

An overview of PST for vibration based fault diagnostics in rotating machinery

T. Haj Mohamad^{1,*} and C. Nataraj^{1,**}

¹Villanova Center for Analytics of Dynamic Systems (VCADS), Villanova University, Villanova, PA 19085, USA. <http://vcads.org>

Abstract. In general, diagnostics can be defined as the procedure of mapping the information obtained in the measurement space to the presence and magnitude of faults in the fault space. These measurements, and especially their nonlinear features, have the potential to be exploited to detect changes in dynamics due to the faults. We have been developing some interesting techniques for fault diagnostics with gratifying results. These techniques are fundamentally based on extracting appropriate features of nonlinear dynamical behavior of dynamic systems. In particular, this paper provides an overview of a technique we have developed called Phase Space Topology (PST), which has so far displayed remarkable effectiveness in unearthing faults in machinery. Applications to bearing, gear and crack diagnostics are briefly discussed.

1 Introduction

Fault diagnostics of practical systems is a very important problem that needs to be solved robustly in order to be able to make giant leaps in reliability and safety. Diagnostics is essentially an epistemological problem that require us to make intelligent inferences based on data, that could be derived from empirical observations or computer models, which are often incomplete, noisy and uncertain. Although there is a rich and varied literature, we feel that many of the diagnostic techniques in use are quite *ad hoc* and heuristic, resulting in lack of general applicability¹.

This paper presents innovative and rigorous techniques involving the nonlinear characteristics in a computational intelligence setting to diagnose changes in complex systems. Our approach consists of developing diagnostic methods using a combination of nonlinear dynamic analysis and computational intelligence techniques. In this paper, several applications are chosen with sufficient generality to be able to apply to a host of disciplines.

The theoretical approaches are validated using data from fault simulators at Villanova University and Case Western University; we also validate our algorithms using experimental data from practical machinery provided by United Technologies Research Center (UTRC, USA) and Federal University of Uberlândia (Brazil).

*e-mail: tajmoha@villanova.edu

**e-mail: nataraj@villanova.edu

¹We wish to acknowledge the excellent contributions by our esteemed colleagues; we have intentionally omitted references to their work in order to keep the paper short and focused on outlining our approach; kindly refer to the detailed literature surveys included in our publications listed in the bibliography.

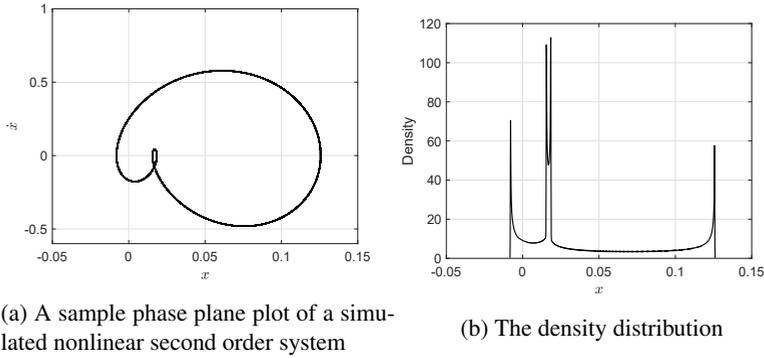


Figure 1: An example of the Phase Space Topology method [4]

The rest of the paper is organized as follows. Section 2 describes a family of methods that were originated and derived by our team: Phase Space Topology (PST) and Extended Phase Space Topology (EPST). In Section 3, we present a recent investigation of EPST for bearing defect analysis. Section 4 summarizes some of the applications that were investigated in order to generalize the applicability of our developed methods. Finally, Section 5 concludes the paper.

2 Extended phase space topology method

We first developed the method of Phase Space Topology [1–3], which is based on the transformation of phase space into the density space, which is characterized with quantitative measures. It was shown that, depending on the geometry and shape of the phase space, the density diagram contains peaks of various heights and sharpness at multiple locations, an example of which is shown in Fig. 1. This stems from the fact that the dynamical system occupies more time at specific regions of the space causing higher densities in those regions. The properties of the peaks in the density diagrams including the location, height and sharpness of the peaks were used as features in the initial approach. Despite the success of this approach, the need to search for the peaks in the density diagrams made it difficult or sometimes even impractical to implement, especially for systems with noisy or more complex phase space patterns. This led to EPST, which is a continuation of our development of the PST family of algorithms. EPST is based on approximating the density distribution with Legendre polynomials, the details of which are described below.

