Principal Component Analysis Anomaly Detector for Rotor Broken Bars

Mohammed Obaid Mustafa, George Georgoulas, and George Nikolakopoulos

Abstract—In this article a method for the detection of broken rotor bars in asynchronous machines operating under full load is presented. Unlike most Motor Current Signature Analysis (MCSA) approaches, which operate in the frequency domain, our method operates in the time domain. The scheme is based on the use of a Principal Component Analysis (PCA) fault/anomaly detector. PCA is applied on the three stator currents to subsequently calculate the $\mathcal Q$ statistic which is employed for detecting the presence/absence of a fault. The efficiency of the proposed scheme was experimentally evaluated using different fault severity levels, ranging from 1/4 of a broken bar to three broken bars. The obtained results indicate that the method can detect the caused asymmetry with a very restricted amount of data.

I. Introduction

The main goal of fault detection and isolation is to effectively detect faults and accurately identify the failed component in the shortest time possible. Fault detection leads to reduction in maintenance costs and prevents unscheduled downtime, which results in an increased system availability. Realising that many faults can be detected in an early stage led to a gradual shift from the traditional Time Based Maintenance (TBM) and Corrective Maintenance (CV) to Condition Based Maintenance (CBM). CBM was embraced by the engineering community since it gives engineers the ability to perform more effective maintenance and many approaches have been proposed especially for the monitoring of electric drives.

Among the various machines engaged in electric drives, asynchronous machines play a prominent role in industrial settings due to their advanced safety, efficiency, reliability, and performance [1], [2]. As all elecro-mechanical systems, asynchronous machines can also break down following the occurrence of a fault that eventually culminates to a failure. These faults are mainly categorized as stator, rotor and bearings faults. This work focuses on rotor faults and more specifically on the detection of broken rotor bars.

One of the reasons for broken rotor bar failures is that large starting currents in an induction motor occur when cooling is at a minimum and this results in thermal and mechanical stresses being at a maximum [3]. Other factors that can cause the breakage of a rotor bar are thermal, magnetic, residual, dynamic, environmental, and stresses [4], [5].

Emails: mohoba@ltu.se, geonik@ltu.se, and georgoul@ kic.teiep.gr Corresponding Author's Email: geonik@ltu.se

Rotor faults are not the most frequent faults occurring in induction machines [5] but have received the most attention in the literature regarding the proposed techniques for their detection. One of the reasons is that rotor cage faults can lead to shaft vibration and thus bearing failures and air gap eccentricity faults, etc. [6].

The rotor bars can be partially or completely broken during the operation of a squirrel cage induction motor. Once a bar breaks the condition of the neighbouring bars also deteriorates progressively due to the increased stresses. To prevent such accumulative destructive process, the problem should be detected as early as possible, that is, when the bars are beginning to break [7].

In the related scientific literature, various methods have been proposed for the detection of broken rotor bars, relying on a variety of monitored quantities such as flux [8], torque [9], vibrations [10], [11], voltages [12] and power [13], [14]. However the most popular methods are those relying on monitoring of the current, usually referred to as Motor Current Signature Analysis (MCSA) methods. MCSA methods usually operate in the frequency domain trying to isolate characteristic components that are generated by the asymmetry caused by the broken bars.

One of the potential limitations of these methods is that they require the collection of quite long data segments for accurate detection of the faulty components. Moreover they are not able to to work under no-load conditiona [15]. Lately another set of methods, which operates during start—up (Transient MCSA) have been proposed as a supplement to the more conventional MCSA [16], [17], [18].

On the other hand a fault detection and diagnosis based on uncertainty bounds violation has been efficiently utilized for broken rotor bar of induction motor [19], [20], [21].

The novelty of this article stems from the establishment of a method based on the application of Principal Component Analysis (PCA) for the detection of broken bars. Although the majority of the current MCSA methods operate in the frequency domain, the proposed method operates in time domain and it is capable of detecting rotor asymmetries using only a small sample of data. The results from the experimental verification of the proposed scheme are very promising, indicating the potential use of the method for real time implementations.

