The Pennsylvania State University

The Graduate School

A PROBABILISTIC FRAMEWORK FOR FAULT DETECTION IN INDUCTION MOTORS

A Thesis in

Electrical Engineering

by

Rohan Samsi

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

December 2006

The thesis of Rohan Samsi was reviewed and approved* by the following:

Jeffrey S Mayer
Associate Professor of Electrical Engineering
Thesis Co-Advisor, Co-Chair of Committee

Asok Ray
Distinguished Professor of Mechanical Engineering
Thesis Co-Advisor, Co-Chair of Committee

Heath F Hofmann
Associate Professor of Electrical Engineering

Kon-Well Wang
William E. Diefenderfer Chaired Professor in Mechanical Engineering

W. Kenneth Jenkins
Professor of Electrical Engineering
Head of the Department of Electrical Engineering

*Signatures are on file in the Graduate School.

Abstract

Online monitoring of induction motor health is of increasing interest, as the industrial processes that depend on these motors become more complex and as the performance to cost ratio of monitoring technology (e.g. sensors, microprocessors) continues to increase. Much effort has been directed towards developing methods that use conventional signal processing and pattern classification techniques. This thesis addresses the main issues of detecting electrical and mechanical faults using the information provided by current and vibration sensors, within a probabilistic framework. The faults studied in this work are stator voltage imbalances (electrical), rotor bar failures (electro-mechanical) and bearing outer race failure (mechanical). These failures are representative of almost all the failures occurring in induction machines. Voltage imbalance has been the longest studied problem. These arise either due to the faulty topology of the electrical system or stator winding faults. Rotor bar faults have accounted for about 10% of failures in induction machines, they are the most difficult type of failures to detect. Bearing failure accounts for 40% of all known failures. There is no direct instrumentation known to detect these kinds of failures.

The framework developed provides a common solution methodology for the detection of all these different faults. The methodology utilizes a combination of machine modeling concepts, along with wavelet, and symbolic dynamic analysis to ensure early detection. Additionally the sensor fusion technique developed, presents a probabilistic approach to the problem of bearing faults in induction motors. The method extends the D-Markov process to combine the information from both electrical and mechanical sensors. This provides accurate detection, with a low false alarm rate. The technique has been simulated using the magnetic equivalent circuit method and experimentally validated on 2-hp squirrel cage induction motors.

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to begin by thanking my family. My parents, brother and my wife. They have stood by me at every turn in my life. I thank them for their support, patience, understanding and belief in me. This work although meant to be an individual effort, would not have been possible without the active support of a lot of people. My advisors Profs. Ray and Mayer have only had words of encouragement for me. They made me understand that graduate school is more than just academics and research. It is about developing an attitude towards life, an attitude that lets you take on any challenge life throws at you. I am grateful for all their support. I would also like to thank my committee members Profs. Wang and Hofmann. Their constructive criticism has gone a long way in shaping my dissertation.

I would also like to thank Prof. Charlie Croskey for giving me an opportunity to be associated in a teaching capacity with the EE department. Teaching is an experience like no other, and I can safely say that I have learnt more than any of my students during my time as a teacher. Lastly, I would like to thank all my friends at Penn State. It has been a great experience and I doubt anything else will match it.

This work has been supported in part by the U.S. Army Research laboratory and the U.S. Army Research Office under Grant No. DAAD19-01-1-0646.

Chapter 1

Introduction

1.1 Motivation

Online monitoring of induction motor health is of increasing interest. There are many benefits to assessing the health of machinery in operation, such as reliability and productivity. A run-to-failure approach is acceptable for non-critical processes. But for critical components, reliability must be ensured. Unexpected breakdowns could have disastrous consequences. On the other hand, down-time for scheduled maintenance reduces productivity, but ensures reliability. Condition-based maintenance facilitated by online health monitoring is often considered the best solution.

There are several aspects to developing and implementing an online health monitoring system. Signals that are likely to manifest a relatively easily detected fault signature and that are readily acquired through sensors and supporting instrumentation must be identified. Basic signal processing techniques to enhance the fault signature and suppress the dominant system dynamics and noise must be considered. A detection or decision making process that follows the signal processing and quantifies the severity of the fault must be developed.

Each of these aspects presents significant challenges. For example, in an induction motor the magnetic flux distribution is probably the best indicator of stator winding fault. Monitoring the flux distribution would require a high-quality magnetic sensor to be mounted inside the motor. This solution, although effective, is not practical, because of the cost of the sensor would be prohibitively high and so would be the cost of installing it. A better solution would be to monitor the currents flowing into the machine and other such signals which can be easily acquired. The problem does not end with this; the acquired signal needs to be analyzed for fault information. There has been much research conducted in this area. Typically the fault signature embedded in the motor signals, such as the stator currents, are 50-80 dB smaller than the signals themselves. Moreover, these signals are load-dependent, and locally non-stationary in the frequency domain. To make matters worse, no two machines are identical in their characteristics, even if they are from the same assembly line. With the fault signals being of such small magnitude, manufacturing variations could be very easily mistaken for faults. Thus, a need arises for a solution that is able to address all the issues mentioned above. An ideal solution would be inexpensive, easily implemented, able to localize and quantify the fault information accurately, and able to accommodate manufacturing variations in the machines. The goal of this research was to approach this ideal.

1.2 Approach

A framework for on-line detection of induction machine faults is shown in Fig(1.1). Typically, only the stator currents are used as inputs, because the sensors to measure these signals do not require any modifications to be made to the motor. If one is willing to go through the trouble and expense of mounting accelerometers, then vibration signals can be used instead of or in addition to the current signals.

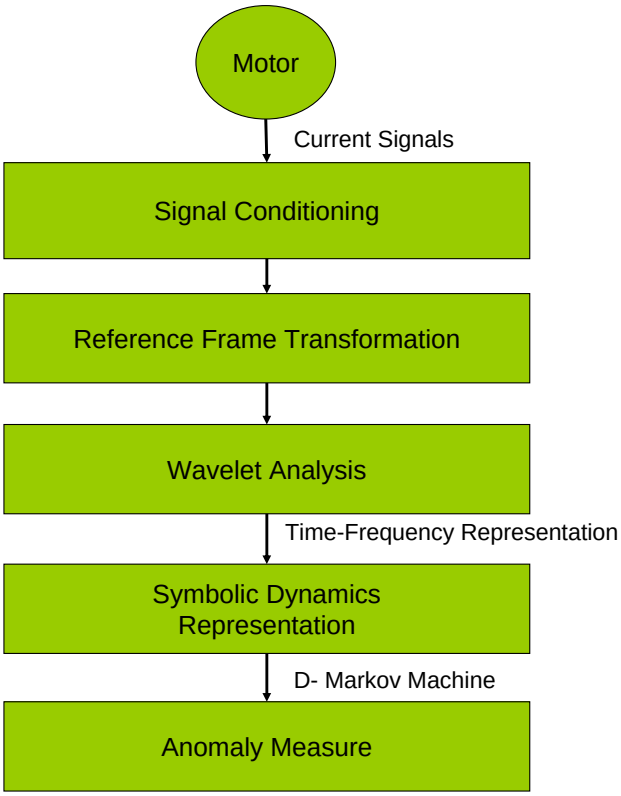

Figure 1.1. Framework

The Parks vector modulus is computed from the current signals to convert the instantaneous 3-phase stator currents, $\mathbb{R}^3$ vectors are converted to an orthogonal reference frame ($\mathbb{R}^2$) via Parks transformation and then into a single value ($\mathbb{R}$) via the modulus operator. The resulting time-series is then analyzed using suitable wavelet transforms. This is followed by a non-linear symbolic analysis and Markov machine construction. It allows for compression of information and gives an accurate measure of the fault. This fault/anomaly measure leads to an estimate of the health of the machine.

As with most useful frameworks, the individual processes can be selected and tuned more or less independently. Off-line analysis such as physics-based modeling

of the fault and power spectral analysis helps in selecting suitable vector transformations and the wavelet analysis to be performed. This is followed by partitioning the wavelet coefficients along the ordinate. The information from different scales of the wavelet can be included using a stacking process, which is described in the next chapter. Each partition is assigned a symbol, and thus the stator current signals corresponding to the fault are abstracted to a symbol sequence. The philosophy being that a change in the symbol sequence is an indication of fault. This change has to be quantified, this is done with the help of a D-Markov machine. This construction helps in associating a statistical behavior with the symbol strings. The change in the statistical behavior from the normal operating condition can be quantified using a standard metric to measure the severity of the fault.

1.3 Contributions

The primary contribution of this thesis is:

- A unified statistical framework for the detection of different faults in induction machines. Important features of this framework include:
  - It provides early detection with a low-false alarm rate.
  - Its implementation can be inexpensive.
  - It is applicable to any induction machine, irrespective of size and rating.
  - It is also load-independent.
  - Under the framework, a technique has been developed for sensor fusion to combine information from diverse sensors, to provide a more robust fault measure. It extends the properties of the D-Markov machine to merge information in the probability space.

1.4 Induction Motor Construction, Operation and Modeling

Induction motors are so-called because the currents in the rotor circuits are induced by the relative motion between these circuits and a rotating magnetic field produced by the stator winding currents. To expand on this explanation and to define terms used throughout this thesis, the construction, operation and modeling of the induction motor are described briefly in this section.

1.4.1 Construction

The stator consists of a core of stacked, insulated, iron laminations with Polyphase windings of insulated copper wire filling the slots in the core, as shown in Fig(1.2). The particular arrangement of the stator winding turns depends upon the intended application of the machine but in general, the turns in each phase are distributed so as to produce an airgap magnetic motive force (mmf) that is distributed symmetrically around one or more magnetic axes or poles. For a three-phase two-pole machine, the respective magnetic axes of a_s, b_s and c_s phases are separated by 120° as shown in Fig(1.2). The rotor, like the stator, consists of a core of stacked, insulated, iron laminates. Either insulated copper windings or aluminum bars fill the slots. The distinction between the use rotor windings and rotor bars leads to the common classification of induction machines as either wound rotor or squirrel cage.

- *Wound rotor*: In this case, the rotor winding is made of wire and has the same number of poles as the stator winding. Each phase of the rotor winding connected to a slip ring on the shaft. Carbon brushes connect the slip rings to an external circuit such as a variable resistor that permits changing the

motor's slip rate. Compared to squirrel cage rotors, wound rotor motors are more expensive and require much more maintenance. The ability to change the rotor circuits in operation, however, allowed the wound rotor motor to be used in variable speed applications before the advent of compact power electronic devices.

- *Squirrel cage rotors*: The squirrel cage takes its name from its shape - a cylindrical cage comprised of axial bars terminated in annular rings. The bars and endrings are typically cast aluminum. The vast majority of the rotor currents flow through the bars rather than the higher-resistance and usually varnished laminates. Very low voltages and very high currents are typical in the bars and end rings.

1.4.2 Operation

The operating speed of an ac induction motor is related to the synchronous velocity of the rotating magnetic field produced by the stator winding currents and to the slip of the rotor winding necessary to induce rotor currents to produce torque sufficient to drive the mechanical load.

The synchronous velocity is related to the frequency of the ac supply and the number of poles in the stator winding by the following equation:

$$N_s = \frac{120F}{P} \tag{1.1}$$

where

N_s = Synchronous speed, in revolutions per minute

F = AC power frequency in Hz

P = Number of poles per phase winding

Actual RPM (revolutions per minute) for an induction motor will be less than this calculated synchronous speed by an amount known as slip that increases with the torque produced. The slip of the AC motor is calculated by:

$$S = (N_s - N_r)/N_s$$

where

N_r = Rotational speed, in revolutions per minute (RPM)

S = Normalized slip, 0 to 1

The rotor velocity matches the synchronous velocity only if there is no mechanical load. When loaded, standard motors have between $2 - 3\%$ slip, special motors may have up to 7% slip, and a class of motors known as torque motors are rated to operate up to 100% slip (0 RPM/full stall).

In operation, the induction motor can be viewed as a transformer with a rotating secondary - when the rotor is not rotating in synchronism with the magnetic field, large rotor currents are induced; the large rotor currents magnetize the rotor and interact with the stator's magnetic fields to bring the rotor into synchronization with the stator's field. An unloaded induction motor at synchronous speed will only consume electrical power to maintain rotor speed against friction and resistance losses; as the mechanical load increases, so will the electrical load - the electrical load is inherently related to the mechanical load. This is similar to a transformer, where the primary's electrical load is related to the secondary's electrical load. Induction motors employ an armature that rotates within a fixed stator, with a small air gap between the two.

1.4.3 Mathematical Model of the Three-phase Induction Motor

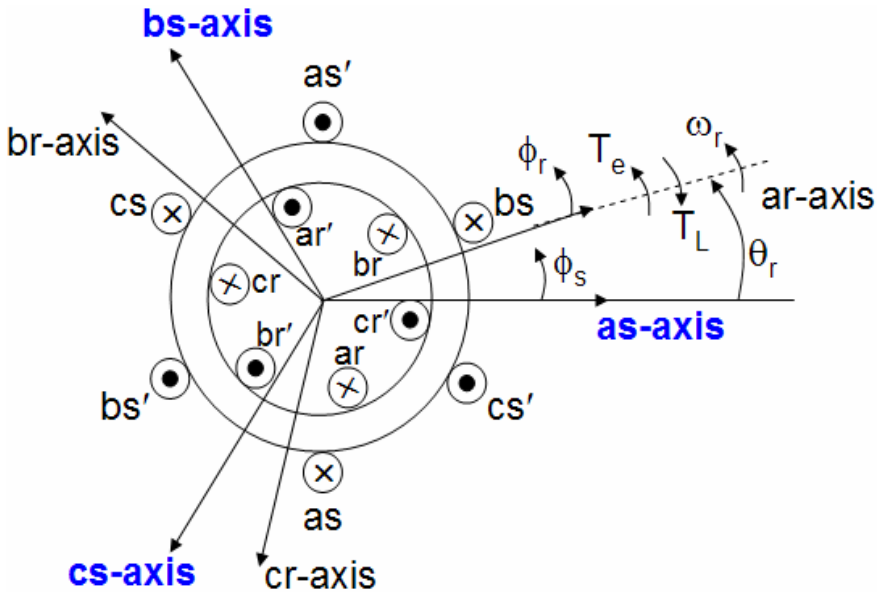

Figure 1.2. A 2-pole, 3-phase Induction Motor

When modeling and induction motor it is common to abstract away the details of the stator winding distribution. Fig(1.2) gives the representation of a 2-pole, 3-phase induction motor. The three phases are denoted by a, b and c, the stator and the rotor variables as s and r respectively. T_e and T_L are the electromagnetic torque and the load torque. ω_r and θ_r are the rotor velocity and position. The three-phase stator and rotor windings are shown in Fig(1.3). The stator voltage equations are expressed:

$$v_{as} = i_{as} r_s + \frac{d\lambda_{as}}{dt}$$

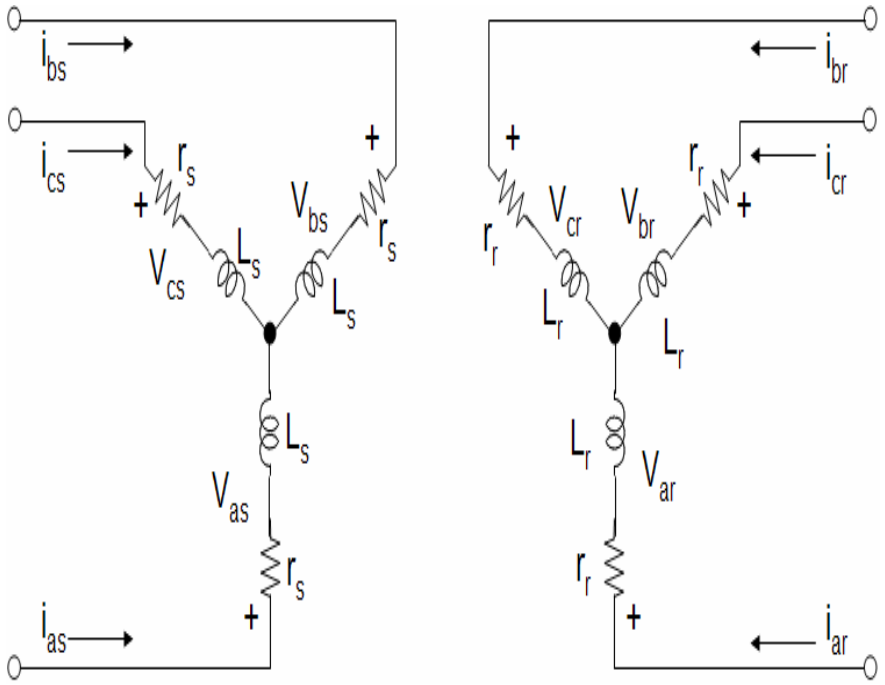

Figure 1.3. Lumped parameter models of the Stator and Rotor circuits

$$\begin{aligned} v_{bs} &= i_{bs}r_s + \frac{d\lambda_{bs}}{dt} \\ v_{cs} &= i_{cs}r_s + \frac{d\lambda_{cs}}{dt} \end{aligned}$$

The rotor voltage equations are expressed:

$$\begin{aligned} v_{ar} &= i_{ar}r_r + \frac{d\lambda_{ar}}{dt} \\ v_{br} &= i_{br}r_r + \frac{d\lambda_{br}}{dt} \\ v_{cr} &= i_{cr}r_r + \frac{d\lambda_{cr}}{dt} \end{aligned}$$

For a squirrel cage rotor, or a wound rotor with shorted windings the voltages applied to the rotor circuits are zero. Since the machine is assumed to symmetrical, all mutual inductances L_{ab}, L_{bc} and L_{ca} are assumed to be identical. The flux

linkages can be expressed as:

$$\begin{bmatrix} \lambda_{as} \\ \lambda_{bs} \\ \lambda_{cs} \\ \lambda_{ar} \\ \lambda_{br} \\ \lambda_{cr} \end{bmatrix} = \begin{bmatrix} \mathbf{L_{ss}} & \mathbf{L_{sr}}(\theta_\mathbf{r}) \\ \mathbf{L_{sr}^T}(\theta_\mathbf{r}) & \mathbf{L_{rr}} \end{bmatrix} \begin{bmatrix} i_{as} \\ i_{bs} \\ i_{cs} \\ i_{ar} \\ i_{br} \\ i_{cr} \end{bmatrix}$$

where

$$\mathbf{L_{ss}} = \begin{bmatrix} L_s & L_{ab} & L_{ab} \\ L_{ab} & L_s & L_{ab} \\ L_{ab} & L_{ab} & L_s \end{bmatrix}$$

and

$$\mathbf{L_{rr}} = \begin{bmatrix} L_r & L_{ab} & L_{ab} \\ L_{ab} & L_r & L_{ab} \\ L_{ab} & L_{ab} & L_r \end{bmatrix}$$

These inductances do not change with a change when the motor is in operation. While

$$\mathbf{L_{sr}}(\theta_\mathbf{r}) = \begin{bmatrix} L_{sr}cos(\theta_r) & L_{sr}cos(\theta_r + \frac{2\pi}{3}) & L_{sr}cos(\theta_r - \frac{2\pi}{3}) \\ L_{sr}cos(\theta_r - \frac{2\pi}{3}) & L_{sr}cos(\theta_r) & L_{sr}cos(\theta_r + \frac{2\pi}{3}) \\ L_{sr}cos(\theta_r + \frac{2\pi}{3}) & L_{sr}cos(\theta_r - \frac{2\pi}{3}) & L_{sr}cos(\theta_r) \end{bmatrix}$$

gives the mutual inductance between the stator and the rotor windings, which does depend on the rotor position.

This three-phase model can be transformed into a different form using a so-called as a reference-frame transform as shown in Eqn(1.2). In this particular

case, we use the stationary reference frame transform. Sometimes it is useful to transform these variables to a rotating reference frame. In the case of a three-wire system, with a floating neutral, the i_0 or zero sequence currents are identically zero. All of the currents can be expressed as the direct axis i_d and the quadrature axis i_q currents. This transformation can be extended to voltages and flux linkages.

$$\begin{pmatrix} i_d \\ i_q \\ i_0 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ 0 & \frac{\sqrt{3}}{3} & -\frac{\sqrt{3}}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} i_a \\ i_b \\ i_c \end{pmatrix} \tag{1.2}$$

Using the Parks vector transform, the electromagnetic torque generated by a three-phase induction motor is derived in Eqn(1.3).

$$T_e = \frac{3P}{4}\frac{M}{L_r}(\lambda_{rd} i_{sq} - \lambda_{rq} i_{sd}) \tag{1.3}$$

The model developed in this section is useful when considering the operating characteristics of the induction motor. In other forms, for example in the state-space form, it can be used for control applications. However, for the purpose of fault detection, taking a model based approach could prevent the early detection of failures. Besides, most of the machines used in industry are not truly symmetric. Many of their windings are short-corded to save on material. In addition, manufacturing variations also can loss of symmetry.

1.5 Induction Motor Failure Modes

The most common failure modes of the induction motor are associated with the stator windings, the rotor cage and the bearings [1] [2]. The distribution among these failure modes is summarized in Fig(1.4).

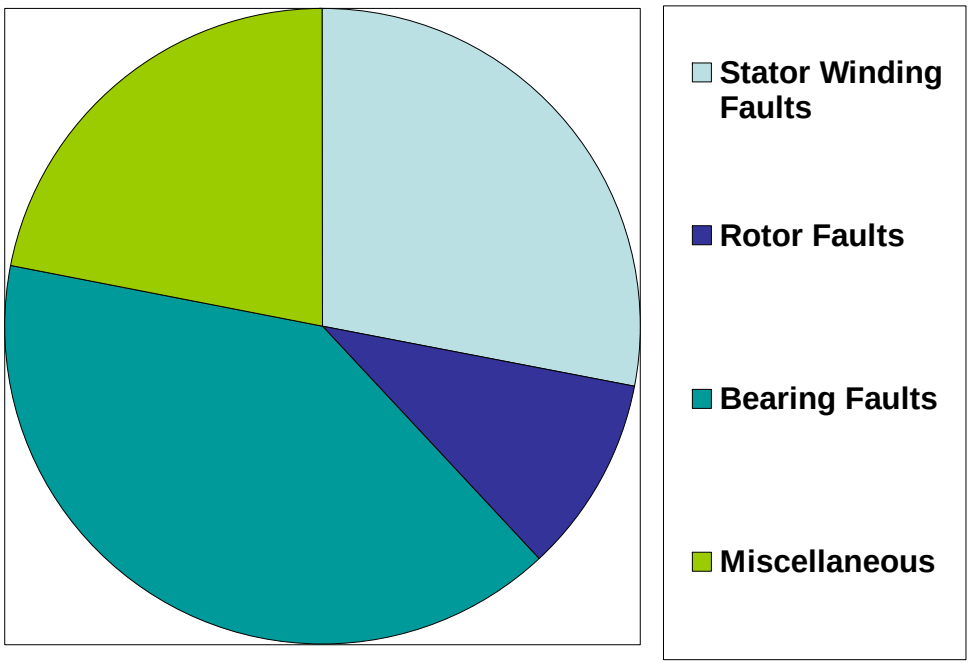

Figure 1.4. A statistic on the various failures in induction motors

1.5.1 Stator Winding Faults

Stator winding faults ultimately involve the short circuiting or open circuiting of the stator windings. Short circuits are more common and stem from the failure of winding insulation between turns within a phase coil, between turns coils of different phases, or between a turn and the stator core. Failure of stator winding insulation is usually the result of various stress, which can be divided into four categories: thermal, electrical, mechanical and environmental.

As a rule of thumb, for every 10° C increase in operating temperature above rated, the insulation life is halved. Typical causes of elevated operating temperature include: overloading, unbalanced phase voltages, obstructed ventilation and

high ambient temperature. Voltage levels in excess of the winding rating will eventual lead to breakdown of the dielectric. Typical causes of excessive voltages include: switching induced overvoltage, lightening and variable frequency drives. Relative motion between neighboring turns or between coil bundles and the stator can lead to abrasion of the insulation. Relative motion of the conductors can result from magnetic forces, vibration or shock. Many environmental factors can lead to stator winding faults. Moisture, chemicals, abrasion and foreign objects cause the most damage.

1.5.2 Rotor Cage Faults

The rotor cage is subject to much higher currents and much lower voltages than the stator winding and is usually not insulated. Consequently, the most common failure mode is an open or broken rotor bar. The various stresses that lead to a broken rotor bar(s) can be divided into three categories: thermal, mechanical and residual.

The causes of elevated operating temperature for the stator winding identified previously are exacerbated for the rotor was due to the relatively poor heat transfer from the rotor. Moreover, the construction of rotor and rotor cage tends to produce some extreme temperature distribution. The rotor bars are subjected to magnetic forces, vibration and shock to an even greater degree than the stator windings. Stresses can be present in any plane (radial, axial or tangential) and, are normally not harmful to the rotor as long as they do not cause any significant change in the rotor geometry. If any of these stresses do result in a change in the rotor geometry, they usually take place during the transition between idle and full-load thermal conditions and cause vibration problems, which might not be noticed running no load.

1.5.3 Bearing Faults

Bearings are common to all machines. Rolling element bearings have several components, which we shall study in a later chapter. Each components is prone to failure. There are several causes for bearing fault, some of them are enumerated below.

1. Mechanical damage: Abusive handling can results in nicks and dents that are harmful, particularly when located in regions tracked by rolling elements. Displaced metal particles generated by nicks and scuffing-type damage introduce secondary effects when they dislodge and indent the raceway. Permanent indentation created by rolling element overload is called brinelling, a type of damage that usually results in failure. Brinelling may occur by an overload mechanism, such as dropping the bearing or improper mounting techniques.

2. Wear damage: Wear generally results in gradual deterioration of bearing components, which in turn leads to loss of dimensions and other associated problems. Failure by wear does not mean that bearings will be removed solely because of change in fit or clearances. Secondary conditions arising from wear can become predominant failure mechanism. Stress raisers could be generated that may sever as sites for crack initiation.

