Juan José GONZÁLEZ DE LA ROSA[1], **Agustín AGÜERA-PÉREZ**[1], **José Carlos PALOMARES-SALAS**[1], **Antonio MORENO-MUÑOZ**[2]

Research Group PAIDI-TIC-168. Area of Electronics.

Amplitude-Frequency Monitoring of Power Quality Transients using Higher-Order Statistics and Self-Organizing Neural Networks

Streszczenie. *W artykule opisano automatyczną metodę klasyfikacji jakości energii w stanach przejściowych z uwzględnieniem amplitudy, częstotliwości i wartości ekstremalnych. W pierwszym etapie przeprowadzane są pomiary statystyczne dla stałej amplitudy i częstotliwości uwzględniające klastry 2D i właściwości trzeciego i czwartego rzędu towarzyszące anomaliom. Następnie uwzględniana jest geometria sieci. Po tym etapie włączany jest moduł sztucznej inteligencji bazujący na sieciach neuronowych.* ***(Monitorowanie jakości energii w stanach przejściowych przy użyciu statystyki wyższego rzędu i sieci neuronowych)***

Abstract. *In this paper a smart automatic classification of PQ transients is performed attending to their amplitudes and frequencies, and the extreme of higher-order cumulants. Feature extraction stage is double folded. First, these statistical measurements reveal the hidden geometry for a constant amplitude or frequency, conforming the 2D clustering grace to the third and fourth-order features associated to each signal anomaly, coupled to the 50-Hz power line. Precisely the main contribution of the work is the novel finding that the maxima and the minima of the higher-order cumulants distribute according to curves families, each of which associated to the transient's frequency or amplitude. Given a statistical order, each datum in a curve corresponds to the initial amplitude (or constant frequency), and to a couple of extremes (min-max) associated to the statistical estimator. The random grouping along each curve reveals the a priori hidden geometry, linked to the subjacent electrical phenomenon. Secondly, the regular surface grid in the input space (amplitude-frequency) experiments a transformation to the output space which is developed by the higher-order statistics. Once the geometry in the feature space has been found, we show the computational intelligence modulus, based in Self-Organizing Maps, which performs satisfactory learning along each frequency and amplitude curve. Performance of a four-neuron network with different geometries is shown, confirming the curves' patterns.*

Słowa kluczowe: jakość energii, sztuczne sieci neuronowe, statystyka wyższego rzędu.
Keywords: Higher-Order Statistics (HOS); Artificial Neural Networks (ANN); Power-Quality (PQ); Self Organizing Maps (SOM); Transients.

Introduction

Companies are endeavoring in innovation and new technologies to monitor and control Power Quality (PQ) anomalies, because today's equipment, and automated manufacturing devices are highly sensitive to the power line signal's imperfections (PQ events), making unaffordable the production costs. Malfunctioning not only has to be thereby detected, but also predicted and undoubtedly diagnosed to identify the cause and prevent the system from a similar shock; thus conveying the idea of a smart automatic detection. As a consequence, design improvement would be reflected *a posteriori* in an enhancement of the industrial production [1], [2],[3].

Solving a PQ incidence implies the acquisition and monitoring myriads of vast data records from the system under test, along with an automated detection and classification procedure, which allows the identification of the cause of the anomalies. These perturbations can be considered as transients, and are intrinsically non-stationary, so it is necessary a battery of observations (sample registers) to obtain a reliable characterization. The goal of the signal processing modulus is to get a feature vector from each data record, to drive the input to the computational intelligence unit, which assumes the role of classifier. The algorithms in the measurement systems for this purpose have been traditionally based in spectral analysis and wavelet transforms, and complementary second-order methods based on the independence of the spectral components and the time evolution of the spectrum. Other tools are threshold-based functions, linear classifiers and Bayesian networks. These algorithms are highly sensitive to noise processes; thereby in most real-life situations do not perform as expected with synthetics [4],[5].

During the last decade an emerging tendency, rooted in higher-order statistics (HOS), has brought innovative results dealing with PQ analysis [6, 7, 4], and other fields of Science and Technology [8, 9, 10]. These advances are based in the following hypothesis. Without perturbation, the 50-Hz of the voltage waveform exhibits a constant statistical behavior (stationarity), generally Gaussian. Deviations can be detected and characterized via HOS; non-Gaussian processes need at least 3^{rd} and 4^{th}-order statistical characterization in order to be completely targeted, because 2^{nd}-order moments and cumulants are not capable of differentiating non-Gaussian events; e.g., the problem of distinguishing between a transient of long duration (the so called oscillatory, within a signal period) and a short duration transient (or impulsive transient, 25 per cent of a cycle), has been outcome under controlled conditions in [5], and the idea of differentiating between healthy signals and signals with transients was pointed out and accomplished in [11]. This problem was also previously announced and described in [12], and falls into the HOS ensemble in the following sense. The short transient could also bring the 50-Hz voltage to zero instantly and, generally affects the sinusoid dramatically. By the contrary, the long-duration transient is considered a modulating signal (the 50-Hz signal is the carrier), and it is associated to load charges [12]. Consequently, for a given time record, and once the analysis has been done, we can suggest the PQ type phenomena, and perhaps the origin of the fault in the energy distribution system.

Thus, from the basis of the former research, this paper conveys the idea of the inverse problem in PQ analysis. That is, given a set of features, to guess the rest of the electrical anomaly's characteristics, and occasionally, to target the faulty element in the energy chain or plant. More precisely, the present work is centered in the higher-order characterization of transients according to the maxima and the minima of the 3^{rd} and 4^{th}-order central cumulants at zero lags (directly related to the skewness and kurtosis). It is shown that the geometrical distribution of these extremes adopt the form of curves families depending on the transient's frequency and initial amplitude. For a constant frequency (amplitude) curve, we concentrate on amplitude (frequency) fluctuations; the coordinates of a point are the skewness maxima and minima (for the 3^{rd} order) and the homologous kurtosis' ex-

tremes (for the 4^{th}-order). Each 2D datum in a curve is the computational result over a time-domain register under test, which in turn is thought to contain a single transient, coupled to the undistorted sine wave. For each constant-frequency curve 50 transients have been selected with an uniformly distributed random initial amplitude. Similarly, for each constant-amplitude curve the frequency has been swept linearly within a given range.