The rest of the paper is structured as follows. Section 2 presents a general introduction to anomaly detection and more specific guidelines for the use of PCA within the anomaly detection framework of induction machines. In Section 3 the experimental set—up is described, while the results of the

M. O. Mustafa 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

G. Georgoulas is with the Laboratory of Knowledge and Intelligent Computing (KIC) Department of Computer Engineering, Technological Educational Institute of Epirus, Arta, Greece

proposed method are analysed in Section 4. Finally, Section 5 concludes the paper also offering some insight for future work and future extensions of the method.

II. FAULT/ANOMALY DETECTION USING PCA

Fault/anomaly detection is the simplest but usually the first, fundamental step, in a condition monitoring system. In other words, if an error occurs, first it should be detected if there is something wrong (faulty) and then try to identify which fault has occurred [16].

In general, anomaly detectors try to detect a deviation from normality and exploit the available information coming from normal operating conditions. These detectors are also called novelty detectors [22], [23], while the term anomaly detector is reserved for the case that the new data deviate from a well-defined notion of normal behaviour and the process is referred to as anomaly detection [24].

There are a number of different methods that can be applied on the anomaly detection problem (boundary methods, density methods, reconstruction methods etc.). In this work we employ probably the simplest of the reconstruction methods, the PCA anomaly detector method.

A. PCA

PCA is the most widely used method for dimensionality reduction [25], [26]. It was introduced over a hundred years ago by Karl Peterson and up to now it has found many applications in multivariate data analysis [27].

PCA extracts a set of orthogonal vectors called principal components (also called loading vectors) which are usually sorted in such a way such as the first principal component accounts for most of the variance in the data, the second principal component for most of the remaining variance and etc. PCA is summarised in the following table (Note: It is interesting that the same algorithm for PCA can be derived following different paths, e.g. the maximum variance formulation or the minimum error formulation):

TABLE I
PRINCIPAL COMPONENTS EXTRACTION ALGORITHM

$$\bar{x}_j = \frac{1}{N} \sum_{i=1}^{N} x_{ij}$$

STEP 2 Subtract this mean from the original variable:

$$\acute{x}_{ij} = x_{ij} - \overline{x}_j$$
, $i = 1, \dots, N$ and $j = 1, \dots, d$

STEP 3 Calculate the covariance matrix of the zero mean data matrix whose elements are given by:

$$S_{nm} = \acute{x}_{ij} \acute{y}_{ij}, n, m = 1, ..., d$$

STEP 4 Calculate the eigenvalues and its corresponding eigenvectors of the covariance matrix S

B. PCA for anomaly Detection

The utilization of PCA for anomaly detection is based on the assumption that PCA can separate the observed *N*-dimensional space (input) into two subspaces: a) a subspace spanned by the

first M principal components, which captures the systematic trends of the process/system and b) a subspace spanned by the remaining N-M principal components that capture essentially the noise component [28]. The latter is usually called the residual space. In case of a fault it is assumed that a deviation will take place, which will be manifested in the residual space.

Therefore the residual space can be monitored using the Q statistic (which point-wise is an indication of the distance of the specific data point from the principal component plane [29]) given by:

$$Q = r^T r, (1)$$

with

$$r = (I - VV^T)x, (2)$$

where x is the input vector and r is the residual vector, which is being derived by projecting the input vector x onto the residual space. Moreover, V is an $N \times M$ matrix whose columns correspond to the principal components associated with the M largest eigenvalues. By utilising a threshold for Q, deviations from the normal operation can be detected.

C. PCA for Detection of broken bars in asynchronous machines

In the case of asynchronous machines, the multiple inputs correspond to the three stator currents. In the healthy case as it can be easily observed in Figure 1, for a perfect machine and also by assuming that we have a perfect power supply, two principal components capture the systematic trends (rotation) of the system leaving a one dimensional residual space. Therefore it comes natural to select the number of retaining principal components equal to two.

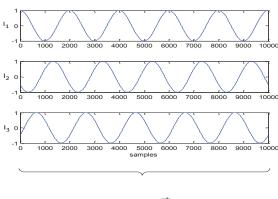
In our implementation we therefore used the first two principal components to model the normal operation and we used the averaged Q value over a predefined window for robustness as it will be shown in Section 4, even though the online version could also be applied.

