Generalized q analysis of log-periodicity: Applications to critical ruptures

Wei-Xing Zhou[1] and Didier Sornette[1,2,3]
[1]*Institute of Geophysics and Planetary Physics, University of California, Los Angeles, California 90095*
[2]*Department of Earth and Space Sciences, University of California, Los Angeles, California 90095*
[3]*Laboratoire de Physique de la Matière Condensée, CNRS UMR 6622 and Université de Nice–Sophia Antipolis, 06108 Nice Cedex 2, France*

(Received 25 January 2002; revised manuscript received 11 April 2002; published 14 October 2002)

We introduce a generalization of the q analysis, which provides a nonparametric tool for the description and detection of log-periodic structures associated with discrete scale invariance. We use this generalized q analysis to construct a signature called the (H,q) derivative of discrete scale invariance, which we use to detect the log-periodicity in the cumulative energy release preceding the rupture of five pressure tanks made of composite carbon-matrix material. We investigate the significance level of the spectral Lomb periodogram of the optimal (H,q) derivative. We confirm and strengthen previous parametric results that the cumulative energy release exhibits log-periodicity before rupture. However, our tests to use this method as a scheme for the prediction of the critical value of the stress at rupture are not encouraging.

DOI: 10.1103/PhysRevE.66.046111 PACS number(s): 05.70.Jk, 81.40.Np

I. INTRODUCTION

The fracture of materials is a catastrophic phenomenon of considerable technological and scientific importance. However, a reliable identification of precursory signatures of impending failure is lacking in most cases. A notable exception has been found [1,2] in the analysis of acoustic emissions recorded during the pressurization of spherical tanks of kevlar or carbon fibers preimpregnated in a resin matrix wrapped up around a thin metallic liner (steel or titanium) fabricated and instrumented by Aérospatiale-Matra, Inc. (now EADS). A recent thorough analysis [3] of the seven acoustic emission recordings of seven pressure tanks that were brought to rupture has unambiguously characterized the acceleration of the acoustic energy rate dE/dt and found it to be in agreement with a power-law ''divergence'' expected from the critical point theory proposed in Ref. [4]. In addition, strong evidence of log-periodic corrections was found [3] that quantify the intermittent succession of accelerating bursts and quiescent phases of the acoustic emissions on the approach to rupture. Reference [3] also proposed an improved model accounting for the crossover from the noncritical to the critical region close to the rupture point exhibits interesting predictive potential. The critical rupture concept, confirmed by other experiments [5], opens the road toward industrial applications involving heterogeneous materials such as fiber composites, rocks, concrete under compression, and materials with large distributed residual stresses [6].

However, a time-to-failure behavior following a power law $dE/dt \propto (t_c - t)^{-\alpha}$ does not provide a reliable and unique signature: fits of noisy data by such power laws are notoriously unreliable; for instance, an error of 1% in the determination of t_c usually leads to errors of tens of percent for the exponent α. In addition, the determination of t_c is very sensitive to the presence of noise. In order to improve the determination of t_c, the existence of log-periodic oscillations has been found to be useful [1,7,8] and has been used for the implementation of prediction schemes [1–3] with reasonable success.

The demonstration of the existence of log-periodic corrections to the power law is of both fundamental and practical interest. From a fundamental point of view, log-periodicity signals a spontaneous hierarchical organization of damage with an approximate geometrical set of characteristic scales. A possible mechanism involves a succession of ultraviolet instabilities of the Mullins-Sekerka type [9] (see also [10] for a review). More generally, the presence of log-periodicity signals the partial breaking of continuous scale invariance into discrete scale invariance, which requires that the underlying field theory be nonunitary [11]. From a practical viewpoint, log-periodicity may help to lock in the fit on experimental data to obtain a better precision on the recovery of the critical rupture time t_c [1–3].

However, most of the evidence of log-periodicity in rupture results from parametric fits of the experimental or numerical data by a log-periodic power-law formula, except for Ref. [7], which introduced a ''canonical'' averaging method to extract the log-periodic signal directly. Parametric fits suffer from two problems: (i) the formula cannot avoid some simplification that, for instance, omits the presence of harmonics and/or other structures; (ii) parametric fits are delicate due to possible degeneracies and, in addition, their statistical significance (i.e., added value) is difficult to estimate. It is thus important to develop further nonparametric tests. Our goal here is to present a nonparametric method that will turn out to be very powerful in identifying log-periodicity in noisy data.

Our method is based on the concept of a q derivative, the inverse of the Jackson q integral [12], which is the natural tool [13,14] for describing discrete scale invariance. Indeed, q derivatives can be identified with the generators of fractal and multifractal sets with discrete dilatation symmetries [14]. Nowhere differentiable functions that characterize fractal or multifractal sets turn out to be perfectly well behaved under the q derivative. Discrete renormalization-group equations, whose general mathematical solutions are power laws with complex exponents (and hence exhibit log-periodicity), can be seen as merely Jackson q integrals of regular functions of

the decimated degrees of freedom. Jackson q integrals constitute the natural generalization of regular integrals for discretely self-similar systems [14]. The way that the Jackson q integral can be related to the free energy of a spin system on a hierarchical lattice was explained in [13].

In Sec. II, we introduce the q derivative and generalize it to take into account anomalous scaling. We discuss the main properties of the generalized q derivative that will be useful for our analysis of this paper. Section III presents our analysis with the generalized q derivative of the acoustic emission data, for the cumulative energy release, obtained during the pressurization of spherical tanks of kevlar or carbon fibers preimpregnated in a resin matrix wrapped up around a thin metallic liner (steel or titanium). For comparison, we use exactly the same set of seven acoustic emission recordings of the seven pressure tanks used in the previous study reported in [3]. Our new results confirm strongly the existence of log-periodicity with a much enhanced confidence. Section IV tests a scheme using the generalized q derivative for prediction purposes. Here, we find disappointing results: the parametric approach in [3] turns out to be more powerful. Section V concludes.

