t)] + C_2(t_c - t)^\beta \sin[\omega \ln(t_c - t)]. \tag{4.10}$$

Now LPPLS function presents four linear parameters, $\{A, B, C_1, C_2\}$, and three non linear parameters, $\{t_c, \omega, \beta\}$.

As a consequence, the calibration procedure is slightly modified as follows

$$\{\hat{t}_c, \hat{\beta}, \hat{\omega}\} = \arg \min_{t_c,\beta,\omega} F_1(t_c, \beta, \omega), \tag{4.11}$$

where $F_1(t_c, \beta, \omega)$ is now

$$F_1(t_c, \beta, \omega) = \min_{A,B,C_1,C_2} SSE(t_c, \beta, \omega, A, B, C_1, C_2). \tag{4.12}$$

In analogy with the original procedure, (4.12) is solved by OLS and (4.11) by non linear optimization method.

4.1.1. Ordinary Least Square Formulation

The Ordinary Least Square (OLS) method is used in the optimization procedure to deduce the candidate optimal values of the linear parameters, $\{\hat{A}, \hat{B}, \hat{C}_1, \hat{C}_2\}$.
In particular, the OLS method aims to solve the following problem:

$$F_1(t_c, \beta, \omega) = \min_{A,B,C_1,C_2} SSE(t_c, \beta, \omega, A, B, C_1, C_2). \tag{4.13}$$

The fundamental hypothesis behind its application, however, is that residuals ϵ_i in (4.2) are a White Noise process, i.e. $\epsilon \sim \mathcal{WN}(0, \sigma^2)$.
Equation (4.13) has a unique solution obtained for fixed values of the nonlinear parameters:

$$\begin{pmatrix} n & \sum f_i & \sum g_i & \sum h_i \\ \sum f_i & \sum f_i^2 & \sum f_i g_i & \sum f_i h_i \\ \sum g_i & \sum f_i g_i & \sum g_i^2 & \sum g_i h_i \\ \sum h_i & \sum f_i h_i & \sum g_i h_i & \sum h_i^2 \end{pmatrix} \begin{pmatrix} \hat{A} \\ \hat{B} \\ \hat{C_1} \\ \hat{C_2} \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum y_i f_i \\ \sum y_i g_i \\ \sum y_i h_i \end{pmatrix} \tag{4.14}$$

where

$$\begin{aligned} y_i &= \ln p_i, \\ f_i &= |t_c - t_i|^\beta, \\ g_i &= |t_c - t_i|^\beta \cos[\omega \ln(t_c - t_i)], \\ h_i &= |t_c - t_i|^\beta \sin[\omega \ln(t_c - t_i)]. \end{aligned}$$

As discussed previously, the reduction from four to three nonlinear parameters considerably improves the computational complexity of the model.

4.1.2. Genetic Algorithm

The Genetic Algorithm (GA) is an algorithm inspired by Darwin's "survival of the fittest" and its theory was developed by John Holland in 1975. It aims at mimicking the natural selection in biological systems, which is governed by survival of the fittest individuals, breeding and mutation.
The main advantage of this method is that it does not require any particular constraint about the continuity or smoothness of the cost function, so it can be applied to a wide range of class of functions.
Moreover, it prevents the calibration to get trapped in local minima, so it is particularly suitable for the LPPL case. This feature is determined by mutations that play a key role in the evolution of a species, since they may increase its probability of survival, as well as introduce less favorable characteristics.

Its procedure (Fantazzini et al., 2011) consists of four steps:

1. *Initial population:* Consider a population where each member is represented as a vector of the three nonlinear coefficients, $[t_c, \beta, \omega]$. Each parameter is randomly

drawn from a uniform distribution with a pre-specified range, then for each member compute the corresponding estimates of linear parameters, $[A, B, C_1, C_2]$, and the value of the cost function.

2. *Selection mechanism:* From the initial population, select a percentage $\%_s$, the parents population, according to the lower values of the cost function.

3. *Breeding mechanism:* A percentage $\%_b$ of offsprings is obtained by randomly drawing two parents, without replacement, and taking the arithmetic mean of them in order to obtain a new offspring solution. Note that one parent can generate offsprings with different partners, so betrayals are allowed.

4. *Mutation mechanism:* In order to allow new regions of the search space to be explored so that premature convergence in local minima is avoided, a percentage $\%_m = 1 - \%_s - \%_b$ is implemented by adding random noise variables to the value of each non linear parameter. The effect of the noise random variable on the previous solution is very important, indeed large perturbations can actually prevent the algorithm from finding the optimal solution, while small variations can leave the search trapped inside a local minimum.

5. *Merging mechanism:* The three groups are then merged together obtaining a population having the same size of the initial one.

The procedure is then repeated until some termination criteria are met.

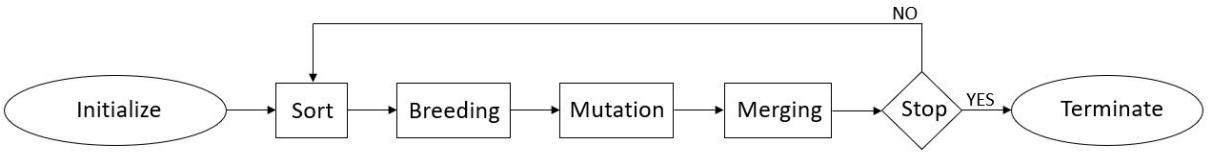

Figure 4.2: Genetic Algorithm procedure.

In the next chapter, some practical information on the setup adopted is given in order to ease the replication of the results obtained. Nonetheless, it can be interesting to report some considerations:
First of all, the setup of the percentages, $\%_s$, $\%_b$, $\%_m$, that we recall has to be in such a way that $\%_s + \%_b + \%_m = 1$, is a delicate matter. Indeed, higher values of $\%_s$ could cause little change in the population across the generations and so a slower convergence, while lower values mean that fewer members are kept for the next generation.
Mutation and Breeding are the two drivers of the Genetic Algorithm, but, while Mutation is able to explore new regions of the search space, Breeding searches new solutions locally, as it is the arithmetic mean of parent solutions. An higher $\%_m$ percentage results in more useless results, since mutated solutions often have higher values in the cost function, while an higher $\%_b$ percentage increases the risk to trap the algorithm in a local minimum.

A general rule for the ideal proportion of $\%_s$, $\%_b$ and $\%_m$ does not exist, since they have to be set specifically for each problem.
Finally, the size of the population and the maximum number of iterations are fundamental in order to allow the GA to find the optimal solution, but take into account that high values require an higher computational effort.

4.2. Other variants and generalizations

In this section, some interesting variations to the original formulation of the calibration procedure are proposed with the aim of finding more reliable and consistent results, once these methods are put into practice.
In particular, we propose the Generalized Least Square method (GLS) as a generalization of the Ordinary Least Square one and the Maximum Likelihood Estimation method which provides us with interval estimations of the parameters instead of punctual results.

4.2.1. Generalized Least Squares Formulation

Even if the OLS version is the most adopted in literature, it is just a raw attempt in the non trivial estimation of the parameters.
Indeed, the Ordinary Least Square method assumes that errors are homoskedastic and uncorrelated, but for the short term dynamics of the price movement it is observed that residuals actually show autocorrelation and heteroskedasticity (Fantazzini et al., 2011).

We now present a more developed approach which relies on Generalized Least Square (GLS) method. It was first proposed in (Bianchetti et al., 2018) and widely illustrated in (Scaringi, 2016).

The GLS approach aims to solve the same problem of the OLS case, hence recall that during a bubble price regime the following relation holds

$$\ln[p_i] = \mathrm{LPPL}(t_i; A, B, C_1, C_2, t_c, \beta, \omega) + \epsilon_i, \tag{4.15}$$

where in this case $\{\epsilon_i\}$ follows an $AR(1)$ process, namely

$$\epsilon_i = \rho\epsilon_{i-1} + u_i. \tag{4.16}$$

Note that the reformulation of the model presented for the OLS calibration procedure is still adopted, therefore the calibration of the model requires the estimation of four linear parameters, $\{A, B, C_1, C_2\}$, and three nonlinear parameters, $\{t_c, \beta, \omega\}$.

Moreover, the optimization is still divided in two different stages

$$\{\hat{t}_c, \hat{\beta}, \hat{\omega}\} = \arg\min_{t_c,\beta,\omega} F^{GLS}(t_c, \beta, \omega), \tag{4.17}$$

where the cost function $F^{GLS}(t_c, \beta, \omega)$ corresponds to

$$F^{GLS}(t_c, \beta, \omega) = \min_{A,B,C_1,C_2} SSE^{GLS}(t_c, \beta, \omega, A, B, C_1, C_2). \tag{4.18}$$

It is important to underline that residuals now do not follow anymore a White Noise process, so we cannot use the previous cost functions for the calibration.
The new approach relies on Cochrane-Orcutt procedure specifically applied to the LPPL case:

1. First, consider the relation (4.15) and express it in the linear form previously adopted for the OLS version

$$y_i = A + Bf_i + C_1 g_i + C_2 h_i + \epsilon_i, \tag{4.19}$$

   where

$$\begin{aligned} y_i &= \ln p_i, \\ f_i &= |t_c - t_i|^\beta, \\ g_i &= |t_c - t_i|^\beta \cos[\omega \ln(t_c - t_i)], \\ h_i &= |t_c - t_i|^\beta \sin[\omega \ln(t_c - t_i)]. \end{aligned}$$

2. Apply the OLS method producing a first estimation of the linear parameters, $\{\hat{A}, \hat{B}, \hat{C}_1, \hat{C}_2\}$. However, since residuals ϵ_i are now distributed as an AR(1) process, it is required to estimate the parameter ρ too.

3. In order to compute this additional parameter, construct the process of estimated residuals by subtraction. Namely,

$$\hat{\epsilon}_i = y_i - \hat{A} - \hat{B}f_i - \hat{C}_1 g_i - \hat{C}_2 h_i. \tag{4.20}$$

   The AR(1) dynamics still hold, hence

$$\hat{\epsilon}_i = \rho \hat{\epsilon}_{i-1} + u_i, \tag{4.21}$$

   with u_i being a White Noise.

4. Apply Yule-Walker formula for autoregressive processes (Adhikari et al., 2013, pag. 19), in order to deduce

$$\hat{\rho} = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)}, \qquad \hat{\gamma}(k) = \mathrm{Cov}(\epsilon_i, \epsilon_{i+k}).$$

   $\hat{\gamma}(k)$ is the sample autocovariance of order k and $\hat{\gamma}(0)$ is the sample variance.

