

Power transients characterization and classification using higher-order cumulants and neural networks

Juan-José González de-la-Rosa^{1,3}, Antonio Moreno Muñoz^{2,3}, A. Luque³

¹University of Cádiz, Electronics Area, EPSA, Av. Ramón Puyol S/N, E-11202, Algeciras-Cádiz, Spain
Phone: +34 956 028020, Fax: +34 956 028001
E-mail: juanjose.delarosa@uca.es

²University of Córdoba, Electronics Area, EPS, Campus Rabanales, E-18071, Córdoba-Spain
E-mail: amoreno@uco.es

³Research Group PAI-TIC168-Computational Instrumentation and Industrial Electronics

Abstract – This paper deals with power-quality (PQ) event characterization using higher-order cumulants. Their maxima and minima are the main features, and classification is based in competitive layers. We concentrate on differentiating two types of transients (short duration and long duration). By measuring the fourth-order cumulants' maxima and minima, we build the two-dimensional feature measured vector. Cumulants are calculated over high-pass filtered signals, to avoid the 50-Hz signal. We have observed that the minima and maxima produce clusters in the feature space for 4th-order cumulants; third-order cumulants are not capable of differentiate these two very similar PQs. The experience sets the foundations of an automatic procedure.

Keywords – Competitive layers, Cumulants, Higher-Order statistics, Neural networks, Power quality, Transient detection.

I. INTRODUCTION

Power quality event detection and classification is gaining importance due to worldwide use of delicate electronic devices. Things like lightning, large switching loads, non-linear load stresses, inadequate or incorrect wiring and grounding or accidents involving electric lines, can create problems to sensitive equipment, if it is designed to operate within narrow voltage limits, or if it does not incorporate the capability of filtering fluctuations in the electrical supply [1, 2].

The solution for a PQ problem implies the acquisition and monitoring of long data records from the energy distribution system, along with an automated detection and classification strategy which allows to identify the cause of these voltage anomalies. Signal processing tools have been widely used for this purpose, and are mainly based in spectral analysis and wavelet transforms. These second-order methods, the most familiar to the scientific community, are based on the independence of the spectral components and evolution of the spectrum in the time domain. Other tools are threshold-based functions, linear classifiers and Bayesian networks. The goal of the

signal processing analysis is to get a feature vector from the data record under study, which constitute the input to the computational intelligence modulus, which has the task of classification.

Some recent works bring a different strategy, based in higher-order statistics (HOS), in dealing with the analysis of transients within PQ analysis [2] and other fields of Science [3]. Without perturbation, the 50-Hz of the voltage waveform exhibits a Gaussian behavior. Deviations from Gaussianity can be detected and characterized via HOS. Non-Gaussian processes need third and fourth order statistical characterization in order to be recognized. In other words, second order moments and cumulants could be not capable of differentiate non-Gaussian events.

The situation described matches the problem of differentiating between a transient of long duration named fault (within a signal period), and a short duration transient (25 per cent of a cycle). This one could also bring the 50-Hz voltage to zero instantly and, generally affects the sinusoid dramatically. By the contrary, the long-duration transient could be considered as a modulating signal (the 50-Hz signal is the carrier). These transients are intrinsically non-stationary, so it is necessary a battery of observations (sample registers) to obtain a reliable characterization.

The main contribution of this work consists of the application of higher-order central cumulants to characterize PQ events (could be seen as a complement of [2]), along with the use of a competitive layer as the classification tool. Results reveal that two different clusters, associated to both types of transients, can be recognized in the 2D graph. The successful results convey the idea that the physical underlying processes associated to the analyzed transients, generate different types of deviations from the typical effects that the noise cause in the 50-Hz sinusoid voltage waveform.

The paper is organized as follows: Section II summarizes the main equations of the cumulants used in the paper. Section

III recalls the competitive layer's foundations, along with the *Kohonen* learning rule. The experience is described in Section IV, and the conclusions are drawn in Section V.

