

RESEARCH

Open Access

Blind estimation of statistical properties of non-stationary random variables

Ali Mansour^{1*}, Raed Mesleh² and el-Hadi M Aggoune²

Abstract

To identify or equalize wireless transmission channels, or alternatively to evaluate the performance of many wireless communication algorithms, coefficients or statistical properties of the used transmission channels are often assumed to be known or can be estimated at the receiver end. For most of the proposed algorithms, the knowledge of transmission channel statistical properties is essential to detect signals and retrieve data. To the best of our knowledge, most proposed approaches assume that transmission channels are static and can be modeled by stationary random variables (uniform, Gaussian, exponential, Weibull, Rayleigh, etc.). In the majority of sensor networks or cellular systems applications, transmitters and/or receivers are in motion. Therefore, the validity of static transmission channels and the underlying assumptions may not be valid. In this case, coefficients and statistical properties change and therefore the stationary model falls short of making an accurate representation. In order to estimate the statistical properties (represented by the high-order statistics and probability density function, PDF) of dynamic channels, we firstly assume that the dynamic channels can be modeled by *short-term stationary but long-term non-stationary random variable (RV)*, i.e., the RVs are stationary within unknown successive periods but they may suddenly change their statistical properties between two successive periods. Therefore, this manuscript proposes an algorithm to detect the transition phases of non-stationary random variables and introduces an indicator based on high-order statistics for non-stationary transmission which can be used to alter channel properties and initiate the estimation process. Additionally, PDF estimators based on kernel functions are also developed. The first part of the manuscript provides a brief introduction for unbiased estimators of the second and fourth-order cumulants. Then, the non-stationary indicators are formulated. Finally, simulation results are presented and conclusions are derived.

Keywords: Wireless communication; Dynamic transmission channel; Higher-order statistics; Cumulants; Cumulative distribution function; Kernel density estimators; Hermite basis set; Spline functions; Characteristic functions; Non-stationary signals

1 Introduction

Unlike discrete transmission channels which have limited practical value, continuous random transmission channels are widely accepted and used to illustrate practical concepts [1]. In numerous proposed algorithms, statistical properties or probability density function (PDF) of transmission channels are assumed to be perfectly known or already estimated [2-6]. To analyze the performance of various wireless transmission schemes, Simon and Alouini in [7] have used channel statistical models and PDF properties. It is widely believed that wireless

transmission channels can be modeled using stationary random variables [1,8,9]. We should highlight the fact that various PDFs have been used in the context of wireless communications, such as Gaussian, Rayleigh, log-normal, exponential, one-sided Gaussian distribution, Hoyt, Weibull, Rice, Nakagami- m , $\alpha - \mu$, and $\eta - \kappa$ (see [10-21] and the references therein). It is worth pointing out that parametrical estimation methods can be used if the channel model (or a PDF for the real random variable (RV)) is selected or identified. Additionally, non-parametrical estimation methods can also be found in the literature [22-27]. However, these methods are hard to be implemented and mostly time consuming and they strongly relate to the proposed models and applications. For this main reason, a generic method and few required

*Correspondence: mansour@ieee.org

¹Lab STIC - ENSTA Bretagne, 2 Rue François Verny, Brest 29200, France
Full list of author information is available at the end of the article

assumptions about the transmission channel are proposed and discussed in this manuscript.

The recent spread of cellular systems (smart sensors, mobile phones, base stations, satellites, surveillance devices, traffic radars, etc.) has increased the complexity of processing algorithms as well as the model of transmission channels. In fact, in various applications, transceiver units are in motion. For such systems, the transmission channels can no longer be considered as static channels, i.e., they cannot be modeled by stationary random variables. To overcome this difficulty, most researchers assume that the transmission channels are static during the processing time^a [28-30].

In this manuscript, the proposed models are improved by considering that transmission channels are dynamic channels that can be modeled as non-stationary variables. It is worth mentioning that for antennas mounted on moving vehicles or drones, the transmission channels can be approximated by almost static or slowly evolutionary dynamic channels over certain periods. The transition among these periods can be very fast. In addition, during two adjacent periods, transmission channels can have completely different statistical properties and/or PDFs. This scenario illustrates a typical scenario of a mobile phone used in a moving car which is moving among big buildings in a modern city. It can represent as well the scenario of a drone flying at low altitude in a mountainous area.

In what follows, we consider that transmission channels can be modeled as short-term stationary signals and long-term non-stationary signals (many natural and physical signals belong to this category, such as speech signals [31,32]). Herein, we develop an algorithm to accurately estimate the transition times. The algorithm is based on high-order statistics (HOS). In several study cases, HOS are shown to be more promising than the second-order stochastic methods, namely, power, variance, covariance, and spectra. In fact, HOS have been used to solve many recent and important telecommunication problems [33,34], such as blind identification or equalization, blind separation of sources, and time delay estimation [35-38]. It is worth mentioning that most of these HOS algorithms are only based on the second- and the fourth-order statistics [39]. Once the transition times are estimated, we should be able to estimate the PDF of the local stationary random variable (i.e., during a short-term period).

It is well known that any random variable can be completely described by its PDF [40]. In many cases, an accurate estimation of the PDF of physical parameters of interest (which are random variables) is essential to achieve our goals. This is while an accurate estimation of PDFs is still challenging to researchers. The estimation of PDFs is a relatively old problem that has been

considered since the beginning of the last century with the rise and the development of new and modern complex systems (radars, sonar, and wireless communication systems). However, in late 1950s, systematic mathematical approaches have been proposed. One of the pioneering work in this field is the work presented by Rosenblatt in [41]. Parzen in [42] formulated the estimation of PDFs using kernel approaches. Later on, new approaches emerged to overcome the specification of diverse applications such as the sum of Gamma densities in Risk Theory [43] or the sum of exponential random variables in wireless communication [44]. Other researchers have focused on the choice of the kernel and smoothing functions [45,46]. The estimation of PDFs using orthogonal series has been introduced by Schwartz in [47]. In more recent work, wavelets have been used as a non-parametric estimation of PDFs [48]. Similar to previous work, Engel used Haar's series to estimate PDFs [49]. Using autoregressive (AR) models, Kay in [50] proved that PDFs can be estimated by appealing the theory of power spectral density (PSD). More recent work reintroduced the estimation of PDF using a windowed Fourier transform [51] or Hermite's orthonormal basis set [52]. We should mention that all the abovementioned work assume that the random variables are stationary. In countless applications, the stationarity assumption may be invalid [53]. To the best of our knowledge, there is no such PDF estimator for non-stationary random variables. Hence, it is the aim of this paper to propose a PDF estimator based on a smooth kernel PDF estimator that assumes known transition times.

2 Mathematical model and background

In this paper, we consider the case of real non-stationary random variables^b. We also assume (1.1 in Appendix 1) that random variables consisting of transmission channels are stationary by part [54,55]. While this assumption is less strict than the most widely used assumption of stationary transmission channels, its consideration is justified in the case of moving receivers or transmitters.

Hereinafter, $\Pr(A)$ stands for the probability of the event A . Let x_1, \dots, x_n stand for n independent realization of a continuous RV X . In addition, let us denote by $f_X(x)$, $F_X(x)$, $\hat{f}_X(x)$, and $\hat{F}_X(x)$ respectively as the PDF, the cumulative distribution function (CDF) of X , and their estimated functions:

$$F_X(x) = \Pr(X \leq x) = \int_{-\infty}^x f_X(x)dx. \quad (1)$$

In the following, $k(x)$ denotes a kernel function of x , and the mathematical expectation (i.e., the mean) of a real RV X is denoted by

$$m_X = E\{X\} = \int_{\mathbb{R}} x f_X(x)dx, \quad (2)$$

where E stands for the mathematical expectation. By definition [56], the q th-order moment μ_q of a stochastic signal X is

$$\mu_q = E(X^q) = \int_{\mathbb{R}} x^q f_X(x) dx. \quad (3)$$

We mentioned earlier that a real random variable X is completely described by its PDF $f_X(x)$. In addition, a random variable can also be described by its first $\Phi_X(u)$ or second $\Psi_X(u)$ characteristic functions as follows:

$$\Phi_X(u) = E\{e^{jXu}\} = \int_{\mathbb{R}} e^{jxu} f_X(x) dx. \quad (4)$$

$$\Psi_X(u) = \ln(\Phi_X(u)), \quad (5)$$

where $j^2 = -1$. While the second characteristic function can be directly obtained from the first one, the first one is considered as the inverse Fourier transform of the PDF. Theoretically, PDFs may be obtained by the Fourier transform of $\Phi_X(u)$. The Taylor series of the two characteristic functions yield other important statistical functions which are the moments and the cumulants [56]:

$$m_k(X) = E\{X^k\} = (-j)^k \left. \frac{\partial^k \Phi_X(u)}{\partial x^k} \right|_{u=0}, \quad (6)$$

$$\text{Cum}_k(X) = (-j)^k \left. \frac{\partial^k \Psi_X(u)}{\partial x^k} \right|_{u=0}. \quad (7)$$

According to the theorem of Leonov-Shiryayev [56], the r th-order cumulant of X can be calculated from its moments, using the following formula^c [53,57]:

$$\begin{aligned} \text{Cum}_r[X] &= \text{Cum}[X, \dots, X] \\ &= \sum (-1)^{k-1} (k-1)! E[X^{v_1}] E[X^{v_2}] \dots E[X^{v_p}]. \end{aligned} \quad (8)$$

Using this relationship, we can evaluate the fourth-order cumulant of X as

$$\begin{aligned} \text{Cum}_4[X] &= E[X^4] - 4E[X]E[X^3] - 3E^2[X^2] + 12E^2[X]E[X^2] \\ &\quad - 6E^4[X]. \end{aligned} \quad (9)$$

For a zero-mean stochastic signal (please see [58] and the cited references therein),^d the second-order cumulant is equal to the second-order moment and therefore the fourth-order cumulant becomes

$$\text{Cum}_4[X] = E[X^4] - 3E^2[X^2]. \quad (10)$$

The HOS are much easier to be estimated than PDF, as will be shown in the next section. For this reason, our non-stationary transition indicator is based on the second-

and fourth-order statistics. In fact, the transition of non-stationary random variable can generate discontinuity in its HOS. As the variance of any signal should be different than zero, then it can be used to identify the existence of signals. On the other hand, normalized signals have the same unit variance. Therefore, the variance is not enough to identify the transition in a general case. Besides that, the fourth-order cumulants of Gaussian signals are zeros. For these reasons, we developed a non-stationary transition indicator based on the variance and the fourth-order cumulant.

