IOP Conference Series: Materials Science and Engineering

PAPER • OPEN ACCESS

Significance of Incorporating Chebfunction Coefficients for Improved Machine Fault Diagnosis

To cite this article: T K Sachin Saj *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **561** 012090

Significance of Incorporating Chebfunction Coefficients for Improved Machine Fault Diagnosis

***Sachin Saj T K, Neethu Mohan, Sachin Kumar S and Soman K.P**

Center for Computational Engineering & Networking (CEN), Amrita School of Engineering, Coimbatore, Amrita VishwaVidyapeetham, India.

E-mail: *sachin96saj@gmail.com

Abstract. For any industry the efficiency and performance of rotating machinery/mechanical systems is a major concern. Bearings and gears are two essential parts in a rotating machinery and any defects in these components can lead to a major breakdown of the system thus causing large economical loss for the company. An appropriate machine condition monitoring system is essential in such scenarios for identifying the health of the machines. Therefore in this paper fault diagnosis of rotating mechanical systems is performed as a feature dependent-pattern classification problem. The machine is made to run in different good as well faulty conditions and the vibration signals are collected. Then chebfunction coefficients are extracted from the vibration signals as part of the feature extraction process. Finally, the extracted features are classified using regularized least squares (RLS) for identifying the good and faulty bearing as well as gear conditions of the machine. The evaluations are performed using different kernel functions and the average accuracy reported is 98% for bearing and gear data. The various experiments performed claims that the proposed system can be used for real-time fault diagnosis in rotating mechanical systems with sufficient accuracy.

1. Introduction

From last two decades there is an extensive increase in manufacturing industries since themanufacturing capabilities of the machines have increased exponentially [1, 2, 3]. Bearings andgears are two important parts in any rotating mechanical systems and faults in these parts canlead to complete breakdown of the system. This can cause a huge economic loss for the company.One of the primary failure mode of bearings and gears are localized defects such as cracks [4].The detection and classification of the faults/defects become a crucial challenge in the modernscenario. Machine condition monitoring helps us in continuously monitor the machine partssuch as bearing, gears etc. It gives indications of future failures, so that proper maintenancemeasures can be done prior itself [5, 6]. This will avoid huge break down of the machinery thussaving millions of rupees for the company. A machine condition monitoring system continuouslycollect the data, process the data and interpret the data [7]. The three main stages in machinefault diagnosis system are 1) signal acquisition-the collection of good/faulty data from rotatingmachines belongs to this stage, 2) feature extraction-this stage consists of the identification ofsuitable features and its extraction from collected data, 3) pattern classification- identificationof normal and abnormal signals by feeding the extracted features to suitable classifier belongsto this final stage.Vibration analysis is the most successful method that is used in machine condition monitoringsystems [8]. Every machines will generate certain kind of vibrations during their operation, evena good condition machine. The only difference is that the frequencies of vibrations producedby good conditioned machines and fault existing machines are different and this is the principleof machine condition monitoring system using vibration signals. Additional benefit of usingvibration signals is that it contains vital information related to the internal forces of the systemthus it can provide much more accurate results.

This paper proposes a feature dependent-machine condition monitoring system using chebfuncoefficients and regularized least squares (RLS) for machine fault diagnosis. In the proposedsystem, the collected vibration signals from different good and faulty conditions are subjectedto the feature extraction stage using chebfun approximation system. The suitable chebfuncoefficients are taken as the features and are classified using RLS classifier. Various experimentsusing different kernel functions are performed to evaluate the effectiveness of the proposed faultdiagnosis system. The remaining section of the paper is organized as follows: Section 2 givesa brief explanation about chebfun approximation system and RLS classifier. The experimentalsetup for vibration signal acquisition is discussed in Section 3 and Section 4 explains thetwo datasets used for experiments. Section 5 narrates the proposed system for machine faultdiagnosis and Section 6 describes the major results of the proposed system. Finally conclusions are given in section 7.