5. Now, the following quasi difference relation holds

$$y_i - \hat{\rho} y_{i-1} = \hat{A}(1-\hat{\rho}) + \hat{B}(f_i - \hat{\rho} f_{i-1}) + \hat{C}_1(g_i - \hat{\rho} g_{i-1}) + \hat{C}_2(h_i - \hat{\rho} h_{i-1}) + (\epsilon_i - \hat{\rho}\epsilon_{i-1}), \tag{4.22}$$

or more clearly

$$y_i^* = A^* + \hat{B} f_i^* + \hat{C}_1 g_i^* + \hat{C}_2 h_i^* + \epsilon_i^*, \qquad i = 2, ..., n \tag{4.23}$$

with

$$\begin{aligned} A^* &= \hat{A}(1-\hat{\rho}), \\ f_i^* &= f_i - \hat{\rho} f_{i-1}, \\ g_i^* &= g_i - \hat{\rho} g_{i-1}, \\ h_i^* &= h_i - \hat{\rho} h_{i-1}, \\ \epsilon_i^* &= \epsilon_i - \hat{\rho}\epsilon_{i-1} = u_i. \end{aligned}$$

6. Finally, ϵ_i^* is a White Noise process, therefore we can compute OLS estimation of the linear parameters, $\{A^*, \hat{B}, \hat{C}_1, \hat{C}_2\}$

$$SSE^{GLS}(t_c, \beta, \omega, A, B, C_1, C_2) = \sum_{i=2}^{n} [y_i^* - A^* - \hat{B} f_i^* - \hat{C}_1 g_i^* - \hat{C}_2 h_i^*]^2. \tag{4.24}$$

7. Once obtained the linear parameters $\{\hat{A}, \hat{B}, \hat{C}_1, \hat{C}_2\}$, where

$$\hat{A} = \frac{A^*}{1-\hat{\rho}},$$

solve the optimization problem (4.18) iterating the previous steps with the new estimates of the linear parameters, until ρ reaches convergence.

8. Eventually, solve (4.17) via Genetic Algorithm, which holds identically as in the OLS case, or an equivalent non linear optimization method.

The algorithm has to be repeated until the parameter ρ achieves convergence. Usually this happens quite soon, approximately between five and ten iterations, as shown in Figure 4.3.

In the next chapter, we will show how this more advanced method provides more reliable results with respect to the OLS case.

4.2.2. Maximum Likelihood Formulation

Sornette and others developed in 2016 a new bubble detection approach, with the idea to find an alternative to the punctual estimation of the parameters (Filimonov et al., 2016).
This new approach, called Maximum Likelihood Estimation, is able to provide interval

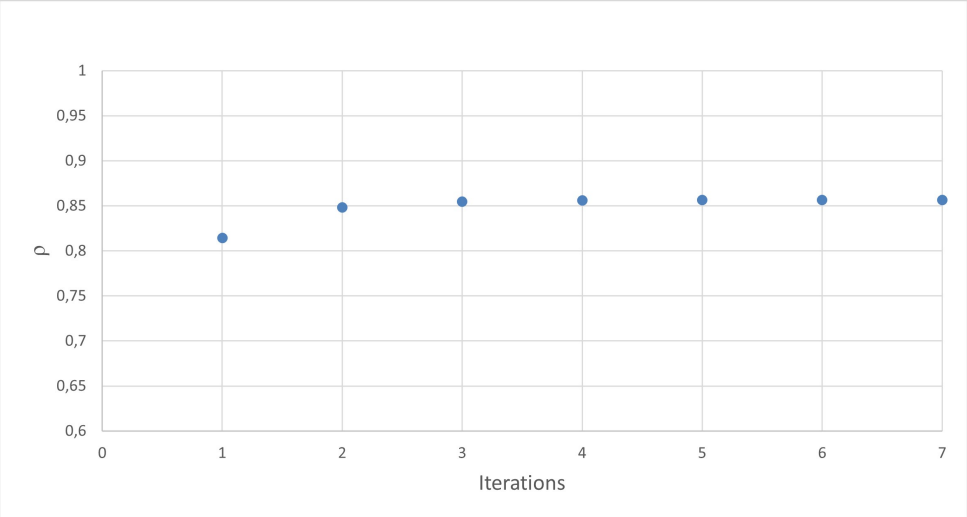

Figure 4.3: ρ convergence in GLS method.

estimations, giving a specific focus on the critical time t_c and addressing the others only as nuisance parameters.

First, we recall the well known relation between historical logprices and model prices

$$\ln p(t_i) = \mathrm{LPPL}(t_i; t_c, \psi) + \epsilon(t_i),$$

where $\psi = [A, B, C_1, C_2, \beta, \omega]$.

This approach relies on the fundamental assumption that the residuals $\epsilon(t_i)$ are normally distributed. Under this hypothesis, the Likelihood function has a well known form:

$$L(t_c, \psi, s) = \frac{1}{(2\pi s)^{n/2}} e^{-\frac{SSE(t_c,\psi)}{2s}}, \tag{4.25}$$

where n is the number of points in the time serie and s is the sample variance of $\epsilon(t_i)$. This method consists in estimating the parameters of a statistical model given historical observations, by finding the parameter values that maximize the likelihood of making these observations.

As underlined many times in the previous chapter, in practical applications of LPPL model the most important result is the value of the critical time t_c, which identifies the forecast of the burst of the bubble.
The MLE approach shares this view as it considers t_c as the main parameter and the others just in an ancillary role.

In the Bayesian approach, the elimination of nuisance parameters is a well known problem which can be generally solved integrating them out. However, it is required to specify a priori the distribution of all the parameters, calculating the posterior and then

integrating out the nuisance parameters ψ in order to derive the posterior marginal distribution of t_c.
In the case of LPPL model, this procedure cannot be implemented directly as there is no information to provide a priori distribution for the six nuisance parameters of the model. Nonetheless, a simpler alternative consists in the so called *Profile Likelihood Function*:

$$L_p(t_c) = \max_{\psi} L(t_c, \psi) = L(t_c, \hat{\psi}), \tag{4.26}$$

where $\hat{\psi}$ is the MLE of the nuisance parameters at each fixed value of the parameter of interest t_c. Namely,

$$\hat{\psi}_{t_c} = \arg\max_{\psi} L(t_c, \psi). \tag{4.27}$$

The profile likelihood approach is technically identical to the following procedure:

1. For fixed values of the nonlinear paramteres t_c, β and ω find via OLS method

$$F_1(t_c, \beta, \omega) = \min_{A,B,C_1,C_2} SSE(A, B, C_1, C_2, t_c, \beta, \omega).$$

2. Now, instead of minimizing the three nonlinear parameters at the same time, keep t_c fixed and compute

$$F_2(t_c) = \min_{\beta,\omega} F_1(t_c, \beta, \omega).$$

3. Eventually,

$$\hat{t}_c = \arg\min_{t_c} F_2(t_c).$$

In general, such extra subordination dramatically reduces the number of local extrema of the cost function. Moreover, it allows one to analyze the whole profile of the cost function $F_2(t_c)$ with all the extrema detected and the corresponding values of $\beta(t_c)$ and $\omega(t_c)$.
For sake of completeness, let us give a formula for the estimation of $\hat{s}_{t_c}$ too

$$\hat{s}_{t_c} = \frac{SSE(t_c, \hat{\psi})}{n} = \frac{F_2(t_c)}{n}. \tag{4.28}$$

As anticipated above, this approach is able to go over the idea of punctual estimation of parameters, since it is able to provide time intervals where the burst of the bubble is more probable to be detected.
It is interesting how this new approach is able to overcome the limits of the so called *Sornette Bounds*, built both on mathematical considerations but also on previous bubble detection studies.
The main critic to the use of *Sornette Bounds* is that they heavily rely on previous results, which is also the reason for different versions of them through the years and authors, for example compare (Filimonov et al., 2016), (Bingcun et al., 2018), (Fantazzini et al., 2011).

Anyway, past experience may not contain all possible situations and it is not possible to be completely sure about the reliablity of the limits chosen.

Even if the profile likelihood approximation is often treated as a regular likelihood, actually it is not a genuine likelihood function. In particular, it treats the nuisance parameters at fixed $\hat{\psi}_{t_c}$ as if they were known. It may thus overstate the amount of information about t_c and under certain conditions it can provide unstable estimates with respect to small changes in the observed data.

In order to overcome this limitation of the profile likelihood, an adjusted version has been proposed, namely the *Modified Profile Likelihood*. It consists in adding an extra factor $M(t_c)$, called modulating factor, to the profile likelihood

$$L_m(t_c) = M(t_c)L_p(t_c) = \left|I(\hat{\psi}_{t_c})\right|^{-\frac{1}{2}} \left|\frac{d\hat{\psi}}{d\hat{\psi}_{t_c}}\right| L_p(t_c), \tag{4.29}$$

where we recall that $\hat{\psi}_{t_c}$ is the MLE estimation of the nuisance parameters for fixed t_c. $I(\hat{\psi}_{t_c})$ is the Observed Fisher Information matrix on ψ assuming t_c is known

$$I(\hat{\psi}_{t_c}) = -\frac{d^2 \ln L(t_c, \psi)}{d\psi\, d\psi^\top}\bigg|_{\psi=\hat{\psi}_{t_c}},$$

while $\frac{d\hat{\psi}}{d\hat{\psi}_{t_c}}$ is the Jacobian matrix of the full MLE of the nuisance parameters ψ, with respect to their MLE calculated for a fixed vaue of t_c. Finally, $|\cdot|$ denotes the absolute value of a matrix determinant.

