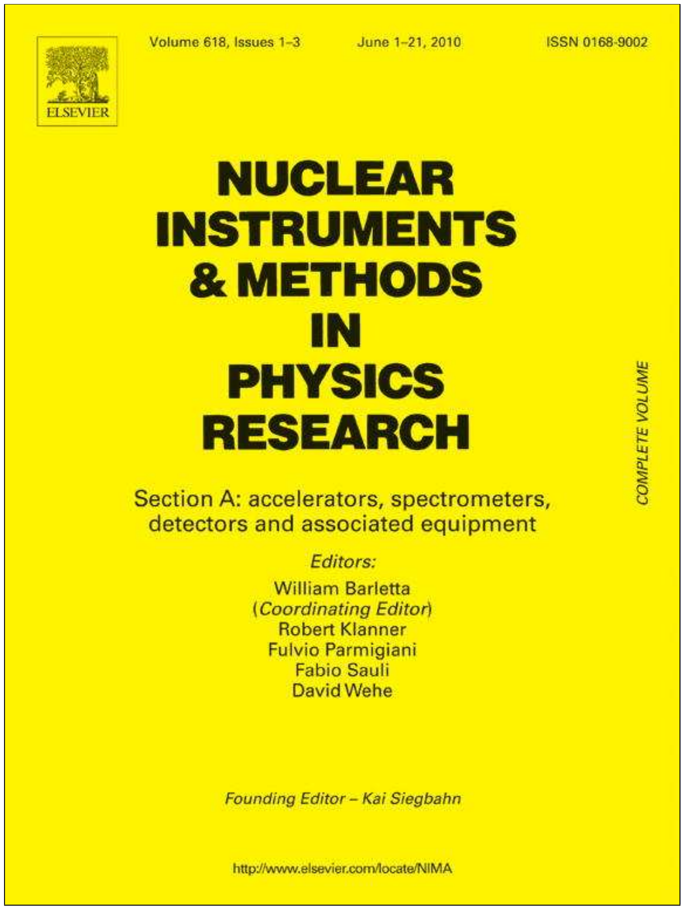
Volume 618, Issues 1–3
June 1–21, 2010
ISSN 0168-9002
ELSEVIER
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH
Section A: accelerators, spectrometers, detectors and associated equipment
Editors:
William Barletta
(Coordinating Editor)
Robert Klanner
Fulvio Parmigiani
Fabio Sauli
David Wehe
COMPLETE VOLUME
Founding Editor – Kai Siegbahn
http://www.elsevier.com/locate/NIMA

Nuclear Instruments and Methods in Physics Research A 618 (2010) 306–314

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

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

Noise-level determination for discrete spectra with Gaussian or Lorentzian probability density functions

Netzer Moriya *

siOnet Ltd. – Applied Modeling Research, Herzelia 46445, Israel

ARTICLE INFO

Article history:
Received 24 November 2009
Received in revised form
8 February 2010
Accepted 14 February 2010
Available online 23 February 2010

Keywords:
Probability density function
Adaptive
Noise
Stochastic process
Gaussian
Cauchy-distribution
Non-stationary
Variance

ABSTRACT

A method, based on binomial filtering, to estimate the noise level of an arbitrary, smoothed pure signal, contaminated with an additive, uncorrelated noise component is presented. If the noise characteristics of the experimental spectrum are known, as for instance the type of the corresponding probability density function (e.g., Gaussian), the noise properties can be extracted. In such cases, both the noise level, as may arbitrarily be defined, and a simulated white noise component can be generated, such that the simulated noise component is statistically indistinguishable from the true noise component present in the original signal. In this paper we present a detailed analysis of the noise level extraction when the additive noise is Gaussian or Lorentzian. We show that the statistical parameters in these cases (mainly the variance and the half width at half maximum, respectively) can directly be obtained from the experimental spectrum even when the pure signal is erratic. Further discussion is given for cases where the noise probability density function is initially unknown.

1. Introduction

As any measured data-signal or experimental spectrum is always contaminated by noise to some extent, the nature of the noise component, corrupting a pure signal in stochastic processes, is of central relevance and is recently being acknowledged by a growing portion of the scientific and engineering communities. In many situations, the careful treatment of noise-induced, non-equilibrium phenomena, represented by nonlinear index-series signals (e.g., time or channel) and the prediction of the noise-level characteristics in such physical stochastic systems, may be of great value [1]. This is particularly true for dynamical processes [2–4] where nonlinear behavior is expected, and may seriously alter any estimation of the states of the system, if not actually cause a total divergence of the model parameters. Such conditions are common in many nonlinear systems when modeled by recursive or adaptive methods such as Wiener or extended Kalman filtering.[1]

In recent years, the estimation of noise level in time series signals has gained noticeable attention, especially in cases where the measured signals or spectra are known to be highly non-stationary, as for instance in electroencephalogram (EEG) spectra and similar biological fields, geophysics and chaotic systems. Refs. [6–17] (and references therein) present the relevant literature for a comprehensive background on some practical and theoretical aspects of noise-level estimation approaches in nonlinear systems, that are subject to non-stationary time evolution. In particular, the authors of Refs. [6,7] present methods to estimate the noise component in noisy time series based on computing vectors of logarithmic displacement [6] or time-dependent exponent curves [7]. However, such methods lack the generality needed for most cases or need the original signal to be known.

General background in the fields of stochastic processes and noise characteristics may be found in standard textbooks (i.e., [18,19]), however, the general problem of estimating the noise level given a non-stationary, nonlinear and initially unknown signal is usually not taken into consideration. Most other methods to extract the noise properties or reduce the effect of noise in experimental measurements that use for example, polynomial fit approaches, such as, *moving average* (MA) [20,21], *moving weighted average* (MWA), *local polynomials method* (LP) [22], *locally weighted scatter plot smoothing* (LOWESS) [23] and other related methods [24], all suffer from the incomplete initial knowledge concerning the degree of the polynomial fit applicable for the numerical manipulation on one hand and the ill-posed conditions of the calculations when using high order polynomials on the other hand. In most cases these limitations prevent the extraction of valuable information pertaining to the noise characteristic in the contaminated signal or experimental spectrum.

Some general methods for estimating the noise statistics such as the *Bayesian estimation* method, *maximum likelihood estimation*,

* Tel.: +972 9 956 9893; fax: +972 9 957 5336.
E-mail address: netzer@si-O-net.com

[1] See for example [5, Chapter 6].

0168-9002/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2010.02.106

innovation correlation and other case-specific approaches have been well known for a long time, while some other methods have recently been published and are briefly reviewed in Refs. [25–28]. Fundamental aspects concerning the incorporation of external estimation of the noise statistics into adaptive filtering algorithms are discussed in Refs. [26–28]. These methods all suffer, however, from either numerical calculation inefficiency, if implemented in real-time applications, or need to be fine-tuned for the specific case under investigation.

