



A Review of Phase Space Topology Methods for Vibration-Based Fault Diagnostics in Nonlinear Systems

T. Haj Mohamad¹ · Foad Nazari¹ · C. Nataraj¹

Received: 23 January 2019 / Revised: 2 May 2019 / Accepted: 9 June 2019 / Published online: 19 June 2019
© The Author(s) 2019, corrected publication 2019

Abstract

Background 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.

Purpose We have been developing some interesting techniques for fault diagnostics with gratifying results.

Methods 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.

Keywords Phase Space Topology · Bearing diagnostics · Gear fault diagnostics · Crack detection · Machine learning

Introduction

Fault diagnostics of practical systems is a very important problem that needs to be solved robustly to be able to make giant leaps in reliability and safety. Diagnostics is essentially an epistemological problem that requires us to make intelligent inferences based on data, which could be derived from empirical observations or computer models, and are often incomplete, noisy and uncertain. Although there is a rich and varied literature,¹ we feel that many of the diagnostic techniques in use are quite ad hoc and heuristic, resulting in lack of general applicability.

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 demonstrate applicability 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).

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

✉ T. Haj Mohamad
thajmoha@villanova.edu
http://vcads.org

Foad Nazari
foad.nazari@villanova.edu

C. Nataraj
c.nataraj@villanova.edu

¹ Villanova Center for Analytics of Dynamic Systems (VCADS), Villanova University, Villanova, PA 19085, USA

¹ We wish to acknowledge the excellent contributions in fault diagnostics by many of our esteemed colleagues; we have had to omit references to their work to keep the paper short and focused on an outline of our approach; kindly refer to the detailed literature surveys included in our publications listed in the bibliography.

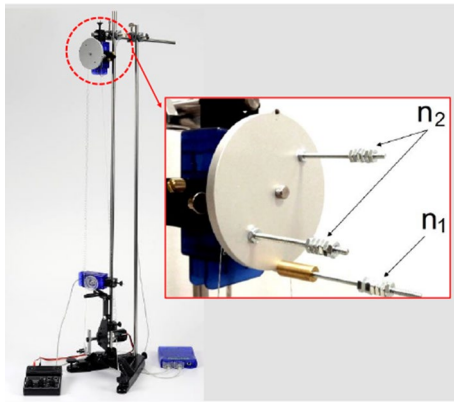


Fig. 1 The nonlinear pendulum experimental setup

Phase Space Topology Family of Methods

A Brief History of PST Development

The PST family of methods, which is based on characterizing the system phase space using quantitative measures, first originated in our previous work [1]. It should be noted that other studies [2–5] have also studied extraction of information from the system's phase space for damage detection, although they have specialized it for the problem of structural health monitoring (SHM). They are indeed superficially similar to our method, but significant differences do exist. Most of these methods require exciting the structure by a chaotic input to extract a single feature and are directed toward SHM applications. In addition, the PST methods outlined below introduce original concepts such as Kernel density-based features, density approximation and orthogonal basis functions as a feature extraction technique.

We first focused on a nonlinear pendulum (shown in Fig. 1), which was diagnosed using a set of four features extracted from the phase plane trajectory of the system to characterize the nonlinear response in the periodic regime. The extracted features were related to the phase portrait eccentricity, harmonic amplitude and excursion variation. Finally, an Artificial Neural Network (ANN) was built using the numerical data for the estimation of parameters of a defective nonlinear pendulum set up with acceptable accuracy.

Kernel Density-Based Features

In Samadani et al. [6], an expansion to the previous study was developed to include the multi-periodic response regime of the pendulum which was characterized using an additional set of features extracted from the constructed density distribution of the phase space. To demonstrate how the Kernel density is estimated, 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, $\hat{\cdot}$, indicates that it is an estimate, and the subscript indicates that its value can depend on h).

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

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

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

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 follows:

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

An example of the periodic and multi-periodic phase planes along with their corresponding density distributions of the nonlinear pendulum is shown in Fig. 2. 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. 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 additional feature set that was extracted from the density distribution included the location, the height and the sharpness of the peaks in the density profile. The results showed that the introduced features are fully capable of characterizing the nonlinear response of the system in the periodic and multi-periodic domains with high complexities. It also demonstrated that the developed parameter estimation approach works effectively for both analytical and experimental data.

