Physica A 390 (2011) 4057–4072

Contents lists available at SciVerse ScienceDirect

Physica A

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

Effects of coarse-graining on the scaling behavior of long-range correlated and anti-correlated signals

Yinlin Xu [a,b], Qianli D.Y. Ma [c,d], Daniel T. Schmitt [a], Pedro Bernaola-Galván [e], Plamen Ch. Ivanov [a,c,e,*]

[a] Center for Polymer Studies and Department of Physics, Boston University, Boston, MA 02215, USA

[b] College of Physics Science and Technology, Nanjing Normal University, Nanjing 210097, China

[c] Harvard Medical School and Division of Sleep Medicine, Brigham & Women's Hospital, Boston, MA 02215, USA

[d] College of Geography and Biological Information, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

[e] Departamento de Física Aplicada II, Universidad de Málaga, 29071 Málaga, Spain

ARTICLE INFO

Article history:
Received 26 January 2011
Received in revised form 6 May 2011
Available online 14 July 2011

Keywords:
Coarse-graining
Detrended fluctuation analysis
Scaling behavior

ABSTRACT

We investigate how various coarse-graining (signal quantization) methods affect the scaling properties of long-range power-law correlated and anti-correlated signals, quantified by the detrended fluctuation analysis. Specifically, for coarse-graining in the magnitude of a signal, we consider (i) the Floor, (ii) the Symmetry and (iii) the Centro-Symmetry coarse-graining methods. We find that for anti-correlated signals coarse-graining in the magnitude leads to a crossover to random behavior at large scales, and that with increasing the width of the coarse-graining partition interval Δ, this crossover moves to intermediate and small scales. In contrast, the scaling of positively correlated signals is less affected by the coarse-graining, with no observable changes when $\Delta < 1$, while for $\Delta > 1$ a crossover appears at small scales and moves to intermediate and large scales with increasing Δ. For very rough coarse-graining ($\Delta > 3$) based on the Floor and Symmetry methods, the position of the crossover stabilizes, in contrast to the Centro-Symmetry method where the crossover continuously moves across scales and leads to a random behavior at all scales; thus indicating a much stronger effect of the Centro-Symmetry compared to the Floor and the Symmetry method. For coarse-graining in time, where data points are averaged in non-overlapping time windows, we find that the scaling for both anti-correlated and positively correlated signals is practically preserved. The results of our simulations are useful for the correct interpretation of the correlation and scaling properties of symbolic sequences.

1. Introduction

Certain complex physical and biological systems have no characteristic scale and exhibit long-range power-law correlations. Due to nonlinear mechanisms controlling the underlying interactions, the output signals of complex systems are also typically nonstationary, characterized by embedded trends and heterogeneous segments with different local statistical properties. Traditional methods such as power-spectrum and auto-correlation analysis [1–3] are not suitable for nonstationary signals. To address this problem, detrended fluctuation analysis (DFA) was developed to more accurately quantify long-range power-law correlations embedded in a nonstationary time series [4–6]. In addition to quantifying scale-invariant features, DFA has also been used to detect the characteristic scales in non-homogeneous signals [7,8].

* Corresponding author at: Center for Polymer Studies and Department of Physics, Boston University, Boston, MA 02215, USA.
E-mail address: plamen@buphy.bu.edu (P.Ch. Ivanov).

0378-4371/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2011.05.015

The DFA method has been successfully applied to quantify the output dynamics of various physical and biological systems including meteorology [9], climate temperature fluctuations [10,11], river flow and discharge [12], economics [13,14], neural receptors in biological systems [15], DNA [4,16,17], cardiac dynamics [18–22], and human gait [23–25]. DFA has been also utilized to identify transitions across different states of the same system characterized by different scaling behavior, e.g., the scaling exponent for heart-beat intervals discriminates between healthy and sick individuals [18], wake and sleep state [26], and across sleep stages [19,22] and circadian phases [21]. It has been shown that the DFA scaling exponent obtained from symbolic DNA sequences relates to the evolutionary degree of various organisms [16].