Difference filtering, and in particular binomial filters, are well known and are described in references such as [29,30]. Recently, a general approach to derive an explicit expression for the probability density function of the *m*'th order numerical differentiation of a purely stochastic variable was presented [31], based on principles similar to binomial filtering, however, allowing both additions (convolution) and subtractions of relevant distributions to be performed. It has been shown that the statistical characteristics of the original (white-) noise component, following *m*'th-order differentiation can be obtained analytically based on the original probability properties of the observed (measured) signal in a simple manner. In the present contribution we focus on the statistical treatment of the noise extraction procedure and present a detail analysis of the noise level estimation when the additive noise is either Gaussian (normal distribution) or Lorentzian (Cauchy distribution) in nature. We show that the statistical parameters in these cases (namely the variance σ^2, for a normal distribution or the half-width at half-maximum (HWHM) γ, for the Lorentzian) can directly be obtained from the noisy signal even when the pure signal is erratic. In this paper we assume that the unknown pure signal is square-integrable and further, can be approximated by a (arbitrarily high) *m*'th-degree polynomial function (see general remarks on the order of the polynomial degree below), and suggest the use of high-order differentiation of the noisy signal in order to derive the statistical parameters of the noise component of such signal, i.e., a pure signal corrupted with additive white noise. The main aim of this analysis is to obtain the first moments of the noise component in the random signal, and in particular, for dynamic, non-stationary processes, where the statistical noise parameters may be functions of time that corrupt an unknown deterministic pure signal component. The approach taken in this paper would be to time-window the sequenced data signal and treat the sliced signal as an independent identical distributed random variable in the analysis proposed, so as to allow any time-varying statistical property of the initial noise component to be developed.

The outline of the paper is as follows: In Section 2 we briefly review the general method and demonstrate that the proposed approach allows for the estimation of the statistical parameters of the original distribution and further, to simulate the noise contribution of the original stochastic signal so that the simulated noise component is statistically indistinguishable from the true contribution of the noise in the originally observed data signal. In Section 3, a discussion follows on the ability to extract the original statistical moments of the noise component in the contaminated signal (such as its variance or higher moments). It is demonstrated in Section 4 that the noise level can be obtained in real-time (or off-line) and with only a few assumptions on the nature of the signal itself. The central parts of the paper are Sections 3 and 4, where the statistical results are derived and demonstrated for the normal (Gaussian) and Cauchy distribution (Lorentzian) cases. We further apply the method to experimental results and extract the noise level in some typical Rutherford Backscattering spectra (RBS). The paper is concluded with some general remarks in Section 5 where we also mention some of the limitations of the approach and discuss the main assumptions that are needed for consideration.

2. The stochastic high order numerical differentiation

We first review the basic idea in difference filtering and in particular binomial filters (see for instance Refs. [29,30]), however, generalizing the method to allow both additions (convolution) and subtractions of relevant distributions to be performed. We consider the random component in a stochastic process $\xi(n_i)$ with n_i, the collection of stochastic events and refer to the case where the stochastic variable represents a random signal with a known probability distribution function (see Ref. [31] for detailed assumptions and conditions on the signals and mathematics). A differentiating operator, operating on a signal, may then be defined with respect to the index of the signal data points in their sequenced order (or equivalently, treating the signal as a time series vector with a unit time step). By this, one may realize that a differentiation procedure, of the first order, is equivalent to subtracting the element n_i from the element n_{i+1} in the stochastic signal. Since, in such a random set of points, each point is totally independent of all other points in the set and controlled only by the mutual statistics that they all belong to (the sample space, i.e., all points (i,j) are uncorrelated for $i \neq j$), the equivalence to subtracting the element n_i from the element n_{i+1} in the noise signal would be the subtraction of two independent random variables with identical statistical distribution (IID). In contrast to the case of the first numerical differentiation, where one could assume that all individual data points were uncorrelated, higher order numerical differentiation involves correlated expressions that may lead, in the general case, to non-trivial expressions for the resultant probability functions.

Using the above definitions (and those listed in Ref. [31]) and referring to some arbitrary random variable function $V(n_i,\xi)$, considered here as the *original data spectrum* with ξ, the stochastic random variable, one can now derive the second order numerical differentiation index series $V^{(2)}(n_i^{(2)},\xi^{(2)})$, with $\xi^{(2)}$ referring to the unknown stochastic random variable corresponding to the second-order numerical differentiation vector by realizing that $n_i^{(1)}=n_i-n_{i-1}$ and $n_{i+1}^{(1)}=\ n_{i+1}-n_i$ so that $n_i^{(2)}=n_{i+1}^{(1)}-n_i^{(1)}=n_{i+1}-2n_i+n_{i-1}$ (we use the parenthesized superscript (m) to denote differentiation of order m). These expressions imply that the probability density function of the second order numerical differentiation is the equivalent *pdf* of the sum of three independent, random variables, all with similar, yet, non-identical, probability density functions (InID).

We now recall that given two independent random variables: $\xi_1,\xi_2 \in \mathbb{R}^k$, with μ and ν corresponding to their respective distribution functions and f and g denoting their respective density functions, the distribution of the sum $\xi_1+\xi_2$ is the convolution $\mu * \nu$ operation between the corresponding distribution functions and the analogue density function of the sum equaling the convolution integral denoted by $f*g$ in the index/time domain. When the distribution of $\xi_1-\xi_2$ is needed, the mathematical operation is no longer that which is referred to as the convolution and so we use the notation $*^{(-)}$ and the explicit integral is to be performed. We will exploit both of these operations in the following. Using the notation $f_{\xi_1,\xi_2,\xi_3}=f_{\xi_1} *^{(-)} f_{\xi_2} * f_{\xi_2}$ (i.e., the density function of the subtraction of the two stochastic variables with probability density functions f_{ξ_1} and f_{ξ_2} calculated first and then added to the density function of the random variable f_{ξ_3}, implies that the second order numerical differentiation may be written as[2]

$$f_{n_i^{(2)}} = (f_{n_i} *^{(-)} f_{(-2n_i)}) * f_{n_i} \tag{1}$$

[2] This directly follows from the second order differentiation step $n_i^{(2)}=n_{i+1}^{(1)}-n_i^{(1)}=n_{i+1}-2n_i+n_{i-1}$.