Even if the Modified Profile Likelihood is a reliable approximation of the genuine likelihood, it is extremely difficult to compute. In particular, the Jacobian Matrix in (4.29).
For this reason, we will use an approximated version, but first it is useful to express the Jacobian as

$$J(t_c) = \left|\frac{d\hat{\psi}}{d\hat{\psi}_{t_c}}\right| = \frac{\left|I(\hat{\psi}_{t_c})\right|}{\left|C(t_c, \hat{\psi}_{t_c}; \hat{t}_c, \hat{\psi})\right|}, \tag{4.30}$$

where $C(t_c, \hat{\psi}_{t_c}; \hat{t}_c, \hat{\psi})$ is a matrix composed by the second order derivatives of a log-likelihood $L(t_c, \hat{\psi}_{t_c}; \hat{t}_c, \hat{\psi}, a)$ with a new parameter such that $\{\hat{t}_c, \hat{\psi}, a\}$ is a sufficient statistic of the model

$$C(t_c, \hat{\psi}_{t_c}; \hat{t}_c, \hat{\psi}) = \frac{d^2 \ln L(t_c, \hat{\psi}_{t_c}; \hat{t}_c, \hat{\psi}, a)}{d\hat{\psi}_{t_c}\, d\hat{\psi}^\top}.$$

The computation of the second-order matrix C is a non trivial task, hence the following approximation (Filomonov et al., 2016) is adopted

$$C(t_c, \hat{\psi}_{t_c}; \hat{t}_c, \hat{\psi}) \approx \Sigma(t_c, \hat{\psi}_{t_c}; \hat{t}_c, \hat{\psi}), \tag{4.31}$$

where Σ is the covariance matrix of score functions.

Finally, we get the approximated formulation of $L_m(t_c)$ that we anticipated before

$$L_m(t_c) \approx \frac{\left|I(\hat{\psi}_{t_c})\right|^{\frac{1}{2}}}{\left|\Sigma(t_c, \hat{\psi}_{t_c}; \hat{t}_c, \hat{\psi})\right|} L_p(t_c). \tag{4.32}$$

To express it in a clearer way, omitting terms that do not depend on t_c

$$L_m(t_c) \propto \frac{(\hat{s}_{t_c})^{\frac{-(n-p-2)}{2}} \left|\sum_{i=1}^{n} \frac{d^2 LPPLS(t_i; t_c, \psi)}{d\psi\, d\psi^T}\right|_{\psi=\hat{\psi}_{t_c}}^{1/2}}{\left|\sum_{i=1}^{n} \frac{dLPPLS(t_i; t_c, \psi)}{d\psi}\right|_{\substack{t_c=t_c \\ \psi=\hat{\psi}_{t_c}}} \frac{dLPPLS(t_i; t_c, \psi)}{d\psi^T}\Big|_{\substack{t_c=\hat{t}_c \\ \psi=\hat{\psi}}}\right|}, \tag{4.33}$$

where $p = \dim \psi = 6$, $\hat{\psi}_{t_c}$ is a vector of MLE estimates of nuisance parameters for a fixed value of t_c, while $\hat{s}_{t_c}$ is given by (4.28).

Let us define the rectangular $n \times p$ matrix

$$X_{ij}(t_c, \psi) = \frac{dLPPLS(t_i; t_c, \psi)}{d\psi_j}, \tag{4.34}$$

and the $p \times p$ matrix

$$H_{ij}(t_c, \psi) = \sum_{k=1}^{n} (\ln p(t_k) - LPPLS(t_k; t_c, \psi)) \frac{d^2 LPPLS(t_k; t_c, \psi)}{d\psi_i \psi_j}, \tag{4.35}$$

where ψ_j is the j-th element of the vector of the nuisance parameters, $[A, B, C_1, C_2, \beta, \omega]$.

Eventually, we can rewrite formula (4.33) as a function of matrices X and H as

$$L_m(t_c) \propto \frac{\left|X^T(t_c, \hat{\psi}_{t_c}) X(t_c, \hat{\psi}_{t_c}) - H(t_c, \hat{\psi}_{t_c})\right|^{1/2}}{\left|X^T(\hat{t}_c, \hat{\psi}) X(t_c, \hat{\psi}_{t_c})\right|} (\hat{s}_{t_c})^{-(n-p-2)/2}. \tag{4.36}$$

Extended formulas of the partial derivatives used in (4.34) and (4.35) are reported in Appendix A.

For practical applications, it is more useful to consider the *Relative Modified Profile Likelihood*, which takes value in $[0; 1]$

$$R_m(t_c) = \frac{L_m(t_c)}{\max_{t_c} L_m(t_c)}. \tag{4.37}$$

Thanks to it, we can finally identify the likelihood interval at 5% cutoff

$$LI(t_c) = \left\{ t_c : R_m(t_c) = \frac{L_m(t_c)}{L_m(\hat{t}_c)} > 0.05 \right\}. \tag{4.38}$$

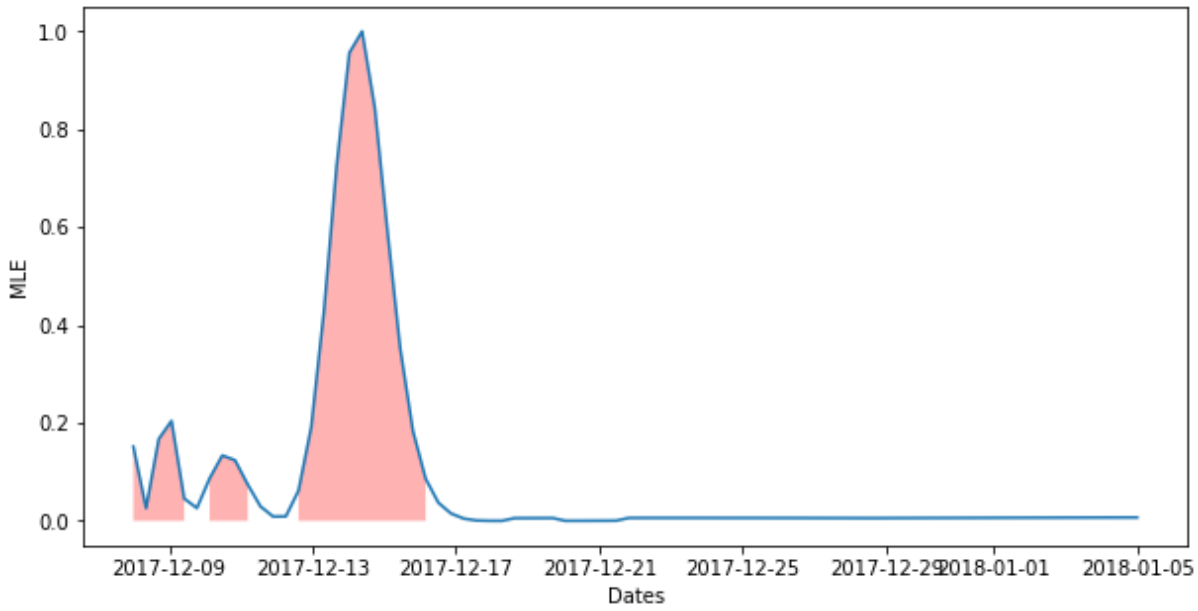

Figure 4.4: Example of Relative Modified Profile Likelihood function, $R_m(t_c)$*, with likelihood intervals at* 5% *cutoff,* $LI(t_c)$.

This is a confidence interval for the critical time t_c, where the punctual estimation corresponds to $\hat{t}_c$ such that $R_m(\hat{t}_c) = 1$.

5. Numerical Results

In this chapter, we will provide numerical results on real financial time series, using the methods presented in Chapter 4. In particular, we will deal with two different kind of analysis:
First, a post-mortem analysis where crashes in the time series are studied after they already occurred. Specifically, we will analyze Bitcoin historical prices between December 2016 and January 2018, when Bitcoin prices skyrocketed from 1'000 USD to almost 20'000 USD.
Secondly, we will carry on a real-time bubble detection study with the aim to forecast crashes caused by bubble behaviours. Again, we will focus on Bitcoin's time series in December 2017, in order to predict the crash that halved Bitcoin's price in roughly one month.

5.1. Post-mortem Analysis

A post-mortem analysis on crashes that already occurred clearly does not have a forecasting ambition, but instead it aims to analyze the time series in order to identify drawdown and drawup periods and, in particular, to establish whether the drawup phase was caused by a normal growth or by an irrational behaviour.
In order to do that, we will follow the procedure proposed in (Demos, Sornette et al., 2018):

1. Identify the most relevant peak dates in the time series, in order to apply the bubble detection method on the subsequent crashes.

2. Distinguish between drawdown and drawup periods within the time series and evaluate their magnitude.

3. Establish whether the drawup periods were governed by a bubble behaviour or by normal growth.

This approach has been applied to Bitcoin's historical prices between December 1^{st} 2016 and January 16^{th} 2018, to try to replicate the results obtained in (Bianchetti et al., 2018), where coherent bubble signals were detected in the second half of December 2017.

Nonetheless, first we need to introduce some theoretical concepts, that are required in the analysis presented above.

5.1.1. Theoretical background

In this section, we will present the Epsilon Drawdown/Drawup method, used to distinguish within the time series between drawdown and drawup phases, and the Lagrange Regularisation approach for the determination of bubble start times, applied on drawup phases in order to estimate the optimal start date of a bubble behaviour, if there exists any.

5.1.1.1. Epsilon Drawdown/Drawup method

The Epsilon Drawup/Drawdown method is applied on historical price series, in order to distinguish between drawdown and drawup periods.
It was first developed by Sornette and others while dealing with post-mortem analysis of financial bubbles (Demos, Sornette et al., 2018). Their idea was to find a set of peak dates where the asset experienced a consistent growth phase and test the Log Periodic Power Law model, extensively presented in Chapter 3 and 4, in order to establish if the consequent crashes would have been caused by a bubble behaviour.

The Epsilon Drawup/Drawdown method consists of the following steps:

1. Calculate the *daily logreturns*

$$r_i = \ln p(t_i) - \ln p(t_{i-1}) = \ln \frac{p(t_i)}{p(t_{i-1})} \qquad \forall i \geq 1, \tag{5.1}$$

   where $p(t_i)$ is the closing price at time t_i, $t_i = t_0 + i\Delta_t$ and $\Delta_t = 1$ day.

2. The first date t_0 is the beginning of a drawup (drawdown) phase, indexed by t_s, if $r_1 > 0$ ($r_1 < 0$). Then, compute the *cumulative return up to* t_i as

$$p_{s,i} = \sum_{k=s}^{i} r_k = \ln p(t_i) - \ln p(t_s) \qquad \forall t_i > t_s. \tag{5.2}$$

3. For any t_i verify whether the current drawup (drawdown) phase is still active or not, by computing the *largest deviation* $\delta_{s,i}$ of the price trajectory from its previous maximum (minimum)

$$\delta_{s,i} = \begin{cases} \max\limits_{s \leq k \leq i} p_{s,k} - p_{s,i} & \text{if drawup,} \\ p_{s,i} - \min\limits_{s \leq k \leq i} p_{s,k} & \text{if drawdown.} \end{cases} \tag{5.3}$$

   If the deviation exceeds a predefined tolerance ϵ at time t_i, namely

$$\delta_{s,i} > \epsilon, \tag{5.4}$$

   the procedure has to be stopped. Indeed, the tolerance ϵ quantifies how much the price is allowed to move in the direction opposite to the drawup/drawdown behaviour.