III. EXPERIMENTAL SETUP

The experimental setup consisted of a three phase induction motor, a DC generator acting as load and a Data Acquisition (DAQ) instrumentation. The overall setup is depicted in Figure 2.

The Portable DAQ device from National Instruments (NI) was used to measure and collect the three phase stator currents under full load current conditions. The rated values of DAQ are 5A r.m.s current and 400V r.m.s. voltage. The data acquisition routines were developed using LabView. The characteristics of the motor under study were: 4 poles, 1.1Kw, 380V, 50 Hz, rated power, 1410 rated speed, and 2.75A rated current.

For the broken rotor bar cases, the motor had been artificially damaged by drilling a part of a rotor's bar. Partially broken rotor bar cases were produced by drilling 4 mm and then 8 mm out of 17 mm in the same rotor bar as presented in Figure 3 in a side view. The drilling has not been performed



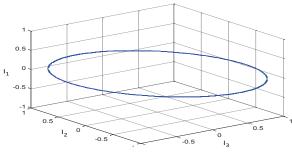


Fig. 1. Ideal normal operation. The three currents depicted separately and as part of the multivariate input.

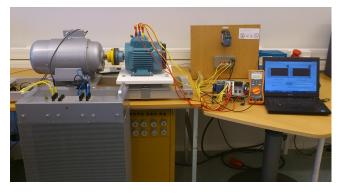


Fig. 2. Experimental setup for evaluating the broken bar fault detection

completely, from one side to the other (17mm), since the aim was to evaluate the proposed scheme in a more difficult to identify situation (partially damaged broken bar).

Additional to these partially broken bar conditions, three more cases have been considered: a one fully broken rotor bar, two broken rotor bars, and three broken rotor bars. The case with the three broken bars is presented in Figure 4 whith the "complete" three drillings.

IV. EXPERIMENTAL RESULTS

For the evaluation of the proposed fault detection scheme, several experimental runs have been performed with the considered squirrel cage induction motor and for the aforemen-



Fig. 3. Partial broken rotor bars cases for three phase asynchronous motor



Fig. 4. Fully broken rotor bar cases for three phase asynchronous motor

tioned six different conditions. Stator currents were sampled under full load at a sampling frequency of 1600Hz. In all cases the data coming from the normal operating condition were used to build the "normal" operating model.

We experimented with various signal lengths for the detection phase (from just one period to 5 periods). The following figures 5-7 we show the boxplot of the Q value averaged over three different window sizes (32, 96 and 160 samples, corresponding to 1 3 and 5 periods of the stator current respectively) and for 30 replications (30 different signal segments for each one of the six different conditions and at full load). In order to have a detection alarm we need to set a detection threshold. The standard implementation involves the calculation of a threshold based on the required confidence level as it was proposed in [30]. However this threshold in our preliminary experiments appeared too conservative. Therefore we resorted to the three-sigma rule [31], i.e. we ran an extra set of 30 replications for the normal case and we estimated the standard deviation of the Q statistic; then a fault was detected once the corresponding (mean) Q value exceeded the (estimated) mean ${\it Q}$ value for the normal cases augmented by three times the estimated standard deviation.

With this threshold, some false alarms have been also indicated. By further increasing the threshold to six (instead of three) standard deviations (six-sigma), perfect detection performance has been achieved. At this point it should be noted that by a simple eye inspection it is almost obvious that the adoption of a higher threshold could also be used. A zoomed version of Figure 7 with the "six sigma" threshold is presented in Figure 8. The overall results are presented in Tables 2 and 3 while the overal procedure is summarised as follows:

A) Model building phase:

- 1) Acquire k segments of the three currents each one of them of size N
- 2) For each segment i (i = 1..k)estimate the principal components and the corresponding mean Q_i
- 3) Estimate the mean value Q_{mean} and the standard deviation Q_{std} of the mean Q_i
- 4) Select a threshold value equal to $Q_{mean} + 6Q_{std}$