3 Unbiased and adaptive HOS estimators

Let X be a zero-mean stochastic signal where x_i is an event (or a signal sample) of X ($1 \leq i \leq N$). The classic estimator of the r th-order moment of X is given by

$$\hat{\mu}_r = \frac{1}{N} \sum_{i=1}^N x_i^r. \quad (11)$$

It is easy to verify that (11) is an unbiased estimator of the r th-order moment of X (i.e., $E[\hat{\mu}_r] = \mu_r$). To estimate the fourth-order cumulant of X , we can derive an estimator from the Leonov-Shiryayev formula (8):

$$\widehat{\text{Cum}}_4[X] = \hat{\mu}_4 - 3\hat{\mu}_2^2. \quad (12)$$

It is established [59,60] that the estimator in (12) is a biased estimator and the estimation error decreases proportional to $\frac{1}{N}$. In fact, using (11) and (12) we can write

$$\widehat{\text{Cum}}_4[X] = \hat{\mu}_4 - 3\hat{\mu}_2^2 = \frac{1}{N} \sum_{i=1}^N x_i^4 - \frac{3}{N^2} \sum_{i,j=1}^N x_i^2 x_j^2, \quad (13)$$

and the estimator expectation becomes

$$\begin{aligned} E[\widehat{\text{Cum}}_4[X]] &= \mu_4 - \frac{3}{N} (\mu_4 + (N-1)\mu_2^2) \\ &= \text{Cum}_4[X] - \frac{3}{N} (\text{Cum}_4[X] + 2\mu_2^2). \end{aligned} \quad (14)$$

Using the Leonov-Shiryayev formula, one can develop an unbiased estimator for the cumulant as^e

$$\widehat{\text{Cum}}_4[X] = \frac{N+2}{N(N-1)} \sum_{i=1}^N x_i^4 - \frac{3}{N(N-1)} \sum_{i,j=1}^N x_i^2 x_j^2. \quad (15)$$

It is worth mentioning that the above estimators are only unbiased for stationary signals. In case of non-stationary signals, adaptive estimators should be developed.

Let $\widehat{\mu}_r\{k\}$ be the estimator of the r th-order moment at the k th iteration, and we can develop the following adaptive estimators for the second- and the fourth-order statistics:

$$\begin{aligned}\widehat{\mu}_2\{k\} &= \frac{(k-1)\widehat{\mu}_2\{k-1\} + x_k^2}{k}, \\ \widehat{\mu}_4\{k\} &= \frac{(k-1)\widehat{\mu}_4\{k-1\} + x_k^4}{k}, \\ \widehat{\text{Cum}}_4[X]\{k\} &= \frac{k+2}{k-1}\widehat{\mu}_4\{k\} - \frac{3k}{k-1}(\widehat{\mu}_2\{k\})^2,\end{aligned}\quad (16)$$

for $1 < k \leq N$. This algorithm is simple and converges quickly in the case of stationary signals. Another adaptive estimator of a fourth-order cross-cumulant ($\text{Cum}_{13}(X, Y)$) which is more suitable for non-stationary signals is given by

$$\begin{aligned}C_N &= \frac{N-2}{N}\gamma C_{N-1} + \frac{1}{N}\gamma \mu_{13}(X, Y) + \frac{N+2}{N(N-1)}\gamma x_N y_N^3 \\ &\quad - 3\gamma x_N y_N \mu_{02}(X, Y) - 3\gamma y_N^2 \mu_{11}(X, Y) + (1-\gamma)x_N y_N^3 \\ &\quad - 3(1-\gamma)x_N y_N \mu_{02}^2,\end{aligned}\quad (17)$$

where $0 < \gamma < 1$ is another forgetting factor.

4 PDF estimators

To get an accurate estimation of $f_X(x)$, one should use available observations or measurements (i.e., x_1, \dots, x_n). Using Fubini's theorem [61]^f, Rosenblatt in [41] proved that any estimator $\hat{f}_X(x)$ of $f_X(x)$ based on the set of x_i , $\{i = 1, \dots, n\}$, is biased. However, Parzen in [42] proved that the $F_n(x)$ is an unbiased estimator $\hat{F}_X(x)$ of the CDF, i.e., $E\{F_n(x)\} = F_X(x)$:

$$F_n(x) = \frac{1}{n} \text{Card}(x_i \leq x),$$

where Card denotes the cardinal of a σ -algebra set. Using the fact that $f_X(x) = \frac{dF_X(x)}{dx}$, Parzen proposed an asymptotically unbiased estimator of the PDF as the sum of scaled and shifted kernel functions:

$$f_n(x) = \int_{-\infty}^{\infty} \frac{1}{h} k\left(\frac{x-y}{h}\right) dF_n(y) \approx \frac{1}{nh} \sum_{i=1}^n k\left(\frac{x-x_i}{h}\right),$$

given that $\lim_{n \rightarrow \infty} h(n) = 0$ and the kernel $k(x)$ and $xk(x)$ are Borel's functions [62]^g in L1 function space, which satisfy few conditions such as [41,47]

$$\int_{\mathbb{R}} k(x) dx = 1.$$

In [42], Parzen suggests few kernel functions including histograms which are kernel estimators when

$$k(x) = \frac{u(x+1) - u(x-1)}{2},$$

with $u(x)$ being the Heaviside step function.

Using the definition in (4), the PDF $f_X(x)$ can be approximated by using the Fourier transform of an estimated characteristic function [53]. In fact, the average over a set of realization is an unbiased estimator of the mathematical expectation of a RV. Therefore, the first characteristic function can be approximated using the following equation:

$$\hat{\Phi}_X(u) = \frac{1}{n} \sum_{i=1}^n e^{j2\pi u x_i}. \quad (18)$$

In [51], the authors mentioned that the Fourier transform of the last equation cannot converge. However, it is possible to obtain the Fourier transform of the product between $\hat{\Phi}_X(u)$ and a weighting window $w(x)$. In this case the PDF estimator is given by

$$\hat{f}_X(x) = \int_{\mathbb{R}} \hat{\Phi}_X(u) w(u) e^{-j2\pi u x} du \approx \frac{1}{n} \sum_{i=1}^n w(x - x_i).$$

The above equation can be considered as a kernel estimator. Window length and parameters are the key factors which can affect the performance of the estimator [51]. To avoid the problem of this approach and apply the well-established theory of power spectral density, Kay in [50] proposed an AR model of order p to estimate the PDF. Parameters of his AR model are estimated using Yule-Walker equations and Levinson recursion.

Using Gram-Charlier series to expand the PDF in terms of the normal density and its derivatives, Bowers in [43] approximated a risk theoretic distribution by the sum of Gamma RVs. In his expansion, the polynomials which multiply the PDF of a zero-mean normalized normal PDF are the Hermite polynomials (see Appendix 2). This study was generalized by Schwartz in [47].

According to Vannucci [24], there are different types of non-parametric density estimators: delta sequence estimators (such as kernel, histogram, and orthogonal series estimators) and penalized maximum likelihood estimators such as spline estimators. She also proved that scaling functions give a good approximation of smooth functions, while wavelets^h can deal with functions which have local fluctuations. In order to practically apply wavelet estimators, a truncation should be made [24]. Therefore, the overall performance depends on the truncation order. An optimal choice of such order can be determined by minimizing the integrated square error or the intergraded

mean square error. In real-world applications, another problem can arise related to the evaluation of the scaling and the mother functions at arbitrary points [24]. In [49], Engel used Haar's functions to estimate the PDF and shows that his estimator is equivalent to a histogram on certain dyadic intervals.

Recently, the authors of [51,52] have compared their approaches to the abovementioned ones. In fact, Xie and Wang in [51] showed that an estimator based on the Fourier transform of the characteristic function gives similar results to the AR model PDF estimator proposed earlier by Kay in [50]. They also showed that in certain cases, histograms can provide similar overall performance to previously mentioned estimators. However, Howard in [52] showed that the Hermite basis estimator can slightly achieve better results than the estimator based on the Fourier transform of the characteristic function proposed in [51]. For these reasons, we only consider the histogram, the Hermite basis estimator and the smooth kernel density estimator proposed by Bowman in [45,46]. Our simulations have shown that smooth kernel density has slightly better performance than the previous mentioned ones (see Section 7).

Finally, two useful error measures are proposed to evaluate the overall performance of various estimators [41,47,51,52]:

- The integrated error

$$\epsilon_1 = E\{|f_X(x) - \hat{f}_X(x)|\}$$

- The root of the mean-squared error

$$\epsilon_2 = \sqrt{E\{|f_X(x) - \hat{f}_X(x)|^2\}}$$

5 Representation of non-stationary PDF

Real-world signals can be modeled by wide or quasi non-stationary random variables [39,63]. In this case, the PDF is a time-dependent function $f_X(x, t)$ where t can be a vector representing discrete time instants. Let us consider for example a wireless transmission channel $h(t)$ which can be modeled by a quasi stationary RV [1]. Let us assume now that $h(t)$ has a Weibull PDF for the first period, a uniform PDF for the second period, then a Gaussian RV in its last parts. In this case, it becomes clear that the PDF of $h(t)$ cannot be represented by simple curve as function of t or X . Hence, the PDF of non-stationary RV should be plotted as a function of a RV X and t (see Figure 1).

In fact, classic PDF estimators cannot provide an accurate estimation of the non-stationary signal PDF. By applying a smooth kernel density estimator over 10,000 independent realization of $h(t)$, a bimodal PDF was obtained (see Figure 1b). It is clear that the obtained PDF cannot correctly represent the PDF shown in Figure 1a.

6 Non-stationary transition indicator

In order to estimate the PDF parts of quasi non-stationary RV, one can easily use the estimator developed in the previous section to detect the number and the size of the stationary parts. Once these parts are well identified and their sizes are relatively enough to estimate a PDF, we can apply any classic PDF estimator. To make this estimation more robust, cumulants of different order (mainly the second and the fourth order) can be used to identify such parts. In fact, Gaussian signals are characterized by their zero fourth-order cumulant (see Appendix 1). In these case, the fourth order cannot separate two adjacent Gaussian parts. However, if these two parts represent two normal RVs with different means or variances, then the second or the first moment can be used to identify these parts. If the two Gaussian parts have the same mean and variance then they can describe the same RV due to the basic assumptions about the whiteness of the samples (i.e., the realization of X).

Hereinafter, we consider that the non-stationary RV is formed by successive segments of stationary RV. In order to estimate the PDF of the non-stationary signal, an estimator of the transition times should be developed. In order to simplify our discussion and gain insight, a generic case is considered. Let us consider a zero-mean non-stationary signal $X(t)$ made by four parts of stationary random variables as shown in Figure 2:

1. The first part contains 8,000 samples of uniform random variable included in $[-1, 1]$.
2. The second part is made of 6,000 samples of zero-mean and unite variance Gaussian signal.
3. The 10,000 samples of a uniform random variable in $[-2, 2]$ formed the third part.
4. The fourth part is a zero-mean Gaussian signal with a standard deviation of $\sqrt{2}$.