3. Lubricant deficiency: Lubrication problems may be associated with an inadequate lubricant, an inefficient lubricating system, or a combination of these conditions. Ideally, rolling elements are separated from the raceways by a film of lubricant, thus minimizing wear of bearing components. Initial stages of wear involve the plastic deformation of grinding furrow asperities, which, in subsequent cold-working, fracture to produce extremely fine platelets of

steel. These cold-worked, hard particles, which also contain carbides, serve as abrasive media. After sometime, original grinding furrows in the rolling element tracks are worn smooth to produce a glazed condition. Continued operation will generally lead to a deterioration that manifests itself as a frosted condition and, sequentially, to scuffing.

4. Crack damage: Cracking of bearing components may originate as a function of operating stress conditions via overload or cyclic loading (fatigue). Additionally, manufacturing-related cracks may derive from the steel making process and/or working process. Improper heat treatment and grinding are are frequently the causes of cracks.

5. Corrosion damage: Corrosion initiated failures are often difficult to recognize in bearings. Applications involving moisture-laden environments may result in surface oxidation and produce rust particles and pits. These particles are potential media for producing rapid wear via abrasion, and the pits can often become sites for crack initiation. Bearings that operate in an environment where water is absorbed in the lubricant maybe subjected to pitting corrosion by hydrolysis.

6. Electric Arc Damage: Improperly grounded electrical equipment can subject bearings to mechanical and metallurgical damage. When current seeking ground passes through bearings, it tends to arc between non-contacting rolling elements and raceways. Electric arcing also causes metallurgical properties to be altered to significant depths, since it effects intense, highly localized temperatures which melt the surface.

1.5.4 Complex Failure Modes and Detection

When analyzing failures, it is often difficult to identify which of the conditions listed previously was the cause of the initial problem. A simple example illustrates this point. A random wound motor is frequently started, and due to excessive coil movement, it sustains a minor turn-to-turn short within one coil. As this condition progresses, excessive heating is generated within the shorted coil, resulting in an insulation deterioration and, eventually, in a partial short to ground through the slot liner. Depending on the type of motor protection, the motor may continue to run thereby increasing the damaged area until the phase or ground insulation is destroyed. At this point, a direct phase-to-phase fault or ground fault occurs, and the motor is quickly dropped off the line.

Inspection could reveal all five modes of failure, but the turn-to-turn condition was the initial problem, and the others were the result of the problem. A turn-to-turn failure is usually very difficult to recognize due to the destructive nature of the final fault conditions. Additionally, it is hard to detect failures in the signals due fault levels being extremely small and non-stationary as shown in Fig(1.5). In this figure two runs of the same motor with a broken rotor bar are shown. This is where the wavelet transforms are used to extract the fault information as shown in Fig(1.6). However, wavelet transforms can also sometimes miss this fault information due to inappropriate scale selection or variation in operating conditions. In addition due to variations during manufacture two seemingly identical motors might have a slight variation in their behavior.

1.5.5 Simulation of Failure Modes in the Laboratory

The preceding study of failure modes are simulated using three experiments in the laboratory listed below:

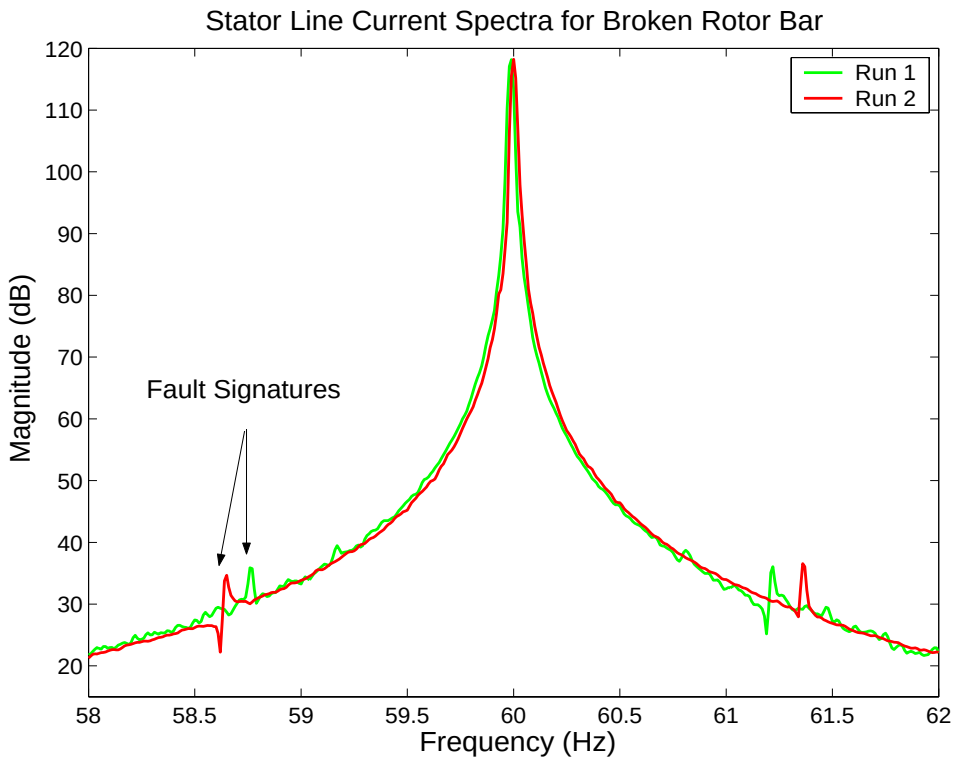

Figure 1.5. Variations in Fault Signatures for Broken Rotor Bars

- Stator Voltage Imbalance: These are caused by unbalanced stator voltages. It has also been used to simulate, insulation failure between windings in the stator; which has similar characteristics [3].

- Rotor Bar Breakage: A loss of one of the rotor bars causes an uneven current distribution. This eventually leads to failure. This is the hardest fault to detect, since the fault location is not accessible, and thus hard to instrument. Damage was introduced by drilling a hole into the rotor surface at a bar location.

- Outer Race Faults (Bearings): Bearing faults can occur in different areas of the bearing. However, all these faults manifest themselves in a similar manner. In this thesis, we shall look at faults concerning the outer race of

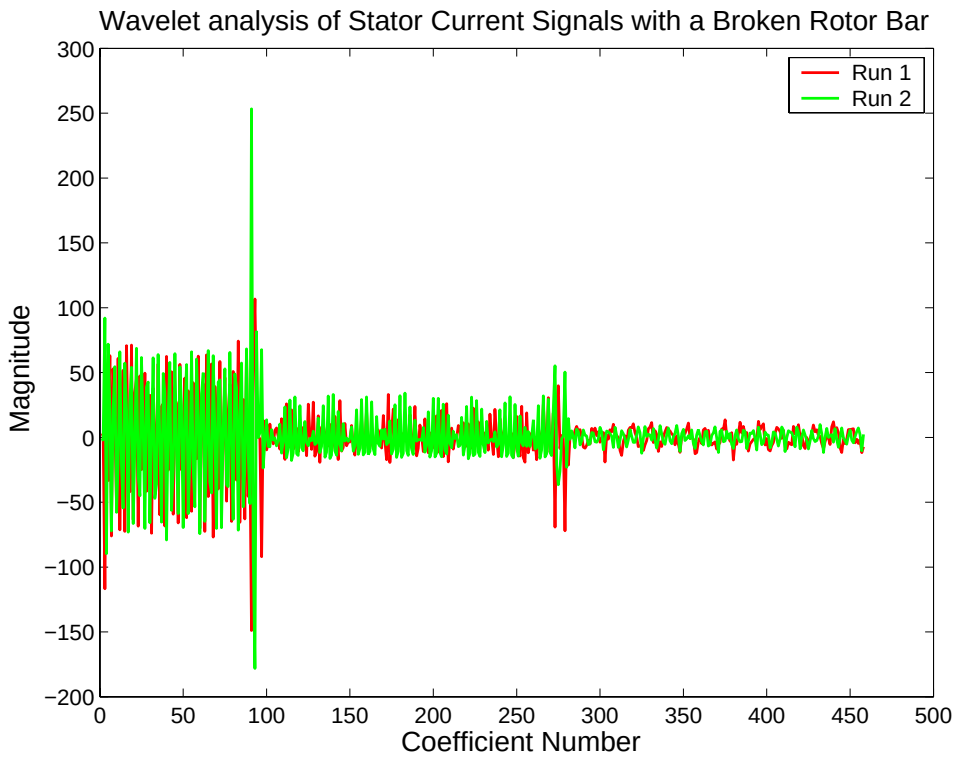

Figure 1.6. Wavelet transforms are used to extract the non-stationary waveform

the bearings alone. The damage was introduced mechanically, using drills and end-mills.

1.6 State-of-the-Art in Induction Motor Health Monitoring

Most health monitoring that have been proposed or developed commercially can be classified into two broad types: model based and non-model based techniques, as shown in Fig(1.7).

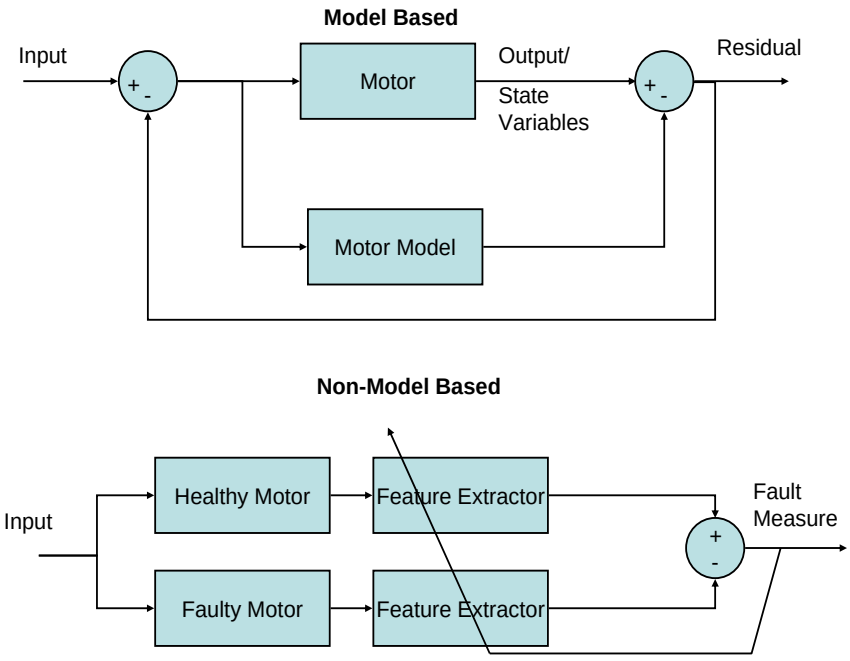

Figure 1.7. Comparison of Model Based and Non-Model Based Techniques

1.6.1 Model-Based Techniques

Generally, these techniques are more effective at detecting a failure after it has occurred than during an incipient stage. A residual is generated based on the difference between the model and the actual behavior of the motor. Once the residual crosses a certain limit, a fault is said to have occurred. While simple and straight-forward to implement, this technique is limited by the nature of the model. Models usually do not account for slight imperfections during manufacture, or even intentional asymmetries. They also do not account for supply imbalances and other environmental conditions, which often exist during normal operation.

Some of these techniques are given below:

- *Luenberger Observers and Extended Luenberger Observers* [4]: These observers are used to estimate the rotor flux which is then used to calculate the

electromagnetic torque developed. During faulted operation, the frequency characteristics of this torque will be different from the nominal condition.The extended Luenberger observer can also be used for a non-linear model of the motor. The poles of the estimation dynamic matrix are chosen to stabilize this estimation. This model is observable as long as the supply frequency is not zero.

- *Extended Kalman Filters* (EKF) : Induction motors are often described mathematically by sixth order state-space equations with four electrical state variables (currents and fluxes), a mechanical state variable (rotor velocity) and a parameter (rotor resistance), and two control input variables (stator voltages) [5]. The hypothesis on which detection is based is that the apparent rotor resistance of an induction motor will increase when a rotor bar breaks. To detect broken bars, measurements of stator voltages and currents are processed by an Extended Kalman Filter (EKF) for the simultaneous estimation of rotor speed and rotor resistance. In the estimated value of rotor resistance exceeds the nominal value by a prescribed amount, it is assumed that a broken bar has been detected. In the EKF approach, the state covariance matrix is adequately weighted, leading to a better states estimation dynamic. Its main advantage is the correct rotor resistance estimation even for an unloaded induction motor. As part of this estimation process, it is necessary to compensate for the thermal variation in the rotor resistance. In fact, a difficulty with broken rotor bar detection is that a variation in the rotor temperature can cause significant variation in rotor resistance. In this case, the rotor resistance, so determined, must be referred and compared to the same temperature in normal operating point. The main drawbacks with this method are that it is not able to differentiate between load changes

and parameter changes, hence there is a significant chance of false alarms. Clearly, temperature changes are also not accounted for. It also does not have the ability to detect faults early enough. Similar methods have been used to estimate the time period for bearing fault identification [6].

- *Recursive least squares*: Recursive least squares techniques have been used repeatedly in the past for the purposes of parameter and state estimation [7]. Some researchers have used these techniques in sensor speed control techniques as well [8]. For fault detection, certain parameters are identified and the variation in these parameters is then tracked. For certain kinds of faults, there will be large variations in some parameters [9]. The disadvantage, in addition to the ones faced by EKF, is the fact that the modeling of the machine does not allow for all kinds of faults to be accommodated, these including non-parametric faults. A good comparison for all such methods applied to the problem of bearing fault detection is given in [10].

1.6.2 Non-Model-Based Techniques

Non-model based techniques rely completely on sensor data for fault detection. The research presented in this thesis can be associated with these techniques. These techniques do not rely on models which are typically low order, to keep the analysis simple. Non-Model-based techniques compare a feature that is believed to carry fault information of each, a healthy and a faulty motor. The difference in the features is quantified as a fault/anomaly measure. By doing this higher order dynamics and design imperfections become a part of the analysis. However, these techniques are often plagued by the problems of computational complexity and a higher risk of false alarms as compared to model based techniques. Some of these techniques used to date are:

- *Negative sequence impedance detectors* [11]: Based on the measurements of currents and voltages, effective negative sequence values are calculated. These then are trended. If the negative sequence shows an increasing trend, it is assumed that the motor is deteriorating rapidly. However, experimental data shows that the faults are detected only after they have increased beyond 0.28 pu.

- *Voltage mismatch detectors* [12] [13]: To supplement negative sequence currents deterioration, this method uses voltage mismatches to detect incipient failure in the stator winding:

$$\begin{aligned} V_{a1} &= z_{11} I_{a1} + z_{12} I_{a2} \\ V_{a2} &= z_{21} I_{a1} + z_{22} I_{a2} \end{aligned}$$

  The nominal way of detecting faults is to measure the values of z_{11}, z_{12}, z_{21} and z_{22} by conducting two tests under regular operating condition. If internal deterioration, such as an inter-turn coil fault or phase-to-phase leakage, has developed, the z_{xy}parameters will have changed from their normal values at that speed and there will be a mismatch between the measured positive and negative sequence voltages and the corresponding calculated values. This mismatch may then be used as a measure of internal motor deterioration. The disadvantage of this method is that it is extremely sensitive to load and speed conditions. It is consistent at constant load and constant speed.

- *Signal spectral analysis* [14]: When broken rotor bars are present, harmonic fluxes are produced in the air gap which introduce harmonic components in the motor current waveform. The motor waveform can be readily converted from a time domain to the frequency domain using Fourier analysis, and the

amplitude of each of the frequency components can be evaluated to identify problems in both the motor and the driven equipment. A significant drawback is the susceptibility to noise and the non-stationarity of the signals. The signals remain stationary only if the operating conditions load, velocity and temperature remain constant. Also, the fact that the fault information is contained in the side-band of frequencies, whereas the main frequency is at least 50 dB higher than the side-bands, makes it extremely hard to extract this information. There are various other ways of processing these signals, the most popular among these are wavelets, these have been explored in [15] [16].

- *Fuzzy Logic and Neural Networks*: Although these techniques have been used in the past for field-oriented control [17], recently they have been applied to fault detection as well. Under the fuzzy logic observer scheme, a number of local linear observers are designed and the state estimation is given by a fuzzy fusion of local observer outputs [18]. The diagnostic signal - a residual is the difference between the estimated and the real system outputs. Chow et al. [19] [20] have used several techniques using neural networks. These include radial basis functions, multi-layer perceptrons and neural/fuzzy systems. Some of these methods are compared in the results section. These methods suffer from the problem of excessive computational sensitivity. For bearing faults, these methods have been combined with wavelet processing to better extract the signal [21].

1.7 Outline

This thesis is comprised of six chapters including this one. In the interest of brevity, this thesis presents the results in a staggered manner, as explained below. Chapter 2 describes in detail the framework developed and its individual processes using a simple example of mixed sinusoids to illustrate key points. Chapter 3 has the results obtained in detection of stator voltage imbalances. In this chapter, we begin by looking at simulation results obtained from MEC-Magnetic Equivalent circuits, and imbalance created in the machine. The framework is validated for individual values of imbalance. Nuances of the wavelet analysis, particularly the difference between the DWT and CWT, are considered. Chapter 4 discusses the specifics of the detection of broken rotor bars from a statistical point of view. Changes in the D-Markov machine structure are explored. Chapter 5 looks at the detection of bearing faults, using a sensor combination approach. In this chapter, aspects such as convergence are also discussed. Chapter 6 summarizes the achievements and shortcomings of the framework. It also provides directions for further research.

Chapter 2

Methodology

This chapter describes the framework in detail. Often, only a small part of the energy of the signals to be considered (current or vibration) can be attributed to fault conditions. There is also substantial level of noise and unmodeled dynamics associated with the underlying signal. Conceptually, signal processing for anomaly detection can be classified in two categories: (i) isolation of signal features that have a high degree of correlation with one or more anomalies; and (ii) application of a decision-making algorithm to classify the set of measured signal features to identify a particular type of anomaly or the absence of an anomaly. These two categories may not be clearly distinguishable in many fault detection and identification techniques such as neural networks [22]. The problem of feature extraction from time series data for machine health monitoring has been addressed by many researchers [4] [23]. We shall use a similar pattern classification approach, applied to our specific faults.

The theme of anomaly detection is built upon the concepts of Symbolic Dynamics [24] Finite State Automata [25], and Pattern Recognition [26] as a means to qualitatively describe the (fast-time-scale) dynamical behavior in terms of symbol sequences [27]. Appropriate phase-space partitioning of the dynamical system

yields an alphabet to obtain symbol sequences from time-series data [28]. Then, tools of computational mechanics [29] are used to identify statistical patterns in these symbolic sequences through construction of a (probabilistic) finite-state machine from each symbol sequence. Transition probability matrices of the finite-state machines, obtained from the symbol sequences, capture the pattern of the system behavior by information compression. For anomaly detection, it suffices that a detectable change in the pattern represents a deviation of the nominal behavior from an anomalous one. The state probability vectors, which are derived from the respective state transition matrices under the nominal and an anomalous condition, yield a vector measure of the anomaly, which provides more information than a scalar measure such as the complexity measure [30].

Inferences are made regarding the occurrence of *slow-time-scale* anomalies (e.g., breakdown in stator winding insulation and rotor bar failure) based on changes in the behavior of the *fast-time-scale* process (e.g., dynamics of the stator current signal).

The framework relies on this separation of time-scales as is shown in Fig(2.2) and makes two simple assumptions:

- *Assumption 1*: Failure occurs over a long period of time and is monotonically increasing with time.

- *Assumption 2*: Fault dynamics are stationary during at the monitoring time scale.

The first assumption illustrates the inability of the framework to identify faults that manifest over a short period of time, instead the framework developed addresses faults that take many months or years to develop. This assumption in itself allows us to detect faults early in the life-cycle of the induction motor. The monotonicity of the failure assumes that the machine is not repaired during the monitoring

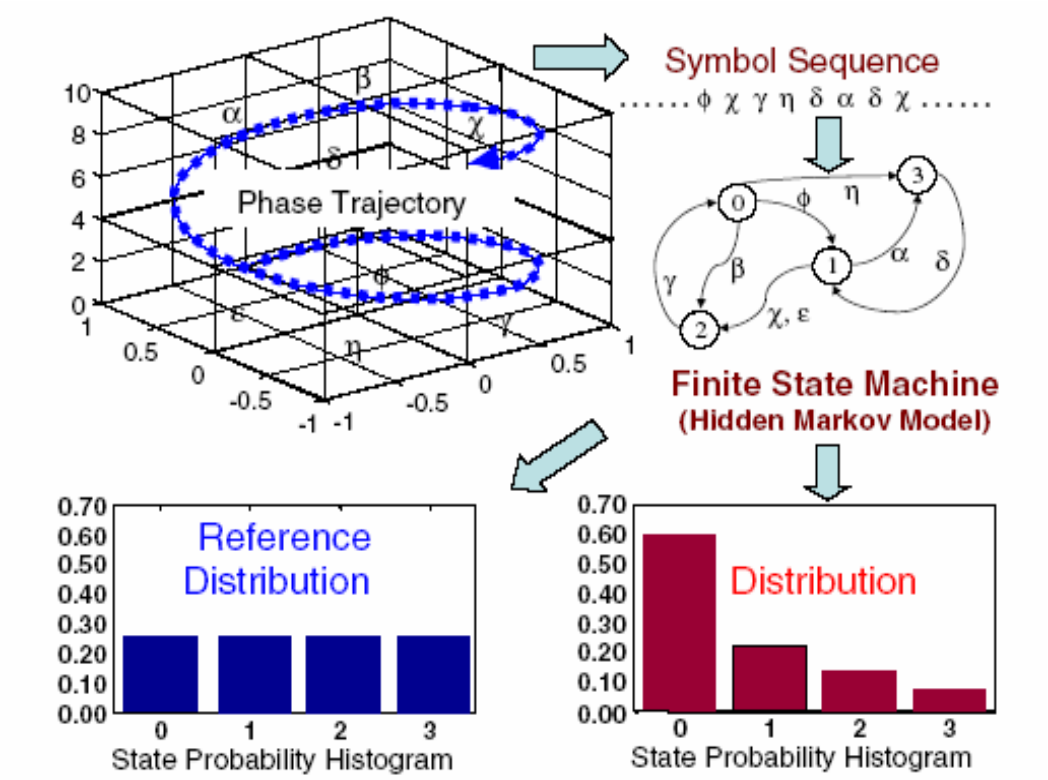

Figure 2.1. Continuous dynamics to statistical distribution

periods. This assumption is not valid if the damaged area is repaired. The second assumption is motivated by the fact that failure occurs over a long period of time. Thus although the fault process maybe a continuous process and increasing in the slow-time / failure time scale (T) we can safely assume that it is stationary or non-increasing in the monitoring time scale (τ). The framework relies on collecting data at periodic intervals for short amounts of time (τ) and comparing this with the healthy machine as the fault slowly manifests. Symbolic dynamics is used to study the changes in patterns between several time epochs. A change in the patterns gives an indication of an anomaly or fault.

Application of symbolic dynamics to anomaly detection in complex systems has

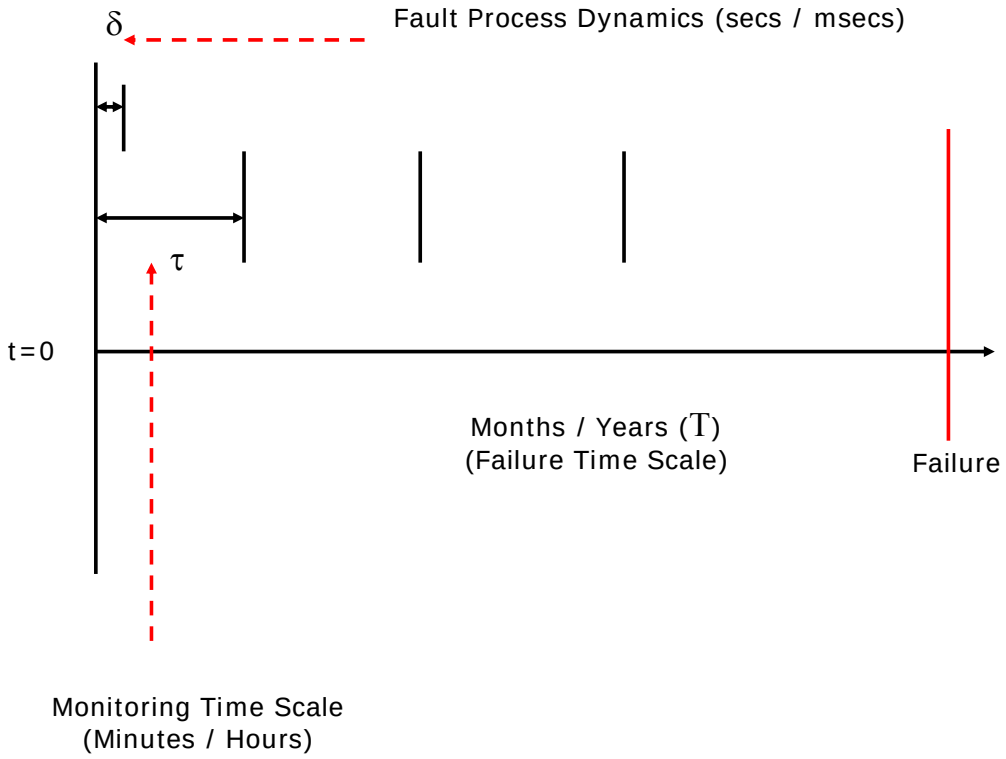

Figure 2.2. Different time scales in framework

been recently introduced in the literature [31]. The key concept in the proposed method is *Symbolic Time Series Analysis* (*STSA*) [32] that requires both temporal and spatial discretization of the pertinent measurement data to construct a sequence of symbols; the symbol sequence is treated as a transform of the original time series data from the phase space into a symbol space such that no significant information is lost [33]. The symbol sequence is used to construct a variant of a Hidden Markov Model (HMM) [34] of the locally stationary (i.e., fast-time-scale) dynamics in the form of a finite state automaton. A fixed-structure finite-memory state machine, known as the D-Markov machine [31], has been used. In this thesis, we shall use the D-Markov machine as the basis for the statistical framework.

For the method to be effective in the detection of motor faults, it is necessary to customize the method.

The focus is on experimental validation of the anomaly detection methodology, where the source of possible anomalies includes stator voltage imbalance and rotor bar breakage (Note: Machine damage evolves at a time scale that is several orders of magnitude slower than the electromechanical process dynamics of induction machines). The next section describes each step in the common methodology employed, for the detection of all slowly manifesting faults.

