

# The Use of a Multilabel Classification Framework for the Detection of Broken Bars and Mixed Eccentricity Faults Based on the Start-Up Transient

George Georgoulas, Vicente Climente-Alarcon, *Member, IEEE*,  
 Jose A. Antonino-Daviu, *Senior Member, IEEE*, Ioannis P. Tsoumas, *Member, IEEE*,  
 Chrysostomos D. Stylios, *Senior Member, IEEE*, Antero Arkkio,  
 and George Nikolakopoulos, *Member, IEEE*

**Abstract**—In this paper, a data-driven approach for the classification of simultaneously occurring faults in an induction motor is presented. The problem is treated as a multilabel classification problem, with each label corresponding to one specific fault. The faulty conditions examined include the existence of a broken bar fault and the presence of mixed eccentricity with various degrees of static and dynamic eccentricity, while three “problem transformation” methods are tested and compared. For the feature extraction stage, the start-up current is exploited using two well-known time–frequency (scale) transformations. This is the first time that a multilabel framework is used for the diagnosis of co-occurring fault conditions using information coming from the start-up current of induction motors. The efficiency of the proposed approach is validated using simulation data with promising results irrespective of the selected time–frequency transformation.

Manuscript received July 1, 2016; revised October 24, 2016; accepted November 28, 2016. Date of publication December 8, 2016; date of current version April 18, 2017. This work was supported in part by the Spanish MINECO and FEDER program in the framework of the “Proyectos I + D del Subprograma de Generación de Conocimiento, Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia” under Grant DPI2014-52842-P and in part by the Horizon 2020 Framework program DISIRE under the Grant Agreement 636834. Paper no. TII-16-0649. (*Corresponding author: G. Georgoulas.*)

G. Georgoulas and G. Nikolakopoulos are with the Control Engineering Group Department of Computer Science, Electrical, and Space Engineering, Luleå University of Technology, SE-97187 Luleå, Sweden (e-mail: geogeo@ltu.se; geonik@ltu.se).

V. Climente-Alarcon and A. Arkkio are with the Department of Electrical Engineering and Automation, Aalto University, FI-00076 Aalto, Finland (e-mail: viclial@iee.org; antero.arkkio@aalto.fi).

J. A. Antonino-Daviu is with the Instituto Tecnológico de la Energía, Universitat Politècnica de València, 46022 Valencia, Spain (e-mail: joanda@die.upv.es).

I. P. Tsoumas is with ABB Corporate Research, Baden-Dättwil 5405, Switzerland (e-mail: ytsoumas@iee.org).

C. D. Stylios is with the Laboratory of Knowledge and Intelligent Computing, Department of Computer Engineering, TEI of Epirus, 47100 Arta, Greece (e-mail: stylios@teiep.gr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TII.2016.2637169

**Index Terms**—Fault diagnosis, induction motors, multilabel classification.

## I. INTRODUCTION

RECENTLY, the field of electrical machines has received increasing attention, toward the development and application of fault detection techniques [1], [2]. The driving forces behind this have been the following: 1) the reduction of maintenance cost and the prevention of significant costly unscheduled downtimes; 2) the development of novel fault detection techniques that can revolutionize the existing commonly used approaches in the field; and 3) the reduction of the cost for applying embedded fault detection solutions in lab experimental cases or bigger industrial setups.

Among the most common faults encountered in the area of induction motors are the opening or shorting of one or more of the stator’s phase windings [3] and the presence of broken rotor bar(s) or cracked rotor’s end-rings [4], air-gap irregularities [5], bearing faults, and eccentricity faults [6].

For the detection of these faults and more specifically, for the case of broken bars and eccentricity, various input signals have been used quite successfully, such as induced voltages [7], vibration signals [8], currents and vibration signals [9], and instantaneous angular speed or power [10]. However, methods that rely only on the use of currents, like the motor current signature analysis (MCSA) [11], [12] are usually preferred mainly due to their noninvasive nature. The underlying philosophy of those methods is to detect the presence of specific components created by the fault.

For the case of broken bars the frequencies of the fault components are given by

$$f_b = (1 \pm 2 \cdot k \cdot s) f_s, \quad k = 1, 2, \dots \quad (1)$$

where  $f_s$  is the fundamental frequency and  $s$  is the slip.

Among the various components, the lower (or left) sideband harmonic (LSH), for  $k = 1$ , and taking the minus sign in (1)

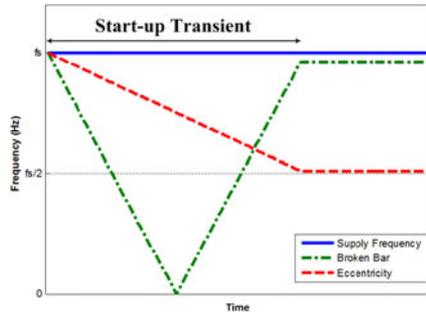


Fig. 1. Approximate evolution of the instantaneous frequency of faulty components during start-up.

[13], is the most extensively used one in the fault detection literature.

In the case of eccentricity, and more specifically, in the case of mixed eccentricity, which is the most commonly encountered type, another group of components, with frequencies given by (2), is expected [11]

$$f_{ecc} = \left(1 \pm m \left(\frac{1-s}{p}\right)\right) f_s, \quad k = 1, 2, \dots \quad (2)$$

where  $p$  is the number of pole pairs.

As it can be seen, both frequency groups depend on the slip value. This observation triggered the development of methods that rely on the analysis of these components during start-up, a transient at which the slip varies from one to a value usually close to zero within a short period of time. This rapid evolution makes the faulty components “draw” quite distinctive patterns in the time–frequency plane. In the case of broken bars the LSH will show up as a V pattern on the time–frequency plane, while in the case of mixed eccentricity and  $m = p/2$ , one of the relevant frequency components will start at being equal to the supply frequency and will end up having a frequency just above half the supply frequency (assuming a value of slip close to 0). These characteristic patterns are depicted in Fig. 1.

Various methods under the term transient MCSA have been proposed for the detection and quantification of the aforementioned components [14]–[17]. Among the different proposed approaches for the diagnosis of these faults or combination of them, the investigation of single-fault diagnosis methods has dominated the literature. However, the simultaneous presence of electromechanical faults is the rule and not the exception in industrial settings. Even though the faults rarely occur/start simultaneously, during the course of operation of an induction motor, one type of faults can trigger the occurrence of another type or a second fault may occur, while the machine operates under the presence of another fault.

Therefore, lately, scenarios where more than one fault can be present at the same time were investigated [18]–[20]. In [18], a filter bank combined with high-resolution spectral analysis was used for the estimation of faulty components caused by multiple faults (broken bars, eccentricity, and bearing faults). In [19], a cascade of neural networks was trained to recognize different fault scenarios, including simultaneous faults. In [20], different faults were investigated through the characteristic “fingerprint” that is left on the decomposed start-up current via

discrete wavelet transform (DWT). However, this approach has not been automated yet.