Subsequently, in [7, 8], the same approach was validated using a 3-DOF nonlinear mass–spring–damper system where the diagnostics was treated as a parameter estimation problem. The results showed the effectiveness of the approach in characterizing the system behavior and capturing the dynamical changes while changing six parameters of the system simultaneously. 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 new advances continuing the development of PST, based

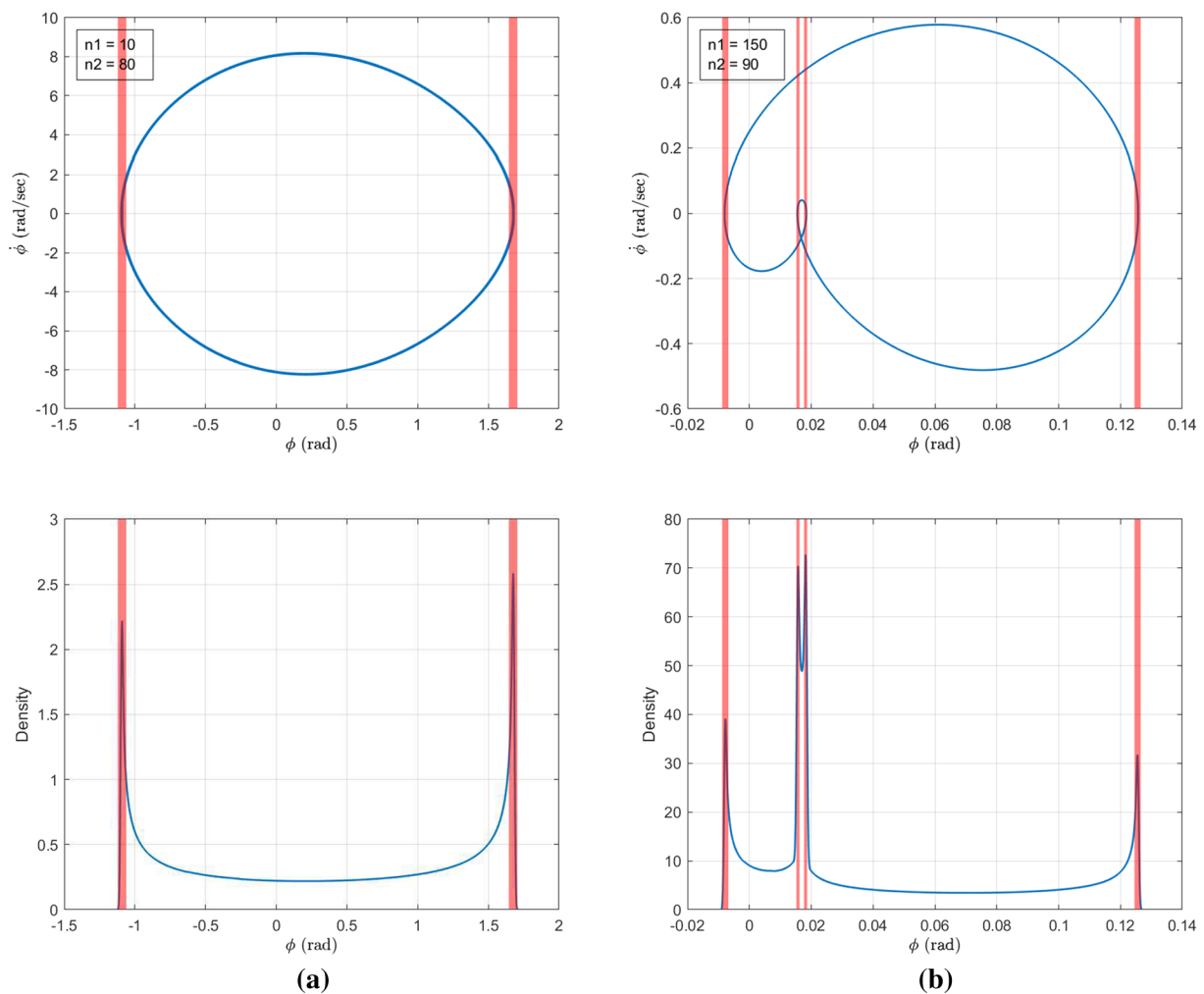


Fig. 2 The phase plane and the corresponding density distribution for **a** periodic domain response and **b** multi-periodic domain response

on approximating the density distribution with an orthogonal basis, the details of which are described below.