The results reveal the strong relationship between the frequency and the amplitude of a concrete transient and its associated higher-order features. This is demonstrated via a mathematical transformation which maps the 2D (amplitude-frequency) input space into two output 2D spaces which relate higher-order features to the frequency and the amplitude. Concretely, the main contribution of the paper is the novel finding that the maxima and the minima of the higher-order cumulants distribute along these families of curves of constant frequency or constant amplitude. A couple of extremes (min-max) of the higher-order statistical estimator belongs to a couple of curves (constant-frequency and constant-amplitude), therefore the image position is given by the cross point of the curves. The random grouping along each curve reveals the *a priori* hidden geometry, in turn linked to the subjacent electrical anomaly.

The computational intelligence stage shows the performance of competitive layers and self-organizing maps, classifying amplitudes within a constant frequency curve and *vice versa*. We conclude that a double characterization of the transient can be performed based in the position in the higher-order spaces, once the distribution of iso-frequency lines and iso-amplitude lines is known.

The paper is structured as follows. The following section explains the fundamentals and the importance for PQ monitoring. Higher-Order Statistics are outlined then, to be followed by a summary on competitive layers and self-organizing networks. Finally, results are presented and conclusions are drawn.

Power quality events' characterization

As more and more electronic equipments enter the residential zones and business environment, the subjects related to PQ and its relationship to vulnerability of installations is becoming an increasing concern to the end users. Particularly has arisen and increased the need to protect sensitive electronic equipment from damaging over-voltages. 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 [12, 13, 2, 14].

The two main regulated aspects of PQ are the following:

- Technical PQ, which includes: Continuity of supply or reliability (long interruptions) and Voltage quality (voltage level variations and voltage disturbances).
- Commercial services associated to the wires (such as the delay to get connected to the grid, etc.) as well as commercial services for energy retail to regulated customers.

Assessment of voltage quality and power disturbances involves looking at electromagnetic deviations of the voltage or current from the ideal single-frequency sine wave of constant amplitude and frequency. A consistent set of definitions can be found in [15]. Regulation in European countries proposes to use the standard EN-50160 to define the voltage quality ranges. This norm actually describes the electricity through the technical characteristics that it has to fulfill to be considered as a compliant product. But there are a lot of undefined aspects; besides the fact that most of the regulator has yet to publish the technical criteria to measure and control all the voltage quality characteristics and decide what would be the penalization. The fact is that the only voltage quality aspect that is now enforced is the maximum voltage level variation settled to ± 7 $\%$ (which is actually different to the ± 10 $\%$ fixed on the EN-50160). But even this aspect is not yet controlled and there is not any defined procedure to determine if the limit has been reached.

On the other hand, the presence of disturbances on power distribution also affect the energy efficiency of the system. As far as energy efficiency is concerned in a power distribution system, the two dominant factors in PQ are its unbalanced and harmonic distortion. In an electrical installation when single-phase loads (especially those with non-linear characteristics), are not evenly and reasonably distributed among the three-phases of the supply, we are in the presence of unbalance. Voltage unbalance in a three-phase system causes three-phase motors to draw unbalanced current. This phenomenon causes additional power losses in conductors and motors and can cause the rotor of a motor to overheat.

Among all categories of electrical disturbances, the voltage sag (dip) and momentary interruption are the nemeses of the automated industrial processes. Voltage sag is commonly defined as any low voltage event between 10 and 90$\%$ of the nominal RMS voltage lasting between 0.5 and 60 cycles. Momentary voltage interruption is any low-voltage event of less than 10$\%$ of the nominal RMS voltage lasting between 0.5 cycles and 3 seconds. In medium voltage distribution networks, voltage sags are mainly caused by power system faults. Fault occurrences elsewhere can generate voltage sags affecting consumers differently according to their location in the electrical network. Even though the load current is small compared to the fault current, the changes in load current during and after the fault strongly influence the voltage at the equipment terminals. It has been discovered that the 85$\%$ of power supply malfunctions attributed to poor PQ are caused by voltage sag or interruptions of fewer than one second duration. Starting large motors can also generate voltage sags, although usually not so severe. In comparison with interruptions, voltage sags affect a larger number of customers and for some customers voltage sags may cause extremely serious problems. These can create problems to sensitive equipment if it is designed to operate within narrow voltage limits, or it does not have adequate ride-through capabilities to filter out fluctuations in the electrical supply.

Over-voltage is an RMS increase in the AC voltage, at the power frequency, for durations greater than a few seconds, and can be the result of a programmed utility operation, or the effect of an external eventuality [16]. Under normal operating conditions, the steady-state voltage is regulated by the utility within a limits band accepted by the EN-50160. Deviations from these limits are rare, and the utility can actuate readily to correct them, if known their occurrence, by acting on conventional distribution technologies, such as tap-changing transformers [17].

However, under the typical operating conditions of a power system there is risk of damaging due to a momentary excess of voltage. Although by themselves they would be described as "abnormal", it is possible to distinguish be-

tween surges and swells. A surge is an over-voltage that can reach thousands of volts, lasting less than one cycle of the power frequency, that is, less than 16 milliseconds. A swell is longer, up to a few seconds, but does not exceed about twice the normal line voltage.

Power system surges, based on waveform shapes, can be classified into "oscillatory transients" and "impulsive transients" [13, 15] and they are the goal of the present research work. Oscillatory transient surges show a damped oscillation with a frequency range from 400 Hz to 5 kHz or more. Impulsive transient surges present a fast rise time in the order of 1 ns-10 μs over the steady state condition of voltage, current or both, that is unidirectional in polarity (primarily either positive or negative), reaching hardly twice the peak amplitude of the signal. They are damped quickly, presenting a frequency range from 4 kHz to 5 MHz, occasionally reaching 30 MHz.

Categorization of electrical transients based on waveform shapes and their subjacent causes (or events) has been studied in [12], and a few previous studies [6, 7] using HOS for feature extraction of electrical signals have shown the possibility of distinguish transients based on details beyond the second-order. In a real-life 50-Hz power line signal, it is very common to find these transients. In Fig. 1 we show an example of anomalous real-life signal, which includes a transient which is not classified into short-duration and long-duration. It is shown the computation results of three higher-order time-domain statistics in order to introduce them qualitatively. The second-order estimator operates as an increase-of-power detector, showing the bumps associated to the increase of power, which in turn are associated to the anomalies of the power-line sine wave, but the third and fourth-order sliding cumulants have to be interpreted deeper. The most intuitive method is to calculate their maxima and minima.