In this paper, the presence of a broken bar, as well as mixed-eccentricity faults, is examined within the framework of multilabel classification [21]–[24]. Multilabel classification is a framework that allows instances to belong to more than one category/class at the same time, a situation that arises in many real-life applications, including the case of equipment suffering more than one fault. This is, to the best of our knowledge, the first time that such a framework is explicitly employed for the diagnosis of co-occurring faults in induction motors during the start-up operation. Therefore, the main contribution of the proposed work is the introduction of the multilabel framework as a means to cope with the occurrence of more than one fault.

As in almost all classification problems, the multilabel framework uses extracted features instead of the original raw measurements. For the specific application, the time–frequency representation of the start-up current is used as a means for feature extraction. The advantage of examining the start-up current is that especially the pattern of the frequency components related to the rotor fault is very different from those due to other failures or phenomena. Hence, its detection constitutes a very reliable indicator of the presence of the fault. A fault detection algorithm relying on the start-up cannot be misled by the presence of other failures or phenomena irrelevant to the fault, like in [25].

Furthermore, the approach proposed in this paper has, on the one hand, the advantage that allows the aforementioned faulty components to “draw” a distinctive pattern on the time–frequency plane, while, on the other hand, poses the challenge of adequately handling the high dimensionality of the representation in the time–frequency plane. In order to show the robustness and the general applicability of the method, two well-known time–frequency (scale) transforms, the short-time Fourier transform (STFT) and the continuous wavelet transform (CWT) are tested.

To cope with the high dimensionality of the output of both the STFT and the CWT, a number of steps are involved, including the application of a two-dimensional (2-D) variant of the piecewise aggregate approximation (PAA) and then the application of principal component analysis (PCA) for further dimensionality reduction. The final diagnosis is accomplished within the multilabel classification framework using simple conventional classifiers. The whole procedure is tested using simulations coming from a suitable model of the induction machine, which can accommodate various types of faults.

The rest of the paper is structured as follows. Section II presents the proposed approach and the methods used in each one of the involved steps, while Section III summarizes the evaluation procedure, along with a description of the involved simulation model, followed by the achieved results. Finally, Section IV concludes the paper.

## II. PROPOSED APPROACH

The proposed approach, which relies on the multilabel framework, requires the transformation of the original raw current measurement to a more suitable representation. The extraction of this more suitable representation builds upon a method for

broken bar fault diagnosis using the time–frequency representation of the start-up [26]. In [26], the time–frequency representation was produced using the STFT. However, the multilabel classification offers a general framework, which is not so much affected by the selected time–frequency representation. To prove that, an alternative time–scale representation based on the CWT is also tested.

Therefore, the approach proposed in [26], is extended to meet the specific characteristics of the faults to be tackled with the multilabel case as it is explained in the following subsections. For completeness, each stage is presented in brief in the following subsections, while the multilabel stage, which is the core of the proposed approach, is presented in more detail since it is novel in the field of induction motor fault diagnosis.

### A. Transient Isolation

The proposed overall method tries to take advantage of the time evolution of the components generated by the fault. Therefore, the isolation of the end of the transient is important. However, contrary to the case of broken bars, where at the end of the transient the faulty components get close and are usually masked by the fundamental/supply frequency, in the case of mixed eccentricity the faulty component at steady state (assuming a small value of the slip) lies quite apart from the fundamental/supply frequency. Therefore, in this paper, unlike the case in [26], the window of the current retained for further processing was selected to be approximately 1.5 times the duration of the transient, while the end of the transient is detected using a moving window as in [26].

### B. Time–Frequency Representation

There are many tools that can be used for the analysis of the time–frequency evolution of signals. Among them, STFT and CWT are probably the most popular in engineering applications.

1) *STFT*: STFT is probably the simplest and most widely used method when it comes to transforming a signal to its time–frequency representation. Therefore, even though it has a number of limitations, inherited by the application of Fourier transform to a finite length signal  $x(\tau)$ , it has already proven its usability in the case of single, broken bar faults [26].

Formally, the STFT is given by (3), where  $w(t)$  represents the window function

$$X(t, \omega) = \int_{-\infty}^{\infty} x(\tau) w(\tau - t) e^{-j\omega\tau} d\tau. \quad (3)$$

In this paper, the time–frequency content below the fundamental frequency is considered. For that range, the components created by the broken bar and the eccentricity would have instantaneous frequencies that would approximately look like the ones in Fig. 1.

As it can be seen in Fig. 2, which depicts the spectrogram of different fault scenarios, generated by the simulation model described in Section III (normalized and measured in decibels), even though the ideal scenario is not met, both the V pattern and the component created by the eccentricity can be spotted. When, however, they co-occur (bottom of Fig. 2), these patterns are a bit distorted.

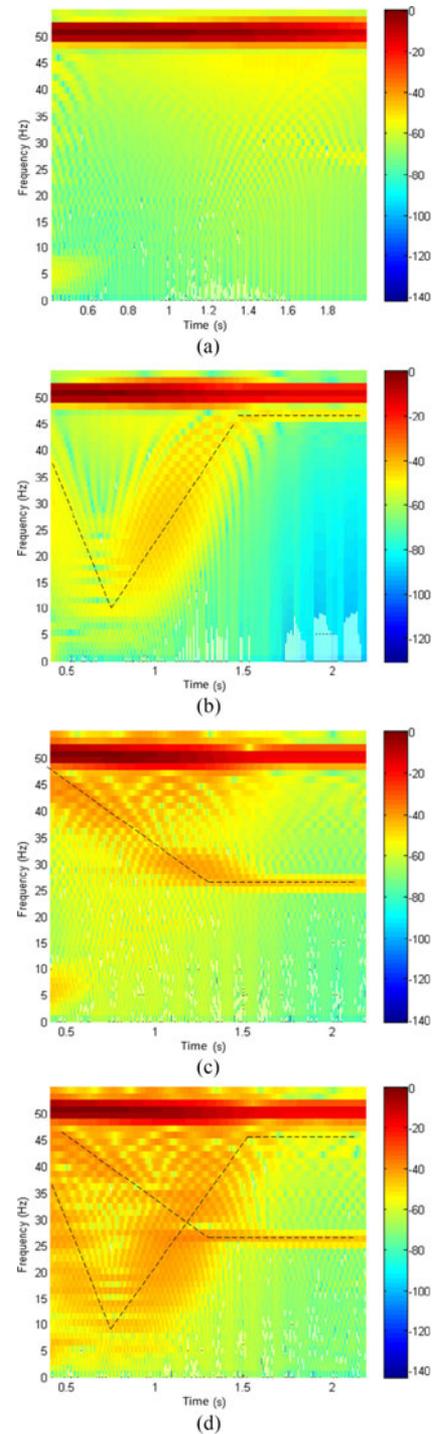


Fig. 2. Time–frequency plane for different conditions, simulated using the model of Section III. From top to bottom, a healthy motor, a motor with one broken bar, a motor with mixed eccentricity (35% static and 35% dynamic), and finally, a motor with both mixed eccentricity (35% static and 35% dynamic) and one broken bar.