Recent Contributions to the Development of PST

Density Distribution Approximation

A feature set is said to be ideal when it is able to reconstruct the original data. This assumption motivates the approximation of the density distribution of the phase space using a series of orthogonal basis where the coefficients of these basis will be used as features.

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

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

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

$$(m + 1)P_{m+1}(z) = (2m + 1)zP_m(z) - mP_{m-1}(z), \tag{5}$$

where the first two terms are given by

$$P_0(z) = 1, \quad P_1(z) = z. \tag{6}$$

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

$$f(z, \beta) = \sum_{j=1}^m \beta_j P_j(z). \tag{7}$$

Letting

$$Z_{ij} = \frac{\partial f(z_i, \beta)}{\partial \beta_j} = P_j(z_i), \tag{8}$$

the estimated coefficients are given by

$$\hat{\beta} = (Z^T Z)^{-1} Z^T y_d \tag{9}$$

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:

$$f = Z\hat{\beta} \tag{10}$$

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

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

where $V = (y_d - f)$ 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 $E\{.\}$ is the expected value.

Artificial Neural Network (ANN)

ANNs are a set of computational models that are designed to recognize a relation or a pattern between inputs and outputs. Nowadays, they are widely used for regression, classification, prediction and optimization problems. ANN consists of an interconnected group of nodes called artificial neurons and each node has a corresponding weight. The knowledge of the system is stored as a set of ANN connection weights. The process of mapping inputs to outputs is called training, which is performed by adjusting the weights as learning proceeds. One of the most used algorithms in training ANN is the backpropagation algorithm, which optimizes the weights of the ANN 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 then computed and propagated backwards. These errors are traced back to each associated neuron to update the weights.

In the last decade, neural networks have received considerable attention in the field of fault diagnostics. One of the main issues in ANNs is selecting an optimal feature set that can represent the system in different operating domains. The general trend in many applications is extracting a large number of features to be used. On the one hand, it works in some scenarios because the nonlinearity of systems makes it hard to define the relevant information that is needed. On the other hand, for applications where the available data are small with respect to the dimensionality of the system, extracting all features might be impossible [9]. In many cases, some of the extracted features are redundant or irrelevant and one can achieve better performance by discarding such variables [10].