2.1 Framework for Fault Detection in Induction Motors

The problem is to quantify different faults, both electrical and mechanical, occurring in three-phase induction motors. Fig(2.3) shows the various steps involved in the framework, using either the stator current or motor vibration signals. The analysis begins with data acquisition of the stator current signals from the test apparatus. After sampling and quantization, time-frequency analysis of the collected data is performed to extract the fault information. Ideally, time-frequency analysis would be performed using the Continuous Wavelet Transform (CWT); however, to reduce the computational burden, the Discrete Wavelet Transform (DWT) is also used. In either case, the transform is customized to the type and nature of the fault.

While both CWT and DWT provide time and frequency information, they are computed using discrete sequences. It is important to understand the difference, which is in the form of outputs from several filters for the DWT whereas, in the case of the CWT, it is obtained as a series of convolutions for every time shift.

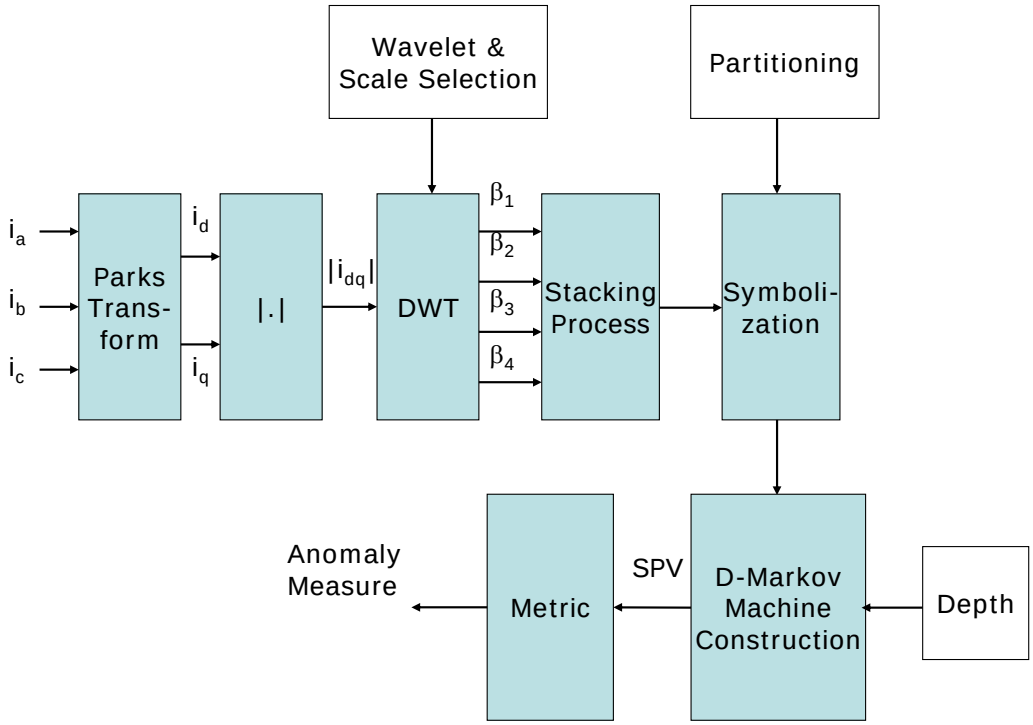

Figure 2.3. Framework

These coefficients are then stacked end-to-end, to ensure that the entire fault spectrum is included. This is followed by partitioning, which involves segmentation of the wavelet coefficient space of the nominal data set into finitely many mutually exclusive and exhaustive regions. Each region is assigned a symbol and the set of all symbols is called the alphabet. At each sampling instant, a symbol from the alphabet is generated.

The sequence of wavelet coefficient data is then represented by a fixed-structure finite-state automaton, first introduced by Ray, called the D-Markov machine [31]. The symbol sequence represents the statistical behavior of the process dynamics, and if a fault occurs, the statistical behavior is very likely to be changed. Ac-

cordingly, analysis of the resulting symbol sequence over the D-Markov machine yields a different statistical behavior. The state probability vectors, which are directly derived from the D-Markov machine, capture the evolving statistical behavior of the symbol sequences at different (slow-time-scale) epochs. Deviation from the nominal to an anomalous behavior of the dynamical system is identified as a change in the statistical behavior of the symbol sequence.

A vector representation of anomaly is obtained from the state probability histogram that, in turn, is derived from the respective transition matrices of the finite state automaton under the nominal and anomalous conditions. An appropriate metric is chosen to obtain a scalar quantification of the deviation of the anomalous probability histogram from the nominal one. Thus, the fault is detected and identified from an appropriately chosen metric that quantifies the vector difference of the current probability distribution and the nominal probability distribution as a scalar anomaly measure.

The subsequent sections describe in detail each element of the framework. This common framework is developed for the detection of most of the induction motor faults, indeed, any slowly manifesting faults in complex systems. As an aid to understanding the framework, let us consider an example. Let us assume the base signal to be a sinusoid with frequency 60 Hz and let there be two other (fault) signals of 50 and 10 Hz, representing the manifestation of the same fault in different frequency bands. Also, let the magnitudes (mean-peak value) of these signals be 1, 0.5 and 0.02 respectively. Fig(2.4) shows both the fault signatures, as well as the mixed signal.

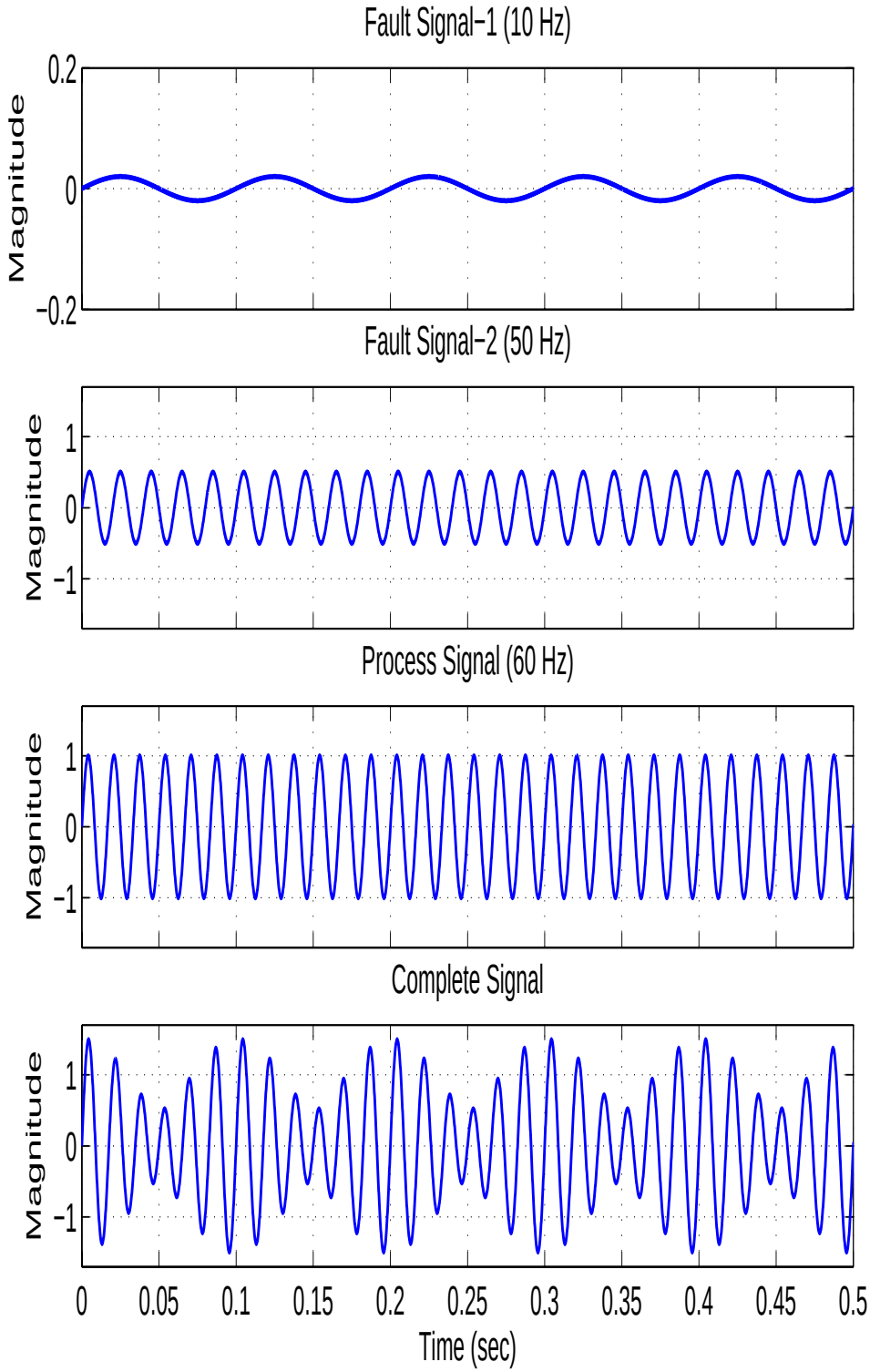

Figure 2.4. Fault signature and signal with embedded fault

2.2 Wavelet Transform of Time Series Data

Processing of the raw data, also known as pre-processing, involves collection of data, and extraction of meaningful information from it. To achieve this, wavelet analysis is used since it provides both time and frequency information [35]. The Continuous Wavelet Transform (CWT) has been used for the off-line analysis, which is given for a function $f(t)$ as:

$$CWT_{\alpha,\beta} = \alpha^{-1/2} \int_{-\infty}^{\infty} f(t)\psi(\frac{t-\beta}{\alpha})dt \tag{2.1}$$

where $\psi(t)$ is the mother wavelet with $\alpha > 0$ being the scale and β being the time shift. Therefore, the wavelet transform coefficients are a function of both time and scale. In our approach, the wavelet basis and the scales at which these transforms are to be evaluated, are chosen according to the nature of the anomaly. This serves the dual purpose of extracting relevant fault information in the best possible manner as well saving valuable computational resources.

2.2.1 Wavelet Basis Selection

The wavelet basis must be chosen off-line. Several properties of a wavelet determine its suitability as a basis. Much depends on the characteristics of the signal such as frequency trends, discontinuities etc. A selection criteria is developed in this thesis. The basis or mother wavelet is probably the most important choice to be made in this regard. There are several families of wavelets, some of them are shown in Fig(2.5).

The following properties are relevant to the problem of fault signature identification.

- *Time-Frequency localization*: While time localization allows identification

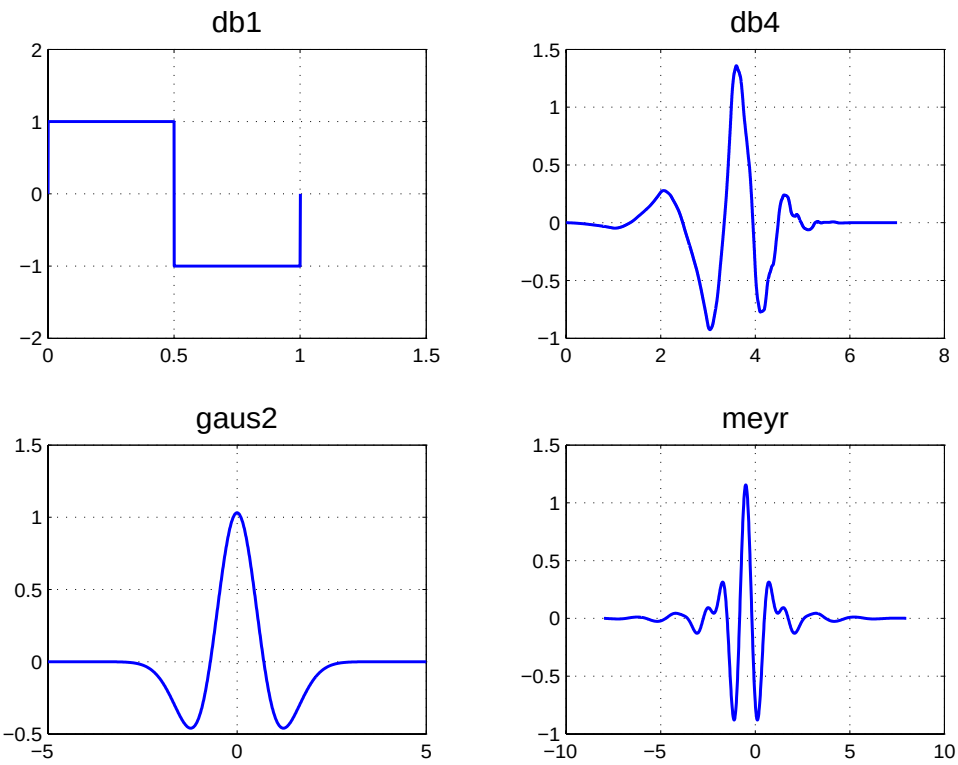

Figure 2.5. Comparison of Common Wavelets

of sudden changes in temporal behavior, frequency localization is useful for identification of faults in narrow frequency bands. For example, the wavelet basis 'db1' (or 'Haar') is an extreme example of good time localization, whereas the 'gaus2' wavelet is well localized in frequency. Localization refers to the ability to separate trends between signals. A wavelet that is well localized in frequency can tell apart two signals that are very closely spaced in the frequency domain. The same extends to wavelets well localized in time, with respect to signals that vary only slightly locally [36]. As we shall see in the following sections, the underlying fault signatures tend to be sinusoidal in nature, but transient in time, and, hence the need for wavelet that has good localization in time and frequency.

- *Vanishing moments*: Vanishing moments of the wavelet are related to its

ability to suppress lower order dynamics. For example, if the wavelet basis has $p+1$ vanishing moments, it will suppress all polynomials of degree p. So a wavelet basis with large number of vanishing moments will capture the change in the "higher-order" terms of a power-series decomposition. Since faults are usually associated with higher-order terms, this will ensure that even small faults will not go undetected. The wavelet 'db1' has 1 vanishing moment, while 'db4' has 4 vanishing moments. Hence, 'db4' can be used to detect faults that are non-linear in nature.

- *Regularity*: This property is derived from vanishing moments. This defines the smoothness of the wavelet basis and helps in separating the frequency components in the fault signals. 'db4' has good regularity of 1.27.

- *Symmetry*: This helps in factoring out phase information that can lead to false alarms. Enforcing this condition causes the wavelet to have complex phase. In fact, except the 'Haar/db1' wavelet, there is no other compactly supported wavelet that is symmetric [37]. The wavelet 'gaus2' is also symmetric but it is not compactly supported, consequently it cannot be implemented in real-time, whereas the wavelet 'db4'is compactly supported and is nearly symmetric.

Many wavelets may satisfy one or more of these desirable properties. In such a case, it is advantageous to choose a wavelet basis that is coherent with the signal. A measure of coherence between the signal and wavelet may be obtained from the cross-correlation, which is defined as:

$$\Gamma_{f,\psi_\alpha} = \frac{< f, \psi_\alpha >}{||f||_2 ||\psi_\alpha||_2} \tag{2.2}$$

where f is the signal and ψ_α is the suitably scaled wavelet. $< f, \psi_\alpha >$ is the inner

product between the vectors f and ψ_α and $||f||_2$ is the ℓ_2 norm of f. Fig(2.6) illustrates this idea of matching the signal with the wavelet. It is not always possible to correlate the underlying fault signature and the wavelet. In such cases, the wavelet should be correlated with a sinusoid at the fault frequency. If the fault frequency cannot be accurately determined then the best guess will suffice, following this the scales are spread around that particular fault frequency.

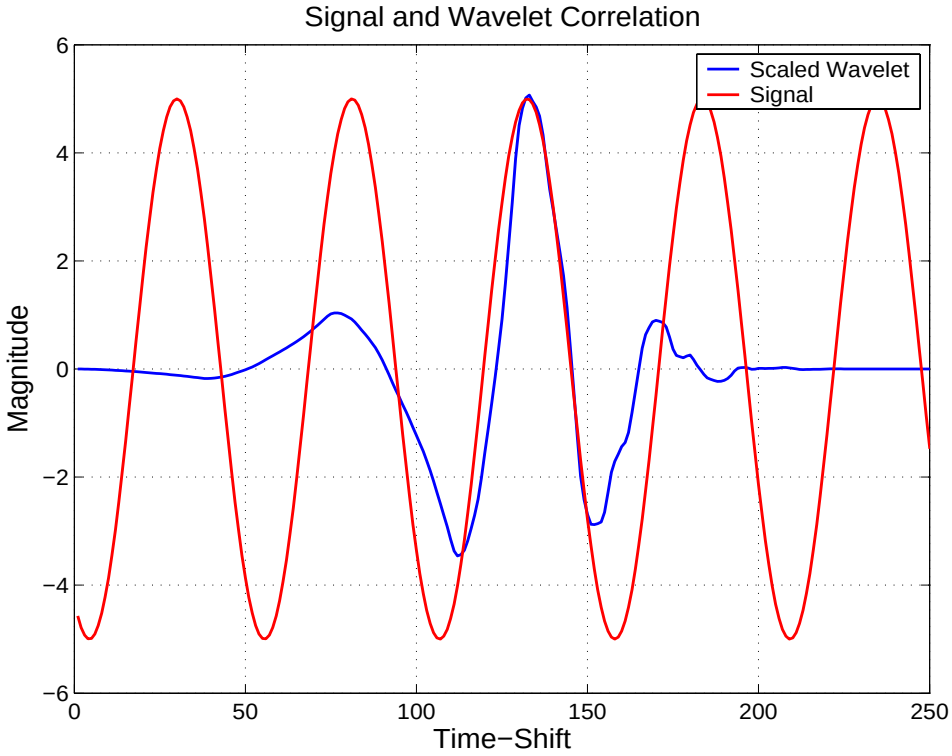

Figure 2.6. Correlation of the Fault Process with the Wavelet

Table(2.1) correlates some of the wavelets from the Daubechies family with the fault signal considered in the example (10 Hz sinusoid).

After considering all of the above factors, the most suitable wavelet for this application is found to be 'db4'. It has a high number of vanishing moments, good regularity, compact support and is nearly symmetric. Once the wavelet has been fixed, the scale(s) at which the analysis is to be performed must be determined.

Table 2.1. Cross Correlation Values

Wavelet	Γ
$db3$	0.3353
$db4$	0.3414
$db5$	0.3108
$db6$	0.2355
$db7$	0.2145
$db8$	0.2065
$db10$	0.2551

2.2.2 Wavelet Scale Selection

For every wavelet, there exists a certain frequency called the center frequency F_c that has the maximum modulus in the Fourier transform of the wavelet. The pseudo-frequency f_p of the wavelet at a particular scale α is given by the following formula [38]

$$f_p = \frac{F_c}{\alpha \ \Delta t} \tag{2.3}$$

The wavelet coefficients of a signal have the maximum magnitude at the scale corresponding to the pseudo-frequency. Hence, it would be appropriate to choose scales such that the pseudo-frequency corresponds to the frequency of interest. Multiple scales can be chosen as per Eqn(2.3). As stated earlier, real-time computation of the CWT may not be always be feasible. In keeping with the objectives of online health monitoring of induction motors, the Discrete Wavelet Transform (DWT) is used to address the real-time problem.

2.2.3 Discrete Wavelet Transform

The Discrete Wavelet Transform (DWT) provides a computationally simple alternative to evaluation of Eqn(2.1). It is expressed as

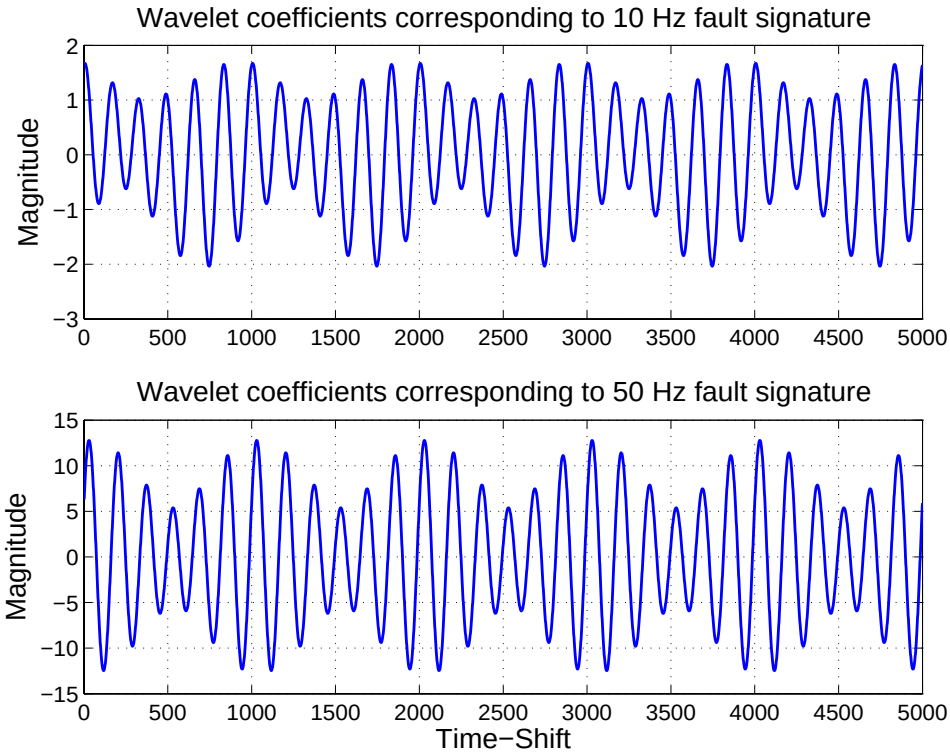

Figure 2.7. Extraction of both fault signals using wavelet expansion at different scales

$$C(j,k) = 2^{-j/2} \int_{-\infty}^{\infty} f(t)\psi(2^{-j}t - k)dt \tag{2.4}$$

The 'detail' coefficients of the DWT are defined by Eqs (2.5) and (2.6). These coefficients are obtained by using a filter-bank implementation as shown in Fig(2.8).

$$D_j(t) = \sum_{k\in\mathbb{Z}} C(j,k)\psi_{j,k}(t) \tag{2.5}$$

$$D_j(t) = \sum_{k\in\mathbb{Z}} 2^{-j/2} \int_{-\infty}^{\infty} f(t)\psi(2^{-j}t - k)dt\psi_{j,k}(t) \tag{2.6}$$

In this application, only orthogonal wavelets are used, since they form tight frames and hence reduce the computational load. Unless the signal to be analyzed exhibits fractal behavior, discrete time analysis at dyadic scales yields satisfactory

performance.

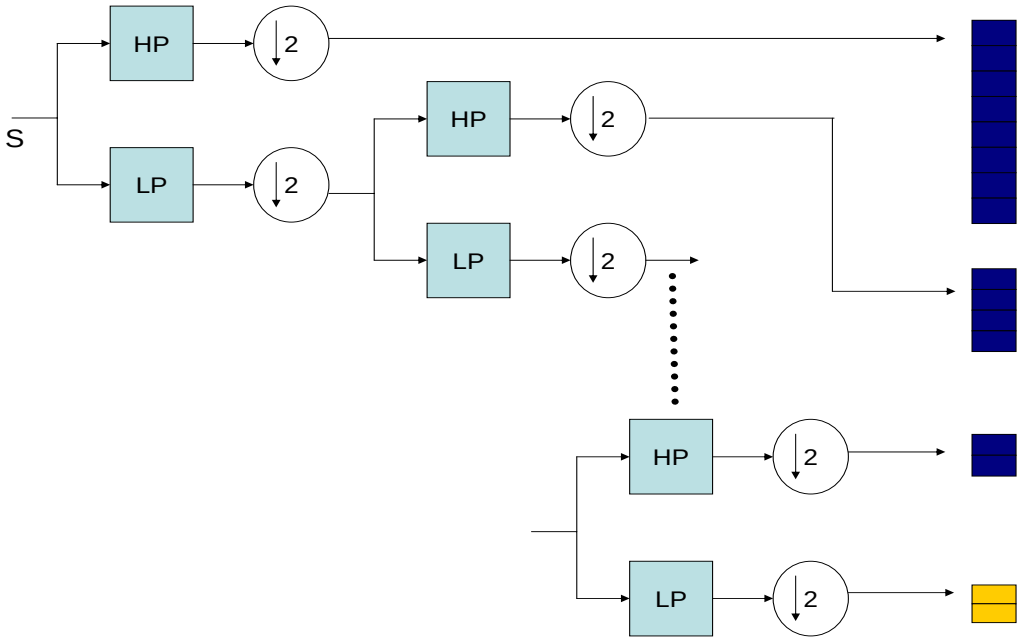

Figure 2.8. Filter Bank Implementation

Enforcing the orthogonality conditions may cause the wavelet transform to lose its shift-invariance [39]. This is a cause of major concern, because most of the waveforms analyzed in this paper are self-similar in some ways and have time shifts between them, and there is no synchronization between them. This problem is alleviated by utilizing symbolic dynamics for extraction of the statistical properties of the underlying process. The following sections describe the statistical framework in detail.

2.2.4 Stacking

For maximum effectiveness, the wavelet analysis is performed over all the scales associated with a particular type of fault. This information is two-dimensional,

containing the magnitude at each scale as well as time-shift. This information can be stacked to obtain a single time-series-like one-dimensional data set. Before we look at different ways of stacking the coefficients, C_b^a where a denotes the scale number, and b the time-shift; such that $a \in (\alpha^1\alpha^2 \dots \alpha^m)$ and $b \in (\beta_1\beta_2 \dots \beta_n)$. There are two ways of stacking:

- Scale-Based: For each scale the time shift data is gathered in a contiguous block and concatenated with the next scale at which this analysis is performed. Hence, the scale-series data as it is called looks like: $C_1^1 C_2^1 \dots C_n^1 C_n^2 C_{n-1}^2 \dots\dots C_2^m C_1^m$ This kind of stacking looks at continuous chunks of time for a particular scale, allowing for the depth (discussed in the subsequent section) to be increased. This means that the process has a finite history, and the depth D can be increased only until the next scale is reached.

- Time-Shift: For each time-shift coefficients at all the scales are evaluated and stacked. Thus, the time-shift series looks like: $C_1^1 C_1^2 \dots C_1^m C_2^m C_2^{m-1} \dots\dots C_n^2 C_n^1$. DWT coefficients when stacked for each scale, correspond to time-shift stacking.