II. HIGHER-ORDER CUMULANTS

High-order statistics, known as cumulants, are used to infer new properties about the data of non-Gaussian processes [4–6]. The relationship among the cumulants 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 [4, 5, 7, 8]:

$$\begin{aligned} Cum(x_1, \dots, x_r) = \sum (-1)^{p-1} \cdot (p-1)! \cdot E\left\{ \prod_{i \in v_1} x_i \right\} \\ \cdot E\left\{ \prod_{j \in v_2} x_j \right\} \cdots E\left\{ \prod_{k \in v_p} x_k \right\}, \end{aligned} \quad (1)$$

where the addition operator is extended over all the set of v_i ($1 \leq i \leq p \leq r$) and v_i compose a partition of $1, \dots, r$.

Let $\{x(t)\}$ be an r th-order stationary random process. The r th-order cumulant is defined as the joint r th-order cumulant of the random variables $x(t), x(t+\tau_1), \dots, x(t+\tau_{r-1})$,

$$\begin{aligned} C_{r,x}(\tau_1, \tau_2, \dots, \tau_{r-1}) \\ = Cum[x(t), x(t+\tau_1), \dots, x(t+\tau_{r-1})]. \end{aligned} \quad (2)$$

Considering $\tau_1 = \tau_2 = \tau_3 = 0$ in Eq. (2), we have some particular cases:

$$\gamma_{2,x} = E\{x^2(t)\} = C_{2,x}(0) \quad (3a)$$

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

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

Equations (3) are measurements of the variance, skewness and kurtosis of the distribution in terms of cumulants at zero lags (the central cumulants). 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.

III. COMPETITIVE LAYERS: A BRIEF SUMMARY

The neurons in a competitive layer distribute themselves to recognize frequently presented input vectors. The competitive transfer function accepts a net input vector \mathbf{p} for a layer (each neuron competes to respond to \mathbf{p}) and returns neuron outputs of 0 for all neurons except for the winner, the one associated with the most positive element of net input. If all biases are 0, then the neuron whose weight vector is closest to the input vector has the least negative net input and, therefore, wins the competition to output a 1.

The winning neuron will move closer to the input, after this has been presented. The weights of the winning neuron are adjusted with the *Kohonen* learning rule. Supposing that the i th-neuron wins, the elements of the i th-row of the input weight matrix (\mathbf{IW}) are adjusted as shown in Eq. (4):

$$\mathbf{IW}_i^{1,1}(q) = \mathbf{IW}_i^{1,1}(q-1) + \alpha [\mathbf{p}(q) - \mathbf{IW}_i^{1,1}(q-1)], \quad (4)$$

where \mathbf{p} is the input vector, q is the time instant, and α is the learning rate parameter. The *Kohonen* rule allows the weights of a neuron to learn an input vector, so it is useful in recognition applications. Thus, the neuron whose weight vector was closest to the input vector is updated to be even closer. The result is that the winning neuron is more likely to win the competition the next time a similar vector is presented. As more and more inputs are presented, each neuron in the layer closest to a group of input vectors soon adjusts its weight vector toward those inputs. Eventually, if there are enough neurons, every cluster of similar input vectors will have a neuron that outputs 1 when a vector in the cluster is presented, while outputting a 0 at all other times. Thus, the competitive network learns to categorize the input vectors it sees.

IV. EXPERIMENTAL RESULTS

The aim 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 cumulants. Each vector's coordinate corresponds to the local maximum and minimum of the 4th-order central cumulant. 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, 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 (5th-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 50 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. 1 and Fig. 2 show an example of signal processing analysis of two sample registers corresponding to a long-duration and a short-duration events, respectively.

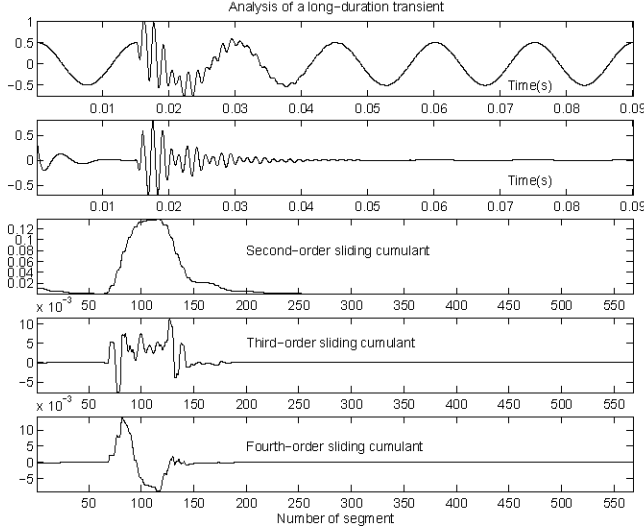


Fig. 1. Long-duration transient analysis. From top to bottom: the original data record, the filtered sequence, 2nd-3rd-4th-order central cumulants sliding windows, respectively.

