Commun Nonlinear Sci Numer Simulat 15 (2010) 637–647

Contents lists available at ScienceDirect

Commun Nonlinear Sci Numer Simulat

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

Universal distribution function for the strongly-correlated fluctuations: General way for description of different random sequences

R.R. Nigmatullin *

Theoretical Physics Department, Kazan State University, Kremlevskaya str., 18 Kazan, 420008 Tatarstan, Russian Federation

ARTICLE INFO

Article history:
Received 18 April 2009
Received in revised form 4 May 2009
Accepted 4 May 2009
Available online 12 May 2009

PACS:
02.50.Sk
0.5.45.+b
02.60.Ed, Pn
06.20.Dk
07.05.Kf
87.10.+b

Keywords:
Universal distribution function of relative fluctuations
Procedure of the optimal linear smoothing
Quantitative "reading" of nano-noises

ABSTRACT

It has been proved that for the strongly-correlated fluctuations there is a universal distribution function for the relative fluctuations (UDFRF). The analytical form of this function follows from the solution of some types of the functional equations. For obtaining the UDFRF a procedure of the optimal linear smoothing (POLS) has been developed. This procedure based on criterion of the minimal relative error helps to separate correctly a possible trend (the "low-frequency" curve, defined as the *generalized mean value* curve or trend) from the "high-frequency" (HF) fluctuations, defined as a random sequence of *relative fluctuations* with zero trend. A universal treatment procedure outlined in this paper helps to find an *optimal* trend, separate it from the relative HF fluctuations and read them *quantitatively*. The statistics of the fractional moments outlined in this paper helps "to read" the found trends and express them in terms of the fitting parameters if the model for their description is *absent*. These new possibilities can be applied for description of different noises (quantum fluctuations, for example) that always present on the scale ($10^{-6} \div 10^{-9}$ m). Quantitative reading of these noises with their subsequent classification is important for every developing nanotechnology that it has a possibility to be applied in this range of scales.

1. Introduction

In analysis of any random sequence we have three different types of errors:

1. The measurement errors related to equipment used.
2. The uncontrollable errors related to the model (physical, mathematical) chosen.
3. The uncontrollable errors related to treatment procedure.

The basic question can be formulated as follows.

Is it possible to eliminate the errors related to points 2 and 3? Recent investigations of the author of this paper show that answer can be definitely *positive*.

The author of this paper suggests the unique (not having similar analogies) methods related to a quantitative "reading" of an arbitrary random sequence having different (technical, geological, economical, medical, and etc.) origin. These suggested methods have the following remarkable properties:

* Tel.: +7 8432360612.
E-mail address: nigmat@knet.ru

1007-5704/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.cnsns.2009.05.019

P1. These methods are *noninvasive* i.e. they contain only *controllable* errors related to transformations of the random sequence considered. Thanks to this property they have doubtless advantages before invasive methods as: the classical Fourier-transform (the well-known Gibbs oscillations phenomenon), wavelet transformations and other recently developed methods [1–6] that contain uncontrollable errors and correlations [7].

P2. Using new methods any random sequence can be read "quantitatively" and, if it is necessary, can be compared with another sequence with the usage of a'*universal*' set of the reduced (fitting) parameters.

P3. These suggested methods are completely free from any *a priori* (model) assumption related to statistical nature of the random sequence analyzed.

These three basic properties make these methods as indispensable and universal tool for analysis of random sequences having different statistical nature.

At present time in accordance with ideology of their construction they can be divided on three independent parts.

1.1. The procedure of the optimal linear smoothing (POLS)

This method helps to find the optimal and smoothed trend (pseudo-fitting) function and divide it from their relative fluctuations. Let us suppose that the random sequence considered contains large-scale fluctuations (trend) and high-frequency fluctuations, which are usually determined as a "noise". In order to separate those from each other we use the procedure of the optimal *linear* smoothing (POLS) with the use of Gaussian kernel. This procedure is defined as

$$\tilde{y} = Gsm(x, y, w) = \frac{\sum_{j=1}^{N} K\left(\frac{x_i - x_j}{w}\right) y_j}{\sum_{j=1}^{N} K\left(\frac{x_i - x_j}{w}\right)}, \quad K(t) = \exp\left(-t^2/2\right). \tag{1}$$

Here the function $K(t)$ defines the Gaussian kernel, the value w defines the fixed width of the smoothing window. The set y_j (j = 1,2,...,N) defines the initial noisy sequence. In spite of the fact that there are many smoothing functions imbedded in many mathematical programs this chosen function has two important features: (a) the transformed smoothed function (1) is obtained in the result of *linear* transformation and does not have uncontrollable error; (b) the value of the smoothing window (w) is adjustable (fitting) parameter and accept *any* value. This function in a certain sense can be considered as a *pseudo-fitting function*, which is not associated directly with a specific model describing the desired process. The value of the optimal window w_{opt} is chosen from the condition

$$\begin{aligned} &\Delta n_j = y_j - Gsm(x, \tilde{y}, w), \\ &\tilde{y}_{w'} = Gsm(x, y_w, w'), \quad w' < w, \\ &\min(RelErr) = \left(\frac{stdev(|y_{w'} - y_w|)}{mean(y_w)}\right) \cdot 100\%. \end{aligned} \tag{2}$$

This procedure automatically decreases the value of the initial fluctuations and helps to find the optimal value of the parameter w minimizing the value of the relative error. In many model calculations realized this optimal value w_{opt} *does* exist that helps to find the optimal smoothed curve (trend) describing the large-scale fluctuations. The desired trend minimizing the value of the relative error is described by expression

$$\tilde{y} = Gsm(x, y, \tilde{w}), \tilde{w} \equiv w_{opt}. \tag{3}$$

After calculation of the optimal trend it becomes possible to divide initial random sequence on two parts: (a) the optimal trend expressed by relationship (3) and (b) detrended sequence representing the values of the *relative fluctuations*, which is expressed as

$$srf = y - \tilde{y}. \tag{4}$$

Here *srf* defines the detrended sequence of the relative fluctuations. This method helps to divide initial noise on two parts (trend defined here as the pseudo-fitting function) and relative fluctuations. An example of application of the POLS on real data is illustrated by Figs. 1(a–c). The IR spectra of the mercaptophenyl diazonium (MCP) chemically adsorbed on some golden electrodes defined as AU(1–3) data was measured. Real data are presented by Fig. 1(a). Fig. 1(a) shows the plot presented by expression (2). The value of the minimal error corresponds to the first local minima and expression (3) determines the desired trend at w = 0.05. It is obvious that the global minima (shown on Fig. 1(a) on the left) corresponds the smoothed procedure with the value of optimal window closed to zero value. Fig. 1(a) shows that the POLS can be applied to the relative fluctuations of the first order. This procedure is useful when the large fluctuations in expression (4) after the first subtraction cannot be totally eliminated. Nontrivial examples of application of the POLS for analysis of real data can be found in papers [8,9].