



## Efficient feature extraction technique for diagnosing broken bars in three-phase induction machines



Luisa H. B. Liboni<sup>a,b,\*</sup>, Rogério Andrade Flauzino<sup>a</sup>, Ivan Nunes da Silva<sup>a</sup>, Eduardo C. Marques Costa<sup>c</sup>

<sup>a</sup> Department of Electrical and Computer Engineering, EESC – University of São Paulo, CP 359, CEP 13566-590 São Carlos, SP, Brazil

<sup>b</sup> Department of Electrical and Computer Engineering, IFSP – Federal Institute of Education, Science, and Technology of São Paulo, CEP 14160-290 Sertãozinho, SP, Brazil

<sup>c</sup> Department of Energy and Automation Engineering, EP – University of São Paulo, CEP 05508-900 São Paulo, SP, Brazil

### ARTICLE INFO

#### Article history:

Received 29 August 2017

Received in revised form 28 November 2018

Accepted 1 December 2018

Available online 03 December 2018

#### Keywords:

Broken bars

Fault diagnosis

Induction machine

Orthogonal decomposition

Support Vector Machines (SVMs)

Feature extraction

### ABSTRACT

An original diagnostic method is proposed for identifying broken rotor bars based on a new technique for decomposing electric signals obtained by measurements, the Orthogonal Component Decomposition technique. This decomposition shows to be an efficient fault-specific feature extractor, and a broken bar simulation is used to determinate the relationships between the decomposition products and the fault phenomena. The efficiency of the information extraction is evaluated by Kruskal-Wallis variance analyses and Support Vector Machines on experimental signals, where the fault occurrence is detected, and the fault severity, given by the number of broken bars, is also diagnosed. The experimental signals are obtained by measurements from induction machines operating with different torque levels and driven either directly by the grid or frequency inverters. The proposed diagnostic method does not depend on frequency analysis and spectral resolution and comprises the Orthogonal Component Decomposition and Support Vector Machines. Furthermore, the results demonstrate the effectiveness of the information extraction since high diagnostic accuracy is achieved in low-torque situations and in a broad range of slip values. Additional analyses support that the proposed decomposition should be further investigated for many other applications on event detection.

© 2018 Elsevier Ltd. All rights reserved.

### 1. Introduction

Induction machines (IMs) are one of the main electromechanical devices used in industries and are widely used in electric power systems in wind-power generators and bulb hydro-power turbines [1]. With advances in power electronics and the advent of high-performance microprocessors, three-phase induction machines (TIMs) have been used in various types of industrial applications, including those with variable speed requirements.

Although IMs present good efficiency and long life with low maintenance costs, they are dynamic systems subjected to mechanical and electrical faults, which may lead to interruptions in manufacturing processes and economic losses. By knowing when an electric machine requires maintenance, costly unscheduled shutdowns and the inefficient use of electric power could be avoided.

Faults in rotor bars constitute about 10–20% of the total faults in TIMs [2], which leads to a significant reduction in the life cycle of

these devices, decreases average torque and mechanical vibration [3–5]. Moreover, broken bars in TIMs represent a cumulative and destructive event, since neighboring bars also deteriorate due to mechanical and thermal stresses [2,3,6,7]. In this context, broken bars should be early detected by a fault identification system.

Fault identification methods are widely applied for diagnosing abnormal events in the operation of several dynamic systems. These methods are based on the monitoring of different variables and signal processing techniques [8]. Different types of variables were used in several efforts for fault detection and classification in TIMs, such as stator voltage and current [9–14], active and reactive power [15–17], torque profile [4,18], and vibration analysis [19].

In this sense, signal processing of such monitored variables is essential for a reliable and accurate diagnostic in any electrical system [3,20]. Therefore, results from signal processing must contain powerful fault-characterization indices for intelligent classifiers or statistical methods to complete the diagnostic process. Signal processing methods range from time and frequency domains to time-frequency domain [21], such as statistical measurements [20], Hilbert Transform [14], Fast Fourier transform (FFT), and Wavelet Transform (WT) [3,22–24].

\* Corresponding author at: Department of Electrical and Computer Engineering, IFSP – Federal Institute of Education, Science, and Technology of São Paulo, CEP 14160-290 Sertãozinho, SP, Brazil.

E-mail address: [luisa.liboni@usp.br](mailto:luisa.liboni@usp.br) (L. H. B. Liboni).

Most studied feature-extraction techniques for the detection of broken rotor bars are based on the spectral analysis (MCSA) of stator currents. [13,24–27]. Broken rotor bars cause rotor asymmetries and distortion in stator currents, which incurs in some fault-characteristic frequencies in the spectrum of such currents. Fault-specific lower and upper sidebands in the stator current spectrum are found at  $f_{1,2} = (1 \pm 2ks)f_s = (1 \pm kf_b)$ , where  $k$  is an integer,  $s$  is the slip,  $f_s$  is the supply frequency, and  $f_b$  is called the sideband frequency and is given by  $2sf_s$ . The exact location of these sidebands in the spectrum depends directly on the slip, and their amplitudes depend on the physical position of broken bars, number of broken bars and also on the load profile [28–31].

A well-known fact in the diagnosis of TIMs is that it is harder to identify a broken bar with frequency-based analysis of stator currents when the machine is lightly loaded, because of the proximity of the sidebands with the fundamental frequency and the small disturbance caused by the fault at low-load conditions [10,6,32]. Hence, the accuracy of frequency-based techniques depends on spectral resolution and needs a large data acquisition interval to minimize leakage [3], which could obscure the fault-specific frequency sidebands. In the same way, diagnostic methods based on the analysis of stator-current envelopes also present identification difficulties in low-torque conditions [10].

Moreover, when frequency-domain techniques are used and slip increases, frequencies  $f_{1,2}$  get pushed away from the fundamental frequency. As a selected number of frequency points around the fundamental is used as inputs to classifiers, diagnostic accuracy could drop, since diagnostic systems could lose track of sidebands by just using data points that are close to the fundamental frequency. Hence, to bypass this drawback on high-slip situations, processing algorithms need to dynamically choose some frequency data points by having prior knowledge of the speed of the machine.

Statistical methods can undertake the detection and classification task of whether a machine is in a healthy or faulty condition after fault-specific features were extracted by the processing of monitored variables. However, several approaches combine signal processing tools with concepts of intelligent systems, such as neural networks [33,34], fuzzy logic [35,36], Support Vector Machines (SVMs) [37], particle swarm optimization [38], genetic algorithms [39], and Bayesian classifiers [10,18].

Several efforts in fault detection have been made using the combination of signal processing and intelligent systems. Nonetheless, the majority of these methods use frequency-domain transforms, mainly based on the MCSA. Frequency-domain transforms are even used on those methods based on the analysis of instant active and reactive powers [16,15].

This paper describes an efficient feature extraction technique for diagnosing broken rotor bars based on a time-domain mathematical technique for decomposing electric signals, named the Orthogonal Component Decomposition (OCD), that can overpass some of the difficulties already discussed. The OCD technique was successfully used in [40] for detection of high-impedance faults in distribution systems, and a careful presentation of the novel technique was made in [41]. Differently from other very useful signal transforms and decompositions, as the FFT, in which signals are decomposed into some predefined orthogonal basis, the OCD technique does not decompose signals in pre-established basis but uses the signals that are being analyzed as the basis for the decomposition. Accordingly, the decomposition products are highly sensitive to the system operation and are strongly related to the system operating parameters and power consumption. In fact, this information on electric parameters and power consumption allow the tool to track different events, such as faults, changes in load and operating conditions in a dynamic electrical system.

Based on the aspects mentioned above, the main contributions of the proposed method are: i) Fault analysis is made in the time domain so that frequency-domain transforms are not required. With signal decomposition in the time domain, a heuristic understanding of the fault process can be made by observing the trends of the products of the decomposition; ii) The decomposition is not based on a system of predefined basis, i.e., the decomposition basis is determined directly from the electric signals being analyzed. The decomposition products are sensitive to the system operation. Therefore, it is shown that the OCD is an effective fault-specific feature extractor even when machines are lightly loaded, which eases the classification made by intelligent systems or statistical methods; iii) With the proposed method, by using the new signal processing tool, the high-slip tracking problem is overcome since the method is processed on the time-domain. Even when machines have a high slip, a reliable fault signature is extracted without prior knowledge of the speed of the machine.

Variance analyses were performed to demonstrate the effectiveness of the OCD technique as a fault-specific feature extractor. Moreover, to present a complete diagnostic method, SVM classifiers were trained offline with experimental data in order to diagnose faulty motors. Experimental data consist of electric signals that were processed with the OCD technique. The machine learning algorithm could then identify the faulty machines as well as the severity of the fault, which is considered as the number of broken bars.

## 2. Mathematical and computational concepts

### 2.1. Orthogonal Component Decomposition technique (OCD)

The OCD is a new signal processing technique designed for electric power systems, mainly assisting the identification of nonlinear processes and events, such as faults and changes in the operating condition of these systems. The technique is a method that decomposes the electric signals being analyzed into the own signals and not into pre-established basis. Therefore, the resulting decomposition products, named the orthogonal components, are sensitive to the system operation and power consumption.