II. THE GENERALIZED q ANALYSIS AND LOG-PERIODICITY

A. Definition

Let us take some $q \in (0,1) \cup (1,\infty)$. The q derivative of an arbitrary function $f(x)$ is defined as

$$D_q f(x) = \frac{f(x) - f(qx)}{(1-q)x}. \tag{1}$$

For $q \to 1$, the definition (1) recovers the usual definition of a derivative.

For $q \neq 1$, $D_q f(x)$ is more than just a derivative: it compares the relative variations of $f(x)$ and of x when x is magnified by the finite factor q. It is thus intuitive that the q derivative tests the scale invariance property of the function $f(x)$. As we said above, it was actually shown by Erzan and Eckmann [14] that the q derivative is the natural tool for describing discrete scale invariance, since a fixed finite q compares $f(x)$ with $f(qx)$ at x magnified by a fixed factor, and thus it also compares $f(qx)$ with $f(q^2x)$, $f(q^2x)$ with $f(q^3x)$, etc. When x is taken as the distance from a critical point, $D_q f(x)$ thus quantifies the discrete self-similarity of the function $f(x)$ in the vicinity of the critical point. From the definition (1), it is clear that

$$D_{1/q} f(x) = D_q f(x/q). \tag{2}$$

It is thus enough to study $D_q f(x)$ for $q \in (0,1)$ to derive its values for all q's.

The necessary and sufficient condition for a function $f(x)$ to be homogeneous with the order ψ is

$$D_q f(x) = \frac{q^\psi - 1}{q-1} \frac{f(x)}{x}. \tag{3}$$

This expression suggests the introduction of a generalized q derivative that we call the (H,q) derivative, such that the dependence in x of $D_q^H f(x)$ disappears for homogeneous functions, for the choice $H = \psi$. Consider, therefore, the following definition:

$$D_q^H f(x) \overset{\triangle}{=} \frac{f(x) - f(qx)}{[(1-q)x]^H}, \tag{4}$$

such that $D_q^{H=1} f(x)$ recovers the standard q derivative $D_q f(x)$. For a power-law function $f(x) = Bx^m$, $D_q^{H=m}[Bx^m] = B(1-q^m)/(1-q)^m$ is constant. For a statistically homogeneous function $f(x) \overset{d}{=} Bx^m$, $D_q^{H=m} f(x) \overset{d}{=}$ const, where the symbol $\overset{d}{=}$ means equality in distribution.

The generalized (H,q) derivative has two control parameters: the discrete scale factor q devised to characterize the log-periodic structure and the exponent H introduced to account for a possible power-law dependence, i.e., to correct for the existence of trends in log-log plots.

B. Application to log-periodic functions

Since Erzan and Eckmann [14] showed that the q derivative is the natural tool for describing discrete scale invariance, it is natural to study the properties of the (H,q) derivative of the simplest function exhibiting discrete scale invariance, namely a power law decorated by a simple log-periodic function,

$$y(x) = A - B\tau^m x^m + C\tau^m x^m \cos(\omega \ln x), \tag{5}$$

where the presence of a phase $\phi = \omega \ln \tau$ has been absorbed in the definition of x, $0 < C < B$, $\omega = 2\pi f$ is the angular log-frequency, and f is the log-frequency (not to be confused with the previous function f used in the previous subsection). This equation, where x is interpreted as the normalized distance $x = (p_c - p)/\tau$ to a critical point p_c, has been used in several works to describe material rupture [1,2,8,3], precursory patterns of large earthquakes [15], rock bursts [16], aftershocks [17,18], and speculative bubbles preceding financial crashes [19].

The (H,q) derivative of $y(x)$ is

$$D_q^H y(x) = x^{m-H}[B' + C' g(x)], \tag{6}$$

where

$$B' = -\frac{B\tau^m(1-q^m)}{(1-q)^H}, \quad C' = \frac{C\tau^m}{(1-q)^H}, \tag{7}$$

and

$$g(x) = C_1 \cos(\omega \ln x) + C_2 \sin(\omega \ln x), \tag{8}$$

with $C_1 = 1 - q^m \cos(\omega \ln q) > 0$ and $C_2 = \sin(\omega \ln q)$. The special choice $H = m$ gets rid of the power law and the (H,q) derivative $D_q^H y(x)$ is, up to an additive constant, proportional to the pure log-periodic function $g(x)$.

For the special choices $q=e^{-n2\pi/\omega}$, where n is a positive integer corresponding to choosing q equal to one of the preferred scaling factors of the log-periodic function, we obtain $C_1=1-q^m$ and $C_2=0$. Therefore, the (H,q) derivative offers a approach for detecting the preferred scaling factors of the discretely scale invariant function by a measure of its phase: those values of q such that the phases of the $(H=m,q)$ derivative are that of a pure cosine should qualify as the preferred scaling factors. Such phases, for instance, can be measured by the Hilbert transform. Here, we do not pursue this possibility, which will be explored in another presentation.

It is easy to show that $g(x)$ is extremal at x_m solution of

$$\omega \ln(x_m)=n\pi+\arctan(C_2/C_1), \tag{9}$$

where $n=0,\pm1,\pm2,\ldots$, and that the extreme values of $g(x)$ are

$$g(x_m)=\pm\sqrt{C_1^2+C_2^2}. \tag{10}$$

The amplitude of $D_q^H y(x)$ is then

$$A=x_m^{m-H}C'\sqrt{C_1^2+C_2^2}, \tag{11}$$

while the successive extreme values are

$$D_m=(B'\pm C'\sqrt{C_1^2+C_2^2})(x_m)^{m-H}, \tag{12}$$

where B' and C' are defined in Eqs. (7). By fixing H close to m, one can in principle obtain the amplitude A_m to be constant as a function of x. This value $H=m$ should provide theoretically the most significant log-periodic component, quantified, for instance, by the largest peak of the Lomb periodogram. However, in practice, the noise embedded in the data may distort the log-periodic oscillations and the most significant log-periodicity may occur for $H\neq m$. The introduction of H can then be seen more generally as affording a convenient detrending scheme.