4. Once the procedure has been stopped at time t_i, identify the actual end of the current drawup (drawdown) phase, t_e, which actually corresponds to the highest (lowest) price found:

$$e = \begin{cases} \arg\max_{s \le k \le i} p_{s,k} & \text{if drawup,} \\ \arg\min_{s \le k \le i} p_{s,k} & \text{if drawdown.} \end{cases} \tag{5.5}$$

The algorithm is then restarted setting $t_s = t_e + \Delta_t$, until the whole time serie is divided in drawups and drawdowns. Note that a drawup phase is always followed by a drawdown and viceversa, by construction of $\delta_{s,i}$.

Finally, a set of peak times $\{t_1^p, t_2^p, ...\}$ is obtained taking the right extremum of each drawup interval. These dates can be regarded as peaks of candidate bubbles.

It is important to underline, however, that the results obtained are very sensitive to the tolerance ϵ adopted, so it is expressed as function of the volatility σ and a constant multiplier ϵ_0:

$$\epsilon = \epsilon_0 \cdot \sigma, \tag{5.6}$$

in particular, the volatility σ corresponds to the annualized standard deviation of the logreturns of the time series. Namely

$$\sigma = \sqrt{252} \cdot s_r, \qquad s_r = \frac{\sum_{i=1}^{n} (r_i - \bar{r})^2}{n-1}, \tag{5.7}$$

with $\bar{r}$ mean of the logreturns, $\{r_i\}$.

In order to find a coherent set of peak dates, the Epsilon Drawup/Drawdown method is repeated for different values of $\{\epsilon_0\}_j$ for $j = 1, ..., N_\epsilon$.

For each ϵ_j, a corresponding set of peak dates is found with the algorithm presented above

$$\Omega_j = \{t_1^p, t_2^p, ...\}_j \qquad j = 1, ..., N_\epsilon. \tag{5.8}$$

The collection of all the elements that occur at least once is then the union of all the sets Ω_j

$$\Omega = \bigcup_{j=1}^{N_\epsilon} \Omega_j. \tag{5.9}$$

To find the most valuable peak dates in the collection Ω, we count the number of times each element t_i^p occurred over all trials:

$$N_{t_i^p} = \sum_{j=1}^{N_\epsilon} I_j(t_i^p), \tag{5.10}$$

where $I_j(t_i^p)$ is the Indicator function

$$I_j(t_i^p) = \begin{cases} 1 & \text{if } t_i^p \in \Omega_j, \\ 0 & \text{otherwise.} \end{cases} \tag{5.11}$$

Then, each $N_{t_i^p}$ is normalized by dividing it for the total number of tested pairs N_ϵ. Values are grouped in the set Λ, where each element corresponds to the fraction of occurrences of each peak date with respect to the total number of trials

$$\Lambda = \{n_{t_i^p} = \frac{N_{t_i^p}}{N_\epsilon}, t_i^p \in \Omega\}. \tag{5.12}$$

Finally, to select the most relevant peak dates, a threshold Λ_T is set and only peak dates that occur more frequently than Λ_T are taken. Namely

$$\Omega_T = \{t_i^p : n_{t_i^p} > \Lambda_T, t_i^p \in \Omega\}. \tag{5.13}$$

These are the candidate peaks where the LPPL model is tested in order to detect potential bubble behaviours of the underlying asset.

5.1.1.2. Lagrange Regularisation approach

The LPPL model is considered to be reliable in the modelization of log-price dynamics only if the underlying asset has entered a bubble behaviour. For this reason, trying to calibrate the LPPL parameters on windows corresponding to phases of normal price growth could lead to spurious estimates.
As a consequence, it is important to develop a method able to identify the beginning of a bubble behaviour and to apply calibration techniques only on time windows starting after this date.
The Lagrange Regularisation approach, presented for the first time in (Demos, Sornette, 2017), is an attempt to provide a solution to this problem.

As recalled in Chapter 4, calibration techniques are usually made on many different time windows, in order to find consistency in the results obtained.
In the case of post-mortem analyses, time windows are selected with fixed end time at the peak of the drawup period, varying the start date across windows of different length.
The crucial idea behind the Lagrange Regularisation Approach is that, for fits conducted with fixed end time t_2 but at increasing window lengths Δt_i, i.e. moving backward the start date t_1, the Average Sum of Squared Errors

$$\chi^2(t_1) = \frac{SSE(t_1)}{N} = \frac{\sum_{i=1}^{N}(\ln p_i - LPPL(t_i; t_c, \psi))^2}{N}, \tag{5.14}$$

where we recall $\psi = [A, B, C_1, C_2, \beta, \omega]$, exhibits an approximate linear behaviour as a function of the window length.

In other words, small windows tend to give smaller χ^2 values.
Eventually, note that $SSE(t_1)$ in (5.14) corresponds to Formula (4.4) in the original calibration technique.

If we choose the optimal fit t_1^* as the one minimizing $\chi^2(t_1)$,

$$t_1^* = \arg\min_{t_1} \chi^2(t_1), \tag{5.15}$$

small windows would be favoured, hence they would be more likely to be selected.

Therefore, in order to compare fits, $\chi^2(t_1)$, performed for different window sizes and determine the optimal one, we add a penalty term addressing for the size of the fit window (Demos, Sornette, 2017). Namely,

$$\chi^2_\lambda = \chi^2 - \lambda(t_2 - t_1), \tag{5.16}$$

hence the optimal fit t_1^* is obtained as

$$t_1^* = \arg\min_{t_1} \chi^2_\lambda(t_1), \tag{5.17}$$

where λ is estimated empirically via linear regression of $\chi^2(t_1)$ on $(t_2 - t_1)$.
This correction is required in order to consider detrended residuals over all window sizes. At this point, we take the smallest residual as start date of the bubble behaviour.

All the data preceding the bubble start date need to be excluded from calibration purposes, as they could return factitious results.

5.1.2. Bitcoin's analysis between December 2016 and January 2018

In this section, we will try to replicate the results reported in (Bianchetti et al., 2018) on Bitcoin historical prices between December 1^{st} 2016 and January 16^{th} 2018, represented in Figure 5.1.

Without discussing in detail the intrinsic value of Bitcoin as digital currency or commodity, for which we recall (Antonopoulos, 2017) and (Ametrano, 2016), we will just deal with its time series, in order to assert whether Bitcoin entered an irrational behaviour and to try to identify the beginning of the bubble period.
As anticipated above, we will follow the approach proposed in (Demos, Sornette et al., 2018), that we recall for simplicity:

1. Detection of peak dates via Epsilon Drawup/Drawdown method.
2. Partition between drawup and drawdown methods.
3. Identification of bubble phases via Lagrange Regularisation approach.

Figure 5.1: Bitcoin's historical prices from 1^{st} *December 2016 to January* 16^{th} *2018.*

1. Detection of peak dates
The key idea behind the detection of peak dates within the time series is the assumption that a bubble is identified as a price run-up.
In this scenario, a drawup is just a sequence of positive returns that may be interrupted by negative returns smaller than a pre-specified tolerance, ϵ. Analogously, a drawdown is a sequence of negative returns that may be interrupted by positive returns no larger in amplitude than ϵ.
To find the most relevant peak dates in the time series, we apply the Epsilon Drawdown/Drawup method, whose detailed presentation is reported in the previous section.

First of all, we compute the annualized volatility of the time series. It is interesting to note how Bitcoin appears to be much more volatile compared to other traditional asset classes during the same period of analysis. At the same time, it appears to be in line with other criptocurrencies like Ether.

Asset Class	Volatility
Bitcoin	76.09%
World Equity Index	6.80%
Global HY Index	5.40%
Gold	9.98%
Ether	81.28%

We apply the Epsilon Drawup/Drawdown method for different values of ϵ_0, $\{\epsilon_0 = 0.1, 0.11, ..., 0.59, 0.6\}$ with a frequency threshold $\Lambda_T = 0.6$. The found peak dates, such that $\Lambda > \Lambda_T$, are:

Peak dates	Λ
04/01/2017	0.65
11/06/2017	0.92
01/09/2017	0.75
16/12/2017	0.86

Column Λ represents the frequency of each peak, i.e. the fraction of occurrences with respect to the total number of trials.
The most remarkable peak in the time series is surely the one in the second half of December 2017, where the price of Bitcoin skyrocketed to almost 20'000 USD, however there are three other candidates that require to be analyzed. Figure 5.2 shows the detected peaks along the time series.

Figure 5.2: Bitcoin's peak dates found via Epsilon Drawdown/Drawup method.

2. Partition between drawdown and drawup periods

Following the Epsilon Drawdown/Drawup method, we can distinguish between drawdown and drawup periods. A drawdown shall be a correction regime or a sharp crash, in any case it starts the day after a peak is reached and it ends the day the price reaches its minimum value over the period from the beginning of the drawdown up to the next peak.
Analogously, periods starting the day after the end of a drawdown and finishing with a peak date are identified as drawup periods. Within them, we can find phases governed either by normal growth or by an irrational behaviour.

Drawdown phases	Return	Duration
05/01/2017; 11/01/2017	−18.75%	7 days
12/06/2017; 16/07/2017	−25.75%	35 days
02/09/2017; 14/09/2017	−25.94%	13 days
17/12/2017; 16/01/2018	−43.99%	31 days

The table above summarizes the drawdown periods found, which are also represented in Figure 5.3.
Column *Return* reports the negative performances registered, while *Duration* indicates their length. It is interesting to see how prices could either enter a correction regime lasting many weeks or have a violent crash with the asset loosing a consistent percentage of its value in a couple of weeks.