The fourth-order statistics of the uniform parts are given as follows:

$$\begin{aligned} E[X^2] &= \frac{A^2}{3}, \\ E[X^4] &= \frac{A^4}{5}, \\ \text{Cum}_4[X] &= \frac{-2A^4}{15}, \end{aligned} \quad (19)$$

where A is the maximum amplitude of X . Figure 3 shows that the proposed adaptive estimators can yield a good estimation of the HOS. However, these estimators are not quite enough to estimate the HOS and the transition times with a wide accuracy range. In fact, the estimators shown in Figure 3 can be considered as

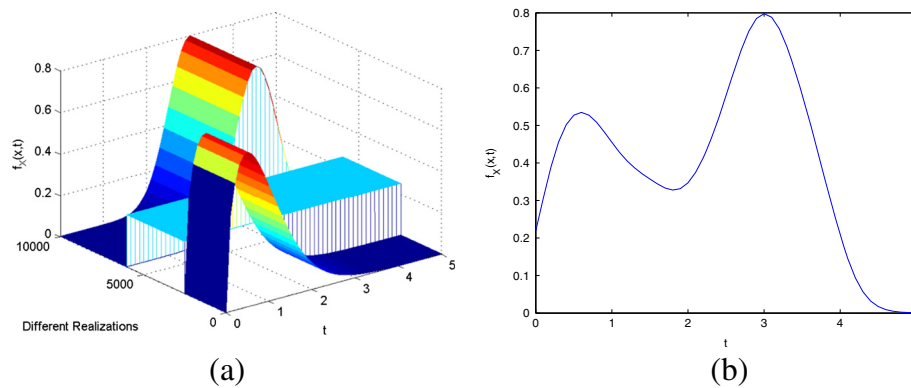


Figure 1 Theoretical and estimated PDF of non-stationary RV. The estimation is done using smooth kernel density estimators. **(a)** Theoretical PDF. **(b)** Estimated PDF.

noisy signals. To reduce the noise level, we adopted two steps:

1. First, the obtained signals are filtered using the smoothing polynomial regression filter proposed by Savitzky-Golay [64-66]. The main idea of the Savitzky-Golay filter is to apply a FIR filter such that its coefficients minimize the mean-squared approximation error. This filter is used to filter biomedical signals such as EEG signals (see Figure 4).
2. Second, the noise in the filtered signals are slightly reduced using a Walsh-Hadamard transform which is a generalization of discrete Fourier transform [67]. This transformation is also used in biomedical signals to filter and compress ECG signals (see Figure 5).

In order to smooth the HOS estimator and achieve better detection of the transition times, the signal of the HOS estimator was firstly filtered using a Savitzky-Golay filter then a Walsh-Hadamard transform truncation filtering techniques is applied. Experimental results showed

that the filtered signals are much more smooth and useful to reach our goal (see Figure 6).

While the filtered signals are much better smoothed than the raw signals, they are still relatively noisy to obtain an accurate estimate of transition time. In fact, Figure 6 shows that the obtained filtered estimators are still suffering from the previously mentioned drawbacks, i.e., the inaccuracy in the estimation of the HOS and the transition times.

To address the previously mentioned problems, we developed the following algorithm:

- First, the PDF of the filtered HOS estimators are obtained using the kernel PDF estimator (see Figure 7). Then the maximum values of the obtained PDF are recorded as the coefficients of a vector called maxHOS.
- The coefficients of the vector maxHOS are introduced as the center of clusters. By minimizing Euclidean distance among the samples of the filtered HOS estimator, all samples of filtered estimated

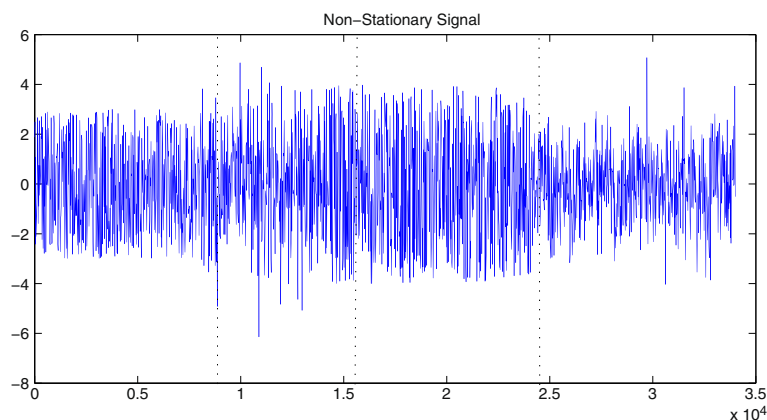
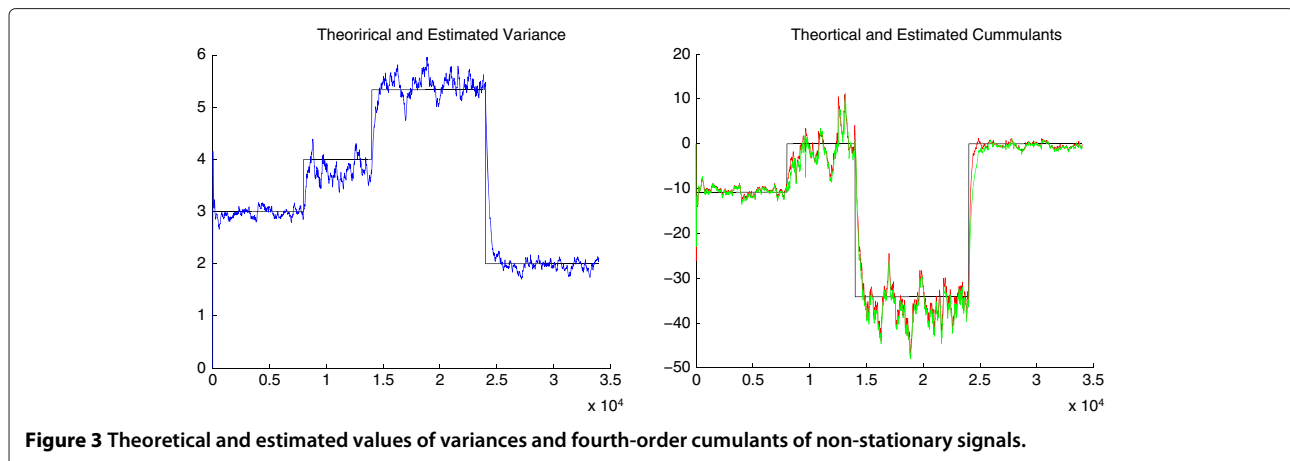


Figure 2 The considered non-stationary signal.

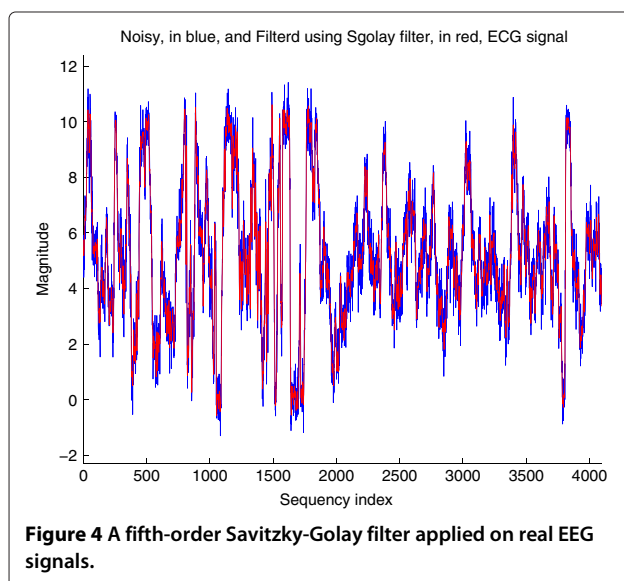


cumulants are clustered in rectangular signals which represent theoretical values of the HOS (see Figure 8).

- Various simulations have been conducted. Our simulations showed that sometimes the rectangular clustered signals suffer from local narrow spurious error windows (see Figure 8). These spurious error windows are normally very narrow. Hence, by supposing that the channel is not a highly dynamic one, i.e., any channel parameter cannot change more than one time in a short period (for example, the short period can be the symbol duration). In this case, one can easily eliminate these windows. In order to clean out the clustered rectangular signals and achieve an accurate non-stationary indicator, the derivative of these signals is evaluated (see Figure 9).
- Let us assume 1.2 in Appendix 1 that the stationary parts of the channels have more than 1,000 samples (in many cases, 500 samples were enough to obtain

good results). In this case, each of local narrow spurious error windows generates two Dirac delta functions with equal values and opposite signs. Based on 1.2 in Appendix 1, the two delta functions should be close to each other within 1,000 samples. Using this fact, we developed and implemented a recursive filtering procedure called continuity process (CP) to eliminate these spurious windows (see Figure 10). The Matlab code is given in Appendix 3.

- Figure 10 shows that the impact of the local narrow spurious error windows is completely eliminated. However, our non-stationary indicator still suffers from two-step transition problem, i.e., the transition between two valid states of the cluster rectangular signals is not immediate (this case can be shown in Figure 8 around the 14,000th and 24,000th samples). Using 1.2 in Appendix 1 and the fact that the two-step transition problem generates two Dirac delta functions with different values but sharing the same sign, we developed another process called two-step transition process to deal with this problem (see Appendix 3).
- Finally, a clear and accurate non-stationary indicator is obtained. Our final indicator is the output of an 'OR' gate applied on the non-stationary indicator of various HOS (in our case, we used two HOS, i.e., the second and the fourth cumulants). The final indicator is shown in Figure 11. In Figure 12, the synoptic of our proposed algorithm using smoothing filters and nearest-neighbor clustering algorithm is presented.



7 PDF estimator for non-stationary signals

In the previous section, the transition between two stationary parts has been well identified. Hereinafter, we assumed that the different stationary segments of the signal are already known and just the PDF of the non-stationary signal should be estimated. To evaluate their algorithms, the authors of [50-52] generate a RV X with

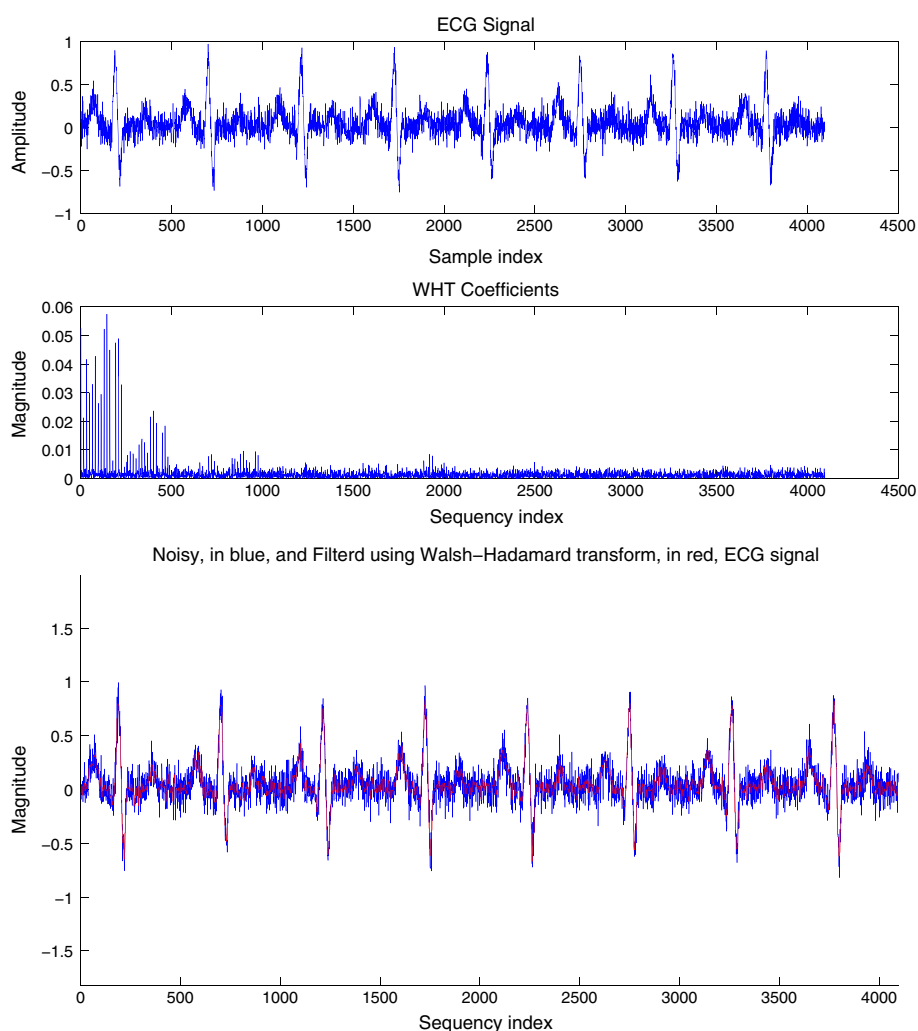


Figure 5 A Walsh-Hadamard transform applied on simulated ECG signals. Noisy, in blue, and filtered using Walsh-Hadamard transform; in red, ECG signal (top). ECG signal and WHT coefficients (bottom).