B) Detection phase

- 1) Acquire a new segment of the three currents
- 2) Use the estimated principal components to estimate the corresponding mean Q value
- 3) Compare the estimated mean Q value with the threshold. If the value exceeds the threshold then declare a fault, otherwise no fault has occurred

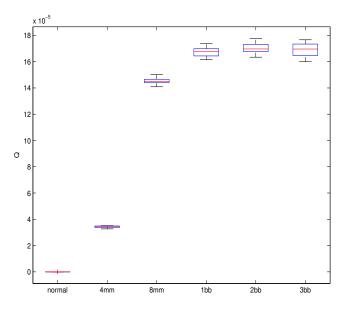


Fig. 5. Q values estimated using 32 samples (1 period)

As it can be seen the anomaly is easily identified. Moreover, by the utilization of a larger set of samples (more periods), the value becomes more consistent. On the other hand it is obvious that with the current setting no severity assessment can be performed since the feature is not monotonically increasing.

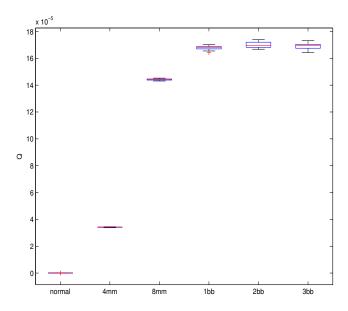


Fig. 6. Q values estimated using 96 samples (3 periods)

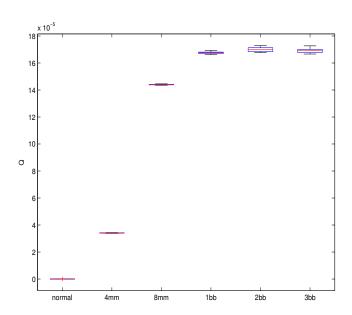


Fig. 7. Q values estimated using 160 samples (5 periods)

TABLE II
DETECTION RESULTS FOR THE THREE SIGMA THRESHOLD

	32s		96s		160s	
Time(s)	healthy	faulty	healthy	faulty	healthy	faulty
Correctly Assigned	27	150	22	150	30	150
False Alarms	3	0	8	0	0	0

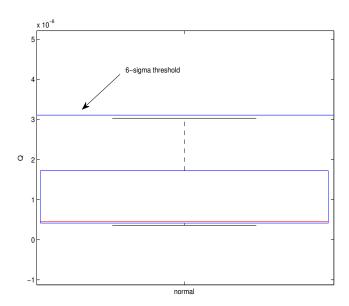


Fig. 8. Detection results for the six sigma threshold

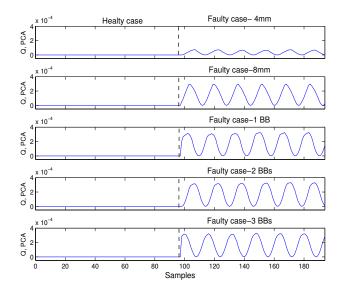


Fig. 9. The Q statistic monitored online

TABLE III
DETECTION RESULTS FOR THE SIX SIGMA THRESHOLD

Time(s)	32s		96s		160s	
	healthy	faulty	healthy	faulty	healthy	faulty
Correctly Assigned	30	150	30	150	30	150
False Alarms	0	0	0	0	0	0

Furthermore the pure online version, using each individual sample instead of the mean Q value over predefined segments, as it can be seen in Figure 9 could also be applied (the subplots of Figure 9 were created by artificially concatenating healthy with faulty data). However the fluctuation could cause some problems to the detection scheme.

For comparison reasons we also present the results of applying the periodogram spectrum estimation method for the case of 50 periods (1600 samples) Figure 10. As it can be directly observed even for such a long window size the less pronounced faults are difficult to be detected.