The OCD is based on two principles derived from linear algebra concepts. A careful presentation of the OCD can be found in [41]. However, as the technique was recently proposed, a brief summary is presented in this paper.

#### 2.1.1. Principle 1 – decomposition into orthogonal quantities

Consider two continuous, finite, and periodic functions, from an electrical system:  $i(t)$  and  $v(t)$ . Thus, a derivative function  $\dot{i}(t)$  and an integrative function  $\tilde{i}(t)$  can be expressed in (1) and (2), respectively. These functions can be considered as vectors belonging to the  $\mathbb{R}^\infty$  vector space.

$$\dot{i}(t) = \frac{di(t)}{dt} \quad (1)$$

$$\frac{d\tilde{i}(t)}{dt} = i(t) \quad (2)$$

The purpose of the OCD is to outcome products that are highly sensitive to the electrical system parameters and power consumption. Therefore, a decomposition as formulated in (3) would surely be very interesting for this purpose. One can observe that Signal  $v(t)$  is decomposed into four parts. Voltage quantity  $v^p(t)$  is proportional to the system current and hence carries out a resistance ( $R$ ) information through coefficient  $\gamma_1$ . Furthermore, coefficient  $\gamma_2$  in quantity  $v^{q1}(t)$  carries out an inductance ( $L$ ) information, since this component is proportional to the derivative function of the

current. Finally,  $\gamma_3$  carries out a capacitance ( $\frac{1}{C}$ ) information since  $v^q(t)$  is proportional to the integrative function of the current. The last quantity in (3) is made necessary as the signal  $v(t)$  in  $\mathbb{R}^\infty$  may not be spanned by just the three quantities mentioned. Consequently, the last quantity of the decomposition carries out information that can not be modeled by passive electric elements.

$$v(t) = \underbrace{\gamma_1 i(t)}_{v^p(t)} + \underbrace{\gamma_2 \dot{i}(t)}_{v^q(t)} + \underbrace{\gamma_3 \ddot{i}(t)}_{v^d(t)} + v^d(t) \quad (3)$$

It is essential that the components in which a signal is decomposed are linearly independent and orthogonal to each other, in order to the  $\gamma$  coefficients in (3) to be independent. In [41] it is shown that a continuous, finite, and periodic function  $f(t)$  is mutually orthogonal to its derivative function  $\dot{f}(t)$  and integrative function  $\tilde{f}(t)$ , which, are not orthogonal to each other, being collinear (anti-parallel) if  $f(t)$  is purely sinusoidal.

Accordingly, an orthogonalization process, as shown in (4) is necessary to make all quantities in (3) independent. The new orthogonal triplet, which  $v(t)$  will be decomposed into, is now given by  $i(t)$ ,  $\dot{i}(t)$ , and a new vector  $i^\perp(t)$ , which is obtained by a Gram-Schmidt orthogonalization.

$$i^\perp(t) = \tilde{i}(t) - \frac{\langle \tilde{i}(t), \dot{i}(t) \rangle}{\|\dot{i}(t)\|^2} \dot{i}(t) \quad (4)$$

After the orthogonalization process, the decomposition can be rewritten as formulated in (5). Thus, all the quantities of the decomposition are entirely independent, where  $\tilde{i}(t)$  was substituted by  $i(t)^\perp$ .

$$v(t) = \underbrace{\gamma_1 i(t)}_{v^p(t)} + \underbrace{\gamma_2 \dot{i}(t)}_{v^q(t)} + \underbrace{\gamma_3 i^\perp(t)}_{v^{q^\perp}(t)} + v^d(t) \quad (5)$$

It is important to state that if the signals are purely sinusoidal, then  $\dot{i}(t)$  and  $\tilde{i}(t)$  are anti-parallel and  $i^\perp(t)$  is null. Therefore, the existence of the latter highlights reactive phenomena in non-purely sinusoidal systems, being an efficient index for detecting nonlinear events in power systems [41,40].

Component  $v^d(t)$  is the signal remainder and is proved to be orthogonal to the other decomposed quantities [41]. It can be calculated as shown in (6).

$$v^d(t) = v(t) - \underbrace{\gamma_1 i(t)}_{v^p(t)} - \underbrace{\gamma_2 \dot{i}(t)}_{v^q(t)} - \underbrace{\gamma_3 i^\perp(t)}_{v^{q^\perp}(t)} \quad (6)$$

All the steps made to achieve the decomposition formulated in (5) can also be made to decompose  $i(t)$  as demonstrated in (7). Therefore, in an electrical system, each pair of current and voltage results in eight orthogonal functions that are the products of the OCD and are named orthogonal components.

$$i(t) = \underbrace{\alpha_1 v(t)}_{i^p(t)} + \underbrace{\alpha_2 \dot{v}(t)}_{i^q(t)} + \underbrace{\alpha_3 v^\perp(t)}_{i^{q^\perp}(t)} + i^d(t) \quad (7)$$

For the pair  $i(t)$  and  $v(t)$ , eight orthogonal components are obtained:  $i^p(t)$ ,  $i^{q^\perp}(t)$ ,  $i^q(t)$ ,  $i^d(t)$ ,  $v^p(t)$ ,  $v^{q^\perp}(t)$ ,  $v^q(t)$ ,  $v^d(t)$ . The  $\gamma$  and  $\alpha$  coefficients can be calculated from the projection of the function to be decomposed onto the triplet vectors as demonstrated in (8) and (9).

$$\gamma_i = \frac{\langle v(t), x(t) \rangle}{\|x(t)\|^2} \quad (8)$$

$$\alpha_i = \frac{\langle \dot{i}(t), y(t) \rangle}{\|y(t)\|^2} \quad (9)$$

when  $i = 1$ ,  $x(t) = i(t)$  and  $y(t) = v(t)$ , when  $i = 2$ ,  $x(t) = \dot{i}(t)$  and  $y(t) = \dot{v}(t)$  and, finally, when  $i = 3$ ,  $x(t) = i^\perp(t)$  and  $y(t) = v^\perp(t)$ . Moreover,  $\langle w(t), z(t) \rangle = \frac{1}{T} \int_{t-T}^t w(\tau) z(\tau) d\tau$ , and  $\|z(t)\|^2 = \langle z(t), z(t) \rangle$ , which is the square of the RMS value of  $z(t)$ .

### 2.1.2. Principle 2 – decomposition onto the plane of line voltages

Let  $v_A(t)$ ,  $v_B(t)$ , and  $v_C(t)$  be continuous functions, periodic on  $T$  and vectors of  $\mathbb{R}^\infty$  space. These functions represent phase voltages of any three-phase system, where,  $v_{AB}(t)$ ,  $v_{BC}(t)$ , and  $v_{CA}(t)$  are the line voltages of this system.

As line voltages of a three-phase system are obtained by the difference between two phase voltages, signals  $v_{AB}(t)$ ,  $v_{BC}(t)$ , and  $v_{CA}(t)$  are filtered functions since any noise equally incident on phase voltages is eliminated.

In [41] it is shown that continuous and periodic line voltages of a three-phase system are always linearly dependent regardless of the content of phase voltages. Therefore, line-voltage vectors in  $\mathbb{R}^\infty$  are positioned in a plane, which spans signals filtered from common-mode noises.

In order to filter a given signal from common-mode noises, one can project the signal onto the plane of line voltages. For example, phase and neutral voltages  $v_A(t)$ ,  $v_B(t)$ ,  $v_C(t)$ , and  $v_N(t)$  and phase and neutral currents  $i_A(t)$ ,  $i_B(t)$ ,  $i_C(t)$ , and  $i_N(t)$  can be projected on the plane of line voltages.

To accomplish such projection task, the given signal must be projected onto an orthogonal basis of the plane of line voltages. This basis can be generated by adopting an arbitrary line voltage as the first basis vector. The second basis vector should be orthogonal to the first. Therefore, the following plane basis is obtained:

$$\begin{cases} v_{b_1}(t) = v_{AB}(t) \\ v_{b_2}(t) = v_{BC}(t) - \frac{\langle v_{AB}(t), v_{BC}(t) \rangle}{\|v_{AB}(t)\|^2} v_{AB}(t) \end{cases} \quad (10)$$

After a phase signal is decomposed into the plane of line voltages, one can understand the phase signal as the sum of two components, as shown in (11). For example, take  $v_A(t)$ ,  $v_B(t)$ ,  $v_C(t)$ , and  $v_N(t)$ . The first component is the planar component, i.e., the one that was decomposed into the plane denoted as  $v_a(t)$ ,  $v_b(t)$ ,  $v_c(t)$ , and  $v_n(t)$ . The second component is an orthogonal to the plane quantity, denoted as  $v_a(t)$ ,  $v_b(t)$ ,  $v_c(t)$ , and  $v_n(t)$ . These last components are not spanned by the plane basis and can contain other noises and nonlinear contents.

$$v_X(t) = \underbrace{\beta_x v_{b_1}(t) + \delta_x v_{b_2}(t)}_{v_x(t)} + v_x \quad (11)$$

where  $x$  can assume  $a$ ,  $b$ ,  $c$  and  $n$ .