An important point is that for depth $D = 1$ both time-shift and scale-based stacking lead to the exact same state probabilities, as we shall see in the later chapters. The most important property that stacking allows, is for the wavelet to include a band of frequencies, as well as the local behavior of the signal at each of the frequencies. Thus in our example, we have two scales, one corresponds to 10 Hz and the other to 50 Hz. In several cases the fault frequencies are known only approximately or in a range. In such cases the scales may be spread around the best estimate of the fault frequencies, such that it would cover the entire bandwidth of the fault signature. Stacking would help include all the fault information for the D-Markov machine to analyze. The stacked coefficients thus obtained are indicative

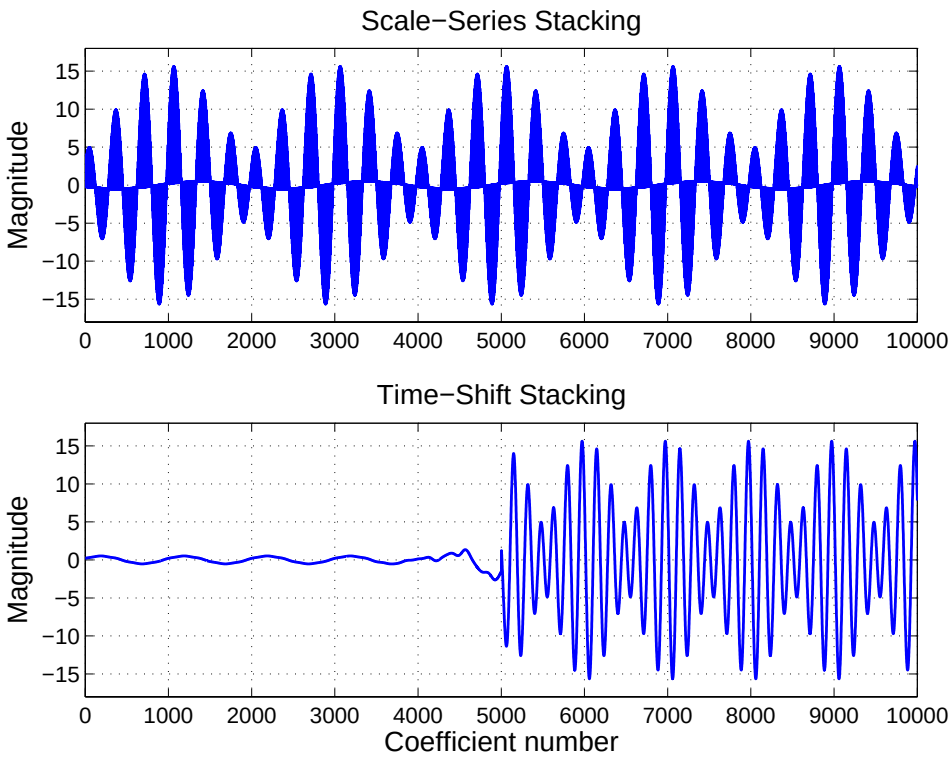

Figure 2.9. Scale-Series and Time-Shift stacking

of the fault. However, there is a drawback to stacking. The data obtained for each of the motor runs can be prohibitively large. For example, at a sampling rate of 1 kHz, with 2 scales for 20 seconds, approximately $40,000$ coefficients are generated. If this data were to be represented using double precision arithmetic, the resultant data structure would be extremely large and cumbersome. The next few sections address this computational problem, as well as illustrate how D-Markov machines can be generated from this data, in order to change the problem from a deterministic detection to a statistical one.

2.3 Partitioning of the Wavelet Coefficient Space

The space of continuous wavelet coefficients is partitioned to generate symbols. The aspects of symbol generation with wavelet coefficients, such as alphabet size are described in detail in [40].

Both CWT [41] and DWT [42] are used in the pre-processing stage, for obtaining symbols. Off-line analysis is performed a-priori, for scale and basis selection. Also the data collection interval is decided beforehand. The CWT scales are chosen in a spread pattern empirically. While the DWT scales are dyadic due to the filter-bank structure, and are centered around a particular scale of interest. After the coefficients are stacked to obtain the scale/time-shift series. The data series corresponding to the nominal condition is partitioned.

The partitioning allows us to extract relevant physical information about the observed dynamics, while reducing the computational load [43]. This involves identification of various artifacts of nonlinear systems using a Finite State Automaton (FSA), without completely characterizing the system. A discrete symbolic description is derived from the stacked coefficient series obtained from the wavelet analysis. The wavelet transform is essentially a linear operator, hence, this is analogous to partitioning the time series data. The resulting space of the map f associated with the flow through wavelet space is divided into cells, so as to obtain a coordinate grid for the underlying dynamics.

A compact set $\Omega \in \mathbb{R}^n$ within which the motion under quasi-stationary operations is circumscribed, is identified with the wavelet space itself. The encoding of Ω is accomplished by introducing a partition $\mathbf{B} = (B_0, B_1, B_2, B_{n-1})$consisting of n mutually exclusive and exhaustive subsets, i.e., $B_j \bigcap B_k = \oslash, \forall j \neq k; \bigcup_{j=1}^{n} B_j = \Omega$. The dynamical system describes an orbit $O = (x_0, x_1, x_2, ...x_n, ...)$, which passes through or touches various elements of the partition $\mathbf{B}$. Let us denote the index of

the domain $B_i \in \mathbf{B}$ visited at k^{th} sample as the symbol $\sigma_k \in \Sigma$. The set of symbols $\Sigma = (\sigma_1, \sigma_2.....\sigma_{n-1})$ labelling the partition elements is called the alphabet.

This forms a mapping from the wavelet space to the space of symbols. Such a mapping is called Symbolic Dynamics, as it attributes a legal (i.e., physically admissible) symbol sequence to the system dynamics. As the size of each cell is finite and the cardinality of the alphabet Σ is finite, any such symbol sequence represents, through iteration, a trajectory that has the compact support Ω. Hence, the modeling has to be restrictive and the worst case must also be considered. In general, a dynamical system would generate only a subset of all possible sequences of symbols as there are many illegal (i.e., physically inadmissible) sequences. The grammar (i.e., set of rules) that determines legality of symbol strings in the alphabet Σ may change with the occurrence of the fault, which as discussed earlier occurs in the slow time scale. Symbolic dynamics can be viewed as coarse graining of the wavelet space and hence subjected to (possible) loss of information. However, the essential robust features are expected to be preserved in the symbol sequences through an appropriate partitioning of the wavelet space [27]. This paper presents two methodologies for partitioning shown in Fig(2.10).

2.3.1 Uniform Partitioning

In this approach, the maximum and minimum of the scale series are evaluated and the ordinates between the maximum and minimum are divided into equal-sized regions. These regions are mutually disjoint and thus form a valid partition. Each region is then labeled with one symbol from the alphabet. If the data point lies in a particular region, it is coded with the symbol associated with that region. Thus, a sequence of symbols is created from a sequence of stacked wavelet coefficients.

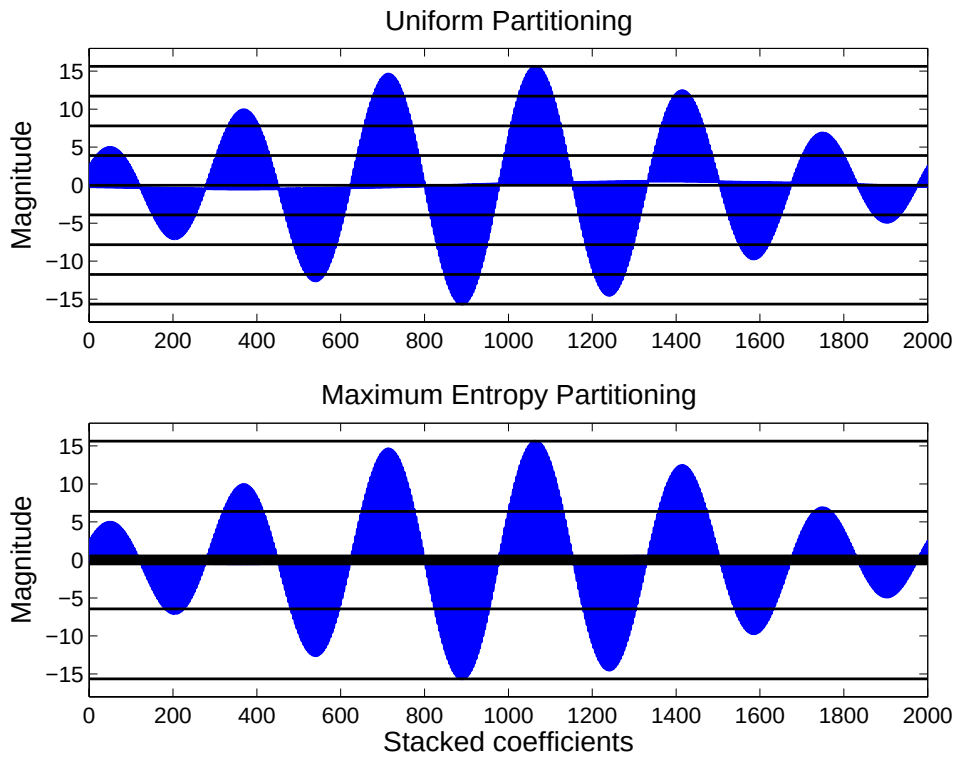

Figure 2.10. Uniform and Maximum Entropy Partitioning

2.3.2 Maximum-Entropy Partitioning (ME)

The Maximum Entropy (ME) scheme of partitioning introduced in [40] is presented briefly below. Let N be the length of the data and $|\Sigma|$ be the cardinality of the (finite) alphabet. The data is sorted in the ascending order. Starting from the first point in the sorted set, every consecutive segment of length $\lfloor \frac{N}{|\Sigma|} \rfloor$ forms a distinct element of the partition, where $\lfloor x \rfloor$ represents the greatest integer less than or equal to x. Each partition is then labeled with one symbol from the alphabet. If the data point lies in a particular region, it is coded with the symbol associated with that region. Thus, a sequence of symbols is created from a sequence of scale series data.

With such a partition, a region with large information content is allocated more symbols and hence a finer partitioning is achieved in such a region. Similarly, a

region with sparse information content is allocated fewer symbols and hence a coarser partitioning is achieved in such a region. Thus, even small changes in the system behavior are reflected in the symbol sequence obtained under ME partitioning, which is performed off-line based on the data set under the nominal (i.e., healthy) condition; this partition remains invariant for all conditions.

2.4 D-Markov Machines for Symbolic Time Series Analysis

After obtaining the symbol sequence, a finite state automaton is constructed. The purpose of a finite state automaton is to represent the dynamical behavior of the underlying process. A finite state automaton, known as the D-Markov machine, has been used to represent the symbolic process [31] in this paper.

The core assumption here is that the symbolic process can be approximated, to a desired level of accuracy, as a D^{th}order Markov chain, where $D \in \mathbb{N}$, and $\mathbb{N}$ is the set of natural numbers. The states are chosen as symbol strings of length D from the symbol sequence from the alphabet Σ. With the cardinality $|\Sigma|$ of the alphabet and the depth D, the number of states in the D-Markov machine is less than or equal to $|\Sigma|^D$. The state machine moves from one state to another upon occurrence of an event symbol. All symbol sequences for which the last D symbols are the same, lead to the same state.

Having constructed the D-Markov machine, the transition probabilities π_{ij}^{σ} between the states, upon occurrence of the symbol $\sigma \in \Sigma$, are determined, where i is the state of origin and j is the state transitioned to. As the system trajectory evolves, different states are visited with different frequencies. The number of times a state is visited as well as the number of times a particular symbol is

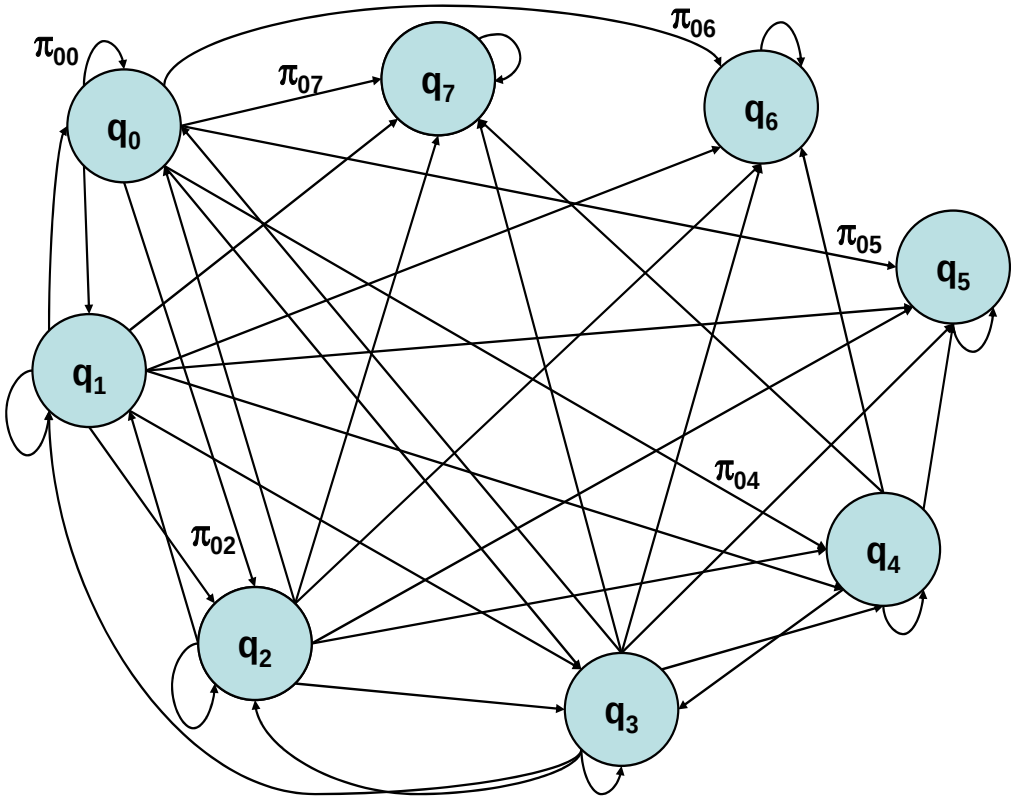

Figure 2.11. Example of a D-Markov Machine

received, is counted. The state probabilities as well as the state-to-state transition probabilities are calculated for each state in this way. The probabilities are obtained by frequency counts. Thus $\pi_{ij}^{\sigma} \approx \frac{n_i^{\sigma}}{N_i}$, where n_i^{σ} is the number of times the symbol σ occurred at the state i and N_i is the total number of occurrences of symbols at the state i, i.e $N_i = \sum_k n_i{}^{(\sigma_k)}$. The matrix comprised of all the transitional probabilities between the various states is called the state transition matrix Π. The row sum of the Π matrices should always be unity and hence Π is a stochastic matrix. For the example considered in the previous sections, the transition probabilities are given in Table(2.2). We can see that from this table

that the transition probabilities, exist from state q_0 to itself ($\pi_{00} = 0.5254$) and to state q_1 ($\pi_{01} = 0.4746$). Whereas, the transitions from q_0 to q_3, q_4, q_5, q_6 and q_7 are not defined, and hence their transition probabilities are null. This is also intuitive, since q_0 and q_1 are adjacent states in the partitioned ordinate, and q_0 is lowest interval in the partition. Similar behavior is shown by state q_7. The states q_3,q_4 and q_5 are the most well connected states. The transition matrix shown in the table is for the nominal data. As the fault manifests itself, the probabilities of transition will change. But for a particular type of fault, they will change in a particular way. This helps in isolating the faults, and also allows for slight accommodation to be able to account for the differences, between various machines in the same class. In the case of Maximum-Entropy (ME) partitioning, the state transition matrix is more well connected, due to the different sized partitions. The elements of the

Table 2.2. Π-Matrix with Uniform Partitioning

	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7
q_0	0.5254	0.4746	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
q_1	0.3751	0.2467	0.3751	0.0032	0.0000	0.0000	0.0000	0.0000
q_2	0.0000	0.4979	0.0000	0.3449	0.1572	0.0000	0.0000	0.0000
q_3	0.0000	0.0100	0.5778	0.0000	0.3998	0.0125	0.0000	0.0000
q_4	0.0000	0.0000	0.0100	0.4032	0.0000	0.5855	0.0012	0.0000
q_5	0.0000	0.0000	0.0000	0.1549	0.3399	0.0073	0.4979	0.0000
q_6	0.0000	0.0000	0.0000	0.0000	0.0030	0.3633	0.2887	0.3450
q_7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.5028	0.4972

state probability vector $\mathbf{p}$ denote the probability of occurrence of states of the D-Markov machine. This is obtained by computing the left eigenvector of the Π matrix, corresponding to the eigenvalue 1.0, shown in Table(2.3). Since $\mathbf{p}$ satisfies the stationary Markov property

$$\mathbf{p}\Pi = \mathbf{p} \tag{2.7}$$

With ME partitioning, in the nominal case, the states have identical probabilities

Table 2.3. **p**-Vector

	0 Ω
q_0	0.1254
q_1	0.1586
q_2	0.1185
q_3	0.1004
q_4	0.1001
q_5	0.1204
q_6	0.1641
q_7	0.1126

of occurrence $1/N$, where N is the number of states. The salient features of D-Markov machines are discussed next.

2.4.1 Discrete and Multinomial Distributions

The counting process forms a discrete distribution [44], $D(\sigma_n) = \sum_{k=1}^{n} P(\sigma_k)$, where $P(\sigma_k)$ defined over $k = 1, 2, \ldots, N$ form the state-probability vector. As can be seen the method is completely data driven, as more data is obtained, the stochastic modeling of the underlying process improves. It can also be modeled as a multinomial distribution, which can be thought of as a die-experiment. To elaborate, let us assume we had a fair die with six faces and a number on each of the faces. The probability of occurrence of each of the states or numbers in the case would be equal. However, if there was a problem with the die (a fault), then the probabilities of certain numbers occurring would increase. After repeated trials, we maybe able to estimate these occurrence probabilities. There is a problem here, it is not possible to take measurements of motor signals and equate them to throwing a die. This is where the framework developed, uses the stationary Markov property given in Eqn(2.7), to yield a vector **p** which satisfies the properties of a

multinomial distribution given by Eqn(2.8).

$$P(\sigma_1 = x_1, \ldots, \sigma_n = x_n) = \frac{N!}{\prod_{i=1}^{n} x_i!} \prod_{i=1}^{n} \theta_i^{x_i} \tag{2.8}$$

where x_i are positive integers. In this case since the events $\sigma_1, \sigma_2, \ldots, \sigma_n$ are all mutually exclusive events, with $P(\sigma_1 = x_1) = \theta_1, \ldots, P(\sigma_n = x_n) = \theta_n$. These requirements for multinomial distributions are satisfied by D-Markov machines, since each of the symbols are generated from mutually-exclusive partitions. Multinomial distributions are 0-Markov processes, i.e. all the states are independent of the previous state. This is true for the SPV's ($\mathbf{p}$) for each of the induction machine.

2.4.2 Robustness to Noise

The D-Markov machine has the ability to reject momentary or shot noise. This can be easily illustrated by considering the following example. Let there be a disturbance in the measured signal for a short amount of time. This will cause some of the transition probabilities to change. Let us assume that the transition probability from i to j changes from $\pi_{ij}^{(\sigma_\ell)}$ to $\tilde{\pi}_{ij}^{(\sigma_\ell)}$, this is caused by a change in $n_i^{(\sigma_\ell)}$ to $\tilde{n}_i^{(\sigma_\ell)}$, but since N_i is a large number the difference $|\pi_{ij}^{(\sigma_\ell)} - \tilde{\pi}_{ij}^{(\sigma_\ell)}| < \epsilon$ where $\epsilon \to 0$ as $N_i \to \infty$. This will be the case when the data set is large. If this disturbance is occurring more often, then it cannot be categorized as noise. It is either part of the system dynamics or is a fault process. The property of noise rejection reduces the noise due to the end-effects [39]. These end-effects are generated due to the convolution of the signal with the decomposition filters. While stacking the output from different filters, these end effects become part of the signal that is analyzed by the D-Markov machine. If these end-effects are not averaged out, the resulting

analysis will have artifacts that are not representative of the underlying signal.

2.4.3 Sensor Fusion by Way of Mixed Multinomials

This section demonstrates the combination of information from the current and vibration sensors to demonstrate a more robust fault detection capability. The combination of different multinomials can be achieved thus [45]. If there are two multinomial distributions $\mathbf{X} \sim M(n_1; p_1, p_2, \ldots, p_k)$ and $\mathbf{Y} \sim M(n_2; \tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_k)$, the combined distribution is defined as

$$\mathbf{Z} \sim (n_1 + n_2; \frac{n_1 p_1 + n_2 \tilde{p}_1}{n_1 + n_2}, \frac{n_1 p_2 + n_2 \tilde{p}_2}{n_1 + n_2}, \ldots, \frac{n_1 p_k + n_2 \tilde{p}_k}{n_1 + n_2}) \tag{2.9}$$

Each of the distributions $\mathbf{X} \sim M(n_1; p_1, p_2, \ldots, p_k)$ and $\mathbf{Y} \sim M(n_2; \tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_k)$ have $k-1$ independent variables. The combined distribution still has $k-1$ independent variables, each of the variables has been weighted and renormalized. The sample space remains the same, i.e., the number of events possible is the same k. Thus, we can combine the information from two disparate sensors. These combined distributions are further checked for consistency, and the fault information. The details of these are given in Chapter 5. In that chapter, vibration and current signals are combined to detect bearing failure. Since, the current and vibration signals have entirely different fault spectra. Also, the vibration signals are typically more noisy and sensitive, they cannot be directly combined, or combined in the wavelet domain. The framework, allows for their combination in the probability space.

2.4.4 Anomaly Measure and Detection

The D-Markov machine described in Section 2.4 is capable of representing the statistical behavior of a dynamical system under faulty/anomalous conditions. In order to quantify changes in the probability distribution, a measure $\mathbf{M}$ must be introduced. The state transition probabilities are dependent on the dynamics of the induction motor as reflected in the symbol sequence. This is a factor in detecting a fault/anomaly because perturbations in the system dynamics cause significant changes in the state transition probabilities, though such changes may also be dependent on how the wavelet space is partitioned (if it is uniformly partitioned or not). The induced norm of the difference between the state transition matrices Π_0 and Π, respectively for the nominal and anomalous state machines G_0 and G , may be used as a measure of anomaly,i.e. $\mathbf{M}(G_0, G) = ||\Pi_0 - \Pi||$. Such a measure, adopted in [46], was found to be very effective. Since D-Markov machines have a fixed state structure, the SPV associated with all state machines will have the same dimension. Hence, they are used for vector representation of the anomaly, leading to the anomaly measure some of these metrics are given below.

- *Norms*: In the finite dimensional case they are all equivalent, and have different interpretations. For example the 2-norm has a physical interpretation of energy. $\mathbf{M}(G_0, G) = ||\mathbf{p_0} - \mathbf{p}||$ gives the difference between p_0 and p the nominal and anomalous SPV's.

- *Standard deviation*: The standard deviation given by

$$Std = \sqrt{\sum_{k=1}^{n} P(\sigma_k)(\sigma_k - \mu)^2}$$

  where $\mu = \frac{1}{N}\sum_{k=1}^{n} P(\sigma_k)$ is the population mean. This measure is equivalent

to the 2-norm.

- *Kullback-Liebler Distance*: $M = -\sum_{i=1}^{i=k} {p_i}^f log_2 \frac{{p_i}^f}{{p_i}^{nom}}$. Where k is the number of states in the SPV, ${p_i}^{nom}$ is the probability of the i^{th} state for the nominal case. Likewise ${p_i}^f$ are the probabilities for the fault/anomalous condition. However, this measure is however not symmetric, thus it does not satisfy all the properties of a metric in a vector space.

The properties described above help us in associating a robust anomaly measure with each of the faults described. The experiments performed to validate these concepts are given in the next three chapters.

Chapter 3

Stator Voltage Imbalance

In this chapter, the problem of stator voltage imbalance is studied. The various causes and effects of stator voltage imbalance have already been described in Chapter-1. This type of fault occurs at the terminals, when instead of a balanced three phase supply, an imbalance occurs. In real life, this imbalance usually occurs in one of the three phases. Stator voltage imbalance is similar to stator winding degradation type of faults. Because it is created easily, it is one of the most widely studied faults. This chapter begins with an introduction to the fault under study and the underlying physical process. This is followed by an analysis of data obtained from a simulation technique called MEC (Magnetic Equivalent Circuits). This is followed by a more in depth analysis by comparing the partitioning techniques as well as other techniques, using data obtained from the induction machine.