2) *CWT*: A wavelet  $\psi(t)$  is a localized waveform that satisfies certain conditions [27]. Due to its localized nature, it is very popular for applications involving transient phenomena or phenomena having abrupt changes. The CWT of a signal is given by

$$T(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \cdot \psi^* \left( \frac{t-b}{a} \right) dt \quad (4)$$

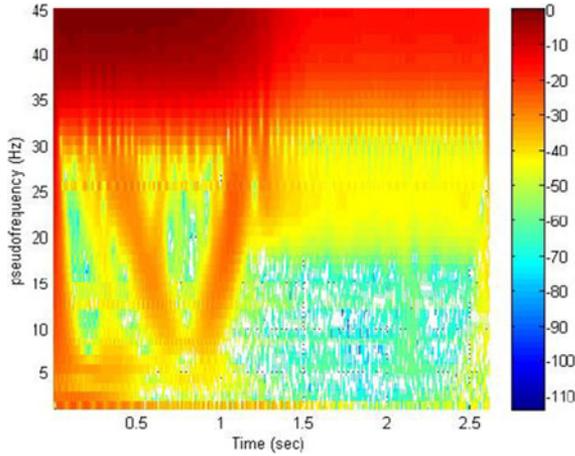


Fig. 3. Time-(pseudo)frequency plane for a motor with both mixed eccentricity (35% static and 35% dynamic) and one broken bar created using the Morlet complex wavelet.

where  $a$  is the scale parameter and  $b$  is the translation parameter and the asterisk indicates the complex conjugate. The relative contribution of the signal energy is provided by the scalogram  $E(a, b)$  defined in (5), which creates an alternative representation to the one created by the STFT

$$E(a, b) = |T(a, b)|^2. \quad (5)$$

While wavelets can be both real and complex, for time–frequency analysis, complex Morlet wavelets (6) are usually preferred [28]

$$\psi(t) = \frac{1}{\sqrt{\pi f_b}} \exp\left(j2\pi f_c t - \frac{t^2}{f_b}\right) \quad (6)$$

with  $f_c$  being the central frequency of the wavelet and  $f_b$  the bandwidth. Furthermore,  $f_c$  can be used to transform scale to “pseudo-frequency” given by (7), which is a notion more familiar to engineers

$$f = \frac{f_c}{a} f_{\text{samp}} \quad (7)$$

where  $f_{\text{samp}}$  is the sampling frequency in hertz. Fig. 3 depicts the result of the application of the CWT to a simulated motor with one broken bar and mixed eccentricity (35% static and 35% dynamic). This is the same case as the one depicted on the bottom of Fig. 2. The V pattern caused by the broken bar is clearly visible, where the eccentricity-related pattern is more smeared. However, the steady-state component is clearly visible. In the following section, it will be shown that both representations can be used for further processing of the start-up current.

### C. Dimensionality Reduction

Time–frequency transforms create a time–frequency representation, practically an image, which has quite high dimensionality. In order to alleviate the “curse of dimensionality problem,” two dimensionality reduction techniques are employed.

1) *2-D Piecewise Aggregate Approximation*: PAA was originally developed for one-dimensional signals [29], [30] but

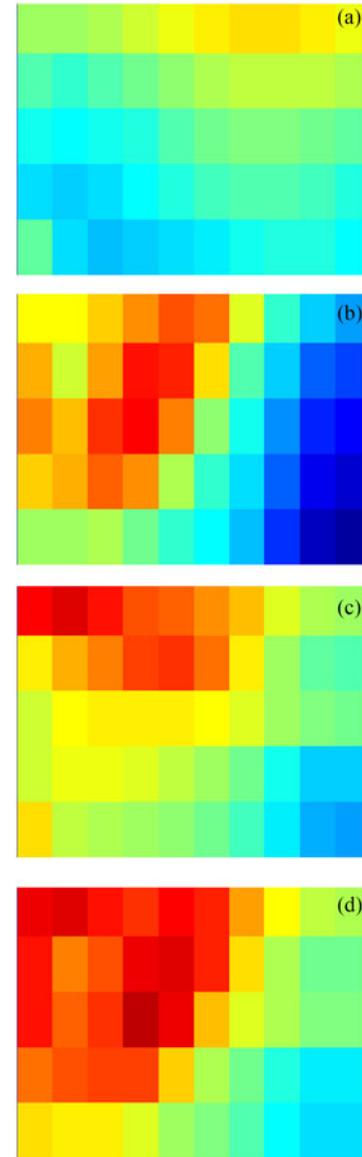


Fig. 4. 2-D PAA corresponding to the cases depicted in Fig. 2. From top to bottom, a healthy motor, a motor with one broken bar, a motor with eccentricity (35% static and 35% dynamic), and finally, a motor with both mixed eccentricity (35% static and 35% dynamic) and one broken bar.

since then a 2-D variant has also been used for reducing the dimensionality of images [26], [31].

In the 2-D case, the original image  $I$  with dimensions  $N \times M$  is turned into an image  $\bar{I}$  of lower dimensions  $p_1 \times p_2$  where

$$\bar{I}(i, j) = \frac{p_1 p_2}{NM} \sum_{x=\frac{M}{p_1}(i-1)+1}^{\frac{M}{p_1}i} \sum_{y=\frac{N}{p_2}(j-1)+1}^{\frac{N}{p_2}j} I(x, y). \quad (8)$$

Practically, (8) indicates that we replace a block of the image with its mean values. Fig. 4 depicts the corresponding PAA representations of the spectrograms of Fig. 2 (using only the frequency components below 45 Hz, since the components of interest lie in the lower part of the spectrogram). Similar representations are created also for the case of CWT.

**2) Principal Component Analysis:** Even after the application of PAA, the resulting image has a dimension (for this specific implementation) of  $10 \times 5$ , which means that a feature vector of 50 elements will be created. Therefore, a second stage is employed to further reduce the dimensionality. In the field of dimensionality reduction, many algorithms have been proposed over the past decade. However, it turns out that PCA [32] can be quite competitive compared to other more advanced methods, when it comes to real-life problems [33].

PCA transforms the original data by projecting them to just a few of the eigenvectors of their covariance matrix that correspond to the largest eigenvalues. This way, dimensionality reduction is achieved. Furthermore, the new features are uncorrelated, which can sometimes lead to an increased classification performance [34].

#### D. Multilabel Classification

Standard classification algorithms assume that a given instance (also referred to as example or data point) is associated with a single label  $\lambda_i$  from a set of disjoint labels  $\Lambda$  [21]. In other words, each instance belongs to just one class  $\lambda_i$ . However, in some domains, such as text categorization, instances can actually be associated with a set of labels  $Y \subseteq \Lambda$ , or using a pattern classification terminology, instances can belong to more than one classes. This kind of problems falls under the umbrella of multilabel classification and can be addressed basically by two approaches: 1) problem transformation and 2) algorithm adaptation.