2. Background

2.1. Chebfun Approximation Function

Chebfun (Cheb function) system is a recently developed platform that can be used for performingnumerical computation with functions [9]. Chebfun approximation offers high speed and preciseaccuracy for computations. In chebfun approximation, functions are represented by chebyshevpolynomial expansions. The chebyshev points are computed by preserving machine precision(roughly to machine precision is about 15-digit accuracy) [10]. The chebfun system dealswith simple linear algebra problems to complicated non-linear problems. The example forconstructing chebfun for a given function is illustrated as follows.

```
F = chebfun('sin(10*x)')
plot(F)
```

Figure 1 represents the chebfun approximation corresponds to sin(10x) in continuous domain.From figure 1 it is clear that the default interval for chebfun approximation is [-1 1].

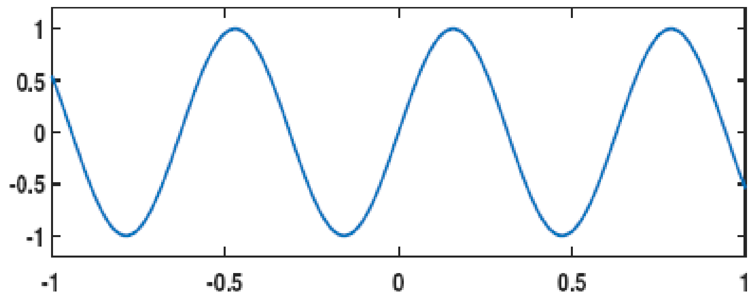

Figure 1: Chebfun illustration of sin(10x)

```
length(F)
Ans => 34
```

The length of chebfun object `F' specifies that this function is represented by using 34 chebyshevpoints, that is the resulting chebyshev polynomial has a degree of 34. Figure 2 represents thevisualization of this 34 chebyshev points. For a given function the corresponding chebfun is apolynomial Interpol ant using chebyshev points. It is defined as in equation 1.

$$x_i = \cos\left(\frac{\pi i}{n}\right), 0 \le i \le n \quad (1)$$

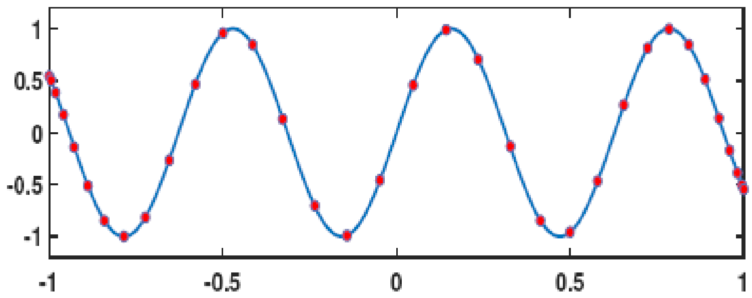

Figure 2: Representation of 34 chebyshev points

The distribution of points towards the end is more compared to the middle, and at -1 and1 there are high number of points representing the signals. Recently, chebfun system is beingused in domains such as control problems in aeronautics [11], power system [12], speech [13] etc.

2.2. Regularized Least Squares Classifier

Regularized least squares (RLS) classifier is one of the basic algorithm used for supervisedlearning [14]. In RLS, the training of the data is done by solving a system of linear equations.In this proposed method, the chebfun coefficients generated as our features is given to the RLSclassifier. For an n dimensional data vector, X = $\{x_1; x_2::::::x_n\}$ with corresponding labelsY = $\{y_1; y_2::::::y_n\}$, RLS predicts the output by computing the weight matrix (W) [2]. Theobjective function is defined as,

$$\min_{W \in R^{n*T}} = \frac{1}{n} \|Y - WX\|_F^2 + \lambda \|W\|_F^2 \tag{2}$$

Where λ is the control parameter. The weight matrix can be finally obtained as,

$$W = (X^T X + \lambda I) X^T Y \tag{3}$$

The detail explanation related to the derivation of RLS algorithm is given in [15, 16].