Following the above arguments for higher numerical differentiation, it can now be easily deduced that the m'th order numerical differentiation of a random variable derived from an arbitrary statistically defined variable, can be obtained by noting that the corresponding weights that dictate the numerical differentiation expressions are given by the non-zero elements (with the zeros omitted) of following matrix denoted here as the *Stochastic-Derivative matrix*, S_k^m [31], (shown here for $m=9$):

$$S_k^m = \left[\begin{array}{ccccccccccccccccc}
 & & & & & & & & 1 & & & & & & & & \\
 & & & & & & & 1 & & -1 & & & & & & & \\
 & & & & & & 1 & & -2 & & -1 & & & & & & \\
 & & & & & 1 & & -3 & & 3 & & -1 & & & & & \\
 & & & & 1 & & -4 & & 6 & & -4 & & 1 & & & & \\
 & & & 1 & & -5 & & 10 & & -10 & & 5 & & -1 & & & \\
 & & 1 & & -6 & & 15 & & -20 & & 15 & & -6 & & 1 & & \\
 & 1 & & -7 & & 21 & & -35 & & 35 & & -21 & & 7 & & -1 & \\
1 & & -8 & & 28 & & -56 & & 70 & & -56 & & 28 & & -8 & & 1
\end{array}\right] \quad (2)$$

where k denotes the columns of the Stochastic-Derivative matrix.

It can easily be verified that each of the elements is given by

$$S_j^{(m)} = (-1)^{j+1}\binom{m}{j} \quad (3)$$

where $\binom{m}{j}$ represents the elements of the binomial coefficients (see also in Refs. [29,30]).[3]

Note that the above expression was derived in order to account for the weight of the individual contributions in the integrals that represent the additions or subtractions of the stochastic variables involved in the differentiation process (i.e., the elements of S_j^m and represents the factors that are needed in order to calculate the probability density function of the differentiated set, and not the differentiated noise values themselves (that increase rapidly in value, as will be shown in the following). In terms of a summation of the individual elements needed to account for the probability density function of the m'th order numerical differentiation, one may write:

$$f_m = \sum_{j=1}^{m} S_j^{(m)} f(z) \quad (4)$$

with $f(z)$ representing the probability density function of the original random variable. It can be shown [31] that for the general case:

$$F_Z(z) = \iint_{D_z} \prod_{j=1}^{m} f_{\zeta_j}\, d\xi_1\, d\xi_1 \ldots d\xi_N = \iint_{D_z} \left[\prod_{j=1}^{m} S_j^m f_j(\xi)\right] d\xi^{(m)} \quad (5)$$

where $F_Z(z)$ is the probability distribution function of the new random variable Z over its volume of existence D_z. Since the density function used here is the same for all individual elements of the multiplication term under the integral, only weighted properly by the suitable elements of the stochastic derivative matrix, this can symbolically be written as

$$F_{n_i}^{(m)} = \iint_{D_z} \left[\prod_{j=1}^{m} S_j^m f(\xi)\right] d\xi^{(m)} \quad (6)$$

where $F_{n_i}^{(m)}$ denotes the probability distribution function of $\partial^m V(n_i,\zeta)/\partial i^m$ that can easily be evaluated to derive the respective density function, recalling that the term $\prod_{j=1}^{m} S_j^m f(\xi)$ really represents the integrals that are to be performed for the additions or subtractions of the stochastic variables involved in the differentiation process, in correspondence with the original probability function weighted accordingly, based on the Stochastic-Derivative matrix.

As a general remark, we note that the only assumption set on the random variable in the above derivation, is the assumption that it represents a pure random signal (with no deterministic component) with an arbitrary probability distribution function ($f(z)$).

It should also be mentioned that the above arguments, although mathematically simple, are general and are not restricted to any particular probability density function of the observation.

In what follows, we focus our discussion on the specific cases where the probability density functions of the noise statistics in the experimental spectrum is either Gaussian or based on Cauchy distribution function. Other cases will be discussed in the last section of this paper.

In the example of the Gaussian case, the analysis yields relatively straightforward expressions as the Gaussian *pdf* belongs to the few probability functions that adds and subtract into similar functions.[4] We therefore consider a Gaussian distribution, where ζ is referred to as the random variable, $f(\zeta) = N(0,\sigma_0^2)$, i.e., a Gaussian distribution where the first moment is equal to zero and the variance is given by σ_0^2 as an illustrative probability density function (the numerical differentiation of the following with mean values other than zero is straightforward).

For Gaussian case, the corresponding expression is of the form:

$$\frac{d^m N(0,\sigma_0^2)}{di^m} = N(0,\beta(m)\sigma_0^2) \quad (7)$$

with $\beta(m)$ given by the sum of the squares of the elements of the $m+1$'s row in the Stochastic-Derivative matrix.

Using Eq. (6) and the arguments above, one can derive the probability density function of a zero mean Normal distribution for any differentiation of order m of a discrete random signal [31]. For instance, the analytical expressions for the first, second and fifth numerical differentiation are simply:

$$\frac{d}{di}N(0,\sigma_0^2) = N(0,(1^2+1^2)\sigma_0^2) = N(0,2\sigma_0^2)$$

$$\frac{d^2}{di^2}N(0,\sigma_0^2) = N(0,(1^2+2^2+1^2)\sigma_0^2) = N(0,6\sigma_0^2)$$

$$\frac{d^5}{di^5}N(0,\sigma_0^2) = N(0,(1^2+5^2+10^2+10^2+5^2+1^2)\sigma_0^2) = N(0,252\sigma_0^2) \quad (8)$$

[3] The Stochastic-Derivative matrix S_j^m, as defined above, is in fact a variant of Pascal Triangle.

[4] Note that for $f(X) = N(0,\sigma_x^2)$ and $f(Y) = N(0,\sigma_y^2)$, the corresponding $f(Z = X \pm Y) = N(0,[\sigma_x^2+\sigma_y^2])$.

In a similar manner and based on the Cauchy distribution[5]:

$$f(X) = L(\mu = 0,\gamma) = \frac{\gamma/\pi}{x^2+\gamma^2} \tag{9}$$

where μ is the location parameter and γ is the half-width at half-maximum (HWHM) of the distribution, the following can easily be obtained:

$$\frac{d}{di}L(0,\gamma_0) = L(0,(1+1)\gamma_0) = N(0,2\gamma_0)$$

$$\frac{d^2}{di^2}L(0,\gamma_0) = L(0,(1+2+1)\gamma_0) = L(0,4\gamma_0)$$

$$\frac{d^5}{di^5}L(0,\gamma_0) = L(0,(1+5+10+10+5+1)\gamma_0) = L(0,32\gamma_0) \tag{10}$$

3. Noise level estimation

To demonstrate the proposed motivations for the use of high-order numerical differentiation of a stochastic signal for noise level extraction, we now refer to the case where uncorrelated noise, either due to the experimental set-up or to the process itself (or both these sources), is added to the signal and is blurring the pure data signal. It is the aim of the following to demonstrate how to extract a simulated noise component such that the simulated noise is statistically identical to the noise part in the original experimental signal.

In this section, we present a detailed analysis, followed by a numerical simulation, for the extraction of the noise level in an arbitrarily given noisy signal where the noise component is either Gaussian or Lorentzian. Further, we show, in the example described in a later section, that the extracted statistical variables (the variance in the case of normally distributed additive noise component and the HWHM where Cauchy distribution is considered) are stable even when the differentiation order is extremely high.

