Early warning signals of financial crises
with multi-scale quantile regressions of Log-Periodic Power Law Singularities

Qun Zhang\textsuperscript{a,b}, Qunzhi Zhang\textsuperscript{b}, Didier Sornette\textsuperscript{b}

\textsuperscript{a}School of Business Administration, South China University of Technology, China
\textsuperscript{b}Department of Management, Technology and Economics, ETH Zurich, Switzerland

Abstract

We augment the existing literature using the Log-Periodic Power Law Singular (LPPLS) structures in the log-price dynamics to diagnose financial bubbles by providing three main innovations. First, we introduce the quantile regression to the LPPLS detection problem. This allows us to disentangle (at least partially) the genuine LPPLS signal and the a priori unknown complicated residuals. Second, we propose to combine the many quantile regressions with a multi-scale analysis, which aggregates and consolidates the obtained ensembles of scenarios. Third, we define and implement the so-called DS LPPLS Confidence\textsuperscript{TM} and Trust\textsuperscript{TM} indicators that enrich considerably the diagnostic of bubbles. Using extensive synthetic signals, a detailed analysis of the “S&P 500 1987” bubble and the application to 16 historical bubbles, we show that the quantile regression of LPPLS signals contributes useful early warning signals. The comparison between the constructed signals and the price development in these 16 historical bubbles demonstrates their significant predictive ability around the real critical time when the burst/rally occurs.

Keywords: Financial bubble, Log-periodic power law singularity (LPPLS), Quantile regression, Early warning signals, Time scale, Probabilistic forecast

JEL: C14, C21, C53, G01, G17

1. Introduction

The present article aims at extending the approach pioneered in Sornette et al. (1996), Johansen et al. (1999), Johansen et al. (2000), Sornette (2009), Johansen and Sornette (2010), and Sornette and Cauwels (2015a) to develop testable diagnostics of financial bubbles. Useful summaries of our starting point can be found in Sornette and Woodard (2010), and Sornette and Cauwels (2014, 2015b) as well in the resources and contributions of the Financial Crisis Observatory at \url{http://www.er.ethz.ch/financial-crisis-observatory.html}.

In a nutshell, we follow the above cited authors in taking the existence of a transient faster-than-exponential price growth as a signature of bubbles (Hüsler, Sornette & Hommes, 2013; Leiss, Nax & Sornette, 2015; Sornette & Cauwels, 2015a). The advantage of this definition of a bubble is that it does not rely on the estimation of what is a fundamental value (see e.g., Siegel, 2003), which is at best defined within a factor of two (Black, 1986). Normal times are characterised by an approximate constant return (or price growth rate). This is nothing but the statement that the average price trajectory is a noisy exponential that reflects the power of compounding interests. As the simplest embodiment of this noisy exponential growth, the Geometrical Brownian Motion model is the starting point of more sophisticated models in financial mathematics and financial engineering. Based on the analyses of many historical bubbles, Sornette et al. (1996), Johansen et al. (1999, 2000), and Leiss et al. (2015) have proposed that there are transient regimes during which the price growth rate (return) grows itself, which translates into a super-exponential time dynamics. The underlying mechanisms involving positive feedbacks, also called procyclical processes, tend to increase and accelerate the deviation from an equilibrium. The resulting super-exponential price trajectories are inherently unsustainable and often burst as crashes or strong corrections.

The Log-Periodic Power Law Singularity (LPPLS) model has been proposed as a simple generic parameter-
isation to capture such super-exponential behavior (Sornette, Johansen & Bouchaud, 1996; Johansen, Sornette & Ledoit, 1999; Johansen, Ledoit & Sornette, 2000; Sornette, 2009), which is inspired from physics (and is sometimes referred to as part of econophysics (Abergel, Aoyama, Chakrabarti, Chakraborti & Ghosh, 2015)). The main point is that positive feedbacks generically lead to finite-time singularities (Ide & Sornette, 2002; Sornette, 2002; Sornette & Cauwels, 2015b). The positive feedbacks can be of many types, such as option hedging, portfolio insurance strategies, margin requirements, as well as the imitation and herding behavior in psychology. Moreover, once one takes into account either the existence of a discrete hierarchy of group sizes (Zhou, Sornette, Hill & Dunbar, 2005) or the interplay between nonlinear value investors and nonlinear trend followers, and the inertia between information flow and price discovery (Ide & Sornette, 2002), the finite-time singular behavior is decorated by accelerating oscillations, whose simplest generic form is log-periodic, which is a hallmark of the symmetry of discrete scale invariance (Sornette, 1998). In summary, these considerations explain our focus on the detection of financial bubbles modelled as transient LPPLS price trajectories.

We thus follow up on these previous efforts to diagnose financial bubbles and their terminations by proposing several innovations. First, we apply the quantile regression to the LPPLS calibration problem. Rather than fitting a given log-price time series by a single LPPLS model, quantile regressions provide a family of calibrated curves indexed by the probability level $q$. Scanning $q$ between 0 and 1 allows us to disentangle (at least partially) the genuine LPPLS signal from the a priori unknown complicated residuals. The standard least squares or maximum likelihood estimation procedures are vulnerable to the existence of outliers (Barroda & Roberts, 1974). In contrast, quantile regressions alleviate some of the statistical problems that have plagued the literature: error in variables, sensitivity to outlier and non-normal error distributions (Barnes & Hughes, 2002). It provides a descriptive approach reporting more than just the expected mean of a conditional distribution, but may also discover more complete structures without imposing global distributional assumptions on the residuals. The prediction inference associated with quantile-based estimates has an inherent distribution-free character since they are influenced only by the local behavior of the underlying distribution near the specified quantile (Koenker, 2005). The different $q$-dependent LPPLS fits provide a bundle of possible scenarios that are compatible with different weights of the residuals supposed to decorate the theoretical driver.

Then, we propose to combine the many quantile regressions with a multi-scale analysis. This leads to the development of ensemble forecasting that combines a grid of quantile-based estimators into a final aggregated predictor, improving the robustness of a single estimator (Aráujo & New, 2007). We introduce the Quantile-Violin plots and the $dt$-Violin plots as powerful representations of the enormous amount of information generated by scanning the quantile levels and the time scales.

Finally, we define and implement the so-called DS LPPLS Confidence™ and Trust™ indicators, which provide an aggregation and consolidation of the wealth of generated information and we put them at work to diagnose 16 historical bubble cases.

These innovations have the ultimate goal of being inserted into an early warning system that could be run by a central bank, say, to inform it towards appropriate counter measures (Sornette, 2002; Scheffer et al., 2009; Roubini & Mihm, 2010; Lenton et al., 2012; Scheffer et al., 2012). This last step of constructing an explicit early warning system is not investigated here, as we focus on the introduction and tests of our new metrics and methodology for a better characterisation of the LPPLS signals.

We proceed as follows. Section 2 presents the LPPLS model and gives an overview on some theoretical aspects of the standard ordinary least square regression (referred to as the $L^2$ norm calibration) and of quantile regressions. Section 3 presents the metrics, the methodology and a battery of tests performed on synthetic time series. In particular, we introduce the Quantile-Violin plots and the $dt$-Violin plots as efficient presentations of the multi dimensional metrics. Section 4 applies the tools presented in section 3 to the S&P 500 bubble that burst in October 1987. Section 5 extends section 4 to three other historical financial bubbles, providing interesting elements of comparison. Section 6 introduces the DS LPPLS Confidence™ and Trust™ indicators. Section 7 applies all the above tools and metrics to 16 historical financial bubbles and compare the indicators with the price time series. Section 8 summarises our main conclusions.

