HIGHER-ORDER STATISTICS INTERPRETATION. APPLICATION TO POWER-QUALITY CHARACTERIZATION

Juan José González de la Rosa, África Luque
Univ. Cádiz. Electronics Area. Research Group PAI-TIC-168, EPSA. Av. Ramón Puyol S/N. E-11202-Algeciras-Cádiz-Spain

Carlos G. Puntonet, J. M. Górriz
Univ. Granada. Dept. ATC, ESII. Periodista Daniel Saucedo. E-18071-Granada-Spain

Antonio Moreno Muñoz
Univ. Córdoba. Electronics Area. Research Group PAI-TIC-168
Campus Rabanales. A. Einstein C-2. E-14071-Córdoba-Spain

Keywords: Higher-Order statistics, Signal processing, Transient Characterization, Power-Quality.

Abstract: In this paper we perform a practical review on higher-order statistics interpretation. Concretely we focuss on an unbiased estimate of the *4*th-order time-domain cumulants. Some synthetics involving classical noise processes are characterized using this unbiased estimate, with the goal of checking its performance and to provide the scientific community with another result, dealing with the interpretation of this signal processing tool. A real-life practical example is presented in the field of electrical power quality event analysis. The work also aims to present a set of general advice in order to save memory and gain speed in a real signal processing frame, dealing with non-stationary processes.

1 INTRODUCTION

Gaussian processes are completely characterized by the autocorrelation sequence and its associated Fourier transform, the power spectrum. In the power spectrum estimation, the information regarding the phase of the frequency components of the signal is not present. The information in the power spectrum is essentially the same as in the autocorrelation (Nandi, 1999).

However, there are numerous situations where we have to look beyond the autocorrelation in order to get extra information regarding deviations from the Gaussian behavior and nonlinear characterization. These additional characteristics help us distinguish among apparently similar measurement data sequences; therefore getting the complete characterization of the process.

Data sequences, and their associated power spectra, which have been obtained by multiplying more than two time-series are called higher-order statistics (HOS). Their associated Fourier transforms are called poly-spectra. They contain additional information regarding the phase of the frequency components of the signals under study (Nikias and Petropulu, 1993).

The power spectrum (second-order spectrum) is a particular case of higher-order spectra. The third-order spectrum is called the bi-spectrum and the fourth-order spectrum is called the try-spectrum. They are defined to be the Fourier transforms or the third and the fourth-order cumulant sequences, respectively.

Poly-spectra are defined as the higher-order moment spectra and cumulant spectra can be defined for both deterministic signals and random processes. Moment spectra can be very useful in the analysis of deterministic signals (transient and periodic), whereas cumulant spectra are of great importance in the analysis of stochastic signals.

The motivation of the poly-spectral analysis yields in three applications: (a) To suppress Gaussian noise processes of unknown spectral characteristics; the bi-spectrum also suppress noise with symmetrical probability distribution, (b) to reconstruct the magnitude and phase response of systems, and (c) to detect and characterize nonlinearities in time-series.

In this paper we show the application results dealing with the characterization of random processes,

José González de la Rosa J., Luque Á., G. Puntonet C., M. Górriz J. and Moreno Muñoz A. (2007).
HIGHER-ORDER STATISTICS INTERPRETATION. APPLICATION TO POWER-QUALITY CHARACTERIZATION.
In *Proceedings of the Second International Conference on Signal Processing and Multimedia Applications*, pages 72-78
DOI: 10.5220/0002136100720078

following the indications in (Nandi, 1999) and in (Nikias and Petropulu, 1993). A real example involving power quality event analysis is then studied to show the difference in dealing with real time-series instead of synthetics. Computational intelligence based in neural classifiers are pointed as the classification strategy, once the transients have been characterized.

The paper is structured as follows. Section 2 includes the motivation of HOS and an amenable mathematical foundation. In Section 3 a time-domain fourth-order cumulant analysis is applied to time series regarding realizations of noise processes. Section 4 presents the real-life analysis of electrical transients. Finally, conclusions are shown in Section 5.

2 HOS IN TIME AND FREQUENCY DOMAINS

2.1 A Preliminary Example

To motivate the use of HOS a first application is outlined. We consider two cosine waves of the same frequency with a phase shift of $\pi/2$ radians. The signals and their autocorrelation plots are depicted in Fig. 1.