For simplicity we thus assume that the arbitrary noisy signal is given by P=S+N, with N being the additive, uncorrelated noise added to the pure, smooth and continuous signal S. Here, the term "smooth signal" should be understood as only limited by the high frequencies inherently contained in the noise spectral characteristics. Let us further assume that on the interval of validity of S, one can approximate the stochastic signal P (for instance, in the Least Mean Square sense) by an m'th-degree polynomial function that may belong to a complete monomial basis. Note that no restrictions were set on the order of the numerical differentiation m that can, in principle, arbitrarily be chosen. This can be proved to be possible for any bounded, smoothed and continuous function S (via Weierstrass approximation theorem.[6]

Assuming the above, the following treatment is fully justified for the general case where the deterministic part of the noisy signal is finite (bounded) and smoothed (at least square integrable) so that the approximation:

$$\frac{d^{m+1}}{di^{m+1}}P = \frac{d^{m+1}}{di^{m+1}}N \tag{11}$$

may be used, since the m'th numerical differentiation of S under the above assumptions is constant, and thus vanishes for higher orders.[7] However, it should be noted (see also below) that the assumption of IID of the elements in the stochastic data set is now only approximated, as the deterministic part (if it exists) of the original signal may cause some dependence for the case of a signal with a deterministic component. The effect, however, is small and can be neglected as the order of differentiation increases. It is thus clear that the purpose of the differentiation is to terminate the effect of the "pure" signal (which is, of course, an unobservable component in the stochastic system). However, the differentiation does leave the differentiated noise contribution almost untouched (as we have assumed additive noise) and when assuming that a prior knowledge about the type of the noise is at hand (e.g., Gaussian, Lorentzian, etc.), the above analysis may suggest a general method to obtain the statistical parameters of the original noise components (i.e., before the differentiation). In this respect, one may interpret the proposed method as passing the noisy signal through a series of high-pass filters that iteratively reduce the power of the pure signal while maintaining that of the additive (assumed white) noise component.

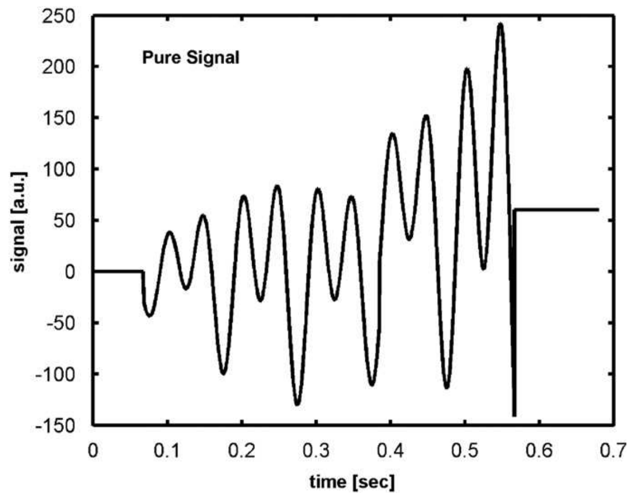

Fig. 1. The "Pure" signal in the example given in the text (the abbreviation [a.u.] in the figure indicates "arbitrary units").

Now, if the characteristics of the statistical properties of the high-order numerical differentiation of the original noise are derived (i.e., $N^{(m+1)}$), the probability density function that statistically describes the initial noise, subject to high-order numerical differentiation), in terms of the parameters (assumed to be unknown) of the statistical nature of the noise (assumed to be known), can obtained and thus the noise-level in the original signal P can be deduced.

It should also be mentioned that the effect of possible outliers in the original noise component, is drastically suspended as such events are usually propagated to the external regions of the axis in the differentiated domain and have only minor contribution on the *pdf* under study.

4. Simulated and experimental examples

To demonstrate the above, we now refer to a detailed example. At the end of this section we apply the principles of the method to

[5] Note also here that for $f(X) = L(0,\gamma_x)$ and $f(Y) = L(0,\gamma_y)$, it follows directly that $f(Z = X \pm Y) = L(0,[\gamma_x+\gamma_y])$.

[6] Also for non-smoothed and discrete function based on the *unisolvence theorem* for interpolating n nodes $> m+1$; see for example the classical proof by K. Weierstrass, Mathematische Werke, Bd. III, Berlin 1903, pp. 1–17. Can also be found in most textbooks on Functional Analysis.

[7] This approach can easily be generalized also for orthogonal, complete-basis functional systems other than polynomials.

some experimental results (typical RBS signals). For the sake of clarity, we list the following steps:

1. We first define an arbitrary signal ("pure") that was manipulated so that rapid structural characteristics are pronounced. Fig. 1 shows the arbitrary, pure deterministic component in the original signal to be used for later reference.
2. We then add a random, white and unbiased noise signal originated from a Gaussian or Cauchy probability density function (*pdf*) as shown in Fig. 2 (in this example $\sigma_0 = 14$ and $\gamma_0 = 5$).
   The noise component itself, as a function of time, is shown in Fig. 3.
3. To verify the original statistical parameters, we have performed a histogram calculation of the noise obtained in the previous step and perform a LMS fit to the Normal and Lorentzian distributions (see Fig. 4). The original σ_0 (Normal) and γ_0 (Cauchy) were indeed confirmed.Fig. 5
4. shows the "experimental" noisy signal that resulted by adding the noise contribution to the "pure" signal. As can easily be seen, the separation between the experimental spectra and the noise components in the combined sets of data points is blurred.
5. We next perform a 50'th order numerical differentiation of the data set and show the obtained result in Fig. 6 as the histogram of the resulting signal, LMS-fitted to a Gaussian and Lorentzian correspondingly. The re-normalization of the variance of the fitted Gaussian and the HWHM of the Cauchy distribution based on expressions (8) and (10) above (with m =50) have reviled σ and γ values to less than 1% deviation from the original value of σ_0 and the corresponding value of γ_0.
6. The corresponding differentiated signals are shown in Fig. 7. Note the huge spread of the spectrum.
7. In Fig. 8 we show the development of the normalized σ and γ as functions of the order of numerical differentiation up to m =200. It can be clearly seen that the obtained values are very close to the original statistical value of the initial noise.