3.1 Fault Process

Fig(3.1) shows imbalanced line currents, which are a cause of the stator voltage imbalance. This imbalance can be represented in the symmetric components, by the

negative sequence components, again shown in Fig(3.1). A basic transformation of the original stator currents yields that the fault information lies in the range of twice the line frequency. The analysis uses the Parks vector transformation, which is obtained by performing a reference frame transform to a stationary reference frame. To detect stator imbalances the line currents represented by I_a, I_b and I_c are used. As such these line currents provide little or no information, this is especially true for small values of imbalance. For small increases in the stator imbalance, an appreciable change is not noticed in the line currents. I_+ and I_- are

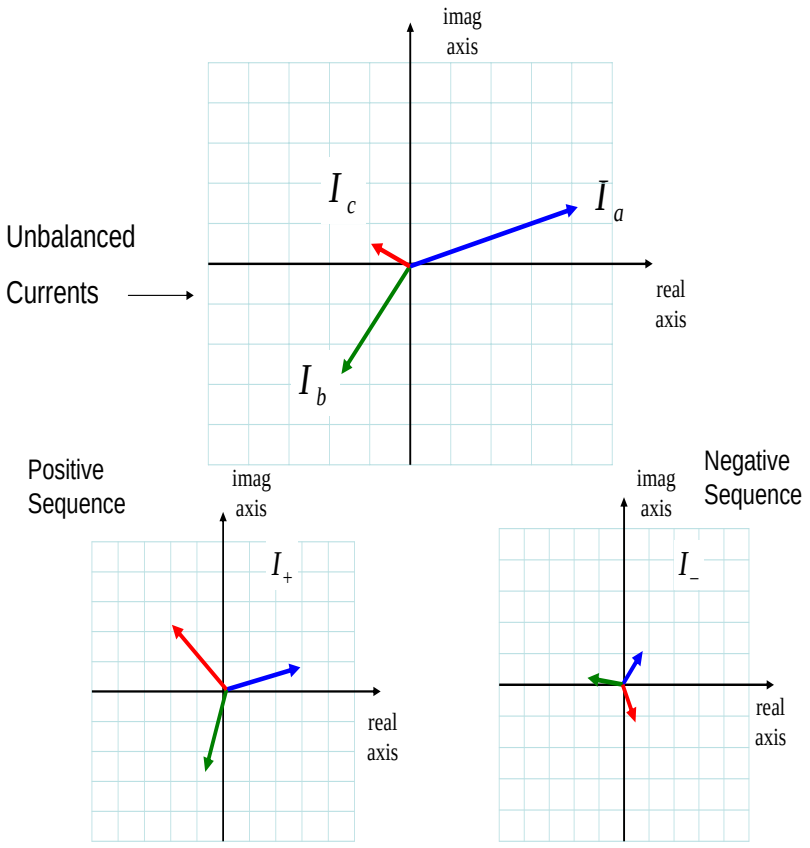

Figure 3.1. Line Currents

the positive and negative sequence components, shown in Fig(3.1). The theory of

symmetrical components allows for the representation of all three line currents by

$$I_a = I_+ cos(\omega t) + I_- cos(\omega t) \tag{3.1}$$

$$I_b = I_+ cos(\omega t - \frac{2\pi}{3}) + I_- cos(\omega t - \frac{4\pi}{3}) \tag{3.2}$$

$$I_c = I_+ cos(\omega t - \frac{4\pi}{3}) + I_- cos(\omega t - \frac{2\pi}{3}) \tag{3.3}$$

The imbalance is usually not large, it is assumed that $I_- \ll I_+$. For this unbalanced condition, the direct and quadrature axis currents are [47]

$$i_d = (I_+ + I_-)\frac{\sqrt{6}}{2} cos(\omega t) \tag{3.4}$$

$$i_q = (I_+ - I_-)\frac{\sqrt{6}}{2} sin(\omega t) \tag{3.5}$$

and the Parks vector modulus then becomes

$$i_d^2 + i_q^2 = \frac{3}{2}(I_+^2 + I_-^2) + 3I_+ I_- cos(2\omega t) \tag{3.6}$$

The fault signature resides most clearly in the twice the line frequency 2ω. The next section uses a simulation example to shows the efficacy of the framework, in detecting the onset of voltage imbalance.

3.2 Simulation Studies

The method used to simulate the machine is the Magnetic Equivalent Circuits (MEC) technique. A more detailed description is given in Appendix-D and in [48] and [49]. In brief, the simulation technique is a trade-off between traditional lumped parameter models, which are insufficient to capture the dynamics, particularly the change in dynamics due to small anomalies; and the computationally

more expensive FEM (Finite Element Models), which accounts for the geometry of the machines as well as minute interactions between all the components.

Magnetic Equivalent Circuit (MEC) models are used for analyzing flux density and inductances in simple electromechanical systems. In this form of analysis, an analogy is drawn between the mmf, flux and reluctance of the magnetic circuit formed by a high-permeability core and the emf, current, resistance of an electric circuit. The actual process of modeling a magnetic circuit involves breaking the high-permeability core into segments for which a reasonable equivalent reluctance can be calculated from the length, cross-sectional area, and permeability of material in each segment. The end points of these segments become nodes in an analog electric circuit, in which node voltages, branch currents, and branch resistances correspond to magnetic potentials, fluxes, and reluctances, respectively, in the magnetic circuit. The sources or forcing functions in this analog circuit are voltage (emf) sources that correspond to mmf produced by currents flowing in windings associated with with the actual magnetic circuit. Modified nodal analysis can be applied to the analog circuit to compute the magnetic potentials and fluxes throughout the magnetic circuit.

3.2.1 Simulation Results

The details of the machine simulated are in Appendix-D. Imbalance was introduced by changing the stator voltage by changing the input voltage on one of the phases. The voltage was reduced on one of the phases in steps, from the nominal voltage by 0.01, 0.02, 0.05, 0.1, 0.15 and 0.2 pu. For increasing amounts of imbalance, the negative sequence components increase in magnitude. Fig(3.2) shows the line currents for the unbalanced and 0.2 pu imbalance. Even for such a large amount of imbalance, there is no visible change in the current signals. Thus, it is not hard to

imagine that, it will be extremely difficult to detect small amounts of imbalances.

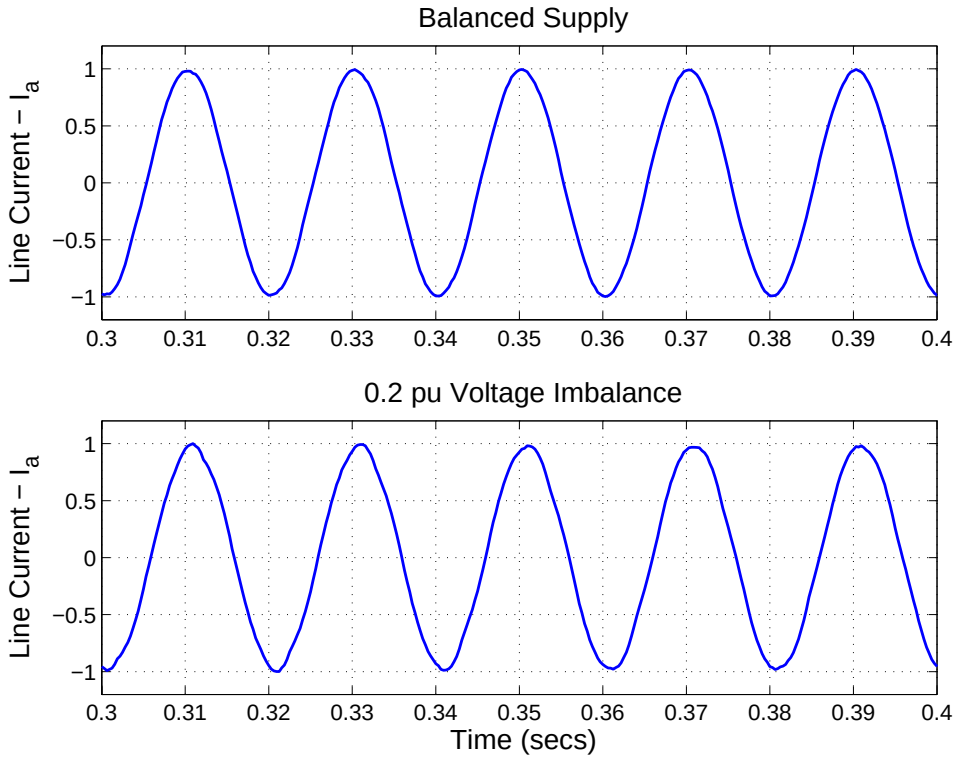

Figure 3.2. Line Currents

A much easier technique is to look at the Park's vector modulus. These curves plot the direct axis current (i_d) against the quadrature axis current (i_q). This will allow us to look at the steady-state characteristics, since ideally for a balanced condition these curves should be a circle. However, due to the machine geometry, this is usually not the case. This is shown in Fig(3.3). The detection problem is now more complicated. This is where model-based fault detection techniques suffer their major drawback. Since, model-based detection assumes symmetrical ideal behavior. In other words the Parks modulus curves would be circular for the balanced voltages, which they are not. Additionally, the residual generated would be large in the balanced case and thus further increasing imbalance is likely to be

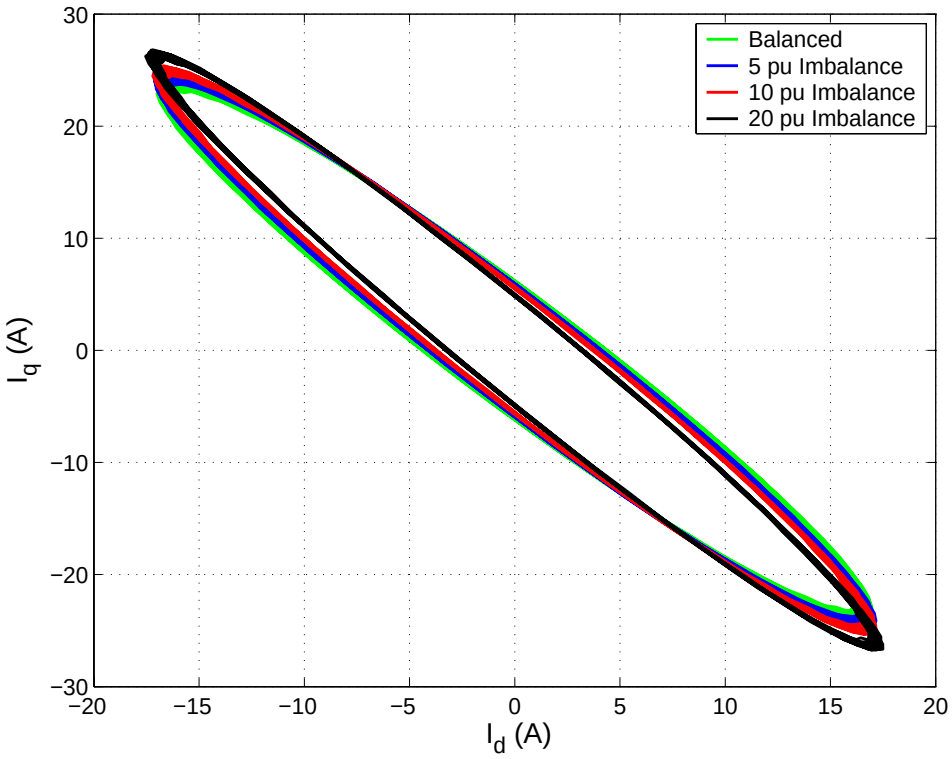

Figure 3.3. I_d vs I_q

imperceptible.

This imbalance information is extracted by performing the wavelet analysis (CWT) at the scales 5.7 and 11.4. At a sampling rate of 800 Hz, and the operating line frequency of 50 Hz, would indicate that the pseudo frequencies corresponding to these scales are 100 Hz and 200 Hz. At each scale the CWT is evaluated for the entire data-set, which comprises of 2 seconds of data. The stacked scale-series data for both the balanced and 0.2 pu imbalance is shown in Fig(3.4).

The figure indicates a change in the wavelet coefficient magnitude between the balanced and 0.2 pu imbalance. These stacked wavelet coefficients are then partitioned (uniformly) and represented as a symbol string. The differences in the occurrences of each of the symbols is represented in the state transition matrices of the D-Markov machine shown in Tables (3.1) and (3.2).

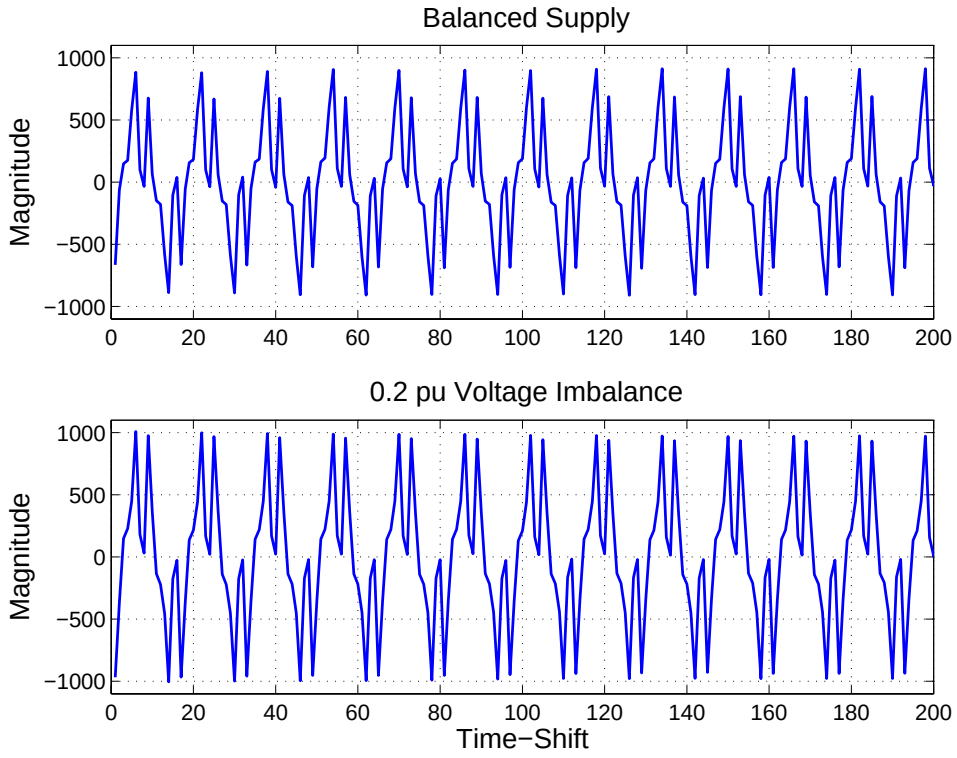

Figure 3.4. Stacking (scale-based), partitioning and comparison

Table 3.1. Π-Matrix for balanced operation

	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7
q_0	0.0067	0.0000	0.0000	0.9933	0.0000	0.0000	0.0000	0.0000
q_1	0.4950	0.0033	0.0033	0.4983	0.0000	0.0000	0.0000	0.0000
q_2	0.0000	0.5714	0.2857	0.0000	0.0000	0.0000	0.1429	0.0000
q_3	0.0013	0.1987	0.0027	0.2040	0.3946	0.0013	0.1785	0.0188
q_4	0.0000	0.1984	0.0027	0.3954	0.2011	0.0013	0.1997	0.0013
q_5	0.0000	0.0000	0.0000	0.5000	0.0000	0.0000	0.5000	0.0000
q_6	0.0000	0.0000	0.0000	0.0000	0.4690	0.0000	0.0138	0.5172
q_7	0.0000	0.0000	0.0000	0.0061	0.9758	0.0000	0.0182	0.0000

Further to extract stationary behavior we look at the state visit probabilities given in Table(3.3). Simply by looking at the State Probability Vectors (SPV), we can see that the peripheral states q_0 and q_7 have increased probability of occurrence as the imbalance increases. This is consistent with the progression of imbalance.

Table 3.2. Π-Matrix for unbalanced operation

	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7
q_0	0.0000	0.0000	0.5067	0.4933	0.0000	0.0000	0.0000	0.0000
q_1	0.3333	0.3333	0.3333	0.0000	0.0000	0.0000	0.0000	0.0000
q_2	0.4837	0.0065	0.0131	0.0098	0.4804	0.0033	0.0000	0.0033
q_3	0.2301	0.0017	0.2561	0.4810	0.0277	0.0000	0.0017	0.0017
q_4	0.0313	0.0000	0.0000	0.0039	0.3906	0.2656	0.0215	0.2871
q_5	0.0000	0.0000	0.0060	0.4401	0.0000	0.1377	0.1497	0.2665
q_6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0308	0.0000	0.9692
q_7	0.0000	0.0000	0.0000	0.0033	0.4884	0.4917	0.0099	0.0066

As the imbalance increases, the magnitude of the wavelet coefficients focussed on the fault should also increase. After partitioning the peripheral states in the D-Markov machine will see an increased probability of occurrence of those states.

Table 3.3. **p**-Vectors for MEC with Uniform Partitioning

	Balanced
q_0	0.0625
q_1	0.1250
q_2	0.0029
q_3	0.3104
q_4	0.3088
q_5	0.0008
q_6	0.1208
q_7	0.0688

	0.05 pu Imbalance
q_0	0.1217
q_1	0.0658
q_2	0.0029
q_3	0.3071
q_4	0.3117
q_5	0.0008
q_6	0.0646
q_7	0.1254

	0.2 pu Imbalance
q_0	0.1242
q_1	0.0017
q_2	0.1275
q_3	0.2408
q_4	0.2133
q_5	0.1392
q_6	0.0271
q_7	0.1263

It is important to quantify the difference between the different levels of imbalance. This imbalance can be treated as a parametric variation. As the parameter (imbalance) increases, a good fault detection technique should also indicate a linear increase in damage/anomaly measure. It is important to understand that although the damage process is non-linearly increasing in time, the measure should reflect the actual fault magnitude. Hence a linear graph would be the most ideal. The metric chosen was the 1-norm, in fact all norms for finite dimensional vector spaces are equivalent [50], hence any norm can be used. Fig(3.5) shows the pu imbalance

and corresponding anomaly measure.

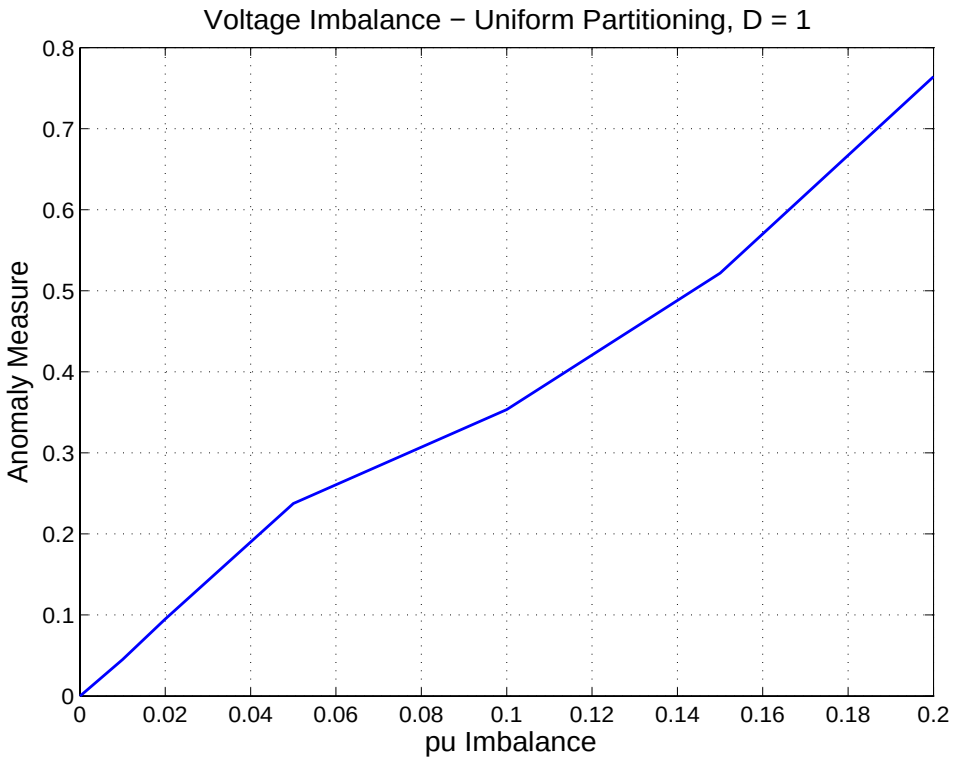

Figure 3.5. Anomaly Measure

The analysis in this section was performed using the uniform partitioning, using the data from MEC simulation. In the following sections we will use the data obtained from the induction motor experiment.

3.3 Experimental Results

In this section the effect of partitioning will be explored. Particularly, we will explore the ME partitioning methods and look at the different metrics used for fault detection. We will also look at depth (D) considerations and how they effect the detection process. Subsequent sections base their analysis on experimental data collected from the motor-dyne test bed, details of which are given in Appendix-

A. Current signals are utilized for the detection process. The experiment was performed at different load levels, from no-load to full-load. The stator imbalance was created by adding a resistor to one of the phases of the motor. The resistance was increased for every run, the values were 0, 1, 5, 10, 14 and 50 Ω. By putting a series resistor, the terminal voltage is affected in one of the phases, reducing it. The amount of stator voltage reduction depends on the load level. The best results shown here were at full-load conditions. The tests were performed for increasing levels of imbalance. Fig(3.6) shows the Park's vector modulus curves for increasing imbalance, notice that even for the balanced the curve is not completely circular.

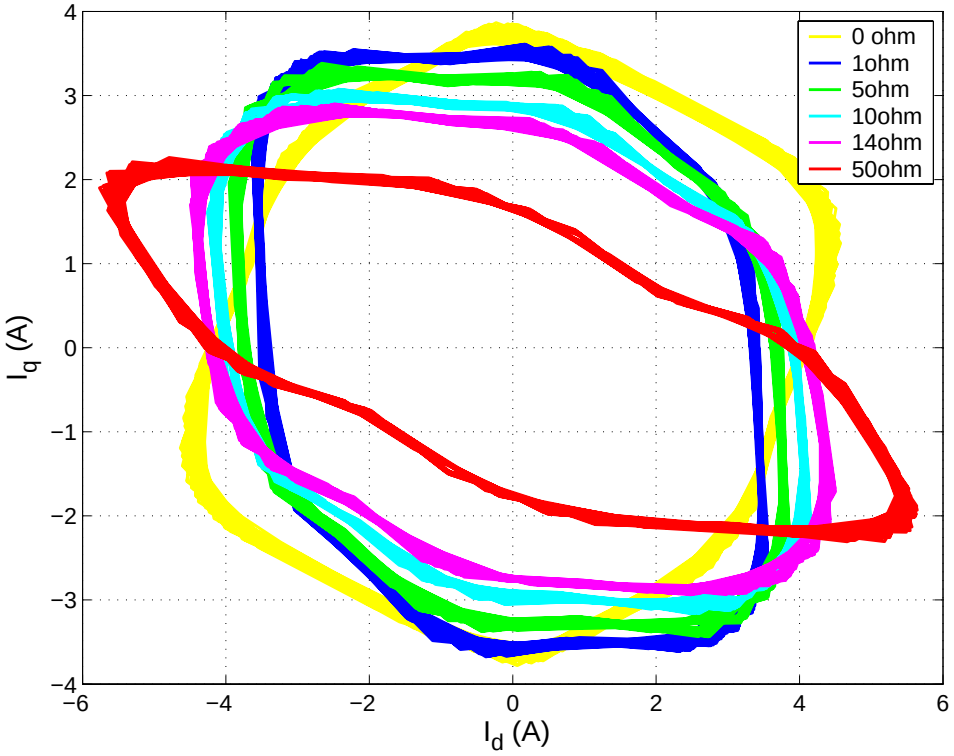

Figure 3.6. I_d vs I_q for increasing resistance levels in one phase

After the data was acquired, the Park vector modulus was computed; subsequently, wavelet analysis was performed on the transformed data. Unlike in the previous sections the analysis was performed using the filter-bank implementation

of the DWT. Decomposition of the signal was performed up to the third level. The outputs of the high-pass filters from the first three levels of wavelet transformation, which also give the detail coefficients, were time-shift stacked. The real strengths of the DWT analysis, namely the computational efficiency and the dyadic scaling are put to good use in this example. Computationally the DWT is more efficient than the Fast Fourier Transform (FFT) [36]. Additionally, due to the dyadic nature of the filter bank harmonics can also be analyzed. The wavelet chosen in this case was the 'db4' wavelet. The wavelet coefficients were obtained at the scales 0.5, 0.25 and 0.125, which were chosen according to Eqn(3.6). Fig(3.7) shows the stacked wavelet coefficients for the balanced supply and 14 Ω of imbalance. Both the partitioning techniques are demonstrated. It is interesting to note that the amount of data obtained successively decreases for each dyadic scale, thus the maximum weight is given to the fundamental scale chosen. There is also an artifact introduced at the ends of each scale, due to the convolution of the filters with the incoming signal. These end-effects as they are known to give false alarms. This problem is mitigated by the shot noise rejection property of the D-Markov machine, as is seen in the previous chapter.

Table(3.4) gives the nominal Π matrix, while the Table(3.5) and Table(3.6), show the Π matrices for 5 Ω and 50 Ω imbalances respectively. It can be seen that the structure of these matrices changes significantly, with the increase in imbalance. Several transitions, which had no chance of occurring; for example, from q_0 to q_1 for the nominal case, occur in the case of the 50 Ω imbalance. For most real systems this does not take place, hence the process remains within a set of states that do not have self-transition probabilities of unity.