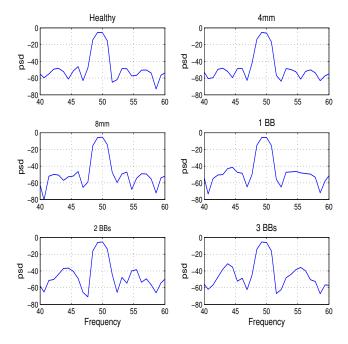


Fig. 10. Periodogram Power Spectral Density Estimate for healthy and faulty cases

V. CONCLUSIONS

In this work we presented a PCA approach to the detection of rotor broken bars under full load conditions. The method was experimentally tested and the results are promising since deviation from normality even for incipient faults can be quite easily spotted. The method in its current form cannot be used for severity assessment since the Q statistic is decreasing at some point. In order to alleviate that we will check the potential use of the T^2 statistic. A potential barrier for its acceptance from the industry, despite the very promising results, is the requirement of three current sensors. To eliminate the need for three current sensors we are currently investigating the use of other variants of PCA, the dynamic PCA approach (using in other words embedding to reconstruct the state space) which does not require the use of three sensors. Moreover since the results come from a laboratory experiment, further investigation is needed using data coming from an industrial

setting in order to assess the robustness of the method under a real life environment.

REFERENCES

- O. Ondel, E. Boutleux, and G. Clerc, "A method to detect broken bars in induction machine using pattern recognition techniques," *IEEE Transactions On Industry Applications*, vol. 42, no. 4, pp. 916–923, JULY/AUGUST 2006.
- [2] V. F. Pires, J. F. Martins, and A. J. Pires, "Eigenvector/eigenvalue analysis of a 3d current referential fault detection and diagnosis of an induction motor," *Energy Conversion and Management*, vol. 51, pp. 901–907, 2010.
- [3] L. A. G. Escudero, O. D. Perez, D. M. Sotelo, and M. P. Alonso, "Robust condition monitoring for early detection of broken rotor bars in induction motors," *Expert Systems with Applications*, vol. 38, pp. 2653–2660, Sept. 2011.
- [4] A. H. Bonnett and G. C. Soukup, "Rotor failures in squirrel cage induction motor," *IEEE Industry Applications*, vol. IA-22, no. 6, pp. 1165–1173, Nov./Dec. 1986.
- [5] A. H. Bonnett and C. Yung, "Increased efficiency versus increased reliability," *IEEE Industry Applications*, vol. 14, no. 1, pp. 29–36, 2008.
- [6] P. Zhang, Y. Du, T. G. Habetler, and B. Lu, "A survey of condition monitoring and protection methods for medium voltage induction motors," *IEEE Trans. Energy Convers*, vol. 47, no. 1, pp. 34–46, Jan. 2011.
- [7] M. R. Mehrjou, N. Mariun, M. H. Marhaban, and N. Misron, "Rotor fault condition monitoring techniques for squirrel-cage induction machine-a review," *Mechanical Systems and Signal Processing*, vol. 25, no. 8, pp. 2827–2848, Nov. 2011.
- [8] I. Rodriguez, R. Alves, and V. Guzman, "Analysis of air gap flux to detect induction motor faults," *Proceedings of the 41st International of Universities Power Engineering Conference, UPEC06.*, vol. 2, pp. 690 – 694, 2006.
- [9] A. M. d. Silva, R. J. Povinelli, and N. A. O. Demerdash, "Rotor bar fault monitoring method based on analysis of air-gap torques of induction motors," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 4, pp. 2274–2283, 2013.
- [10] A. Khezzar, M. Oumaamar, and a. H. R. M. Hadjami, M. Boucherma, "Induction motor diagnosis using line neutral voltage signatures," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 11, pp. 4581–4591, 2009.
- [11] W. Thomason and P. Orpin, "Current and vibration monitoring for fault diagnosis and root cause analysis of induction motor drives," *Proceedings of the Thirty-First Turbomachinery Symposium*, 2002.
- [12] N. M. Elkasabgy, A. R. Eastham, and G. E. Dawson, "Detection of broken bars in the cage rotor on an induction machine," *IEEE Transactions on Industrial Electronics*, vol. 28, no. 1, pp. 165–171, 1992.
- [13] Z. Liu, X. Yin, Z. Zhang, D. Chen, and W. Chen, "Online rotor mixed fault diagnosis way based on spectrum analysis of instantaneous power in squirrel cage induction motors," *IEEE Transactions on Energy Conversion*, vol. 19, no. 3, pp. 485–490, 2004.
- [14] M. H. Benbouzid, "A review of induction motors signature analysis as a medium for faults detection," *IEEE Transactions on Industrial Electronics*, vol. 47, no. 5, pp. 984–993, 2000.
- [15] K. M. Siddiqui and V. Giri, "Broken rotor bar fault detection in induction motors using wavelet transform," *International Conference on Computing, Electronics and Electrical Technologies [ICCEET]*, vol. 1, pp. 95–99, March 2012.
- [16] G. Georgoulas, M. O. Mustafa, I. P. Tsoumas, J. A. Daviu, V. C. Alarcon, C. D. Stylios, and G. Nikolakopoulos, "Principal component analysis of the start-up transient and hidden markov modeling for broken rotor bar fault diagnosis in asynchronous machines," *ELSEVIER, Expert Systems with Applications*, vol. 40, pp. 7024–7033, 2013.
- [17] J. Antonino-Daviu, M. Riera-Guasp, J. Roger-Folch, and M. Palomares, "Validation of a new method for the diagnosis of rotor bar failures via wavelet transform in industrial induction machines," *IEEE Trans. Ind. Appl.*, vol. 42, no. 4, pp. 990–996, 2006.
- [18] G. Georgoulas, I. Tsoumas, J. Antonino-Daviu, V. Climente-Alarcn, C. Stylios, E. Mitronikas, and A. Safacas, "Automatic pattern identification based on the complex empirical mode decomposition of the startup current for the diagnosis of rotor asymmetries in asynchronous machines," *IEEE Transaction on Industrial Electronics*, vol. 61, no. 9, pp. 4937 – 4946, Sep. 2013.