The first one involves the transformation of the multilabel classification problem into one or more single-label classification problems, which can be tackled with any conventional classification algorithm. In other words, the original data set is transformed into (usually) more than one data sets, with the number of data sets depending on the transformation method, where each data point is labeled with just one label. The labeling also depends on the transformation method as it is explained in the following sections.

The second one, as its name implies, relies on the adaptation of classification algorithms to directly handle multiple labels per instance. In other words, the classifiers are trained using more than one label per data point and are meant to produce a set of labels for each one of the new unseen data.

In this paper, only methods from the problem transformation approaches are investigated: the label powerset (LP) approach, the binary relevance (BR), and the calibrated label ranking calibrated label ranking (CLR) [21], [23].

**1) Label Powerset:** The LP transformation simply treats each unique label set encountered in the training set as one of the classes in a multiclass classification scheme. Therefore, after the transformation, any conventional multiclass algorithm can be used. For example, let's consider the following hypothetical data sample shown in Table I. The original dataset has three labels  $\{\lambda_1, \lambda_2, \lambda_3\}$  (e.g., in a fault detection framework for induction motors, the labels could be {broken bar, eccentricity, bearing fault}). The transformed data shown in Table II will have four classes. Note that in the second column, we no longer

**TABLE I**  
ORIGINAL DATASET

Examples	Label Set
1	$\{\lambda_1, \lambda_2\}$
2	$\{\lambda_1\}$
3	$\{\lambda_1, \lambda_3\}$
4	$\{\lambda_3\}$
5	$\{\lambda_2, \lambda_3\}$

**TABLE II**  
LABEL POWERSET

Examples	Label
1	$\lambda_{1-2}$
2	$\lambda_1$
3	$\lambda_{1-3}$
4	$\lambda_3$
5	$\lambda_{2-3}$

**TABLE III**  
BINARY RELEVANCE SETS

Data Set ( $\lambda_1$ )		Data Set ( $\lambda_2$ )		Data Set ( $\lambda_3$ )	
Example	Label	Example	Label	Example	Label
1	$\lambda_1$	1	$\lambda_2$	1	$\neg\lambda_3$
2	$\lambda_1$	2	$\neg\lambda_2$	2	$\neg\lambda_3$
3	$\lambda_1$	3	$\neg\lambda_2$	3	$\lambda_3$
4	$\neg\lambda_1$	4	$\neg\lambda_2$	4	$\lambda_3$
5	$\neg\lambda_1$	5	$\lambda_2$	5	$\lambda_3$

$\neg\lambda_i$  corresponds to NOT  $\lambda_i$ .

have sets but single labels, i.e.,  $\lambda_{1-2}$  is a single label, which means that this label/class corresponds to labels  $\lambda_1$  and  $\lambda_2$  (e.g., {broken bar and eccentricity}) of the original data set.

LP has the advantage of taking into consideration potential correlations between the labels. On the other hand, it has the disadvantage that it can become computationally intractable for large label sets, with many of the newly created classes having very few training data. For this specific application, due to the small number of classes, computational complexity is not an issue.

**2) Binary Relevance:** BR creates  $q$  new data sets, where  $q$  is the number of unique labels, in the original data set, as shown in Table III. Each new data set is used to train a (binary) classifier that learns to discriminate the corresponding label/class against the rest of the classes (one vs all approach). In other words, an instance either has a specific fault or not. Furthermore, BR does not suffer from the complexity problem of the LP method, but on the other hand, it cannot take into consideration potential correlations between the labels.

**3) Calibrated Label Ranking:** CLR [35] is an extension of the Ranking by Pairwise Comparison (RPC) method. While BR is a “one-versus-all” approach, RPC is a “one-versus-one” approach having binary classifiers to discriminate between two labels at a time. RPC works by first creating  $q(q - 1)/2$  datasets,

**TABLE IV**  
RANKING BY PAIRWISE COMPARISON SETS

Data Set $(\lambda_1, \lambda_2)$		Data Set $(\lambda_1, \lambda_3)$		Data Set $(\lambda_2, \lambda_3)$	
Example	Label	Example	Label	Example	Label
1	x <sup>a</sup>	1	$\lambda_1, \neg\lambda_3$	1	$\lambda_2, \neg\lambda_3$
2	$\lambda_1, \neg\lambda_2$	2	$\lambda_1, \neg\lambda_3$	2	x
3	$\lambda_1, \neg\lambda_2$	3	x	3	$\neg\lambda_2, \lambda_3$
4	x	4	$\neg\lambda_1, \lambda_3$	4	$\neg\lambda_2, \lambda_3$
5	$\neg\lambda_1, \lambda_2$	5	$\neg\lambda_1, \lambda_3$	5	x

<sup>a</sup>x corresponds to an instance not included in the “pairwise” dataset.

one for each pair of labels  $(\lambda_i, \lambda_j)$ ,  $i, j \in [1, q]$ ,  $i \neq j$ . Each one of these data sets contains instances from the original dataset that were annotated by one of the two labels of the pair under investigation but not by both of them. The datasets corresponding to RPC for our original example are shown in Table IV.

As it can be seen, not all instances from the original dataset are part of the newly created dataset. When a new instance example is to be classified, all trained classifiers are invoked and the votes obtained by them are used to create a ranking for the labels. In other words, and as its name implies, RPC creates a ranking of the labels. Thus, PRC does not automatically provide a set of labels, and it needs a function to decide how many of the top-ranked labels should be selected to annotate the unknown example. CLR solves this problem of bipartitioning a ranking list by introducing an additional virtual label  $\lambda_0$ . As in the case of BR, a new dataset is created for each label  $\lambda_i$ , corresponding to the pair  $(\lambda_i, \lambda_0)$ . For each instance that contains the label  $\lambda_i$ , the virtual label is considered false and vice versa. The number of votes received for the virtual label acts as a breaking point of the ranking into relevant (to be assigned to the unknown instance) and irrelevant labels.

In general, it should be highlighted that CLR has problems when dealing with datasets having many labels, since the number of pairwise comparisons can become very large. However, for this specific application, as for the case of LP, this is not an issue.

### E. Basic Classifiers

As it mentioned earlier, the problem transformation methods have the advantage that common classifiers can be used once the transformed data sets are created. As it was shown, BP and CLR result in a number of binary classification problems, while LP creates just one data set with many labels (multiclass problem). In this paper, for both the multiclass classification problem, as well as for the binary classification problems, two simple classifiers, a linear one and a nonlinear one, are tested. Both of these two classifiers can be used either as binary or as multiclass classifiers.

1) *Minimum Mahalanobis Distance Classifier*: One of the simplest linear classifiers is the minimum Mahalanobis distance (MMD) classifier [37]. The MMD classifier has the advantage that it does not require the tuning of any parameters.