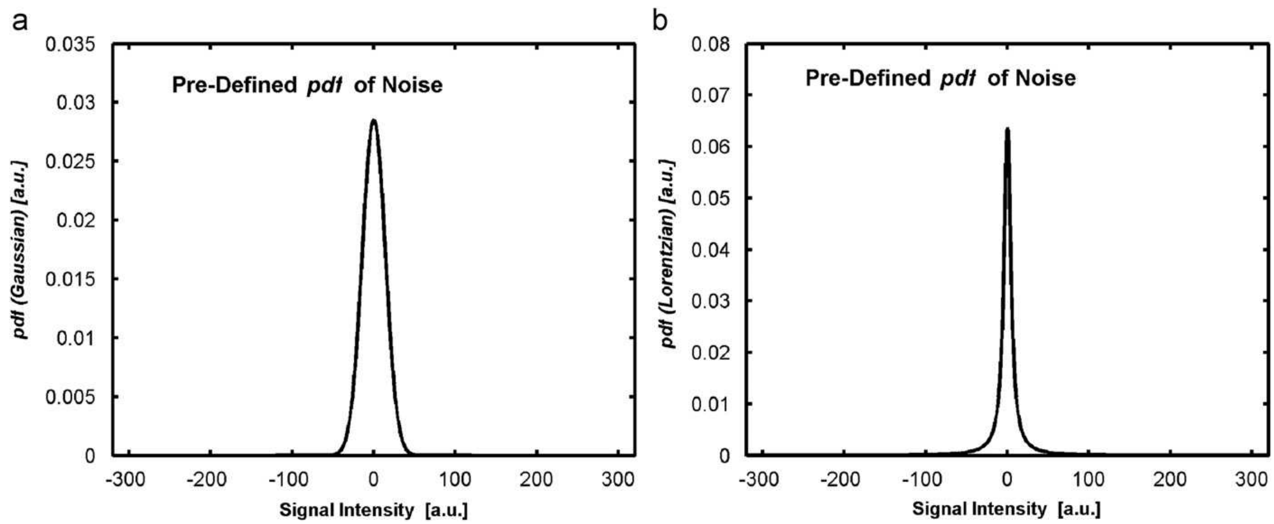

Fig. 2. The Gaussian (a) and Lorentzian (b) functions used as the model to generate a Normal and Cauchy distributed random signals in the example given in the text (the abbreviation [a.u.] in the figure indicates "arbitrary units").

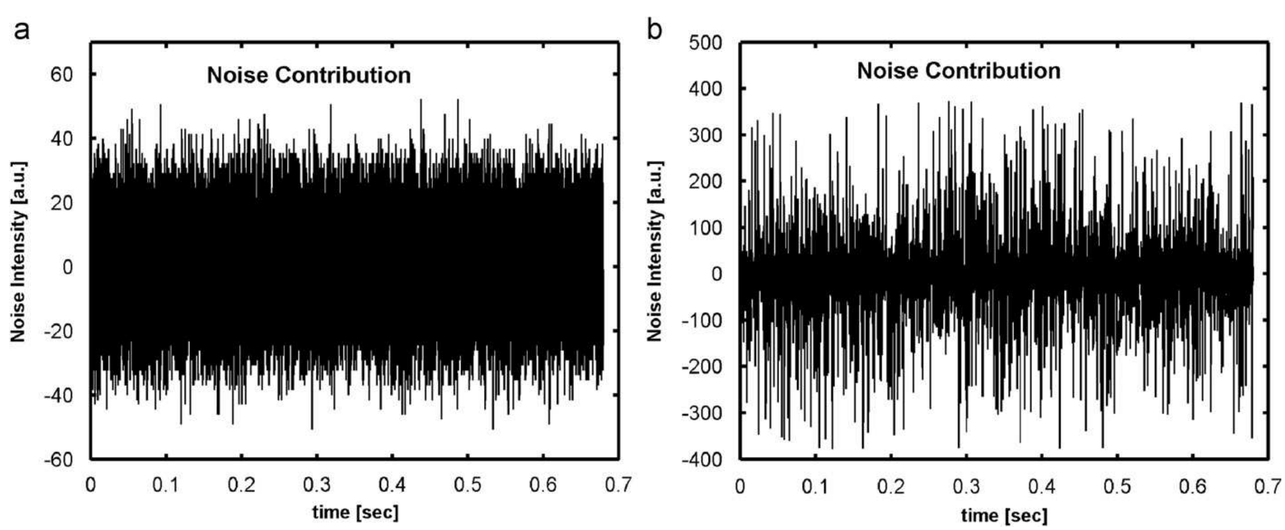

Fig. 3. A Normal (a) and Lorentzian (b) distributed random signals, representing the white noise contribution, derived from the model Gaussian and Lorentzian functions (Figs. 2 a,b) in the example given in the text (the abbreviation [a.u.] in the figure indicates "arbitrary units").

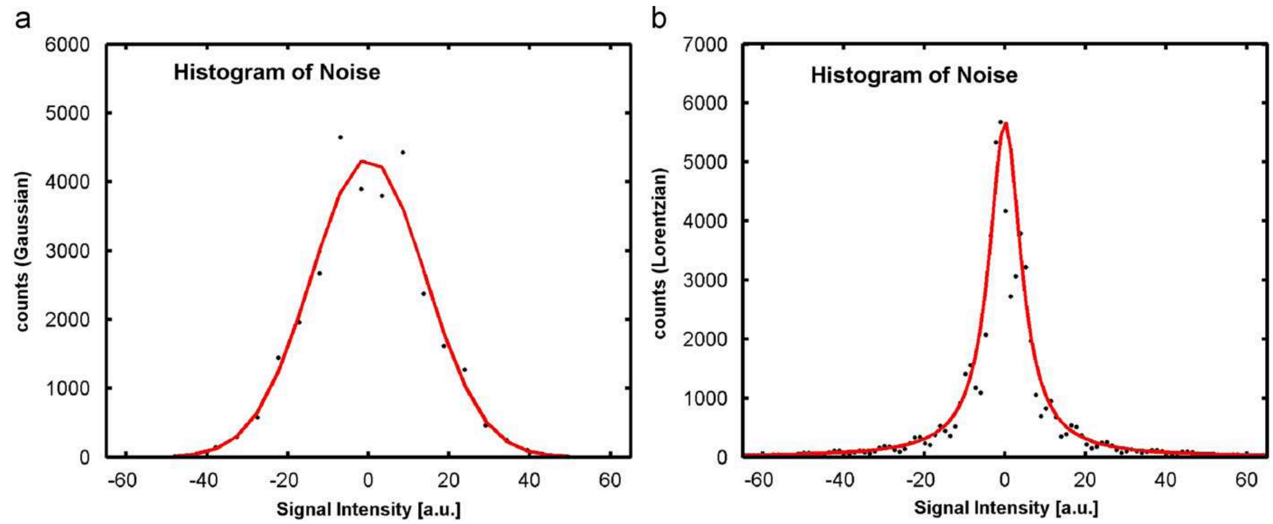

Fig. 4. The histogram of the noise signal (Fig. 3) fitted to a Gaussian (a) and Lorentzian (b) functions in the LMS sense. Histogram counts are dotted while fits are marked as solid lines (the abbreviation [a.u.] in the figure indicates "arbitrary units").