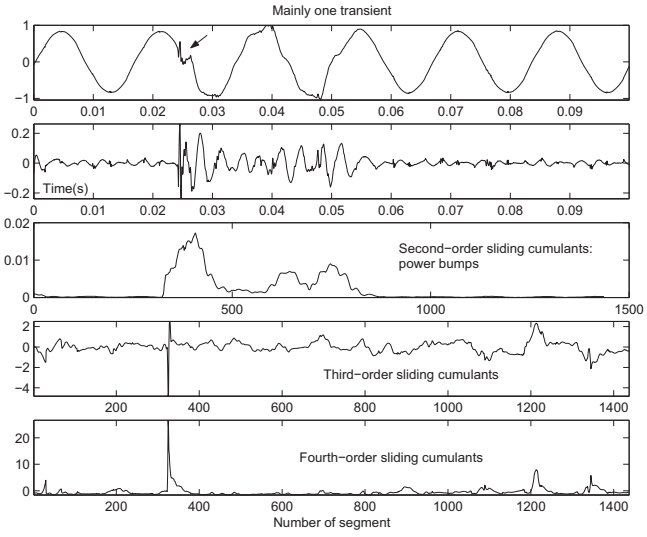

Fig. 1. A transient coupled in the power line 50-Hz sine wave, and the computation of time-domain higher-order statistics. The record has been previously normalized and high-pass filtered with the goal of extracting the anomaly.

Once the foundations of PQ have been presented, in the following Section higher-order statistics in the time-domain are introduced in order to present the signal processing guts, along with a basic example which shows the performance of the statistical estimators, which have been used in the computation of the cumulants. This example also motivates the use of HOS in time-series characterization.

Higher-Order Statistics

Higher-order cumulants are used to infer new properties about the data of non-Gaussian processes [18, 10]. In multiple-signal processing scenarios it is frequently used the *Leonov-Shiryaev* formula [19, 20], which states the combinational relationship among the cumulants of *r* stochastic signals, $\{x_i\}_{i\in[1,r]}$, and their associated moments of order *p*, $p \leq r$.

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

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

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

(2)
$$C_{r,x}(\tau_1,\tau_2,\ldots,\tau_{r-1}) = Cum[x(t), x(t+\tau_1),\ldots,x(t+\tau_{r-1})]$$

The second-, third- and fourth-order cumulants of zero-mean *x(t)* can be estimated via [21]:

(3a)
$$C_{2,x}(\tau) = E\{x(t)\cdot x(t+\tau)\}$$

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

(3c)
$$\begin{aligned} &C_{4,x}(\tau_1,\tau_2,\tau_3) \\ &= E\{x(t)\cdot x(t+\tau_1)\cdot x(t+\tau_2)\cdot x(t+\tau_3)\} \\ &- C_{2,x}(\tau_1)C_{2,x}(\tau_2-\tau_3) \\ &- C_{2,x}(\tau_2)C_{2,x}(\tau_3-\tau_1) \\ &- C_{2,x}(\tau_3)C_{2,x}(\tau_1-\tau_2) \end{aligned}$$

By putting $\tau_1 = \tau_2 = \tau_3 = 0$ in Eq. (3), it is obtained:

(4a)
$$\gamma_{2,x} = E\{x^2(t)\} = C_{2,x}(0)$$

(4b)
$$\gamma_{3,x} = E\{x^3(t)\} = C_{3,x}(0,0)$$

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

The expressions in Eq. (4) are estimates of the variance, skewness and kurtosis of the probability distribution in terms of the central cumulants at zero lags.

The normalized kurtosis and skewness are defined as $\gamma_{4,x}/(\gamma_{2,x})^2$ and $\gamma_{3,x}/(\gamma_{2,x})^{3/2}$, respectively. Normalized quantities 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*). In the experimental section, results are obtained by using sliding cumulants, i.d. a moving (sliding) window in the time domain over which to compute

each cumulant (*3*rd-order and *4*th-order central cumulants for zero time-lags).

To show the relevance of HOS an illustrative example is arranged. Four noise processes (Gaussian, uniform, exponential and Laplacian) previously catalogued in [18], and indistinguishable from the second-order perspective (autocorrelation sequence), are presented in this subsection in order to illustrate the importance of higher-order cumulants. The *3*rd and *4*th-order cumulants are computed according to the unbiased estimate given in [5]. We consider a 2048-point sample register for each random set of data. The four identical autocorrelation sequences (not presented in figure) contrast to the third and fourth-order ones, where substantial differences are observed, specially those corresponding to zero time lags. This result can be seen in Fig. 2, where the *3*rd and the *4*th-order cumulants sequences are depicted. The theoretical values of the *3*rd and *4*th-order cumulants at zero time-lag are respectively: [0 (Gaussian), 0 (Uniform), 2 (Exponential), 0 (Laplacian)] and [0 (Gaussian), -1 (Uniform), 6 (Exponential), 12 (Laplacian)], according to [18]. The difference between the theoretical and the experimental values is due to the lack of averaging (only one sample register is considered). The convergency of the estimate is assured in [5].

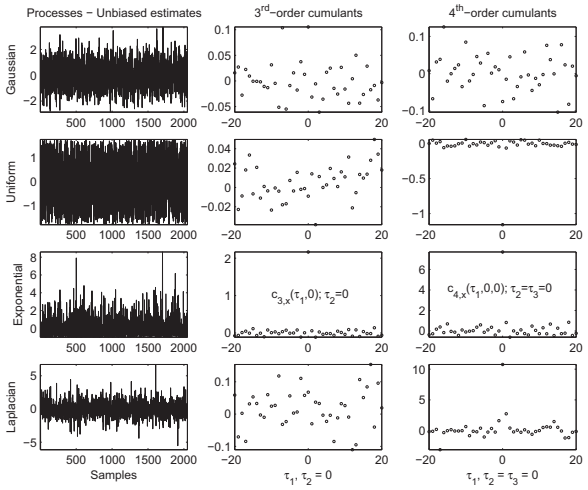

Fig. 2. *3*rd and *4*th-order cumulant sequences for the four noise processes.