Figure 1 shows $y(x)$ defined in Eq. (5) as a function of the distance x to the critical point with $A=1260$, $B=300$, $C=6$, $m=0.3$, $\omega=5.4$, and $\phi=0$. Figure 2 shows its generalized q derivative with $q=0.5$ for $H=0.2$, $H=0.3$, and $H=0.4$, respectively. This generalized q derivative has been calculated by using the incorrect assumption that the critical point is at $x=1$ in order to also show the distortion resulting from an error in the determination of the critical point. This distortion becomes important when x is not large compared to 1. We observe that the amplitude of the oscillations of $D_q^H y(x)$ increases with H when going towards the critical point $x=0$, in agreement with the prediction (11).

In the next section, using the (H,q) analysis, we test for the presence of log-periodicity in the cumulative energy release obtained in the experimental recordings of the seven pressure tanks used in the previous study reported in [3].

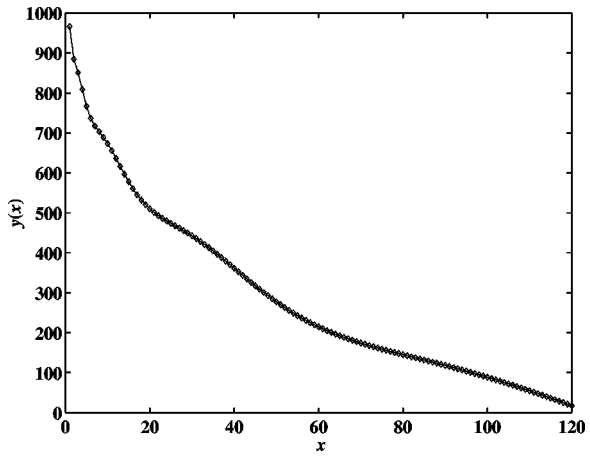

FIG. 1. Plot of $y(x)$ defined by Eq. (5) as a function of the pressure-to-rupture x with $A=1260$, $B=300$, $C=6$, $m=0.3$, $\omega=5.4$, and $\phi=0$. We generated 120 evenly spaced data points with x between 1 and 120, to mimic the cumulative energy release of a real acoustic emission experiment. The variables x and y are dimensionless.

III. INVESTIGATION OF THE SIGNIFICANCE OF LOG-PERIODICITY IN ACOUSTIC EMISSION ENERGY RELEASE USING THE (H,q) ANALYSIS

A. Data sets

In order to illustrate our new method of (H,q) analysis for the description of log-periodic oscillations in complex systems, we apply it to critical ruptures. We revisit the data of acoustic emissions recorded during the pressurization of spherical tanks of kevlar or carbon fibers preimpregnated in a resin matrix wrapped up around a thin metallic liner (steel or titanium) fabricated and instrumented by Aérospatiale-Matra, Inc. (now EADS). We use the same notations to label the data sets as used in Ref. [3]. The true value p_c of the pressure

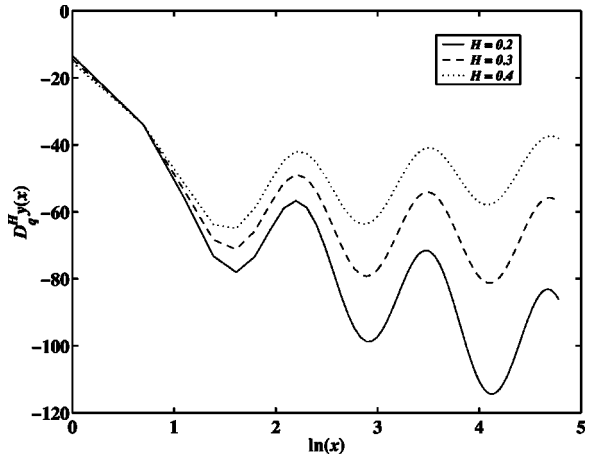

FIG. 2. The generalized q derivative of $y(x)$ shown in Fig. 1 with $q=0.5$ for $H=0.2$, $H=0.3$, and $H=0.4$, respectively. The calculation assumes a critical point at $x_c=1$ while the synthetic function has its genuine critical point at $x_c=0$. This incorrect value of the critical point in the calculation of the generalized q derivative is responsible for the distortions observed for $\ln(x)<1$. The generalized q derivative and x are dimensionless.

at which rupture occurred is used in order to define the pressure-to-failure $p_c - p$ quantifying the distance to the critical rupture point. Note that all pressures are expressed in units of bars.

In our analysis, we exclude data sets of pressure tanks No. 5 and No. 7, because for these two data sets, $p_{\rm last} \ll p_c$ (the last data point recorded is for a pressure far below the critical rupture value). This leads to few oscillations and low statistical significance. For instance, data set No. 5 has $p_c = 797$, $p_{\rm last} = 713$, and $p_{\rm min} = 110.5$ so that the number of log-periodic oscillations is $f \ln(p_c - p_{\rm min})/\ln(p_c - p_{\rm last}) \approx 3$, where we made no truncation on the low-pressure end and used a large frequency $f = 1.42$ fitted in Ref. [3] (see Table 4 in that reference) instead of the average frequency $f = 0.86$. The data set No. 7 presents even fewer oscillations. This leads to insufficient statistical significance and we are thus left to analyze five data sets: No. 1, No. 2, No. 3, No. 4, and No. 6.