Coefficients  $\beta$  and  $\delta$  in (11) are calculated by the projection of the phase signal onto the plane basis vectors:

$$\beta_x = \frac{\langle v_X(t), v_{b_1}(t) \rangle}{\|v_{b_1}(t)\|^2} \quad (12)$$

$$\delta_x = \frac{\langle v_X(t), v_{b_2}(t) \rangle}{\|v_{b_2}(t)\|^2} \quad (13)$$

### 2.1.3. Obtaining orthogonal components

In three-phase electric systems, both principles are used to obtain orthogonal components, as initially proposed in [42]. Signals are first filtered, i.e., decomposed into planar and non-planar components by using the second principle. Then each voltage/current pair obtained can be further decomposed into electrical information by using the first principle.

For example, a planar phase-A current  $i_a(t)$  can be decomposed into the triplet vectors  $v_a(t)$ ,  $\dot{v}_a(t)$ , and  $v_a^\perp(t)$  of the planar phase-A voltage. Likewise, a non-planar phase-A voltage  $v_a(t)$  can be

decomposed into the triplet vectors  $i_a(t)$ ,  $\dot{i}_a(t)$ , and  $i_a^\perp(t)$  of non-planar phase-A current  $i_a(t)$ . Thus, 64 time-dependent orthogonal signals are obtained from (14) with  $x = a, b, c, n, \bar{a}, \bar{b}, \bar{c}, \bar{n}$ .

One of the heuristic justifications of why the decomposition is sensitive to the system operation and parameters is fundamentally tied to the fact that the  $\gamma$  and  $\alpha$  coefficients are linked to resistances, capacitances, inductances and the remaining quantities are linked to noise and nonlinear contents. It is important to observe that the orthogonal components are highly sensitive to nonlinear events because the event changes the system parameters and the power consumption of the system.

$$\left\{ \begin{array}{l} v_x^p(t) = \frac{\langle v_x(t), \dot{i}_x(t) \rangle}{\|\dot{i}_x(t)\|^2} \dot{i}_x(t) \\ i_x^p(t) = \frac{\langle \dot{i}_x(t), v_x(t) \rangle}{\|v_x(t)\|^2} v_x(t) \\ v_x^{q\parallel}(t) = \frac{\langle v_x(t), \dot{i}_x(t) \rangle}{\|\dot{i}_x(t)\|^2} \dot{i}_x(t) \\ i_x^{q\parallel}(t) = \frac{\langle \dot{i}_x(t), v_x(t) \rangle}{\|v_x(t)\|^2} v_x(t) \\ v_x^{q\perp}(t) = \frac{\langle v_x(t), i_x^\perp(t) \rangle}{\|i_x^\perp(t)\|^2} i_x^\perp(t) \\ i_x^{q\perp}(t) = \frac{\langle i_x(t), v_x^\perp(t) \rangle}{\|v_x^\perp(t)\|^2} v_x^\perp(t) \\ v_x^d(t) = v_x(t) - v_x^p(t) - v_x^{q\parallel}(t) - v_x^{q\perp}(t) \\ i_x^d(t) = i_x(t) - i_x^p(t) - i_x^{q\parallel}(t) - i_x^{q\perp}(t) \end{array} \right. \quad (14)$$

## 2.2. Relationship between orthogonal components and the phenomena of broken bars

In order to overcome problems with frequency transformations, such as spectral resolution and leakage, one can obtain time-dependent metrics from the orthogonal components to be used as inputs to a classification system. For such, it is necessary to understand the relationship between the orthogonal components and the phenomena of broken bars.

In [43], it is shown that the instantaneous active and reactive power consumptions of a TIM oscillate when a phenomenon of broken bars occurs. More specifically, it is shown that the ratio between the maximum value and the mean value of these oscillations is approximately proportional to the number of broken bars in the rotor. In the case of healthy machines, the instantaneous active and reactive power consumption is maintained constant. As the orthogonal components are related to power consumption and based in [44,17,43], a new index obtained from the orthogonal components is proposed to be used as inputs to classification systems.

It is verified that the index obtained by the ratio between the maximum RMS value of an orthogonal components and its mean RMS value within an acquisition window is approximately constant for a given number of broken bars. Therefore, this index is monotonic with the fault severity. In order to understand this index, take the component  $i_a^p(t)$  as an example, which is not yet projected onto the plane of line voltages:

$$i_a^p(t) = \langle i_a(t), v_A(t) \rangle \frac{v_A(t)}{\|v_A(t)\|^2}. \quad (15)$$

Assuming that phase voltage  $v_A(t)$  is given by:

$$v_A(t) = \sqrt{2}V \cos(\omega t), \quad (16)$$

the stator current of a machine with broken bars is given by [17]:

$$i_a(t) = \sqrt{2}I \cos(\omega t - \varphi) + \sqrt{2} \sum_{k=1}^{\infty} \left[ I_{b1k} \cos((\omega - k\omega_{bk})t - \varphi_{b1k}) + I_{b2k} \cos((\omega + k\omega_{bk})t - \varphi_{b2k}) \right], \quad (17)$$

where  $\omega$  is the angular frequency of the machine,  $\omega_b$  is the angular frequency induced by the broken bar fault,  $I$  and  $\varphi$  are respectively the RMS value and phase of the fundamental current. Moreover,  $I_{b1k}$ ,  $I_{b2k}$ ,  $\varphi_{b1k}$  and  $\varphi_{b2k}$  are the RMS values and phase of the fault-induced sidebands.

From calculating  $\langle i_a(t), v_A(t) \rangle$ , one can observe that part of the result is composed by a component regarding the motor in a healthy condition and other component regarding the motor in a faulty condition:

$$\langle i_A, v_A \rangle = \frac{1}{T} \int_{t-T}^t i_A(t) v_A(t) d\tau = \langle i_A(t), v_A(t) \rangle_{\text{healthy}} + \langle i_A(t), v_A(t) \rangle_{\text{faulty}} \quad (18)$$

The quantity  $\langle i_A(t), v_A(t) \rangle_{\text{healthy}}$ , given by:

$$\langle i_A(t), v_A(t) \rangle_{\text{healthy}} = \frac{1}{T} \int_{t-T}^t \sqrt{2}I \cos(\omega t - \varphi) \sqrt{2}V \cos(\omega t) \tau = VI \cos(\varphi) \quad (19)$$

is the active power, as can be seen in (19). This term is time-invariant, and hence, if there is no fault, the dot product  $\langle i_A(t), v_A(t) \rangle$  is constant within a given period  $[t, t+T]$ . Therefore, this constant term will multiply a normalized purely sinusoidal signal  $\frac{v_A(t)}{\|v_A(t)\|^2}$  and the RMS values of  $i_A^p(t)$  will be constant in time. If there is a fault, one parcel of the dot product will be constant  $\langle i_A(t), v_A(t) \rangle_{\text{healthy}}$ . However, the other parcel  $\langle i_A(t), v_A(t) \rangle_{\text{faulty}}$  will oscillate since the sidebands will make the dot product oscillate within a given period  $[t, t+T]$ . This oscillating term multiplied by the normalized purely sinusoidal signal  $\frac{v_A(t)}{\|v_A(t)\|^2}$  will cause a modulation in the orthogonal component. Therefore, the RMS values of  $i_A^p(t)$  will oscillate within  $[t, t+T]$ , as shown in Figs. 1 and 2. The mean of the RMS value is related to the parcel  $\langle i_A(t), v_A(t) \rangle_{\text{healthy}}$ , which is practically constant for any number of broken bars. The amplitude of the oscillation is related to the amplitude of the sidebands considered inside the term  $\langle i_A(t), v_A(t) \rangle_{\text{faulty}}$ , thus to the number of broken bars. Finally, the slip of the motor will only influence the period of the oscillations since  $\omega_{bk}$  is considered inside the dot product.