Competitive layers and Self-Organizing Maps

Neurons in a competitive layer distribute themselves to recognize frequently presented inputs. The competitive transfer function accepts a net input vector **p** for a layer (each neuron competes to respond to **p**) and returns neuron outputs of 0 for all neurons except for the winner, which the one associated with the most positive element of net input. If all biases are 0, then the neuron whose weight vector is the closest to the input array has the least negative net input and, consequently, 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 (0.9 in the present case). Supposing that the *i*th-neuron wins, the elements in the *i*th-row of the input weight matrix (**IW**) are adjusted as shown in Eq. (5):

$$\text{(5)}\quad \mathbf{IW}_i^{1,1}(q) = \mathbf{IW}_i^{1,1}(q-1) + \alpha\left[\mathbf{p}\,(q) - \mathbf{IW}_i^{1,1}(q-1)\right],$$

with **p** the input vector, q the time instant, and α the learning rate. The neuron whose weight vector was closest to the input vector is updated and located even closer. The result is that the winning neuron is more likely to win the competition the next time a similar input is presented. Meanwhile new inputs are presented, each neuron in the layer closest to a group of input vectors soon adjusts its weights 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.

Self-Organizing Maps (SOM) learn to classify feature input vectors according to how they are grouped in the input space. SOM differ from competitive layers in that neighbor-neurons learn to recognize neighboring sections of the input space. Thus, SOM learn both the distribution (as do competitive layers) and topology of the input vectors they are trained on. Consequently, instead of updating only the winning neuron, all neurons in its neighborhood are updated using the *Kohonen* rule. The neurons in the layer of a SOM are arranged originally in physical positions according to a topology function. A distance function allows the calculation of the distances between neurons. Thus, for the *i*th neighboring neuron, in the *q*th instant, we have the weight vector **w**, given in Eq. (6):

$$\text{(6)}\quad \mathbf{w}_i(q) = \mathbf{w}_i(q-1) + \alpha\left[\mathbf{p}\,(q) - \mathbf{w}_i(q-1)\right].$$

Thus, when a vector is presented, the weights of the winning neuron and its closest neighbors move toward it. Consequently, after many presentations, neighboring neurons will have learned vectors similar to each other.

Data arrangement and feature extraction: input and output spaces

In this paper we have focussed in oscillatory transients, which is a very common anomaly, usually associated to parasitic discharges. The decaying law for the amplitude is exponential, and each transient vanishes within less than half a cycle (2 ms), corresponding to the power line. A sample frequency of 0.1 ms is chosen. Each time-series under test for computation performance contains eight cycles of the power-line, with one transient to be targeted.

Computation is explained over Fig. 3. For each register, the *3*th and *4*th-order cumulants are computed according to the following procedure. For a given statistical order (three or four) the feature extraction algorithm computes a cumulant over a sliding window, whose length is equal to two periods of the power line (20 ms), and then it jumps to the following starting point (overlapping 98 %). Besides, each *n*th-order cumulant, $Cum_{n,x}[i]$, associated to the *i*th computation window has been normalized by $(Cum_{2,x}[i])^2$, so that to give a real statistical characterization. When the computation swept along a signal finishes, the maximum and the minimum are calculated for all the sliding windows results. Consequently, each signal is characterized by its third and fourth-order extremes. The constant-value zone which has been remarked in Fig. 3, establishes another feature for the higher statistical orders. This fact is object for future works.

As mentioned before, the computation results in the feature extraction stage consist of two couples of extremes, associated to the *3*rd and *4*th-order cumulants, respectively. Since each pair can be interpreted as coordinates in a 2D space, we will consider that the mathematical procedure

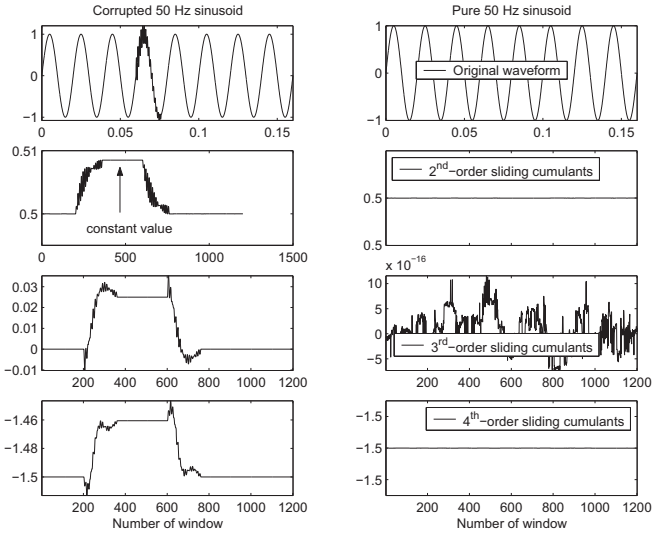

Fig. 3. Example of computation on a 1000 Hz transient, which shows the maxima and the minima associated to each cumulant time-series and the constant value zone, between the extremes, which will be used in future work.

characterizes each evaluated signal as two points, in two 2D spaces. One of them projected into the space defined by *3*rd-order cumulants' extremes (maxima and minima) and the other one into the *4*th-order analogous space. These spaces are called output spaces hereinafter.

In order to test if transients' characteristics can be inferred-tracked from the output spaces, an experiment is proposed. Transient's amplitude and frequency will be modulated and their projections will be analyzed. Hence, families of transients are generated and organized in two groups attending to: the named constant-frequency-variable-amplitude (CFVA) and constant-amplitude-variable-frequency (CAVF) groups.

A CFVA family of transients comprises a battery of 50 transients with equal frequencies and randomly selected initial amplitudes, distributed according to a uniform probability distribution between 0.1 and 0.9. We have tested the CFVA families associated to seven frequency values (200, 300, 500, 1000, 1500, 2000, 2500 Hz). In the same way, nine CAVF families, with amplitudes (0.1, ... , 0.9 V) and random frequencies between 200 - 2500 Hz, have been also tested.

Both for the CFVA and CAVF groups, *3*rd and *4*th order statistics are computed. As amplitude and frequency define a 2D space, it will be interpreted as the input space in order to describe the whole proposed methodology as a spatial transformation from $\Re^2$ to $\{\Re^2,\Re^2\}$.

Results

The results are organized in four couples of graphs (each figure contains a couple of graphs). Each couple is composed of the 2D graph and its associated 3D. The 3D graph allows a better understanding of the data assemblies, since the z-axis represents the random variable in each case. CFVA graphs are depicted in Fig. 4 and Fig. 5, and the homologous CAVF are Fig. 6 and Fig. 7.

The random conception of the experiment allows to appreciate that data are naturally arranged along their corresponding curves. These curves may be interpreted as isolines, since they are traced using a fixed parameter. Additionally, this finding means that the frequency (or amplitude) of the transient confines it to a curve along which, its amplitude (or frequency) determines the higher-order 2D coordinates (maxima and minima of the skewness and the kurtosis). This fact conveys the idea of an automatic learning process.