We follow [3] and use the acoustic emission data recorded for pressures sufficiently high so that a noticeable acceleration in the cumulative energy release takes place. This choice is not crucial at all, since the (H,q) analysis is not sensitive to the points far from the critical point. We also exclude the last six points nearest to p_c because they contain the largest noise and may suffer from finite-size effects that may lead to serious distortions. The data were recorded with an approximately evenly spaced sampling, so that the sampling frequency close to the critical point $(p \to p_c)$ becomes comparable to or even smaller than the underlying log-frequency in the $\ln(p_c - p)$ axis. Again, this choice is not crucial and tests performed with different numbers of points removed give similar results.

B. Detection of log-periodicity

In the analysis, q ranges from 0.1 to 0.9 with spacing 0.1, while H takes values from -0.9 to 0.9 with spacing 0.1. This defines $9 \times 19 = 171$ parameter pairs (H,q). For each parameter pair (H,q), we calculate the generalized q derivative of the cumulative energy release. We then perform a spectral analysis of the generalized q derivative using the Lomb periodogram method [20], in order to test for the statistical significance of possible log-periodic oscillations. For each (H,q) pair, the highest peak $P_N(H,q)$ and its associated angular log-frequency $\omega(H,q)$ in the Lomb periodogram are obtained. The basic criterion used to identify a log-periodic signal is the strength of the Lomb periodogram analysis, i.e., the height of the spectral peaks.

We shall not present all results for all five data sets because they are completely similar. A description of the typical results obtained for data set No. 1 and presented in Figs. 3–6 will be sufficient. Figure 4 shows the dependence of the highest peak $P_N(H,q)$ in each Lomb periodogram as a function of H and q, while Fig. 3 gives the associated angular log-frequencies $\omega(H,q)$. In Fig. 4, the pairs (H,q) near $(-0.2, 0.1)$ give the largest $P_N = 56.3$, which should imply the most significant log-periodic oscillations in the generalized q derivative. However, the associated $\omega \approx 1.5$ in the region **B** of Fig. 3 is dangerously too low as it corresponds to only one oscillation in the signal. It is obvious that the spectral analysis cannot distinguish in this case any underlying oscillations from a global trend of the (H,q) derivative. There is an additional caveat: the log-frequencies in the region **B** are close to the most probable (log-periodic) angular frequency $\omega^{\rm mp} = 1.8$, resulting solely from the most probable noise. It was indeed shown [17] that noise-decorating power laws may lead to artifactual log-periodicity with a most probable frequency corresponding roughly to 1.5 oscillations over the whole range of analysis. In the present context, this defines the most probable (log-periodic) angular frequency $\omega^{\rm mp}$ by the following formula:

$$\omega^{\rm mp} = \frac{2\pi \times 1.5}{\ln(p_c - p_{\rm min}) - \ln(p_c - p_{\rm max})}, \tag{13}$$

where the acoustic emission signal is recorded from the pressure $p_{\rm min}$ to $p_{\rm max}$. In the case of experiment No. 1, this leads to $\omega^{\rm mp} = 1.8$. Thus, a value of $\omega \leqslant \omega^{\rm mf}$ corresponds to at most 1.5 oscillations in the plot of the generalized q derivative as

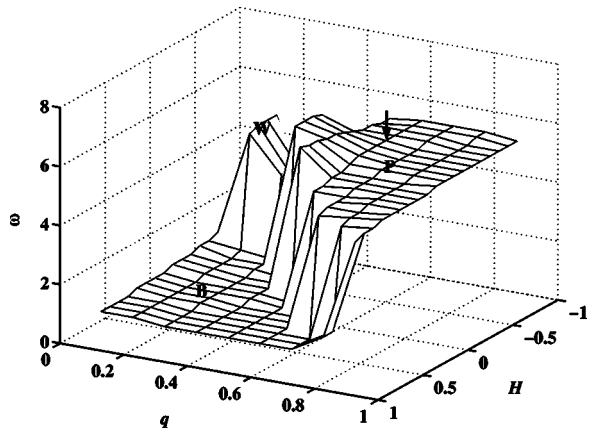

FIG. 3. Dependence of the angular log-frequency $\omega(H,q)$ of the most significant peak in each Lomb periodogram of the (H,q) derivative of the cumulative energy release before the rupture of tank No. 1. The wedge **W** and the bottom **B** are excluded by the criteria discussed in the text. The optimal pair $(-0.5, 0.6)$ is indicated by an arrow in the platform **P**. The variables are dimensionless.

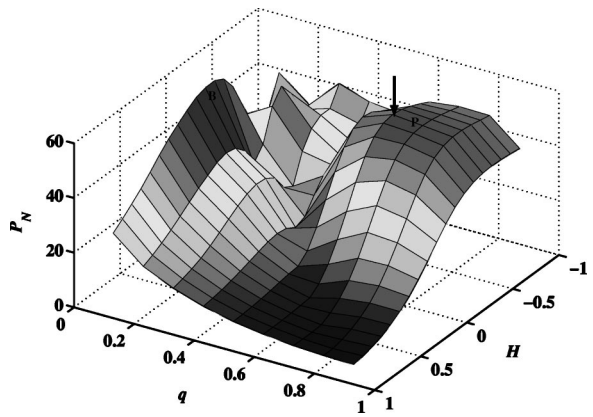

FIG. 4. Dependence of the height $P_N(H,q)$ of the most significant peak in each Lomb periodogram of the (H,q) derivative of the cumulative energy release before the rupture of tank No. 1. The variables are dimensionless.