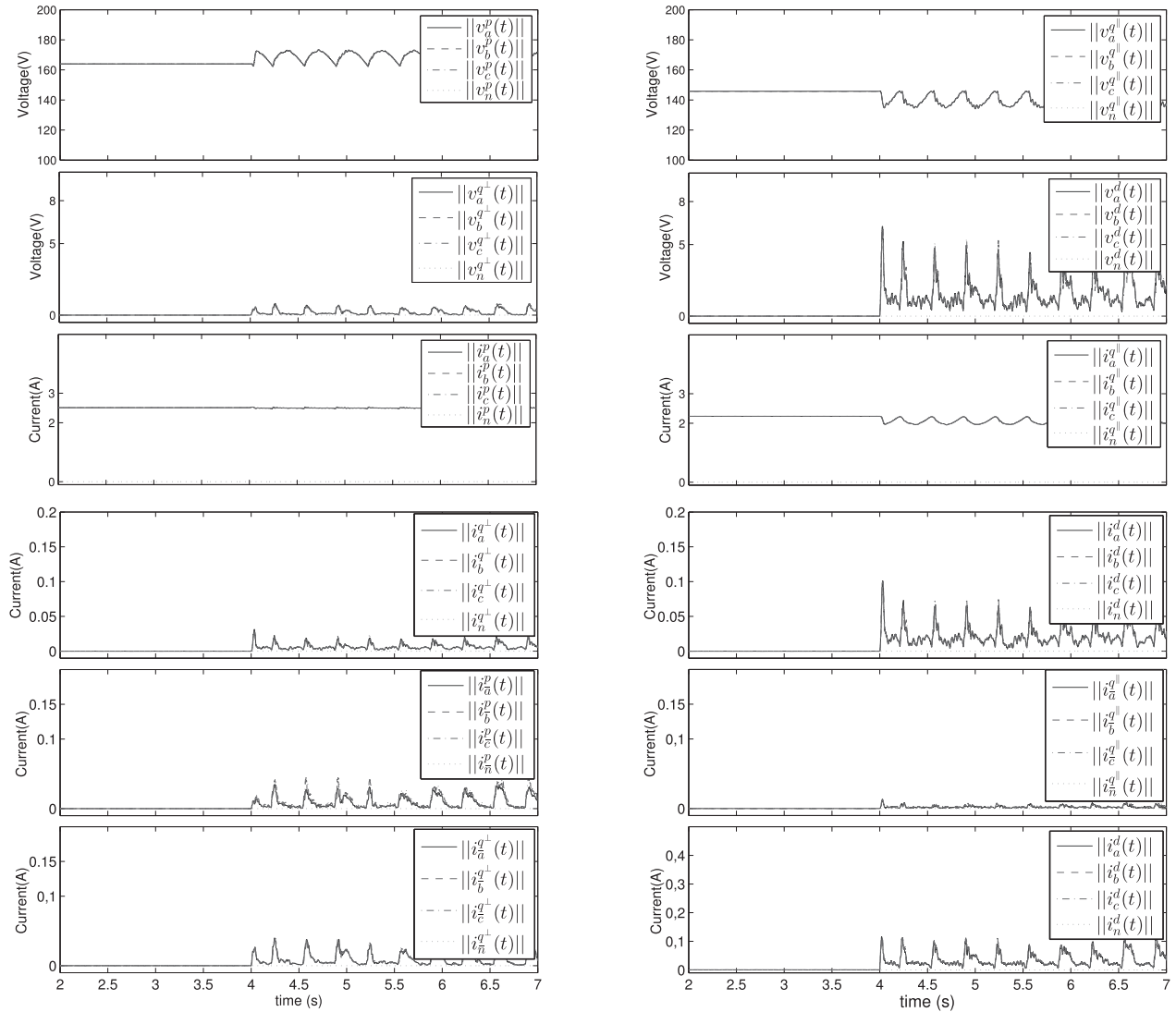
When a fault occurs, the projection of  $i_A^p(t)$  onto the plane will have a non-unitary ratio between its maximum RMS value and mean RMS value, which can be calculated as:

$$R = \frac{\max \|i_A^p(t)\|}{\text{mean} \|i_A^p(t)\|}. \quad (20)$$

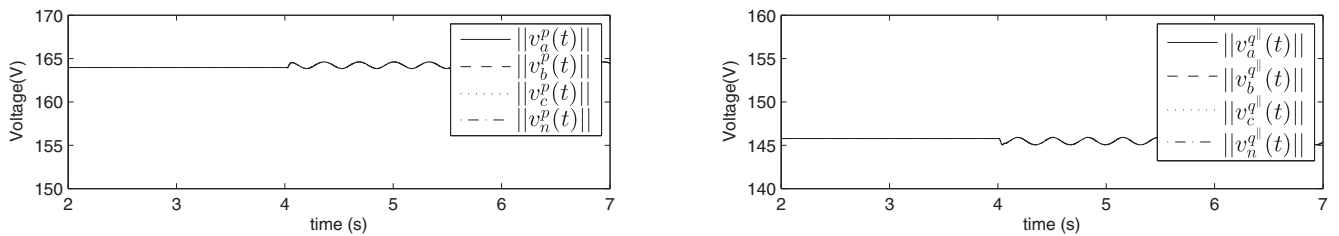
The index  $R$  is very interesting since it is independent of the load torque, this is, all motors on a healthy condition with any value of load torque have  $R = 1$  for the orthogonal components.

To observe the dynamic relationship between orthogonal components and a broken bar fault, a TIM with four poles was simulated (2 HP, 380 V/60 Hz, 3.6 A, 28 rotor bars). The simulation was based on the rotor asymmetry mesh model proposed in [45]. In this simulation, the slip frequency was at 1.53 Hz and torque was approximately 8.0 Nm (rated torque). At simulation time  $t = 4$  s a 4 adjacent broken bar fault (4 bb) was provoked (14.3% of the total number of bars) and simulation ran until  $t = 10$  s. Stator currents and supply voltages were then processed with the OCD.

Fig. 1 shows the RMS values  $\|v_x^p(t)\|$  and  $\|i_x^p(t)\|$  of the orthogonal components  $v_x^p(t)$  and  $i_x^p(t)$  with respect to time. It is noticeable that the RMS values of components, with  $x = a, b, c$  respond to the



**Fig. 1.** Orthogonal component RMS value in time for a simulated 2-hp 60-Hz four-pole 380-V 3.6-A 28-rotor-bars IM. A 4 bb (14.3% of the total number of bar) fault is provoked in  $t = 4$  s. Slip frequency is at 1.53 Hz and torque is approximately 8.0 Nm (rated torque).



**Fig. 2.** RMS values of some orthogonal components with respect to time for a simulated 2-hp 60-Hz four-pole 380-V 3.6-A 28-rotor-bars IM. A 1 bb (3.6% of the total number of bar) fault is provoked in  $t = 4$  s. Slip frequency is at 1.53 Hz and torque is approximately 8.0 Nm.

fault phenomena and oscillates with frequency  $f_b = 2sf_s$ , because of the stator current modulation imposed by the fault-characteristics sidebands, which in this case, is approximately 3 Hz. In this case, phases a, b, c have the same behavior, as expected, since the current sidebands are induced in all phases by the modulation in the rotating magnetic flux. By comparing Figs. 1 and 2, where some components are shown for the case of a 1 broken bar fault (1 bb) (3.6% of the total number of bars), it is

noticeable that the amplitude of this oscillations increases with the number of broken bars.

Orthogonal components derived from neutral voltages are null, because there is no supply-voltage imbalance and although usually IMs have no neutral connection, the OCD processing tool formulation expects the decomposition of these signals. In this way, to use the proposed technique, a theoretical definition was adopted:  $v_n(t) = v_a(t) + v_b(t) + v_c(t)$  and  $i_n(t) = i_a(t) + i_b(t) + i_c(t)$ .

Components  $\|v_x^q(t)\|$  and  $\|i_x^q(t)\|$  also react to the fault and components  $\|v_x^+(t)\|$  and  $\|i_x^+(t)\|$  become non-zero. These components highlight the reactive nature of broken bars [16]. Although these signals are not power signals,  $\|v_x^+(t)\|$  and  $\|i_x^+(t)\|$  are related to the existence of reactive parameters in non-purely sinusoidal systems. Non-planar voltage components,  $\|v_x^p(t)\|$ ,  $\|v_x^q(t)\|$ ,  $\|v_x^{q+}(t)\|$ , and  $\|v_x^d(t)\|$  are null and not shown because all voltages are symmetrical and purely sinusoidal. Therefore, no voltage quantities are unfiltered. Differently, non-planar currents  $\|i_x^p(t)\|$ ,  $\|i_x^q(t)\|$ ,  $\|i_x^{q+}(t)\|$ , and  $\|i_x^d(t)\|$  are non-zero when the fault is provoked because current modulation is present. Moreover, components representing non-passive electric elements are null until the occurrence of the fault.

It can be seen that components representing non-passive elements, and components representing active and reactive elements have oscillations highly related to the fault. When real signals are analyzed, non-planar voltage orthogonal components are non-zero, because of noise, and orthogonal components derived from neutral voltages are non-zero because of possible phase-voltage imbalances.

Considering that the load torque is kept constant, the acquisition time of voltage and current signals must be large enough to contain at least one period of oscillations of the RMS values of the orthogonal components so that the calculation of the  $R$  index could be made correctly. As the RMS values oscillate with frequency  $fb = 2sf_s$ , the minimum acquisition time can be easily calculated a priori by using the slip value when the motor is free of load torque.

In Fig. 3, Pearson's absolute correlation of the  $R$  index of some orthogonal components with respect to the state of the simulated motor (healthy or faulty with 1 bb, 2 bb, and 4 bb), in several load torques, is plotted for an acquisition window of 3 s. It can be observed that the indices maintain an almost-fixed correlation value with respect to load torque. Therefore, when machines are lightly loaded, orthogonal components still highlight the fault occurrence. The acquisition window of 3 s is sufficient to contain at least an integer number of oscillations for the lowest load torque condition (1 Nm).

### 2.3. Feature selection and dimensionality reduction

In problems regarding pattern classification, attribute selection is used to hand the most relevant features or variables to a given

classifier. Feature selection also promotes a dimensionality reduction of the input space, providing a set of more compact and relevant data, and sometimes improving the performance of intelligent classifiers.

Various techniques for reducing feature space can be found, for example, Principal Components Analysis (PCA), in which, new highly correlated attributes are obtained as linear combinations of the original ones. A disadvantage of this technique is that no unique attribute can be discarded.

In this paper, a correlation-based feature selection (CFS) algorithm addresses the attribute selection problem through a correlation-based analysis for selecting a good set of  $R$  indices from all the 64 orthogonal components. This attribute set is used by the classification system to point out a broken bar phenomenon and its severity. A good attribute set contains features that are highly correlated with the class, in this case, the faulty machines, but not greatly correlated with each other [46].