A number of remarks have to be made for these graphs. Attending to the CFVA graphs both for *3*rd and *4*th order (4 and 5) it is seen that as the amplitude decreases (moving towards the origin) exists a major data concentration nearby the origin, it becomes more difficult to distinguish a transient by its frequency, due to the close separation of the curves. For the *3*th-order curves, as the frequency increases curves confuse each other, thereby making difficult frequency identification. This fact is also observed in the *4*th-order family, but we have found an exception, in the curve corresponding to 2500 Hz. Higher frequencies curves would overlap the former high-frequency curves (below 2500 Hz).

If we pay attention to CAVF curves (Fig. 6 and Fig. 7), each effect observed in CFVA curves has a "mirror" or symmetrical phenomenon. Hence, as the amplitude decreases (moving towards the origin) the random distributions are enclosed or confined to shorter ranges. This fact was expected because CFVA curves are closer nearby the origin, limiting the range of dispersion. In the same way, as the frequency increases (moving right within each curve), extremes tend to overlap each other and to concentrate along a definite trend line, coincident with the observed overlapping of high-frequency curves in CFVA. A particular effect of CAVF curves is that as the frequency decreases points are scattered from the principal alignment observed in the main trend.

In the former graphs it is observed that each transient's frequency or amplitude is associated to a specific curve. Therefore, a point result of the intersection between a CFVA and a CAVF curve should define a transient with a specific frequency and amplitude. In other words, an input-space point, defined by a particular frequency and amplitude, must have associated an image within both CFVA and CAVF specific curves. This fact conveys the idea of a mapping or space projection, in which coordinates of the input space (amplitude-frequency) have an specific associated point in the output spaces.

In order to visualize the transformation, a mesh (grid) is mapped from the input space to the output ones (Fig. 8). Each cross-point, with a known amplitude and frequency, is mapped and then connected to their consecutive neighbors, using iso-frequency lines (continuous line in Fig. 8) and iso-amplitude line (discontinuous line in Fig. 8). Obviously, the resulting structure reminds the graphs obtained from the random experiment, but assembling the amplitude and frequency information.

The direct mapping has been explained with the goal of showing that the higher-order features associated to a given transient adopt spatial concrete patterns. Collaterally, it is shown that the inverse mapping action is equally described, and it would allow the inference of the transient's characteristics (amplitude and frequency) from the basis of the measurement of the higher-order features (skewness and kurtosis), i.e. from the output spaces, it would be possible to infer any frequency and amplitude in the studied range. This is a future work. However, the objective of the ANNs in the present paper is to corroborate the subjacent geometry, thereby to show that these output spaces' geometries can be learned, in order to develop an automatic classification system in a future research stage.

To start showing the engine of the artificial intelligence stage, based on neural classifiers, we have selected a representative situation, shown in Fig. 9. We have chosen the

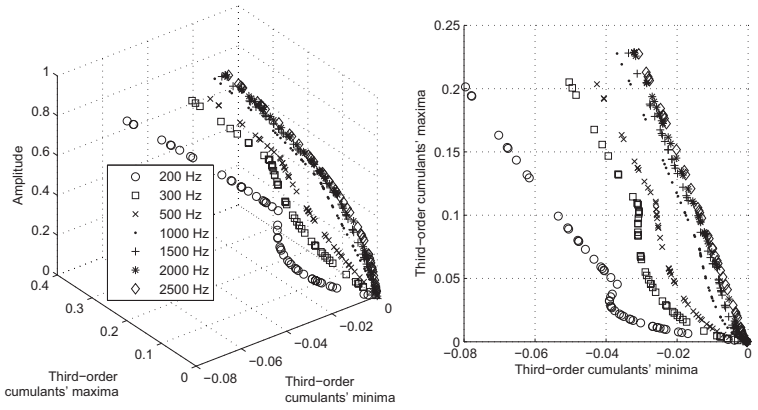

Fig. 4. Constant frequency curves corresponding to the third-order extremes. Each curve is associated to a frequency.

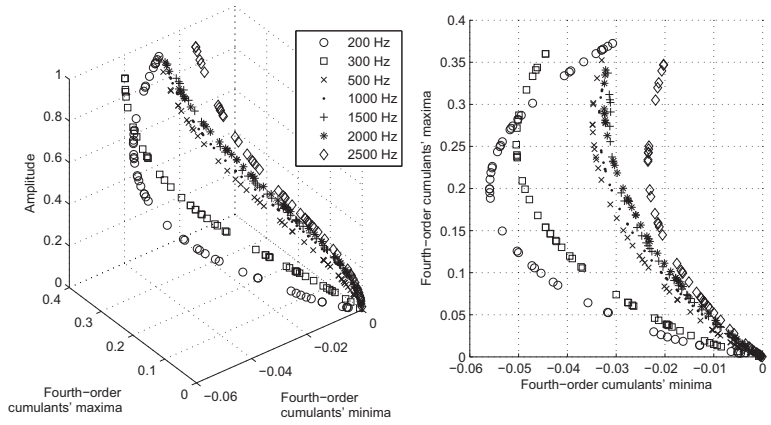

Fig. 5. Constant frequency curves corresponding to the fourth-order extremes. Each curve is associated to a frequency.

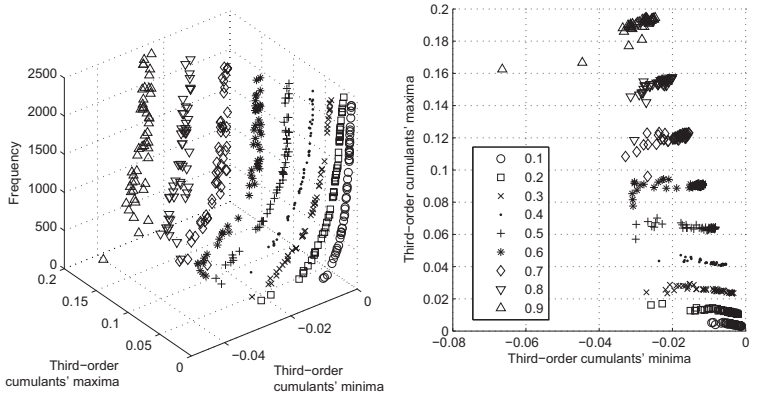

Fig. 6. Constant amplitude curves corresponding to the third-order extremes. Each curve is associated to an initial amplitude of the transient.