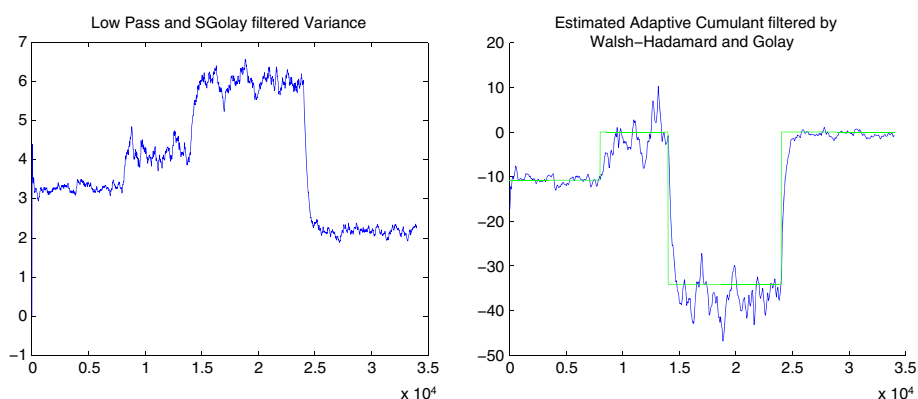


Figure 6 Estimated variance and fourth-order cumulant filtered by a Savitzky-Golay filter and a truncation of its Walsh-Hadamard transform. Low-pass and Savitzky-Golay filter variance (left). Estimated adaptive cumulant filtered by Walsh-Hadamard and Golay (right).

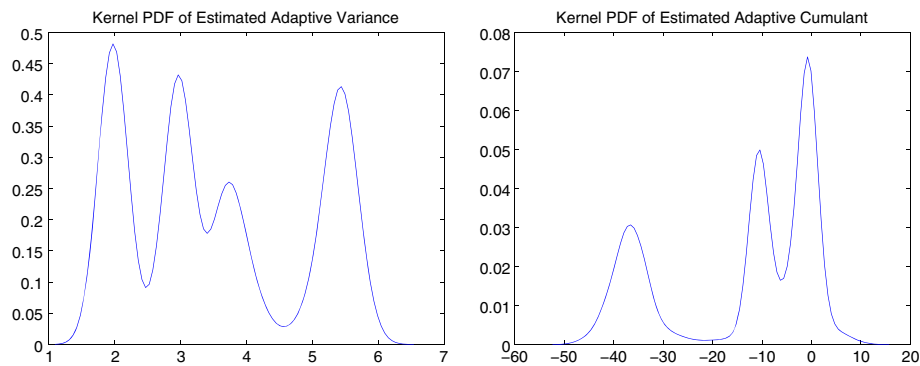


Figure 7 Kernel PDF estimator of filtered variance and adaptive cumulant. Kernel PDF of estimated adaptive variance (left). Kernel PDF of estimated adaptive cumulant (right).

its PDF as a moderated sum of PDF of two uncorrelated normal RVs:

$$f_X(x) = \sum_{i=1}^2 \frac{p_i}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(x-m_i)^2}{2\sigma_i^2}\right) \quad (20)$$

where m_i and σ_i are, respectively, the mean and the standard deviation of the normal RV X_i and $p_1 + p_2 = 1$ are weighting parameters. It was mentioned in [52] without any proof that the previous PDF shown in equation (20) 'is consistent with a random sum of independent RVs':

$$X = \sum_{i=1}^M X_i \quad (21)$$

where, for example, $M \in 1, 2$, and $\Pr[M = 1] = p_2$. We proved (see Appendix 4) that the sum should be just over two independent RV and we give the statistical properties of the newly obtained RV.

It is well known that the overall performance of a histogram depends on the number of samples as well as the theoretical PDF. Performances can deteriorate if the number of samples is not large enough. On the other hand, if

realization can be repeated many times, then an accurate estimation of the PDF can be obtained as the average of all obtained histograms (in Figure 13b, the average is done using ten iterations). For a large number of samples, the histogram performs quite well (see Figure 13c).

Our simulations show that the Hermite basis set estimator suffers from two major inconveniences:

1. It depends on the value of k_s ,
2. It can generate a negative function, see Figure 14.

The smooth kernel estimator proposed in [46] seems to overcome the previous two mentioned drawbacks (see Figure 15).

8 Conclusions

In this manuscript, a transition indicator for non-stationary signals is presented. The new indicator is based on HOS of quasi non-stationary variables (the random variables are considered stationary by parts). To estimate the HOS, unbiased adaptive exponential estimators are presented.

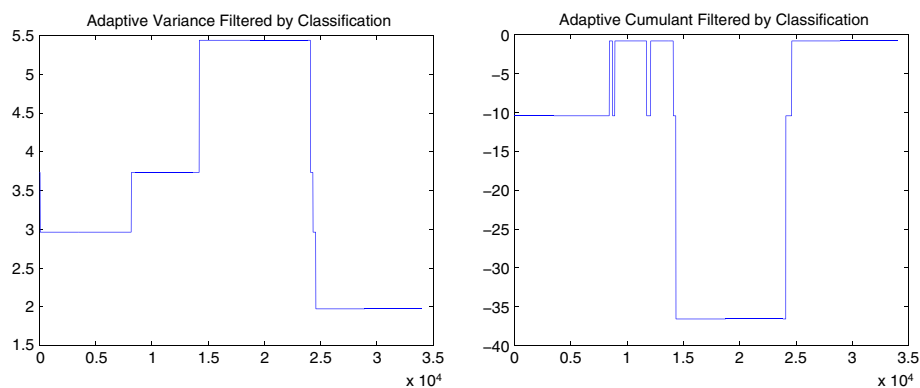
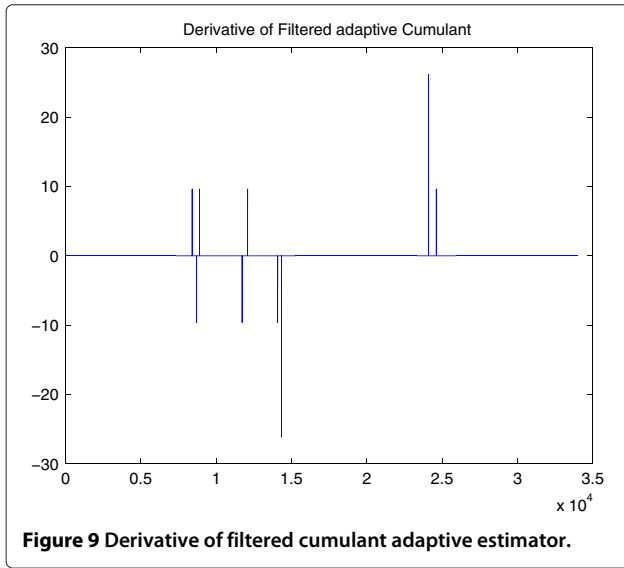
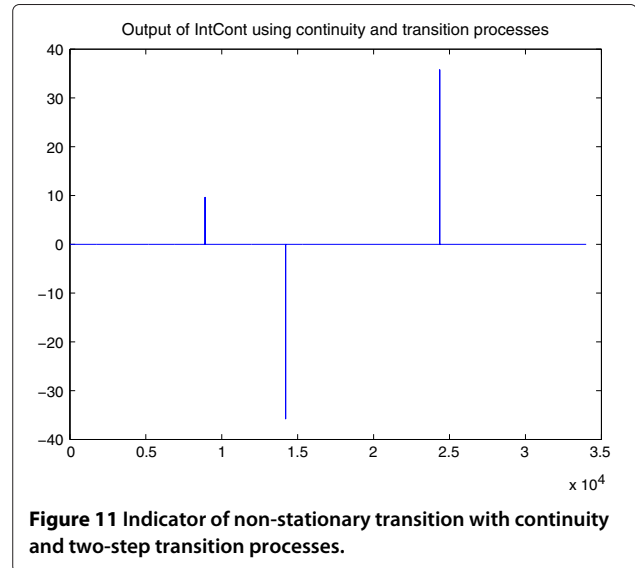
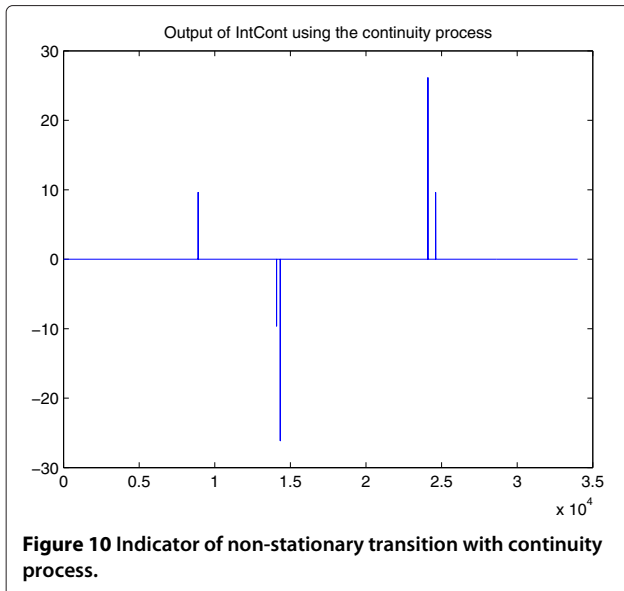


Figure 8 Clustering algorithm applied on estimated variances filtered by a Savitzky-Golay filter. Adaptive variance filtered by classification (left). Adaptive cumulant filtered by classification (right).



To reduce the noise level of the HOS estimators, we used a cascade filtering procedure based on Savitzky-Golay filters and a truncation of its Walsh Hadamard transform. Then rectangular signals representing the theoretical values of HOS are obtained by using clustering algorithm using the maxima of kernel PDF estimator and minimizing Euclidean distances.

Simulation studies show that the obtained rectangular signals can suffer from local narrow spurious error windows which can be eliminated using a continuity assumption 1.2 in Appendix 1 and a continuity cleaning procedure called continuity process. In addition to these local narrow spurious error windows, estimated signals suffer another artifact called the two-step transition problem. After solving this problem using assumption 1.2



in Appendix 1 and a continuity procedure, an accurate transition indicator of non-stationary signals is achieved. Simulation studies corroborate the performance of our proposed algorithm and the accuracy of our non-stationary transition indicator.

Finally, a survey of major PDF estimators is done. A comparative study is also presented and discussed. The advantages and drawbacks of major methods are highlighted and a theoretical study is provided. Simulation results show a slight advantage of smooth kernel estimator methods. It is worth mentioning that the histogram with a large number of samples is still one of the simplest and efficient estimators. The case of non-stationary process was considered and a PDF estimation approach was discussed.