The classifier assigns a new instance  $\mathbf{x}$  to class  $i$ , which maximizes the following quantity:

$$i = \arg \max_{l=1 \dots L} \left\{ 2 \ln (P(\omega_l)) - (\mathbf{x} - \mu_l)^T \mathbf{C}^{-1} (\mathbf{x} - \mu_l) \right\} \quad (9)$$

where  $P(\omega_l)$  is the *a priori* probability of occurrence of class  $\omega_l$ ,  $\mu_l$  is the average vector of class  $\omega_l$ , and  $\mathbf{C}$  is the covariance matrix of the training data, which is considered the same for all classes. Despite its simplicity, the MMD classifier is quite competitive when dealing with real-life problems [38].

2) *Nearest Neighbor Classifier*: The nearest neighbor (NN) classifier belongs to the family of “lazy” classifiers [39], while the term “lazy” comes from the fact that no training takes place. The NN, in its basic implementation, stores all training data, and once a new instance is to be classified, it assigns to it the label of its closest training instance. NN can be quite effective, especially in low-dimensional spaces because it can create nonlinear boundaries between the different classes.

## III. EVALUATION PROCEDURE OF THE PROPOSED MULTILABEL FRAMEWORK

In order to test the effectiveness of the multilabel classification procedure to deal with the diagnosis of induction faults during start-up, a number of stator current waveforms were generated by the simulation model described in the next section. These waveforms were processed using both STFT and CWT and then, after the feature extraction stage, were utilized for the three different multilabel approaches presented in Section II.

### A. Simulation Model

The analytical model used to obtain the simulated waveforms follows the magnetic coupled circuits approach and can be summarized for a machine having  $m$  stator and  $n$  rotor phases in  $m+n$  equations in the form of

$$[U_s] = [R_s][I_s] + \frac{d[\Psi_s]}{dt} \quad (10)$$

$$[0] = [R_r][I_r] + \frac{d[\Psi_r]}{dt} \quad (11)$$

where the first equation is applied to each of the stator circuits whereas the second is valid for the rotor ones, with  $[U_s]$  the vector of stator voltages,  $[I_s]$  the vector of stator currents,  $[I_r]$  the vector of rotor currents,  $[\Psi_s]$  the vector of stator fluxes,  $[\Psi_r]$  the vector of rotor fluxes,  $[R_s]$  the stator resistance matrix, and  $[R_r]$  the rotor resistance matrix. The flux linkages are

$$[\Psi_s] = [L_{ss}][I_s] + [L_{sr}][I_r] \quad (12)$$

$$[\Psi_r] = [L_{sr}]^T [I_s] + [L_{rr}][I_r] \quad (13)$$

where  $L_{ss}$  is the stator inductance matrix,  $L_{sr}$  is the stator–rotor inductance matrix, and  $L_{rr}$  is the rotor inductance matrix. The computation of self and mutual inductances for the previous equations has to be accurately done in order to properly show the effects of the fault. This has been carried out by taking into account the position of all the conductors along the airgap, using the yoke flux instead of the airgap flux and utilizing the circular

convolution to compute this magnitude from the conductors' distribution [40].

Broken bars are simulated by increasing the resistance of the respective element up to a level where the respective current is negligible. Eccentricity is reproduced by modifying the airgap's permeance value in a single step through altering the length of the airgap according to two sinusoidal distributions relative to the rotor, one independent of its instantaneous angle with respect to the stator  $\theta_r$  for simulating dynamic eccentricity, and the other that unwinds its movement for static eccentricity. The computation of the inductances is performed for all the rotor positions  $\theta_r$ . The computed values are stored in a lookup table, which is used in the simulation.

The machine model is completed with the equation for the calculation of the electromagnetic torque and Newton's second law for rotation. Since under mixed eccentricity all the inductance matrices vary with the rotor position, this has to be reflected in the coenergy derivative when computing the torque [41] compared to [40]

$$T_m = \frac{1}{2} \left[ [I_s^t] \frac{\partial [L_{ss}]}{\partial \theta_r} [I_s] + [I_s^t] \frac{\partial [L_{sr}]}{\partial \theta_r} [I_r] + [I_r^t] \frac{\partial [L_{sr}^T]}{\partial \theta_r} [I_s] + [I_r^t] \frac{\partial [L_{rr}]}{\partial \theta_r} [I_r] \right] \quad (14)$$

$$T_m - T_L = J \frac{d\Omega}{dt} = J \frac{d^2 \theta_r}{dt^2} \quad (15)$$

where  $T_m$  is the motor's electromagnetic torque,  $T_L$  is the load torque,  $J$  is the moment of inertia, and  $\Omega$  is the angular speed of the motor.

As formulated, this model reproduces accurately the frequencies of the flux and current components associated to the faults. However, since no saturation is considered, their amplitudes are magnified. In order to avoid this, the airgap permeance variation is scaled to approximate the fault harmonic's amplitudes obtained from experimental testing of the same machine. Although more rigorous approaches can be included in the model [41], [42], this suffices for the purpose of this paper.

## B. Experimental Procedure—Results

In order to test the performance of the multilabel classification approach in fault diagnosis of induction motors, a number of experiments were conducted. Using the simulation model described in the previous section, a motor with two pairs of poles is simulated for the following scenarios: 1) healthy motor; 2) motor with one broken bar; 3) motor with mixed eccentricity; and d) motor with mixed eccentricity and one broken bar. For each condition, ten different start-ups are generated with different durations. For the case of mixed eccentricity, various degrees of static and dynamic eccentricities are simulated with the minimum value being equal to 5% and the maximum equal to 35%.

For the case of the BR transformation, a hierarchical classification approach is imposed, taking into consideration that the healthy category cannot appear simultaneously with any other class. In other words, an example cannot be both healthy and

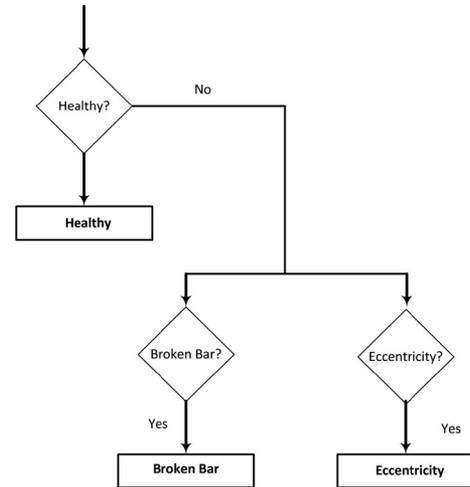


Fig. 5. Schematic of the hierarchical approach in the case of BR method.