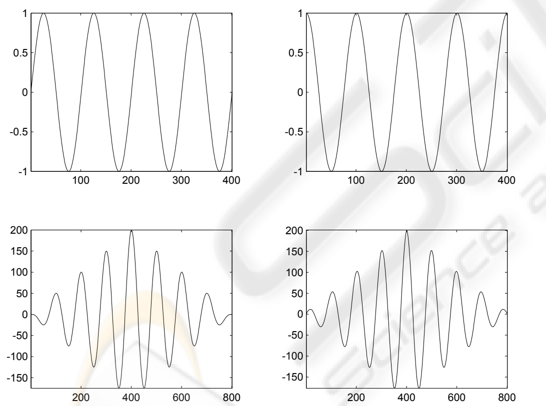

Figure 1: Two sinusoids with the same frequency, shifted $\pi/2$ radians (Two upper graphs). They have the same second-order statistics (the autocorrelation sequence).

Figs. 2 and 3 show the contour plots associated to the third-order cumulant and the bi-spectrum of the former cosine sequences. Qualitatively (but also in almost every practical case) we only have to pay attention at the coarser differences between graphs. The main bump in the center of Fig. 2, which corresponds to the situation of zero phase shift between the sinusoids, disappears in Fig. 3, giving rise to others, secondary bumps. Thanks to the bi-spectrum, these extra bumps indicate by the way the presence of a phase shift.

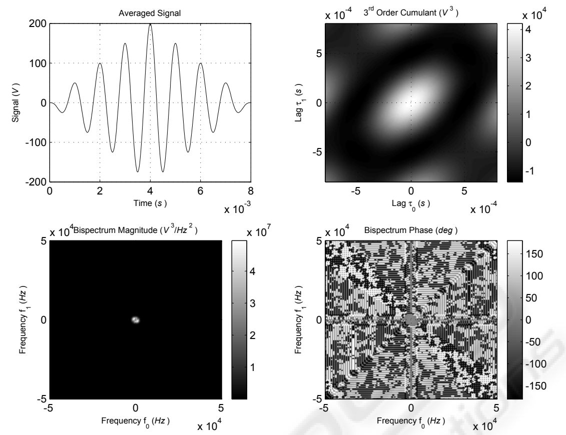

Figure 2: Third-order statistics of the autocorrelation function (top-left) corresponding to a sinusoid with no phase shift. A main *bump.* in the center is observed in the cumulant plot (top-right).

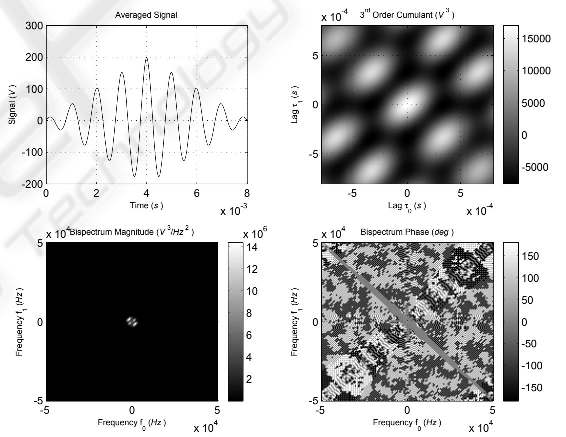

Figure 3: Third-order statistics of the autocorrelation function (top-left) corresponding to a sinusoid with a constant 90 phase shift. Several *bumps.* appear in the cumulant plot (top-right).

Let's revise some mathematical foundations.

2.2 Cumulants and Moments

High-order statistics, known as cumulants, are used to infer new properties about the data of non-Gaussian processes (Hinich, 1990; Mendel, 1991; Nikias and Petropulu, 1993). Before cumulants, due to the lack of analytical tools, such processes had to be treated as if they were Gaussian. Cumulants and their associated Fourier transforms, known as poly-spectra, reveal information about amplitude and phase, whereas second order statistics (power, variance, covariance and spectra) are phase-blind (Mendel, 1991; Swami et al., 2001; De la Rosa et al., 2005; De la Rosa and Ruzzante, 2007,).