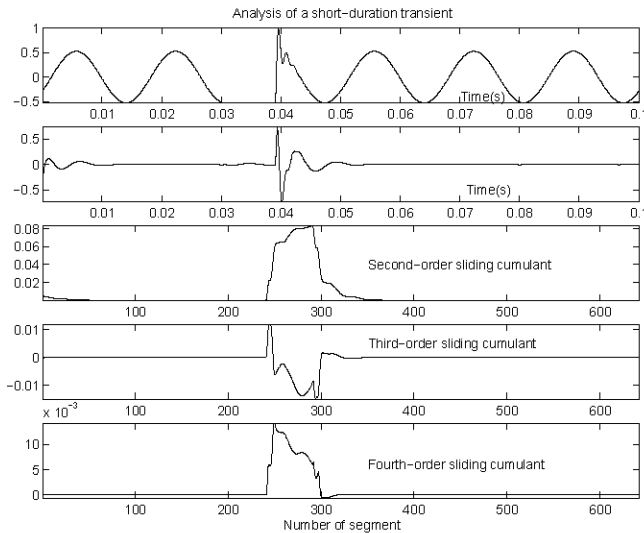


Fig. 2. Short-duration transient analysis.

The second-order cumulant sequence corresponds to the variance, which clearly indicates the presence of an event. 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 calculated. An unbiased estimator of the cumulants have 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 3rd-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 graph in Figs. 2 and 1), but also for the different location of minima, which suggest a clustering for the points.

The following graphs are intended for use as a visual comparison of the cumulants' sequences for the two classes of events. Fig. 3 shows the difficulty of getting a criterion in order to distinguish one transient from the other, if we attend to the maximum and minimum of the sliding cumulant. Besides, the shapes are very similar (for both types of events) and highly dependent on the window length, we have chosen to compute the sliding cumulants. On the contrary, Fig. 4 conveys the idea of an easier method of detection, because the shapes are quite different, along with the possibility of getting different extreme points. So, we focuss on 4th-order cumulants' sequence.

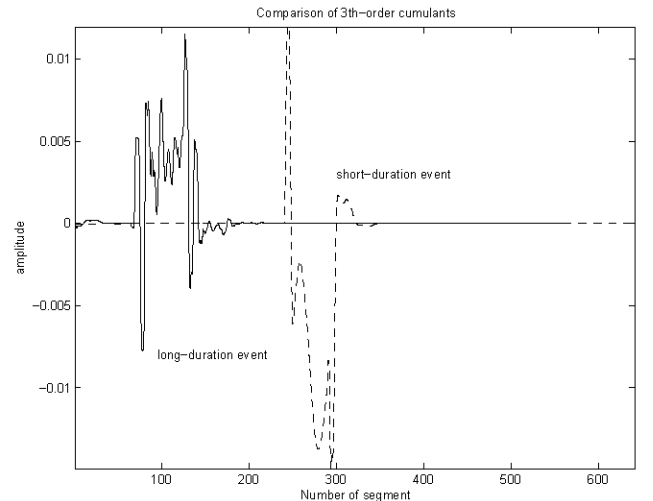


Fig. 3. Comparison of 3rd-order cumulants for two types of transients.

For each sample register the sliding 4th-order cumulants' sequence are calculated. For each sequence, the maximum and the minimum are detected and selected as a point in the feature space. Fig. 5 presents the results of the training stage, using

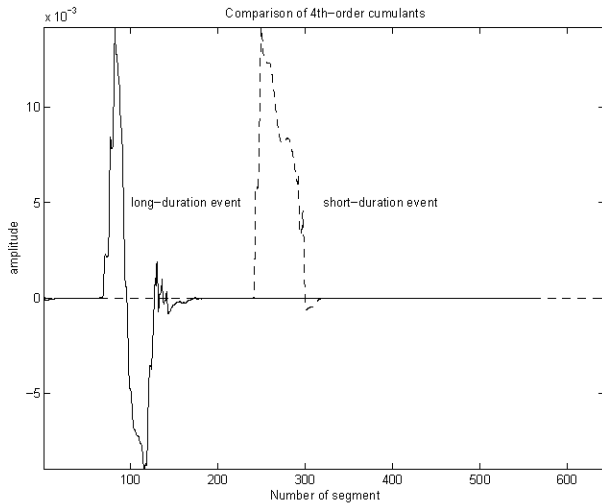


Fig. 4. Comparison of 4th-order cumulants for two types of transients.