2.1 Kernel density estimation

Let $X=(x_1, x_2, \dots, x_n)$ be an independent and identically distributed sample data drawn from a distribution with an unknown density function f . The shape of this function can be estimated by its kernel density estimator (the hat, ^ indicates that it is an estimate, and the subscript indicates that its value can depend on h).

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (1)$$

Here, $h > 0$ is a smoothing parameter called the bandwidth, and $K(\cdot)$ is the kernel function which satisfies the following requirements.

$$\int_{-\infty}^{\infty} K(u) du = 1, \quad K(-u) = K(u) \quad \forall u \quad (2)$$

There is a range of kernel functions that can be used, including uniform, triangular, biweight, triweight, Epanechnikov and normal. Due to its conventional and convenient mathematical properties, we use the normal density function in our approach, defined as the following:

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} \quad (3)$$

2.2 Density distribution approximation

Let x be a state of the system and $y_d = \hat{f}_h(x)$, its density computed using the kernel density estimator. y_d is then approximated with Legendre orthogonal polynomials. Legendre polynomials can directly be obtained from Rodrigues' formula which is given by:

$$P_m(x) = \frac{1}{2^m m!} \frac{d^m}{dx^m} [(x^2 - 1)^m], \quad m = 0, 1, 2, \dots \quad (4)$$

It can also be obtained using Bonnet's recursion formula:

$$(m + 1)P_{m+1}(x) = (2m + 1)xP_m(x) - mP_{m-1}(x) \quad (5)$$

where the first two terms are given by:

$$P_0(x) = 1, \quad P_1(x) = x \quad (6)$$

The coefficients of the Legendre polynomials are obtained by using the least squares method assuming the following linear regression model:

$$f(x, \beta) = \sum_{j=1}^m \beta_j P_j(x) \quad (7)$$

Letting

$$X_{ij} = \frac{\partial f(x_i, \beta)}{\partial \beta_j} = P_j(x_i), \quad (8)$$

the estimated coefficients are given by:

$$\hat{\beta} = (X^T X)^{-1} X^T y_d \quad (9)$$

The coefficients $\hat{\beta}$ constitute the features in our approach that can be used in classification or regression problems. The approximated density using Legendre Polynomials is then calculated using the following:

$$f = X\hat{\beta} \quad (10)$$

Root mean square error (RMSE) and Pearson's correlation coefficient (PCC) are calculated to compute the quality of the fit using the following equations:

$$\text{RMSE} = \sqrt{\frac{1}{N} Z Z^T}, \quad \text{PCC} = \frac{a^T b}{\sqrt{(a^T a)(b^T b)}} \quad (11)$$

where, $Z = (y_d - f)$ is the residual vector, N is the number of points in the density function, $a = (y_d - E\{y_d\})$ and $b = (f - E\{f\})$, where $E\{\cdot\}$ is the expected value.

2.3 Artificial neural network

Artificial neural networks (ANN) are a set of algorithms that are designed to recognize a relation or a pattern between inputs and outputs. ANN consists of an interconnected group of nodes called artificial neurons and each node has a corresponding weight that adjusts as learning proceeds. ANN can be used to solve regression problems for a continuous output or classification problems for a discrete output. One of the most used algorithms in training ANN is the backpropagation algorithm, which is a popular method for optimizing the weights of the ANN in order to correctly map inputs to outputs. It works by propagating an input forward through the network layers to the output layer where the calculated output is compared with the desired output. The error values of the calculated output and the desired output are computed and propagated backwards. These errors are traced back to each associated neuron in order to update the weights.

3 Example application: bearing diagnostics

The algorithm, a flowchart of which is illustrated later in Fig. 3, is best explained with an example application; in this case, we choose the bearing diagnostics problem. Many traditional bearing fault detection techniques involve pattern recognition, which is effective only at one operating speed and requires retraining the classifier each time the rotational shaft speed changes because of the dependence of the dynamic response of the system and the rotational speed. This limitation motivates the need for a new method that is effective under variable operating speeds. The current study investigates different bearing configurations under two operating conditions: (A) constant operating speed and (B) variable operating speed.