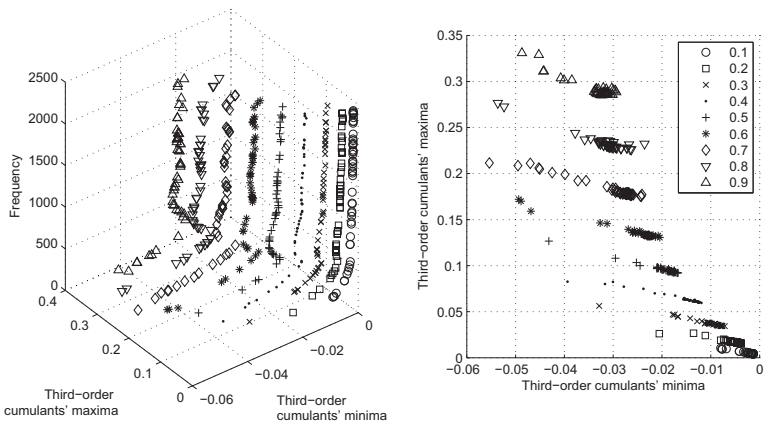

Fig. 7. Constant amplitude curves corresponding to the fourth-order extremes. Each curve is associated to an initial amplitude of the transient.

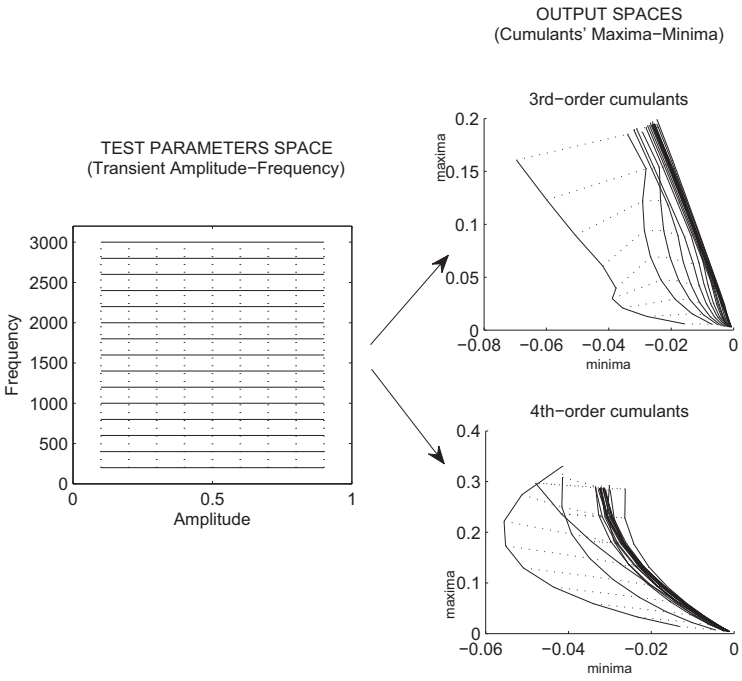

Fig. 8. Mapping process which shows the transformation of the input space into the output spaces, making the projections to the 3rd and 4th orders output spaces.

200 Hz-curve, as it is clear that high-frequency transients are difficult to be distinguish. A six neuron model has been chosen for the competitive layer (CL) training, and two common topologies have been chosen for the SOM neural network to be trained (hexagonal and random). Training results are almost the same for the 30 training sessions and for both topologies, then establishing the independence from the learning space. The SOM conception is based in a structure of 2×3 cell array. The election of these six starting centroids is totally free. Then, given a frequency, moving along a constant-frequency curve, we try to cluster the transient's amplitudes into six assemblies.

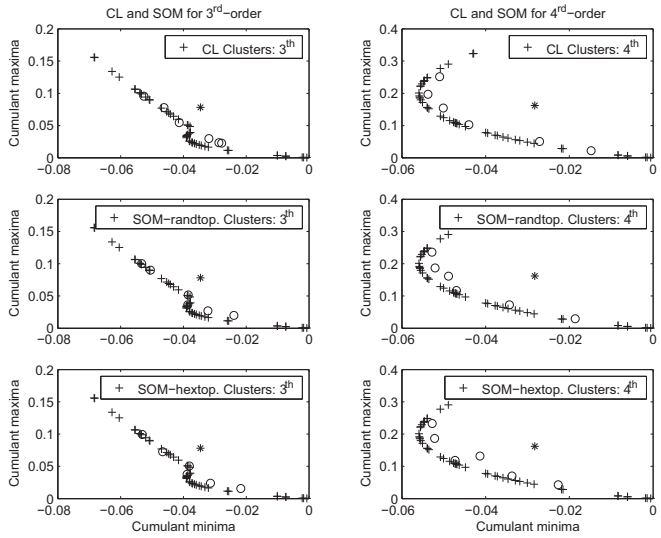

Fig. 9. Clusters in the 200-Hz curve. Left (right) column: CL and SOM performance for third-order (fourth).

Both for three and four orders, the separation between classes (inter-class distance) is well defined in this 2D feature graph for the SOM topologies, and only a fuzzy situation appears in the 3rd order graph. Consequently, it is possible to classify amplitudes into six classes or sets. The correct configuration of the clusters is a posteriori corroborated during the simulation of the three ANN models (CL and the two SOM topologies), in which we have obtained an approximate classification accuracy of 95 %.

Hereinafter we center on SOM performance, considering three topologies and four distance's functions for each of them. The following figures show the results for different SOM topologies, where it have selected 4 clusters for each ANN configuration, enough for our purpose. It can be seen that, if we maintain the frequency constant (e.g. 200 Hz), and making variable amplitude, for both 3rd and 4th order, all the ANN topologies are capable to locate the centroids alongside each data trace (crosses in the curves of Fig. 10 and Fig. 11). This process has been repeated for a frequencies range between 100 Hz and 4,500 Hz in steps of 100 Hz to both 3rd and 4th order. The "randtop" topology with "boxdist" distance's function is the one which best fits the cloud of measurements for the higher-order features (skewness and kurtosis), in 80.0 % and 75.5 % of the cases, respectively, so it is the configuration elected.