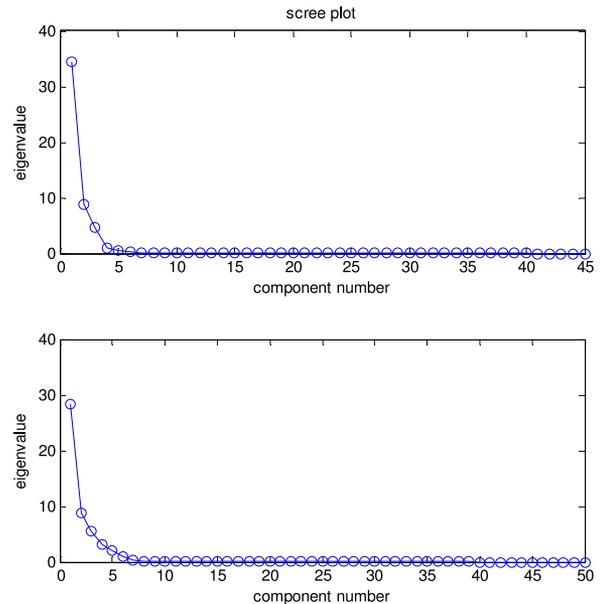


Fig. 6. Scree plot of the principal components indicating that most of the data's variance is summarized by the first few principal components.

belong to any of the fault classes at the same time. Therefore, first a binary classifier decides whether the new sample is healthy or not, and if it is not healthy, the standard BR approach is invoked to decide to which class(es) the new sample belongs. This procedure is illustrated in Fig. 5.

The processing of the stator currents was performed by STFT and CWT, to verify the robustness of the multilabel classification framework. In both sets of experiments (STFT and CWT), during the final stage of the dimensionality reduction process, four different values for the number of retained components are tested (two, three, four, and five). As it can be seen from the scree plots (the plots of the eigenvalues versus the index of the eigenvector) [43] in Fig. 6 for both the STFT and the CWT, the first few principal components seem to summarize all the

**TABLE V**  
CLASSIFICATION RESULTS—STFT

PCs <sup>a</sup>	MMD Classifier			1 – NN Classifier		
	LP	BR	CLR	LP	BR	CLR
2	90%	92.5%	92.5%	95%	92.5%	95%
3	100%	97.5%	95%	97.5%	97.5%	97.5
4	100%	97.5%	97.5%	97.5%	97.5%	100%
5	100%	97.5%	97.5%	97.5%	97.5%	100%

<sup>a</sup>PCs—principal components.

**TABLE VI**  
CLASSIFICATION RESULTS—CWT

PCs	MMD Classifier			1 – NN Classifier		
	LP	BR	CLR	LP	BR	CLR
2	100%	100%	100%	100%	100%	98.8%
3	100%	100%	100%	100%	100%	98.8%
4	100%	100%	97.5%	100%	100%	98.8%
5	100%	100%	100%	100%	100%	100%

relevant information. The leave-one-out (loo) testing procedure is used for every experiment and all the results of this study in terms of absolute accuracy are summarized in **Tables V** and **VI** for the case of STFT and CWT, respectively. At this point, it should be noted that since the findings presented in [20] are based on eye inspection, no direct comparison of the proposed method with [20] is possible.

From **Tables V** and **VI**, it can be observed that the CWT representation is better fitted for the analysis of the start-up current for the diagnosis of combined faults when compared to the use of STFT. However, STFT also leads to very high performance if more than three principal components are retained. In terms of the retained principal components, three or more is a reasonable choice. The MMD classifier also has a slightly better performance compared to the 1 – NN classifier. As it was pointed out in [38], MMD can be very competitive when it comes to real-life problems. Regarding the effectiveness of the three different multilabel schemes (LP, BR, and CLR), LP is slightly better with the other two schemes having similar performance. It should also be noted that in case of fault combinations that are not present in the database, e.g., when data from single bearing faults are available, the LP method will not be able to recognize them. This is because the LP method has to have training data from all the potential combinations. Therefore, on the one hand, LP has a slightly superior performance compared to the other two investigates methods, but on the other hand, the price to be paid is the need for a data set with a complete representation of all the potential combination of labels (faults).

#### IV. CONCLUSION

A multilabel classification framework for the diagnosis of multiple faults in induction motors was presented. This is the first time that such an approach is employed for the case of induction motors during the start-up; therefore, direct compar-

ison with similar schemes is not possible. As it was presented, the proposed scheme was applied in the cases of broken bars and mixed-eccentricity with promising results. The three transformation schemes (LP, BR, and CLR) yielded comparable results, with the LP method outperforming the other two in most experiments, irrespective of the selected time–frequency representation. However, as it was stated in the previous section, LP cannot cope with the combination of faults that it has not been explicitly trained to recognize. Therefore, in real-life applications, any of the other two methods could also be considered, at least until a representative database, including all faults, is built. Regarding the number of retained principal components, three to five seems to be a good choice for this specific setting.

The proposed approach is a data-driven approach. Therefore, further investigation is needed involving more data as well as data coming from experimental settings and, more importantly, data coming from industrial settings before the method can be applied to real-life situations. Further investigation is also required regarding the selection of “optimal” parameters of the feature extraction stages, even though the method seems to work with high levels of accuracy for a range of parameters and for different time–frequency representations. On the other hand, the method is generic enough so as to be applied to different scenarios. More importantly, the multilabel framework can be applied to different application areas where multiple states can occur concurrently, without any major modifications. However, the selection of the most appropriate multilabel classification scheme is application dependent.

In future work, the proposed approach will be tested with the inclusion of other fault conditions, such as bearing faults and winding shorts. Furthermore, different severity levels within a specific fault class will be investigated, e.g., broken bar faults with more one broken bars. Moreover, the method will be tested using experimental data. Finally, regarding the specific application, which relies on the frequency content at different parts of the time–frequency plane, a multi-instance classification approach, which does not use a single feature vector to describe each condition, will be tested [44].

#### REFERENCES

- [1] A. S. Nandi and H. Toliyat, “Condition monitoring and fault diagnosis of electrical machines—a review,” *IEEE Trans. Energy Convers.*, vol. 20, no. 4, pp. 719–729, Dec. 2005.
- [2] G. Georgoulas *et al.*, “Principal component analysis of the start-up transient and hidden Markov modeling for broken rotor bar fault diagnosis in asynchronous machines,” *Expert Syst. Appl.*, vol. 40, no. 17, pp. 7024–7033, 2013.
- [3] O. Mustafa, G. Nikolakopoulos, and T. Gustafsson, “Stator winding short circuit fault detection based on set membership identification for three phase induction motors,” in *Proc. 20th Mediterranean Conf. Control Autom.*, 2012, pp. 290–296.
- [4] P. M. Santos, M.B.R. Correa, C.B. Jacobina, E.R.C. da Silva, A.M.N. Lima, G. Didiery, H. Raziky, and T. Lubiny, “A simplified induction machine model to study rotor broken bar effects and fault detection,” in *Proc. 37th IEEE Power Electron. Specialists Conf.*, pp. 1–7, 2006.
- [5] G. G. Acosta, C. J. Verucchi, and E. R. Gelso, “A current monitoring system for diagnosing electrical failures in induction motors,” *Mech. Syst. Signal Process.*, vol. 20, no. 4, pp. 953–965, Oct. 2004.
- [6] S. Nandi, T. C. Ilamparithi, L. Sang, and H. Doosoo, “Detection of eccentricity faults in induction machines based on nameplate parameters,” *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 1673–1683, Jul. 2010.