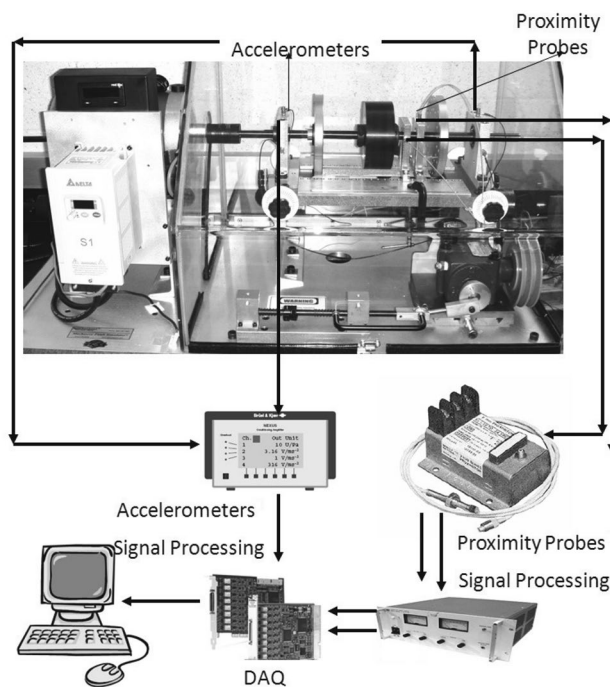


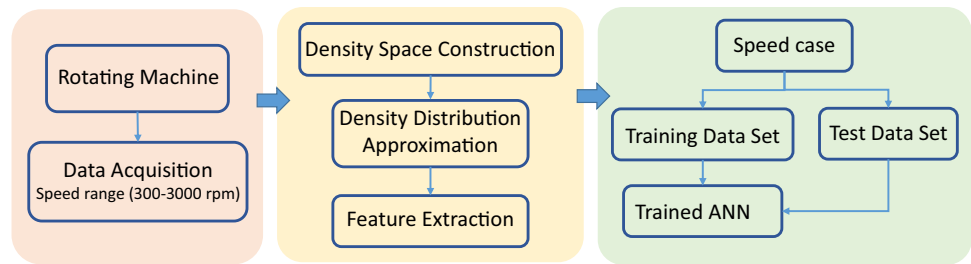
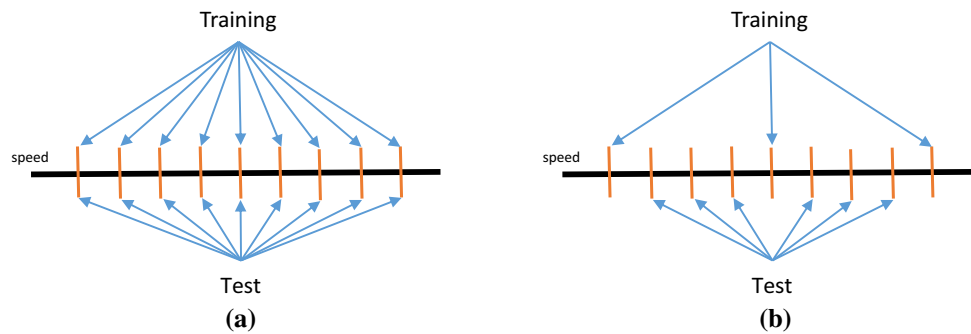
Fig. 3 Rotating fault simulator machine

Our goal has always been to extract just enough number of features that are relevant and provide rich information about the system. This can be achieved by combining our understanding of nonlinear dynamics in the developed algorithms. In this study, it is demonstrated that the selected number of features extracted by PST method provides very high level of information about the faults and the nonlinear effects on the system.

Example Application: 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. In particular, rolling element bearing defects are one of the major sources of breakdown in rotating machinery. The rotating fault simulator machine shown in Fig. 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.

The PST algorithm is best explained with an example application; in this case, we choose the bearing diagnostics problem. A flowchart of the algorithm is illustrated in Fig. 4.

Fig. 4 Flowchart of bearing diagnostics algorithm**Fig. 5** Speed configurations for training and testing the classification model **a** known speed sets and **b** bounded speed sets

Many traditional techniques that have been developed for bearing fault detection involve pattern recognition, which is only effective for an a priori defined operating condition set, shaft rotational speed in this case. Therefore, the classification model is required to be retrained each time the operating condition (rotational shaft speed) changes because of the dependence of the dynamic response on them. The lack of effective techniques under variable operating speeds is a well-established need, which motivates the development of this method.

In this study, different rotational speed configurations between the training and testing sets of the classification model are investigated in (A) known speed domains and (B) bounded speed domains. To achieve this goal, the classification problem was initially trained on a set of speeds and then the classifier was tested on the *same* set of speeds. The classifier was trained and tested at the same 19 rotational speeds. In the second case, the bearing fault diagnostic approach was generalized to variable operating speeds within a given range of speeds (i.e., 300–3000 rpm). In this case, the classifier was trained on one set of speeds and then tested on another *different* set of speeds in which the speed test domains were completely bounded by the training speed sets as shown in Fig. 5. The analysis of both of the aforementioned procedures is described in the sequel in detail.

Case (A): Known Speed Domains

To study a variety of different bearing defects under various rotational speeds (300–3000 rpm), a rotating fault simulator

machine, shown in Fig. 3, was investigated. Four bearing conditions were investigated: healthy bearings (H), bearings with inner race defects (IR), outer race defects (OR) and ball defects (B). Vibration data were measured using two GE/Bently Nevada 7200 series proximity probe sensors in orthogonal directions and installed close to the bearing housing.

Legendre polynomials were used to approximate the estimated density distribution of the horizontal vibration signal for every rotational speed and bearing condition. The Legendre polynomial order was selected based on the least polynomial order that gives a reliable fit between the estimated and approximated density functions. To compute the quality of the fit, root mean square error and Pearson's correlation coefficient were calculated. Based on these metrics, Legendre polynomials of order 20 were used to approximate the estimated density functions. Coefficients for the Legendre polynomials were computed for each of the 760 signal samples using the least squares method as shown in Eq. (9). The computed coefficients for each case, which represent the feature vector for that case, were saved in a 21-element vector (using only the horizontal vibration signal). Because of high impact of the rotation speed on the response of the dynamic behavior, it was used as an additional feature, making the total number of features equal to 22. From all available sampled data (760 samples), 50% of them (380 samples) were used to train the ANN, 25% of them (190 samples) were used to validate the training algorithm and the remaining 25% (190 samples) were used to test the classifier. The backpropagation algorithm was used for training

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

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

the ANN. An ANN with ten hidden neurons was considered for this problem.