In order to achieve this goal, the classification problem was initially performed by training and testing the classifier on the same set of speeds. The classifier was trained at 19 rotational speeds, and then tested on the same set of rotational speeds. The second step involved generalizing this diagnostic approach to variable operating speeds. In this step, the classifier was trained on one set of speeds and then tested on another *different* set of speeds. The detailed description of the analysis of both of the above mentioned procedures is provided in the sequel.

3.1 Case A: Constant operating speed

A rotating fault simulator machine, shown in Fig. 2, is employed to study a variety of different bearing defects under various rotational speeds (300-3000 rpm). Four bearing conditions were investigated: healthy bearings (H), bearings with inner race defects (IR), outer race defects (OR) and ball defects (B). Proximity probe sensors were used to measure the vibration signal of the shafts in two orthogonal directions.

The density function of the horizontal vibration signal for every speed and bearing condition was approximated using Legendre Polynomials. The order of the polynomial was selected based on the best fit between the estimated density function and the approximated density function. Root mean square error and Pearson's correlation coefficient were calculated to compute the quality of the fit. Legendre polynomials of order 20 were used to approximate the estimated density functions. The coefficients of the Legendre polynomials were computed for each of the 760 sampled signals using the least squares method as shown in Eq.9. The computed coefficients for each case were saved in a vector of 21 arrays (using only the horizontal vibration signal), which was used as a feature input to train an ANN classifier. Since the rotation speed has a high impact on the response of the dynamic behavior, it was used as an additional feature, making the total number of features equal to 22. With

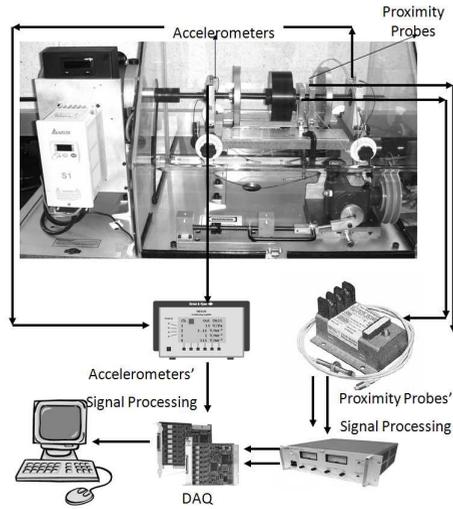


Figure 2: Rotating fault simulator machine

760 sampled data available, 50% of the sampled data (380 samples) were used for training the ANN, 25% of the sampled data (190 samples) were used for validation of the training algorithm and the remaining 25% (190 samples) were used to test the classifier. The total number of selected neurons was 10 and the backpropagation algorithm was used to train the artificial neural network.

The performance of the classification model is presented by means of confusion matrices. In general, in a confusion matrix, the predicted classes are compared with the actual classes. Each row of the matrix represents the results of prediction for the corresponding class at that row, while each column represents the actual class. The elements in the main diagonal of the matrix represent the correct classified prediction for each corresponding class. These elements are known as true positives. For a specific row, all elements excluding the element on the main diagonal are the misclassified prediction for the corresponding class, which are known as false positives. A false negative for a specific class is defined as the summation of elements on its corresponding column, excluding the element on the main diagonal.

The classifier performance can be analyzed using certain evaluation matrices derived from the confusion matrix, such as accuracy, sensitivity and precision. Table 1 shows the predictions for training and test data using the neural network classifier. As can be seen, the classifier has been able to predict all defects with 100% accuracy, 100% precision and 100% sensitivity with no misclassification. This result is remarkable for several reasons. Firstly, it shows that combining the EPST method with the proximity sensor data can resolve the challenges in identifying faults at low rotational speeds (below 10 Hz). Secondly, no *a priori* knowledge of the system was included in the features. This suggests that the EPST approach can be conveniently applied to diverse dynamical systems in an automated process, with minimal need for adaptation and reliance on expert knowledge about the system. Conventional bearing analyses search for specific characteristics of the system such as ball pass frequencies but this study did not require any additional analysis because the method functioned well without it. We note with caution that it may well be the case that other operating conditions require additional feature combinations. Finally, no feature ranking or feature selection al-

Table 1: Prediction of various defect classes for case (A)

		Test Confusion Matrix				Sensitivity	Precision
Output Class	H	44 23.2%	0 0.0%	0 0.0%	0 0.0%	100.0%	100.0%
	IR	0 0.0%	51 26.8%	0 0.0%	0 0.0%	100.0%	100.0%
	OR	0 0.0%	0 0.0%	48 25.3%	0 0.0%	100.0%	100.0%
	B	0 0.0%	0 0.0%	0 0.0%	47 24.7%	100.0%	100.0%
		H	IR	OR	B	Overall Accuracy 100.0%	
		Target Class					

gorithm [5] was employed to select the optimal feature set. Due to the fact that the effect of coefficients in function approximation decreases by increasing the number of orthogonal functions, the calculated coefficients are *naturally* ranked by their order of significance.