2. Model and calibrations

2.1. Log-Periodic Power Law Singularity (LPPLS) model

The Johansen-Ledoit-Sornette (JLS) model (Johansen, Sornette & Ledoit, 1999; Johansen, Ledoit & Sornette, 2000) assumes that the asset price $p(t)$ follows
a standard diffusive dynamics with varying drift $\mu(t)$ in the presence of discrete discontinuous jumps,

$$ \frac{dp}{p} = \mu(t)dt + \sigma(t)dW - \kappa dt, $$

(1)

where $\sigma(t)$ is the volatility and $dW$ is the increment of a Wiener process (with zero mean and variance equal to $dt$). The term $dJ$ represents a discontinuous jump such that $J = 0$ before the crash and $J = 1$ after the crash occurs. The loss amplitude associated with the occurrence of a crash is determined by the parameter $\kappa$. Each successive crash corresponds to a jump of $J$ by one unit. The dynamics of the jumps is governed by a crash hazard rate $h(t)$. Since $h(t)dt$ is the probability that the crash occurs between $t$ and $t + dt$, conditional on the fact that it has not yet happened, we therefore have the expectation $E_t[dJ] = 1 \times h(t)dt + 0 \times (1 - h(t)dt) = h(t)dt$. By the no-arbitrage condition leading to the condition that the price process is a martingale ($E_t[\frac{dp}{p}] = 0$, neglecting the risk free rate), it leads to $\mu(t) = kh(t)$.

Under the assumption of the JLS model, the crash hazard rate aggregated by the noise traders with herding behaviors has the following dynamics:

$$ h(t) \approx B_0[|t_c - t|^m + B_1|t_c - t|^m \cos(\omega \ln|t_c - t| + \phi)]. $$

(2)

Using $\mu(t) = kh(t)$, we obtain the dynamics of the expectation of the logarithm of the price in the form of the Log-Periodic Power Law Singularity (LPPLS) model:

$$ E[\ln p(t)] = A + B|t_c - t|^m + C|t_c - t|^m \cos(\omega \ln|t_c - t| + \phi), $$

(3)

where $t_c$ denotes the most probable time for the burst of the bubble, in the form of a crash for example. The constant $A = \ln[p(t_c)]$ gives the terminal log-price at the critical time $t_c$. $B$ and $C$ respectively control the amplitude of the power law acceleration and of the log-periodic oscillations. The exponent $m$ quantifies the degree of super-exponential growth. The log-periodic angular frequency $\omega$ is related to a scaling ratio $\lambda = \exp\left(\frac{2\pi}{\omega}\right)$ of the temporal hierarchy of accelerating oscillations converging to $t_c$. Finally, $\phi \in (0, 2\pi)$ is a phase embodying a characteristic time scale of the oscillations. Expression (3) is the first-order log-periodic correction to a pure power law for an observable exhibiting a singularity at $t_c$ (Gluzman & Sornette, 2002; Sornette, 2009).

Previous calibrations of the LPPLS specification (3) to the log-price development during a number of historical financial bubbles have suggested to qualify fits based on the parameters of the LPPLS model belonging to the following intervals (Johansen & Sornette, 2010; Jiang, Zhou, Sornette, Woodard, Bastaiaens & Cauwels, 2010; Filimonov & Sornette, 2013): $m \in [0.1, 0.9]$, $\omega \in [6, 13]$, $|C| \leq 1$, $B < 0$. In our explorations, we have found that relaxing the search space to the larger intervals $m \in [0, 2]$ and $\omega \in [1, 50]$ does not change the results significantly, particularly for the calibrated critical times within statistical fluctuations. The results we report below have thus been obtained for the larger search ranges $m \in [0, 2]$, $\omega \in [1, 50]$. Given the starting and ending dates $t_{\text{start}}$ and $t_{\text{end}}$ of the fitting window, we define $dt := t_{\text{end}} - t_{\text{start}}$ as the duration of the fitting window. The critical time $t_c$ is searched in the interval $[t_{\text{end}} - \eta dt, t_{\text{end}} + \eta dt]$, with $\eta$ typically equal to 0.20.

### 2.2. The optimization problem using the standard Ordinary Least Squares (OLS) method

Filimonov and Sornette (2013) suggested to expand the cosine term of Eq. (3) with $C_1 = C \cos \phi$, $C_2 = -C \sin \phi$ to obtain a representation with 4 linear and 3 nonlinear parameters, providing a substantial gain in efficiency and stability of the calibration. This leads to rewrite Eq. (3) as:

$$ \ln E[\ln p(t)] = A + B|t_c - t|^m + C_1|t_c - t|^m \cos(\omega \ln|t_c - t|) + C_2|t_c - t|^m \sin(\omega \ln|t_c - t|). $$

(4)

The optimization problem with the standard Ordinary Least Squares (OLS) method aims to minimize the sum $F(t_c, m, \omega, A, B, C_1, C_2)$ of squared residuals between the log-price $\ln p(t_i), i = 1, 2, ..., N$ and the function (4), where

$$ F(t_c, m, \omega, A, B, C_1, C_2) = \sum_{i=1}^{N} \left( \ln p(t_i) - A - B|t_i - t_c|^m - C_1|t_i - t_c|^m \cos(\omega \ln|t_i - t_c|) - C_2|t_i - t_c|^m \sin(\omega \ln|t_i - t_c|) \right)^2. $$

(5)

### 2.3. The optimization problem using the Quantile Regression calibration method

Intuitively, the OLS calibration method is finding the best fit “in mean”. In other words, the parameters are adjusted so that the function to calibrate is the closest to the mean of the noisy realisation of the log-price, where the mean should be considered conceptually to occur over many realisations of the noise decorating the supposed theoretical function (4). If the noise is not normally distributed and exhibits heavier tails, the OLS calibration may be contaminated by large deviations of the
noise from the mean. Then, fitting the data to the function that is the closest to the median of the noisy realisation of the log-price may be more adequate and lead to more stable estimations. It is well known that this amounts to replacing the $L^2$ norm (sum of the square of the differences) in Eq. (5) by the $L^1$ norm (sum of the absolute value of the differences). Quantile regressions amount to generalizing the minimisation of the $L^1$ norm and provide not just a single best fit to the median but a bundle of best fits to the different quantile realisations of the noise around the theoretical LPPLS function (4).

First, let us recall that the $q$th quantile of a random variable $Y$ with distribution function $F_Y(y) = P(Y \leq y)$ is defined as:

$$Q_q(y) = \inf \{y | F_Y(y) \geq q, \ q \in (0, 1) \}.$$  

Let us define the $q$-dependent norm

$$\rho_q(e_t) = \begin{cases} 
-(1-q)e_t & \text{if } e_t < 0, \\
q e_t & \text{if } e_t \geq 0.
\end{cases}$$

For $q = 1/2$, $\rho_{1/2}(e_t) = \frac{1}{2}|e_t|$, so minimizing $\rho_{1/2}(e_t)$ is nothing but minimising the $L^1$ norm.

Quantile regression corresponds to finding the quantile-dependent parameters $\{ \hat{t}(q), \hat{m}(q), \hat{\omega}(q), \hat{A}(q), \hat{B}(q), \hat{C}_1(q), \hat{C}_2(q) \}$ that minimise the function

$$S_q(t, m, \omega, A, B, C_1, C_2) = \sum_{i=1}^N \rho_q[\ln p(t_i) - \ln E[p(t_i)]] ,$$

and $\ln E[p(t_i)]$ is given by function (4). In other words, for each quantile level $q$, we obtain a set of $q$-dependent calibrated parameters

$$\{ \hat{t}(q), \hat{m}(q), \hat{\omega}(q), \hat{A}(q), \hat{B}(q), \hat{C}_1(q), \hat{C}_2(q) \} = \arg\min_{t, m, \omega, A, B, C_1, C_2} S_q(t, m, \omega, A, B, C_1, C_2) .$$

From the definition (7), one can see that the quantile regression is an asymmetrically weighted $L^1$-based regression, where the asymmetry is governed by the value $q$. The special case $q = 1/2$ is symmetric and recovers the $L^1$ norm calibration, as already mentioned. For $q \neq 1/2$, by construction of (7), the best fit corresponds statistically to $q \cdot 100\%$ of the data points $\{ \ln p(t_i), i = 1, 2, \ldots, N \}$ to be below the theoretical curve $\ln p_q(t)$ and $(1-q) \cdot 100\%$ of the data points to be above it. Thus, for $q > 1/2$ (resp. $q < 1/2$), most of the data points are below (resp. above) the calibrated curve $\ln p_q(t)$, putting it above (resp. below) the median fit. Scanning $q$ provides a family of calibrated functions that are sensitive to different parts of the statistical fluctuations supposed to decorate the theoretical generating process (4).