The algorithm has an appropriate correlation measure called symmetric uncertainty, based on entropy metric, and a heuristic search strategy. The search strategy performs a greedy conservative forward search through the space of attribute subsets [46].

### 2.4. Support Vector Machines

SVM algorithms are used for pattern classification, regression, and novelty detection. It has been extensively used in the area of machine diagnostics [47]. They can be considered as learning algorithms with a hidden layer and supervised training based on structural risk minimization, arising from statistical learning theory [48]. In pattern classification problems the algorithm maximizes the margin, which is the minimal distance between a hyperplane separating the classes and the data points closest to this hyperplane. The most important vectors that define the hyperplane position are called support vectors (SVs). The tool can be used in high-dimensional input spaces, exhibiting good generalization. SVM learning involves the optimization of a convex function and does not have local minima, and few parameters are required for tuning the learning machine [48].

Data must be mapped into a new larger input space so that the problem in this new space becomes separable. This mapping is done by a kernel function. Usually, in very high-dimensional input space, the linear function, shown in (21), is the kernel of choice.

The original SVM mathematical formulation is called *hard-margin SVM*, in which data are not permitted inside the margin. However, generalization errors can be reduced by using the

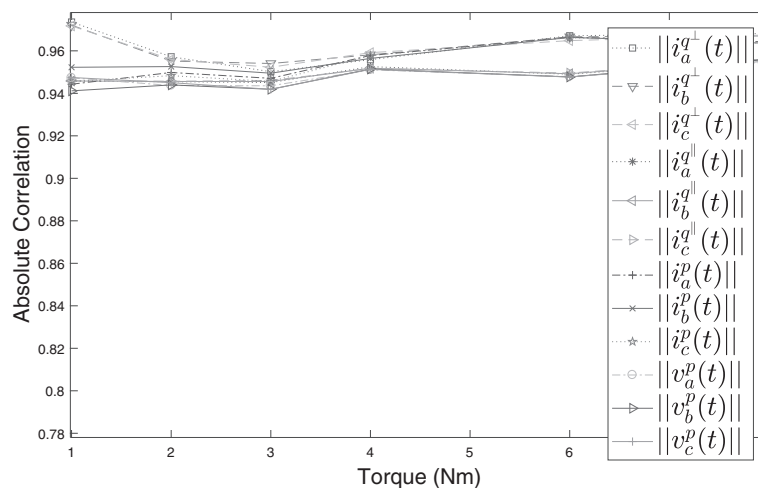


Fig. 3. Correlation between the  $R$  index of some orthogonal components and the state of the motor as a function of load torque.

concept of *soft margins*. According to [48], with the soft-margin mathematical formulation, some data are allowed to violate the margin constraints, in such cases, some of the SVs may lie within the margin bounds. When using the soft-margin approach, the regularization parameter  $C$  controls the trade-off between achieving a low error on the training data and minimizing the generalization error. Moreover, in the case of an imbalance between the amount of data in different classes, an asymmetric soft-margin parameter can be used to avoid biased learning.

In this paper, different SVM configurations are used to classify fault occurrence and fault severity, more specifically, multi-class classification SVMs with Gaussian Kernel as shown in (22). All parameters were selected via a stratified cross-validation analysis [49].

$$K(u, v) = u' \cdot v \quad (21)$$

$$K(u, v) = \exp(-\gamma|u - v|^2) \quad (22)$$

### 3. Experimental setup and methodology

To evaluate the proposed strategy, a large experimental dataset was obtained from healthy and faulty machines to achieve different operating conditions, since load, fault, and supply frequency influence the majority of diagnostic systems.

The test bench allows the connection of TIMs (1.0 HP power, 220 V/380 V supply voltage, 4 poles, 60 Hz, 34 rotor bars, and 4.1 Nm rated torque) to a direct-current (DC) generator (2 kW power, 190 V rated field voltage, and 250 V rated armature voltage) that works as a load. Four healthy machines were used in the experiments:

- Normal machine with rotor interchangeability: machine with healthy rotor, which was also adapted to receive faulty rotors in order to simulate faults;
- Factory machine: healthy machine without any kind of intervention;
- High-performance machine: healthy machine with same nameplate information but with different construction materials and;
- Normal machine: machine with healthy rotor.

Four different faulty rotors were used in the experiments: one broken bar (1 bb) (2.9% of the total number of bars), two adjacent broken bars (2 bb) (5.9% of the total number of bars), 4 symmetrical broken bars ( $2 \times 2$  bb) (11.8% of the total number of bars), consisting of two pairs of adjacent broken bars in diametrically opposed positions, and four adjacent broken bars (4 bb) (11.8% of the total number of bars).

The analysis of several different machines is important because it makes the classifier robust to real situations in the industry. Although the machines and rotors are similar in nameplate information, circuit parameters change for each one of them.

Therefore, classifiers are trained to be robust to some parameter variations, as they should be in an industrial situation, where machines are modified and reassembled because of coil rewindings, bearing substitutions and parts replacement.

Stator current and supply voltage were acquired in steady-state condition for 3 s in order to constitute the datasets, as shown in Fig. 4. This acquisition time is sufficient to calculate the  $R$  indices correctly. Experiments were carried out for two study cases: direct-driven machines and inverter-driven machines. For the direct-driven dataset, experiments were running at 60 Hz, and each instance was obtained from a ‘machine + load’ combination, with load torques being 0.5 Nm, 1.0 Nm, 2.0 Nm, 3.0 Nm, 3.5 Nm, and 4.0 Nm. For the inverter-driven dataset, experiments were

|                                  | condition               | motor            | load (Nm) |                                    | condition               | motor            | load (Nm) |
|----------------------------------|-------------------------|------------------|-----------|------------------------------------|-------------------------|------------------|-----------|
|                                  |                         |                  |           |                                    |                         |                  |           |
|                                  |                         |                  |           |                                    |                         |                  |           |
| Direct-driven database instances | healthy instances       | normal           | 0.5       | Inverter-driven database instances | healthy instances 40 Hz | normal           | 0.5       |
|                                  |                         |                  | 1.0       |                                    |                         |                  | 1.0       |
|                                  |                         |                  | 4.0       |                                    |                         |                  | 4.0       |
|                                  |                         | high performance | 0.5       |                                    |                         | high performance | 0.5       |
|                                  |                         |                  | 1.0       |                                    |                         |                  | 1.0       |
|                                  |                         |                  | 4.0       |                                    |                         |                  | 4.0       |
|                                  | faulty instances        | 1 bb rotor       | 0.5       |                                    | faulty instances 40 Hz  | 1 bb rotor       | 0.5       |
|                                  |                         |                  | 1.0       |                                    |                         |                  | 1.0       |
|                                  |                         |                  | 4.0       |                                    |                         |                  | 4.0       |
|                                  |                         | 4 bb rotor       | 0.5       |                                    |                         | 4 bb rotor       | 0.5       |
|                                  |                         |                  | 1.0       |                                    |                         |                  | 1.0       |
|                                  |                         |                  | 4.0       |                                    |                         |                  | 4.0       |
|                                  | healthy instances 60 Hz | normal           | 0.5       |                                    | healthy instances 60 Hz | normal           | 0.5       |
|                                  |                         |                  | 1.0       |                                    |                         |                  | 1.0       |
|                                  |                         |                  | 4.0       |                                    |                         |                  | 4.0       |
|                                  |                         | high performance | 0.5       |                                    |                         | high performance | 0.5       |
|                                  |                         |                  | 1.0       |                                    |                         |                  | 1.0       |
|                                  |                         |                  | 4.0       |                                    |                         |                  | 4.0       |
|                                  | faulty instances 60 Hz  | 1 bb rotor       | 0.5       |                                    | faulty instances 60 Hz  | 1 bb rotor       | 0.5       |
|                                  |                         |                  | 1.0       |                                    |                         |                  | 1.0       |
|                                  |                         |                  | 4.0       |                                    |                         |                  | 4.0       |
|                                  |                         | 4 bb rotor       | 0.5       |                                    |                         | 4 bb rotor       | 0.5       |
|                                  |                         |                  | 1.0       |                                    |                         |                  | 1.0       |
|                                  |                         |                  | 4.0       |                                    |                         |                  | 4.0       |

Fig. 4. Database instances are formed by 3 s of acquired signals in steady state condition.

running at 40 Hz, 45 Hz, 50 Hz, 40 Hz, 45 Hz, 50 Hz, and 55 Hz and each instance was obtained from a ‘machine+supply frequency+load’ ‘machine + supply frequency + load’ combination, with load torques also being 0.5 Nm, 1.0 Nm, 2.0 Nm, 3.0 Nm, 3.5 Nm, 0.5 Nm, 1.0 Nm, 2.0 Nm, 3.0 Nm, 3.5 Nm, and 4.0 Nm. The inverter was controlled by V/Hz algorithm. Fig. 5 shows the laboratory workbench.

Acquisition frequency for the direct-driven dataset was 3.84 kHz and 100 kHz for the inverter-driven dataset. The voltage signals of inverter-driven motors were initially filtered before the calculations of the OCD took place. This initial filtering is necessary since the voltages from the inverter are switched signals. In order to process signals with the OCD technique, voltages need to be continuous and periodic. Therefore, a low-pass finite impulse response (FIR) filter with cut-off frequency of 300 Hz was applied to voltage signals when machines were driven by the inverter. This low-pass filtering stage can also be accomplished by passive linear filters. This low-pass filtering stage can also be accomplished by passive linear filters.