3. Experimental Setup

In this paper we have used two different experimental setups: A) Bearing fault simulator system, B) Helical gear simulator system for generating the vibration signals.

3.1. Bearing Fault Simulator System

The experimental setup shown in figure 3 is the bearing fault simulator connected to sensors[5, 8]. This setup consists of variable DC motor of max 0.5 HP with a speed ranging from 0 to3000 rpm. A shaft of diameter 30 mm is placed to motor through an aluminum flexible coupling.The reason for doing this is to minimize the effects caused by the misalignment and transmissionof vibration from the motor. The setup consists of a control panel for the entire monitoring ofthe system. The details about this fault simulator setup is available in [2, 5, 8].

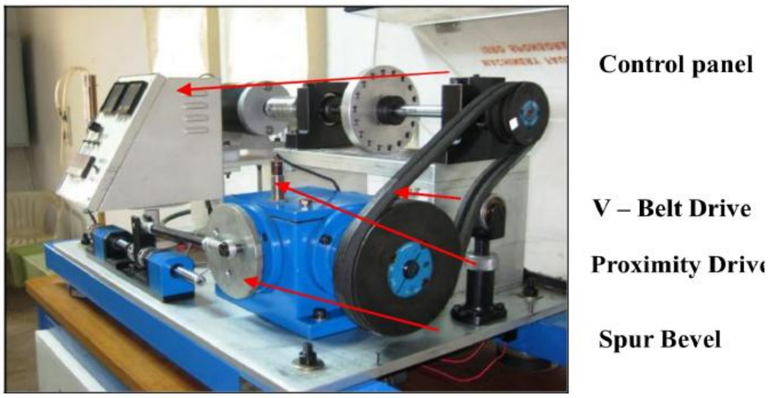

Figure 3: Experimental setup for bearing fault simulator

3.2. Helical Gear Fault Simulator System

Helical gear is one of the most important gear used for transmission as it produces large amountof thrust [17]. This fact motivated us to use gear fault system for condition monitoring. For theexperiment, a two-stage helical gearbox of 5 HP is used. A 3-phase induction motor is used todrive this two-stage helical gearbox. An inverter drive is used to control the motor speed. The first stage gearbox speed is about 80 rpm and the speed of the smaller gear shaft is 1200 rpm.This shaft is attached to DC motor that generates about 2 KW power. Hence the final resultantload on the gearbox is 2.6 HP that is only 52% out of its rated power of 5 HP. The whole setupis mounted on I beams. The vibration signal coming out of the setup is collected through theaccelerometers connected to the gearbox.

4. Vibration Data Acquisition

This section explains details about the vibration data generated using the experimental setupdescribed in Section 3.

4.1. Bearing Fault Dataset (Dataset-1)

The bearing fault dataset (dataset-1) consists of 4 classes such as A: good condition, B: innerrace fault condition (IRF), C: outer race fault condition (ORF), D: inner and outer race faultcondition (IORF). Figure 4 represents the examples of signal corresponds to these 4 generatedclasses. The signals are generated with sampling frequency of 12000 Hz and sampling length of8192 for all conditions.

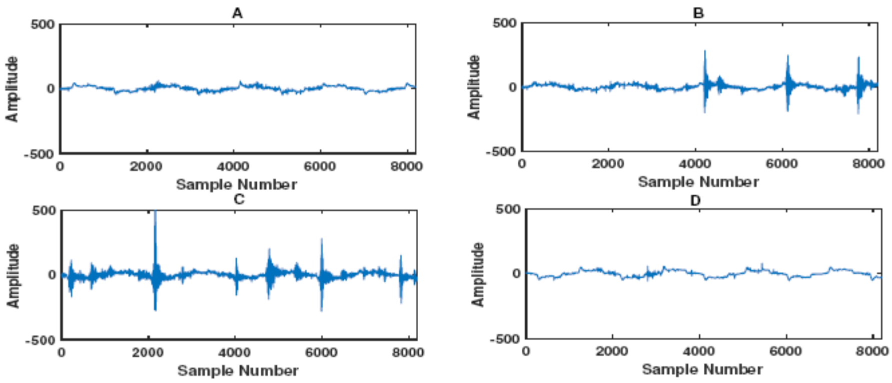

Figure 4: Examples of signals corresponding to four classes in dataset-1 A) Good condition, B)IRF, C) ORF, D) IORF