The performance of the selected classification model is presented using confusion matrices. A confusion matrix compares between the results of the predicted and the actual classes. Each row of the matrix shows the prediction results for the corresponding class at that row while each column shows the actual class. The main diagonal elements of the matrix show the correct classified prediction for each corresponding class, which are known as true positives. For each row of the matrix, all elements except the main diagonal elements are the misclassified prediction for the corresponding class, which are known as false positives. For each specific class, the summation of elements on its corresponding column excluding the main diagonal elements are called false-positive elements. The performance of the classifier can be evaluated using certain evaluation matrices derived from the confusion matrix, such as sensitivity, precision and overall accuracy. Table 1 shows the predictions for test data using the neural network classifier for case (A).

As can be observed, the classifier has predicted all defects with no misclassification and with 100% sensitivity, 100% precision and 100% overall accuracy. These results are remarkable for a variety of reasons. First, it demonstrates that combination of the proximity sensor data and the PST features can address the challenges in fault identification at low rotational speeds (below 10 Hz). Second, no a priori knowledge of the system was integrated in extracted features. Therefore, the PST approach can be applied to diverse dynamical systems in an automated fashion with minimal adaptation or dependence on expertise. In conventional bearing diagnostic techniques, specific characteristics of the system such as ball pass frequencies are searched for; however, the proposed methodology does not require any additional analysis and functioned well without it. We should be cautious to note that other operating conditions may need additional feature combinations. Finally, no feature selection or ranking algorithm [11] was used here to select the optimal feature set. Since the coefficients decrease by increasing the number of orthogonal functions, the calculated coefficients are ranked *naturally*.

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

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

Case B: Bounded Speed Domains

For this case, density distributions were constructed using both horizontal and vertical vibration data for every speed and bearing condition. Then using Legendre polynomials of order 20, the estimated density functions were approximated. The feature set was constructed using the first 15 Legendre polynomial coefficients in each direction in addition to the rotational speed of the shaft making the total number of features equal to 31. The feature vector was then used as an input for training the ANN classifier. The classifier was modeled with ten neurons in the hidden layer and trained using backpropagation algorithm. For training, the classification model, four rotational speeds were selected including the machine operating range boundaries (300 and 3000 rpm) and two middle speeds (1200 and 2400 rpm). For these speeds, the available vibration data of 160 total samples at different bearing conditions were used for training and the remaining 600 samples obtained at the other speeds (e.g., 420, 600, ..., 2820 rpm) were used for testing the trained classifier.

The classification results for the test data are represented as a confusion matrix in Table 2. As can be observed, the classifier overall accuracy is 96.7% with 20 misclassifications out of 600 predictions, which indicates a high prediction rate of classifying the four bearing conditions. Most of the misclassifications corresponded to bearings with ball defects. To evaluate effectively, the precision and sensitivity parameters were calculated and presented in Table 2. In summary, the classifier had both high precision and sensitivity for the majority of the bearing conditions. The classification model worked virtually perfectly in distinguishing between healthy and defective bearings since there were zero misclassifications.

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.

Table 3 Summary of other investigations

Component	Issue and failure	Sensor data	Operating condition	Algorithm used
Nonlinear pendulum	Parameter estimation	Rotary motion sensor	Variable lumped mass n_1 and number of nuts n_2	Recurrence and statistical analysis [12], PST [1]
Bearing	Outer race, inner race and roller	Proximity probes and accelerometers	Variable speed and motor load	PST [13–15], mapped density [16], envelope analysis [17], statistical time analysis [14, 17]
Gear	Single crack, multiple cracks and missing tooth	Accelerometers and tachometer	Constant speed and motor load	PST [18], RP [19]
Shaft	Propagated cracks	Proximity probes	Variable speed	PST [20]
Servo actuator	Spool friction, permanent magnet	Position, velocity and pressure	Variable	PST [21]

In the following subsections, a brief introduction to each system along with the main contributions are described. Table 3 summarizes the main contributions that were achieved.

Other Studies in Bearing Diagnostics

We have performed various investigations on the bearing set up (shown in Fig. 3) to develop robust techniques to diagnose bearings. In [11, 22–25], 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 [26–29]. Accelerometer 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 [16, 30], we introduced the mapped density method 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 had 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 [13–15, 31], the PST method was introduced. In this work, we performed 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 [32], Gottwald and Melbourne's 0-1 test and the Higuchi fractal dimension [33].