The ME partitioning on the other hand, by construction tries to equalize the probabilities of occurrence of each of the symbols. This leads to state transition

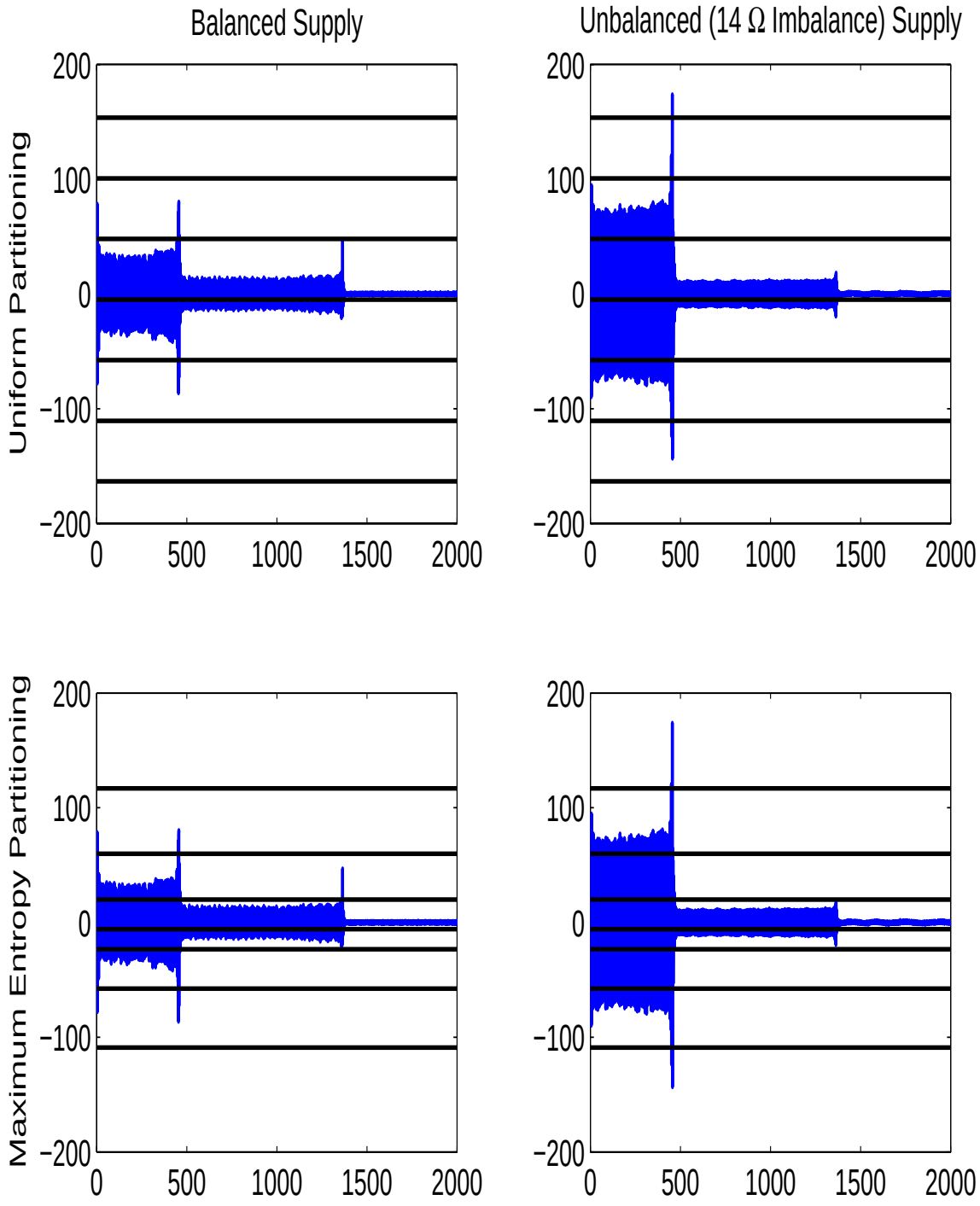

Figure 3.7. Wavelet Coefficients along with partitions

Table 3.4. Π Matrix for R = 0 Ω with Uniform Partitioning

	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7
q_0	0.4373	0.5125	0.0502	0.0000	0.0000	0.0000	0.0000	0.0000
q_1	0.1366	0.1234	0.0621	0.1954	0.3752	0.1068	0.0004	0.0000
q_2	0.0091	0.0492	0.0576	0.2005	0.2174	0.2769	0.1595	0.0298
q_3	0.0000	0.0300	0.0865	0.3411	0.1685	0.1817	0.0536	0.1386
q_4	0.0000	0.1457	0.1098	0.2683	0.3012	0.1455	0.0295	0.0000
q_5	0.0000	0.1421	0.2484	0.2185	0.3026	0.0419	0.0398	0.0067
q_6	0.0000	0.1037	0.2032	0.2666	0.0142	0.1066	0.1961	0.1096
q_7	0.0000	0.0000	0.2613	0.2651	0.0000	0.0000	0.1109	0.3627

Table 3.5. Π Matrix for R = 5 Ω with Uniform Partitioning

	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7
q_0	0.8428	0.1113	0.0444	0.0000	0.0000	0.0008	0.0008	0.0000
q_1	0.1022	0.0383	0.0483	0.1466	0.2132	0.1966	0.1716	0.0833
q_2	0.0032	0.0327	0.0713	0.1710	0.2699	0.0693	0.0785	0.3042
q_3	0.0000	0.0021	0.2280	0.3731	0.2263	0.1704	0.0000	0.0000
q_4	0.0000	0.0363	0.1792	0.2543	0.4076	0.1120	0.0106	0.0000
q_5	0.0004	0.1488	0.4842	0.1770	0.1003	0.0287	0.0405	0.0202
q_6	0.0000	0.3420	0.3106	0.0000	0.0099	0.1490	0.0512	0.1373
q_7	0.0000	0.1525	0.4113	0.0000	0.0000	0.0076	0.0564	0.3722

Table 3.6. Π Matrix for R = 50 Ω with Uniform Partitioning

	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7
q_0	0.2921	0.0109	0.0374	0.0350	0.0280	0.0310	0.0377	0.5281
q_1	0.0348	0.1024	0.0077	0.1686	0.6049	0.0794	0.0021	0.0000
q_2	0.0710	0.0591	0.0736	0.2819	0.4750	0.0105	0.0256	0.0033
q_3	0.0700	0.0508	0.0156	0.1672	0.4268	0.0928	0.1592	0.0178
q_4	0.0693	0.0127	0.0808	0.2965	0.3309	0.0642	0.1128	0.0329
q_5	0.3113	0.3241	0.3124	0.0011	0.0000	0.0000	0.0000	0.0512
q_6	0.1954	0.4100	0.3328	0.0162	0.0000	0.0000	0.0000	0.0456
q_7	0.6066	0.0159	0.0027	0.0060	0.0097	0.0094	0.0091	0.3405

matrices that have transition probabilities from almost every state to every other state, as is shown in Table (3.7). The ME technique is more effective at detecting deviations (anomalies). However, due to its construction, it is not obvious to detect the fault progression by observing the change in the state transition matrices. Tables (3.8) and (3.9) show the progression of the fault. Although there is a change in the matrices, it is hard to relate it to the underlying physical phenomenon.

Table 3.7. Π Matrix for R = 0 Ω with Maximum Entropy Partitioning

	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7
q_0	0.4020	0.0560	0.0189	0.1298	0.2018	0.1330	0.0585	0.0000
q_1	0.0585	0.0550	0.0189	0.1679	0.1311	0.1503	0.2626	0.1557
q_2	0.0163	0.0249	0.1903	0.1788	0.0931	0.1279	0.1605	0.2082
q_3	0.0457	0.1404	0.2194	0.0835	0.1314	0.0780	0.1398	0.1618
q_4	0.1573	0.1305	0.1842	0.0764	0.0959	0.1378	0.1964	0.0214
q_5	0.1353	0.0937	0.0838	0.1823	0.1932	0.2482	0.0525	0.0109
q_6	0.1522	0.2475	0.0777	0.1257	0.1535	0.1225	0.0668	0.0540
q_7	0.0326	0.2520	0.2069	0.0556	0.0000	0.0019	0.0630	0.3879

Table 3.8. Π Matrix for R = 5 Ω with Maximum Entropy Partitioning

	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7
q_0	0.4652	0.0444	0.0475	0.0456	0.0631	0.0831	0.1228	0.1284
q_1	0.0360	0.0726	0.1123	0.0523	0.1175	0.1826	0.0686	0.3581
q_2	0.0052	0.2881	0.0658	0.1851	0.0706	0.2513	0.1340	0.0000
q_3	0.0000	0.1588	0.2828	0.2208	0.1182	0.0610	0.1584	0.0000
q_4	0.0199	0.1905	0.2186	0.1167	0.1753	0.1193	0.1597	0.0000
q_5	0.0819	0.1784	0.0444	0.1214	0.3237	0.2023	0.0317	0.0162
q_6	0.1769	0.4928	0.1319	0.0113	0.0207	0.0666	0.0324	0.0675
q_7	0.2062	0.3739	0.0000	0.0000	0.0000	0.0018	0.0494	0.3688

The state probability vectors give a more intuitive feel for the underlying process. In fact, a suitable metric applied to the difference between the nominal condition, which is the one without imbalance and the ones with imbalance gives us the the anomaly/fault measure.

Table 3.9. Π Matrix for R = 50 Ω with Maximum Entropy Partitioning

	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7
q_0	0.2737	0.0313	0.0207	0.0372	0.1071	0.0356	0.0399	0.4545
q_1	0.1302	0.0746	0.1383	0.1417	0.0583	0.4176	0.0197	0.0197
q_2	0.1752	0.0249	0.1370	0.0125	0.0282	0.5424	0.0507	0.0291
q_3	0.0806	0.0054	0.1242	0.0504	0.0564	0.3116	0.1484	0.2230
q_4	0.0690	0.0682	0.0727	0.0004	0.3194	0.1074	0.1388	0.2240
q_5	0.1452	0.1156	0.1122	0.3675	0.2265	0.0165	0.0000	0.0165
q_6	0.6415	0.2995	0.0019	0.0000	0.0000	0.0000	0.0000	0.0570
q_7	0.6237	0.0470	0.0054	0.0025	0.0044	0.0056	0.0089	0.3025

Table 3.10. **p**-Vectors for R = 0 Ω, 5 Ω and 50 Ω with Uniform Partitioning

	0 Ω
q_0	0.0255
q_1	0.0965
q_2	0.1276
q_3	0.2574
q_4	0.2141
q_5	0.1365
q_6	0.0675
q_7	0.0750

	5 Ω
q_0	0.0514
q_1	0.0720
q_2	0.2130
q_3	0.1874
q_4	0.2113
q_5	0.0949
q_6	0.0445
q_7	0.1256

	50 Ω
q_0	0.2685
q_1	0.0574
q_2	0.0608
q_3	0.1103
q_4	0.1800
q_5	0.0375
q_6	0.0518
q_7	0.2337

Tables(3.10) and (3.11) respectively list the state probability vectors for detecting stator faults using uniform partitioning and ME partitioning. Fig(3.8) displays a histogram representation of state probability vectors in Table (3.10). This visualization shows how the structure of the underlying distribution changes as the fault progresses.

Uniform and ME partitioning techniques were investigated to capture the changes in the voltage imbalance; all three metrics 1-norm, the standard deviation and the Kullback distance were used for computation of the anomaly measure. The results are displayed in Fig(3.9) and Fig(3.10). Since the problem of detecting voltage imbalance can be treated as that of a parameter variation, a more suitable technique is one that shows an essentially linear increase in the anomaly measure

Table 3.11. **p**-Vectors for R = 0 Ω, 5 Ω and 50 Ω with Maximum Entropy Partitioning

	0 Ω		5 Ω		50 Ω
q_0	0.1250	q_0	0.1280	q_0	0.3278
q_1	0.1250	q_1	0.2086	q_1	0.0590
q_2	0.1250	q_2	0.1009	q_2	0.0481
q_3	0.1250	q_3	0.0825	q_3	0.0596
q_4	0.1250	q_4	0.1106	q_4	0.0967
q_5	0.1250	q_5	0.1235	q_5	0.0944
q_6	0.1250	q_6	0.0888	q_6	0.0414
q_7	0.1250	q_7	0.1571	q_7	0.2730

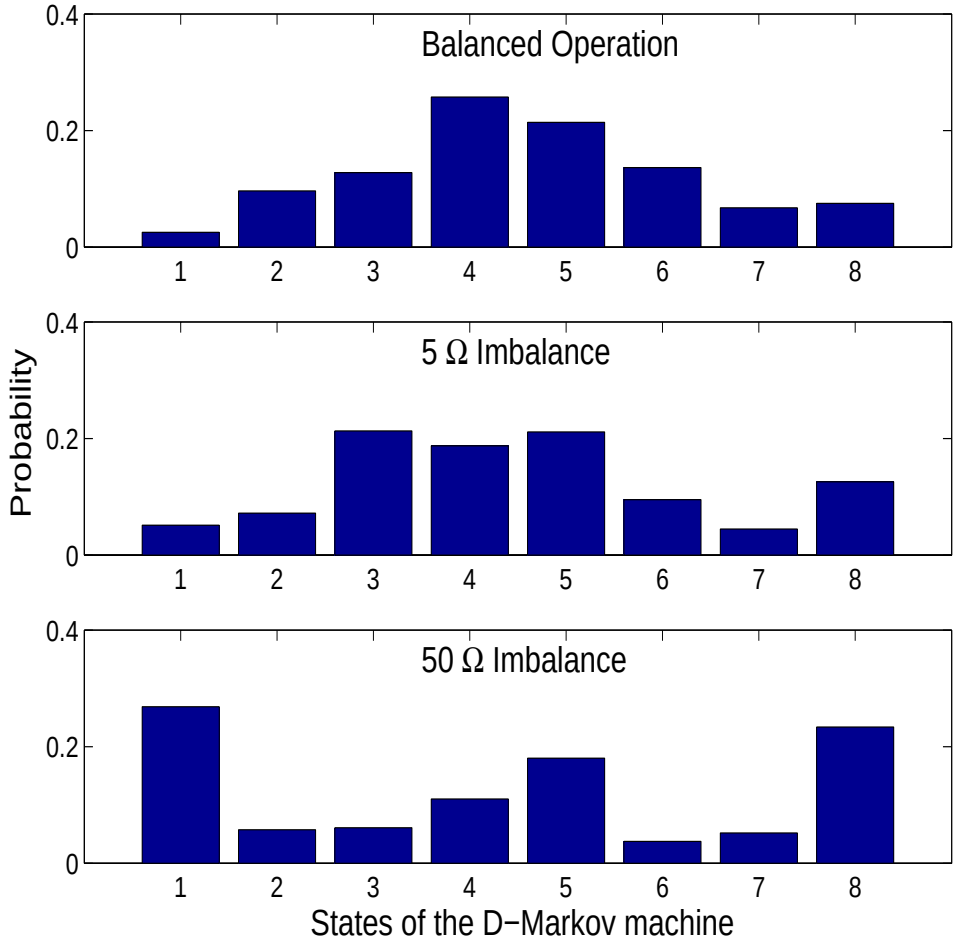

Figure 3.8. Probability histograms under different conditions

with the imbalance. From this perspective, ME partitioning performs better than uniform partitioning and Kullback distance is the best measure. The standard deviation works in the case of the ME partitioning because the nominal case is partitioned such that each of the states has a probability equal to the mean. This does not occur in the case of uniform partitioning and hence it fails in that case.

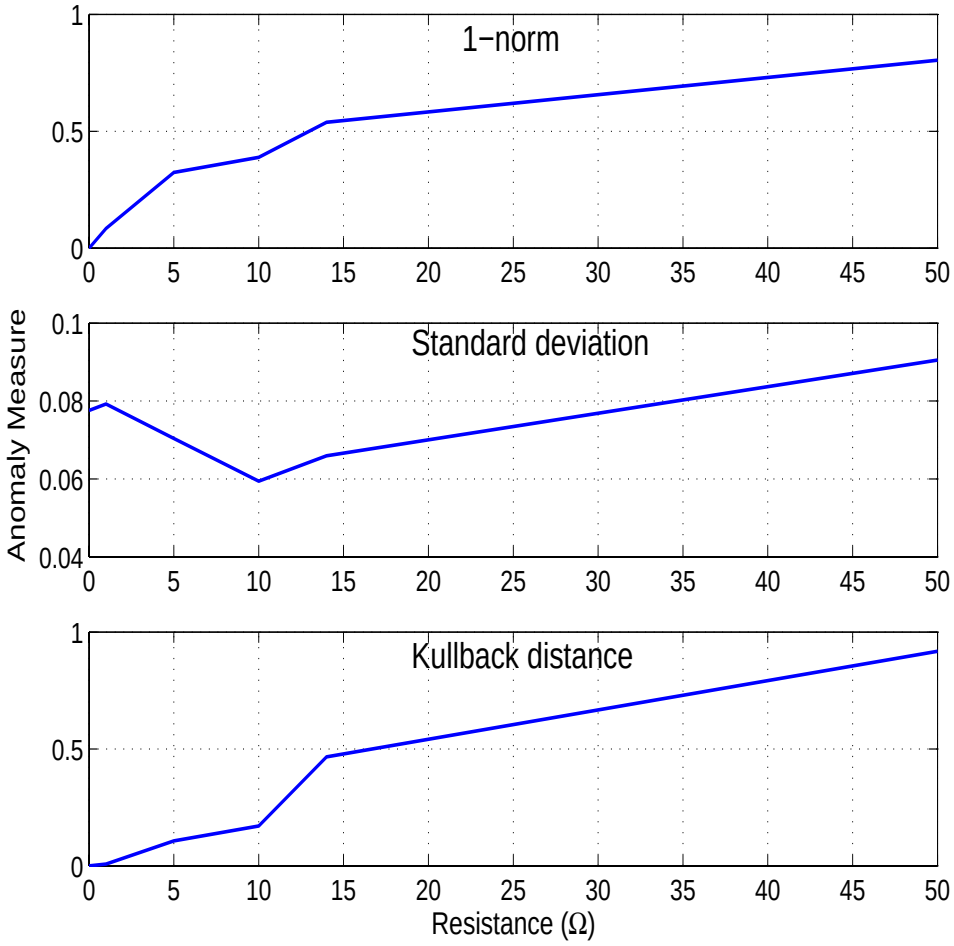

Figure 3.9. Uniform Partitioning with Depth = 1

3.3.1 Comparison with other pattern recognition techniques

This section compares the symbolic dynamics methodology with other techniques, namely PCA (Principal Component Analysis) and neural networks, details of each

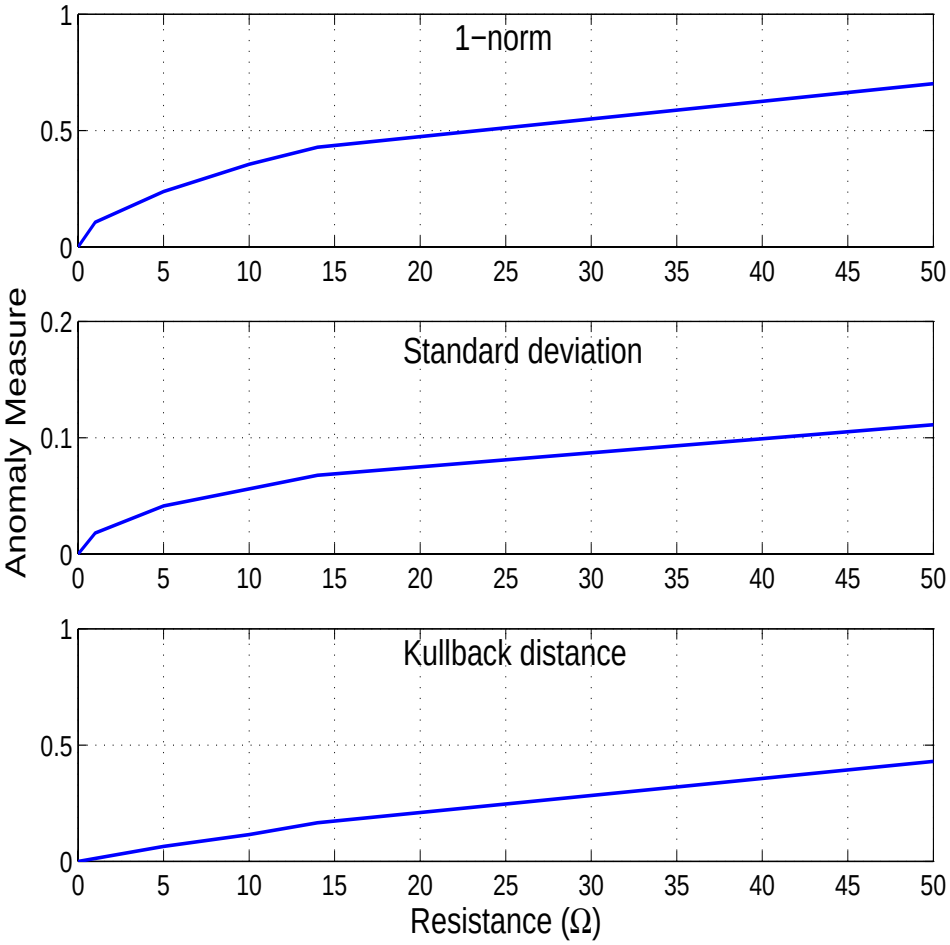

Figure 3.10. Maximum Entropy (ME) Partitioning with Depth = 1

of these techniques is given in Appendix-C. PCA is computationally simple, it essentially finds the changes in orientation of the principal eigen vector of the correlation matrix [51]. Fig(3.11) shows the performance of the PCA method. PCA is not a very effective method for detection of induction motor faults. This is due to the fact that the PCA method only looks at the changes in the direction of the principal eigen vector, which in this case corresponds to f the line frequency.

Neural networks have been used extensively for fault detection in induction machines [52] [20]. The analysis is performed for one such neural network. This neural network has 4 layers excluding the input and output layers. Hidden layers

have 50, 40, 30, 40 neurons respectively. The output layer has 10 neurons. The input layer has just 1 neuron, since the data set is the Parks vector modulus and is one dimensional. The pattern matrix has 30 rows and 200 columns. The neural network is a feed-forward network, trained using error back propagation with the training error being 10^{-5}. Fig(3.11) shows a comparative evaluation of all three methods. The idea is to look for the most linear response, the reason for this has been discussed in the earlier sections. Additionally, at the very least we also must ensure that for increasing levels of imbalance the fault/anomaly measure also increase. The author does not make a claim that this is the optimal neural network for this application. As with any non-linear techniques there exists no standard metric/benchmark for comparison. The purpose of this exercise is to illustrate the robustness of the symbolic wavelet analysis method.

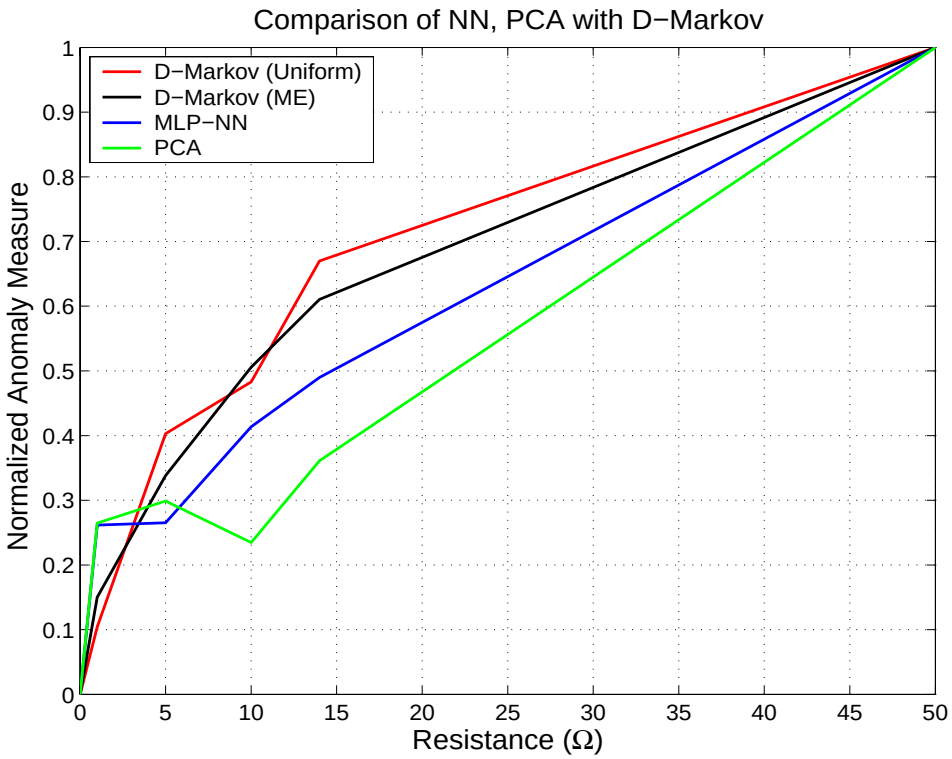

Figure 3.11. Comparison with MLP-NN and PCA

The entire premise of the detection problem is the early detection with a low-false alarm. To this end it is important that the fault anomaly measure be linear in nature, when plotted against the imbalance. Thus, when compared to MLP-NN and PCA, our technique performs better in this particular case. For a similar type of a fault, other methods like neural networks could detect a change only when the imbalance exceeded 0.2 pu which in this case is 28 Ω as reported in [22].

3.4 Summary

This chapter has dealt with the detection of voltage imbalances. The framework developed has been compared using both the CWT and the DWT analysis for simulation and experimental results respectively. The D-Markov machine had a fixed depth of 1 and the number of symbols were fixed to be 8. In the following chapter, we will look at the problem of broken rotor bar detection. We will also use the statistical nature of the framework that has been developed, and not been used thus far. In addition, we shall compare the performance of CWT and DWT.

Chapter 4

Rotor Bar Faults

The previous chapter presented a detailed study of the stator voltage imbalance, which was created externally, and is considered an electrical fault. Data both from a simulation and an experiment were analyzed. Results from the CWT and DWT implementations of wavelets were studied. In this chapter, we study the detection of rotor bar failures, which account for 10% of all induction motor failures. The failure mechanism was discussed in detail in Chapter 1. In this chapter we will look at more advanced analysis, namely the effects of depth and change in the alphabet size. We will also look at the comparison between CWT and DWT as pre-processing tools. Finally, we will also look at a comparison between our technique and techniques used by other researchers. Most importantly, a statistical flavor is demonstrated, that utilizes different sets of data at the same damage level to demonstrate the statistical nature of the anomaly measure obtained. The framework will be validated using data from an MEC simulation as well as data from the motor-dyne test bed. In the later half of the this chapter, inverter driven (V/Hz operation) induction motor operation is also studied. The inverter further complicates the signal isolation problem due to the switching harmonics. In all the cases current signals will be the only signals used in the detection process. This

will allow for the diagnostic unit and sensors to be located away from the motor.