4.2. Gear Fault Dataset (Dataset-2)

The dataset-2 consists of 7 classes of normal and abnormal conditions such as A: 10% fault,B: 20% fault, C: 30% fault, D: 40% fault, E: 80% fault, F: 100% fault, G: expt no load. Thedata is sampled at a frequency of 8.2 kHz and sampling length of 2047. Figure 5 represents theexamples of signals corresponds to these 7 classes.

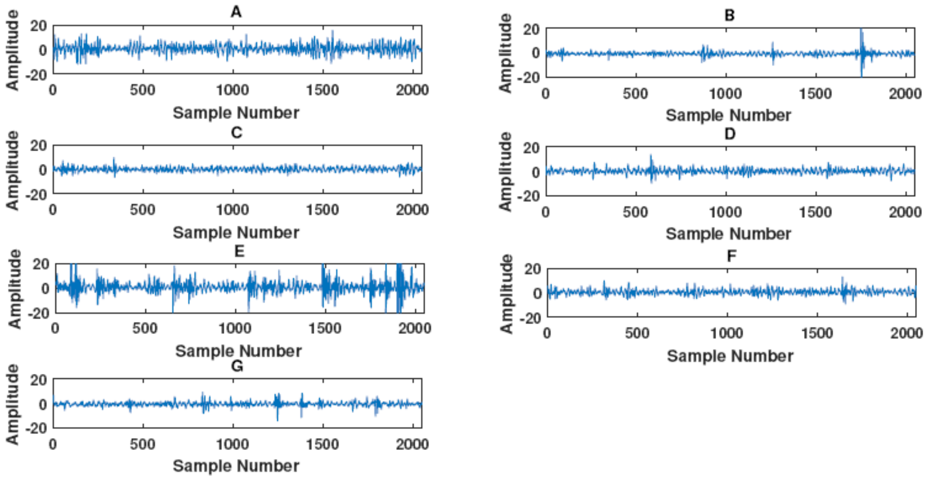

Figure 5: Examples of signals corresponding to 7 classes in dataset-2 A) 10% fault, B) 20%fault, C) 30% fault, D) 40% fault, E) 80% fault, F) 100% fault, G) expt no load

5. Proposed System for Machine Fault Diagnosis Monitoring

The simplified block diagram of the proposed system for machine fault diagnosis is given in figure 6. In the proposed system, the data acquisition stage contains the collection of vibrationsignals from machine fault simulator explained in Section 3. The vibration signals collected arebelongs to two categories: 1) signals corresponds to bearing faults and 2) signals correspondsto gear faults. Further the signals are fed to the feature extraction stage where the chebfunapproximation system is used for extracting the chebfun coefficients. The extracted chebfuncoefficients corresponds to signals in dataset-1 is shown in figure 7. Finally this coefficients aregiven to the trained RLS classifier to identify the good and faulty signals.