To understand the intrinsic dynamics of a given system, it is important to know how its intrinsic nature (e.g. nonstationarities) or external manipulations such as data pre-processing can affect the results. In the previous studies we have investigated the effect of various data artifacts on the DFA scaling analysis of long-range power-law correlated signals. Specifically, we considered different types of nonstationarities associated with different trends present in the signal, e.g., polynomial, sinusoidal and power-law trends [27]. Furthermore, we have studied the effects of nonstationarities that are often encountered in real data or result from "standard" data pre-processing approaches, e.g., signals with segments removed, signals with random spikes as well as signals with different local behavior [28]. We have also investigated the effects of linear and nonlinear filtering of signals [29] and the effects of extreme data loss [30]. Comparative studies of the performance of the DFA method and other scaling analysis methods are presented in Refs. [5,31,32].

In this paper, we focus on the effects of different coarse-graining (i.e., signal quantization) approaches on the scaling properties of correlated signals quantified by the DFA. Different coarse-graining and symbolic analyses have been developed and utilized to investigate complex dynamics generated by various physical [33–38], chemical [39], geological [40,41], biological and physiological [42–45] nonlinear systems.

Coarse-graining in the magnitude of a signal. It consists of the discretization of data and is frequently imposed by the nature of the measurements, i.e., the limitations on the accuracy of instruments, acquisition-data rate or even data-storage requirements. In many situations, coarse-graining is imposed by the intrinsic nature of the data, while in other cases coarse-graining is applied at a latter stage; for example, to compute certain information theory measures derived from Shannon entropy. These functionals can be applied only to symbolic time series because they take as arguments probability distributions (in practice, we never have access to the full distribution function but only to a binned version of it).

One of these functionals, especially suitable when dealing with correlations in time series, is the mutual information. It is well known that mutual information is closely related to the correlation function but, contrary to it, mutual information captures the complete dependence structure, including nonlinear correlations [46,47]. The estimators of mutual information require the data to be binned, i.e., the range of the data is partitioned and each element of the time series is assigned to an interval of the partition. This transformation is equivalent to the coarse-graining procedures described in this study, and depending on the details of such transformation the correlation structure of the signal can be modified. For example, it is known that uniform partitions in some situations can lead to misleading results [48]. The measure of correlations using mutual information has been used in several fields: DNA sequences [49], physiological signals [50], quantum information [51], and complex systems [52,53].

Coarse graining in time (or space). It consists of substituting the values of the signal in an interval by its mean value. It may appear as a direct effect of the measurement (e.g. sonometers or pyranometers integrate the input during the measurement interval) or due to further modification of data. It is common to use this coarse-graining to smooth out the short scale heterogeneities of a signal. It has been recently used to analyze the long-scale structure of DNA [54–57]. Another typical application of coarse-graining in time is the mean-field approximation where one considers the mean effect of the interactions during a given period (in time or space) in order to simplify the problem. Such techniques have been used to study the growth of fractal structures in the framework of the mean-field approximation [58].

The outline of this paper is as follows. In Section 2, (i) we review the Fourier filtering method for generating long-range power-law correlated signals, (ii) we describe the Floor, Symmetry and Centro-Symmetry methods of coarse-graining in magnitude and the coarse-graining in time method, and (iii) we briefly present the DFA method for scaling analysis. In Section 3, we compare the scaling properties before and after coarse-graining for both anti-correlated and positively correlated signals to study the effects of various coarse-graining approaches on the DFA scaling. In Section 4, we summarize our findings.

2. Method

2.1. Fourier filtering method

Using a modified Fourier filtering method [59], we generate stationary uncorrelated, correlated, and anti-correlated signals $x(i)$ ($i = 1, 2, 3, \ldots, N_{max}$) with a zero mean and standard deviation $\sigma = 1$. This method consists of the following steps.

(a) First, we generate a random uncorrelated and Gaussian distributed sequence $\eta(i)$ and calculate the Fourier transform coefficients $\eta(q)$.
(b) The desired signal $x(i)$ must exhibit correlations, which are defined by the form of the power spectrum

$$S(q) = \langle x(q)x(-q)\rangle \sim q^{-\beta}, \tag{1}$$