Before the definitions, it is convenient to remark that cumulants of order higher than 2 are all zero in signals with Gaussian probability density functions. What is the same, cumulants are blind to any kind of a Gaussian process. This is the reason why it is not possible to separate these signals using the statistical approach (Nikias and Petropulu, 1993).

The relationship among the cumulant of *r* stochastic signals, $\{x_i\}_{i\in[1,r]}$, and their moments of order $p, p \leq r$, can be calculated by using the *Leonov-Shiryaev* formula (Mendel, 1991; Nikias and Petropulu, 1993)

$$Cum(x_1, ..., x_r) = \sum (-1)^{p-1} \cdot (p-1)! \cdot E\{\prod_{i\in s_1} x_i\} \cdot E\{\prod_{i\in s_2} x_j\} \cdots E\{\prod_{i\in s_p} x_k\} \quad (1)$$

where the addition operator is extended over all the partitions, like one of the form $(s_1, s_2, \cdots, s_p)$, $p = 1, 2, \cdots, r$; and $(1 \leq i \leq p \leq r)$; s_i is a set belonging to a partition of order *p*, of the set of integers *1,. . . ,r*.

For a zero mean variable, using (1), the second-, third-, and fourth-order cumulants, are particular cases and are given by:

$$Cum(x_1, x_2) = E\{x_1 \cdot x_2\} \quad (2a)$$

$$Cum(x_1, x_2, x_3) = E\{x_1 \cdot x_2 \cdot x_3\} \quad (2b)$$

$$\begin{aligned} Cum(x_1, x_2, x_3, x_4) &= E\{x_1 \cdot x_2 \cdot x_3 \cdot x_4\} \\ &- E\{x_1 \cdot x_2\}E\{x_3 \cdot x_4\} \\ &- E\{x_1 \cdot x_3\}E\{x_2 \cdot x_4\} \\ &- E\{x_1 \cdot x_4\}E\{x_2 \cdot x_3\} \end{aligned} \quad (2c)$$

Let $\{x(t)\}$ be an *r*th-order stationary random real-valued process. The *r*th-order cumulant is defined as the joint *r*th-order cumulant of the random variables *x(t), x(t+*τ_1*),. . . , x(t+*τ_{r-1}*)*,

$$\begin{aligned} &C_{r,x}(\tau_1, \tau_2, \ldots, \tau_{r-1}) \\ &= Cum[x(t), x(t+\tau_1), \ldots, x(t+\tau_{r-1})] \end{aligned} \quad (3)$$

For stationary random processes the *r*th-order cumulant is only a function of *r-1* lags. If $\{x(t)\}$ is non-stationary then the *r*th-order cumulant includes time dependency.

The second-, third- and fourth-order cumulants of zero-mean *x(t)* can be expressed using (2) and (3), via

$$C_{2,x}(\tau) = E\{x(t) \cdot x(t+\tau)\} \quad (4a)$$

$$C_{3,x}(\tau_1, \tau_2) = E\{x(t) \cdot x(t+\tau_1) \cdot x(t+\tau_2)\} \quad (4b)$$

$$\begin{aligned} &C_{4,x}(\tau_1, \tau_2, \tau_3) \\ &= E\{x(t) \cdot x(t+\tau_1) \cdot x(t+\tau_2) \cdot x(t+\tau_3)\} \\ &- C_{2,x}(\tau_1)C_{2,x}(\tau_2 - \tau_3) \\ &- C_{2,x}(\tau_2)C_{2,x}(\tau_3 - \tau_1) \\ &- C_{2,x}(\tau_3)C_{2,x}(\tau_1 - \tau_2) \end{aligned} \quad (4c)$$

By putting $\tau_1 = \tau_2 = \tau_3 = 0$ in Eq. (4), we obtain

$$\gamma_{2,x} = E\{x^2(t)\} = C_{2,x}(0) \quad (5a)$$

$$\gamma_{3,x} = E\{x^3(t)\} = C_{3,x}(0, 0) \quad (5b)$$

$$\gamma_{4,x} = E\{x^4(t)\} - 3(\gamma_{2,x})^2 = C_{4,x}(0, 0, 0) \quad (5c)$$

The expressions in Eq. (5) are measurements of the variance, skewness and kurtosis of the distribution in terms of cumulants at zero lags (the central cmulants).