If we consider the frequency as a variable, results are depicted in Fig. 12 and in Fig.13 for order 3 and 4, respectively. It is observed that, as the frequency decreases (moving to the left) data scatter making difficult SOM performance. For all topologies, one cluster remains outside the natural data tendency. Similarly to the constant frequency case, the am-

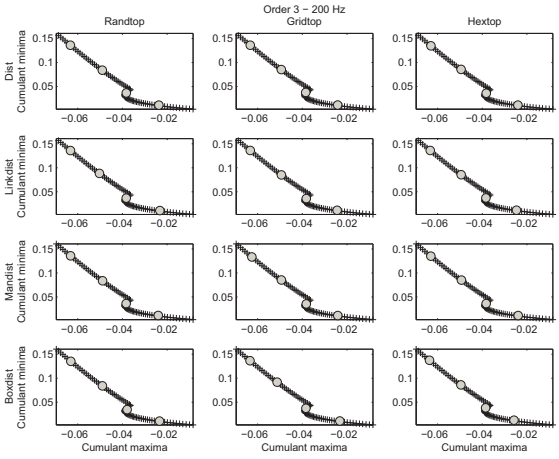

Fig. 10. SOM learning for order 3 using different topologies (Frequency 200 Hz).

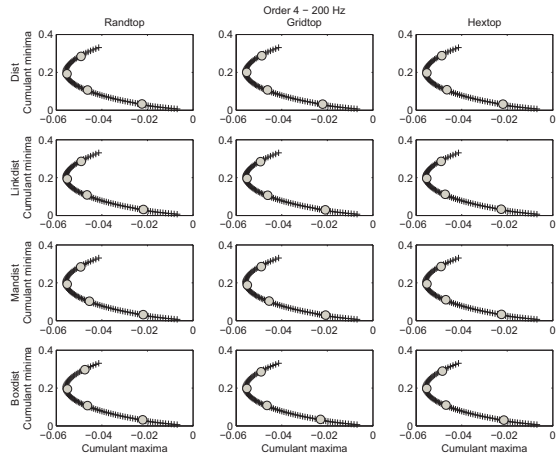

Fig. 11. Different possibilities of SOM learning and topologies for order 4 and for 200 Hz.

plitude was swept between 0.05 and 0.9 V in steps 0.01 V. The results are more dispersed than in constant frequency and we obtained a 61.9 % and 69.7 % for the configuration "randtop-boxdist" (topology - distance's function) *3*rd and *4*th order, respectively. Although these values are lower than constant frequency results, yet remain acceptable for our purposes, so this configuration is chosen again.

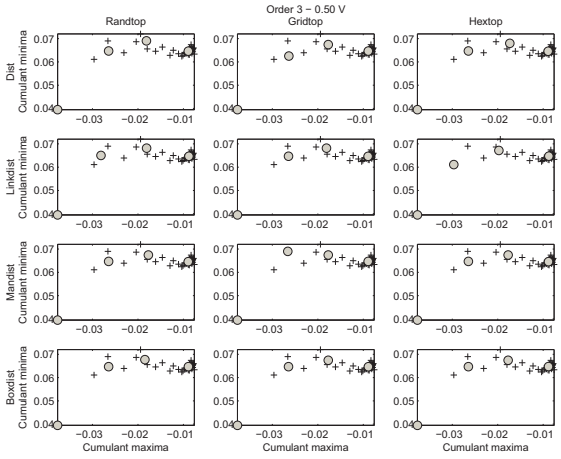

Fig. 12. SOM learning for order 3 using different topologies (Fixed amplitude 0.5 V).

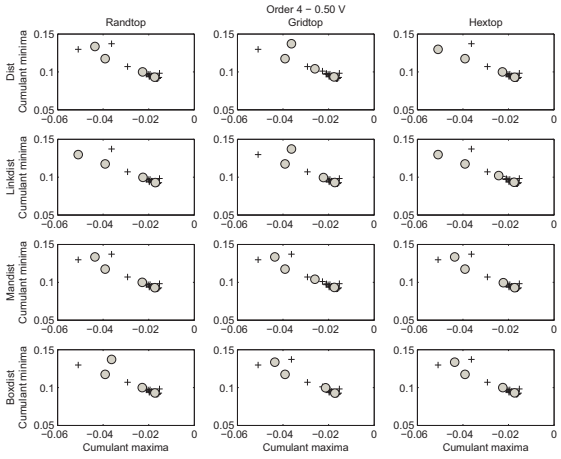

Fig. 13. Different possibilities of SOM learning, for 4rd-order, for a fixed amplitude (0.5 V).

Table 1 and table 2 show the summary of the training results associated to each topology and distance function, both for CFVA and for CAVF. The combination topology-distance which best fits the experimental data is "randtop-boxdist", both for CFVA and the CAVF cases, with percentages of success of 80.00 % and 61.63 % for *3*rd order. This means that this topology is the best fit to the 80.00 % of the tested frequencies and also the 61.63 % of the tested amplitudes.

During the simulation, new signals (randomly selected from our data base) were processed using this methodology. As expected, the accuracy of the classification results increases with the number of data. To evaluate the confidence of the statistics a significance test has been conducted. As a result, the number of measurements is significantly correct. With the respect to the frequency limit, it has been proved that transients of more than 1,500 Hz are not distinguished properly. As final remark, the experiment has been reproduced using a linear-decaying law for the transient, obtaining similar results.

Table 1. SOMs' training performance for different frequencies.

Statistic	Success	Configuration	Percentage
Skewness	36	randtop-boxdist	80.00
	4	randtop-linkdist	8.89
	3	randtop-dist	6.67
	2	randtop-mandist	4.44
Kurtosis	34	randtop-boxdist	75.55
	4	randtop-linkdist	8.89
	4	randtop-dist	8.89
	3	randtop-mandist	6.67

Table 2. SOMs' training performance for different amplitudes.

Statistic	Success	Configuration	Percentage
Skewness	53	randtop-boxdist	61.63
	8	randtop-linkdist	9.30
	13	randtop-dist	15.2
	12	randtop-mandist	13.95
Kurtosis	60	randtop-boxdist	69.77
	13	randtop-linkdist	15.11
	7	randtop-dist	8.14
	6	randtop-mandist	6.98

Conclusion

The results of the present paper indicate that an automatic HOS-based detection and classification system for PQ transients can be implemented. With a low computational cost, the system would process raw data to calculate the sliding skewness and the kurtosis and evaluate the extremes, which characterize each transient in a 2D space, either for the third and the fourth-order cumulants.