- [19] M. O. Mustafa, G. Nikolakopoulos, and T. Guastafsson, "Broken bars fault diagnosis based on uncertainty bounds violation for three phase induction motors," *International Transactions on Electrical Energy* Systems, Dec. 2013.
- [20] —, "Experimental evaluation of a broken rotor bar fault detection scheme based on uncertainty bounds violation," *Proceedings Of The* 39th Annual Conference of the IEEE Industrial Electronics Society ,IECON2013, pp. 5542–5546, Nov. 2013.
- [21] ——, "A fault diagnosis scheme for three phase induction motors based on uncertainty bounds," *Proceedings Of The 38th Annual Conference of* the IEEE Industrial Electronics Society ,IECON2012, pp. 1606–1611, October 2012.
- [22] M. Markou and S. Singh, "Novelty detection: a review-part 1: statistical approaches," *Signal Processing*, vol. 83, no. 12, pp. 2481–2497, 2003.
- [23] —, "Novelty detection: a review-part 2: statistical approaches," Signal Processing, vol. 83, no. 12, pp. 2499–2521, 2003b.
- [24] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: a survey," ACM Computing Surveys, vol. 41, no. 3, pp. 1–58, 2009.
- [25] J. o. W. L. Martinez, A. Martinez, Exploratory Data Analysis with MATLAB, Taylor and F. Group, Eds., 2010.
- [26] C. M. Bishop, Pattern Recognition and Machine Learning, Springer, Ed., 2006.
- [27] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, 4th ed. Academic Press, Ed., 2009.
- [28] L. H. Chiang, R. D. Braatz, and E. L. Russell, Fault detection and diagnosis in industrial systems, Springer, Ed., 2001.
- [29] C. Aldrich and L. Auret, Unsupervised Process Monitoring and Fault Diagnosis with Machine Learning Methods, Springer, Ed., 2013.
- [30] J. E. Jackson and G. S. Mudholkar, "Control procedures for residuals associated with principal component analysis," *Technometrics*, vol. 21, no. 3, pp. 341–349, 1979.
- [31] D. Ruan, K. E. G.Chen, G., and W. G., *Intelligent Data Mining: Techniques and Applications*, Springer, Ed., 2005, vol. 5.