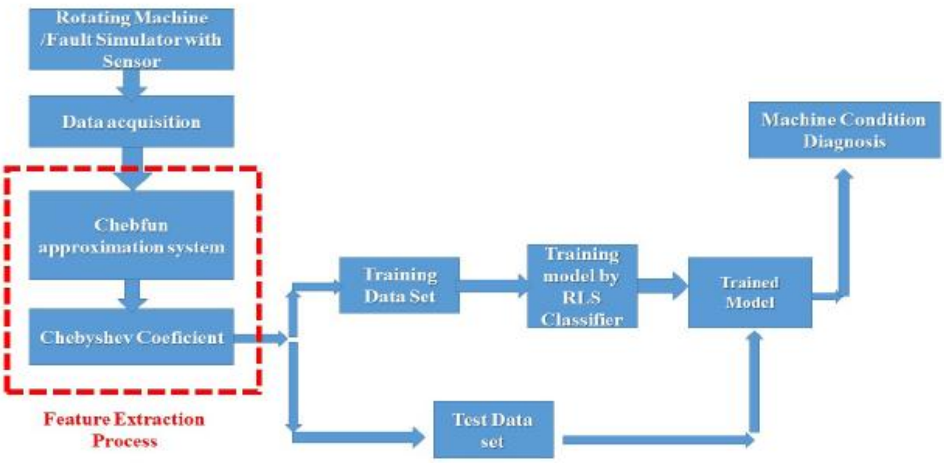

Figure 6: Block diagram of proposed machine fault diagnosis system

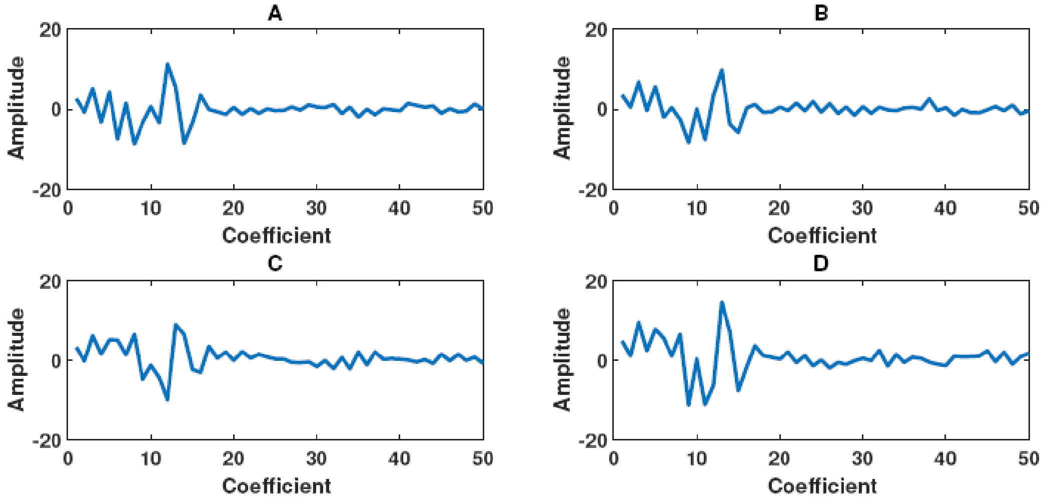

Figure 7: Examples of chebfun coefficients extracted as features using signals from dataset-1 A)Good, B) IRF, C) ORF, D) IORF

6. Results and Discussion

6.1. Evaluations using Bearing Fault Dataset (Dataset-1)

The first data set used for evaluation is bearing fault dataset (dataset-1). 100 signals from eachclass is generated and three splits of data are considered for evaluation. First is 20:80 data splitwhere 20% of data is taken for testing from all the classes and remaining 80% of data is taken fortraining the classifier. The second data split considered is 30:70, where 30% of data is taken fortesting purpose and rest 70% is used for training purpose. The third data split is 40:60, where40% of data is taken for testing and rest for training. In the feature extraction stage, chebfuncoefficients are extracted as features in 10, 20, 30, 40, 50 range and are given to RLS classifier.The performance is evaluated using two kernel functions such as linear and radial basis function(RBF). Two different cross-validation are also tried namely hold-out (HO) and leave-one-out(LOO).