The most important finding is that the transients distribute themselves naturally along curves of constant frequency and constant amplitudes. Hence, once a signal with a coupled transient is analyzed, its image is located in a particular point of the higher-order spaces, determined by the transient's frequency and amplitude. It has been demonstrated that output spaces can be meshed, providing a background structure in which transients' characteristics can be tested, i. e., given the 2D coordinates in the higher-order space, the frequency and the amplitude of the transient under test can be estimated attending to the bounding isolines in its neighborhood. Thus, the worse measurement resolution is given by the cell's dimensions in the output space, being possible in practice a better inference's accuracy within each cell of the output space's grid. For high frequencies the cells in the grid are overlapped, making difficult to infer the frequencies, and consequently limiting the range of learning. On the other hand, but less important, the same phenomenon takes place for low amplitudes.

The CL and SOM configurations have shown that the learning process is topology-independent. Besides, clusters formation for a six-neuron topology shows the feasibility of amplitude discrimination inside each iso-frequency curve and *vice versa*. The fact that different SOM topologies exhibit the same learning performance along the same curve, indicates that the outlined structures in the curves' patterns exist, and the automatic procedure is capable of discover these

arrangements. Thereby, these patterns *a priori* exist, with independence of the scientific procedure used to reveal the subjacent.

Future work is planned with a twofold purpose. First, the objective is to deal with different types of PQ anomalies, like swells and sags. Secondly, we will focus on the development of a more precise learning process, using additional indicators, like *5*th and *6*th-order cumulants or moments, along with the constant value zone, which was forwarded in the example of Fig. 3. On the other hand it is compulsory to mention the possibility of using hybrid learning systems, including fuzzy logic and genetic algorithms or other type of ANN like backpropagation networks.

Acknowledgement

The authors would like to thank the *Spanish Ministry of Science and Innovation* for funding the research projects TEC2009-08988 and TEC2010-19242-C03-03 (SIDER-HOSAPQ). Our unforgettable thanks to the trust we have from the *Andalusian Government* for funding the Research Group PAIDI-TIC-168 in *Computational Instrumentation and Industrial Electronics-ICEI*.

BIBLIOGRAPHY

[1] M. H. J. Bollen, I. Y.-H. Gu, P. G. V. Axelberg, E. Styvaktakis, Classification of underlying causes of power quality disturbances: Deterministic versus statistical methods, EURASIP Journal on Advances in Signal Processing 2007 (Article ID-79747) (2007) 1–17.

[2] A. Moreno, *et al*, Mitigation Technologies in a Distributed Environment, 1st Edition, Power Systems, Springer-Verlag, 2007.

[3] M. Riera-Guasp, J. A. Antonino-Daviu, M. Pineda-Sánchez, R. Puche-Panadero, J. Pérez-Cruz, A general approach for the transient detection of slip-dependent fault components based on the discrete wavelet transform, IEEE Transactions on Industrial Electronics 55 (12) (2008) 4167–4180.

[4] M. V. Ribeiro, C. A. G. Marques, C. A. Duque, A. S. Cerqueira, J. L. R. Pereira, Detection of disturbances in voltage signals for power quality analysis using hos, EURASIP Journal on Advances in Signal Processing 2007 (Article ID-59786) (2007) 1–13.

[5] J. J. G. De la Rosa, A. M. Muñoz, A. Gallego, R. Piotrkowski, E. Castro, Higher-order characterization of power quality transients and their classification using competitive layers, Measurement (Ed. Elsevier) 42 (Issue 3) (2009) 478–484.

[6] J. J. G. De la Rosa, A. Moreno, C. G. Puntonet, A practical review on higher-order statistics interpretation. application to electrical transients characterization, Dynamics of continous discrete and Impulsive Systems-Series B: Applications and Algorithms 14 (4) (2007) 1577–1582.

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

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

[9] J. J. G. De la Rosa, I. Lloret, C. G. Puntonet, R. Piotrkowski, A. Moreno, Higher-order spectra measurement techniques of termite emissions. a characterization framework, Measurement (Ed. Elsevier) 41 (1) (2008) 105–118, available online 13 October 2006.

[10] J. J. G. De la Rosa, R. Piotrkowski, J. Ruzzante, 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) (2007) 1917–1926, available online 10 October 2006.

[11] J. J. G. De la Rosa, A. M. Muñoz, Higher-order characterization of power quality transients and their classification using competitive layers, Przegąd Elektrotechniczny-Electrical Review 10 (Issue 85) (2009) 284–289.

[12] M. H. J. Bollen, E. Styvaktakis, I. Y.-H. Gu, Categorization and analysis of power system transients, IEEE Transactions on Power Delivery 20 (3) (2005) 105–118.

[13] D. Paul, Low-voltage power system surge overvoltage protection, IEEE Transactions on Industry Applications 37 (1) (2001) 223–229.

[14] F. Martzloff, Protecting computer systems against power transients, IEEE Spectrum 27 (Issue 4) (1990) 37–40.

[15] IEEE Recommended practice for monitoring electric power quality, Tech. Rep. IEEE Std. 1159-1995, The Institute of Electrical and Electronics Engineers, Inc. (1995).

[16] IEEE Guide for service to equipment sensitive to momentary voltage disturbances, Tech. Rep. IEEE Std. 1250-1995, The Institute of Electrical and Electronics Engineers, Inc. (April 1995).

[17] A. Moreno, J. Flores, D. Oterino, J. J. G. De la Rosa, Power line conditioner based on ca pwm chopper, in: ISIE 2007, Proceedings of the 2007 IEEE International Symposium on Industrial Electronics, June 2007, 2007, pp. 2454–2456.

[18] C. L. Nikias, A. P. Petropulu, Higher-Order Spectra Analysis. A Non-Linear Signal Processing Framework, Englewood Cliffs, NJ, Prentice-Hall, 1993.

[19] C. L. Nikias, J. M. Mendel, Signal processing with higher-order spectra, IEEE Signal Processing Magazine (1993) 10–37.

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

[21] A. K. Nandi, Blind Estimation using Higher-Order Statistics, 1st Edition, Vol. 1, Kluwer Academic Publishers, Boston, 1999.

Authors: Prof. Juan José González de la Rosa[1],
Prof. Agustín Agüera Pérez[1],
Prof. José Carlos Palomares Salas[1],
Prof. Antonio Moreno Muñoz[2]

[1]Univ. of Cádiz. EPSA. Av. Ramón Puyol S/N. E-11202-Algeciras-Cádiz-Spain
[2]Univ. of Córdoba. Campus de Rabanales. Ed. Leonardo da Vinci. E-14071-Córdoba-Spain
email: juanjose.delarosa@uca.es; amoreno@uco.es