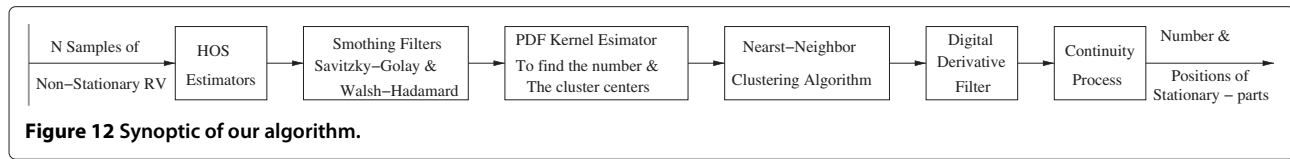
Endnotes

^aHowever, the processing time is not standardized. In fact, different authors claim that the parameters of the channel should remain constant during one frame duration, few hundred symbols, or during the convergence time of their adaptive algorithms.

^bThe fact that the channel is considered as a real channel is not limiting our approach, as a complex Gaussian channel could be represented by its modulus as a Rayleigh channel.

^cThe original formula shows the relationship among the cumulant of r stochastic signals X_i ($i = 1, \dots, r$) and their moments of order p , $p \leq r$:

$$\text{Cum}[X_1, \dots, X_r] = \sum (-1)^{k-1} (k-1)! \times E \left[\prod_{i \in v_2} X_i \right] E \left[\prod_{j \in v_2} X_j \right] \dots E \left[\prod_{k \in v_p} X_k \right], \quad (22)$$



where the addition operation is over all the set of v_i ($1 \leq i \leq p \leq r$) and v_i constitutes a partition of $\{1, \dots, r\}$, [40].

^dIn many applications, the stochastic signal X is a zero-mean signal. This assumption is not a major one and it has been used in many studies (please see [58] and the cited references therein). In fact, by considering this assumption, the mathematical notations can be simplified. However, all the proposed steps can be straightforwardly derived in the case of a non-zero mean transmission channel. Besides that, the mean can be estimated and canceled out from the other equations.

^eFurther details are given in Appendix 1.

^fAccording to Fubini's theorem [61], a double integral defined over two measure spaces $\mathcal{D}_a, \mathcal{D}_b$ of a measurable function $f(x, y)$ can be computed using iterated integrals:

$$\begin{aligned} \iint_{\mathcal{D}_a, \mathcal{D}_b} f(x, y) d(x, y) &= \int_{\mathcal{D}_b} \left(\int_{\mathcal{D}_a} f(x, y) dx \right) dy \\ &= \int_{\mathcal{D}_a} \left(\int_{\mathcal{D}_b} f(x, y) dy \right) dx. \end{aligned}$$

^gA function $f(x)$ is called Borel's function if $\forall y \in \mathcal{A}$ is an open set, the inverse of $f(x)$, and $x = f^{-1}(y)$ is an element of a Borel's set, i.e., Lebesgue measurable set [62].

^hWavelets are basis functions which have quite interesting properties such as their localization in space and frequency [24].

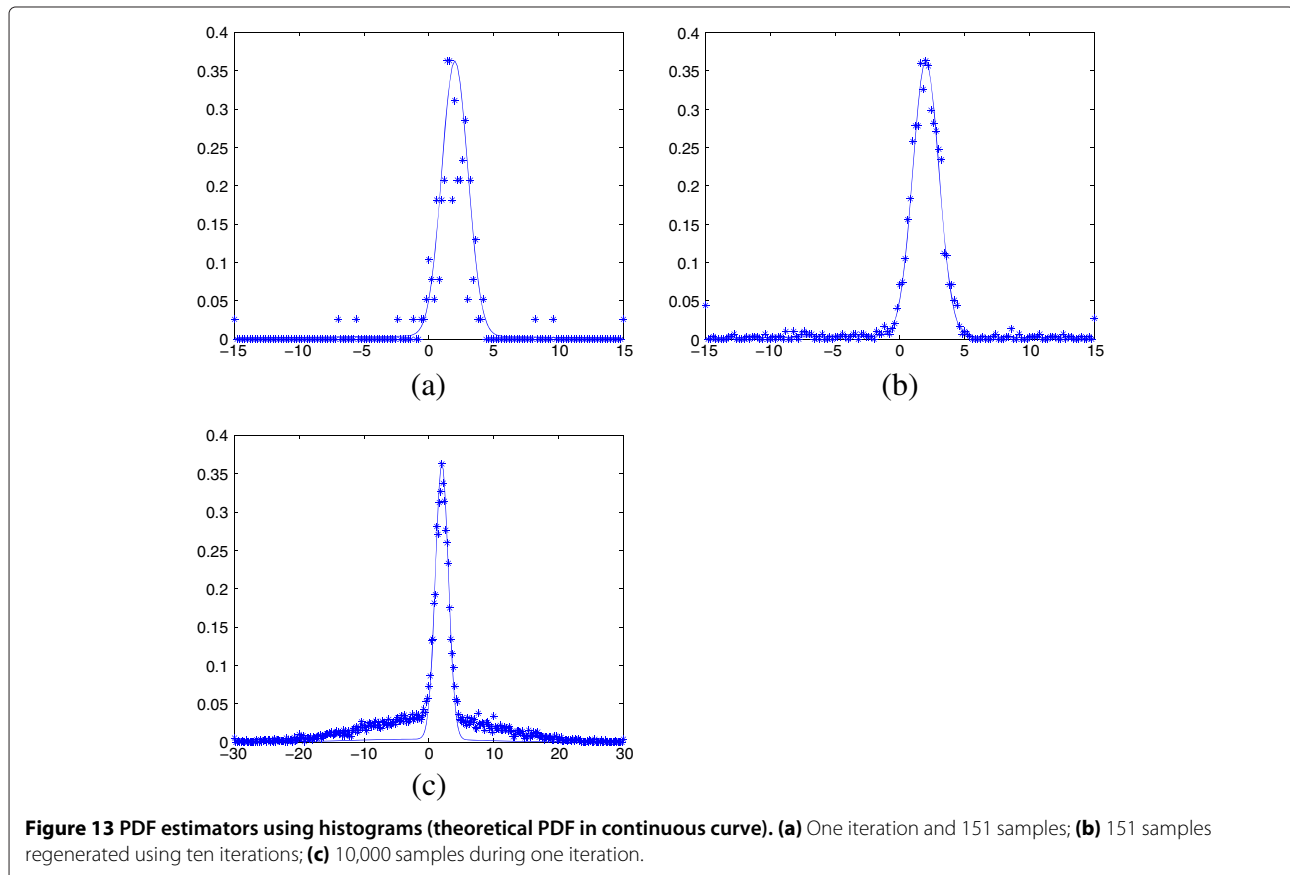
Appendix 1: high-order statistics estimators

In this section, HOS estimators are developed. Generally, HOS estimators can be divided into main families: the arithmetic and the exponential estimators.

1.1 Arithmetic estimators

Let X to be a zero-mean stochastic ergodic signal where x_i is an event (or a signal sample) of X ($1 < i < N$). In this case, the arithmetic estimator of the q th-order moment is given by

$$\widehat{\mu}_q = \frac{1}{N} \sum_{i=1}^N x_i^q. \quad (23)$$



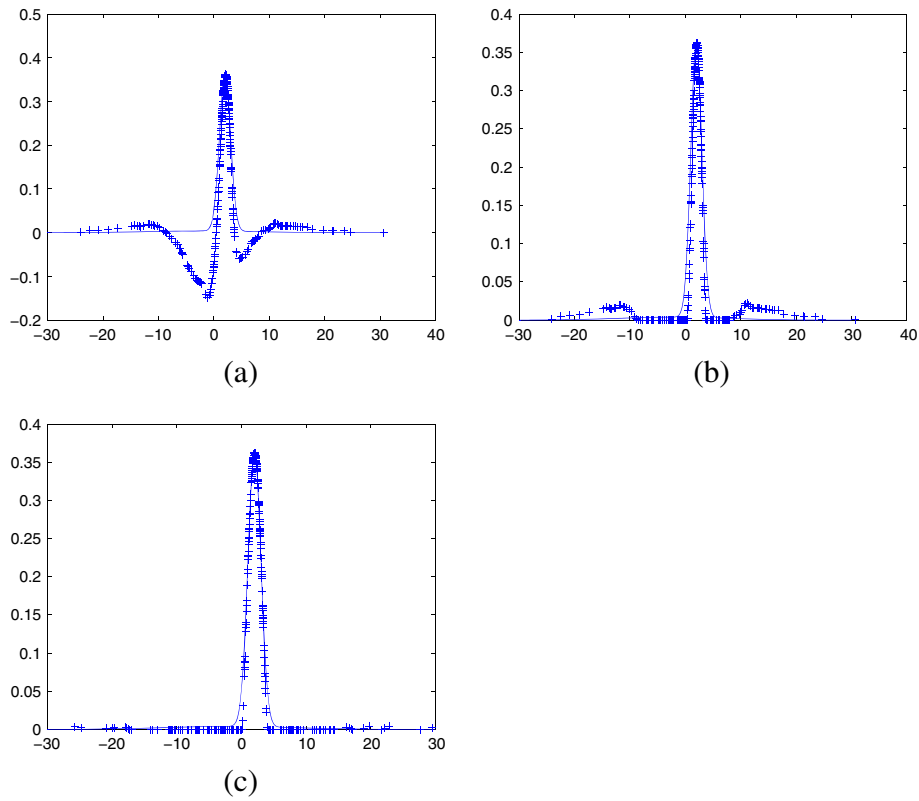


Figure 14 PDF estimators using Hermite's polynomial functions with various values of K_s . Theoretical PDF are in continuous curves. **(a)** $K_s = 4.5$ without correction. **(b)** $K_s = 4.5$ with correction. **(c)** $K_s = 7.5$ without correction.