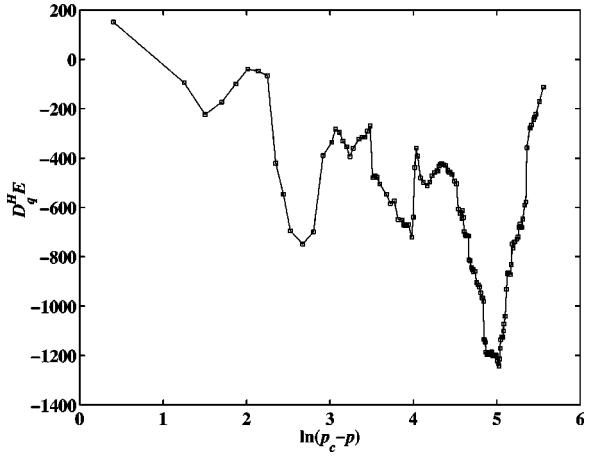

FIG. 5. (H,q) derivative of the cumulative energy release before the rupture of tank No. 1 as a function of the logarithm of the pressure-to-rupture p_c-p with $q=0.6$ and $H=-0.5$. The variables are dimensionless.

a function of $\ln(p_c-p)$ for the whole range of pressure. This is insufficient to be able to conclude about the existence of log-periodicity with good statistical confidence. To qualify the existence of log-periodicity, pairs of (H,q) with $\omega(H,q)\leqslant\omega^{\mathrm{mp}}$ should be discarded, because, for these values, there is a non-negligible probability that the observed log-periodicity may result from noisy fluctuations around the power law and is thus spurious [17].

A physically meaningful region of (H,q) is thus obtained, labeled by **P** (platform) and **W** (wedge) in Figs. 3 and 4. Among these pairs, the optimal pair $(H,q)=(-0.5,0.6)$ corresponds to the highest Lomb peak $P_N=51.7$, indicated by arrows in Fig. 3 and Fig. 4. It lies within the platform **P** shown in Fig. 3. Figure 5 gives the generalized q derivative for this optimal pair $(H,q)=(-0.5,0.6)$, while Fig. 6 plots its corresponding Lomb periodogram. The log-periodic oscillations are found to be very significant. Ideally, we should estimate the probability that random noise, of several plausible standard distributions, creates a false alarm that a periodicity (or log-periodicity) is found in the (H,q) derivative of the signal. This has been done in a systematic way for a large variety of noises, without and with long-range correlation [21]. However, it is difficult to identify what should be the correct null hypothesis of the noise decorating the generalized q derivative. Under the null hypothesis of independent Gaussian noise decorating the signal, the false-alarm probability of the log-periodic component is $\approx 5\times 10^{-21}$. The false-alarm probability for the null hypothesis of independent heavy-tailed noises (say, Lévy stable noise and power-law noise) and weakly correlated noises [say, GARCH(1,1) noise and fractional Gaussian noise (fGn) with the Hurst index less than 0.5] is even lower [21]. If we assume a strongly correlated noise such as fractional Gaussian noise with the Hurst index greater than 0.5, the false-alarm probability increases: for instance, a fractional Gaussian noise with a Hurst index of 0.98 leads to a false-alarm probability of 1% [21].

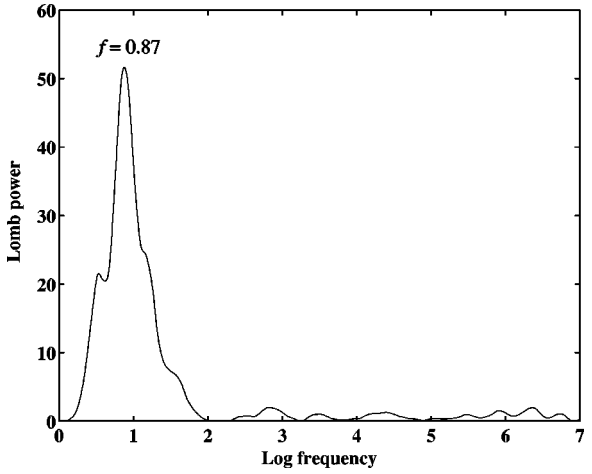

FIG. 6. Lomb power of the (H,q) derivative shown in Fig. 5. The variables are dimensionless.

It is worth noting that the majority of the pairs (H,q) in **P** have similar oscillatory behavior (i.e., similar ω) indicating a robust log-periodic signal. Moreover, we note that a log-periodic angular frequency which is too small may result from noise, as we just said. Similarly, a too high log-periodic angular frequency is also the signature of noise, simply because it is easy to fit noisy data with many oscillations. We should thus also discard the pairs of (H,q) whose ω are too large. Reference [3] used this criterion to discard solutions with $\omega\geqslant 14$. It is not always obvious to determine the value of the angular frequency threshold beyond which solutions should be rejected as noise. Fortunately, the Lomb peak P_N corresponding to a large angular frequency stemming from noise is usually low [21], as we observed in experiments No. 4 and No. 6 (not shown).

We summarize in Table I the results of the (H,q) analysis on the cumulative energy release of data sets No. 1, No. 2, No. 3, No. 4, and No. 6. For each data set, the generalized q analysis was performed between $(p_{\min},p_{\max})$, with a number of points varying between 55 and 164 as indicated in the column "N" in the table. p_c is the true critical pressure at rupture. P_N is the optimal value of the Lomb peak height, and ω (and ω' if it exists) is the corresponding angular log-frequency of the highest peak (respectively the second highest peak). The column "Gaussian" presents the false-alarm probability of the Lomb peak under the null hypothesis of independent Gaussian noise. The column "fGn" evaluates the value of the Hurst index of a fractional Brownian noise which would give a false-alarm probability 1% to get the same peak as in our analysis. The larger this number is, the more significant is the Lomb peak because a high Hurst index is very improbable as it implies an unreasonable long-range dependence and persistence of the noise [21].

These two columns "Gaussian" and "fGn" thus give us a sense of the amplitude of the signal over noise ratio, from the point of view of the statistical significance of the peaks found in the analysis. The conclusion derived from reading these two columns is that it is highly improbable that the obtained peaks result from a Gaussian noise and one would need a very strong persistence in a fractional Brownian noise to simulate the obtained signatures. This holds regardless of

TABLE I. Summary of the results of the (H,q) analysis on the cumulative energy release of data sets No. 1, No. 2, No. 3, No. 4, and No. 6. For each data set, the generalized q analysis was performed between $(p_{\min}, p_{\max})$ with a number of points given in the column N. p_c is the true critical pressure of rupture. The column (H,q) lists the optimal pairs. P_N is the optimal value of the Lomb peak height. ω and ω' are the corresponding angular log-frequencies of the highest and second highest peak. $N_{\rm osc}$ is the number of oscillations. The column ''Gaussian'' presents the false-alarm probability of the Lomb peak under the null hypothesis of independent Gaussian noise. The column ''fGn'' gives the Hurst index of a fractional Gaussian noise that would give a false-alarm probability 1% to obtain the same Lomb peak as in the signal. Pressures are expressed in units of bars.