- [7] N. M. Elkasabgy, A. R. Eastham, and G. E. Dawson, "Detection of broken bars in the cage rotor on an induction machine," *IEEE Trans. Ind. Appl.*, vol. 28, no. 1, pp. 165–171, Jan./Feb. 1992.
- [8] A. Khezzer, M. Oumaamar, M. Hadjami, M. Boucherma, and H. Razik, "Induction motor diagnosis using line neutral voltage signatures," *IEEE Trans. Ind. Electron.*, vol. 56, no. 11, pp. 4581–4591, Nov. 2009.
- [9] V. Trana, B. Yanga, M. Oha, and A. Tanb, "Fault diagnosis of induction motor based on decision trees and adaptive neuro-fuzzy inference," *Expert Syst. Appl.*, vol. 36, no. 2, pp. 1840–1849, 2009.
- [10] S. H. Kia, A. M. Mabwe, H. Henao and G. A. Capolino, "Wavelet based instantaneous power analysis for induction machine fault diagnosis," in *Proc. 32nd IEEE Annu. Conf. Ind. Electron.*, 2006, pp. 1129–1234.
- [11] M. Benbouzid, "A review of induction motors signature analysis as a medium for faults detection," *IEEE Trans. Ind. Electron.*, vol. 47, no. 5, pp. 984–993, 2000.
- [12] B. M. Ebrahimi, J. Faiz, and M. J. Roshtkhari, "Static-, dynamic-, and mixed eccentricity fault diagnoses in permanent-magnet synchronous motors," *IEEE Trans. Ind. Electron.*, vol. 56, no. 11, pp. 4727–4739, Nov. 2009.
- [13] W. T. Thomson and M. Fenger, "Current signature analysis to detect induction motor faults," *IEEE Ind. Appl. Mag.*, vol. 7, no. 4, pp. 26–34, Jul./Aug. 2001.
- [14] M. Pineda-Sanchez, M. Riera-Guasp, J. A. Antonino-Daviu, J. Roger-Folch, J. Perez-Cruz, and R. Puche-Panadero, "Instantaneous frequency of the left sideband harmonic during the start-up transient: A new method for diagnosis of broken bars," *IEEE Trans. Ind. Electron.*, vol. 56, no. 11, pp. 4557–4570, Nov. 2009.
- [15] J. Pons-Llinares, J. A. Antonino-Daviu, M. Riera-Guasp, M. Pineda-Sanchez, and V. Climente-Alarcon, "Induction motor diagnosis based on transient current analytic wavelet transform via frequency B-Splines," *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 1530–1544, May 2011.
- [16] P. Karvelis, G. Georgoulas, I. P. Tsoumas, J. A. Antonino-Daviu, V. Climente-Alarcon, and C. D. Stylios, "A symbolic representation approach for the diagnosis of broken rotor bars in induction motors," *IEEE Trans. Ind. Inf.*, vol. 11, no. 5, pp. 1028–1037, Oct. 2015.
- [17] J. Antonino-Daviu, P. Jover, M. Riera, A. Arkkio, and J. Roger-Folch, "DWT analysis of numerical and experimental data for the diagnosis of dynamic eccentricities in induction motors," *Mech. Syst. Signal Process.*, vol. 21, no. 6, pp. 2575–2589, 2007.
- [18] A. Garcia-Perez, R. de Jesus Romero-Troncoso, E. Cabal-Yepez, and R. A. Osornio-Rios, "The application of high-resolution spectral analysis for identifying multiple combined faults in induction motors," *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 2002–2010, May 2011.
- [19] V. N. Ghate and S. V. Dudul, "Cascade neural-network-based fault classifier for three-phase induction motor," *IEEE Trans. Ind. Electron.*, vol. 58, no. 5, pp. 1555–1563, May 2011.
- [20] J. Antonino-Daviu, P. Jover Rodriguez, M. Riera-Guasp, M. Pineda-Sanchez, and A. Arkkio, "Detection of combined faults in induction machines with stator parallel branches through the DWT of the startup current," *Mech. Syst. Signal Process.*, vol. 23, no. 7, pp. 2336–2351, 2009.
- [21] G. Tsoumakas, I. Katakis, and I. Vlahavas, "Mining multi-label data," in *Data Mining and Knowledge Discovery Handbook*. New York, NY, USA: Springer, 2009, pp. 667–685.
- [22] E. Gibaja and S. Ventura, "A tutorial on multilabel learning," *ACM Comput. Surveys*, vol. 47, no. 3, p. 52, 2010.
- [23] J. Read, "Scalable multi-label classification," Ph.D. dissertation, University of Waikato, Hamilton, New Zealand. Sep. 2010.
- [24] C. M. Vong, P. K. Wong, and W. F. Ip, "A new framework of simultaneous-fault diagnosis using pairwise probabilistic multi-label classification for time-dependent patterns," *IEEE Trans. Ind. Electron.*, vol. 60, no. 8, pp. 3372–3385, Aug. 2013.
- [25] C. Yang, T.-J. Kang, D. Hyun, S. B. Lee, J. A. Antonino-Daviu, and J. Pons-Llinares, "Reliable detection of induction motor rotor faults under the rotor axial air duct influence," *IEEE Trans. Ind. Appl.*, vol. 50, no. 4, pp. 2493–2502, Jul. 2014.
- [26] G. Georgoulas, P. Karvelis, C. D. Stylios, I. P. Tsoumas, J. A. Antonino-Daviu, and V. Climente-Alarcon, "Automatizing the broken bar detection process via short time Fourier transform and two-dimensional piecewise aggregate approximation representation," in *Proc. Energy Convers. Congr. Expo.*, pp. 3104–3110, 2014.
- [27] S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed., Oxford, U.K.: Academic Press, 1999.
- [28] A. Teolis, *Computational Signal Processing with Wavelets*. Boston: Birkhauser, 1998.
- [29] E. Keogh, K. Chakrabarti, M. Pazzani, and S. Mehrotra, "Locally adaptive dimensionality reduction for indexing large time series databases," *ACM SIGMOD Rec.*, vol. 30, no. 2, pp. 151–162, May 2001.
- [30] B. K., Yi and C. Faloutsos, "Fast time sequence indexing for arbitrary Lp norms," in *Proc. VLDB*, 2000, pp. 385–394.
- [31] J. Chen, Y. S. Moon, M. F. Wong, and G. Su, "Palmprint authentication using a symbolic representation of images," *Image Vis. Comput.*, vol. 28, no. 3, pp. 343–351, 2010.
- [32] I. Jolliffe, *Principal Component Analysis*. New York: Wiley, 2005.
- [33] L. J. van der Maaten, E. O. Postma, and H. J. van den Herik, "Dimensionality reduction: A comparative review," *J. Mach. Learning Res.*, vol. 10, pp. 66–71, 2009.
- [34] C. M. Bishop, *Neural Networks for Pattern Recognition*. New York: Oxford Univ. Press, 1995.
- [35] J. Fürnkranz, E. Hüllermeier, E. L. Mencía, and K. Brinker, "Multilabel classification via calibrated label ranking," *Mach. Learning*, vol. 73, no. 2, pp. 133–153, 2008.
- [36] E. Hüllermeier, J. Fürnkranz, W. Cheng, and K. Brinker, "Label ranking by learning pairwise preferences," *Artif. Intell.*, vol. 172, no. 16, pp. 1897–1916, 2008.
- [37] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York: Wiley, 2002.
- [38] D. J. Hand, "Classifier technology and the illusion of progress," *Statist. Sci.*, vol. 21, no. 1, pp. 1–14, 2006.
- [39] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*. San Mateo, CA, USA: Morgan Kaufmann, 2005.
- [40] M. Pineda-Sanchez, V. Climente-Alarcon, M. Riera-Guasp, and R. Puche-Panadero, "Enhanced simulink induction motor model for education and maintenance training," *Int. J. Syst., Cybern. Inf.*, vol. 10, no. 2, pp. 92–97, 2012.
- [41] S. Nandi, "A detailed model of induction machines with saturation extendable for fault analysis," *IEEE Trans. Ind. Appl.*, vol. 40, no. 5, pp. 1302–1309, Sep.–Oct. 2004.
- [42] A. Ghoggal, S. E. Zouzou, M. Sahraoui, H. Derghal, and A. Hadri-Hamida, "A winding function-based model of air-gap eccentricity in saturated induction motors," in *Proc. XXth Int. Conf. Electr. Mach.*, Marseille, France, 2012, pp. 2731–2737.
- [43] W. L. Martinez, A. R. Martinez, and J. Solka, *Exploratory Data Analysis with MATLAB*. London: CRC Press, 2010.
- [44] Z. H. Zhou, M. L. Zhang, S. J. Huang, and Y. F. Li, "Multi-instance multi-label learning," *Artif. Intell.*, vol. 176, no. 1, pp. 2291–2320, 2012.