This estimator assumes that the signal X is stationary over N samples. This estimator is a non-biased estimator (i.e., $E(\widehat{\mu}_q) = \mu_q$) and its variance is given by

$$\text{Var}(\widehat{\mu}_q) = \frac{\mu_{2q} - \mu_q^2}{N}.$$

Clearly, it is a consistent estimator; hence for stationary signals, its variance decreases with an increased number of samples. An arithmetic estimator of the q th-order cumulant can be developed from Equation 8:

$$\widehat{\text{Cum}}_q(X) = \sum (-1)^{k-1} (k-1)! \mu_{v_1} \mu_{v_2} \dots \mu_{v_p}. \quad (24)$$

It is proved [59,60] that the estimator in (24) is a biased consistent estimator where the estimation error decreases proportional to $\frac{1}{N}$:

$$E(\widehat{\text{Cum}}_q(X)) = \sum_{p=1}^q \frac{(-1)^p}{N^{p-1}} (p-1) \times \begin{pmatrix} \mu_q \\ +(N-1)\mu_{v_1}\mu_{v_1} \\ \vdots \\ +(N-1)^{p-1}\mu_{v_1}\dots\mu_{v_p} \end{pmatrix}.$$

A non-biased cumulant estimator can be deduced from the last equation:

$$\widehat{\text{Cum}}_q(X) = \sum_{p=1}^q c_p (-1)^p (p-1) \mu_{v_1} \mu_{v_2} \dots \mu_{v_p}, \quad (25)$$

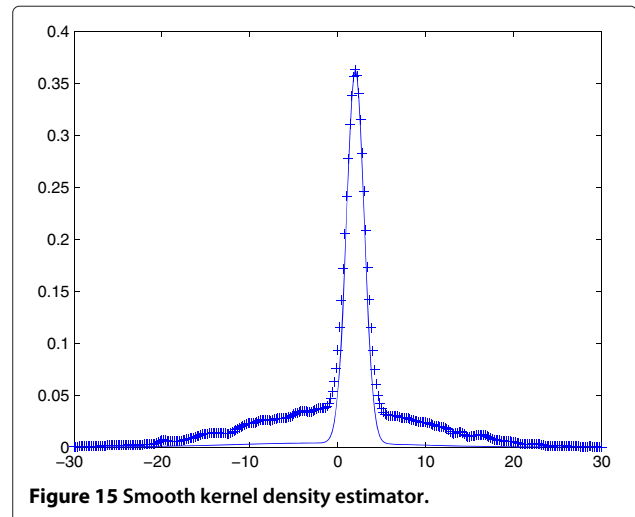


Figure 15 Smooth kernel density estimator.

where the parameters c_p depend on the partitions of the indices v_i . These parameters can be estimated as the solution of q linear equations. Let us consider the fourth-order cumulant:

$$\widehat{\text{Cum}}_4(X) = a\widehat{\mu}_4 - 4b\widehat{\mu}_1\widehat{\mu}_3 - 3c\widehat{\mu}_2^2 + 12d\widehat{\mu}_1^2\widehat{\mu}_2 - 6e\widehat{\mu}_1^4. \quad (26)$$

In order to make the last estimator unbiased, one should solve a linear system of equations obtained by comparing term-to-term the expectation of Equation 26 and the theoretical value given by (9)

$$\begin{aligned} a &= \frac{N^3 + N^2 - 24N + 24}{(N-1)(N-2)(N-3)} & b &= \frac{N(2N^2 - 10N + 9)}{2(N-1)(N-2)(N-3)} \\ c &= \frac{N(N^2 - N - 6)}{(N-1)(N-2)(N-3)} & d &= \frac{N^2(2N - 5)}{2(N-1)(N-2)(N-3)} \\ e &= \frac{N^3}{(N-1)(N-2)(N-3)} \end{aligned} \quad (27)$$

For zero-mean signals, we can easily prove that

$$\begin{aligned} E(\widehat{\text{Cum}}_4(X)) &= \mu_4 - \frac{3}{N}(\mu_4 + (N-1)\mu_2^2) \\ &= \text{Cum}_4(X) - \frac{3}{N}(\text{Cum}_4(X) + 2\mu_2^2). \end{aligned}$$

This means that the following estimator is an unbiased estimator for the fourth-order cumulant of a zero-mean stationary signal X :

$$\widehat{\text{Cum}}_4(X) = \frac{N+2}{N-1}\widehat{\mu}_4 - \frac{3}{N-1}\widehat{\mu}_2^2. \quad (28)$$

For real-time applications, the estimators should be adaptive ones. The estimator (23) is not an adaptive one, but it is easy to derive an adaptive version:

$$\widehat{\mu}_q\{k\} = \frac{1}{k} \sum_{i=1}^k x_i^q = \frac{(k-1)\widehat{\mu}_q\{k-1\} + x_k^q}{k}, \quad (29)$$

where $\widehat{\mu}_r\{k\}$ is the estimator of the r th order moment at the k th iteration.

1.2 Exponential estimators

Exponential estimators are defined as

$$\widehat{\mu}_q = (1 - \lambda_q) \sum_{i=1}^N \lambda_q^{N-i} x_i^q, \quad (30)$$

where $0 < \lambda_q < 1$ represents a forgetting factor. This estimator can be calculated easily in an adaptive way:

$$\widehat{\mu}_q\{k\} = \lambda_q \widehat{\mu}_q\{k-1\} + (1 - \lambda_q) x_k^q. \quad (31)$$

The latest estimator is biased ($E(\widehat{\mu}_q) = (1 - \lambda_q^N) \mu_q$), but it is asymptotically non-biased. The main interest in

such estimator resides on the fact that it can give better estimation for the moments of non-stationary signals. Thus, the closest λ to 1, the more past samples are taking into account. A non-biased exponential estimator can be written as

$$\widehat{\mu}_q = \frac{1 - \lambda_q}{1 - \lambda_q^N} \sum_{i=1}^N \lambda_q^{N-i} x_i^q. \quad (32)$$

Estimator (32) can be also modified to an adaptive version:

$$\widehat{\mu}_q\{k\} = \frac{1}{1 - \lambda_q^N} \left(\lambda_q (1 - \lambda_q^{k-1}) \widehat{\mu}_q\{k-1\} + (1 - \lambda_q) x_k^q \right). \quad (33)$$

An adaptive non-biased estimator of the cumulants could be derived using (22) and (33). To simplify our discussion, the fourth-order cumulant unbiased estimator for zero-mean signals could be developed as

$$\begin{aligned} \widehat{\text{Cum}}_q(X)\{k\} &= \widehat{\text{Cum}}_q(X)\{k-1\} + (1 - \gamma) H_k \\ &\quad \times \left(\widehat{\text{Cum}}_q(X)\{k-1\} \right) \end{aligned}$$

where γ is a forgetting factor and

$$\begin{aligned} H_k \left(\widehat{\text{Cum}}_q(X)\{k-1\} \right) &= x_k^4 - 4x_k^3 \widehat{\mu}_1\{k-1\} - 3x_k^2 \widehat{\mu}_2\{k-1\} \\ &\quad + 12x_k^2 \widehat{\mu}_1^2\{k-1\} - 6\widehat{\mu}_1^4\{k-1\} \\ &\quad - \widehat{\text{Cum}}_q(X)\{k-1\} \end{aligned}$$

1.3 Adaptive unbiased estimators of the fourth-order cumulants

A non-biased estimator of fourth-order cross-cumulants can be obtained from the definition of the cross-cumulants [68]. In fact let us consider K_{22} an estimator of $\text{Cum}_{22}(X, Y)$ defined as

$$K_{22} = \frac{a}{N} \sum_i x_i^2 y_i^2 - \frac{b}{N^2} \sum_{ij} x_i^2 y_j^2 - \frac{2c}{N^2} \sum_{ij} x_i x_j y_i y_j,$$

where a , b , and c should be set in order to make K_{22} a non-biased and consistent estimator. When samples x_i and y_i are independent, one can use similar estimators to these proposed in [69,70]. In the following, we assume that the samples are independent and identically distributed (iid) over time but spatially correlated. In this case, one can prove that K_{22} become a non-biased and consistent estimator:

$$\begin{aligned} E(K_{22}) &= aE(X^2 Y^2) - \frac{b}{N^2} \left(NE(X^2 Y^2) - \frac{2c}{N^2} (NE(X^2 Y^2) \right. \\ &\quad \left. + N(N-1)E(X^2)E(Y^2)) + N(N-1)E(XY)^2 \right), \end{aligned}$$

when $a = \frac{N+2}{N-1}$ and $b = c = \frac{N}{N-1}$. Similarly, one can develop other estimators:

$$K_{13} = \widehat{\text{Cum}}_{13}(X, Y) \quad (34)$$

$$= \frac{N+2}{N(N-1)} \sum_i x_i y_i^3 - \frac{3}{N(N-1)} \sum_{ij} x_i y_i y_j^2$$

$$K_{31} = \widehat{\text{Cum}}_{31}(X, Y) \quad (35)$$

$$= \frac{N+2}{N(N-1)} \sum_i x_i^3 y_i - \frac{3}{N(N-1)} \sum_{ij} x_i y_i x_j^2.$$

To obtain these estimators, signals are assumed stationary. The last assumption cannot be satisfied in our application. Therefore, some modifications should be considered. Let $C_{13}(N) = K_{13}(X, Y)$ be the adaptive estimator of the cumulant 3×1 using N samples, $A_N = \sum_i^N x_i y_i^3$ and $B_N = \sum_{ij}^N x_i y_i y_j^2$. In this case, Equation 35 can be written as

$$N(N-1)C_{13}(N) = (N+2)A_N - 3B_N. \quad (36)$$

Hence, we can prove that

$$\begin{aligned} N(N+1)C_{N+1} &= (N+3)(A_N + x_{N+1}y_{N+1}^3) \\ &\quad - 3\left(B_N + x_{N+1}y_{N+1} \sum_{j=1}^{N+1} y_j^2 + y_{N+1}^2 \sum_{i=1}^{N+1} x_i y_i\right) \\ &= N(N-1)C_N + A_N + (N+3)x_{N+1}y_{N+1}^3 \\ &\quad - 3x_{N+1}y_{N+1} \sum_{j=1}^N y_j^2 - 3y_{N+1}^2 \sum_{i=1}^N x_i y_i. \end{aligned}$$

Finally, the last equation can be modified to

$$\begin{aligned} C_N &= \frac{N-2}{N}C_{N-1} + \frac{1}{N}\mu_{13}(X, Y) + \frac{N+2}{N(N-1)}x_N y_N^3 \\ &\quad - 3x_N y_N \mu_{02}(X, Y) - 3y_N^2 \mu_{11}(X, Y), \end{aligned}$$

where $\mu_{nm}(X, Y) = \frac{1}{N-1} \sum_{i=1}^{N-1} x_i^n y_i^m$ is the estimator of $E(X^n Y^m)$ using $N-1$ samples. Using the last two equations, we can derive the final form of our adaptive fourth-order cumulant estimator:

$$\begin{aligned} C_{31}(N) &= \frac{N-2}{N}C_{31}(N-1) + \frac{1}{N}\mu_{31}(X, Y) + \frac{N+2}{N(N-1)}x_N^3 y_N \\ &\quad - 3x_N y_N \mu_{20}(X, Y) - 3x_N^2 \mu_{11}(X, Y). \end{aligned}$$

Appendix 2: Hermite's polynomial functions

According to [71], Hermite polynomials are real orthogonal polynomials with respect to the weight function $w(x) = e^{-x^2}$. The n th-order Hermite polynomial is defined as follows:

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n e^{-x^2}}{dx^n}.$$

Two hermite polynomials of orders n and m satisfy the following properties:

$$\begin{aligned} \frac{1}{2^p p \sqrt{\pi}} \int_{\mathbb{R}} H_n(x) H_m(x) e^{-x^2} dx &= \delta_m^n \\ \lim_{N \rightarrow \infty} \left(\sum_{n=0}^N H_n(x) \frac{t^n}{n} \right) &= e^{-t^2 + 2tx} \end{aligned}$$

$$H_{n+1}(x) - 2xH_n(x) + 2nH_{n-1}(x) = 0,$$

where δ_m^n is Kronecker's symbol. It is worth mentioning that the above recurrence equation is widely used in practice to estimate the Hermite polynomials. In fact, $H_0(x) = 1$, $H_1(x) = 2x$ so $H_2(x) = 4x^2 - 2$ and so on. In order to use Hermite's basis set, one should define the following two functions:

$$\begin{aligned} b_i^N(x) &= \frac{H_i(x) \exp\left(\frac{-x^2}{2}\right)}{\sqrt{2^i i!} \sqrt{\pi}} \\ b_i(x) &= \frac{1}{\sqrt{k_s}} b_i^N\left(\frac{x}{k_s}\right), \end{aligned}$$

where k_s is the scaling factor, $H_i(x)$ is the Hermite's polynomial of the i th order. Figure 16 shows the first eight normalized Hermite polynomial functions, i.e., $b_i(x)$ when $k_s = 1$.