3. Methodology, metrics and tests on synthetic time series

To illustrate the performance of the OLS and quantile regression methods, we test them on synthetic time series generated with parameters determined by the best fits of the S&P 500 Composite Index over the time period corresponding to the bubble that burst with the crash in October 1987, hereafter referred to as the “S&P 500 1987” bubble. The corresponding LPPLS curve with the true critical normal date $T_c = 1987.10.19$ is augmented by adding a proportional noise generated by the normal distribution $N(0, 0.027)$. Using the synthetic time series, we introduce the metrics and procedures that are later applied to empirical data.

3.1. Comparison of quantile regressions on synthetic noisy time series

Fig. 1 shows three panels for three time windows [1984.07.30, 1987.06.12] (top panel), [1984.09.21, 1987.08.06] (middle panel) and [1984.12.03, 1987.10.16] (bottom panel) representing a bundle of nine coloured quantile-based calibrated curves for $q$ increasing from 0.10 to 0.90 from bottom to top. The blue line is the in-sample synthetic price time series with a window size $dt=750$ trading days, which is extended by the following blue dashed out-of-sample data. The black dashed vertical line in each panel represents the corresponding end date $t_{end}$ of the in-sample window. The red dashed vertical line is the true critical date $T_c = 1987.10.19$ used in generating the synthetic log-price time series. The in-sample standard $L^2$-based fitted curve is also shown as the red thick curve, which is extended by the red dashed thick out-of-sample curve.

For the first and second panels (with [1984.07.30, 1987.06.12] and [1984.09.21, 1987.08.06]), the estimated critical time $\{ \hat{t}_q \} | q = 0.10, 0.20, \ldots, 0.90 \}$ and the $\hat{t}_q$ determined by the standard $L^2$ calibration method are all within 10 days of the true $T_c$. Note that these estimated critical times correspond to the times at which the calibrated curves peak. One can see that the quantile curves cover approximately 80% of the span of the noise amplitude around the LPPLS trajectory, where the lowest and largest $q$ provide loose envelopes of the noisy synthetic time series.

The last panel corresponds to the time window [1984.12.03, 1987.10.16], with $t_{end} = 1987.10.16$ being
3 days before the true $T_c$. Given the final acceleration of the log-price close to $T_c$ described by the LP-PLS model, and in the presence of noise, we expect that the calibration becomes more sensitive to idiosyncratic realisations of the noise. This is illustrated by the divergence between the fitted functions obtained for low $q$’s, large $q$’s and the $L^2$ calibrated curves. Compared with the true value $T_c = 1987.10.16$, the $L^2$ calibrated curve exhibits a significantly delayed prediction $\hat{T}_c = 1987.11.13$. For this specific noise realisation, the quantile regressions cluster into two well-defined families. The lower $q$ fits ($0.1 \leq q \leq 0.5$) predict more delayed $\hat{T}_c$ in the range $[1987.12, 1988.01.19]$. In contrast, the larger $q$ fits ($0.6 \leq q \leq 0.9$) predict $\hat{T}_c$ in the range $[1987.09.21, 1987.10.03]$, even earlier than $T_c$. This illustrates the first advantage of quantile regressions for LPPLS signals, that is, to provide a range of scenarios that can bracket the true value of $T_c$.

More generally, one never knows precisely how the noise entangles with the LPPLS signals. While it is standard to assume additive noise (to the log-price), there is no reason to exclude more complex combinations of the signal and the noise. In such cases, as well as in the presence of model errors (the true generating process is not known and the LPPLS model is only an approximation), quantile regressions provide a useful reading of the influence of the different noise quantile levels on the calibration results.

![Image](image-url)

Fig. 1. Quantile-based LPPLS fitting curves of a synthetic time series for the log-price generated by expression (4) with the parameters calibrated on the S&P 500 1987 bubble, to which is added a noise $\epsilon_t \sim N(0, 0.027)$, in the time window $[1984.07.30, 1987.10.16]$ (bottom panel), $[1984.09.21, 1987.08.06]$ (middle panel) and $[1984.12.03, 1987.10.16]$ (top panel). The size of the calibration window is $dt=750$ trading days. The blue line is the in-sample synthetic price time series, which is extended by the blue dashed out-of-sample data. A coloured bundle of quantile-based fitting curves for $q = 0.10, 0.20, ..., 0.90$ brackets the synthetic time series from bottom up as $q$ increases. We also show the standard $L^2$-based fitting curve (red thick line). The black dashed vertical lines are the corresponding end date $t_{end}$ of the fitting window. The red dashed vertical line is the true critical date $T_c = 1987.10.19$ used in generating the synthetic log-price time series.

3.2. Multi-scenario analysis: combining the information obtained from the different quantile regressions

3.2.1. $\hat{T}_c(q, dt)$ versus $t_{end}$

In a real time situation, $t_{end}$ of the time window corresponds to the end time for one to look at all the available
data. Considering a potential bubble bursting at the true critical time $T_c$, $t_{end}$ is also a “present” time to associate with the LPPLS signals and to predict the bubble ends at $\hat{t}_c$. Intuitively, there is an inherent tradeoff among these three times $t_{end}$, $\hat{t}_c$ and $T_c$. When $t_{end}$ is far from $T_c$, it is unlikely that the existing information is rich enough to provide an accurate prediction $\hat{t}_c$. Conversely, when $t_{end}$ is close to $T_c$, the singular nature of the LPPLS trajectory makes the determination $\hat{t}_c$ sensitive to the idiosyncratic realisation of the noise.

It is thus necessary to study their relationships systematically. We introduce the $t_{end} - \hat{t}_c$ plane as shown in Fig. 2, in which $T_c$ is indicated by the red dashed horizontal and vertical lines. The black diagonal line $\hat{t}_c = t_{end}$ separates the region where the estimated burst is in the future ($\hat{t}_c > t_{end}$, domain above the diagonal) from the region where the estimated burst is in the past ($\hat{t}_c < t_{end}$, domain below the diagonal). Then, one can identify six possible regions (represented by the roman numbers I to VI) associated with the different relationships among $t_{end}$, $\hat{t}_c$ and $T_c$. The grey band represents the searched interval $[t_{end} - \eta dt, t_{end} + \eta dt]$ of $\hat{t}_c$ in the calibration, as explained in section 2.

- **Regions I and VI** are the most desired situations in which the bubble is on-going and the predicted $\hat{t}_c$ is in the future (true positives).

- **Region II** corresponds to a failure of the prediction that purports that the bubble has ended ($\hat{t}_c < t_{end}$), while this is not true ($T_c > t_{end}$). This corresponds to false negatives.

- **Regions III and IV** represent the case where the bubble has already ended and the calibration correctly diagnoses it (true negatives).

- **Region V** is another failure of the prediction, which is opposite to region II. The prediction is that the bubble continues and its critical time $\hat{t}_c$ is in the future ($\hat{t}_c > t_{end}$), while it has truly ended ($T_c < t_{end}$). This corresponds to false positives.

---

**Fig. 2.** $t_{end} - \hat{t}_c$ plane in which the six possible regimes associated with the different relationships between $t_{end}$, $\hat{t}_c$ and $T_c$ are depicted. The true time $T_c$ at which the bubble bursts is indicated by the red dashed horizontal and vertical lines.

Ninety nine values $(\hat{t}_c(q, dt)|q = 0.01, 0.02, ..., 0.99)$ are obtained by quantile regression with $dt = 500$ trading days for various $t_{end}$ sliding in steps of 5 trading days within [1986.05.12, 1988.08.29], for the synthetic noisy LPPLS time series already used in Fig. 1. Fig. 3 show the evolution of the averages (resp. medians) of the set of 99 $q$’s, which are represented by the black squares (resp. red stars). They can be compared with the standard OLS fits shown as the blue triangles. The red dashed horizontal and vertical lines represent the true critical date $T_c = 1987.10.19$. The black dashed diagonal line represents $\hat{t}_c = t_{end}$.

The main message of Fig. 3 is that the OLS method seems to over-perform significantly in the following sense. One can observe that its corresponding $\hat{t}_c$ remains stable and close to the true $T_c$ (horizontal red dashed line) about three months before $T_c$, and this value is also confirmed for about four months after $T_c$, giving a total interval for $t_{end}$ of seven months of very stable and consistent prediction. In contrast, the averages and medians $\hat{t}_c$ obtained from the 99 quantile calibrations exhibit a stable determination of $\hat{t}_c$ close to $T_c$ for about four months but only after $T_c$.