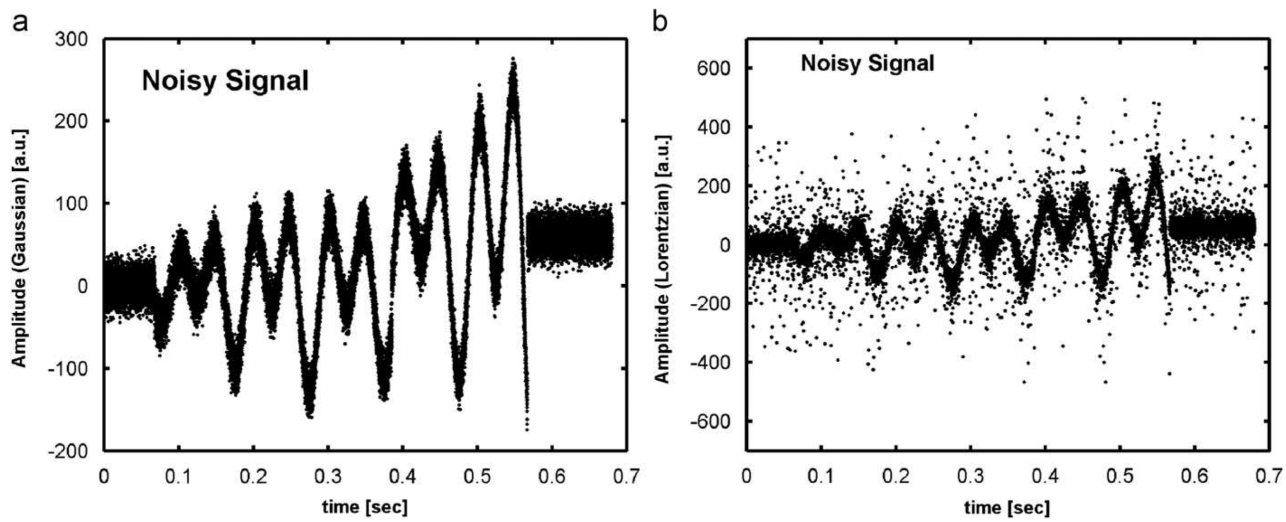

Fig. 5. The "noisy" signals used in the example given in the text and composed by adding the noise component to the "pure" signal (see Figs. 3 and 1 respectively). The abbreviation [a.u.] in the figure indicates "arbitrary units".

When any prior knowledge, with respect to the original probability function of the noise is not at hand (i.e., the type of the corresponding *pdf*), one may, based on the above approach, conduct a simple test, to deduce the specific probability density function of the original stochastic process in the observed data signal by reproducing only the noise component from the noisy signal and further process it so that no unintentional influences, related to the pure signal rather than to the noise, are affecting the probability testing procedure (we do not show this here).

Typical Rutherford Backscattering (RBS) spectra of a Fe-doped $\langle 100 \rangle$ InP substrate, were performed using 1.8 MeV He^{+} ion beam at random direction and along the $\langle 100 \rangle$ lattice axis (channeling direction). Fig. 9 shows both the raw RBS spectra (a) and the noise-level estimates as extracted by applying the HOSD analysis on the RBS spectra (b) assuming uncorrelated normal noise distribution. Differentiations of orders up to 15 are shown in Fig. 9(b) to yield stable results for the entire range of differentiation orders. Note, however, that due to the detailed structure of the spectra near channel 420, low differentiation orders are insufficient in obtaining the no-biased results and only higher orders revile the zero-biased noise level estimates. Comparisons (not shown here) with either polynomial techniques (see in Refs. [20,21]) or other smoothing methods such as Savitzky–Golay filtering [32,33], yielded inferior results due to the inherent difficulty to match a consistent set of parameters so to cover the entire spectral range, even when windowing approach was tested. Also, due to its direct process of the noisy signal to recover the noise component, the HOSD method has shown high numerical efficiency as compared to the other methods tested.

It should also be noted that the theory and method described here are totally insensitive to the sampling rate in the original discrete signal. This is due to the normalization of the time axis (in a time series signal) into an indexed axis where each of the sampled points are given integer indexes (i.e., $1, 2, \ldots, k$). The assumption that the separation between two adjacent points is one unit (i.e., that the sampled points are equally separated in the

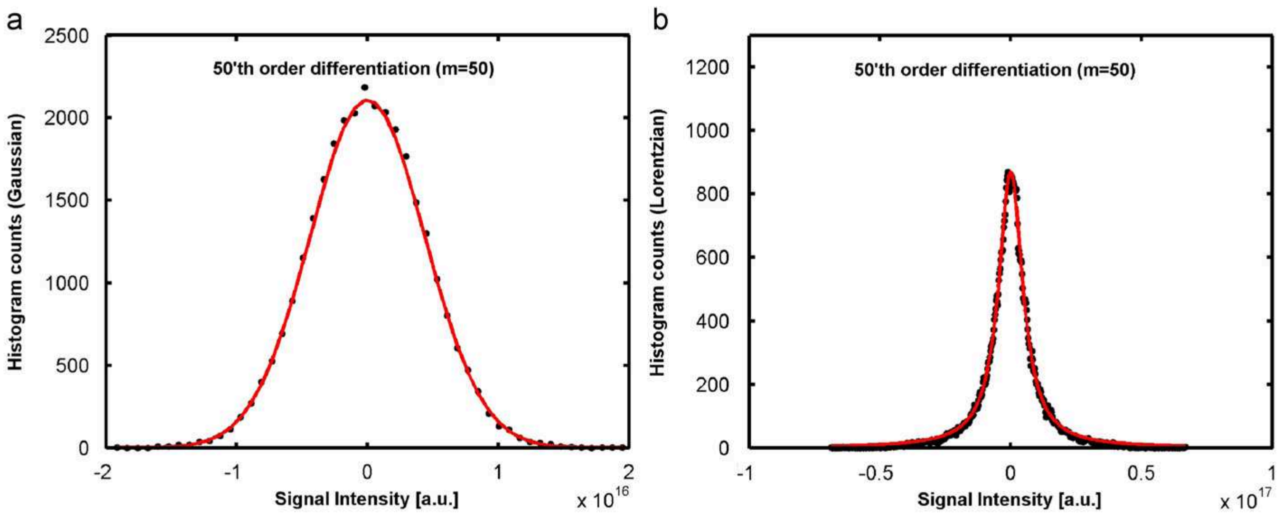

Fig. 6. The histogram of the noisy signal (Fig. 5), after fifty (50) steps of differentiation. Also shown are the Gaussian and Lorentzian fits obtained in the LMS sense (solid). The abbreviation [a.u.] in the figure indicates "arbitrary units".