4.1 Detection of Broken Rotor Bars

The so-called squirrel cage of an induction motor is depicted in Fig(4.1). The cage is comprised of axial rotor bars that terminate in the two end rings. The bars and end-rings carry the induced current. Rotor bars break due to various reasons, as

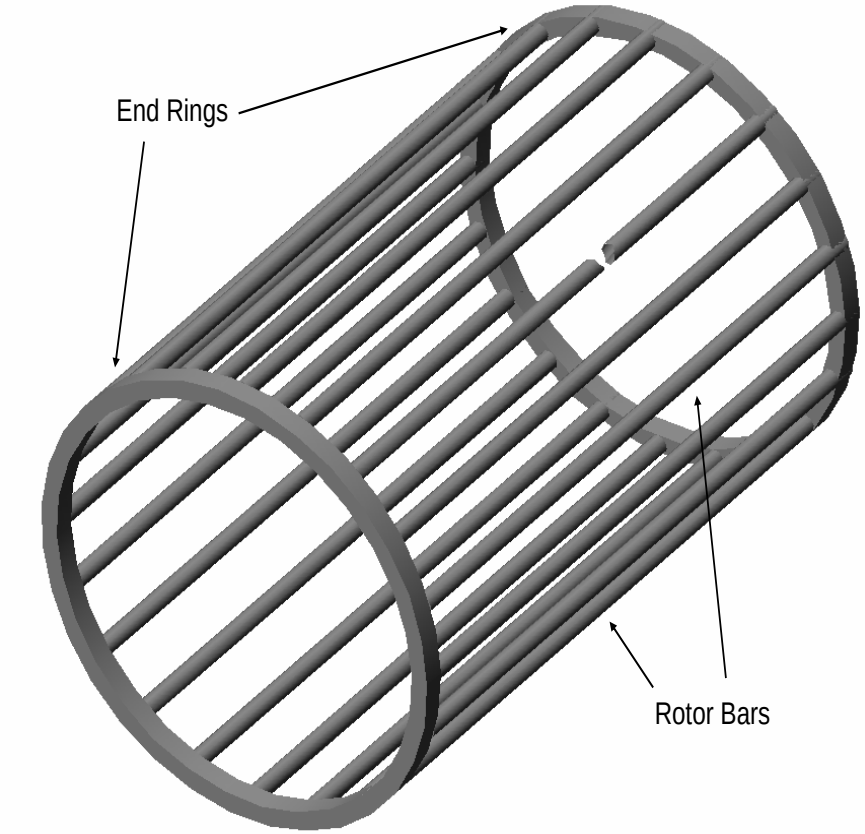

Figure 4.1. Cage with a broken rotor bar

discussed previously in Chapter 1. The broken rotor bars are perhaps the hardest fault to detect and costliest in terms of maintenance. There is also no possibility

of mounting a sensor on the rotor or inside the stator to detect rotor bar failure. However, the redeeming feature of this particular fault is the ability of the machine to continue to operate even with broken rotor bars. We will use the framework to detect broken rotor bars in the induction machines. The breakage of rotor bars causes a change in the electromagnetic field.

It was found that broken rotor bars introduce an additional frequency component in the line current signal [53] [54] [55].

$$F = f\left[\frac{K}{P}(1-s) \pm s\right] \tag{4.1}$$

where f, P and s are, respectively, the electrical supply frequency, number of pole pairs, and the per-unit slip; $K/P = 1, 3, 5, 7, 11, 13, \ldots$ due to the normal winding configuration. For $K/P = 1$, the additional component in the current spectrum is $(1-2s)f$ due to the broken rotor bar and $(1+2s)f$, due to the speed oscillations. These fault signatures naturally will also appear in the i_d and i_q signals. The phase current signals can be written as

$$\begin{aligned}
i_A &= i_f cos(\omega t - \alpha_0) + i_l cos((1-2s)\omega t - \beta_l) + i_r cos((1+2s)\omega t - \beta_r) \\
i_B &= i_f cos(\omega t - \alpha_0 - \frac{2\pi}{3}) + i_l cos((1-2s)\omega t - \beta_l - \frac{2\pi}{3}) + i_r cos((1+2s)\omega t - \beta_r - \frac{2\pi}{3}) \\
i_C &= i_f cos(\omega t - \alpha_0 + \frac{2\pi}{3}) + i_l cos((1-2s)\omega t - \beta_l + \frac{2\pi}{3}) + i_r cos((1+2s)\omega - \beta_r + \frac{2\pi}{3})
\end{aligned}$$

where i_f, i_l and i_r are the peak values of, respectively, the fundamental supply phase current, the lower sideband component at frequency $(1-2s)f$, and the upper sideband component at frequency $(1+2s)f$; three angles α_0, β_l and β_r denote the initial phase angle, respectively, for the fundamental component, the lower sideband component and, the upper sideband component. Here again the

Parks vector modulus is computed:

$$|i_d + ji_q|^2 = \frac{2}{3}(i_f^2 + i_l^2 + i_r^2) + 3i_f i_l cos(2s\omega t - \alpha_0 + \beta_l)$$
$$+3i_f i_r cos(2s\omega t + \alpha_0 - \beta_r) + 3i_r i_l cos(4s\omega t + \beta_r + \beta_l) \quad (4.2)$$

The fault frequencies are clearly concentrated at frequencies of $2sf$ and $4sf$, with no component at the fundamental supply frequency. Although there are bound to be variations in speed and consequently changes in the slip; but if the rough operating point is known, wavelet analysis can be used to zoom in and extract the fault information. The next section focuses on online information extraction using wavelet analysis.

The inverter operation involves analyzing a far richer signal. In addition to the fundamental signal inverters introduce harmonics due to the switching action. The harmonics introduced have the following frequencies $(2m + 1)f$ where $m = 0, 1, 2 \ldots$. Since the motor under study is 3-phase, triple harmonics are eliminated. So the harmonics present in the stator current waveform are$f, 5f, 7f, 11f, 13f, \ldots$. The higher harmonics are weighted by $\frac{1}{m}$ where m is the harmonic number, so harmonics beyond 15 can be ignored. This still leaves fault signatures in several frequency bands.

4.2 Simulation Studies

We used the Magnetic Equivalent Circuit analysis for our simulations, since of the two methods this is computationally less intensive, and yet gives us realistic fault models [48]. The machine simulated had 30 stator teeth and 42 rotor teeth. For the initial part, extensive studies were conducted on an unloaded machine with a number of broken rotor bars. This was done by increasing the resistance of the

rotor bars from their nominal values $0.142m\Omega$ up to 1Ω. This 2000 fold increase in the resistance of the bar effectively emulates breakage of the bar. A line-fed machine was simulated without load; this is acceptable in a simulation since there is no noise and the effects of the damage can be seen even without loading the machine. However, during an experimental analysis better results are obtained when the machine is operated on load.

4.2.1 Simulation Results

The line currents are the signals used to detect the broken rotor bars. From Eqn(4.2), we get a dc component and frequency content of $2sf$ and $4sf$. To eliminate the dc component, we can subtract the mean of the signal, which will give only the frequency component which is indicative of the failure. Alternately, as done in this case, the dc offset is ignored and the wavelet analysis performed only at the scales of interest. In this case the wavelet used was 'db1', which is well localized in time but poorly localized in frequency. The results demonstrate that regardless of which wavelet is chosen, as long as the scales are chosen appropriately the framework is able to detect the fault. This is due to the property of the wavelets to span the entire basis [37]. The scales must be chosen to coincide with the pseudo frequency. The coefficients were stacked end to end (scale-based), this data stream was then converted to a symbol stream with an alphabet size of eight. The depth was fixed at 1, hence there are only 8 states in the D-Markov machine. In some cases the Power Spectral Density (PSD) at the fault frequencies gives an indication of the fault. The PSD of the Parks vector modulus is shown in Fig(4.2). As the number of broken bars increases we can see that the PSD also increased in amplitude in the frequency region of the fault signal. This is not evident in the case of experimental data, and thus the framework is used in the extraction of this

information.

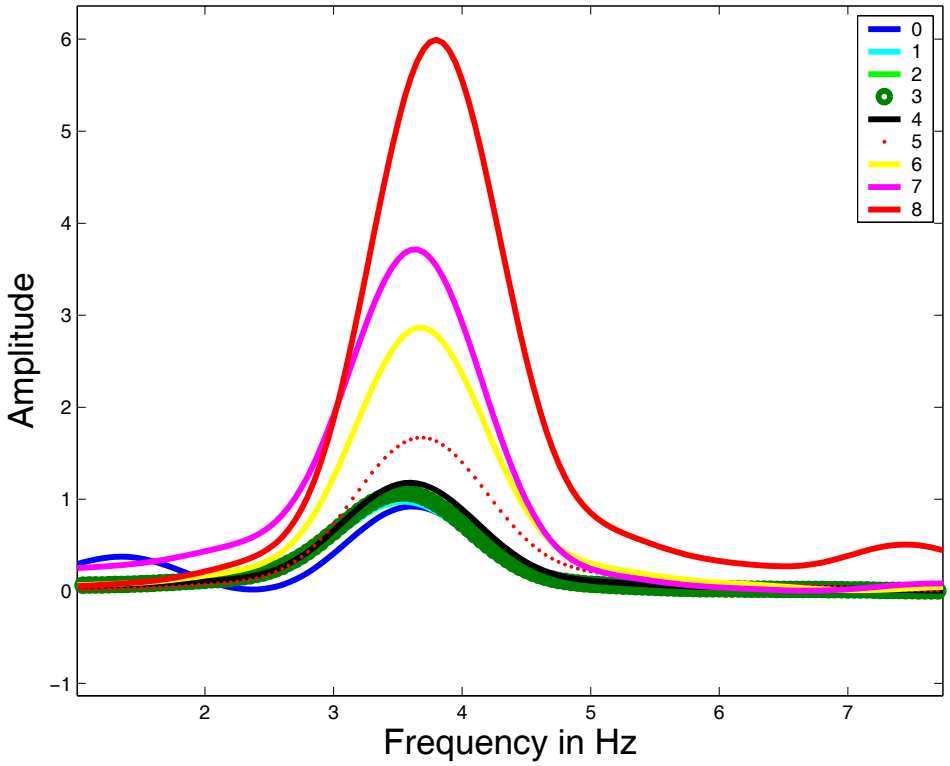

Figure 4.2. Power Spectral Density of the Stator Currents with Rotor Bars Damage

For analyzing the data we compared three techniques Neural Network Multi-Layer Perceptron (NN-MLP), Neural Network Radial Basis Function (NN-RBF) and the D-Markov method. Implementation details for the NN-MLP, NN-RBF are given in Appendix-C. In the first analysis as shown in Fig(4.3), we provided the wavelet coefficients as data to all the three methods. We can see that the D-Markov method is able to detect the anomaly right away. While NN-MLP takes 6 rotor bars to be broken before it can detect the change. All the fault measures were normalized, so that we can compare the methods. In the second analysis as shown in Fig(4.4), we gave the raw data, i.e. Parks vector modulus to all three methods. In this case the NN-MLP and the NN-RBF seem to perform better. However, closer examination reveals that there are inconsistencies. The anomaly measure

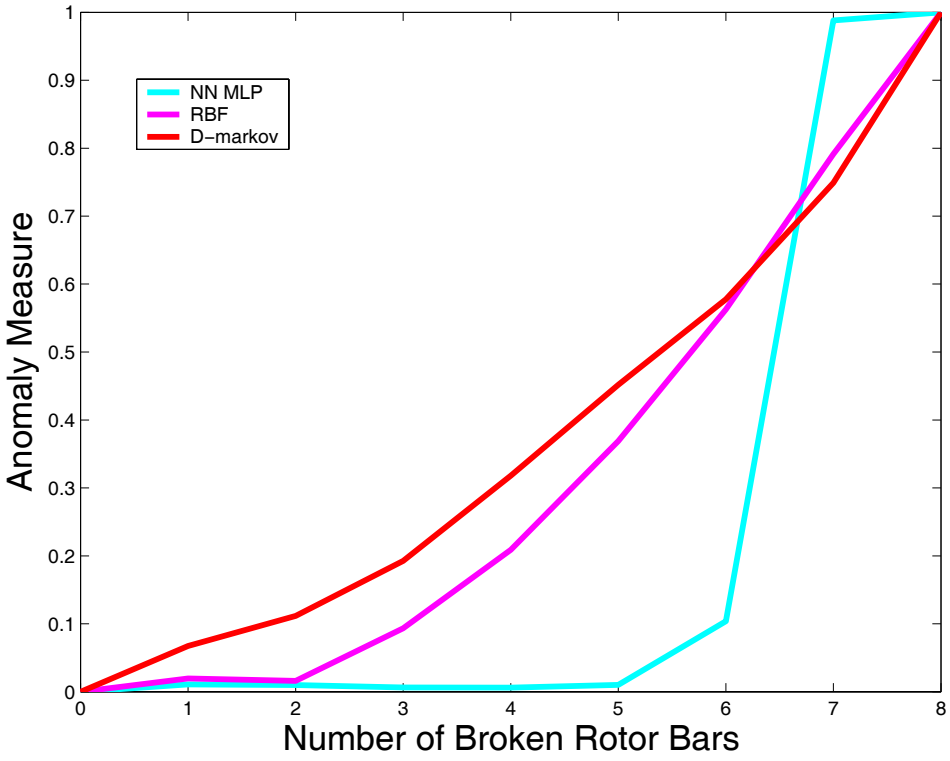

Figure 4.3. Comparison of various methods after Wavelet Analysis

increases and then decreases for increasing fault magnitude. The anomaly measure is like a fault residual, i.e. it should increase monotonically for increasing value of faults. Additionally, this particular problem can also be viewed as parameter variation problem, with the number of failed bars being the parameter. Thus, a linear anomaly curve is the most desirable. We see that the D-Markov method without wavelet analysis gives slightly better results than the D-Markov method with wavelet analysis. This could be due to an improper choice of scales or that we might have missed a higher frequency component of the fault, which is likely possible with the use of the 'db1' wavelet which has a lower number of vanishing moments. This is not the case as we shall see in the next few sections. However, it can be argued that both the D-Markov method, with and without wavelet analysis performed consistently with the increase in the number of broken rotor bars.

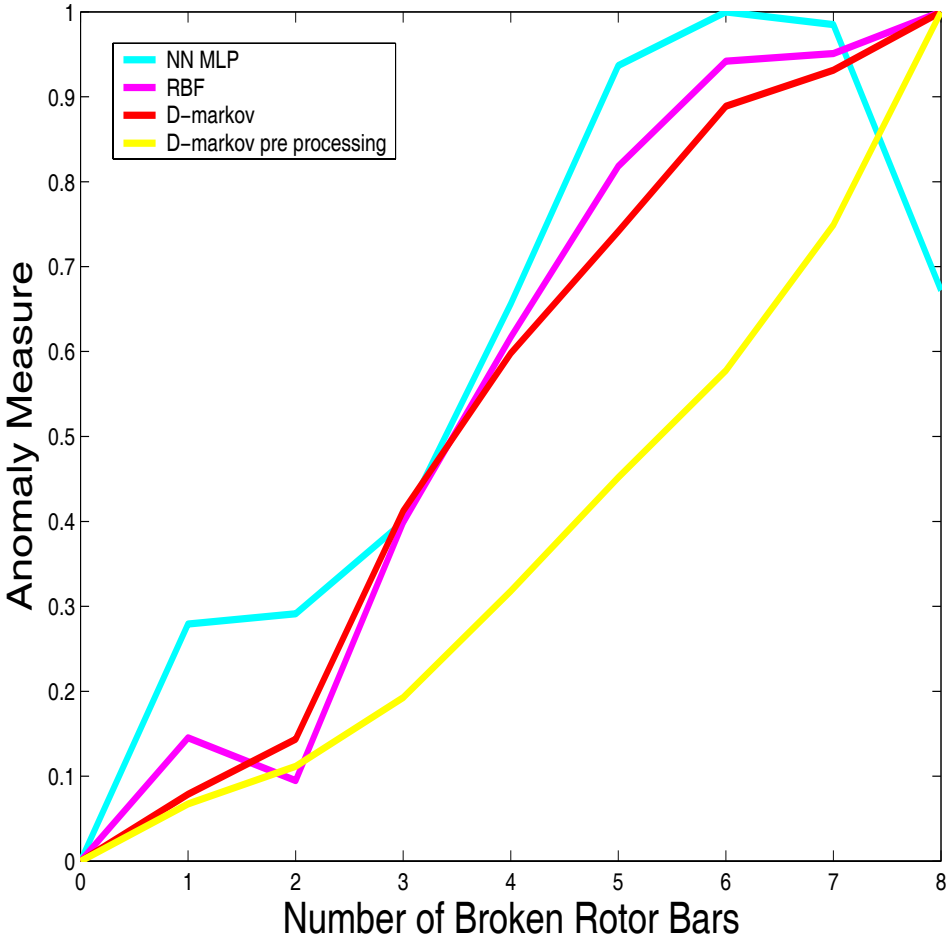

Figure 4.4. Comparison of various methods without Wavelet Analysis

In this section we have looked at the failure of several rotor bars, and the progression of the anomaly as each successive bar fails. However, this MEC simulation does not model the mechanical interaction of the rotor which has a failed rotor bar. Additionally, due to it being a simulation the results do not address the statistical nature of the fault detection framework. This will be motivated and studied in detail for a single bar failure.

4.3 Experimental Results

4.3.1 Line-Fed Motor

The rotor bar damage was created by drilling a hole into the rotor. The drilling of the rotors has an effect similar to that of a breakage of a rotor bar. It removes a parallel path in the rotor circuit. It was found that larger levels of rotor damage generate large amounts of mechanical oscillations; hence, only motors with a single rotor bar damaged were tested. Two motors were tested; one with one broken rotor bar, and a healthy motor. Five runs each of the healthy and faulty motors were performed. The motors were loaded to 3 N-m and data was collected for 100 seconds. For the analysis, 40 seconds of steady-state operating data were used. The wavelet used was 'db4' at decomposition levels 7,8 and 9. These correspond to scales 0.0078, 0.0039 and 0.0020, which are in the neighborhood of the fault spectra given by Eqn(4.2).

Table (4.1), compares ME partitioning with uniform partitioning for the nominal rotor (Nom) and the rotor with a broken rotor bar (Brb). For ME partitioning in the nominal case yields no information because the space is partitioned such that each symbol has equal probability.

Table 4.1. p-Vectors for Nominal (Nom) and Broken Rotor bars (Brb) with Uniform and Maximum Entropy Partitioning

	Uniform			Max Entropy	
	Nom	Brb		Nom	Brb
q_0	0.0049	0.0560	q_0	0.1241	0.3528
q_1	0.0024	0.0316	q_1	0.1241	0.0633
q_2	0.1776	0.2944	q_2	0.1265	0.0560
q_3	0.7518	0.4380	q_3	0.1241	0.0316
q_4	0.0511	0.1046	q_4	0.1241	0.0779
q_5	0.0049	0.0389	q_5	0.1265	0.0706
q_6	0.0049	0.0341	q_6	0.1241	0.0803
q_7	0.0024	0.0024	q_7	0.1265	0.2676

Due to the DWT filter-bank implementation, the pre-processed signal has only 458 data points. This causes the D-Markov machine to be less effective. Additionally, there could be a possibility that the probability distribution does not converge to a stationary distribution. Despite this, the counting process is able to distinguish between the motor with the broken rotor bars and those without.

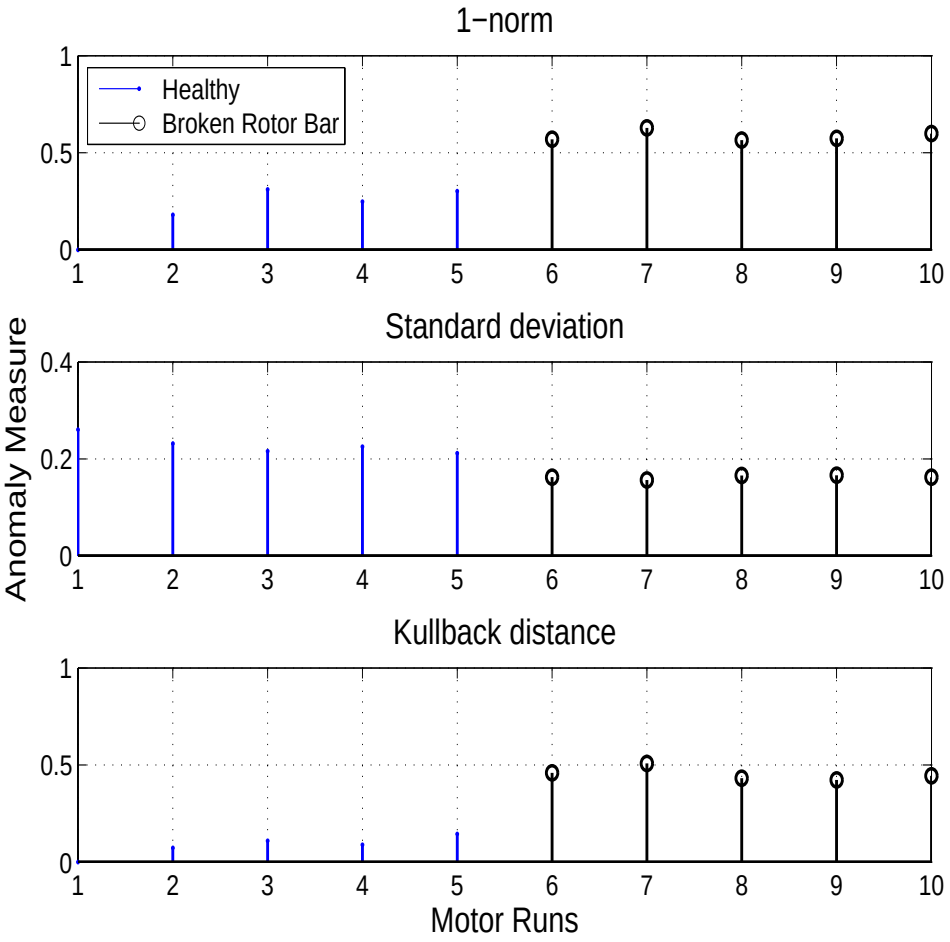

Figure 4.5. Anomaly Measure using 'db4' wavelet and Depth = 1

Both uniform and ME partitioning performed comparably and were able to detect the fault. For brevity only the results from the uniformly partitioned data is presented. At depth $D = 1$ and $\mid \Sigma \mid = 8$, the number of states is 8; it is seen in Fig.(4.5) that all three metrics, 1-norm, standard deviation, and Kullback

distance, are able to differentiate between the healthy and broken rotor bars.

4.3.2 Inverter-Fed Results

Experiments consisted of five runs of the normal motor and five runs of the faulty motor (one broken bar) for analysis. The load on the motor was 3 N-m and the slip was observed to be $\frac{1}{60}$ for all the runs. The stator currents were sampled at $10kHz$ and data was acquired for 45 seconds, resulting in $450,000$ data points for each run. Only 25 seconds of data between the 10^{th} and 35^{th} second were used for the analysis. The 'db4' wavelet was used, as it has the maximum correlation among the wavelets considered. The rotor bar breakage faults manifest at frequencies mentioned in Section(4.1). Substituting these frequencies in place of f_p in Eqn(2.3), the scales are calculated. In DWT, these scales correspond to decomposition levels 11 and 12.

Wavelet decomposition was performed on the acquired data and 'detail' coefficients corresponding to levels 11 and 12 are considered for symbolic analysis. Fig(4.6) show the 'detail' coefficients for run 1 (normal) and run 6 (faulty). As seen from the plots, the coefficients are significantly different from each other and hence are the right choice to distinguish between the normal and faulty motors.

The alphabet size $|\Sigma|$ was chosen to be eight. The 'detail' coefficients, corresponding to normal condition (run 1), are stacked to form the scale series. The scale series is partitioned under ME criterion to obtain the partitions. These partitions remain invariant in the further analysis. With the partition generated, the symbol sequences are generated from wavelet coefficients for all the runs. The state probability vectors of these symbol sequences are then obtained. Figs(4.7) and (4.8) show the histograms of the state probability vectors for runs 1 and 6 respectively.

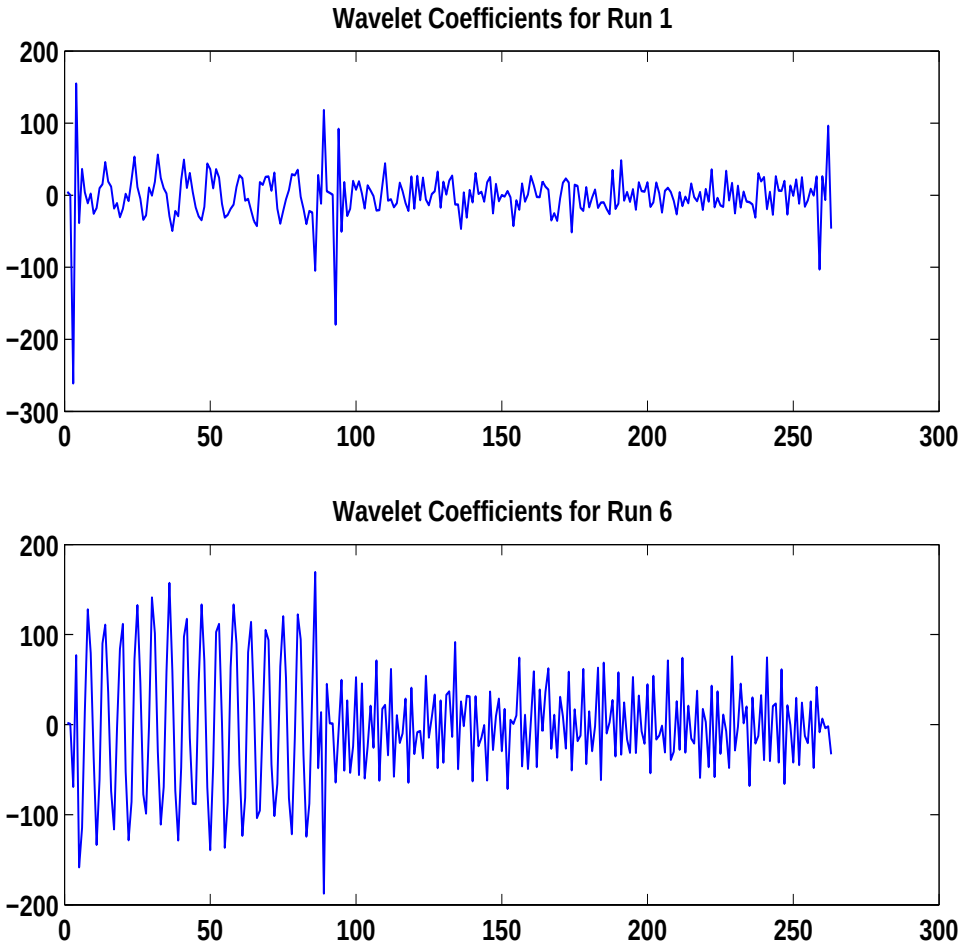

Figure 4.6. Plot of Discrete Wavelet Coefficients

It can be seen clearly from the figures that the probability distributions of the two runs differ substantially from each other. To quantify the difference, the relative entropy or the Kullback distance between the distributions is used as a measure. p^k represents the distribution corresponding to run k and ${p_i}^k$ is the probability of symbol i in the distribution p^k. Runs 1 through 5 correspond to the normal motor while runs 6 through 10 correspond to the faulty motor.

$$M(k) = -\sum_{i=1}^{i=|\Sigma|} {p_i}^k log_2 \frac{{p_i}^k}{{p_i}^1} \tag{4.3}$$

Though an alternative definition of relative entropy is possible, this definition

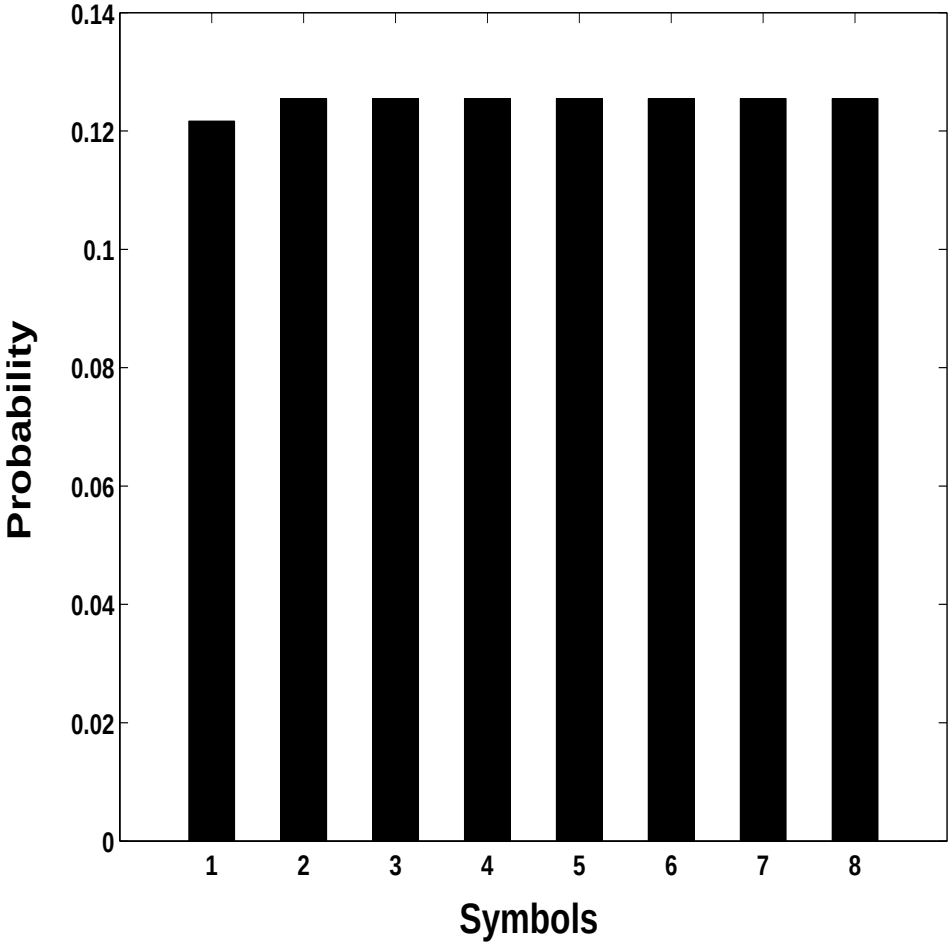

Figure 4.7. Probability Density for Run 1

was chosen in order to avoid generating a measure of ∞. It should be noted that, under ME partition, p^1 has a uniform distribution and hence the measure is always finite. This cannot be guaranteed for the p^k distribution, or for uniform partitioning. Fig(4.9) depicts the measures obtained for all the runs.