The second message is that, when $t_{end}$ is too far from $T_c$, the estimated $\hat{t}_c$ is not stable and systematically underestimates the time of the bubble burst. Moreover, $\hat{t}_c$ is found to move upward proportionally to $t_{end}$ as the later increases. This observation holds for all three estimators (i.e., average, median and OLS fit).
3.2.2. Quantile-Violin representation $q(t_c) - pd f(t_c(q))$ of the ensemble of quantile regression functions

The results of Fig. 3 are far from constituting the whole story since the quantile regressions can give much more than just an average or median tendency. As already pointed out with Fig. 1, several scenarios can be revealed by scanning over the quantiles. In order to capture the wealth of information of these 99 functions obtained for each $t_{\text{end}}$, we introduce a generalisation of the violin plot (Hintze & Nelson, 1998) and call it “Quantile-Violin plot”, in which the standard box plot is complemented by a rotated kernel density plot on its right side, and the corresponding $q$ values are given on the left side. Thus, Fig. 4 presents the Quantile-Violin plots of $\tilde{t}_c$ for the seven values of $t_{\text{end}}$ = 1987.03.19, 1987.04.30, 1987.06.25, 1987.08.06, 1987.10.15, 1988.02.11 and 1988.04.21. Here, the results are determined by quantile regression of the same synthetic noisy LPPLS time series as Figs. 1 and 3, and $dt = 500$ trading days. For a given $t_{\text{end}}$, the right side of the Quantile-Violin plot gives the rotated kernel density function of $\tilde{t}_c$ over the set of 99 quantiles, as well as the descriptive statistics, such as the median (red line), the upper quartile (blue line), the mean (black line) and the lower quartile (brilliant blue line). These values can be read on the scale along the main vertical $t_c$ axis. The left side of the Quantile-Violin plot gives values of $q$ for each $\tilde{t}_c$ contributing to the distribution on the right side, with $q = 0$ on the central axis and $q = 1$ corresponding to the maximum extension to the left. The red dashed horizontal and vertical lines represent the real critical date $T_c = 1987.10.19$.

The first important message of Fig. 4 is that the median and mean of $\tilde{t}_c$ are misleading because its probability density function is multimodal for $t_{\text{end}} < T_c$. It is interesting to see that, when $t_{\text{end}}$ is quite earlier than $T_c$, there are scenarios in which $\tilde{t}_c$ tends to be stable and much closer to the true value than the mean, median or OLS estimates shown in Fig. 3. From mid-1987, one can observe that one of the modes of the distribution of $\tilde{t}_c$ is placed on the correct value $T_c$. When $t_{\text{end}}$ is close to $T_c$, one can also see how sensitive the quantile regressions are as four main scenarios appear corresponding to four modes of the distribution of $\tilde{t}_c$. 

![Quantile-Violin plot](image-url)
We again use 1987.08.06, 1987.10.15, 1988.02.11 and 1988.04.21, determined by quantile regression of the same synthetic noisy LPMLS time series already used in Figs. 1 and 3. We again use \( dt = 500 \) trading days. For each \( t_{\text{end}} \), the right side of the Quantile-Violin plot gives the rotated kernel density function of \( \hat{t}_c \) over the set of 99 quantiles, as well as the descriptive statistics, such as the median (red line), the upper quartile (blue line), the mean (black line) and the lower quartile (brilliant blue line). These values can be read on the scale along the main vertical axis \( t_c \) axis. The left side of the Quantile-Violin gives values of \( q \) for each \( \hat{t}_c \) contributing to the distribution on the right side, with \( q = 0 \) on the axis and \( q = 1 \) corresponding to the maximum extension to the left. The red dashed horizontal and vertical lines represent the real critical date \( T_c = 1987.10.19 \).

3.2.3. \( dt \)-Violin representation \( dt(t_c) - \text{pdf}(t_c(dt)) \) of the ensemble of quantile regression functions

Previous works have shown the importance of a multi-scale analysis (see e.g., Sornette & Zhou, 2006). In our case, for a fixed \( t_{\text{end}} \), this amounts to scan \( t_{\text{start}} \) and redo the analysis for each window. Specifically, we shift \( t_{\text{start}} = t_{\text{end}} - dt \) in steps of 5 trading days, obtaining 126 windows of sizes \( dt = 750, 745, ..., 125 \) trading days. For each window \([t_{\text{start}}, t_{\text{end}}]\), we perform the OLS estimation and the quantile regression of the model (4) on the same synthetic noisy LPMLS time series already used in Figs. 1, 3 and 4, obtaining a set \([\hat{t}_c(q, dt)]q = 0.01, 0.02 ... 0.99\). This procedure is summarised in Fig. 5.

Fig. 4. Quantile-Violin plots of \( \hat{t}_c \)'s for the seven values of \( t_{\text{end}} = 1987.03.19, 1987.04.30, 1987.06.25, 1987.08.06, 1987.10.15, 1988.02.11 \) and 1988.04.21, determined by quantile regression of the same synthetic noisy LPMLS time series already used in Figs. 1 and 3. We again use \( dt = 500 \) trading days. For each \( t_{\text{end}} \), the right side of the Quantile-Violin plot gives the rotated kernel density function of \( \hat{t}_c \) over the set of 99 quantiles, as well as the descriptive statistics, such as the median (red line), the upper quartile (blue line), the mean (black line) and the lower quartile (brilliant blue line). These values can be read on the scale along the main vertical axis \( t_c \) axis. The left side of the Quantile-Violin gives values of \( q \) for each \( \hat{t}_c \) contributing to the distribution on the right side, with \( q = 0 \) on the axis and \( q = 1 \) corresponding to the maximum extension to the left. The red dashed horizontal and vertical lines represent the real critical date \( T_c = 1987.10.19 \).

Analogously to Fig. 4, Fig. 6 presents a synopsis of the results concerning the estimation of \( \hat{t}_c \), but now over the population of the 126 windows for the fixed \( t_{\text{end}} = 1987.08.06 \) and the various \( q \)'s. We further generalise the violin plot (Hintze & Nelson, 1998) in the form of “\( dt \)-Violin plots”. The standard box plot is now complemented by a rotated kernel density plot of \( \hat{t}_c \) over the set of 126 windows on its right side for a fixed \( q \), and the corresponding \( dt \) values are added on the left side. Specifically, Fig. 6 shows seven \( dt \)-Violin plots for the seven values \( q = 0.05, 0.10, 0.20, 0.30, 0.50, 0.80 \) and 0.90. The kernel density distribution of \([\hat{t}_c(dt)]dt = 750, 745, ..., 125 \) trading days] is shown rotated on the right side, as well as some descriptive statistics, such as the median (red line), the upper quartile (blue line), the mean (black line) and the lower quartile (brilliant blue line). The left side of the \( dt \)-Violin plot gives the values of \( dt \) for each \( \hat{t}_c \) contributing to the distribution on the right side. The smallest window size of 125 days is on the central vertical axis of the \( dt \)-Violin plots while the largest window size of 750 days corresponds to the maximum distance to the left.

The following conclusions can be drawn for this particular \( t_{\text{end}} = 1987.08.06 \) (represented by the horizontal black dashed line). First, all quantile fits predict essentially the same values for the mean, median and lower quartile of \( \hat{t}_c \), and these values are about 40-day earlier than the true critical date \( T_c = 1987.10.19 \) (horizontal red dashed line). The quantile predictions differ however on the shape of the density distribution of \( \hat{t}_c \). The quantile fits for \( q = 0.30 \) to 0.80 have significantly heavier tails towards large values of \( \hat{t}_c \), even producing a second local mode about three weeks after \( T_c \). From the perspective of a decision maker, this corresponds to a second possible scenario, which together with the main mode brackets \( T_c = 1987.10.19 \).

The left side of each \( dt \)-Violin plot features dots that
are organised in rays, showing that the predicted $\widehat{t}_c$ form several families, and in each family $\widehat{t}_c$ is an affine function of the size $dt$ of the window of analysis. Thus, the larger $dt$, the longer the horizon over which $\widehat{t}_c$ is expected to occur. Shorter windows are more sensitive to local structures that may announce a proximal turning point. Long windows integrate the LPPLS signals over a large span and are therefore less sensitive to local structures. The dominating mode of the density distribution of $\widehat{t}_c$ is contributed by the leftmost rays, while the fat tail towards large $\widehat{t}_c$ is created by the rightmost rays.