Gear Fault Diagnostics

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. 6. 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). The vibrational signals were recorded using a triaxial accelerometer installed on the test gearbox.

In [19, 34], 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 applied mutual information to rank the extracted features order to obtain an optimal feature set. Results indicated that RQA parameters provide valuable information in characterizing the dynamics of various gear faults to discriminate

**Fig. 6** Gear-train experimental setup

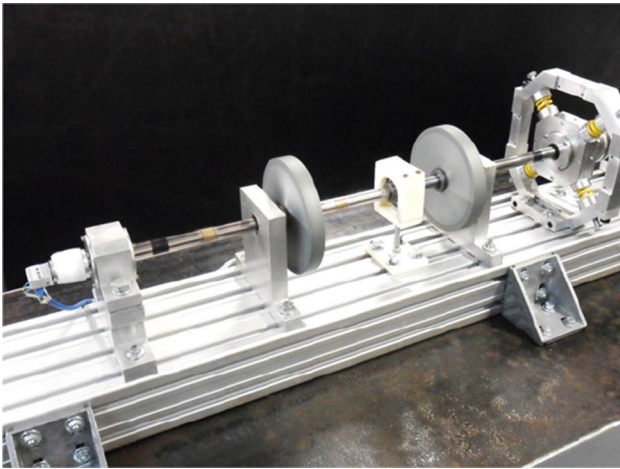


Fig. 7 Crack experimental setup

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 [18], the PST method was applied to detect anomalous behavior and to diagnose various gear defects. Results indicate 99% accuracy in classifying between different gear conditions.

Crack Detection

Experiments were conducted on a crack propagation simulator test rig, shown in Fig. 7, in collaboration with Federal University of Uberlândia, Brazil. This machine consists of a 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 h to produce a fatigue crack, which is the first damage condition. Then the second fatigue crack was produced using the crack propagator for another 24 h period.

In [20], the PST 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 PST-extracted features. Results show 100% performance using this algorithm; it is notable that only three features were necessary to detect cracks and identify the crack level.

Conclusion

This paper summarizes the development of a novel family of methods that originated and developed at VCADS. It discusses briefly some of the applications such as nonlinear

pendulum, gear-train, cracked shafts and electro-hydraulic servo actuators [21]. It also presents a recent investigation of PST for bearing defect analysis. 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 are able to zoom in on these nonlinear aspects of the response and can characterize them effectively.

Acknowledgements This work is supported by the US Office of Naval Research under the Grant ONR N0014-15-1-2311 with Capt. Lynn Petersen as the Program Manager. We deeply appreciate this support and are humbled by ONR's enthusiastic recognition of the importance of this research.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Samadani M, Kwuimy CAK, Nataraj C (2013) Diagnostics of a nonlinear pendulum using computational intelligence. In: Proceedings of the 6th annual dynamic systems and control conference, Palo Alto, CA, 2013
2. Todd M, Nichols J, Pecora L, Virgin L (2001) Vibration-based damage assessment utilizing state space geometry changes: local attractor variance ratio. *Smart Mater Struct* 10:1000
3. Torkamani S, Butcher EA, Todd MD, Park G (2012) Hyperchaotic probe for damage identification using nonlinear prediction error. *Mech Syst Signal Process* 29:457
4. Nie Z, Hao H, Ma H (2012) Using vibration phase space topology changes for structural damage detection. *Struct Health Monit* 11:538
5. George RC, Mishra SK, Dwivedi M (2018) Mahalanobis distance among the phase portraits as damage feature. *Struct Health Monit* 17:869
6. Samadani M, Kwuimy CAK, Nataraj C (2015) Model-based fault diagnostics of nonlinear systems using the features of the phase space response. *Commun Nonlinear Sci Numer Simul* 20:583
7. Samadani M, Kwuimy CAK, Nataraj C (2015) Characterization of phase space topology using density: application to fault diagnostics. In: Annual conference of the prognostics and health management society, Coronado, CA, 2015
8. Samadani M, Kwuimy CAK, Nataraj C (2016) Characterization of the nonlinear response of defective multi-DOF oscillators using the method of phase space topology (PST). *Nonlinear Dyn* 86:2023
9. Verikas A, Bacauskiene M (2002) Feature selection with neural networks. *Pattern Recognit Lett* 23:1323
10. Benoît F, Van Heeswijk M, Miche Y, Verleysen M, Lendasse A (2013) Feature selection for nonlinear models with extreme learning machines. *Neurocomputing* 102:111
11. Kappaganthu K, Nataraj C (2011) Feature selection for fault detection in rolling element bearings using mutual information. *ASME J Vib Acoust* 133:061001

