Physica A 392 (2013) 2442–2446

Contents lists available at SciVerse ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Discussion

Comment on 'multifractal diffusion entropy analysis on stock volatility in financial markets' [Physica A 391 (2012) 5739–5745]

A. Yu. Morozov

Department of Mathematics, University of Leicester, LE1 7RH, UK

ARTICLE INFO

Article history:
Received 12 September 2012
Received in revised form 29 November 2012
Available online 23 January 2013

Keywords:
Diffusion entropy analysis
Multifractal time series
Rényi entropy
Volatility
Anomalous diffusion

ABSTRACT

In their recent article 'multifractal diffusion entropy analysis on stock volatility in financial markets' Huang, Shang and Zhao (2012) [6] suggested a generalization of the diffusion entropy analysis method with the main goal of being able to reveal scaling exponents for multifractal times series. The main idea seems to be replacing the Shannon entropy by the Rényi entropy, which is a one-parametric family of entropies. The authors claim that based on their method they are able to separate long range and short correlations of financial market multifractal time series. In this comment I show that the suggested new method does not bring much valuable information in obtaining the correct scaling for a multifractal/mono-fractal process beyond the original diffusion entropy analysis method. I also argue that the mathematical properties of the multifractal diffusion entropy analysis should be carefully explored to avoid possible numerical artefacts when implementing the method in analysis of real sequences of data.

1. Introduction

Revealing scale invariant properties of complex time series is an important area of research with hundreds of publications emerging every year. In 2002, N. Scafetta and colleagues have developed an interesting method of determining the scaling exponent of stochastic process based on computation of the diffusion entropy. An important particularity of the diffusion entropy analysis (DEA) method is that it allows us to obtain the correct scaling even in the case of strange kinetics, where the second moment is diverging [1,2]. On the other hand, it is well known, that in many cases the real-world time series show a multifractal scaling behaviour [3–5], which cannot be described by a single scaling exponent. Recently, Huang and colleagues [6] have elaborated a technique based on extension of the initial DEA with the aim of being able to reveal multi scaling exponents both for long and short ranged correlations. This idea seems to be inspired by the method of multifractal detrended fluctuation analysis which is an extension of the classical detrended fluctuation analysis with weighting of detrended standard deviations according to their magnitudes [3].

The main idea of Huang and colleagues [6] is now to replace the Shannon entropy with the Rényi entropy when scaling the probability density function of diffusion process constructed using time series. The Rényi entropy is actually a one-parameter family of entropies defined in the continuous case by

$$S(q) = \frac{1}{1-q} \ln \int [p(x)]^q \, dx, \tag{1}$$

where q has the meaning of a weight of different probabilities [7] (similarly, one can define the Rényi entropy in a discrete case); for $q = 1$ we obtain the Shannon entropy. The authors allow the parameter q to vary within a large region

E-mail addresses: am379@leicester.ac.uk, am379@le.ac.uk.

0378-4371/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2012.12.036

(e.g. $-1 < q < 4$) and they suggest that the new method should potentially provide information about the correct scaling behavior across various scales. In other words, the authors claim that if the probability density of a diffusion process admits scaling, we should have

$$S(t,q) \propto \delta(q)\log(t), \tag{2}$$

where t is the diffusion time; the scaling exponent can now depend on q. Based on the introduced method, they further try to reveal short term correlations between very small and extremely large oscillations of volatility of some financial time series. They call the new method the 'multifractal diffusion entropy analysis method', which I will refer to as MF DEA further for brevity.

Below I will show that although the method suggested by Huang et al. might look sound at first glance, due to a number of conceptual drawbacks its implementation can be misleading. In particular, some of the conclusions made by the authors regarding the volatility of the financial market may simply be artefacts of numerical simulations.