3.2 Case B: Variable operating speed

For this part of the study, horizontal and vertical vibration data were used for every speed and bearing condition to construct the density function. The estimated density functions were then approximated using Legendre polynomials of order 20. As in case A, the order of the approximated density function was selected based on root mean square error and Pearson’s correlation coefficient. The first 15 Legendre polynomial coefficients in each direction were used as a feature set. The shaft rotational speed was added to the feature set to produce an input vector of 31 arrays for each sampled data. The feature vector was used as an input to train the artificial neural network classifier. The neural network was modeled with 10 neurons and the back propagation algorithm. The classifier was trained using the extracted features from vibration data for different bearing conditions and for four rotational speeds. The rotational speeds that were selected to train the classifier are: the machine operating range boundaries (300 and 3000 rpm) and two middle speeds (1200 and 2400 rpm). The available vibration data of 160 total samples for different bearing conditions at these speeds were used for training the artificial neural network. The remaining 600 samples obtained at the other speeds (e.g., 420, 600, . . . , 2820 rpm) were used for testing the trained classifier. Figure 3 shows a flowchart for the algorithm of case B.

The classification results represented as a confusion matrix for the test data for four bearing conditions are shown in Table 2. As can be seen, the classifier has 96.7% overall accuracy with 20 misclassifications out of 600 predictions. These results indicate a high prediction rate of the classifier for the four bearing conditions. Most of the misclassified predictions are for bearings with ball defects. For a better understanding of the classifier performance, sensitivity and precision were calculated for each bearing condition and are shown in Table 2. These evaluation matrices represent a measure of the classification performance for each bearing condition.

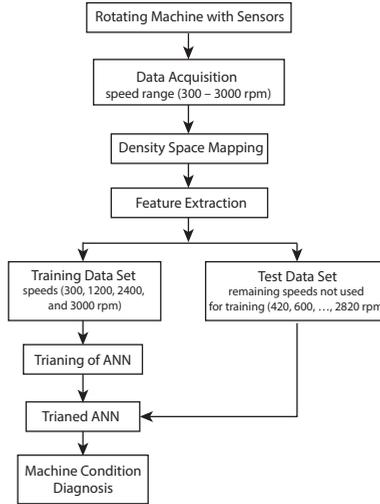


Figure 3: Flowchart of bearing diagnostics algorithm (Case B)

Table 2: Prediction of various defect classes for test data, Case B

Test Confusion Matrix					Sensitivity	Precision	
Output Class	H	150 25.0%	0 0.0%	0 0.0%	0 0.0%	100.0%	100.0%
	IR	0 0.0%	142 23.7%	0 0.0%	0 0.0%	94.7%	100.0%
	OR	0 0.0%	0 0.0%	139 23.2%	1 0.2%	92.7%	99.3%
	B	0 0.0%	8 1.3%	11 1.8%	149 24.8%	99.3%	88.7%
		H	IR	OR	B	Overall Accuracy 96.7%	
		Target Class					

4 Summary of applications

This section presents a summary of some of the applications that we investigated including bearing fault diagnostics, gear fault diagnostics and crack shaft diagnostics. In the following subsections a brief introduction to each system along with the main contributions are described.

4.1 Other studies in bearing diagnostics

Rotating machinery are probably among the most important components in industry. Rotating machines are composed of different sub-systems interacting with each other in a nonlinear fashion; changes in any of these components can significantly affect the overall performance.

Rolling element bearing defects are one of the major sources of breakdown in rotating machinery. The rotating fault simulator machine shown in Fig. 2, that is mentioned in Section 3, is considered to study a variety of different machinery defects under various operating conditions such as rotational speed, load and unbalance. It basically consists of a motor-driven shaft mounted on two bearings. Shafts and bearings with different sizes and conditions can be used. Various vibration sensors can be used such as accelerometers and proximity probes.