Tank	N	$p_{\min}$	$p_{\max}$	p_c	(H,q)	P_N	ω	ω'	$\omega^{\rm mp}$	$N_{\rm osc}$	Gaussian	fGn
No. 1	144	453.5	711.5	713	$(-0.5, 0.6)$	51.7	5.5	/	1.8	4.6	5E-21	0.98
No. 2	138	467.5	661.5	673	$(0.3, 0.2)$	34.4	10.0	5.3	3.3	4.5	2E-13	0.93
No. 3	63	671.5	756.5	764	$(-0.2, 0.4)$	14.5	4.5	9.9	3.8	1.8	5E-5	0.70
No. 4	56	671.5	746.5	756	$(0, 0.6)$	19.0	5.5	17.5	4.3	1.9	3E-7	0.79
No. 6	72	614.5	723.5	734	$(-0.3, 0.8)$	15.7	12.7	5.4	3.9	4.9	1E-5	0.71

the rather small size of the data sets. Since the number of data points is limited, it is important to quantify as much as possible the amplitude of the signal compared to that of the noise. To quantify the impact of noise on the detectability of log-periodicity in signals with short events series, we now estimate the signal-to-noise ratio γ. For this, we assume that the noise is independent of the underlying log-periodic components. Since the events are not badly bunched, following [22], we have

$$P_N = \frac{N}{2}\left(1 + \frac{2\sigma^2}{A^2}\right)^{-1}, \tag{14}$$

where σ is the standard deviation of the noise and A is the amplitude of the signal. The signal-to-noise ratio defined as $\gamma = A/\sigma$ can thus be estimated by

$$\gamma = \left(\frac{4P_N}{N - 2P_N}\right)^{1/2}, \tag{15}$$

where N is the number of events (data points in one experiment). Using this formula, we obtained $\gamma = 2.44$, 1.48, 1.44, 2.52, and 1.35 for the five tanks, respectively. Note that Eq. (15) holds approximately for all types of noises, not only for Gaussian noise [22,21]. These different tests confirm that we are extracting meaningful information, above an (admittedly) large noise level.

We find that all data sets taken together can be described by a fundamental value of the angular log-frequency $\omega = 5.4 \pm 0.5$. To see this, we construct the average Lomb power spectrum over the five Lomb power spectra for the five experiments, which is shown in Fig. 7. The highest peak lies at $f = 0.87$, i.e., at $\omega = 2\pi f = 5.4$. We cannot estimate quantitatively the false-alarm probability of this average Lomb peak since the different data sets have different sizes. We observe that this log-frequency $f = 0.87$ (or $\omega = 5.4$) is found to be very close to either the log-frequency of the highest or of the second highest peak of each individual spectrum of the five data sets. This value $f = 0.87$ corresponds to a preferred scaling ratio $\lambda = e^{1/f} = 3.2 \pm 0.3$.

The log-frequency $f_2 = 1.60$ (or $\omega_2 = 10.1$) of the second highest peak of the average Lomb spectrum can be interpreted as the first harmonic of the fundamental log-frequency $f = 0.87$. This log-frequency corresponds to the highest peak of data set No. 2 and to the second highest peak of data set No. 3. It is not unexpected to find in nonlinear systems that the first harmonic may be a stronger power spectrum than the fundamental frequency. The second highest peak of data set No. 4 can be interpreted as the second harmonics of the fundamental log-frequency. The highest peak of data set No. 6 occurs for a log-frequency almost exactly in the middle range between the first and second harmonics and is thus less credible.

These results taken together give a reliable indication that there is a genuine log-periodicity of the energy release rate before rupture in data sets No. 1, No. 2, No. 3, No. 4, and No. 6. This confirms and strengthens the claim of Ref. [3] that a pure power law fails to parametrize the data but a log-periodic formula does a much better job.

IV. IN-SAMPLE PREDICTION OF RUPTURES

The (H,q) analysis has shown its power for detecting log-periodicity, conditioned on our knowledge of the critical

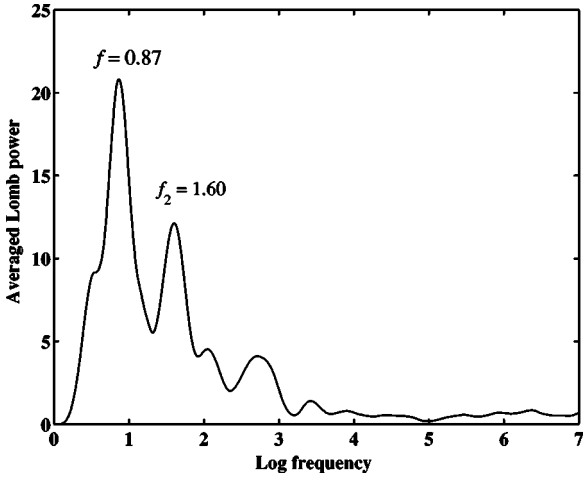

FIG. 7. Averaged Lomb power of the individual Lomb powers of all the five data sets. The highest and the second highest peaks lie at $f = 0.87$ and $f_2 = 1.60$. The variables are dimensionless.