![Fig. 6. Seven dt-Violin plots of $\widehat{t}_c$’s over the set $\{\widehat{t}_c(dt)|dt = 750, 745, ..., 125$ trading days$\}$ for the seven values $q=0.05, 0.10, 0.20, 0.30, 0.50, 0.80$ and $0.90$ and for the fixed $t_{end} = 1987.08.06$ determined by quantile regression of the same synthetic noisy LPPLS time series already used in Fig. 1, 3 and 4. The kernel density distribution of $\{\widehat{t}_c(dt)|dt = 750, 745, ..., 125$ trading days$\}$ is shown rotated on the right side of each dt-Violin plot, as well as the median (red line), the upper quartile (blue line), the mean (black line) and the lower quartile (brilliant blue line). The left side of each dt-Violin plot gives the values of $dt$ for each $\widehat{t}_c$ contributing to the distribution on the right side. The smallest window size of 125 days is on the central vertical axis of the dt-Violin plots while the largest window size of 750 days corresponds to the maximum distance to the left. The black dashed horizontal line indicates $t_{end} = 1987.08.06$. The red dashed horizontal line shows the true critical date $T_c = 1987.10.19$.

Notwithstanding this multi-ray structure of $\widehat{t}_c$ as a function of $dt$, Fig. 7 shows that the average and median of the 99 $q$’s estimates $\{\widehat{t}_c(q, dt)|q = 0.01, 0.02,...0.99\}$ for each $dt$ and for the fixed $t_{end} = 1987.08.06$ are stable. They constantly predict the end of the bubble about one month before the true value $T_c = 1987.10.19$. In contrast, the standard OLS estimation of $\widehat{t}_c$ is very sensitive to the chosen size $dt$ of the window, leading to inconclusive diagnostics. Thus, the quantile regressions introduce stability in the forecasts when they are exploited as an ensemble of scenarios.

**Fig. 6.** Seven $dt$-Violin plots of $\widehat{t}_c$’s over the set $\{\widehat{t}_c(dt)|dt = 750, 745, ..., 125$ trading days$\}$ for the seven values $q=0.05, 0.10, 0.20, 0.30, 0.50, 0.80$ and $0.90$ and for the fixed $t_{end} = 1987.08.06$ determined by quantile regression of the same synthetic noisy LPPLS time series already used in Fig. 1, 3 and 4. The kernel density distribution of $\{\widehat{t}_c(dt)|dt = 750, 745, ..., 125$ trading days$\}$ is shown rotated on the right side of each $dt$-Violin plot, as well as the median (red line), the upper quartile (blue line), the mean (black line) and the lower quartile (brilliant blue line). The left side of each $dt$-Violin plot gives the values of $dt$ for each $\widehat{t}_c$ contributing to the distribution on the right side. The smallest window size of 125 days is on the central vertical axis of the $dt$-Violin plots while the largest window size of 750 days corresponds to the maximum distance to the left. The black dashed horizontal line indicates $t_{end} = 1987.08.06$. The red dashed horizontal line shows the true critical date $T_c = 1987.10.19$.

**Fig. 7.** Predicted critical time $\widehat{t}_c$ as a function of $\{dt = 750, 745, ..., 125$ trading days$\}$ by two methods for the same synthetic noisy LPPLS time series used in Fig. 6: (i) averages (black squares) and medians (red stars) of the 99 $q$’s estimates $\{\widehat{t}_c(q, dt)|q = 0.01, 0.02,...0.99\}$ obtained by quantile regression; (ii) standard OLS estimates (blue triangles). The red horizontal dashed line gives the position of the true critical date $T_c = 1987.10.19$. The black horizontal dashed line shows $t_{end} = 1987.08.06$.

4. Application to the diagnostic of the “S&P 500 1987” bubble and the prediction of its end

Equipped with the tests on synthetic time series presented in the previous section, we now apply them to the S&P 500 1987 bubble that burst on the famous Black Monday 19 October 1987.

4.1. LPPLS quantile regression curves for different quantile probability level $q$

Fig. 8 shows the calibrations by expression (4) using the quantile regression method (8) with (7) for nine quantile probability level $q = 0.10, 0.20,..., 0.90$. The smaller (resp. larger) values of $q$ tend to fit the lowest (resp. highest) part of the empirical price time series, providing together a bundle of fits that seem quite reasonable visually. For the first two panels corresponding...
to $t_{end}$ not to close from the crash, one can observe that, apart from the lowest quantiles that exhibit more variability, the higher values of the quantiles provide consistent fits with estimated values of the critical time $\hat{t}_c$ close to the true value $T_c$. In contrast, the standard $L^2$-based fit tends to overshoot, similarly to the lowest quantiles. The situation reverses for the last panel with $t_{end}$ being very close to the crash, for which most of the quantiles (and the $L^2$-based fit) overshoot significantly by about five months, while the lowest curve for $q = 0.10$ undershoots by approximately two months.

Fig. 8. Nine coloured calibrated curves obtained by the method of LPPLS quantile regression for nine different $q$ of the “S&P 500 1987” bubble in the time window [1984.07.30, 1987.06.12] (top panel), [1984.09.21, 1987.08.06] (middle panel) and [1984.12.03, 1987.10.16] (bottom panel). The in-sample continuous curves are extended out-of-sample as dashed lines with the same colours. The noisy black line is the in-sample empirical price time series, followed by the black dashed out-of-sample data. The red thick line is the standard $L^2$-based fitting curve for comparison, which is extended by the red dashed out-of-sample calibrated curve. The black dashed vertical line shows the value of $t_{end}$ used in the calibration. The red dashed vertical line gives the true $T_c = 1987.10.19$.

4.2. Multi-scale analysis of $\hat{t}_c$ as a function of $dt$ and $q$

Fig. 9 presents the dependence of $\hat{t}_c$ as a function of $t_{start}$ for the fixed $t_{end}$ =1987.08.06. The four lines respectively correspond to the results for three different characteristic values of $q$ and for the standard $L^2$ fit. While the latter method gives quite unstable results, the determination of $\hat{t}_c$ by quantile regressions provides consistent and reasonable results for $t_{start}$ earlier than mid-1985. After this time, the two methods tend to predict values of $\hat{t}_c$ much earlier than $T_c$. 

10
The horizontal red dashed line shows the true critical time \( T_c = 1987.10.19 \).

Fig. 10 shows the dependence of \( \hat{t}_c(q, dt) \) as a function of \( t_{end} \) for two different window sizes of the S&P 500 trading days, obtained from quantile regressions with 99 values \( q = 0.01, 0.02, ..., 0.99 \) for the S&P 500 1987 bubble. The red horizontal and vertical dashed lines represent the true critical time \( T_c = 1987.10.19 \). The black tilted dashed line represents the diagonal line \( t_c = t_{end} \).

The Quantile-Violin plots (represented by \( q(t_c) - pdf(t_c(q)) \)) provide a more in-depth view of the unfolding scenarios obtained by the LPPLS quantile regressions performed with the search ranges \( m \in [0, 2] \), \( \omega \in [1, 50] \) and \( t_c \in [t_{end} - 0.20dt, t_{end} + 0.20dt] \). Specifically, Fig. 11 plots the results for the S&P 500 1987 bubble, where the three panels correspond to \( dt = 500, 750 \) and 1000 trading days, respectively. Each panel contains seven Quantile-Violin plots associated with the seven \( t_{end} = 1987.03.19, 1987.04.30, 1987.06.25, 1987.08.06, 1987.10.15, 1988.02.11 \) and 1988.04.21.

For \( dt = 500 \) days, one can observe the stabilisation for \( t_{end} = 1987.06.25, 1987.08.06 \) and 1987.10.15 of a set of scenarios bracketing the true critical time \( T_c = 1987.10.19 \). Earlier \( t_{end} \)'s predictions are too far from \( T_c \) to have it in their prediction horizon. A qualitatively similar picture emerges for \( dt = 750 \) days, albeit more murky, with a larger spread of the estimated \( \hat{t}_c \)'s. A similar behavior is obtained for the larger time scale \( dt = 1000 \) days, with an even broader set of scenarios around the true \( T_c \).