Normalized kurtosis and skewness are defined as $\gamma_{4,x}/(\gamma_{2,x})^2$ and $\gamma_{3,x}/(\gamma_{2,x})^{3/2}$, respectively. We will use and refer to normalized quantities because they are shift and scale invariant. If *x(t)* is symmetrically distributed, its skewness is necessarily zero (but not *vice versa*); if *x(t)* is Gaussian distributed, its kurtosis is necessarily zero (but not *vice versa*).

2.3 Poly-spectra

We will assume in the following that the cumulant sequences satisfies the bounding condition:

$$\sum_{\tau_1=-\infty}^{\tau_1=+\infty} \cdots \sum_{\tau_{r-1}=-\infty}^{\tau_{r-1}=+\infty} |C_{r,x}(\tau_1, \tau_2, \ldots, \tau_{r-1})| < \infty \quad (6)$$

Under this assumption, the higher-order spectra are usually defined in terms of the *r*th-order cumulants as their (*r*-1)-dimensional Fourier transforms

$$\begin{aligned} &S_{r,x}(f_1, f_2, \ldots, f_{r-1}) \\ &= \sum_{\tau_1=-\infty}^{\tau_1=+\infty} \cdots \sum_{\tau_{r-1}=-\infty}^{\tau_{r-1}=+\infty} C_{r,x}(\tau_1, \tau_2, \ldots, \tau_{r-1}) \\ &\cdot \exp[-j2\pi(f_1\tau_1 + f_2\tau_2 + \cdots + f_{r-1}\tau_{r-1})] \end{aligned} \quad (7)$$

The special poly-spectra derived from (7) are power spectrum (*r*=2), bi-spectrum (*r*=3) and try-spectrum (*r*=4). Only power spectrum is real, the others are complex magnitudes.

Poly-spectra are multidimensional functions which comprise a lot of information. As a consequence, their computation may be impractical in some cases. To extract useful information one-dimensional slices of cumulant sequences and spectra, and bi-frequency planes are employed in non-Gaussian stationary processes (Jakubowski et al., 2002), (De la Rosa et al., 2004,).

3 ESTIMATES AND CHARACTERIZATION OF STATISTICAL PROCESSES

In this section we show the computation results obtained from the application of an estimator. In practice, the computation of the cumulants and the polyspectra is based in estimates. In order to asses the performance of the selected estimator we test it considering a 2048-point sample register for each noise process, as it was catalogued in (Nikias and Petropulu, 1993). These four noise processes are indistinguishable from the second-order perspective, as it is shown in 4.

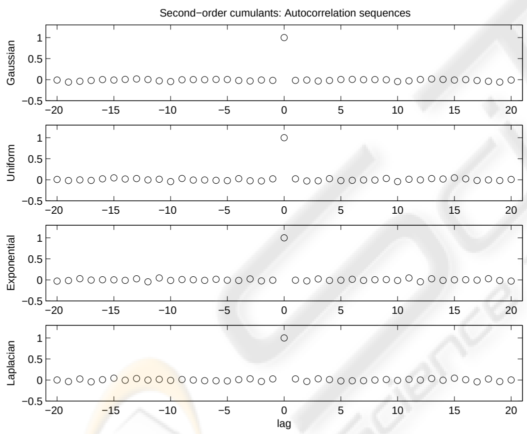

Figure 4: The four second-order cumulants corresponding to 4 sample registers, each of which is a realization of a different noise process. They exhibit the same autocorrelation sequence.

If we look into the fourth-order sequences, substantial differences are observed, specially those corresponding to zero time lags. This can be seen in Figs. form 5 to 8.