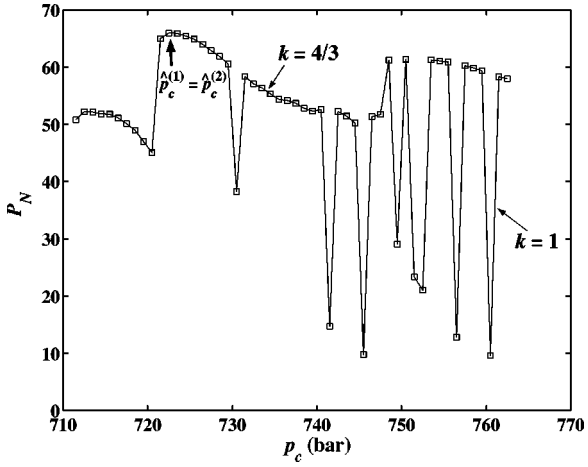

FIG. 8. In-sample prediction of the critical pressure of the rupture of tank No. 1. The optimal Lomb peak height P_N is shown as a function of the presumed critical pressure p_c for experiment No. 1 and for $p_{\max}=705.5$. The highest peak is found for $p_c=\hat{p}_c^{(1)}=722.5$, which is the predicted p_c, to be compared with the true value 713. The two fine arrows indicate the upper threshold of predictable $\hat{p}_c$ estimated by Eq. (18) with $k=4/3$ and $k=1$. The coarse arrow indicates the predicted critical pressures $\hat{p}_c^{(1)}$ and $\hat{p}_c^{(2)}$. The ordinate variable is dimensionless.

pressure p_c at rupture. It is natural to ask whether it can be extended to provide advanced prediction of p_c. To carry out these tests, we use the cumulative energy release which provided the strongest log-periodic signal in the analysis reported in the previous sections.

Our strategy is to use each data set up to a maximum pressure $p_{\max}<p_c$, assume some value for p_c, and perform the same analysis as in Sec. III. For a given presumed critical point p_c, we determine the optimal pair (H,q). For each presumed p_c, we thus obtain the Lomb peak height $P_N(H,q)$ and its associated angular log-frequency $\omega(H,q)$ as a function of H and q. In order to find the optimal (H,q) corresponding to the maximum $P_N(H,q)$, we add the criteria discussed in Sec. III B [see formula (13)] to exclude those (H,q) with $\omega(H,q)\leqslant\omega^{\mathrm{mp}}$. Having determined the optimal $P_N(p_c)$ and $\omega(p_c)$ as functions of the presumed critical pressure p_c, we determine the predicted critical pressure $\hat{p}_c$ by the condition

$$P_N(\hat{p}_c)=\max_{p_c}\{P_N(p_c)\}, \tag{16}$$

without further constraints on $\omega(p_c)$.

In practice, we analyze the function $P_N(p_c;H,q)$ of the three variables p_c, H, and q in the following sequence: we fix a value for p_c and explore the plane (H,q). We then change p_c and redo the exploration of the plane (H,q), and so on. Figure 8 gives the optimal Lomb peak height P_N as a function of the presumed critical pressure p_c for experiment No. 1 and for $p_{\max}=705.5$. The highest peak is found for $p_c=\hat{p}_c^{(1)}=722.5$, which is the predicted p_c, to be compared with the true value 713.

This result is encouraging and we now attempt to improve this prediction skill by requesting that the predicted p_c should not be too far from the last point, i.e., it is nonsensical to predict too far from the "present." To implement this idea, we modify formula (13) and impose $\omega>k\omega^{\mathrm{mp}}$, namely,

$$\frac{\omega}{2\pi}>\frac{1.5k}{\ln(p_c-p_{\min})-\ln(p_c-p_{\max})}, \tag{17}$$

where $k\geqslant 1$ is a "safety factor." This constraint translates into the following condition for p_c:

$$p_c<p_{\max}+\frac{p_{\max}-p_{\min}}{e^{3k\pi/\omega}-1}. \tag{18}$$

This constraint (18) means that there exists an upper threshold beyond which we cannot make a physically meaningful prediction. Since the left-hand side of (18) is monotonically increasing with ω, it is possible to make a prediction much earlier before the critical point, the larger ω is. This is natural since large ω implies more log-periodic oscillations and thus a stronger log-periodic signal. To implement this condition in practice, we take $\omega=5.4$ as the central value of the log-periodic angular frequency. According to this constraint (18), ruptures of tanks No. 5 and No. 7 are unpredictable, since the last point p_{last} is too far from the true critical pressure.

The results are given in Table II. The columns $\hat{p}_c^{(1)}$ and $\hat{p}_c^{(2)}$ list the predicted critical pressure $\hat{p}_c$ and its corresponding upper threshold in the parentheses with the constraint (18) for the two choices $k=4/3$ and $k=1$, respectively. The prediction for experiment No. 1 is very good. For the other

TABLE II. In-sample prediction using the (H,q) analysis on the cumulative energy release of tanks No. 1, No. 2, No. 3, No. 4, and No. 6. Columns $\hat{p}_c^{(1)}$ and $\hat{p}_c^{(2)}$ list the predicted critical pressure $\hat{p}_c$ and its corresponding upper threshold in the parentheses with the constraint (18) for $k=4/3$ and $k=1$, respectively. The meaning of the other columns is the same as in Table I. Pressures are expressed in units of bars.

Tank	Points	$p_{\min}$	$p_{\max}$	ω	p_c	$\frac{p_c-p_{\max}}{p_c}$	$\hat{p}_c^{(1)}$	$\hat{p}_c^{(2)}$
No. 1	139	453.5	705.5	5.5	713	1.1%	722.5(734)	722.5(761)
No. 2	149	452.5	662.5	5	673	1.6%	680.5(681)	686.5(700)
No. 3	65	671.5	759.5	4.5	764	0.6%	763.5(765)	767.5(772)
No. 4	56	671.5	746.5	5.5	756	1.3%	755.5(755)	755.5(763)
No. 6	74	614.5	726.5	6.4	734	1.0%	742.5(744)	758.5(759)

experiments, $\hat{p}_c$ is found to be very close to the upper threshold. The predictions for experiments No. 2, No. 3, and No. 4 are reasonable while the prediction for case No. 6 fails completely.