Fig. 12 is the same as Fig. 11 but for LPPLS quantile regressions performed with the more restrictive search conditions \( m \in [0.1, 0.9], \omega \in [6, 13] \) and \( t_c \in [t_{end} - 0.20dt, t_{end} + 0.20dt] \), which are derived from previous investigations (Johansen & Sornette, 2010; Jiang, Zhou, Sornette, Woodard, Bastaiaensen & Cauwels, 2010; Fil-
Reducing the search space of the two key nonlinear LPPLS parameters $m$ and $\omega$ has two major effects: (i) the distributions of $\hat{T}_c$ tend to be more stable as a function of $t_{end}$ and bracket the true $T_c$ for all cases, except for the earliest $t_{end}=1987.03.19$ at the shortest time scale $dt=500$ days; (ii) the spreads of $\hat{T}_c$ values over the different scenarios are narrower, indicating that the LPPLS quantile regressions provide more precise predictions of the true $T_c$.

Fig. 11. Quantile-Violin plots of $\hat{T}_c$ for the seven values of $t_{end}=1987.03.19$, 1987.04.30, 1987.06.25, 1987.08.06, 1987.10.15, 1988.02.11 and 1988.04.21, determined by quantile regression of the log-price of the S&P 500 1987 bubble with formula (4) with the search space: $m \in [0, 2]$, $\omega \in [1, 50]$ and $t_c \in [t_{end} - 0.20 dt, t_{end} + 0.20 dt]$. The three panels correspond to $dt=500$, 750 and 1000 trading days, respectively. For each $t_{end}$, the right side of the Quantile-Violin plot gives the rotated kernel density function of $\hat{T}_c$ over the set of 99 quantiles, as well as the descriptive statistics, such as the median (red line), the upper quartile (blue line), the mean (black line) and the lower quartile (brilliant blue line). These values can be read on the scale along the main vertical axis $t_c$ axis. The left side of the Quantile-Violin gives values of $q$ for each $\hat{T}_c$ contributing to the distribution on the right side, with $q=0$ on the axis and $q=1$ corresponding to the maximum extension to the left. The red dashed horizontal and vertical lines represent the real critical date $T_c = 1987.10.19$. 
the largest $q = 0.90$ gives a predicted $pdf(\hat{t}_c)$ with its mode very close to $T_c$. The other quantiles have their main mode earlier, roughly in the middle of $t_{end}$ and $T_c$. The third panel for $t_{end} = 1987.10.15$ very close to $T_c$ exhibits a strong bimodal (and a trimodal for the lowest quantiles) structure of the $pdf(\hat{t}_c)$, bracketing $T_c$ associated with two modes. The main mode occurs about one month and a half earlier than $T_c$ for all $q$, while it is two months later than $T_c$ for the largest $q = 0.90$.

The $dt$-Violin plots of $\hat{t}_c$ for the S&P 500 1987 bubble are presented in Fig. 13, where the three panels correspond to $t_{end} = 1987.06.25$, 1987.08.06 and 1987.10.15, respectively. Each panel contains seven $dt$-Violin plots associated with the seven values of $q = 0.05, 0.10, 0.20, 0.30, 0.50, 0.80, 0.90$. Each $dt$-Violin plot is constructed over the statistics obtained over the following set of $dt$ values: $[dt = 750, 745, \ldots, 125 \text{ trading days}]$. This provides an ensemble view of the predicted transition times $\hat{t}_c$ over a large set of window scales.

The top panel for $t_{end} = 1987.06.25$ demonstrates that essentially all $q$’s predictions are approximatively the same, in the sense that the modes of the $pdf(\hat{t}_c)$ are close to the true critical date $T_c = 1987.10.19$. When $t_{end} = 1987.08.06$ approaches $T_c$ as shown in the middle panel, the prediction quality deteriorates with the $pdf(\hat{t}_c)$ both broadening and becoming bimodal. Only
Fig. 13. Seven $dt$-Violin plots of $\hat{t}_c$ constructed as in Fig. 6 for the S&P 500 1987 bubble, associated with the seven values of $q = 0.05, 0.10, 0.20, 0.30, 0.50, 0.80, 0.90$, for $t_{end}=1987.06.25$ (top panel), 1987.08.06 (middle panel) and 1987.10.15 (bottom panel). Each $dt$-Violin plot is constructed over the statistics obtained over the set $\{dt = 750, 745, \ldots, 125$ trading days$\}$. The black dashed horizontal lines in each panel indicates $t_{end}$. The red dashed horizontal line shows the $T_c = 1987.10.19$.

5. Applications to the prediction of the end of four historical bubbles

The previous section has studied the S&P 500 1987 bubble in great details. But this is just one case. We now extend our analysis to three additional historical bubbles listed in Table 1 to explore the ensemble behavior of the prediction of their critical end times over the set $\{t_{end}=1987.08.06, 1987.10.09, 1987.10.09\}$ and over 99 quantiles. We refer to these three additional historical bubbles by the names of the involved markets and the years when they burst. The first one is S&P 500 2007, which was studied in Sornette and Cauwels (2014, 2015a). The second and third one are SSEC 2007 and SSEC 2009, discussed in details in Jiang et al. (2010). For each bubble, we picked one value of $t_{end}$, spanning from one to three months before the crash that terminated the bubble, as given in Table 1.

Table 1
List of four historical bubbles for fixed $t_{end}$, with the corresponding $t_{end}$ used for the analysis of the LPPLS quantile regressions and the time $T_c$ at which each bubble burst.

| Asset & Year of crash | Selected $t_{end}$ | $T_c$  |
|-----------------------|-------------------|--------|
| S&P 500 1987          | 1987.08.06        | 1987.10.19 |
| S&P 500 2007          | 2007.07.25        | 2007.10.09 |
| SSEC 2007             | 2007.09.10        | 2007.10.18 |
| SZSC 2009             | 2009.04.23        | 2009.07.10 |

Fig. 14 shows the medians (red stars) and averages (black squares) of $\hat{t}_c(q, dt)$ as a function of $q$ for the fixed $t_{end}$ given in Table 1, over the population of window sizes spanning $\{dt = 750, 745, \ldots, 125$ trading days$\}$. For the S&P 500 1987 bubble (the first panel) and the S&P 500 2007 bubble (the second panel), the results confirm the previous analysis by showing that the LPPLS quantile regressions provide significantly better predictions than the standard $L^2$ calibration based predictions. Only for the SSEC 2007 bubble (the third panel), we observe significant variations of the medians and averages as functions of $q$. For the medians, we see an approximate plateau for $q$ between 0.50 and 0.80, which slightly overestimates the true $T_c$, but is earlier than the $L^2$ calibration based prediction (blue line). Lower (resp. larger) $q$’s predictions underestimate (resp. overestimate) the true $T_c$. In the case of the SZSC 2009 bubble (the last panel), all quantiles give again consistent predictions for $\hat{t}_c$, which are however too early by about one month. Its $L^2$ calibration based prediction is closer to the true $T_c$, while slightly overestimating it.
Summarising the results of these four cases, the quantile regressions are better than the $L^2$ calibration in two cases, approximately the same in one case and worse in the last case. Just on the basis of the statistics shown in Fig. 14, it seems that the LPPLS quantile regressions may not be significantly better than the $L^2$ calibration. We thus need to recall that the synthetic tests in section 3 have presented that this is not the case. For these four bubbles, Fig. 15 thus strengthens this conclusion by showing again the unstable behaviour of $L^2$ calibrations compared with the LPPLS quantile regressions as a function of the window sizes. For more details, the medians and averages of quantile estimates are shown as functions of $dt$ for the fixed $t_{end}$ given in Table 1, over the population of $q$ values spanning $[q = 0.01, 0.02, ..., 0.99]$. Overall, one can observe a quite erratic behavior of $\hat{t}_c$ for the $L^2$ calibration in Fig. 15, compared to a much more stable behavior for the quantile regressions. The latter exhibit approximate plateaus of stability of the predicted $\hat{t}_c$ as a function of $dt$, which gives confidence in the reliability of the detected LPPLS signal as a function of time scale. This is particularly evident for the S&L 500 2007 bubble (second panel), for which the stable plateau extends almost over the whole range of $dt$.