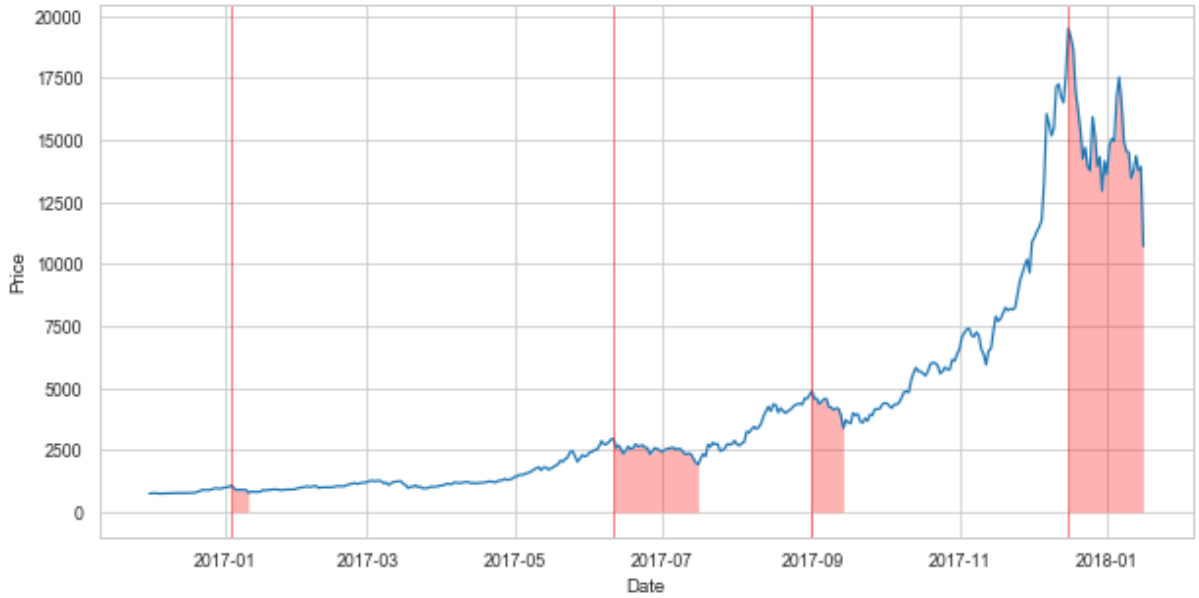

Figure 5.3: Bitcoin's drawdown phases between December 1^{st} *2016 and January* 16^{th} *2018.*

3. Identification of bubble phases

It remains to verify the existence of bubble behaviours during drawup phases and, if there is any, estimate its start date via Lagrange Regularisation approach.
To detect for bubbles we apply the OLS version to many different time windows, setting the right extremum, t_2, at the bubble peak and moving backwards the left one, t_1. Moreover, we impose a minimum window length of at least 30 days.
Drawup periods are synthesized in the table below:

Drawup phases	Return	Duration	Bubble
01/12/2016; 04/01/2017	45.09%	35 days	N
12/01/2017; 11/06/2017	257.42%	151 days	N
17/07/2017; 01/09/2017	124.51%	47 days	N
15/09/2017; 16/12/2017	423.75%	93 days	Y

Analogously to the drawdown case, column *Return* reports the positive performance registered, while *Duration* indicates the length of each period. In addition, the column *Bubble* underlines whether a bubble behaviour is detected or not.

Clear bubble signals are found in the time window spanning from September 15^{th} 2017 to December 16^{th} 2017. In particular, we are able to estimate the start date of the bubble behaviour on October 14^{th} 2017, thanks to Lagrange Regularisation approach. Coherent bubble signals were detected in the second half of December also in (Bianchetti et al., 2018), correctly forecasting the subsequent crash which almost halved Bitcoin prices in approximately one month.

To conclude, Figure 5.4 includes these final results. Specifically, red areas represent drawdowns, grey areas are for normal growth phases and green ones show bubble behaviour.

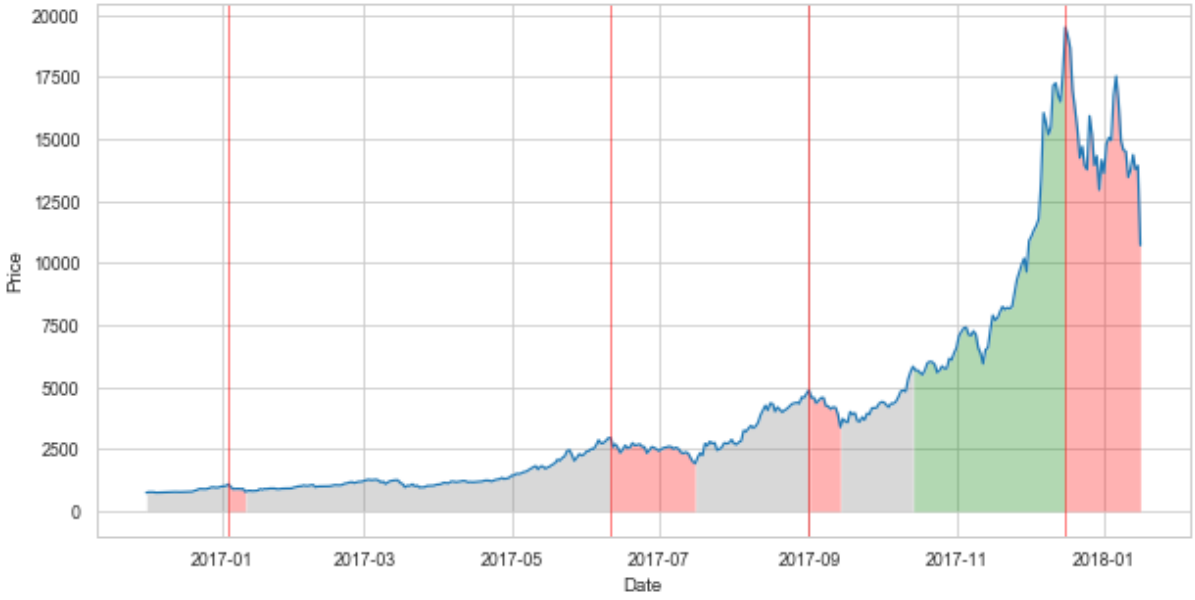

Figure 5.4: Bitcoin's phases between December 1^{st} *2016 and January* 16^{th} *2018.*

Socioeconomic considerations

In this final section, we will briefly report some considerations on the socioeconomic drivers behind Bitcoin's bubble (Demos, Sornette et al., 2018).

The major factor behind the huge increase in Bitcoin's price across 2017 can be recalled to the rising demand from Chinese markets. The main reason behind that was the devaluation of Chinese Yuan, promoted by the People's Bank of China (PBoC) since 2014, in order to raise the competitiveness of exporting firms.
As a consequence to the depreciation of their currency, Chinese investors started buying Bitcoins, which was seen as a safer store of value. Undoubtedly, limitations on

foreign-exchange investments too considerably contributed to this phenomenon.
When Chinese government started observing an outflow of capital from China, the Central Bank of China required Chinese Bitcoin exchanges, which until then were completely unregulated, to comply with the country's financial regulations, as it suspected illegal activities, like money laundering. Then, in February 2017 the PBoC also forbade Bitcoin withdrawals into currencies other than Chinese Yuan.
Nonetheless, when this restrictive measure was finally interrupted in June 2017, there was a really positive reaction which caused a further increase in Bitcoin's price.
In a completely unexpected way, in September 2017 Chinese government decided to stop all trading activities related to the digital coin, reducing sharply trading volumes. Although Bitcoin trading was considered officially dead, investors actually shifted towards unofficial OTC exchanges and other foreign exchanges, which had a meaningful growth during the previous months. In other words, activities on Chinese exchanges were not suspended, but only redirected to other markets.

At this point, Bitcoin had already gained global attention, but, unlike in the past, this brought interest on other cryptocurrencies too. During this period, the cryptocurrency market changed considerably, as shown in Figure 5.5, with Bitcoin reducing its market share from 90% at the beginning of 2017 to roughly 50% in June 2017, despite its huge increase in price.
In the last quarter of 2017, the market capitalization of the whole crypto-market grew more than 400%, sustained by the constant inflow of fresh money. However, with the crash in the second half of December 2017, the capitalization of Bitcoin and other cryptocurrencies dropped sharply, putting Bitcoin's market share to an all time low.

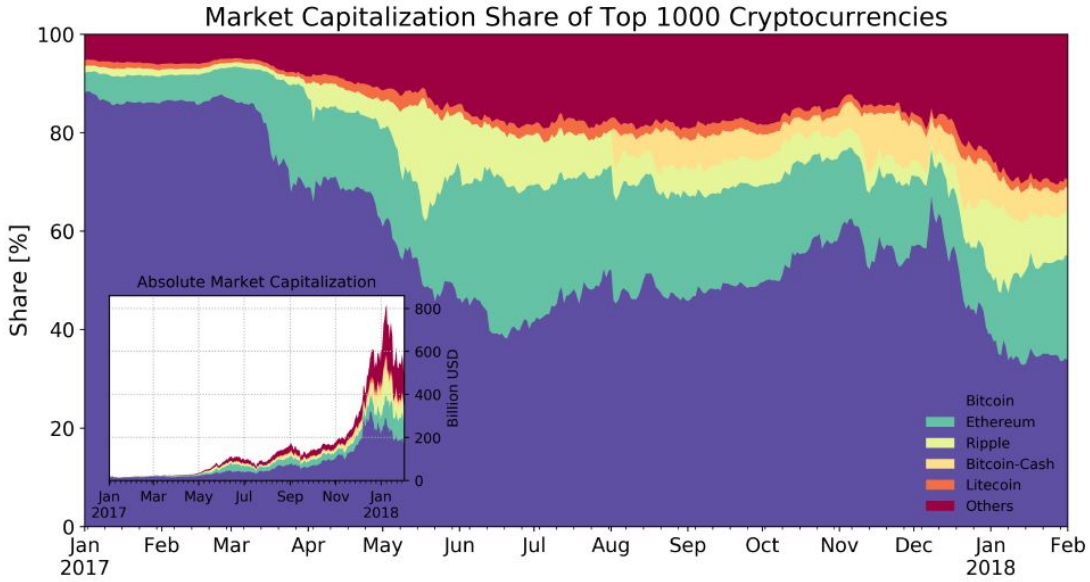

Figure 5.5: Progressive maturation of the Cryptocurrency market during 2017. (Demos, Sornette et al., 2018)

Even though 2017 was a positive year for cryptocurrencies, the launch of new coins and technologies was also perceived by Bitcoin's investors as a threat, spreading the idea that Bitcoin could become obsolete. This is surely a key factor in the violent correction that occurred at the end of 2017.