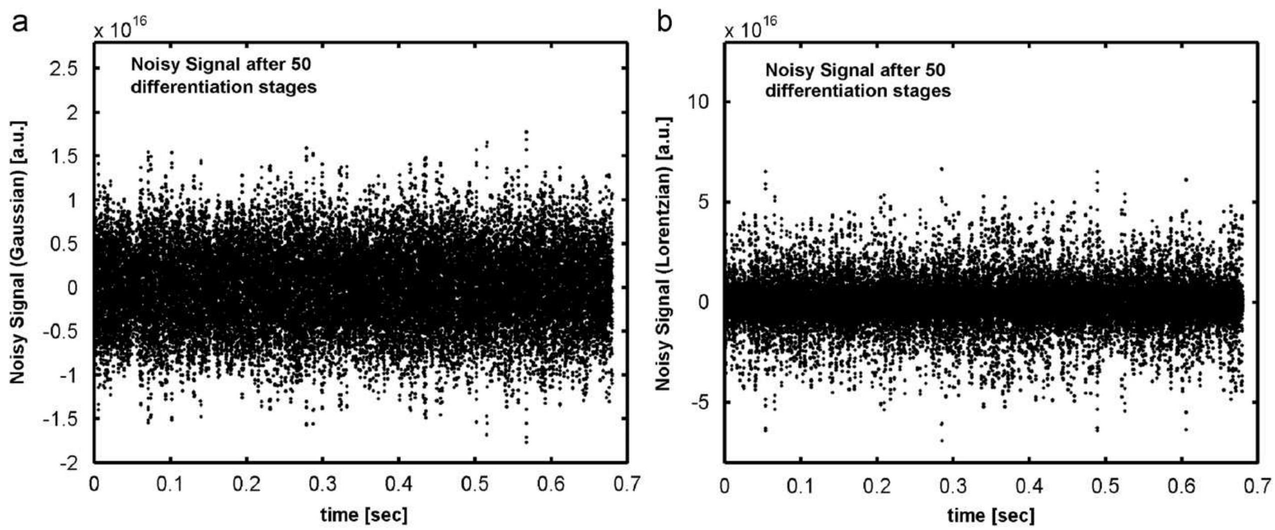

Fig. 7. The distribution of the differentiated signal after 50 steps of differentiation. The abbreviation [a.u.] in the figure indicates "arbitrary units".

index axis), allows the further numerical differentiation of the differentiation process with no need to consider any specific case (with respect to sampling rates) for the sake of the generality of the discussion.

5. Conclusions

We now wish to remark on two basic issues that are of importance for the scope of the proposed method. First, we address the restriction we intentionally set, on the pure signal, i.e., its smoothness and being subject to an m'th-degree polynomial fit. We note that this limitation is only marginal since as long as the fit quality is not worse than the noise variance, (i.e., the residual is within the statistics) this condition will not set severe practical limitations on the degree of the differential order m. Secondly, as was also mentioned above, although the assumption of independence of the individual elements in the stochastic data set is now only approximated, as the deterministic part may cause some dependence between the differentiated points, it can be shown that the effect is small and can surely be neglected as the order of differentiation increases. Also, note the assumption of uncorrelated, additive noise component in the original noisy signal.

In conclusion, we have derived a principle set of expressions for the probability density function of the m'th numerical differentiation of a stochastic variable. It was shown that the statistical characteristics can be obtained analytically based on the original probability parameters of the observed signal for the normal and Cauchy distribution, as a sum of independent, though non-identical, random variables, based on a simple weighting procedure of the original probability density function. We suggest that this allows the estimation of the statistical parameters of the original (assumed white), noise distribution parameters, such as the variance of the probability density function, and further, to simulate the noise contribution in the stochastic system so that

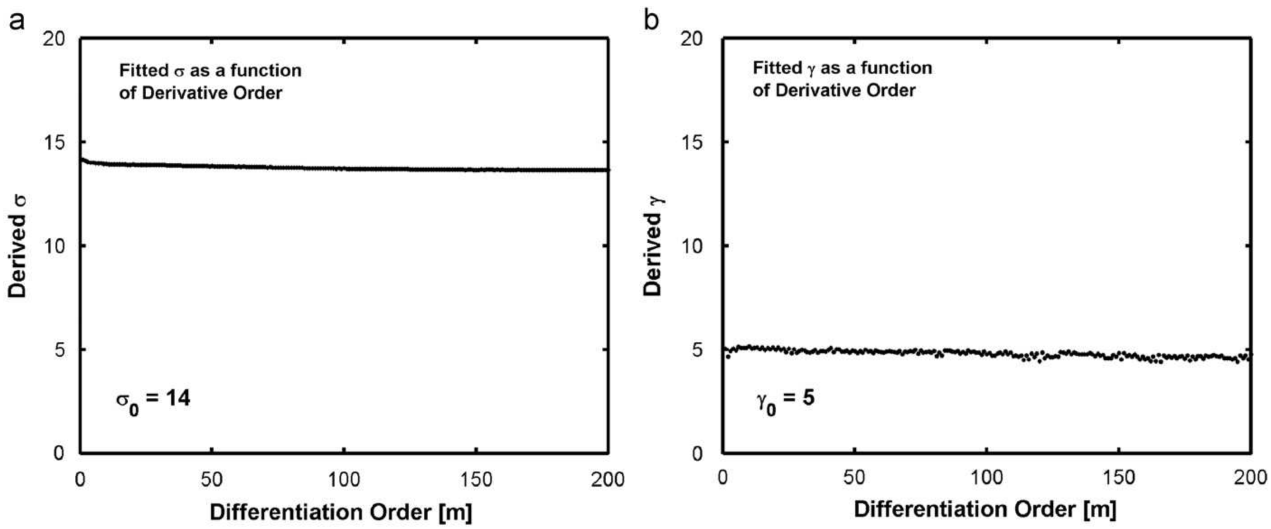

Fig. 8. The re-normalized values of the σ_m (a) and γ_m (b) (with m, the differentiation order) as extracted from all histograms for each of the 200 differentiation steps. Note that all values are close to the initial value of the original σ_0 and γ_0 correspondingly.