As already mentioned, in the presence of the stochastic nature of log-price, significant variability in the predictions is unavoidable, and the advantage of the LPPLS quantile regressions is to provide bundles of scenarios. The relevant analysis on their stability and sensitivity to $t_{end}, dt$ and $q$ is very helpful, as shown in this and previous sections.
6. Consolidated DS LPPLSTM indicators

The previous sections have presented a wealth of measures, summarised through the use of the quantile-Violin and \( dt \)-Violin plots, which represent the ensemble of predictions for a given present time \( t_{\text{end}} \) over the set of quantile levels \( q \) used in the LPPLS quantile regression, and over the set of time scales (i.e., window sizes) \( dt \) used in the calibrations. While informative, the effective use of so many fluctuating and often conflicting signals to inform on the danger for a bubble burst and to trigger an actionable decision remains a challenge. To address this, we propose two indicators that aggregate these signals, inspired from previous works on historic bubbles (Sornette & Zhou, 2006; Jiang et al., 2010; Sornette & Cauwels, 2015a). These two indicators have been briefly discussed to present the ex-ante forecast of the Chinese bubble and its burst that started in June 2015 (Sornette et al., 2015).

(1) The DS LPPLSTM Confidence indicator is the fraction of fitting windows whose calibrations meet the filtering condition 1 in Table 2. It thus measures the sensitivity of the observed bubble pattern to the 126 time windows of duration from 125 to 750 trading days. A large value indicates that the LPPLS pattern is found at most scales and is thus more reliable. If the value is close to one, the pattern is practically insensitive to the choice of \( dt \). A small value of the indicator signals a possible fragility since it is presented in a few fitting windows.

(2) The DS LPPLSTM Trust indicator quantifies the sensitivity of the calibrations to the specific realised instance of the noise in the financial time series. Because the calibration is an attempt to disentangle the LPPLS signal from an unknown realisation of the residuals, we generate bootstrap samples of the original data 100 times and add the residuals to the calibrated LPPLS price that proxy for 100 supposed independent realisations of equivalent price patterns. The DS LPPLS Trust indicator is defined as the median level over the 126 time windows of the fraction among the 100 synthetic time series that satisfy the filtering condition 2 in Table 2. It thus measures how closely the theoretical LPPLS model matches the empirical price time series, 0 being a bad and 1 being a perfect match.

(3) Arithmetic average and geometric average of the DS LPPLSTM Confidence indicator and DS LPPLSTM Trust indicator: combining these two indicators is instructive to join the two types of information on the time
scale over which the LPPLS signal appears and on the quality of the fits.

**Table 2**

Search space and filtering conditions for the qualification of valid LPPLS fits. Within the JLS framework, the condition that the crash hazard rate $h(t)$ is non-negative by definition (Bothmer & Meister, 2003) translates into the value of Damping larger than or equal to 1.

| Item        | Search space | Filtering condition 1 | Filtering condition 2 |
|-------------|--------------|-----------------------|-----------------------|
| $m$         | [0, 2]       | [0.1, 0.9]            | [0.1, 0.9]            |
| $\omega$    | [1, 50]      | [6, 13]               | [6, 13]               |
| $t_c$       | $[t_{\text{end}} - 0.2 dt, t_{\text{end}} + 0.2 dt]$ | $[t_{\text{end}} - 0.15 dt, t_{\text{end}} - 0.2 dt]$ | $[t_{\text{end}} + 0.13 dt, t_{\text{end}} + 0.12 dt]$ |
| Number of oscillation: $m$ | -- | (2.5, + $\infty$) | (2.5, + $\infty$) |
| Damping: $\omega$ | -- | (1.2, + $\infty$) | (1, + $\infty$) |
| Relative error: $\frac{\text{pt}}{\hat{\text{pt}}}$ | -- | [0, 0.035] | [0, 0.14] |

7. Empirical analysis of 16 historical bubbles with the consolidated DS LPPLSTM indicators

In order to provide a more extensive test of the LPPLS quantile regression approach, we construct the DS LPPLS Confidence and Trust indicators described in the previous section, for 16 historical bubbles listed in Table 3. These indicators can then be compared with the price time series to allow a judgement of how well they can be associated with bubbles and their terminations.

**Table 3**

List of the 16 historical bubbles obtained from the previous studies (Sornette, Johansen & Bouchaud, 1996; Johansen & Sornette, 2000, 2001, 2010; Sornette, Demos, Zhang, Cauwels, Filimonov & Zhang, 2015) as well as cases reported at the website of the Financial Crisis Observatory at ETH Zurich (www.er.ethz.ch/financial-crisis-observatory.html). The data was obtained from Thomson Reuters Datastream.

| Asset & Year of crash | Data range | Range of $t_{\text{end}}$ |
|-----------------------|------------|---------------------------|
| S&P 500 1987          | 1984.01.02-1987.11.13 | 1986.11.14-1987.11.13 |
| Chile 1991/1994       | 1987.10.01-2000.12.01 | 1990.08.15-2000.12.01 |
| Venezuela 1997        | 1994.01.03-1999.12.30 | 1996.11.15-1999.12.30 |
| Indonesia 1994/1997   | 1990.01.03-1999.12.30 | 1992.11.17-1999.12.30 |
| Malaysia 1994         | 1991.01.01-1995.12.29 | 1993.11.15-1995.12.29 |
| Thailand 1994         | 1990.01.01-1994.12.30 | 1992.11.13-1994.12.30 |
| Hong Kong 1987/1994/1997 | 1980.01.02-1999.12.31 | 1982.11.16-1999.12.31 |
| S&P 500 2007          | 2004.01.01-2009.12.31 | 2006.11.15-2009.12.31 |
| Sugar prices          | 2002.01.01-2013.12.31 | 2004.11.15-2013.12.31 |
| Brent Oil 2008        | 1990.01.01-2015.04.16 | 1992.11.13-2015.04.16 |
| DJIA 1929             | 1926.01.02-1930.12.31 | 1928.07.07-1930.12.31 |
| Nasdaq Composite      | 1993.01.01-2002.12.31 | 1995.11.16-2002.12.31 |
| Index 2000            | 2004.01.01-2014.12.31 | 2006.11.15-2014.12.31 |
| SSEC 2007/2009        | 2004.01.01-2014.12.31 | 2006.11.15-2014.12.31 |
| SZSC 2007/2009        | 2004.01.01-2014.12.31 | 2006.11.15-2014.12.31 |
| Hong Kong 2007        | 2000.01.03-2015.04.10 | 2003.01.17-2015.04.10 |
| SSEC 2015             | 2011.02.23-2015.05.12 | 2014.01.07-2015.05.12 |

Fig. 16 – Fig. 31 present the price time series of the 16 historical bubbles together with the DS LPPLS Confidence and Trust indicators constructed using (i) the $L^2$ fitting method (green curves) and (ii) the quantile regressions (red curves). Since the Confidence and Trust indicators can be constructed for each quantile level $q$, we choose to present them for $q = 0.10$ as well as for their arithmetic and geometric averages over the 9 deciles ($q = 0.10, 0.20, ..., 0.90$). The geometric average is more conservative than the arithmetic average, since it requires all deciles to give a non-zero signal in order to be non-vanishing. In contrast, the arithmetic average just needs one of the 9 deciles to give a signal. The arithmetic average strives for improving sensitivity (i.e., the proportion of positives that are correctly identified as such) while the geometric average improves specificity (i.e., reducing the false negatives).

Overall, the DS LPPLS Confidence and Trust indicators are found to have strong discriminating power to identify the market regimes during which prices tend to accelerate upward (resp. downward), which are followed by strong corrections (resp. rallies). This conclusion holds both for the $L^2$ fitting method (green curves) and (ii) the quantile regressions (red curves). However, the quantile regressions add to the $L^2$ fitting method by providing in general earlier warning signals. In addition, in particular using the lower quantiles, one can observe a larger sensitivity towards the detection of negative bubbles and the subsequent rebound by using the quantile regressions.
Fig. 16. S&P 500 1987 bubble. The three groups of three panels represent respectively from top to bottom the DS LPPLS Trust indicator, the DS LPPLS Confidence indicator and their product. In all panels, the green line is obtained by using the standard $L^2$ calibration method while the red lines are obtained using quantile regressions. In each group, the top panel is obtained using the first decile $q = 0.10$ quantile regression, the middle panel is the arithmetic average over the 9 deciles $\{q = 0.10, 0.20, ..., 0.90\}$ and the bottom panel is the geometric average over the same 9 deciles $\{q = 0.10, 0.20, ..., 0.90\}$.