Table 1: Classification accuracy for dataset-1 with different data split

Data Split	Kernel / Co-efficient	Accuracy (%)				
		10	20	30	40	50
20 : 80	Linear	97.50	98.54	98.03	98.75	95.00
	RBF _HO	97.50	98.12	95.00	97.50	98.33
	RBF _LOO	96.25	98.36	98.75	96.25	96.00
30 : 70	Linear	96.66	97.50	97.67	96.67	97.50
	RBF _HO	96.66	95.80	98.30	98.33	96.67
	RBF _LOO	96.66	97.50	98.33	96.67	97.50
40 : 60	Linear	96.87	97.50	95.62	96.25	97.50
	RBF _HO	96.87	96.87	96.85	96.25	98.12
	RBF _LOO	97.50	98.12	97.50	95.00	96.87

The classification accuracy for all the data split are given in Table 1. The results are noted foran average of 10 iterations. For 20:80 data split, the highest accuracy of 98.75% is obtained forLinear and RBF kernel. Also, it is observed that the accuracy is highest with 40 and 30 chebfuncoefficients as features. In 30:70 data split, RBF kernel with LOO and HO cross-validation hasattained the highest accuracy of 98.33% with 40 and 30 chebfun coefficients. For 40:60 splitting,the highest accuracy of 98.12% has attained with 50 and 20 chebfun coefficients for RBF kernel.

From this evaluations, it is clear that the proposed system has guaranteed an average accuracyof 98% for bearing fault discrimination.

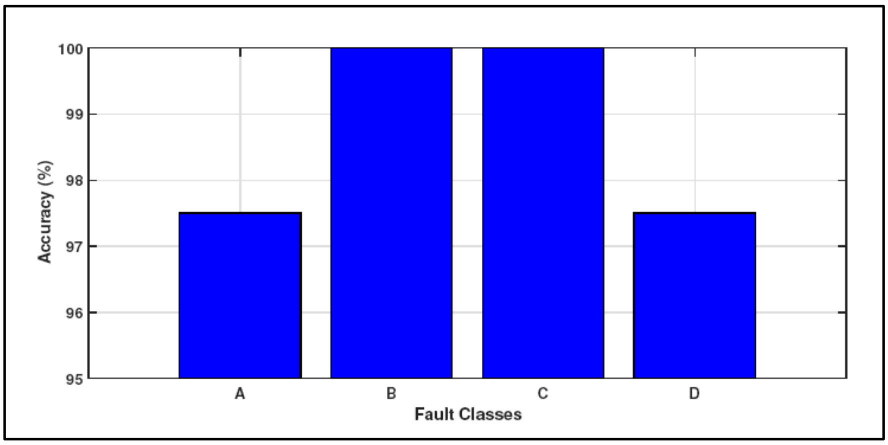

Figure 8: Illustration of class-wise accuracy for dataset-1 A) Good, B) IRF, C) ORF, D) IORF

The bar chart in figure 8 shows the class-wise accuracy obtained for dataset-1 in a singleiteration with linear kernel. Here, the faulty classes (IRF, ORF) are accurately classified with100% result. The good and IORF has attained an accuracy of 97.5%. The highest class-wiseaccuracy obtained confirms the superior performance of the proposed method.

6.2. Evaluations using Gear Fault Dataset (Dataset-2)

The second data set used for evaluation is gear fault dataset (dataset-2). It contains 7 differentgear fault conditions. 60 signals from each class is generated for evaluation. The three splitsof data (20:80, 30:70, 40:60) same as previously explained is considered. In this case also thechebfun coefficients are extracted as 10,20,30,40,50 range and are classified using RLS.classification accuracy for all the data split are tabulated in Table 2. The results are noted foran average of 10 iterations. For 20:80 data split, the highest accuracy of 98.80% is obtainedfor linear kernel with 30 chebfun coefficient as features. In 30:70 data split, linear kernel hasattained the highest accuracy of 96.03% with 20 chebfun coefficients. For 40:60 splitting, thehighest accuracy of 94.04% is achieved with 40 chebfun coefficient and linear kernel. From thisevaluations, it is clear that the proposed system has obtained a highest accuracy of 98% . Theclass-wise accuracy for a single iteration with linear kernel is shown in figure 9. From the figureit is evident that the proposed system is able to classify the fault conditions with very highaccuracy except for class B, (for class B (20% fault) accuracy is 92.75%.