5.2. Real-time bubble identification

In this section, we will deal with the non trivial task of identifying real-time bubbles, i.e. during their development, in order to forecast the subsequent crash. This can be a very useful practice, not only for risk management, but also for valuation and trading purposes.
We will consider again Bitcoin historical time series and, in particular, we will focus on the bubble bursting in the second half of December 2017.

It is important to underline that this is not an easy task. For this reason, we adopt a bubble detection methodology supported by different methods, repeating iterations on many different time windows of varying length.
In order to detect real-time bubbles, we apply the following approach:

1. Fix a present time of analysis, t^*, corresponding to the last available date or, in case of real-time analysis on historical time series, to three days before the bubble peak.

2. Apply the bubble detection OLS version on N windows of various length, keeping fixed the right extremum $t_2 = t^*$ and moving the left one, t_1.

3. Shift backwards t_2, in such a way that $t_2 = t^* - 1$, and repeat point 2).

4. Iterate the process until t_2 has been shifted M times.

5. Compute the fraction of valid bubble signals on the $N \cdot M$ windows inspected and plot the predicted t_c values.

6. Apply also MLE and GLS as supporting methods.

5.2.1. Bitcoin's bubble identification in December 2017

We consider more in detail the Bitcoin's bubble that bursted in the second half of December 2017, applying the real-time bubble detection method presented above.

Specifically, we set t^* on December 13^{th} 2017, three days before the peak reached on December 16^{th}. We span $N = 20$ windows, from a maximum length of 72 days to a minimum of 53 days (Bianchetti et al., 2018). Moreover, we repeat it for $M = 5$ times, i.e. moving backwards t_2 until December 9^{th} 2017.

We get a fraction of valid bubble signals equal to 0.67, meaning that 67 over 100 detected time windows indicate the presence of a bubble behaviour. In particular, the calibrated values of t_c, corresponding to the predicted burst of the bubble are represented in Figure 5.6.
It is observed a cluster on December 17^{th}, correctly predicting the imminent crash.

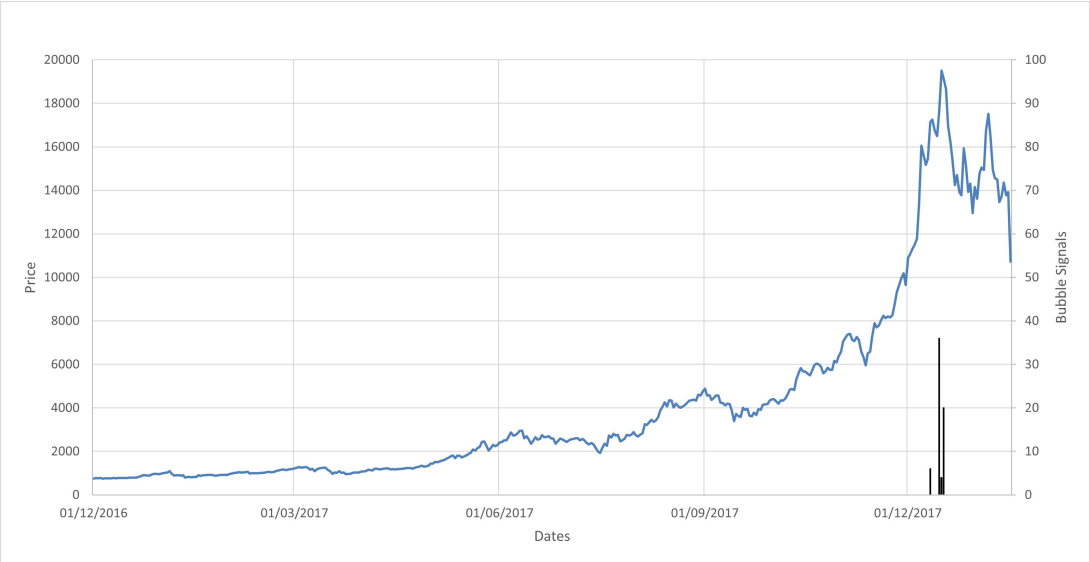

Figure 5.6: Bubble signals detected on December 13^{th} *2017 via OLS version.*

Nonetheless, most of the times signals obtained are not so easy to interpret, for example we could obtain two distinct clusters which seems to forecast two different scenarios.
If we apply the procedure above on December 8^{th}, see Figure 5.7, we obtain a fraction equal to 0.78 with valid bubble signals spanning from December 9^{th} to December 15^{th}.
Similar situations require a more in-depth study, hence we rely both on the Generalized Least Squares and Maximum Likelihood Estimation approaches, presented in Chapter 4.
We recall that the main advantage of the GLS method is to model residuals with more reliable dynamics, while the MLE approach is able to provide interval estimates of the critical time, t_c, where the burst of the bubble is more probable to happen.

Applying the MLE approach together with the GLS version on December 8^{th}, we can compare the shape of the Relative Modified Profile Likelihood, $R_m(t_c)$, with the signals obtained via GLS and OLS version, as shown in Figure 5.8.
One can observe how the MLE approach identifies the most probable critical time, t_c, on December 15^{th}, while bubble signals on December 9^{th} are considered less reliable.

Regarding the GLS version, we have applied the calibration method on the same time windows used in the OLS case, obtaining results in line with MLE forecast, as most of

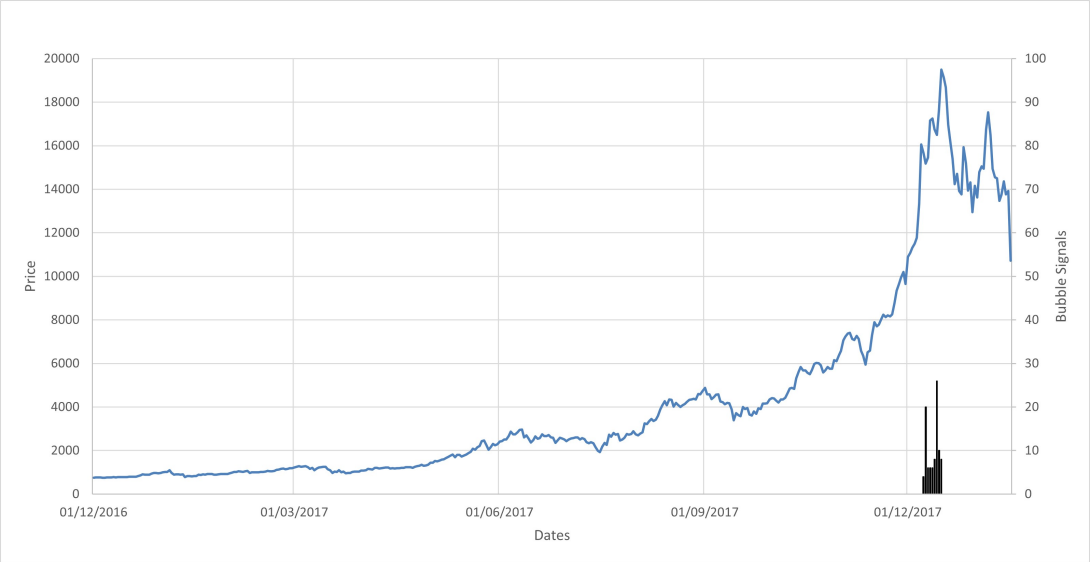

Figure 5.7: Bubble signals detected on December 8th 2017 via OLS version.

the signals appear to be clusterized around December 15th.

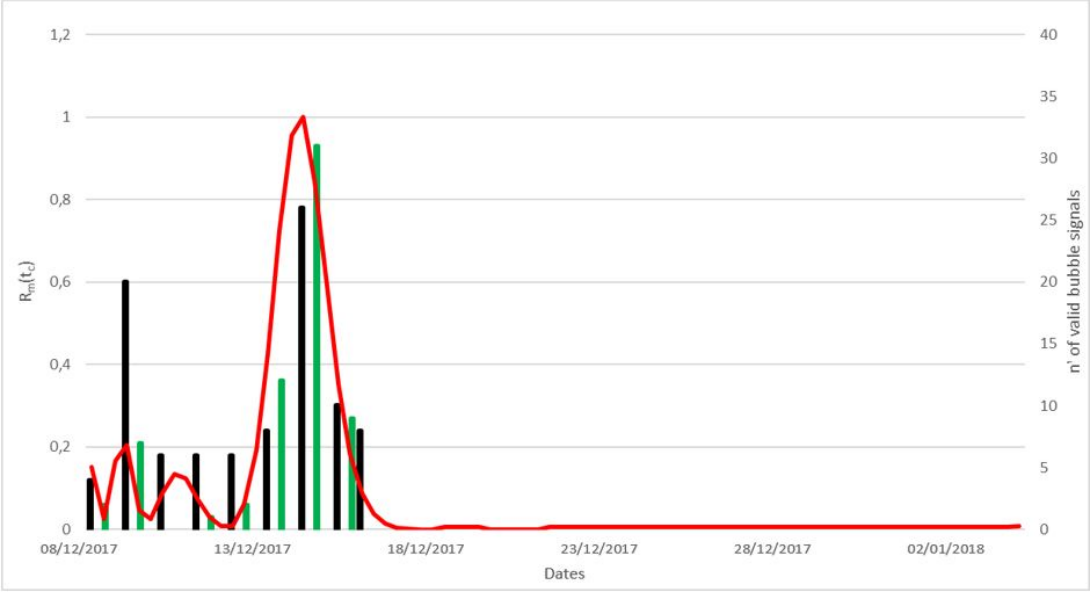

Figure 5.8: Comparison among OLS version (black bars), MLE approach (red line) and GLS version (green bars) on December 8th 2017.

5.2.2. Bitcoin's bubble identification in 2021

After the peak reached in December 2017, Bitcoin's prices remained steadily below 13'000 USD for the following two years. In addition, in line with the drawdowns registered on the main Equity and Credit indices, in March 2020 Bitcoin lost almost half of its market capitalization, as coronavirus was perceived as a global threat and governments were taking restrictive measures to slow down the pandemic.
After that, Bitcoin saw a huge increase in price, from 6'000 USD to 60'000 USD in approximately one year, as shown in Figure 5.9.