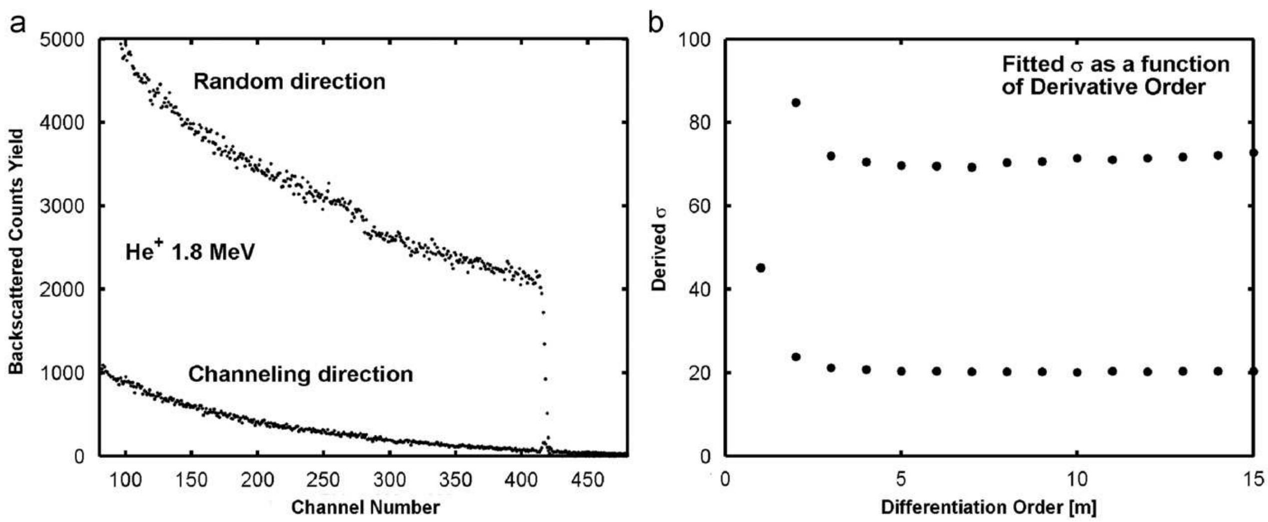

Fig. 9. Typical Rutherford Backscattering (RBS) spectra obtained with 1.8 MeV He^+ ion beam of a Fe-doped ⟨1 0 0⟩ InP substrate at random and channeling directions (a). Shown in (b) are the noise-level estimates as extracted by applying the HOSD method.

the noise component reconstructed is statistically indistinguishable from the true contribution of the noise in the originally observed data signal. Note that although we use the assumption that the deterministic component of the noisy signal is smoothed, which under the practical restriction that the signal is finite in time (as a time-series) and thus can always be approximated by a polynomial function, the order of such approximation is unknown for the general case. The proposed method eliminates any a-priori assumption of the order of the polynomial approximation as a higher degree of derivative can be used to still recover the noise component in a statistical manner. Both simulated signals and experimental spectra were used and yielded good results. Tests to compare the results with other methods to recover the noise component, show that the proposed method is very efficient from numerical perspective (similar to Savitzky–Golay filtering, however, with better performance), due to its direct process of the noise component.

It is interesting to note, that in contrast with the Gaussian distribution density function (and the Cauchy distribution, as discussed above), the general case of arbitrary distribution need not necessarily result in a density function similar to the original function after the differentiation of order m. This may simply be concluded if one experiments with a uniform probability density function which results to obtain a triangular shaped probability function even for m=1. However, recalling that the m'th order numerical differentiation is, in fact, a series of integral terms representing additions and subtraction of the corresponding probability distribution functions, each weighted according to the Stochastic-Derivative Matrix, the derivation of the respective expression may be straightforward. For instance, the probability density function of the first numerical differentiation of the exponential probability distribution function (i.e., Pearson Type X) $f(X) = \lambda e^{(-\lambda x)}$ for $x \geq 0$ is given by $f'(z) = (\lambda/2)e^{-\lambda z}(1-e^{-2\lambda z})$ for $z \geq 0$, i.e., a non-exponential *pdf*.

References

[1] J.P.M. Heald, J. Stark, Phys. Rev. Lett. 84 (4) (2000) 2366.
[2] N. Moriya, Nucl. Instr. and Meth. B53 (1991) 208.
[3] N. Moriya, L.C. Feldman, H.S. Luftman, C.A. King, J. Bevk, B. Freer, Phys. Rev. Lett. 71 (1993) 883.
[4] N. Moriya, L.C. Feldman, S.W. Downey, C.A. King, A.B. Emerson, Phys. Rev. Lett. 77 (1995) 1981.
[5] M.S. Grewal, A.P. Andrews, Kalman Filtering: Theory and Practice Using MATLAB, John Wiley & Sons, New York, 2001.
[6] J.B. Gao, Physica D106 (1997) 49.
[7] C.J. Cellucci, A.M. Albano, P.E. Rapp, R.A. Pittenger, R.C. Josiassen, Chaos 7 (1997) 414.
[8] J. Hu, J.B. Gao, K.D. White, Chaos Solitons Fractals 22 (2004) 807.
[9] J.B. Gao, Phys. Rev. E 63 (2001) 066202.
[10] J.B. Gao, Y.H. Cao, L.Y. Gu, J.G. Harris, J.C. Principe, Phys. Lett. A 317 (2003) 64.
[11] C. Rieke, K. Sternickel, R.G. Andrzejak, C.E. Elger, P. David, K. Lehnertz, Phys. Rev. Lett. 88 (2002) 244102.
[12] A. Provenzale, L.A. Smith, R. Vio, G. Murante, Physica D58 (1992).
[13] D.J. Yu, W.P. Lu, R.G. Harrison, Chaos 9 (1999) 865.
[14] R. Manuca, R. Savit, Physica D99 (1996) 134.
[15] M.B. Kennel, Phys. Rev. E 56 (1997) 316.
[16] T. Schreiber, Phys. Rev. Lett. 78 (1997) 843.
[17] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, J.D. Farmer, Physica D 58 (1992) 77.
[18] A. Papoulis, Probability, Random Variables, and Stochastic Processes, second ed., McGraw-Hill, New York, 1984.
[19] A. Papoulis, Signal Analysis, McGraw-Hill, New York, 1977.
[20] G.E.P. Box, G. Jenkins, Time Series Analysis, Forecasting and Control, Holden-Day Inc., San Francisco, 1970.
[21] P.F. Velleman, D.C. Hoaglin, Applications, Basics and Computing of Exploratory Data Analysis, Duxbury Press, Boston, Mass., 1981.
[22] J. Fan, I. Gijbels, Local Polynomial Modeling and its Applications, Chapman & Hall, London, 1996.
[23] W.S. Cleveland, J. Amer. Statist. Assoc. 74 (4) (1979) 829.
[24] J.F. Shortle, Stoch. Models 21 (1) (2005) 215.
[25] D. Loebis, R. Sutton, J. Chudley, W. Naeem, Control Eng. Practice 12 (2004) 1531 ([Online] Available: 12).
[26] R.K. Mehra, IEEE Trans. Autom. Control 17 (5) (1972) 693.
[27] M. Efe, J.A. Bather, D.P. Atherton, An adaptive Kalman filter with sequential rescaling of process noise, in: Proceedings of the American Control Conference, San Diego, CA, vol. 6, 1999, pp. 3913–3917.
[28] P.R. Belanger, Automatica 10 (3) (1974) 267.
[29] D.A. Forsyth, J. Ponce, Computer Vision: A Modern Approach, Prentice-Hall, Upper. Saddle River, N.J., 2002.
[30] B. Cyganek, J.P. Siebert, An Introduction to 3D Computer Vision Techniques and Algorithms, John Wiley & Sons Inc., United Kingdom, 2009.
[31] N. Moriya, Appl. Math. Model. 30 (9) (2006) 816.
[32] A. Savitzky, M.J.E. Golay, Anal. Chem. 36 (1964) 1627.
[33] W.H. Press, S.A. Teukolsky, W.T. Vettering, B. Flannery, Numerical Recipes in C: The Art Of Scientific Computing, Cambridge University Press, Cambridge, UK, 1992.