Fig. 17. Chile 1991/1994: same as Fig. 16. The negative readings of indicators signal the diagnostic of “negative” bubbles (Siegel, 2003; Sornette & Cauwels, 2015a), whose end corresponds to a “negative crash” (i.e., a rally or rebound).
Fig. 18. Venezuela 1997: same as Fig. 16. The main features here are the identification of a strong negative bubble and its rebound.

Fig. 19. Indonesia 1994/1997: same as Fig. 16.
**Fig. 20. Malaysia 1994:** same as Fig. 16. This bubble exhibits a remarkably “clean” LPPLS pattern, so that all indicators exhibit very strong consistency and precision to identify the peak and subsequent burst.

**Fig. 21. Thailand 1994:** same as Fig. 16.
Fig. 22. Hong Kong 1987/1994/1997: same as Fig. 16. A number of bubbles and crashes occurred over the period shown here, as shown in Fig. 10 in Sornette and Woodard (2010).

Fig. 23. S&P 500 2007: same as Fig. 16.
Fig. 24. Sugar prices: same as Fig. 16.

Fig. 25. Brent Oil 2008: same as Fig. 16. An ex-ante diagnostic of the bubble and forecast of the crash in July 2008 was reported in Sornette, Woodard and Zhou (2009).
Fig. 26. DJIA 1929: same as Fig. 16.

Fig. 27. Nasdaq Composite Index 2000: same as Fig. 16. See Johansen and Sornette (2000) for a detailed account of the dot-com bubble that crashed in 2000.
Fig. 28. SSEC 2007/2009: same as Fig. 16. See Jiang, Zhou, Sornette, Woodard, Bastiaensen and Cauwels (2010) for a detailed description of this bubble and ex-ante diagnostics.

Fig. 29. SZSEC 2007/2009: same as Fig. 28.
Fig. 30. Hong Kong 2007: same as Fig. 16.

Fig. 31. SSEC 2015: same as Fig. 16. An account of the real-time ex-ante diagnostic of this bubble was represented in Sornette et al. (2015).

8. Concluding remarks

We have shown that positive (resp. negative) bubbles followed by large crashes/corrections (resp. rallies) can be identified by diagnosing the existence of log-periodic power law singular (LPPLS) structures in the log-price dynamics. The innovation of the present article includes
(1) the introduction of the quantile regression applied to the LPPLS detection problem, (2) the combination of the many quantile regressions with a multi-scale analysis and (3) the implementation of the DS LPPLS Confidence and Trust indicators that provide an aggregation and consolidation of the wealth of generated information. We have shown both using extensive synthetic signals, a detailed analysis of the S&P 500 1987 bubble and the application to 16 historical bubbles that the quantile regression of LPPLS signals contributes useful early warning signals. The aggregated DS LPPLS Confidence and Trust indicators have been offered to combine the ensemble of signals generated at multi-scales and many quantile levels. The many figures constructed for the 16 historical bubble cases demonstrate their significant predictive ability around the real critical time when the burst/rally occurs. Overall, we find that the quantile regression method improves on the $L_2$ based calibration method by providing richer and more stable scenarios.

References

Abergel, F., Aoyama, H., Chakrabarti, B. K., Chakraborti, A., & Ghosh, A. (2015). Econophysics and data driven modelling of market dynamics. Milan: Springer-Verlag.

Araújo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. Trends in Ecology & Evolution, 22, 42–47.

Barnes, M. L., & Hughes, A. T. W. (2002). A quantile regression analysis of the cross section of stock market returns. Working Paper, Federal Reserve Bank of Boston.

Barroda, I., & Roberts, F. (1974). Solution of an overdetermined system of equations in the $L_1$ norm [F4]. Communications of the ACM, 17, 319–320.

Black, F. (1986). Noise. The Journal of Finance, 41, 529–543.

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

Filimonov, V., & Sornette, D. (2013). A stable and robust calibration scheme of the log-periodic power law model. Physica A: Statistical Mechanics and its Applications, 392, 3698–3707.

Guzman, S., & Sornette, D. (2002). Log-periodic route to fractal functions. Physical Review E, 65, 036142.

Hintze, J. L., & Nelson, R. D. (1998). Violin plots: A box plot-density trace synergism. The American Statistician, 52, 181–184.

Hüsler, A., Sornette, D., & Hommes, C. H. (2013). Super-exponential bubbles in lab experiments: Evidence for anchoring over-optimistic expectations on price. Journal of Economic Behavior & Organization, 92, 304–316.

Iide, K., & Sornette, D. (2002). Oscillatory finite-time singularities in finance, population and rupture. Physica A: Statistical Mechanics and its Applications, 307, 63–106.

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

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

Johansen, A., & Sornette, D. (2000). The Nasdaq crash of April 2000: Yet another example of log-periodicity in a speculative bubble ending in a crash. The European Physical Journal B-Condensed Matter and Complex Systems, 17, 319–328.

Johansen, A., & Sornette, D. (2001). Bubbles and anti-bubbles in Latin-American, Asian and Western stock markets: An empirical study. International Journal of Theoretical and Applied Finance, 4, 853–920.

Johansen, A., & Sornette, D. (2010). Shocks, crashes and bubbles in financial markets. Brussels Economic Review (Cahiers economiques de Bruxelles), 53, 201–253.

Johansen, A., Sornette, D., & Ledoit, O. (1999). Predicting financial crashes using discrete scale invariance. Journal of Risk, 1, 5–32.

Koenker, R. (2005). Quantile regression. 38. Cambridge: Cambridge University Press.

Leiss, M., Nax, H. H., & Sornette, D. (2015). Super-exponential growth expectations and the global financial crisis. Journal of Economic Dynamics and Control, 55, 1–13.

Lenton, T., Livina, V., Dakos, V., Van Nes, E., & Scheffer, M. (2012). Early warning of climate tipping points from critical slowing down: Comparing methods to improve robustness. Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 370, 1185–1204.

Roubini, N., & Mihn, S. (2010). Crisis economics: A crash course in the future of finance. New York: Penguin Press.

Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., Held, H., Van Nes, E. H., Rietkerk, M., & Sugi-hara, G. (2009). Early-warning signals for critical transitions. Nature, 461, 53–59.

Scheffer, M., Carpenter, S. R., Lenton, T. M., Bascompte, J., Brock, W., Dakos, V., Van De Koppell, J., Van De Leemput, I. A., Levin, S. A., Van Nes, E. H. et al. (2012). Anticipating critical transitions. Science, 338, 344–348.

Siegel, J. J. (2003). What is an asset price bubble? An operational definition. European Financial Management, 9, 11–24.

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

Sornette, D. (2002). Predictability of catastrophic events: Material rupture, earthquakes, turbulence, financial crashes, and human birth. Proceedings of the National Academy of Sciences of the United States of America, 99, 2522–2529.

Sornette, D. (2009). Why stock markets crash: Critical events in complex financial systems. Princeton University Press.

Sornette, D., & Cauwels, P. (2014). 1980–2008: The illusion of the perpetual money machine and what it bodes for the future. Risks, 2, 103–131.

Sornette, D., & Cauwels, P. (2015a). Financial bubbles: Mechanism, diagnostic and state of the world. Review of Behavioral Economics, 2.

Sornette, D., & Cauwels, P. (2015b). Managing risk in a creepy world. Journal of Risk Management in Financial Institutions, 8, 83–108.

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

Sornette, D., Johansen, A., & Bouchaud, J.-P. (1996). Stock market crashes, precursors and replicas. Journal de Physique I, 6, 167–175.

Sornette, D., & Woodard, R. (2010). Financial bubbles, real estate bubbles, derivative bubbles, and the financial and economic crisis. Tokyo: Springer-Verlag.

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

Sornette, D., & Zhou, W.-X. (2006). Predictability of large future
changes in major financial indices. *International Journal of Forecasting*, 22, 153–168.

Zhou, W.-X., Sornette, D., Hill, R. A., & Dunbar, R. I. (2005). Discrete hierarchical organization of social group sizes. *Proceedings of the Royal Society of London B: Biological Sciences*, 272, 439–444.