Table 2: Classification accuracy for dataset-2 with different data split

Data Split	Kernel / Co-efficient	Accuracy (%)				
		10	20	30	40	50
20 : 80	Linear	91.66	96.42	98.80	97.61	95.23
	RBF _HO	91.66	92.85	90.47	97.61	91.60
	RBF _LOO	92.85	96.42	90.47	91.66	95.23
30 : 70	Linear	95.23	96.03	92.85	92.85	94.44
	RBF _HO	84.12	85.71	86.50	92.06	92.06
	RBF _LOO	88.50	84.12	92.06	87.30	86.50
40 : 60	Linear	89.55	91.66	82.73	94.04	89.88
	RBF _HO	77.97	86.90	78.57	82.73	81.54
	RBF _LOO	82.73	80.95	85.71	83.92	82.73

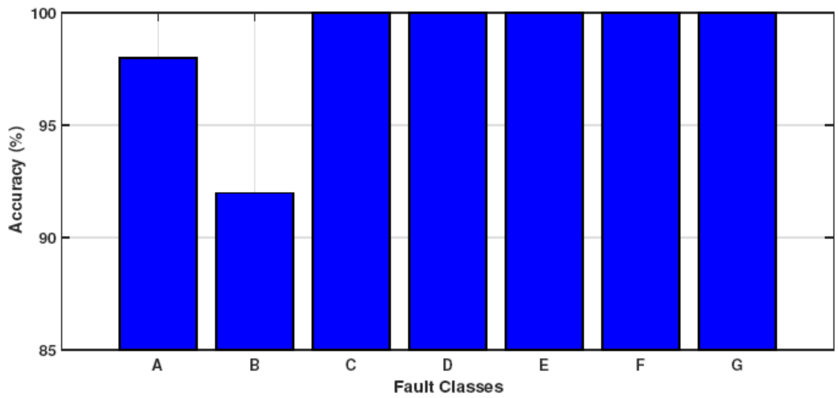

Figure 9: Illustration of class-wise accuracy for dataset-2 A) 10% fault, B) 20% fault, C) 30%fault, D) 40% fault, E) 80% fault, F) 100% fault, G) expt no load

From the experimental evaluations it is concluded that the proposed fault diagnosis systemis accurate. For both bearing and gear dataset, the proposed system has achieved the bestaccuracy with 20:80 data split using 30-50 chebfun coefficient features.

Table 3: Performance comparison with existing approaches in terms of accuracy

Existing Approaches	Fault Category	No. of Classes	Accuracy (%)
Stat + RLS-RBF [2]	Bearing	12	96.67
Cyclo + DT [1]	Bearing	4	94.00
Cyclo + Non-Prob. SMO Linear Polykernel [1]	Bearing	4	97.25
Cyclo + Non-Prob. SMO Linear Polykernel [1]	Bearing	4	96.75
Cyclo + RLS-RBF [1]	Bearing	4	96.45
Stat + eSVC-Linear [5]	Bearing	12	98.03
Stat + nuSVC-Linear [5]	Bearing	12	95.58
Stat+PSVM [8]	Bearing	4	95.00
Stat + J48 [17]	Gear	7	85.10
Stat + NN [18]	Ball Bearing	3	95.00
Chebfun + RLS-Linear (Proposed System)	Bearing	4	98.75
Chebfun + RLS-Linear (Proposed System)	Gear	7	98.80