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).

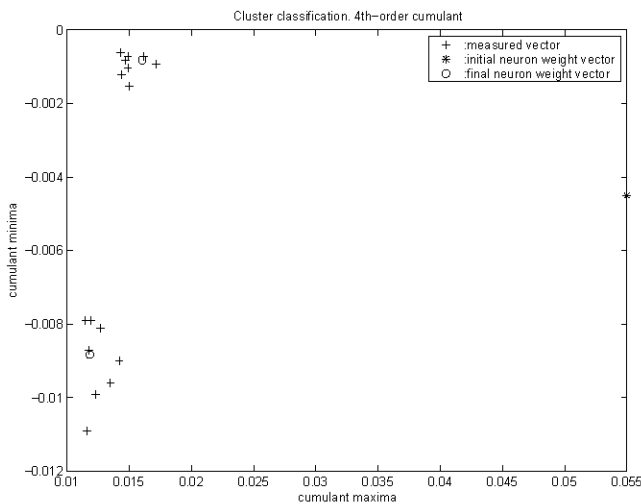


Fig. 5. Competitive layer training results over 20 epochs. Upper cluster: Short-duration PQ-events. Down cluster: Long-duration event.

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 has been conducted. This informs if the number of experiments is statistically significant according to the fitness test [2]. As a result of the test, the number of measurements is significantly correct.

V. CONCLUSIONS AND ACCOMPLISHMENTS

In this paper we have proposed a method to detect and classify two PQ 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 [2] 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.

ACKNOWLEDGEMENT

The authors would like to thank the *Spanish Ministry of Science and Education* for funding the project DPI2003-00878, where noise processes have been modeled and contrasted; and also for supporting the PETRI project PTR95-0824-OP dealing with higher-order statistics. Our unforgettable thanks to the trust we have from the *Andalusian Government* for funding the excellency project TIC-155.

REFERENCES

- [1] A. Moreno, V. Pallarés, J. J. G. de la Rosa, and P. Galisteo, "Study of voltage sag in a highly automated plant," in *MELECON 2006, Proceedings of the 2006 13th IEEE Mediterranean Electrotechnical Conference*.
- [2] Ömer Nezih Gerek and D. G. Ece, "Power-quality event analysis using higher order cumulants and quadratic classifiers," *IEEE Transactions on Power Delivery*, vol. 21, no. 2, pp. 883–889, April 2006.
- [3] J. J. G. de la Rosa, C. G. Puntonet, I. Lloret, and J. M. Górriz, "Wavelets and wavelet packets applied to termite detection," *Lecture Notes in Computer Science (LNCS)*, vol. 3514, pp. 900–907, May 2005, computational Science - ICCS 2005: 5th International Conference, GA Atlanta, USA, May 22–25, 2005, Proceedings, Part I.
- [4] C. L. Nikias and J. M. Mendel, "Signal processing with higher-order spectra," *IEEE Signal Processing Magazine*, pp. 10–37, 1993.
- [5] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications," *Proceedings of the IEEE*, vol. 79, no. 3, pp. 278–305, 1991.
- [6] A. K. Nandi, *Blind Estimation using Higher-Order Statistics*, 1st ed., A. K. Nandi, Ed. Boston: Kluwer Academic Publishers, 1999, vol. 1, no. 1.
- [7] J. J. G. De la Rosa, C. G. Puntonet, and I. Lloret, "An application of the independent component analysis to monitor acoustic emission signals generated by termite activity in wood," *Measurement (Ed. Elsevier)*, vol. 37, no. 1, pp. 63–76, January 2005, available online 12 October 2004.
- [8] C. L. Nikias and A. P. Petropulu, *Higher-Order Spectra Analysis. A Non-Linear Signal Processing Framework*. Englewood Cliffs, NJ, Prentice-Hall, 1993.