The OCD technique was then applied to the signals on both datasets, resulting in 64 time-domain functions. Then, the  $R$  indices on a 3 s window frame were calculated for each orthogonal component, producing 64 attributes. Attributes on both datasets were normalized between 0 and 1. SVMs were then used to classify normal and faulty motors as well as the fault severity.

The performance metric for the diagnostic system was obtained by a 10-fold stratified cross-validation classification repeated 10



Fig. 5. Laboratory workbench with the 1-hp induction machine and data acquisition system.

times in different seeds. Therefore, training data was not used as validation data. Stratified cross-validation ensures that each fold has the right proportion of instances of each class, preventing biased classification. The classes were considered as healthy, one broken bar, two broken bars, and four broken bars. Additionally, different seed repetitions are used to improve upon a repeated holdout, by reducing the variance of the classification estimate, and to prevent the overfitting of the classifiers.

#### 4. Results and discussions

In this study, three-phase 1-hp squirrel-cage induction machines were tested for fault diagnosis purposes. Accordingly, normal machines and rotors with 1 bb, 2 bb,  $2 \times 2$  bb, and 4 bb were tested. In Fig. 6, the RMS values of the orthogonal component  $i_a^p(t)$  is shown with respect to time in different operating conditions for direct-driven motors.

An initial statistical test was made in the direct-driven dataset to evaluate if healthy machines are distinguishable from faulty machines with respect to the 64 attributes of the dataset, which are the Rindices calculated in a 3 s window frame. A Kruskal-Wallis nonparametric one-way analysis of variance (ANOVA) cast doubt on the associated null hypothesis, which suggests that all instances of the dataset are from a same class with a chi-square static of 1070.86 and 5% of significance level. This result shows that there are noticeable differences between faulty and healthy machines with respect to the 64 attributes.

The OCD products were found to be relevant features for classifying normal and faulty machines. Table 1 shows the classification hit-rates for the direct-driven case.

Moreover, the performances of SVM classifiers with a reduced number of inputs are also shown. The selected features were obtained from the CFS algorithm and then used as inputs to the SVM classifiers.

It is observed that very simple classifier models are capable of outputting high-accuracy classification. This fact indicates that the classification problem is separable, since a SVM with linear kernel and low  $C$  parameter, which approximates the classifier to a hard-margin SVM, performs with an accuracy of 100%.

Table 1 also shows the severity classification hit-rates for the direct-driven case.

It is important to point out that  $2 \times 2$  bb and 4 bb configurations are considered the same for severity classification (4 bb),

**Table 1**

Accuracy of fault and severity classification for normal and faulty machines (1 bb, 2 bb,  $2 \times 2$  bb and 4 bb). Direct-driven at 60 Hz and eight different load torques. Database with 45 normal machines and 26 faulty machines.

| Type     | Classifier       | N <sup>o</sup> of Inputs | Accuracy (%) mean (std deviation) |
|----------|------------------|--------------------------|-----------------------------------|
| Fault    | SVM <sup>1</sup> | 64                       | 100.0 (0.0)                       |
| Fault    | SVM <sup>2</sup> | 10                       | 100.0 (0.0)                       |
| Severity | SVM <sup>3</sup> | 64                       | 100.0 (0.0)                       |
| Severity | SVM <sup>4</sup> | 10                       | 99.46 (1.97)                      |
| Severity | SVM <sup>5</sup> | 10                       | 100.0 (0.0)                       |

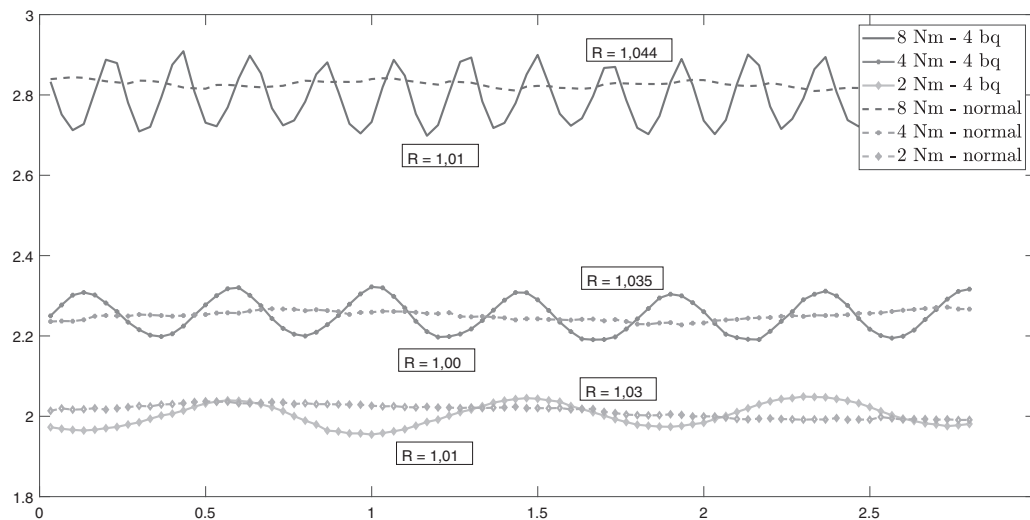
SVM<sup>1</sup> = SVM<sup>2</sup> = SVM<sup>3</sup> = SVM<sup>4</sup> ( $C = 1$ , linear kernel). SVM<sup>5</sup> ( $C = 50$ ,  $\gamma = 1.0$ , Gaussian Kernel).

where the number of broken bars is understood as the severity of the fault. In [50] the effects of adjacent and non-adjacent bar breakages on fault diagnosis were studied and it is known that geometric location of broken bars can affect the classification task. In the case were  $N_b$  non-adjacent broken bars are separated by one pole pitch, this is, by  $180^\circ$  electrical, fault-specific characteristics corresponds to  $N_b$  adjacent bar breakages. Hence, no difficulty should arise in diagnosing the presence of 4 broken bars under these distinct situations.

It is also important to address that attribute selection algorithms can increase classifier performance because of irrelevant and redundant attribute elimination. For a smaller number of inputs, accuracy has not changed significantly and an SVM with higher margin constraints and Gaussian Kernel could be used for the classification. A natural dimensionality reduction could be obtained by using data from only one phase of the motor, which could reduce the number of sensors used and the cost of the diagnostic method. In Table 2, classification performance is shown when Rindices from just phase A are used as inputs to classifiers.

Regarding the uncertainty of the classifiers, the standard deviation measures the spread of the accuracies obtained in the several classification experiments. The smaller the standard deviation, the smaller the uncertainty and thus higher the confidence in the classification and the reliability of the experiment. It can be seen that the variance of the classification is improved upon repeated experiments given that the cross-validation classification is repeated 10 times in different seeds.

The Kruskal-Wallis ANOVA tested in the inverter-driven dataset also casts doubt on the associated null hypothesis, which suggests that all samples were from the same class with a



**Fig. 6.** RMS values and  $R$  indices of the orthogonal component  $i_a^p(t)$  for healthy motors and motors with 4 bb in different load torque conditions.

**Table 2**

Accuracy of fault and severity classification for normal and faulty machines (1 bb, 2 bb,  $2 \times 2$  bb and 4 bb). Direct-driven at 60 Hz and eight different load torques. Database with 45 normal machines and 26 faulty machines. Only indices from phase A are considered.

| Type     | Classifier       | N <sup>o</sup> of Inputs | Accuracy (%) mean (std deviation) |
|----------|------------------|--------------------------|-----------------------------------|
| Fault    | SVM <sup>1</sup> | 16                       | 100.0 (0.0)                       |
| Fault    | SVM <sup>2</sup> | 6                        | 100.0 (0.0)                       |
| Severity | SVM <sup>3</sup> | 16                       | 100.0 (0.0)                       |
| Severity | SVM <sup>4</sup> | 6                        | 100.0 (0.0)                       |

SVM<sup>1</sup> = SVM<sup>2</sup> ( $C = 50$ ,  $\gamma = 0.75$ , Gaussian Kernel).

SVM<sup>3</sup> = SVM<sup>4</sup> ( $C = 50$ ,  $\gamma = 4.0$ , Gaussian Kernel).

**Table 3**

Accuracy of fault and severity classification for normal and faulty machines (1 bb, 2 bb,  $2 \times 2$  bb and 4 bb). Inverter-driven at 40, 45, 50, and 55 Hz and eight different load torques. Database with 64 normal machines and 125 faulty machines. Only indices from phase A are considered.

| Type     | Classifier       | N <sup>o</sup> of Inputs | Accuracy (%) mean (std deviation) |
|----------|------------------|--------------------------|-----------------------------------|
| Fault    | SVM <sup>1</sup> | 16                       | 98.76 (1.02)%                     |
| Fault    | SVM <sup>2</sup> | 5                        | 98.06 (1.31)%                     |
| Severity | SVM <sup>3</sup> | 16                       | 98.76 (1.02) %                    |
| Severity | SVM <sup>4</sup> | 6                        | 98.49 (2.16) %                    |

SVM<sup>1</sup> = SVM<sup>2</sup> ( $C = 100$ ,  $\gamma = 4$ , Gaussian Kernel) SVM<sup>3</sup> = SVM<sup>4</sup> ( $C = 100$ ,  $\gamma = 8$ , Gaussian Kernel).

chi-square statistic 559.39 and 5% significance level. Similarly to the direct-driven case, this result shows that the OCD technique can characterize the fault occurrence even in inverter-driven machines. Hit rates for the inverter-driven dataset are shown in Table 3, where R indices from phase-A signals are used as inputs.