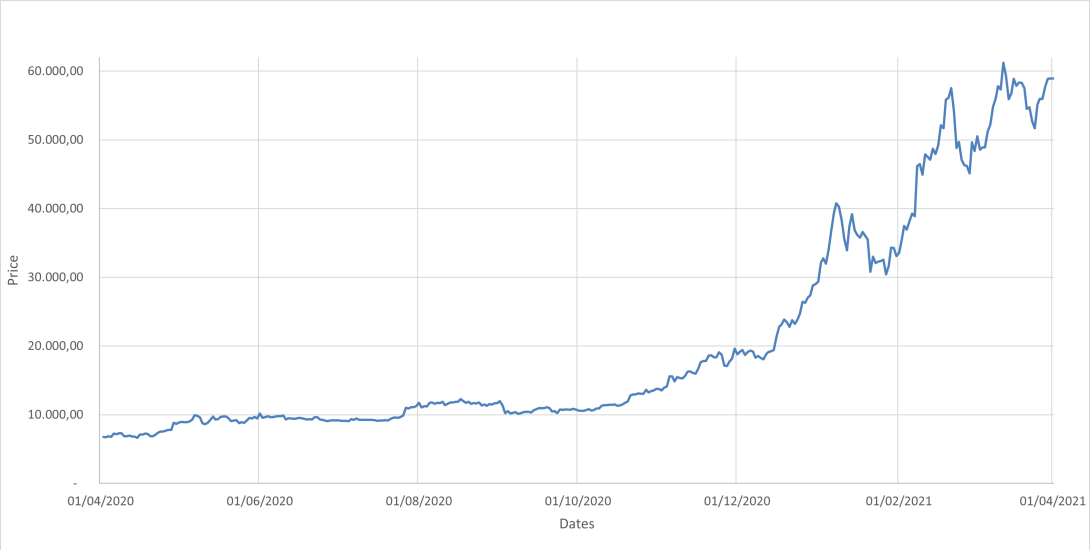

Figure 5.9: Bitcoin's historical prices from April 1^{st} *2020 to April* 1^{st} *2021.*

After this sharp growth, one might ask whether Bitcoin has entered an irrational behaviour or not. To try to give an answer to this question, we apply the same bubble detection method as above to Bitcoin's prices. Results obtained show bubbles signals in different points of the time series.

In more detail, we apply the real-time bubble detection method across $M = 5$ consecutive days, spanning $N = 20$ windows (from a maximum length of 72 days to a minimum of 53 days), from the beginning of December 2020 till the end of March 2021. We located this time period as Bitcoin approached and overcome its all-time high, previously reached in December 2017.

Bubble signals are detected in the first half of January 2021 and in the second half of February 2021, as reported in Figure 5.10. In particular, the first cluster correctly forecasts the correction registered in January 8^{th} 2021 and the second one anticipates the drawdown registered in February 21^{th} 2021. The table below summarizes the relevant data:

Peak date	Drawdown	Duration	Fraction	Identification date
08/01/2021	−25.41%	20 days	0.49	05/01/2021
21/02/2021	−21.55%	8 days	0.24	20/02/2021

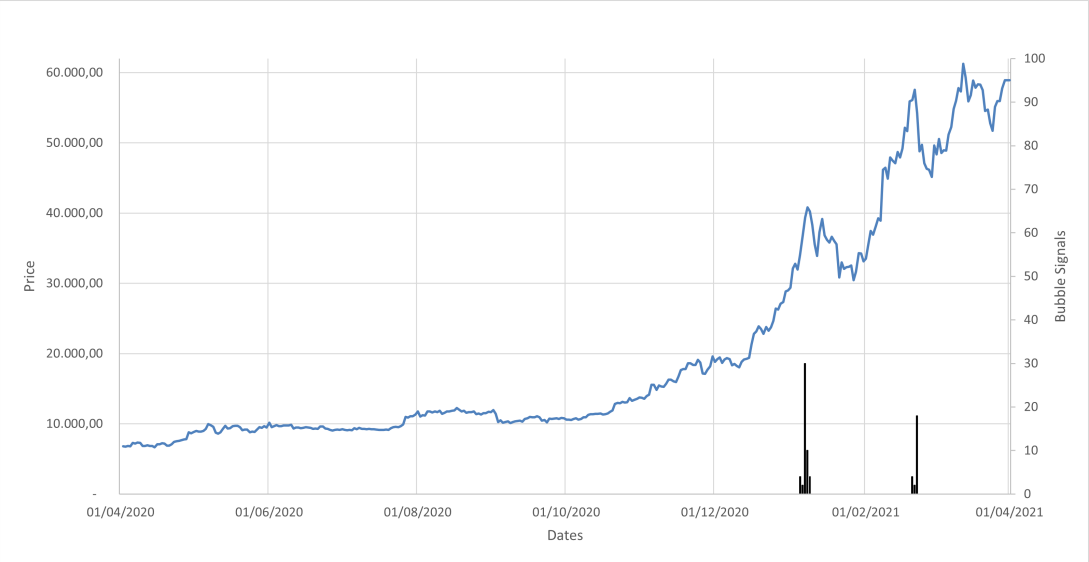

Figure 5.10: Bubble signals detected via OLS version.

Specifically, the first column, *Peak date*, refers to the date when the peak was reached, the second one, *Drawdown*, reports the magnitude of the subsequent correction phase and the third one, *Duration*, its length.
Eventually, *Fraction* indicates the fraction of valid bubble signals and *Identification date* the date on which the bubble detection method was applied.

Considering more closely each bubble cluster and comparing the forecasts obtained in the OLS case with MLE and GLS variants, see Figure 5.11 and Figure 5.12, we can draw similar conclusions.
The first bubble cluster is concentrated between January 7^{th} and January 10^{th}, while the second one is between February 20^{th} and February 23^{th}.

Despite these two sharp corrections, however Bitcoin's price had a quick rebound and now, on March 31^{th} 2021, it is approaching again 60'000 USD. In our view, this is a consequence of the positive news heard on the recognition and adoption of Bitcoin and other cryptocurrencies by the financial community. Nonetheless, it remains a volatile asset and we advise the reader to pay close attention.

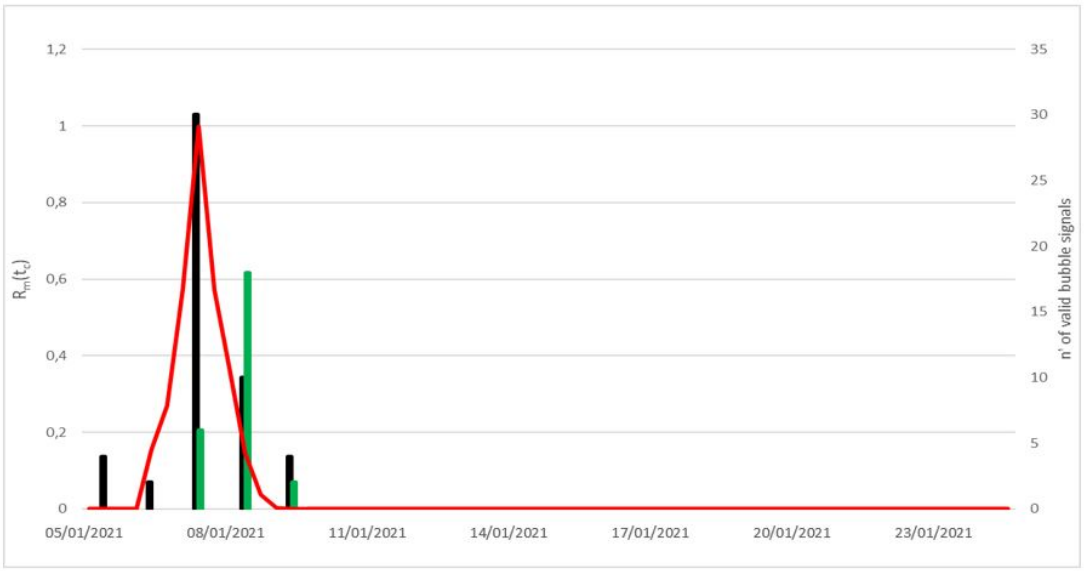

Figure 5.11: Comparison among OLS version (black bars), MLE approach (red line) and GLS version (green bars) on January 8^{th} *2021.*

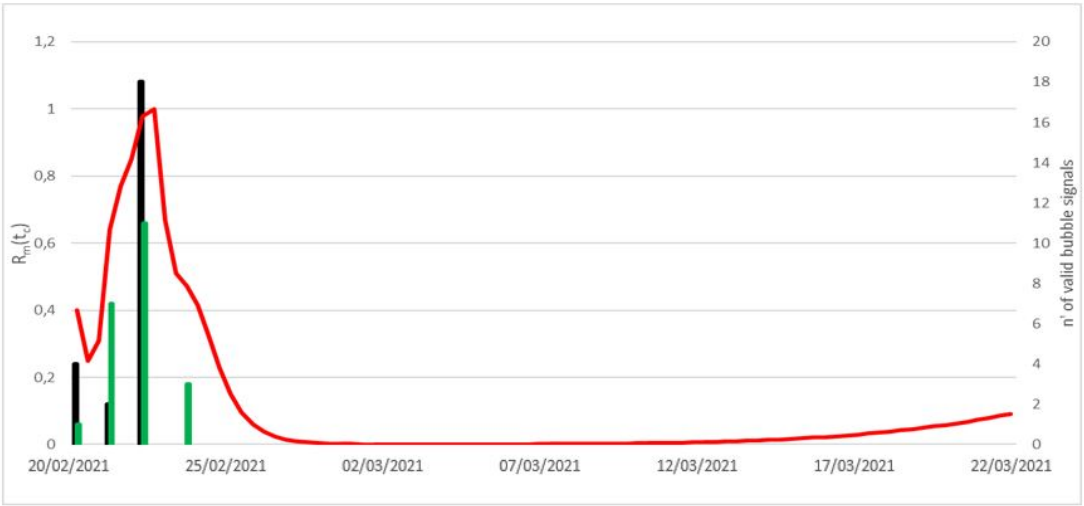

Figure 5.12: Comparison among OLS version (black bars), MLE approach (red line) and GLS version (green bars) on February 21^{th} *2021.*

6. Conclusion

In the current work, we have presented the Log Periodic Power Law (LPPL) model, as a valid tool for the detection and forecast of financial bubbles.
We briefly recall that, according to LPPL model, an asset is governed by a bubble behaviour if its price is characterized by the LPPL function. The seven parameters of the model, $\{A, B, C, t_c, \beta, \omega, \phi\}$, must be calibrated in order to optimally fit the asset prices on a fixed time window. Moreover, if the estimated parameters satisfy specific constraints, the so called Sornette bounds, one can state that the asset is under a bubble regime and the critical time, t_c, represents the candidate crash time.