6.3. Comparison with Existing Methods

The classification performance of the proposed machine fault diagnosis system is compared withvarious existing machine condition monitoring systems and are given in Table 3. The existingsystems used for comparison are stat-RLS [2], cyclo-DT [1], cyclo-SMO [1], cyclo-RLS [1], stat-cSVC [5], stat-nuSVC [5], stat-PSVM [8], stat-J48 [17] and stat-NN [18]. In stat-RLS [2] method,fault diagnosis is done using RLS algorithm with statistical features. In [1], vibration analysisis performed using cyclostationary features with decision tree (cyclo-DT), sequential minimumoptimization (cyclo-SMO), RLS (cyclo-RLS) algorithms. In stat-SVC [5] method, statisticalfeatures with c-SVC and nu-SVC classifiers is proposed for machine fault diagnosis. In stat-PSVM [8], statistical features are classified using proximal support vector machine (PSVM)for machine fault detection. In stat-J48 [17], a decision tree -J48 algorithm using statisticalfeatures with an initial feature selection procedure is adopted for classification purpose. In[18], a neural network (NN) system trained with statistical parameters are used for ball bearingclassification. From Table 3, it is clear that the proposed fault monitoring system achieved ahigher performance with the existing methods. This highlight that the proposed chebfun basedsystem can be used for real time machine fault diagnosis with state-of-the-art classificationaccuracy.

7. Conclusion

In this paper, the significance of chebfun approximation system in vibration signals for machinecondition monitoring is investigated. The vibration signals taken from two different machinefault setup namely bearing fault and gear fault are considered. The collected vibration signals aregiven to chebfun approximation system to extract chebfun coefficients as features. Further theextracted features are given to RLS classifier. The experimental evaluations are performed usingvarious kernel functions and chebfun coefficients. The proposed system has achieved an averageaccuracy of 98% for both bearing and gear fault classification. Also, the performance is comparedwith several other existing approaches and the results are found to be satisfactory. Hence, it isconcluded that the proposed machine fault diagnosis system using chebfun coefficients can beused as an effective tool for real-time machine condition monitoring.

References

[1] S Sachin Kumar, Neethu Mohan, PrabaharanPoornachandran, and KP Soman. Condition monitoring inroller bearings using cyclostationary features. In Proceedings of the Third International Symposium onWomen in Computing and Informatics, pages 690{697. ACM, 2015.

[2] Neethu Mohan, PS Ambika, S Sachin Kumar, MuruganSai M, and KP Soman. Multicomponent faultdiagnosis using statistical features and regularized least squares. International Journal of AppliedEngineering Research, 10(20):19074{19080, 2015.

[3] Luisa F Villa, AnibalRenones, Jose R Peran, and Luis J de Miguel. Statistical fault diagnosis based onvibration analysis for gear test-bench under non-stationary conditions of speed and load. MechanicalSystems and Signal Processing, 29:436{446, 2012.

[4] V Sugumaran and KI Ramachandran. Fault diagnosis of roller bearing using fuzzy classifier and histogramfeatures with focus on automatic rule learning. Expert Systems with Applications, 38(5):4901{4907, 2011.

[5] V Sugumaran, GR Sabareesh, and KI Ramachandran. Fault diagnostics of roller bearing using kernel basedneighborhood score multi-class support vector machine. Expert Systems with Applications, 34(4):3090-3098, 2008.

[6] M Amarnath, V Sugumaran, and Hemantha Kumar. Exploiting sound signals for fault diagnosis of bearingsusing decision tree. Measurement, 46(3):1250{1256, 2013.

[7] AthiraChandran, T Anjali, Neethu Mohan, and KP Soman. Overlapping group sparsity induced conditionmonitoring in rotating machineries. In International Conference on Soft Computing and PatternRecognition, pages 409{418. Springer, 2016.

[8] V Sugumaran, V Muralidharan, and KI Ramachandran. Feature selection using decision tree andclassification through proximal support vector machine for fault diagnostics of roller bearing. Mechanicalsystems and signal processing, 21(2):930{942, 2007.

[9] Zachary Battles and Lloyd N Trefethen. An extension of matlab to continuous functions and operators.SIAM Journal on Scientific Computing, 25(5):1743{1770, 2004.

[10] Rodrigo B Platte and Lloyd N Trefethen. Chebfun: a new kind of numerical computing. In Progress inindustrial mathematics at ECMI 2008, pages 69{87. Springer, 2010.

[11] James F Whidborne. Solving optimal control problems using chebfun. In 2016 UKACC 11th InternationalConference on Control (CONTROL), pages 1-6. IEEE, 2016.