We have performed various investigations on this setup in order to develop robust techniques to diagnose bearings. In [5–9], we have applied conventional methods such as fast Fourier transform, envelope spectrum, and discrete wavelet transform over a span of rotational speeds as well as nonlinear physics-based modeling [10–13]. Accelerometers data were used to extract features to diagnose bearings with inner race, outer race and ball defects. Mutual information was then used as a ranking technique, and the optimal feature subset corresponding to the highest classification accuracy was determined. An overall accuracy of 97.0% was achieved using this procedure.

In [14, 15], we introduced the mapped density method in order to discriminate simultaneous bearing faults under various rotational speeds. In this work we studied the use of the information provided by proximity probe sensors. The method has significant success in fault discrimination for a single and two bearing fault configurations (accuracy of 97% for a single bearing fault and 92% for two bearing faults). Moreover, the results indicate that this method has high performance in distinguishing between different bearing conditions signatures (accuracy of 88%).

In [4, 16, 17], the EPST method was introduced. In this work we perform bearing diagnostics on different rotational speeds domains and obtained very good results (overall accuracy of 96.7%). We have also applied other nonlinear techniques such as recurrence plots [18], Gottwald and Melbourne's 0-1 test and the Huguchi fractal dimension [19].

4.2 Gear-train setup

Gear fault diagnostics is still a challenging task because of the highly nonlinear characteristics of faults and its complex nonstationary dynamics. Our work investigated gear fault diagnostics using vibration data of a helicopter gearbox mock-up system provided by UTRC. The experimental setup is shown in Fig. 4. This work studied multiple test gears with different health conditions such as healthy gears (H) and defective gears with root crack on one tooth (SCD), multiple cracks on five teeth (MCD) and missing tooth (MTD) are studied. The vibrational signals were recorded using a triaxial accelerometer installed on the test gearbox.

In [20, 21], we presented the application of recurrence plots (RPs) and recurrence quantification analysis (RQA) in the diagnostics of various faults in a gear-train system. It also apply mutual information to rank the extracted features in order to obtain an optimal feature set. Results indicate that RQA parameters provide valuable information in characterizing the dynamics of various gear faults in order to discriminate the healthy gear condition from defective conditions. Also, outstanding performance was achieved using RQA parameters to identify various gear conditions with 100% accuracy, 100% recall and 100% precision in detecting multiple cracks and missing tooth conditions.

In [22], the EPST method was applied to detect anomaly behavior and to diagnose various gear defects. Results indicates 99% accuracy in classifying between different gear conditions.

4.3 Crack detection

Experiments were conducted on a Crack Propagation Simulator test rig, shown in Fig. 5, in collaboration with Federal University of Uberlândia, Brazil. This machine consists of a



Figure 4: Gear-train experimental setup

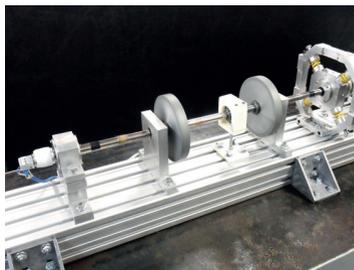


Figure 5: Crack experimental setup

flexible steel shaft mounted on two roller bearings. Two orthogonal proximity probes oriented in the horizontal and vertical directions were used to measure the vibration response of the shaft. The crack propagator was used over a period of 24 hours to produce a fatigue crack, which is the first damage condition. Then the second fatigue crack was produced by using the crack propagator for another 24 hour period.

In [23], the EPST method integrated with mutual information was applied to detect cracks and to identify the level of degradation. Mutual information was used to only select the most relevant EPST extracted features. Results show 100% performance by using this algorithm; it is notable that only three features were necessary to detect cracks and identify the crack level.

5 Conclusion

This paper gives a snapshot of more than a decade's work that was originated and developed here at VCADS. These methods are based on the hypothesis that the nonlinear response of dynamical systems contains valuable information about the system, given the intrinsic nonlinear nature of real systems. We conclude with a strong message that the key to robust and accurate diagnostics is algorithms that zoom in on these nonlinear aspects of the response and can characterize them effectively.