12. Kwuimy CAK, Samadani M, Nataraj C (2014) Bifurcation analysis of a nonlinear pendulum using recurrence and statistical methods: applications to fault diagnostics. *Nonlinear Dyn* 76:1963
13. Samadani M, Mohamad TH, Nataraj C (2016) 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
14. Mohamad TH, Samadani M, Nataraj C (2018) Rolling element bearing diagnostics using extended phase space topology. *J Vib Acoust* 140:061009
15. Mohamad TH, Kwuimy CAK, Nataraj C (2018) Multi-speed multi-load bearing diagnostics using extended phase space topology. In: MATEC web conf. vol 241, p 01011
16. Mohamad TH, Kwuimy CK, Nataraj C (2017) 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
17. Mohamad TH, Nataraj C (2019) Fault identification and severity analysis of rolling element bearings using phase space topology. *Mech Syst Signal Process* (under review)
18. Mohamad TH, Nataraj C (2017) Gear fault diagnostics using extended phase space topology. In: Annual conference of the prognostics and health management society
19. Mohamad TH, Chen Y, Chaudhry Z, Nataraj C (2018) Gear fault detection using recurrence quantification analysis and support vector machine. *J Softw Eng Appl* 11:181
20. Mohamad TH, Cavalini AA, Steffen V, Nataraj C (2019) Detection of cracks in a rotating shaft using density characterization of orbit plots. In: Cavalca KL, Weber HI (eds) Proceedings of the 10th international conference on rotor dynamics—IFTToMM. Springer International Publishing, Cham, 2019, pp 90–104, ISBN 978-3-319-99268-6
21. Mohamad TH, Nazari F, Nataraj C (2018) Model-based condition monitoring of servo valves. In: Proceedings of the annual conference of the prognostics and health management, Philadelphia, PA, 2018
22. Kappaganthu K, Nataraj C, Samanta B (2009) Optimal feature set for detection of inner race defect in rolling element bearings. In: Annual conference of the prognostics and health management society
23. Kappaganthu K, Nataraj C, Samanta B (2011) Feature selection for bearing fault detection based on mutual information. In: Gupta K (ed) IUTAM symposium on emerging trends in rotor dynamics. Springer Netherlands, Dordrecht, 2011, pp 523–533, ISBN 978-94-007-0020-8
24. Kappaganthu K, Nataraj C (2011) Modeling and analysis of outer race defects in rolling element bearings. *Adv Vib Eng* 11:371–384
25. Kappaganthu K (2010) Ph.D. thesis, Villanova University
26. Maraini D, Nataraj C (2018) Nonlinear dynamic analysis of a rotor-bearing system using describing functions. *J Sound Vib* 420:227
27. Kappaganthu K, Nataraj C (2011) Nonlinear modeling and analysis of a rolling element bearing with a clearance. *Commun Nonlinear Sci Numer Simul* 16:4134
28. Harsha SP, Nataraj C (2006) Nonlinear dynamic analysis of high-speed ball bearings due to surface waviness and unbalanced rotor. *Int J Nonlinear Sci Numer Simul* 7:163
29. Harsha S, Nataraj C, Kankar P (2006) The effect of ball waviness on nonlinear vibration associated with rolling element bearings. *Int J Acoust Vib* 11:56
30. Samadani M (2017) Ph.D. thesis, Villanova University
31. Mohamad TH, Nataraj C (2018) An overview of PST for vibration based fault diagnostics in rotating machinery. In: MATEC web of conferences, EDP Sciences, 2018, vol 211, p 01004
32. Kwuimy CA, Samadani M, Kankar PK, Nataraj C (2015) Recurrence analysis of experimental time series of a rotor response with bearing outer race faults. In: ASME international design engineering technical conferences
33. Kwuimy CAK, Mohamad TH, Nataraj C (2018) Using the Gottwald and Melbourne's 0-1 test and the Hugichi fractal dimension to detect chaos in defective and healthy ball bearings. In: MATEC web conf. vol 241, p 01017
34. Kwuimy CAK, Kankar PK, Chen Y, Chaudhry Z, Nataraj C (2015) Development of recurrence analysis for fault discrimination in gears. In: ASME international design engineering technical conferences

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.