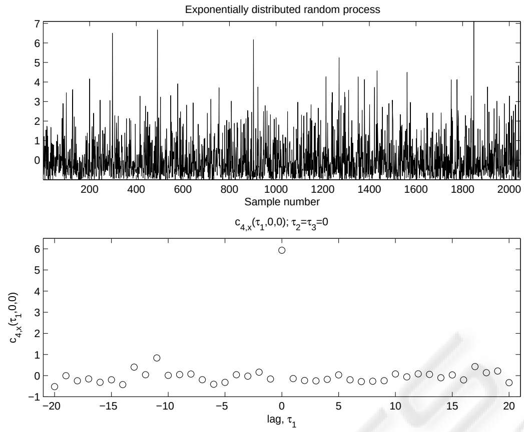

Figure 5: A 2048-point realization of an exponentially distributed noise process. Maximum value (at zero lag) = 5.9326585626518 (theoretical=6).

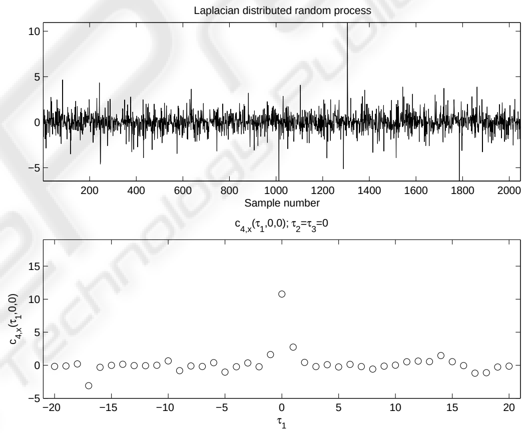

Figure 6: A 2048-point realization of an Laplacian distributed noise process. Maximum value (at zero lag) = 10.7788 (theoretical=12).

4 FEATURE EXTRACTION IN ELECTRICAL POWER EVENT ANALYSIS

The aim of the experiment is to differentiate between two classes of transients (PQ events), named long-duration and short-duration. The experiment comprises two stages. The feature extraction (classification) stage is based on the computation of the cumulants (De la Rosa et al., 2007,). Each vector's coordinate corresponds to the local maximum and minimum of the 4th-order central cumulant. This is the feature-extraction stage. And the classification stage is based on the application of the competitive layer to the feature vectors, in order to obtain two clusters in the feature plane. We use a two-neuron competitive layer,

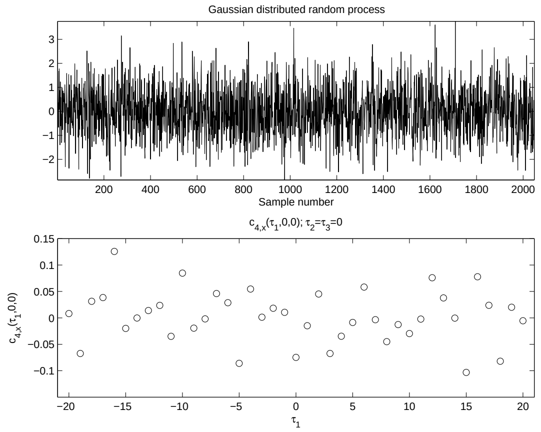

Figure 7: A 2048-point realization of an Gaussian distributed noise process. Extreme value not defined; al the values surrounds zero.

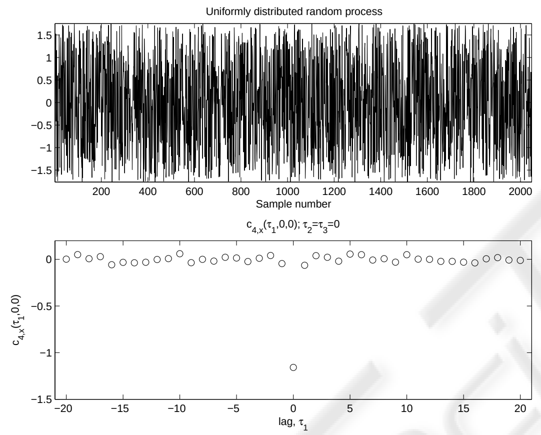

Figure 8: A 2048-point realization of a uniformly distributed noise process. Extreme value (for zero lag) = -1.15845526517794 (theoretical=-1).

which receives two-dimensional input feature vectors in this training stage.