It is noticeable that high accuracies were obtained in the diagnosis of inverter-driven motors. This shows the good classification separability when the OCD technique is used for processing the data. The slight deterioration in comparison with the classification performances in the direct-driven dataset can be explained by the noise induced in the acquisition system. Although the OCD has a filtering step, current clamps were positioned inside the inverter panel, and this may have impaired the classification. A less noisy location could have enhanced the results.

As already discussed, the diagnosis of broken rotor bars is difficult when the machine is lightly loaded. As seen in Section 2.2, RMS values of orthogonal components oscillate when a fault occurs. The frequency and amplitude of these oscillations depend on the load. However, leakage is not present and broken bar detection and identification is still possible from the  $R$  index, which is proportional to the number of broken bars. Moreover, the correlation to the fault of the  $R$  indices is robust to load condition as showed in Fig. 3. For a more in-depth investigation, the global direct-driven dataset was separated into 2 smaller datasets. The first dataset contains data from no-load torque condition to less than 40% of the rated-load torque. The second dataset contains data of 75%–100% of the rated-load torque. These newly formed datasets were tested again with SVM classifiers, as shown in Table 4.

**Table 4**

Accuracy of fault classification for normal and faulty machines (1 bb, 2 bb,  $2 \times 2$  bb and 4 bb). Direct driven at 60 Hz for less than 40% rated torque dataset and for more than 75% rated torque dataset. Only indices from phase A are considered.

| Classifier  | N <sup>o</sup> of Inputs | Accuracy (%) mean (std deviation) |
|-------------|--------------------------|-----------------------------------|
| SVM – < 40% | 16                       | 99.34 (2.13)%                     |
| SVM – < 40% | 16                       | 99.34 (2.11)%                     |
| SVM – < 40% | 5                        | 100.0 (0.0)%                      |
| SVM – > 75% | 5                        | 100.0 (0.0)%                      |

SVM ( $C = 50$ ,  $\gamma = 0.01$ ).

It can be observed that the different levels of load torque did not result in a severe accuracy loss. A two-tailed paired T-test confirms that there are no differences between accuracies in the low torque and high torque condition, with a 5% significance level. This result confirms that by using the machine own signals as a basis for the decomposition in the time domain, even small effects of a broken bar can be perceived. One can also observe that although no significant changes in accuracy have occurred, the SVMs with 16 inputs have higher standard deviation, meaning that some classifiers benefit from the dimensionality reduction.

It is important to stress that other methods could undertake the diagnostic process after the voltage and current signals were processed with the OCD, such as statistical methods: average and kurtosis analysis; or other intelligent systems. As already explained, the choice of using SVMs was based on the fact that these classifiers have good performance in high-dimensional input spaces, what ensures a fair comparison of performances between a large set of attributes and reduced set of attributes selected from feature selection algorithms.

## 5. Conclusions

This paper presents a fault diagnostic method for broken rotor bars in three-phase induction machines based on the OCD. The decomposition products are closely linked to the system operating condition and are heuristically related to electric parameters and power consumption of the system. It is shown that the products of the decomposition are a good set of fault-specific features.

Experimental tests were used to demonstrate the method effectiveness. Supply voltage and stator current of healthy and faulty machines were monitored and then processed with the Orthogonal Component Decomposition. Support Vector Machines, which were optimized via cross-validation, were able to characterize and classify faulty and healthy motors regardless of the load torque.

Fault severity, represented by the number of broken rotor bars, was also identified. Feature selection algorithms show that the products from the decomposition are strongly correlated to a fault occurrence and to the fault severity, but they are also redundant. This novel method should be used together with feature selection algorithms so not to increase computational burden. During the training phase, all 64 components are calculate, however, after selecting the best features, during the use of the classifier, only few components can be calculated. Therefore, with the feature selection, fewer attributes can be used as inputs to the diagnostic system with little change in accuracy.

Overall, this method does not depend on frequency analysis, spectral resolution or the number of frequency data points considered around the supply frequency.

The OCD products are very sensitive to the system operation. This application demonstrates and supports that the novel approach for signal decomposition can be employed in the study of nonlinear events and fault detections on other electromechanical systems. Future studies are then encouraged so that the technique becomes consolidated as a signal processing tool sensitive to other events, such as those related to incipient faults and to the differentiation of types of faults, such as the case of oscillating loads and broken bars.

## References

- [1] M.B. Salles, K. Hameyer, J.R. Cardoso, A.P. Grilo, C. Rahmann, Crowbar system in doubly fed induction wind generators, *Energies* 3 (4) (2010) 738–753.
- [2] A. Ceban, R. Pusca, R. Romary, Study of rotor faults in induction motors using external magnetic field analysis, *IEEE Trans. Industr. Electron.* 59 (5) (2012) 2082–2093.