From a theoretical point of view, we have focused on reviewing all the past literature, with particular attention to the latest publications. The main assumption to the theoretical framework is that the network of irrational agents is modelled as a Hierarchical Diamond lattice, with the idea to mimic the heterogeneity of real financial networks. This choice has revealed to better fit with respect to the Bidimensional Ising model. In addition, we recall that the susceptibility function, χ, which regulates the system response to perturbations, diverges in correspondence of the critical point of the system; from a financial point of view, this means that an external source might act as trigger of a massive selloff from a consistent group of investors, causing a violent fall in the price of the asset.

The calibration of the parameters of the model has been a non trivial task, given the high non linearity of the LPPL function and the presence of many local minima. To solve this issue, we have proposed the original calibration procedure, together with two more recent and reliable variants.
Indeed, while the original calibration procedure relies on the fundamental assumptions that residuals are homoskedastic and uncorrelated, the Generalized Least Squares (GLS) variant opens up to correlated residuals following an autoregressive process. This has proven to provide more accurate forecasts, despite a higher computational effort. In our view, a further development could be carried on by removing the homoskedasticity assumption, by adopting a GARCH process for the residuals.
On the other hand, the Maximum Likelihood Estimation (MLE) method has the great advantage to provide confidence intervals, and not only punctual results, for the estimation of the parameters, providing a measure of the uncertainty related to the estimated critical time.

Calibrating methods have been applied both on post-mortem and real-time analyses. In particular, for post-mortem studies, where crashes in the time series are analyzed

after they already occurred, we have proposed a newly approach, able to distinguish between drawdown and drawup periods, identify periods of bubble activity and also provide an estimate for the beginning of the bubble behaviour. In order to do that, we relied on the Epsilon Drawdown/Drawup method (Demos, Sornette et al., 2018) and the Lagrange Regularisation approach (Demos, Sornette, 2017).
For real-time analyses, where calibration has bubble forecasting purposes, we have implemented a coherent bubble detection method, which combines the three calibrating techniques.
All the methods have been implemented in Python. To this extent, we suggest the adoption of parallel computing algorithms to ease the computational effort required.

Practical applications have been carried on Bitcoin's historical time series.
Specifically, a post-mortem analysis on Bitcoin's prices between December 2016 and January 2018 identifies bubble signals in the second half of December 2017, in line with previous publications (Bianchetti et al., 2018). In addition, some socio-economic considerations have been reported to give the reader a wider view on the drivers of the bubble behaviour and its subsequent crash.
At the same time, we have carried on a real-time analysis on the bubble phase registered at the end of 2017, correctly forecasting the crash that almost halved Bitcoin's market capitalization.
Eventually, we have made a real-time analysis on actual Bitcoin's prices too, to establish whether the sharp increase in price observed throughout the last year was influenced by a bubble behaviour or not. We found evidence of bubble signals in correspondence of the two peaks reached on January 8^{th} 2021 and February 21^{th} 2021, correctly anticipating the subsequent sharp corrections. Despite its strong rebound in price, we conclude that Bitcoin is under extremely volatile conditions and we advise the reader to pay close attention.

Bibliography

[1] Adhikari R., Agrawal R. K. (2013) An Introductory Study on Time Series Modeling and Forecasting. *Lambert Academic Publishing.*

[2] Ametrano F. (2016) Hayek Money: the Cryptocurrency Price Stability Solution. *SSRN Electronic Journal.*

[3] Antonopoulos A. (2017) Mastering Bitcoin. *O'Reilly (2ND Edition, pag. 1-30).*

[4] Bianchetti M., Ricci C., Scaringi M. (2018) Are cryptocurrencies real financial bubbles? Evidence from quantitative analysis. *SSRN Electronic Journal.*

[5] Bianchetti M., Galli D., Ricci C., Salvatori A., Scaringi M. (2016) Brexit or Bremain? Evidence from bubble analysis. *SSRN Electronic Journal.*

[6] Bjork T., (2009) Arbitrage Theory in Continuous Time. *Oxford Finance (3RD Edition, pag. 92-157).*

[7] Bingcun D., Kwangwon A., Tarzia D., Zhang F. (2018) Forecasting Financial Crashes: Revisit to Log-Periodic Power Law. *Wiley Hindawi (Vol. 2018, pag. 1-12).*

[8] Cameron A., Trivedi P. (2005) Microeconometrics, Methods and Applications. *Cambridge University Press (pag. 65-85).*

[9] Demirer R., Demos G., Gupta R., Sornette D. (2017) On the Predictability of Stock Market Bubbles: Evidence from LPPLS Confidence Multi-scale Indicators. *Quantitative Finance (Vol. 19, pag. 843-858).*

[10] Demos G. (2017) Information Geometry and the Dynamic Detection of Financial Bubbles and Crashes. *PhD Thesis. ETH Zurich*

[11] Demos G., Filimonov V., Sornette D. (2016) Modified Profile Likelihood Inference and Interval Forecast of the burst of financial bubbles. *Swiss Finance Institute Research Paper (Vol. 16, pag. 1167-1186).*

[12] Demos G., Sornette D. (2017) Lagrange Regularisation approach to compare nested data sets and determine objectively financial bubbles' inceptions. *Swiss Finance Institute Research Paper (Vol. 18-20).*

[13] Demos G., Gerlach J.C., Sornette D. (2018) Dissection of Bitcoin's Multiscale Bubble History from January 2012 to February 2018. *Royal Society Open Science (Vol. 06, pag. 25-54).*

[14] Fantazzini D., Geraskin P. (2011) Everything You Always Wanted to Know about Log Periodic Power Laws for Bubble Modelling but Were Afraid to Ask. *The European Journal of Finance (Vol.13, pag. 366-391).*

[15] Franchin S. (2018) Financial Bubbles Growth and Burst: from Theoretical Models to Real Case Computational Analysis. *MSc Thesis. Alma Mater Studiorum - Università di Bologna.*

[16] Greene W. (2002) Econometric Analysis. *Prentice Hall (pag. 191-280).*

[17] Hubele N., Montgomery D., Runger G. (2011) Engineering Statistics. *John Wiley & Sons (Ed. 5, pag. 306-344).*

[18] Johansen A., Ledoit O., Sornette D. (2008) Crashes as Critical Points. *International Journal of Theoretical and Applied Finance (Vol. 03, pag. 219-255).*

[19] Johansen A., Sornette D. (2001) Significance of log-periodic precursors to financial crashes. *Quantitative Finance (Vol. 01, pag. 452-471).*

[20] Jordà O., Schularick M., Taylor A., (2015) Leveraged Bubbles. *Journal of Monetary Economics (Vol. 76, pag. 1-20).*

[21] Kim H.Y., Lee W. (2005) Genetic Algorithm Implementation in Python. *Fourth Annual ACIS International Conference on Computer and Information Science.*

[22] Salvatori A. (2016) Stochastic Models for Self-Organized Criticality in Financial Markets. *MSc Thesis. Università degli Studi di Milano.*

[23] Scaringi M. (2016) Financial bubbles: genesis and detection within the JLS model framework. *MSc Thesis. Università degli Studi di Milano.*

[24] Sornette D. (2003) Why stock markets crash: Critical events in complex financial systems. *Princeton University Press, New Jersey (pag. 448).*

[25] Sornette D. (2009) Dragon-Kings, Black Swans and the Prediction of Crises. *Swiss Finance Institute Research Paper (Vol. 09-36).*

A. Appendix

We present the list of partial derivatives required in formula (4.34) for the calculation of the Modified Profile Likelihood (Filomonov et al., 2016)

$$\frac{dLPPLS}{d\beta} = |t_c - t|^{\beta} \ln|t_c - t| \left[B + C_1 \cos(\omega \ln|t_c - t|) + C_2 \sin(\omega \ln|t_c - t|)\right],$$
$$\frac{dLPPLS}{d\omega} = |t_c - t|^{\beta} \ln|t_c - t| \left[- C_1 \sin(\omega \ln|t_c - t|) + C_2 \cos(\omega \ln|t_c - t|)\right],$$
$$\frac{dLPPLS}{dA} = 1,$$
$$\frac{dLPPLS}{dB} = |t_c - t|^{\beta},$$
$$\frac{dLPPLS}{dC_1} = |t_c - t|^{\beta} \cos(\omega \ln|t_c - t|),$$
$$\frac{dLPPLS}{dC_2} = |t_c - t|^{\beta} \sin(\omega \ln|t_c - t|).$$

The second order derivatives required in (4.35) are:

$$\frac{d^2LPPLS}{d\beta^2} = |t_c - t|^{\beta} (\ln|t_c - t|)^2 \left[B + C_1 \cos(\omega \ln|t_c - t|) + C_2 \sin(\omega \ln|t_c - t|)\right],$$
$$\frac{d^2LPPLS}{d\omega\, d\beta} = |t_c - t|^{\beta} (\ln|t_c - t|)^2 \left[- C_1 \sin(\omega \ln|t_c - t|) + C_2 \cos(\omega \ln|t_c - t|)\right],$$
$$\frac{d^2LPPLS}{d\beta\, dB} = |t_c - t|^{\beta} \ln|t_c - t|,$$
$$\frac{d^2LPPLS}{d\beta\, dC_1} = |t_c - t|^{\beta} \ln|t_c - t| \cos(\omega \ln|t_c - t|),$$
$$\frac{d^2LPPLS}{d\beta\, dC_2} = |t_c - t|^{\beta} \ln|t_c - t| \sin(\omega \ln|t_c - t|),$$
$$\frac{d^2LPPLS}{d\omega^2} = -|t_c - t|^{\beta} (\ln|t_c - t|)^2 \left[C_1 \sin(\omega \ln|t_c - t|) + C_2 \cos(\omega \ln|t_c - t|)\right],$$
$$\frac{d^2LPPLS}{d\omega\, dC_1} = -|t_c - t|^{\beta} \ln|t_c - t| \sin(\omega \ln|t_c - t|),$$
$$\frac{d^2LPPLS}{d\omega\, dC_2} = |t_c - t|^{\beta} \ln|t_c - t| \cos(\omega \ln|t_c - t|).$$

All other second order derivatives are equal to zero.
Moreover, we recall that, thanks to Schwarz Theorem, $\frac{d^2LPPLS}{d\psi_i\, d\psi_j} = \frac{d^2LPPLS}{d\psi_j\, d\psi_i} \qquad \forall \psi_i, \psi_j.$