2. General drawbacks of the multifractal diffusion entropy analysis

(i) Firstly, I should emphasize that one needs to properly re-define the MF DEA method for negative values of q. Indeed, for a diffusion process, the probability density function (pdf) (x, t) is determined on an unbounded x-axis; however, for $q \le 0$ the integral in the Rényi entropy (1) will be diverging since $[p(x,t)]^q \to +\infty$, $|x| \to \infty$. Moreover, in the case, where the asymptotic pdf shows a Levy type of dynamics with $p(x,t) \propto x^{-\mu}$, $2 < \mu < 3$, the integral becomes diverging starting already from $q < 1/2$. For those reasons, one needs to introduce the limits for integration, i.e. to consider

$$S(q,t) = \frac{1}{1-q}\ln\int_{F_1(t)}^{F_2(t)}[p(x,t)]^q\,dx. \tag{3}$$

Interestingly, the choice of $F_i(t)$ cannot not be arbitrary: different parameterizations of $F_i(t)$ will result in different scaling coefficients δ for the same q. For time series with a single scaling coefficient, the natural choice of $F_i(t)$ should be $F_i = A_i t^{\delta}$, where δ is the scaling coefficient and A_i are constants. Only in this case we shall obtain the required behavior $S(t) \propto \delta \log(t)$. It is easy to show that in the case of the discrete analogue of (1) we will face the same problem as before: we shall need to cut somewhere an infinite sum of elements.

The amended definition (3) of the diffusion entropy $F_i = A_i t^{\delta}$ will seriously restrict implementation of MF DEA. Indeed, to compute the value of entropy we already need to know the scaling δ of the integration limits, i.e. the scaling of the underlying pdf. However, in this case, we simply do not need to compute S: such computation will not provide us with any important information if we already know the scaling exponent. One can object saying that when computing the diffusion entropy (3) for $q < 1/2$ we can simply use the maximal $x_{\min}$ and the minimal $x_{\max}$ values of the time series under consideration which give an estimate for F_i; actually, this idea was used in Ref. [6]. I found, however, that such a way of computing F_i can be erroneous due to the large deviation of $x_{\min}$ and $x_{\max}$ which can be easily seen from the following example shown in Fig. 1. Here I formally applied the MF DEA from Ref. [6] for $-1 < q < 3$ to different artificial time series with a well-known nature: the classical uncorrelated random walk, correlated random walk, uncorrelated Levy flight and some others. As an example, in Fig. 1(B) I show S computed numerically and plotted versus diffusion time for the positively correlated Brownian motion ($H = 0.75$).

One can see from the figure that generally, the scaling exponents obtained by fitting the graphs for MF DEA decrease with a decrease of q. However, this fact does not indicate any antipersistent correlation between extremely large and extremely low oscillations and multifractality of the process as suggested by Ref. [6]. Rather, this phenomenon happens even for a mono-fractal non-correlated random walk and non-correlated Levy flight, which is definitely a computation artefact. Interestingly, I found that a similar contradiction remains for simple artificial multifractal time series as, for instance, fractional Brownian motion with variable Hurst exponent (not shown result). Consequently, if the method does not successfully work even for simple artificial sequences, one should not expect it to work for more complicated real-world time series. In particular, by re-considering financial time records from Ref. [6], I found that the extreme values of the time series increase in time, but they do not show well pronounced scaling behavior, thus possibly causing errors in the estimation of (3) (result not shown). Extensive numerical simulations show that increasing the length of time series does not qualitatively modify the above results: for $q < 0$ all the scaling coefficients obtained based on the MF DEA method will rapidly drop with a decrease in q. This property does not depend on whether or not time series are positively or negatively correlated, and cannot be used as an indication of antipersistence as suggested in Ref. [6].

(ii) Another important drawback of the MF DEA is that for negative q, the main contribution into the integral is given by the points located at the tail of the pdf and the estimated values of $p(x, t)$ based on the time series is (see Ref. [6]):

$$p \approx \frac{N_1}{N-t+1} \approx \frac{N_1}{N}, \tag{4}$$

where N is the total number of points in the time series, N_1 is the number of points in a histogram cell centered at x. When constructing the histogram from the data, we always have only a few points (often $N_1 = 1$ or 2) in the histogram cells at the tail of the distribution. The eventual error in replacing the true distribution $p(x, t)$ (regardless of its shape) with the values $1/N$ or $2/N$ becomes largely amplified in the case of the power law distribution $p(x) \propto x^{-\mu}$ when we compute $[p(x)]^q$ for