**George Georgoulas** received the Dipl.-Eng. and Dr.-Eng. degrees in electrical and computer engineering from the University of Patras, Patras, Greece, in 1999 and 2006, respectively.

He is currently a Postdoctoral Research Fellow in the Department of Computer Science, Electrical, and Space Engineering, Luleå University of Technology, Luleå, Sweden. He has over 100 publications in international journals, conferences, and books. His current research interests include machine learning, computational

intelligence signal processing, fault detection, diagnosis, failure prognosis, and the development of decision support systems.



**Vicente Climente-Alarcon** (S'11–M'12) received the M.Sc. degrees in chemical and industrial engineering in 2000 and 2011, respectively, and the Ph.D. degree in electrical engineering in 2012, all from the Universitat Politècnica de Valencia, Valencia, Spain.

He was an Assistant Professor in the School of Industrial Engineering, Universitat Politècnica de Valencia, where he was engaged in the area of condition monitoring of electrical machines and externally as a consultant in automation and management of power systems. He is currently a Docent in the Department of Electrical Engineering and Automation, Aalto University, Espoo, Finland. His current research interests include the study of transients in electrical machines.



**Jose A. Antonino-Daviu** (S'04–M'08–SM'12) received the M.S. and Ph.D. degrees in electrical engineering both from the Universitat Politècnica de València, Valencia, Spain, in 2000 and 2006, respectively.

He was with IBM during two years, where he was engaged in several international projects. He is currently an Associate Professor in the Department of Electrical Engineering, Universitat Politècnica de València, where he is involved in developing his docent and research work. He

was an Invited Professor at Helsinki University of Technology (Finland) in 2005 and 2007, Michigan State University (USA) in 2010, Korea University (Korea) in 2014, and the Universite Claude Bernard Lyon 1 (France) in 2015. He has over 100 publications in international journals, conferences, and books. His current research interests include condition monitoring of electric machines, wavelet theory and its application to fault diagnosis, and design and optimization of electrical installations and systems.



**Chrysostomos D. Stylios** (M'97–SM'15) received the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Patras, Patras, Greece, in 1999, and the Diploma in electrical and computer engineering from the Aristotle University of Thessaloniki, Thessaloniki, Greece, in 1992.

He is currently an Associate Professor in the Department of Computer Engineering, Technological Educational Institute of Epirus, Arta, Greece. In the academic year 2014–2015, he

was a Visiting Researcher in the Electrical and Computer Engineering Department, University of Alberta, Canada, and in the Computer Science Department, University of Texas at El Paso. He has authored or coauthored over 160 journal and conference papers and book chapters. His current research interests include fuzzy cognitive maps, soft computing, computational intelligence techniques, signal processing methods, and decision support systems.

Dr. Stylios is a Member of the TC 8.2 and TC 5.4 of IFAC.



**Ioannis P. Tsoumas** (S'04–M'08) received the Dipl.-Eng. and Dr.-Eng. degrees in electrical and computer engineering from the University of Patras, Patras, Greece, in 2000 and 2007, respectively.

From 2008 to 2015, he was with Siemens AG, Industry Sector, Drive Technologies Division, Large Drives, Products Research & Development, Nuremberg, Germany. In 2015, he joined the ABB Corporate Research Center, Baden-Dättwil, Switzerland. His current research inter-

ests include medium-voltage power converters and drives, modulation techniques for power converters, motor–converter interaction, converter and drive system efficiency, and electrical machines and drives condition monitoring and fault diagnosis.

Dr. Tsoumas is member of the Association for Electrical, Electronic and Information Technologies (VDE) and of the Technical Chamber of Greece.



**Antero Arkkio** was born in Vehkalahti, Finland, in 1955. He received the M.Sc. (Tech.) and D.Sc. (Tech.) degrees from Helsinki University of Technology (Aalto University), Espoo, Finland, in 1980 and 1988, respectively.

He is currently a Professor of Electrical Engineering at Aalto University. His current research interests include modeling, design, and measurement of electrical machines.



**George Nikolakopoulos** (S'00–M'06) received the M.S. and Ph.D. degrees in electrical engineering both from the University of Patras, Patras, Greece, in 2000 and 2006, respectively.

He is currently a Professor of Robotics and Automation in the Department of Computer Science, Electrical, and Space Engineering, Luleå University of Technology, Luleå, Sweden. His current research interests include robotics and control applications, with a special focus on aerial robotics, UGVs, and bio-inspired robotics.

He has been the Coordinator of several European (Horizon20020) and National Research Projects. His published scientific work includes more than 170 published international journals and conferences in the fields of his interest.