We analyze a number of 16 1000-point (roughly) real-life registers during the feature extraction stage. Before the computation of the cumulants, two pre-processing actions have been performed over the sample signals. First, they have been normalized because they exhibit very different-in-magnitude voltage levels. Secondly, a high-pass digital filter (*5*th-order Butterworth model with a characteristic frequency of 150 Hz) eliminates the low frequency components which are not the targets of the experiment. This by the way increases the non-Gaussian characteristics of the signals, which in fact are reflected in the higher-order cumulants.

After filtering, a 50-point sliding battery of central cumulants (2nd, 3rd and 4th order) are calculated. The window's width (50 points) has been selected neither to be so long to cover the whole signal nor to be very short. The algorithm calculates the 3 central cumulants over 500 points, and then it jumps to the following starting point; as a consequence we have 98 per cent overlapping sliding windows (49/50=0.98). Thus, each computation over a window (called a segment) outputs 3 cumulants.

Fig. 9 shows an example of signal processing analysis of two sample registers corresponding to a long-duration and a short-duration events, respectively.

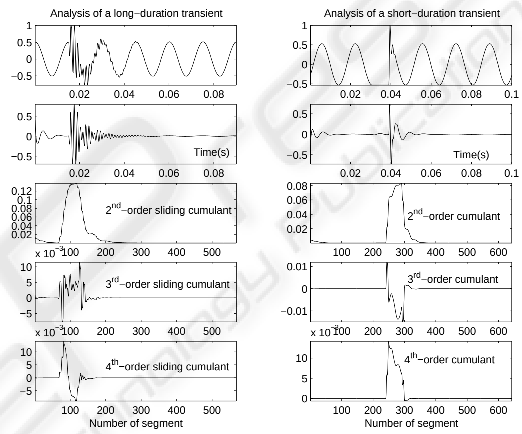

Figure 9: Long-duration vs. short-duration transient analysis. From top to bottom: the original data record, the filtered sequence, 2nd-3rd-4th-order cumulants sliding windows, respectively.

The *2*nd-order cumulant sequence corresponds to the variance, which clearly indicates the presence of an event, due to the excess of power. Both types of transients exhibit an increasing variance in the neighborhood of the PQ event, that presents the same shape, with only one maximum. The magnitude of this maximum is by the way the only available feature which can be used to distinguish different events from the second-order point of view. This may suggest the use of additional features in order to distinguish different types of events.

For this reason the higher-order central cumulants are chosen. An unbiased estimator of the cumulants has been selected. Third-order diagrams don't show quite different clusters if we consider a bi-dimensional space (2 coordinates for each feature vector) because maxima and minima are similar. It is possible to differentiate PQ events from the *3*rd-order perspective if we consider more features in the input vector, like the number of extremes (maxima and minima), and the order in which the maxima and the minima appear as time increases. In this paper we have focussed the experience on a bi-dimensional representation (2-dimensional feature vectors) because we obtain very intelligible 2-D graphs.

Fourth-order sliding cumulants exhibit clear differences, not only for the shape of the computation graph (the bottom graphs in Fig. 9, but also for the different location of minima, which suggest a clustering zone for the points.

Fig. 10 presents the results of the training stage, using the *Kohonen* rule. The horizontal (vertical) axis corresponds to the maxima (minima) value. Each cross in the diagram corresponds to an input vector and the circles indicate the final location of the weight vector (after learning) for the two neurons of the competitive layer. Both weight vectors point to the asterisk, which is the initializing point (the midpoint of the input intervals).

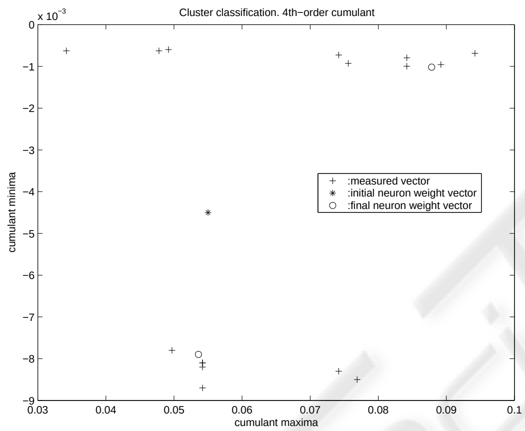

Figure 10: Competitive layer training results over 20 epochs. Upper cluster: Short-duration PQ-events. Down cluster: Long-duration event.