Appendix 3: continuity process algorithm

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% DYcum = IntCont(Ycum,threscon)
%Parameters:
% 1- DYcum indicates the discontinuity of
%    square functions
%
% 2- Ycum is sum of noisy square functions
% 3- threscon is the threshold continuity
%Notes:
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Author: A. Mansour
%Comments:
%Keywords: Cleaning a noisy square
%          indication function
%Warnings:
%
%Discussion:
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% For any comment or bug report, please send
%    an e-mail to mansour@ieee.org
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
function DYcum = IntCont(Ycum,threscon)
DYcum = diff(Ycum);
ndy = find(DYcum ~= 0); %the indices of
non zero value of DYcum

```

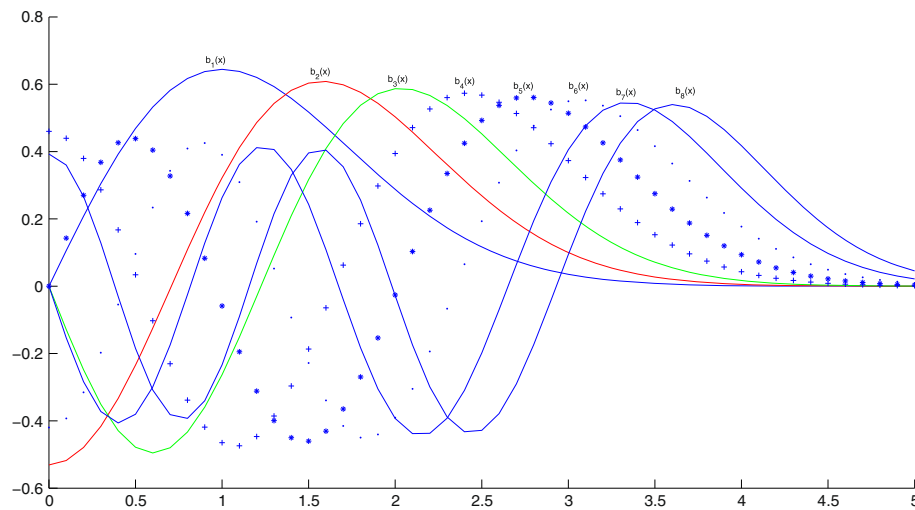


Figure 16 The first eight normalized Hermite polynomial functions.

```

VYcum = DYcum(ndy); % non zero values
      of DYcum
Dndy = diff(ndy); % intervals of change
SInt = find(Dndy < threscon); %indices of
      small intervals
for i = 1 : length(SInt)
    if ((VYcum(SInt(i)) + VYcum(SInt(i)+1)) == 0)
%error estimated windows create two diracs
% of same values but different signs
        DYcum(ndy(SInt(i))) = 0;
%if an error windows is found then the
% change in derivative should be eliminated
        DYcum(ndy(SInt(i)+1)) = 0;
        VYcum(SInt(i)) = 0;
        VYcum(SInt(i)+1) = 0;
    end
end

figure; plot(DYcum);
title('Output of IntCont using the
      continuity process');

%In some cases, the transition between two
%valid states of the Ycum shows two steps.
%This case could creates two diracs of the
%same sign and close to each other. The
%following code is to eliminate this case.
ndy = find(DYcum ~= 0); %the indices of non
      zero value of the modified DYcum
VYcum = DYcum(ndy); % non zero values of
      modified DYcum
Dndy = diff(ndy); % intervals of change
SInt = find(Dndy < threscon); %indices of
      small intervals for i=1:length(SInt)

    if ((VYcum(SInt(i)) * VYcum(SInt(i)+1)) > 0)
%transition error estimated windows create
%two diracs of different values but same
%signs
        CorrectDirac = DYcum(ndy(SInt(i)))
            + DYcum(ndy(SInt(i)+1));
        midledirac = floor((ndy(SInt(i))
            + ndy(SInt(i)+1))/2);
        %an integer index in the middle of
        %the transitions

        DYcum(midledirac) = CorrectDirac;
        DYcum(ndy(SInt(i))) = 0;
        %if an error windows is found then
        %the change in derivative should be
        %eliminated

        DYcum(ndy(SInt(i)+1)) = 0;
        VYcum(SInt(i)) = 0;
        VYcum(SInt(i)+1) = 0;
    end
end