While these results seem rather good, they deteriorate extremely fast as $p_{\max}$ is decreased by a few tens of bars.

V. CONCLUDING REMARKS

In this paper, we have introduced a nonparametric tool for the detection of log-periodicity in complex systems. We have applied this method to analyze the cumulative energy release during the period approaching critical ruptures. We have confirmed and strengthened previous parametric results [3] that the cumulative energy release exhibit log-periodicity before rupture. We remark that the data of the energy release rate should have the same log-periodic structure.

The nonparametric method in its present implementation does not seem reliable for the prediction of the critical pressure at rupture. However, it can be a useful complement and confirmation to the parametric method, consisting in fitting the data to a log-periodic function with a power-law envelope. Rather than analyzing the statistical significance of the residues of such a fit with respect to the presence of log-periodicity, we can use the fitted value of the critical pressure p_c at rupture to perform an (H,q) analysis and extract ω and its associated Lomb power spectrum in order to qualify its statistical significance. This could be used to confirm/deny the quality of the parametric fit.

ACKNOWLEDGMENTS

We are grateful to A. Erzan for a discussion on Jackson's integral and for supplying the corresponding references and to A. Johansen for helpful comments. This work was supported by National Science Foundation Grant No. NSF-DMR99-71475 and the James S. McDonnell Foundation.

[1] J.-C. Anifrani, C. Le Floc'h, D. Sornette, and B. Souillard, J. Phys. I **5**, 631 (1995).
[2] J.-C. Anifrani, C. Le Floc'h, and D. Sornette, Contrôle Industriel **220**, 43 (1999).
[3] A. Johansen and D. Sornette, Eur. Phys. J. B **18**, 163 (2000).
[4] D. Sornette and C. Vanneste, Phys. Rev. Lett. **68**, 612 (1992); C. Vanneste and D. Sornette, J. Phys. I **2**, 1621 (1992); D. Sornette, C. Vanneste, and L. Knopoff, Phys. Rev. A **45**, 8351 (1992).
[5] A. Garcimartin, A. Guarino, L. Bellon, and S. Ciliberto, Phys. Rev. Lett. **79**, 3202 (1997); A. Guarino, A. Garcimartin, and S. Ciliberto, Eur. Phys. J. B **6**, 13 (1998); A. Guarino, S. Ciliberto, and A. Garcimartin, Europhys. Lett. **47**, 456 (1999).
[6] J.V. Andersen, D. Sornette, and K.-T. Leung, Phys. Rev. Lett. **78**, 2140 (1997); D. Sornette, K.-T. Leung, and J.V. Andersen, *ibid.* **80**, 3158 (1998).
[7] A. Johansen and D. Sornette, Int. J. Mod. Phys. C **9**, 433 (1998).
[8] M. Sahimi and S. Arbabi, Phys. Rev. Lett. **77**, 3689 (1996).
[9] Y. Huang, G. Ouillon, H. Saleur, and D. Sornette, Phys. Rev. E **55**, 6433 (1997).
[10] D. Sornette, Phys. Rep. **297**, 239 (1998).
[11] H. Saleur and D. Sornette, J. Phys. I **6**, 327 (1996).
[12] F.H. Jackson, Q. J. Pure. Appl. Math **41**, 193 (1910); Q. J. Math Oxford Ser. **2**, 1 (1951).
[13] A. Erzan, Phys. Lett. A **225**, 235 (1997).
[14] A. Erzan and J.-P. Eckmann, Phys. Rev. Lett. **78**, 3245 (1997).
[15] D. Sornette and C.G. Sammis, J. Phys. I **5**, 607 (1995); H. Saleur, C.G. Sammis, and D. Sornette, J. Geophys. Res. [Solid Earth] **101**, 17661 (1996); A. Johansen, D. Sornette, H. Wakita, U. Tsunogai, W.I. Newman, and H. Saleur, J. Phys. I **6**, 1391 (1996); A. Johansen, H. Saleur, and D. Sornette, Eur. Phys. J. B **15**, 551 (2000); Y. Huang, H. Saleur, and D. Sornette, J. Geophys. Res. [Solid Earth] **105**, 28111 (2000).
[16] G. Ouillon and D. Sornette, Geophys. J. Int. **143**, 454 (2000).
[17] Y. Huang, A. Johansen, M.W. Lee, H. Saleur, and D. Sornette, J. Geophys. Res. [Solid Earth] **105**, 25451 (2000).
[18] M.W. Lee and D. Sornette, Eur. Phys. J. B **15**, 193 (2000).
[19] D. Sornette, A. Johansen, and J.-P. Bouchaud, J. Phys. I **6**, 167 (1996); D. Sornette and A. Johansen, Physica A **245**, 411 (1997); A. Johansen and D. Sornette, Risk **12**, 91 (1999); Eur. Phys. J. B **9**, 167 (1999); A. Johansen, D. Sornette, and O. Ledoit, J. Risk **1**, 5 (1999); A. Johansen, O. Ledoit, and D. Sornette, Int. J. Theor. Appl. Finance **3**, 219 (2000); A. Johansen and D. Sornette, Eur. Phys. J. B **17**, 319 (2000); A. Johansen and D. Sornette, Int. J. Theor. Appl. Finance **4**, 853 (2001); D. Sornette and A. Johansen, Quantitative Finance **1**, 452 (2001).
[20] W.H. Press, S.A. Teulolsky, W.T. Vetterlong, and B.P. Flannery, *Numerical Recipes in Fortran* (Cambridge University Press, Cambridge, 1994).
[21] W.-X. Zhou and D. Sornette, Int. J. Mod. Phys. C **13**, 137 (2002).
[22] J.H. Horne and S.L. Baliunas, Astrophys. J. **302**, 757 (1986).