- [3] M.R. Mehrjou, N. Mariun, M.H. Marhaban, N. Misron, Rotor fault condition monitoring techniques for squirrel-cage induction machine – a review, *Mech. Syst. Signal Process.*, 25 (8) (2011) 2827–2848.
- [4] K. Gyftakis, D. Spyropoulos, J. Kappatou, E. Mitronikas, A novel approach for broken bar fault diagnosis in induction motors through torque monitoring, *IEEE Trans. Energy Convers.*, 28 (2) (2013) 267–277.
- [5] J. Faiz, V. Ghorbanian, B. Ebrahimi, Emd-based analysis of industrial induction motors with broken rotor bars for identification of operating point at different supply modes, *IEEE Trans. Industr. Inf.* 10 (2) (2014) 957–966.
- [6] J. Antonino-Daviu, S. Aviyente, E. Strangas, M. Riera-Guasp, Scale invariant feature extraction algorithm for the automatic diagnosis of rotor asymmetries in induction motors, *IEEE Trans. Industr. Inf.* 9 (1) (2013) 100–108.
- [7] T. Ghanbari, Autocorrelation function-based technique for stator turn-fault detection of induction motor, *IET Sci. Meas. Technol.* 10 (2) (2016) 100–110, <https://doi.org/10.1049/iet-smt.2015.0118>.
- [8] X. Dai, Z. Gao, From model, signal to knowledge: a data-driven perspective of fault detection and diagnosis, *IEEE Trans. Industr. Inf.* 9 (4) (2013) 2226–2238.
- [9] G. Didier, E. Ternisien, O. Caspary, H. Razik, A new approach to detect broken rotor bars in induction machines by current spectrum analysis, *Mech. Syst. Signal Process.* 21 (2) (2007) 1127–1142.
- [10] A. da Silva, R. Pavinelli, N. Demerdash, Induction machine broken bar and stator short-circuit fault diagnostics based on three-phase stator current envelopes, *IEEE Trans. Industr. Electron.* 55 (3) (2008) 1310–1318.
- [11] J. Pons-Llinares, J. Antonino-Daviu, M. Riera-Guasp, M. Pineda-Sanchez, V. Climente-Alarcon, Induction motor diagnosis based on a transient current analytic wavelet transform via frequency B-splines, *IEEE Trans. Industr. Electron.* 58 (5) (2011) 1530–1544.
- [12] M. Pineda-Sanchez, M. Riera-Guasp, J. Roger-Folch, J. Antonino-Daviu, J. Perez-Cruz, R. Puche-Panadero, Diagnosis of induction motor faults in time-varying conditions using the polynomial-phase transform of the current, *IEEE Trans. Industr. Electron.* 58 (4) (2011) 1428–1439.
- [13] J. de Jesus Rangel-Magdaleno, H. Peregrina-Barreto, J. Ramirez-Cortes, P. Gomez-Gil, R. Morales-Caporal, FPGA-based broken bars detection on induction motors under different load using motor current signature analysis and mathematical morphology, *IEEE Trans. Instrum. Meas.* 63 (5) (2014) 1032–1040.
- [14] J. Rangel-Magdaleno, H. Peregrina-Barreto, J. Ramirez-Cortes, I. Cruz-Vega, Hilbert spectrum analysis of induction motors for the detection of incipient broken rotor bars, *Measurement* 109 (2017) 247–255, <https://doi.org/10.1016/j.measurement.2017.05.070>.
- [15] G. Bossio, C. De Angelo, J. Bossio, C. Pezzani, G. Garcia, Separating broken rotor bars and load oscillations on IM fault diagnosis through the instantaneous active and reactive currents, *IEEE Trans. Industr. Electron.* 56 (11) (2009) 4571–4580.
- [16] M. Drif, A. Cardoso, The use of the instantaneous-reactive-power signature analysis for rotor-cage-fault diagnostics in three-phase induction motors, *IEEE Trans. Industr. Electron.* 56 (11) (2009) 4606–4614.
- [17] M. Drif, A. Cardoso, Discriminating the simultaneous occurrence of three-phase induction motor rotor faults and mechanical load oscillations by the instantaneous active and reactive power media signature analyses, *IEEE Trans. Industr. Electron.* 59 (3) (2012) 1630–1639.
- [18] A. da Silva, R. Pavinelli, N. Demerdash, Rotor bar fault monitoring method based on analysis of air-gap torques of induction motors, *IEEE Trans. Industr. Inf.* 9 (4) (2013) 2274–2283.
- [19] V. Climente-Alarcon, J. Antonino-Daviu, F. Vedreno-Santos, R. Puche-Panadero, Vibration transient detection of broken rotor bars by PSH sidebands, *IEEE Trans. Ind. Appl.* 49 (6) (2013) 2576–2582.
- [20] M. Cocconcioni, B. Lazzerini, S. Volpi, Robust diagnosis of rolling element bearings based on classification techniques, *IEEE Trans. Industr. Inf.* 9 (4) (2013) 2256–2263.
- [21] D. Camarena-Martinez, C.A. Perez-Ramirez, M. Valtierra-Rodriguez, J.P. Amezcua-Sanchez, R. de Jesus Romero-Troncoso, Synchrosqueezing transform-based methodology for broken rotor bars detection in induction motors, *Measurement* 90 (2016) 519–525, <https://doi.org/10.1016/j.measurement.2016.05.010>.
- [22] B.M. Ebrahimi, J. Faiz, S. Lotfi-fard, P. Pillay, Novel indices for broken rotor bars fault diagnosis in induction motors using wavelet transform, *Mech. Syst. Signal Process.* 30 (2012) 131–145.
- [23] C. da Costa, M. Kashiwagi, M.H. Mathias, Rotor failure detection of induction motors by wavelet transform and fourier transform in non-stationary condition, *Case Stud. Mech. Syst. Signal Process.* 1 (2015) 15–26.
- [24] T. Yang, H. Pen, Z. Wang, C.S. Chang, Feature knowledge based fault detection of induction motors through the analysis of stator current data, *IEEE Trans. Instrum. Meas.* 65 (3) (2016) 549–558, <https://doi.org/10.1109/TIM.2015.2498978>.
- [25] A. Garcia-Perez, R. de Jesus Romero-Troncoso, E. Cabal-Yepez, R. Osornio-Rios, The application of high-resolution spectral analysis for identifying multiple combined faults in induction motors, *IEEE Trans. Industr. Electron.* 58 (5) (2011) 2002–2010.
- [26] F. Gu, T. Wang, A. Alwodai, X. Tian, Y. Shao, A. Ball, A new method of accurate broken rotor bar diagnosis based on modulation signal bispectrum analysis of motor current signals, *Mech. Syst. Signal Process.* 50–51 (2015) 400–413.
- [27] J.R. Rivera-Guillen, J.J. De Santiago-Perez, J.P. Amezcua-Sanchez, M. Valtierra-Rodriguez, R.J. Romero-Troncoso, Enhanced FFT-based method for incipient broken rotor bar detection in induction motors during the startup transient, *Measurement* 124 (2018) 277–285, <https://doi.org/10.1016/j.measurement.2018.04.039>.
- [28] G. Joksimovic, J. Riger, T. Wolbank, N. Peric, M. Vasak, Stator-current spectrum signature of healthy cage rotor induction machines, *IEEE Trans. Instrum. Meas.* 60 (9) (2013) 4025–4033.
- [29] C. Concar, G. Franceschini, C. Tassoni, A. Toscani, Validation of a faulted rotor induction machine model with an insightful geometrical interpretation of physical quantities, *IEEE Trans. Industr. Electron.* 60 (9) (2013) 4074–4083.
- [30] P. Shi, Z. Chen, Y. Vagapov, Z. Zouaoui, A new diagnosis of broken rotor bar fault extent in three phase squirrel cage induction motor, *Mech. Syst. Signal Process.* 42 (1–2) (2014) 388–403.
- [31] V. Ghorbanian, J. Faiz, A survey on time and frequency characteristics of induction motors with broken rotor bars in line-start and inverter-fed modes, *Mech. Syst. Signal Process.* 54–55 (2015) 427–456.
- [32] S. Bindu, V. Thomas, Diagnoses of internal faults of three phase squirrel cage induction motor – a review, in: *Proc. 2014 Int. Conf. Advances in Energy Conversion Technol. (ICAECT'04)*, 2014, pp. 48–54.
- [33] A. Sadeghian, Z. Ye, B. Wu, Online detection of broken rotor bars in induction motors by wavelet packet decomposition and artificial neural networks, *IEEE Trans. Instrum. Meas.* 58 (7) (2009) 2253–2263.
- [34] W. Sun, S. Shao, R. Zhao, R. Yan, X. Zhang, X. Chen, A sparse auto-encoder-based deep neural network approach for induction motor faults classification, *Measurement* 89 (2016) 171–178, <https://doi.org/10.1016/j.measurement.2016.04.007>.
- [35] R. Romero-Troncoso, R. Saucedo-Gallaga, E. Cabal-Yepez, A. Garcia-Perez, R. Osornio-Rios, R. Alvarez-Salas, H. Miranda-Vidales, N. Huber, FPGA-based online detection of multiple combined faults in induction motors through information entropy and fuzzy inference, *IEEE Trans. Industr. Electron.* 58 (11) (2011) 5263–5270.
- [36] M. Seera, C.P. Lim, S. Nahavandi, C.K. Loo, Condition monitoring of induction motors: a review and an application of an ensemble of hybrid intelligent models, *Expert Syst. Appl.* 41 (10) (2014) 4891–4903.
- [37] D. Matic, F. Kulić, M. Pineda-Sánchez, I. Kamenko, Support vector machine classifier for diagnosis in electrical machines: application to broken bar, *Expert Syst. Appl.* 39 (10) (2012) 8681–8689.
- [38] A. Soualhi, G. Clerc, H. Razik, Detection and diagnosis of faults in induction motor using an improved artificial ant clustering technique, *IEEE Trans. Industr. Electron.* 60 (9) (2013) 4053–4062.
- [39] H. Razik, M. de Rossiter Correa, E. da Silva, A novel monitoring of load level and broken bar fault severity applied to squirrel-cage induction motors using a genetic algorithm, *IEEE Trans. Industr. Electron.* 56 (11) (2009) 4615–4626.
- [40] O.E. Batista, R.A. Flauzino, M.A. de Araujo, L.A. de Moraes, I.N. da Silva, Methodology for information extraction from oscillograms and its application for high-impedance faults analysis, *Int. J. Electr. Power Energy Syst.* 76 (2016) 23–34.
- [41] L.H. Liboni, R.A. Flauzino, I.N. da Silva, E.C.M. Costa, M. Suetake, Efficient signal processing technique for information extraction and its applications in power systems, *Electr. Power Syst. Res.* 141 (2016) 538–548, <https://doi.org/10.1016/j.epsr.2016.08.019>, URL: <http://www.sciencedirect.com/science/article/pii/S0378779616303200>.
- [42] R.A. Flauzino, Identification and location of high-impedance faults in power distribution systems based on orthogonal component decomposition and fuzzy inference (in portuguese) (Ph.D. thesis), SEL/EESC/ University of São Paulo, 2007.
- [43] G.R. Bossio, C.H.D. Angelo, J.M. Bossio, C.M. Pezzani, G.O. Garcia, Separating broken rotor bars and load oscillations on im fault diagnosis through the instantaneous active and reactive currents, *IEEE Trans. Industr. Electron.* 56 (11) (2009) 4571–4580, <https://doi.org/10.1109/TIE.2009.2024656>.
- [44] M. Drif, A.J.M. Cardoso, Stator fault diagnostics in squirrel cage three-phase induction motor drives using the instantaneous active and reactive power signature analyses, *IEEE Trans. Industr. Inf.* 10 (2) (2014) 1348–1360, <https://doi.org/10.1109/TII.2014.2307013>.
- [45] C. Martins Cunha, R. Lyra, B. Filho, Simulation and analysis of induction machines with rotor asymmetries, *IEEE Trans. Ind. Appl.* 41 (1) (2005) 18–24.
- [46] M.A. Hall, Correlation-based feature selection for discrete and numeric class machine learning, in: *Proc. 17th International Conference on Machine Learning*, Morgan Kaufmann, 2000, pp. 359–366.
- [47] P. Bilski, Application of Support Vector Machines to the induction motor parameters identification, *Measurement* 51 (2014) 377–386, <https://doi.org/10.1016/j.measurement.2013.12.013>, 20th World Congress of the International Measurement-Confederation (IMEKO) – Metrology for Green Growth, Busan, South Korea, 2012.
- [48] C. Campbell, Y. Ying, Learning with Support Vector Machines, G – Reference, Information and Interdisciplinary Subjects Series, Morgan & Claypool, 2011.
- [49] I. Witten, E. Frank, Data mining: practical machine learning tools and techniques, second ed., The Morgan Kaufmann Series in Data Management Systems, Elsevier Science, 2005.
- [50] G. Sizov, A. Sayed-Ahmed, C.-C. Yeh, N. Demerdash, Analysis and diagnostics of adjacent and nonadjacent broken-rotor-bar faults in squirrel-cage induction machines, *IEEE Trans. Industr. Electron.* 56 (11) (2009) 4627–4641.