References

- [1] M. Samadani, C.A.K. Kwiimy, C. Nataraj, *Communications in Nonlinear Science and Numerical Simulation* **20**, 583 (2015)
- [2] M. Samadani, C.A.K. Kwiimy, C. Nataraj, *Diagnostics of a nonlinear pendulum using computational intelligence*, in *Proceedings of the 6th Annual Dynamic Systems and Control Conference* (Palo Alto, CA, 2013)
- [3] M. Samadani, C.A.K. Kwiimy, C. Nataraj, *Nonlinear Dynamics* **86**, 2023 (2016)
- [4] T.H. Mohamad, M. Samadani, C. Nataraj, *Journal of Vibration and Acoustics* **140** (2018)
- [5] K. Kappaganthu, C. Nataraj, *ASME Journal of Vibration and Acoustics* **133** (2011)
- [6] K. Kappaganthu, C. Nataraj, B. Samanta, *Optimal Feature Set for Detection of Inner Race Defect in Rolling Element Bearings*, in *Annual Conference of the Prognostics and Health Management Society* (2009)

- [7] K. Kappaganthu, C. Nataraj, B. Samanta, *Feature Selection for Bearing Fault Detection Based on Mutual Information*, in *IUTAM Symposium on Emerging Trends in Rotor Dynamics*, edited by K. Gupta (Springer Netherlands, Dordrecht, 2011), pp. 523–533, ISBN 978-94-007-0020-8
- [8] K. Kappaganthu, C. Nataraj, *Advances in Vibration Engineering* **11** (2011)
- [9] K. Kappaganthu, Ph.D. thesis, Villanova University (2010)
- [10] D. Maraini, C. Nataraj, *Journal of Sound and Vibration* **420**, 227 (2018)
- [11] K. Kappaganthu, C. Nataraj, *Communications in Nonlinear Science and Numerical Simulation* **16**, 4134 (2011)
- [12] S.P. Harsha, C. Nataraj, *International Journal of Nonlinear Sciences and Numerical Simulation* **7**, 163 (2006)
- [13] S. Harsha, C. Nataraj, P. Kankar, *International Journal of Acoustics and Vibrations* **11**, 56 (2006)
- [14] T.H. Mohamad, C.K. Kwuimy, C. Nataraj, *Discrimination of Multiple Faults in Bearings Using Density-Based Orthogonal Functions of the Time Response*, in *ASME 2017 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (American Society of Mechanical Engineers, 2017), pp. V008T12A035–V008T12A035
- [15] M. Samadani, Ph.D. thesis, Villanova University (2017)
- [16] M. Samadani, T.H. Mohamad, C. Nataraj, *Feature Extraction for Bearing Diagnostics Based on the Characterization of Orbit Plots With Orthogonal Functions*, in *ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference* (American Society of Mechanical Engineers, 2016), pp. V008T10A026–V008T10A026
- [17] T.H. Mohamad, C.A.K. Kwuimy, C. Nataraj, *Multi-speed Multi-load Bearing Diagnostics Using Extended Phase Space Topology*, in *International Conference on Structural Nonlinear Dynamics and Diagnosis* (2018)
- [18] C.A. Kwuimy, M. Samadani, P.K. Kankar, C. Nataraj, *Recurrence analysis of experimental time series of a rotor response with bearing outer race faults*, in *ASME International Design Engineering Technical Conferences* (2015)
- [19] C.A.K. Kwuimy, T.H. Mohammad, C. Nataraj, *Using the Gottwald and Melbourne's 0-1 test and the Hugichi fractal dimension to detect chaos in defective and healthy ball bearings*, in *International Conference on Structural Nonlinear Dynamics and Diagnosis* (2018)
- [20] C.A.K. Kwuimy, P.K. Kankar, Y. Chen, Z. Chaudhry, C. Nataraj, *Development of recurrence analysis for fault discrimination in gears*, in *ASME International Design Engineering Technical Conferences* (2015)
- [21] T.H. Mohamad, Y. Chen, Z. Chaudhry, C. Nataraj, *Journal of Software Engineering and Applications* **11**, 181 (2018)
- [22] T.H. Mohamad, C. Nataraj, *Gear Fault Diagnostics Using Extended Phase Space Topology*, in *Annual Conference of the Prognostics and Health Management Society* (2017)
- [23] T. Haj Mohamad, A.A. Cavalini, V. Steffen, C. Nataraj, *Detection of Cracks in a Rotating Shaft Using Density Characterization of Orbit Plots*, in *Proceedings of the 10th International Conference on Rotor Dynamics – IFToMM*, edited by K.L. Cavalca, H.I. Weber (Springer International Publishing, Cham, 2019), pp. 90–104, ISBN 978-3-319-99268-6