The separation between classes (inter-class distance) is well defined. Both types of PQ events are horizontally clustered. The correct configuration of the clusters is corroborated during the simulation of the neural network, in which we have obtained an approximate classification accuracy of 97 percent. During the simulation new signals (randomly selected from our data base) were processed using the method described.

The accuracy of the classification method increases with the number of data. To evaluate the confidence of the statistics a significance test have been conducted. This informs if the number of experiments is statistically significant according to the fitness test (Ömer Nezih Gerek and Ece, 2006). As a result of the test, the number of measurements is significatively correct.

5 CONCLUSION

In this work we have reviewed computation of higher-order statistics. Concretely we have focussed on a *4*-order estimates. We have also proposed a method to detect and classify two electrical power transients, named short and long-duration. The method comprises two stages. The first includes pre-processing (normalizing and filtering) and outputs the 2-D feature vectors, each of which coordinate corresponds to the maximum and minimum of the central cumulants. The second stage uses a neural network to classify the signals into two clusters. This stage is different-in-nature from the one used in (Ömer Nezih Gerek and Ece, 2006) consisting of quadratic classifiers. The configuration of the clusters is assessed during the simulation of the neural network, in which we have obtained an acceptable classification accuracy.

ACKNOWLEDGEMENTS

The authors would like to thank the *Spanish Ministry of Education and Science* for funding the project DPI2003-00878 which involves noise processes modeling, and the PETRI project PTR95-0824-OP involving higher-order statistics. Also thanks to the *Andalusian Government* for the trust put in the research group PAI-TIC-168, and for supporting the excellence project PAI2005-TIC00155, which involves higher-order statistics.

REFERENCES

Hinich, M. J. (1990). Detecting a transient signal by biespectral analysis. *IEEE Trans. Acoustics*, 38(9):1277–1283.

Jakubowski, J., Kwiatos, K., Chwaleba, A., and Osowski, S. (2002). Higher order statistics and neural network for tremor recognition. *IEEE Trans. on Biomedical Engineering*, 49(2):152–159.

Mendel, J. M. (1991). Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications. *Proceedings of the IEEE*, 79(3):278–305.

Nandi, A. K. (1999). *Blind Estimation using Higher-Order Statistics*, volume 1. Kluwer Academic Publichers, Boston, 1 edition.

Nikias, C. L. and Petropulu, A. P. (1993). *Higher-Order Spectra Analysis. A Non-Linear Signal Processing Framework.* Englewood Cliffs, NJ, Prentice-Hall.

Ömer Nezih Gerek and Ece, D. G. (2006). Power-quality event analysis using higher order cumulants and quadratic classifiers. *IEEE Transactions on Power Delivery*, 21(2):883–889.

Swami, A., Mendel, J. M., and Nikias, C. L. (2001). *Higher-Order Spectral Analysis Toolbox User's Guide*.

De la Rosa, J. J. G., Lloret, I., Puntonet, C. G., and Górriz, J. M. (2004). Higher-order statistics to detect and characterise termite emissions. *Electronics Letters*, 40(20):1316–1317. Ultrasonics.

De la Rosa, J. J. G., Moreno, A., Lloret, I., Puntonet, C. G., and Górriz, J. M. (2007). Power transients characterization and classification using higuer-order cumulants and competitive layers. *Lecture Notes in Computer Science (LNCS)*, 4431:782–789. International Conference on Adaptative and Natural Computung Algorithms, ICANNGA 2007 Warsaw, Poland, April 11-14, 2007, Proceedings, Part I; http://icannga07.ee.pw.edu.pl/.

De la Rosa, J. J. G., Puntonet, C. G., and Lloret, I. (2005). An application of the independent component analysis to monitor acoustic emission signals generated by termite activity in wood. *Measurement (Ed. Elsevier)*, 37(1):63–76. Available online 12 October 2004.

De la Rosa, J. J. G. and Ruzzante, R. P. J. (2007). Third-order spectral characterization of acoustic emission signals in ring-type samples from steel pipes for the oil industry. *Mechanical systems and Signal Processing (Ed. Elsevier)*, 21(Issue 4):1917–1926. Available online 10 October 2006.