figure; plot(DYcum);
title('Output of IntCont using continuity
      and transition processes');

```

Appendix 4

The PDF of a moderate sum of two uncorrelated RVs

Let us consider two independent RVs $X_i, i = 1, 2$. It is easy to prove that the proposed random sum can be written as follows:

$$X = \sum_{i=1}^M X_i = X_1 + aX_2,$$

where, $M \in 1, 2$ is a modified Bernoulli RV, and a is a Bernoulli RV with $\Pr(a = 1) = 1 - \Pr(a = 0) = \epsilon$. In this case, one can write:

$$\begin{aligned} F_X(x) &= \Pr(X_1 + aX_2 \leq x) \\ &= \Pr(a = 0) \Pr(X_1 + aX_2 \leq x \mid a = 0) + \Pr(a = 1) \\ &\quad \times \Pr(X_1 + aX_2 \leq x \mid a = 1) \\ &= (1 - \epsilon) \Pr(X_1 \leq x) + \epsilon \Pr(X_1 + X_2 \leq x) \\ &= (1 - \epsilon)F_{X_1}(x) + \epsilon F_{X_1+X_2}(x). \end{aligned}$$

Using the previous equation, one can deduce [53] that

$$f_X(x) = \frac{dF_X(x)}{dx} = (1 - \epsilon)f_{X_1}(x) + \epsilon f_{X_1}(x) * f_{X_2}(x),$$

where $*$ represents the convolution product. If X_1 and X_2 are two uncorrelated mutually Gaussian RVs, then $Z = X_1 + X_2$ is another normal RV with $\mathcal{N}(m_1 + m_2, \sqrt{\sigma_1^2 + \sigma_2^2})$, which proves the proposed statement.

The mean of the new variable X is given by

$$m = E\{X\} = m_1 + m_2.$$

Using the definition of the variance and the properties of a Bernoulli RV, the second moment of the new variable X becomes

$$\begin{aligned} E\{X^2\} &= E\{X_1^2\} + E\{a^2\}E\{X_2^2\} + 2E\{aX_1X_2\} \\ &= \sigma_1^2 + m_1^2 + \epsilon(\sigma_2^2 + m_2^2) + 2\epsilon m_1 m_2. \end{aligned}$$

In this case the variance of X could be obtained as follows:

$$\sigma^2 = \sigma_1^2 + \epsilon\sigma_2^2 + 2\epsilon(1 - \epsilon)m_1 m_2.$$

Competing interests

The authors declare that they have no competing interests.

Authors' information

AM received his M.Sc. and Ph.D. degrees in Signal, Image and Speech Processing from the 'Institut National Polytechnique de Grenoble - INPG' (Grenoble, France) on July 1993 and January 1997, respectively, and his HDR degree (Habilitation à Diriger des Recherches; in the French system, this is the highest of the higher degrees) on November 2006 from the Université de Bretagne Occidentale - UBO (Brest, France). His research interests are in the areas of blind separation of sources, high-order statistics, signal processing, passive acoustics, cognitive radio, robotics, and telecommunication. From January 1997 to July 1997, he held a post-doc position at Lab. de Traitement d'Images et Reconnaissance de Forme, France. From August 1997 to September 2001, he was a research scientist at the Bio-Mimetic Control Research Center (BMC) in Riken, Nagoya, Japan. From October 2001 to January 2008, he was holding a teacher-researcher position at the Ecole Nationale Supérieure des Ingénieurs des Etudes et Techniques d'Armement (ENSIETA) in Brest, France. From February 2008 to August 2010, he was a senior lecturer at

the ECE Department of Curtin University, Perth, Australia. During January 2009, he held an invited professor position at the Université du Littoral Côte d'Opale in Calais, France. From September 2010 till June 2012, he was a professor at University of Tabuk, Kingdom of Saudi Arabia. He also served as the *electrical department head* at the University of Tabuk. Since September 2012, he has been a professor at Ecole Nationale Supérieure de Techniques Avancées Bretagne (ENSTA Bretagne) in Brest, France. He is the author and the co-author of three books. He is the first author of several papers published in international journals. He is also the first author of many papers published in the proceedings of various international conferences. Finally, he was elected to the grade of IEEE Senior Member in February 2006. He has been a member of Technical Program Committees (TPC) and he was a chair, co-chair, and a scientific committee member in various international conferences. He is an active reviewer for a variety of international journals. He was a lead guest editor of *EURASIP Journal on Advances in Signal Processing*.

RM (S'00-M'08-SM'13) holds a Ph.D. in Electrical Engineering from Jacobs University in Bremen, Germany and several years of post-doctoral wireless communication and optical wireless communication research experience in Germany. In October 2010, he joined University of Tabuk in Saudi Arabia where he is now an assistant professor and the director of research excellence unit. His main research interests are in spatial modulation, MIMO cooperative wireless communication techniques, and optical wireless communication. His publications received more than 800 citations since 2007. He has published more than 50 publications in top-tier journals and conferences, and he holds seven granted patents. He also serves as on the TPC for academic conferences and is a regular reviewer for most of IEEE/OSA Communication Society's journals and IEEE/OSA Photonics Society's journals.

HMA received his M.S. and Ph.D. degrees in Electrical Engineering from the University of Washington (UW), Seattle, WA, USA. He is a professional engineer registered in the State of Washington and a senior member of the Institute of the IEEE. He has taught graduate and undergraduate courses in Electrical Engineering at a number of universities in the USA and abroad. He served at many academic ranks including Endowed Chair Professor and Vice President and Provost. He was the winner of the Boeing Supplier Excellence Award. He was also the winner of the IEEE Professor of the Year Award, UW Branch. He is listed as inventor in a major patent assigned to the Boeing Company. His research work is referred to in many patents including patents assigned to ABB, Switzerland, and EPRI, USA. Currently, he is a professor and director of the Sensor Networks and Cellular Systems (SNCS) Research Center, University of Tabuk, Tabuk, Saudi Arabia. He authored many papers in IEEE and other journals and conferences. His research interests include modeling and simulation of large scale networks, sensors and sensor networks, scientific visualization, and control and energy systems.

Acknowledgements

The authors gratefully acknowledge the support for this work by SNCS Research Center at the University of Tabuk under the grant from the Saudi Ministry of Higher Education.

Author details

¹Lab STIC - ENSTA Bretagne, 2 Rue François Verny, Brest 29200, France. ²SNCS Research Center, University of Tabuk, P.O. Box 741, Tabuk 71491, Kingdom of Saudi Arabia.

Received: 1 July 2013 Accepted: 7 February 2014

Published: 20 February 2014

References

1. J Proakis, M Salehi, *Digital Communications* (McGraw-Hill, New York, 2007)
2. V Tarokh, H Jafarkhani, AR Calderbank, Space-time codes high data rate wireless communication: performance criterion and construction. *IEEE Trans. Inf. Theory*. **44**(2), 744 (1998)
3. I Bradaric, AP Petropulu, K Diamantaras, On blind identification of FIR-MIMO systems with cyclostationary inputs using second order statistics. *IEEE Trans. Signal Process.* **51**(2), 434 (2003)
4. JA Srar, KS Chung, A Mansour, Adaptive array beamforming using a combined LMS-LMS algorithm. *IEEE Trans. Antennas Propagation*. **58**, 3545 (2010)
5. R Mesleh, S Ikki, H Aggoune, A Mansour, Performance analysis of space shift keying (SSK) modulation with multiple cooperative relays. *EURASIP J. Adv. Signal Process.* **1** (2012). doi:10.1186/1687-6180-2012-201

6. M Abaza, R Mesleh, A Mansour, A Alfalou, MIMO techniques for high data rate free space optical communication system in log-normal channel, in *International Conference on Technological Advances in Electrical, Electronics and Computer Engineering* (TAECE, Konya, 9–11 May 2013), pp. 161–166
7. M Simon, MS Alouini, *Digital Communication over Fading Channels: a Unified Approach to Performance Analysis* (Wiley, New York, 2000)
8. S Haykin, M Moher, *Communication Systems* (Wiley, New York, 2009)
9. M Abaza, R Mesleh, A Mansour, HA Aggoune, Diversity techniques for FSO communication system in correlated log-normal channels. *Opt. Eng.* **53**(1) (2014). 10.1117/1.OE.53.1.016102
10. VP Skitovic, Linear forms of independent random variables and the normal distribution law. *Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya*. **18**(185) (1954). In Russian
11. M Kendall, A Stuart, *The Advanced Theory of Statistics: Distribution Theory* (Charles Griffin & Company Limited, London, 1961)
12. CL Nikias, M Shao, *Signal Processing with α -Stable Distribution and Applications* (Wiley, New York, 1995)
13. J Sijbers, J Den Dekker, P Scheunders, D Van Dyck, Maximum-likelihood estimation of Rician distribution parameters. *IEEE Trans. Med. Imaging*. **17**(3), 357 (1998)
14. GK Karagiannidis, DA Zogas, SA Kotsopoulos, On the multivariate Nakagami-m distribution with exponential correlation. *IEEE Trans. Commun.* **51**(8), 1240 (2003)
15. MD Yacoub, G Fraidenraich, JCS Santos Filho, Nakagami-m phase-envelope joint distribution. *Electron. Lett.* **41**(5) (2005). doi:10.1049/el:20057014
16. MD Yacoub, G Fraidenraich, HB Tercius, FC Martins, The symmetrical on the $\eta - \kappa$ distribution: a general fading distribution. *IEEE Trans. Broadcasting*. **51**, 504 (2005)
17. G Fraidenraich, M Daoud Yacoub, The $\alpha - \eta - \mu$ and $\alpha - k - \mu$ fading distributions, in *IEEE 9th International Symposium on Spread Spectrum Techniques and Applications* (Manaus, 28–31 August 2006), pp. 255–263
18. S Nadarajah, S Kotz, On the $\eta - \kappa$ distribution. *IEEE Trans. Broadcasting*. **52**(3), 411 (2006)
19. MD Yacoub, The $\alpha - \eta$ distribution: a physical fading model for the Stacy distribution. *IEEE Trans. Vehicular Technol.* **56**(1), 27 (2007)
20. SL Cotton, WG Scanlon, J Guy, The $k - \mu$ distribution: applied to the analysis of fading in body to body communication channels for fire and rescue personnel. *IEEE Antennas Wireless Propagation Lett.* **7**, 66 (2008)
21. MD Yacoub, Nakagami-m phase-envelope joint distribution: an improved model, in *Microwave and Optoelectronics Conference, IMOC'2009*, (Belem, 3–9 November 2009), pp. 335–339
22. DD Cox, A penalty method for nonparametric estimation of the logarithmic derivative of a density function. *Ann. Instit. Stat. Math.* **37**, 271 (1985)
23. J Vidal, A Bonafonte, AR Fonollosa, NF De Losada, Parametric modeling of PDF using a convolution of one-sided exponentials: application to HMM, in *European Signal Processing Conference*, ed. by M Holt, C Cowan, P Grant, W Sandham, *European Signal Processing Conference* (Elsevier, Edinburgh, 1994), pp. 54–57
24. M Vannucci, Nonparametric density estimation using wavelets. Discussion Paper, Duke University, 1998
25. A Archambeau, M Verleysen, Fully nonparametric probability density function estimation with finite gaussian mixture models, in *5th International Conference on Advances in Pattern Recognition Calcutta, ICAPR'2003*, (Allied, New Delhi, Calcutta, 2003), pp. 81–84
26. R Boscolo, H Pan, VP Roychowdhury. *IEEE Trans. Neural Netw.* **15**(1), 55 (2004)
27. F Roueff, T Rydén, Nonparametric estimation of mixing densities for discrete distributions. *Ann. Stat.* **33**(5), 2066 (2005)
28. S Alamouti, A simple transmit diversity technique for wireless communication. *IEEE J. Selected Areas Commun.* **16**(8), 1451 (1998)
29. EG Larsson, P Stoica, *Space-time Block Coding for Wireless Communications* (The Press Syndicate of the University of Cambridge, Cambridge, 2003)
30. V Choqueuse, A Mansour, G Burel, L Collin, K Yao, Blind channel estimation for STBC systems using higher-order statistics. *IEEE Trans. Wireless Commun.* **10**, 495 (2011)
31. A Mansour, C Jutten, N Ohnishi, Kurtosis: definition and properties, in *International Conference on Multisource-Multisensor Information Fusion*, ed. by HR Arabnia, DD Zhu, (Las Vegas, 6–9 July 1998), pp. 40–46
32. A Mansour, C Jutten, What should we say about the kurtosis. *IEEE Signal Process. Lett.* **6**(12), 321 (1999)
33. JM Mendel, Tutorial on higher-order statistics (Spectra) in signal processing and system theory: theoretical results and some applications. *IEEE Proc.* **79**, 277 (1991)
34. A Cichocki, SI Amari, *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications* (Wiley, New York, 2002)
35. JF Cardoso, Source separation using higher order moments, in *Proceedings of International Conference on Speech and Signal Processing 1989, ICASSP 1989*, (Glasgow, 22–25 May 1989), pp. 2109–2212
36. P Comon, Independent component analysis, a new concept *Signal Process.* **36**(3), 287 (1994)
37. A Mansour, C Jutten, A direct solution for blind separation of sources. *IEEE Trans. Signal Process.* **44**(3), 746 (1996)
38. A Hyvärinen, J Karhunen, E Oja, *Independent Component Analysis* (Wiley, New York, 2001)
39. A Mansour, A Kardec Barros, N Ohnishi, Blind separation of sources: methods, assumptions and applications. *IEICE Trans. Fundamentals Electron. Commun. Comput. Sci.* **E83-A**(8), 1498 (2000)
40. A Papoulis, *Probability, Random Variables, and Stochastic Processes* (McGraw-Hill, New York, 1991)
41. M Rosenblatt, Remarks on some nonparametric estimates of a density function. *Ann. Math. Stat.* **27**(3), 832 (1956)
42. E Parzen, On the estimation of a probability density function and mode. *Ann. Math. Stat.* **33**(3), 1065 (1962)
43. LN Bowers, Expansion of probability density functions as a sum of gamma densities with applications in risk theory. *Trans. Soc. Actuaries*. **18 PT 1**(52), 125 (1966)
44. HV Khuong, HY Kong, General expression for pdf of a sum of independent exponential random variables. *IEEE Commun. Lett.* **10**(3), 159 (2006)
45. CH Reinsch, Smoothing by spline functions. *Numer. Math.* **10**, 177 (1967)
46. AW Bowman, A Azzalini, *Applied Smoothing Techniques for Data Analysis* (Oxford University, New York, 1997)
47. SC Schwartz, Estimation of probability density by an orthogonal series. *Ann. Math. Stat.* **38**(3), 1261 (1967)
48. M Vannucci, On the application of wavelets in statistics. Ph.D. thesis, Dipartimento di Statistica, University of Florence, 1996
49. J Engel, Density estimation with, Haar series. *Stat. Probability Lett.* **9**, 111 (1990)
50. S Kay, Model-based probability density function estimation. *IEEE Signal Process. Lett.* **5**(12), 318 (2001)
51. J Xie, Z Wang, Probability density function estimation based on windowed Fourier transform of characteristic function, in *Proceedings of the 2nd International Congress on Image and Signal Processing*, (Tianjin, 17–19 October 2009)
52. RM Howard, PDF estimation via characteristic function and an orthonormal basis set, in *Proceedings of the 14th WSEAS International Conference on Systems*, (Corfu, 23–25 July 2010), pp. 100–105
53. A Mansour, *Probabilités et statistiques pour les ingénieurs: cours, exercices et programmation* (Hermes Science, Paris, 2007)
54. R Mesleh, H Haas, S Sinanovic, CW Ahn, S Yun, Spatial modulation. *IEEE Trans. Vehicule Technol.* **57**(4), 2228 (2008)
55. Z Wang, GB Giannakis, A simple general parameterization quantifying performance in fading channels. *IEEE Trans. Commun.* **51**(8), 1389 (2003)
56. P McCullagh, *Tensor Methods in Statistics* (Chapman and Hall, London, 1987)
57. AN Shirayayev, *Probability* (Springer, London, 1984)
58. MH Vu, Exploiting transmit channel side information in MIMO wireless systems. Ph.D. thesis, Stanford University, 2006
59. S Kotz, NL Johnson, *Encyclopedia of Statistical Sciences* (University of Amsterdam, Amsterdam, 1993)
60. JL Lacoume, PO Amblard, P Comon, *Statistiques d'ordre supérieur pour le traitement du signal* (Mason, Paris, 1997)
61. J Bass, *Cours de Mathématiques: Tome 1, Fascicule 2* (Masson, Paris, New York, Barcelone et Milan, 1977)
62. J Bass, *Cours de Mathématiques: Tome 2* (Masson, Paris, New York, Barcelone et Milan, 1977)
63. A Mansour, M Kawamoto, ICA papers classified according to their applications & performances. *IEICE Trans. Fundamentals Electron. Commun. Comput. Sci.* **E86-A**(3), 620 (2003)

64. A Savitzky, MJE Golay, Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **36**, 1627–1639 (1994)
65. SJ Orfanidis, *Introduction to Signal Processing* (Prentice-Hall, New Jersey, 1995)
66. RW Schafer, What is a Savitzky-Golay filter? *IEEE Signal Process. Mag.*, 111–117 (2011)
67. K Beauchamp, *Applications of Walsh and Related Functions* (Academic, New York, 1984)
68. M Kendall, A Stuart, *The Advanced Theory of Statistics: Design and Analysis, and Time-series* (Charles Griffin & Company Limited, London, 1961)
69. A Mansour, A Kardec Barros, N Ohnishi, Comparison among three estimators for high order statistics, in *Fifth International Conference on Neural Information Processing (ICONIP'98)*, ed. by S Usui, T Omori, (Kitakyushu, 21–23 October 1998), pp. 899–902
70. A Martin, A Mansour, Comparative study of high order statistics estimators, in *International Conference on Software, Telecommunications and Computer Networks*, (Split, Dubrovnik, Venice, 10–13 October 2004), pp. 511–515
71. A Bouvier, M George, F Le Lionnais, *Dictionnaire des mathématiques* (Presses Universitaires De France, Paris, 1993)

doi:10.1186/1687-6180-2014-21

Cite this article as: Mansour et al.: Blind estimation of statistical properties of non-stationary random variables. *EURASIP Journal on Advances in Signal Processing* 2014 **2014**:21.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com