The anomaly measures of runs 1 through 5 are very small and the for runs 6 through 10 the measures are large. Thus it can be inferred that the wavelet based symbolic approach differentiates between normal and faulty motors. Moreover, anomaly measures of the normal motor runs are in a band (< 0.1) quite distinct from faulty runs (> 0.48). Thus there is no possibility of false classification or false

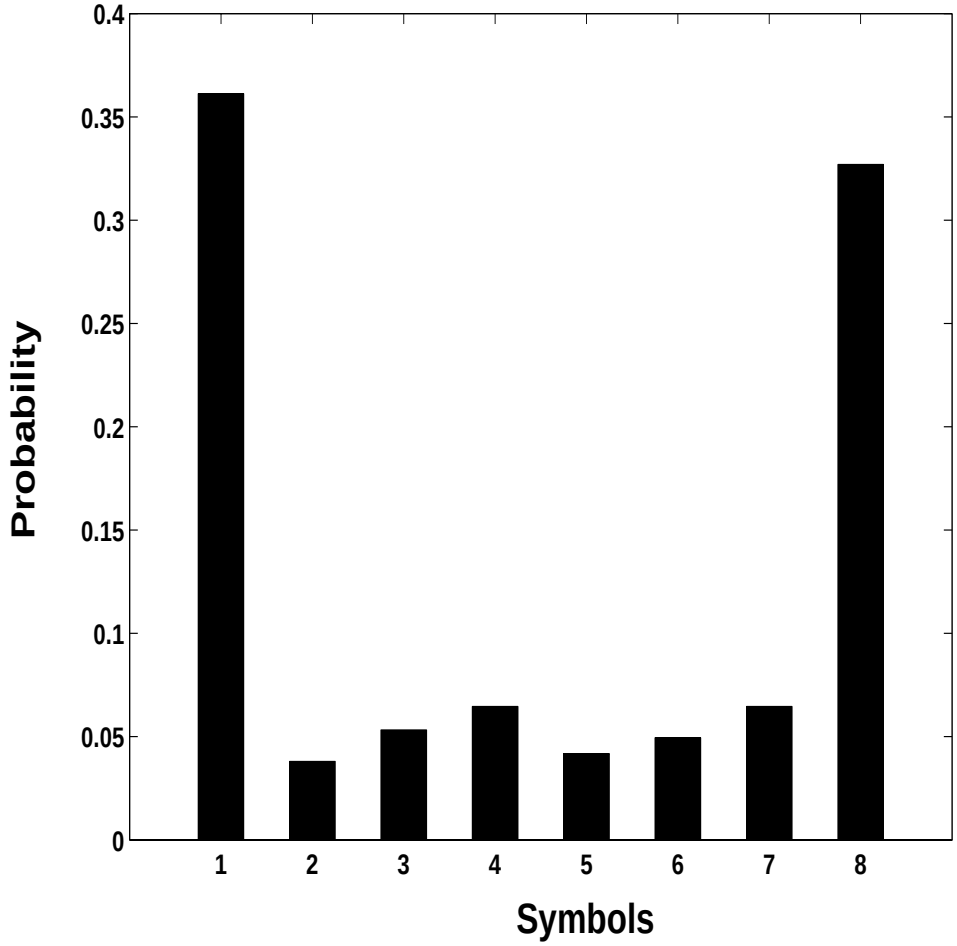

Figure 4.8. Probability Density for Run 6

alarms, which are particularly high in inverter fed machines.

4.3.3 Effect of Depth and Number of Symbols

Changing the number of symbols and depth does have an effect on the structure of the D-Markov machine, and correspondingly on the entire fault detection process. For example, the change in the number of symbols from $|\Sigma| = 8$ to $|\Sigma| = 16$, does increase the accuracy of the process, however, if the number of amount of scale-series data is insufficient there is the danger of the state transition matrix (Π) not converging to stationary behavior. Moreover, the state transition proba-

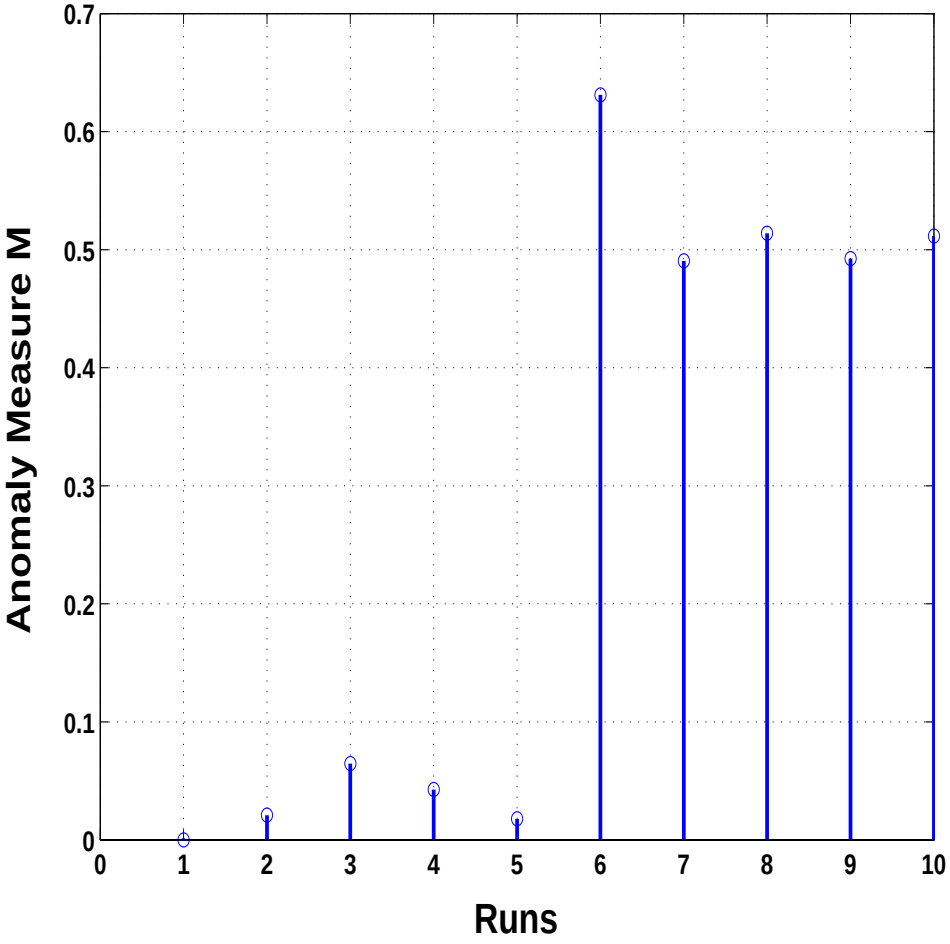

Figure 4.9. Anomaly Measure of Experimental Runs

bilities maybe insignificant and in many cases non-existent thus making the state transition matrix extremely sparse. A good example of this is shown in Table(4.2).

Table(4.3) also has 16 states. In this table each of the states are not equal to a symbol, in fact 2 symbols make a state. Thus the depth $D = 2$, and hence the number of states is $|\Sigma|^D$ where $|\Sigma|$ is the number of symbols. This construction is particularly useful if there is a good understanding of the memory properties of the process. Generally, the matrices obtained tend to be more sparse, due to the large number of illegal transitions (zero probability of transition between many state pairs). In most physical processes with no known Markov property, $D = 1$

Table 4.2. Π-Matrix with Uniform Partitioning for Healthy Motor with 16 symbols and Depth =1

	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7	q_8	q_9	q_{10}	q_{11}	q_{12}	q_{13}	q_{14}	q_{15}
q_0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_2	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_3	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_4	0.00	0.00	0.00	0.00	0.13	0.26	0.26	0.26	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_5	0.00	0.00	0.00	0.00	0.06	0.06	0.52	0.22	0.10	0.00	0.00	0.02	0.02	0.00	0.00	0.00
q_6	0.00	0.00	0.00	0.00	0.02	0.05	0.80	0.09	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_7	0.00	0.02	0.00	0.00	0.13	0.48	0.27	0.00	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_8	0.00	0.00	0.00	0.05	0.10	0.24	0.24	0.29	0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
q_{10}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
q_{11}	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_{12}	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_{13}	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_{14}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
q_{15}	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

performs as well as $D > 1$, without the additional problem of dealing with large state transition matrices.

Table 4.3. Π-Matrix with Uniform Partitioning for Healthy Motor with 4 symbols and Depth =2

	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7	q_8	q_9	q_{10}	q_{11}	q_{12}	q_{13}	q_{14}	q_{15}
q_0	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_1	0.00	0.00	0.00	0.00	0.00	0.67	0.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_2	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_3	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_4	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_5	0.00	0.00	0.00	0.00	0.01	0.94	0.05	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.86	0.10	0.00	0.00	0.00	0.00	0.00
q_7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00
q_8	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_9	0.00	0.00	0.00	0.00	0.00	0.95	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_{10}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
q_{11}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
q_{12}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
q_{13}	0.00	0.00	0.00	0.00	0.00	0.67	0.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
q_{14}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
q_{15}	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00

Table (4.4) shows the state probability vectors for the healthy motors with two different D-Markov machine structures. Notice that our initial assertion that the increasing depth causes some of the states not be visited at all holds true in this case as well. The results for $D = 2$ and $|\Sigma| = 8$ are shown in Fig(4.10), where the number of states is fixed at 64. The state transition matrices as well as the state probability vectors are not shown here due to the prohibitively large size. While the metrics, 1-norm and standard deviation, are seen to have comparable performance. The Kullback distance is ineffective due to the large number of state pairs with 0 transition probabilities, thus causing the measure to go to ∞.

Table 4.4. **p**-Vectors

16 Symbols, Depth =1		4 Symbols, Depth =2	
q_0	0.0024	q_0	0.0000
q_1	0.0024	q_1	0.0073
q_2	0.0000	q_2	0.0000
q_3	0.0024	q_3	0.0000
q_4	0.0560	q_4	0.0049
q_5	0.1217	q_5	0.8659
q_6	0.6350	q_6	0.0512
q_7	0.1168	q_7	0.0073
q_8	0.0511	q_8	0.0024
q_9	0.0000	q_9	0.0488
q_{10}	0.0000	q_{10}	0.0049
q_{11}	0.0049	q_{11}	0.0000
q_{12}	0.0024	q_{12}	0.0000
q_{13}	0.0024	q_{13}	0.0073
q_{14}	0.0000	q_{14}	0.0000
q_{15}	0.0024	q_{15}	0.0000

4.3.4 Comparison with CWT

Fig(4.11) shows the anomaly measure plot using the 'gaus2' wavelet computed using the CWT across scales 100, 150, 200, 250, 300, 350 and 400. The measure used is the 1-norm. The most important distinguishing improvement is noticed in the nominal case the first 10 runs of data, where the deviation from the benchmark data set is minimal.

In this case the methodology is able to perform better than the case of DWT, for several reasons. The most prominent being the amount of data used is significantly greater than the DWT case, which is decimated by 2 for every level of decomposition Fig(2.8). The number of scales of expansion is higher, since in using the CWT, we can choose any number of scales. The wavelet has minimal effect on the detection, although, being symmetric does give the 'gaus2' wavelet an advantage over the nearly symmetric 'db4'.

The biggest disadvantage to this method is the difficulty of real-time implemen-

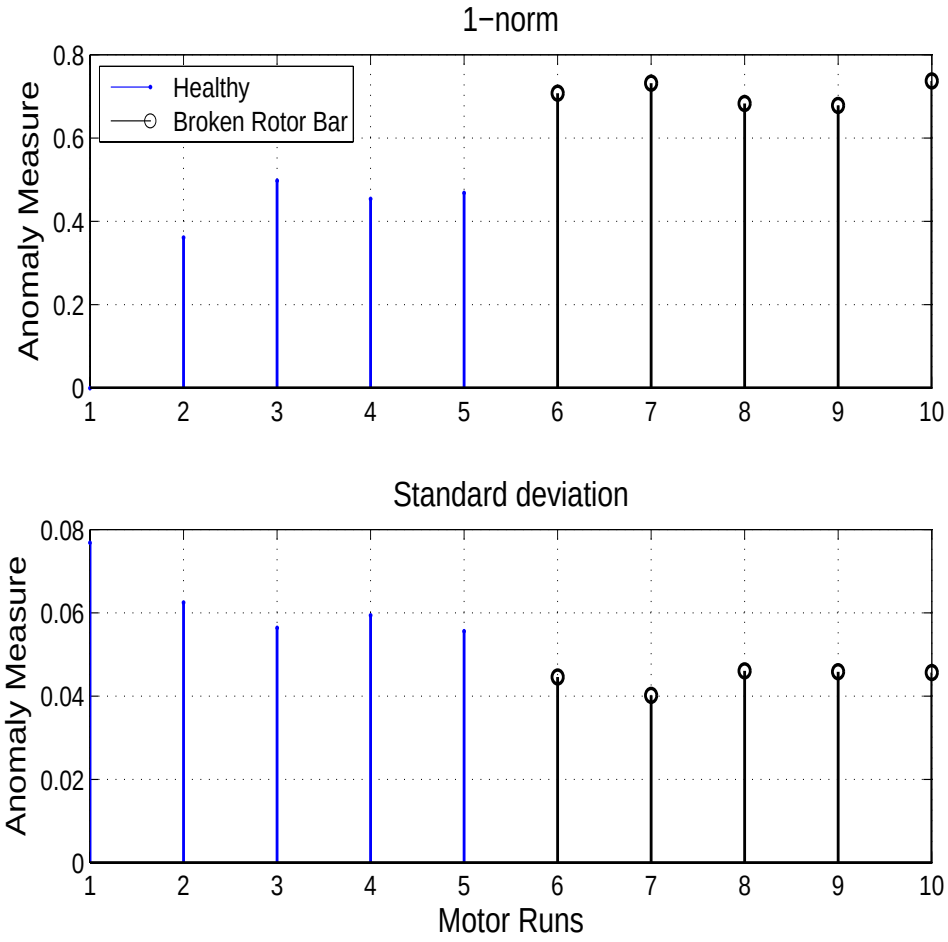

Figure 4.10. Anomaly Measure with Depth = 2, $|\Sigma| = 8$

tation. Since the CWT cannot be computed using a filter bank it is significantly slower. Moreover, there is a larger amount of data to be handled by the D-Markov machine, which need not necessarily help in improving the detection process significantly.

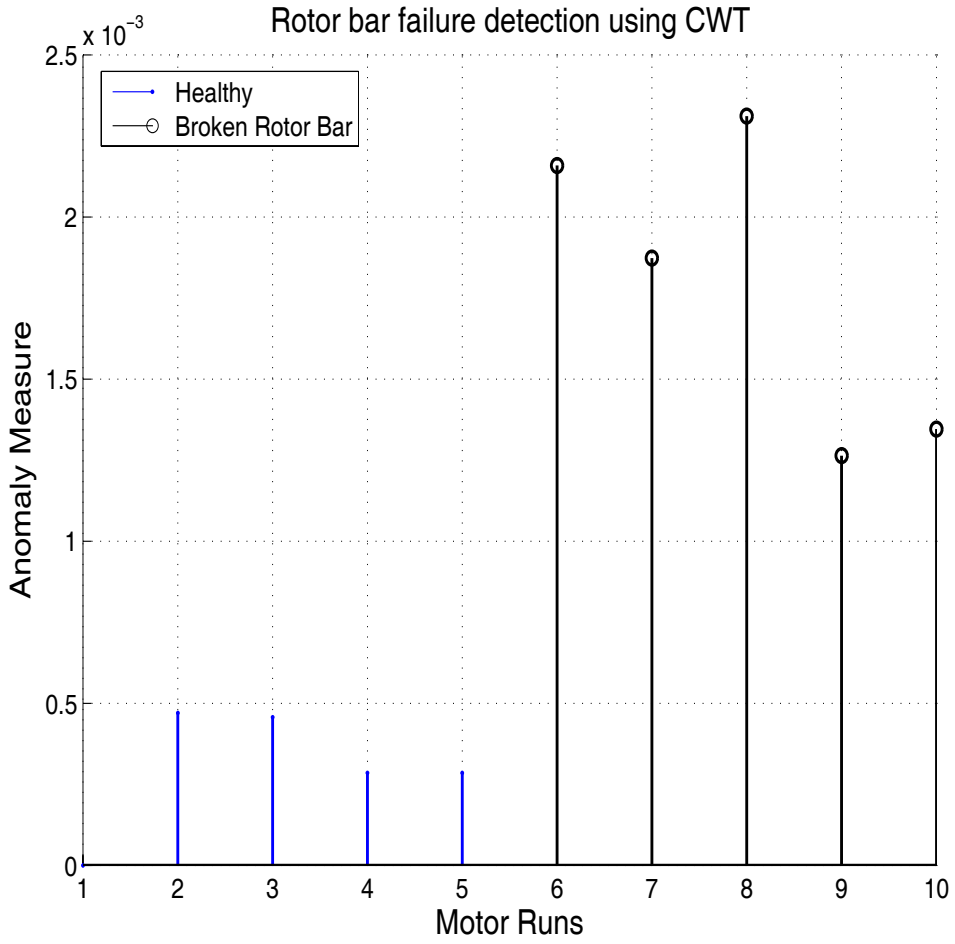

Figure 4.11. Anomaly Measure using 'gaus2' wavelet and Depth = 1, CWT

4.4 Comparison with Other Pattern Recognition Techniques

This section compares the symbolic dynamics methodology with other techniques, namely PCA (Principal Component Analysis) and neural networks. Each of these techniques were given the stacked wavelet coefficients from the line-fed induction motor. PCA is computationally simple; it essentially finds the changes in orientation of the principal eigen vector of the correlation matrix [51]. Details are given in Appendix-C. Fig(4.12) shows the performance of the PCA method. It is clear

that PCA is not an effective method for detection of these faults. This is due to the fact that the PCA method only looks at the changes in the direction of the principal eigen vector, which in this case corresponds to f the line frequency.

Neural networks are an extremely powerful tool for pattern classification. They have been used detect broken rotor bars in induction machines [52]. This paper shows the analysis performed by one such neural network. This neural network has 4 layers excluding the input and output layers. Hidden layers have 50, 40, 30, 40 neurons respectively. The output layer has 10 neurons. The input layer has just a single neuron, since the data set is the stacked wavelet coefficients and is one dimensional. The pattern matrix has 30 rows and 200 columns. The neural network is a feed-forward network, trained using error back propagation with the training error being 10^{-5}. As can be seen in Fig(4.12), the neural network is not able to identify the undamaged rotor. The reason attributed for their failure is the insufficient amount of data. The data contained after the DWT filter bank analysis is only 458 points. This amount of data is insufficient to make an accurate classification in the case of neural networks or PCA. This comparison proves that the D-Markov method possesses the ability to use limited data and make a distinction. There is no doubt that the performance all three techniques will improve when the amount of data collected is large.

The authors do not make a claim that this is the optimal neural network for this application. As with any non-linear techniques there exists no standard metric/benchmark for comparison. The purpose of this exercise is to illustrate the robustness of the symbolic wavelet analysis method.

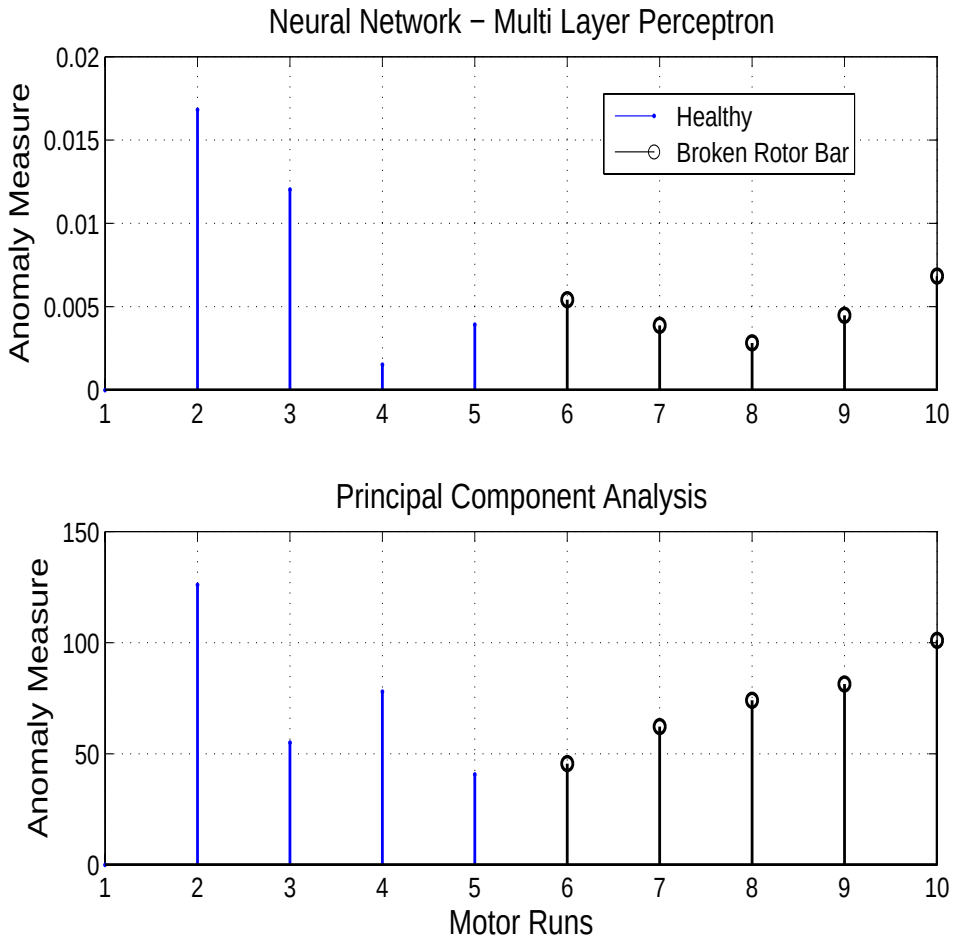

Figure 4.12. Comparison with PCA and Neural Networks

4.5 Summary

In this chapter, the detection of broken rotor bars in squirrel cage induction motors was explored. A simulation model and experimental data were utilized in demonstrating the efficacy of the model. The statistical nature of the framework is demonstrated, where the fault metric has a range instead of a deterministic value. Aspects of the framework such as depth, symbol size and the comparison between the CWT and DWT is shown. The method is also compared to other pattern comparison techniques.

In the following chapter, we shall look at the problem of bearing degradation.

We will also address issues like data sufficiency, which have not been dealt with in this chapter. A technique for sensor fusion is also developed, that allows for a more robust fault measure.

Chapter 5

Bearing Fault Detection

Bearing faults account for 40% of all failures in induction motors [56]. This is probably due to the fact that they carry the weight of the rotor. Also, improper installation like misalignments have adverse effects on them. Most of the bearings found on induction motors are of the rolling element type. There are several parts that make up the bearing, the balls (rolling elements), the inner race, the outer race and the cage, which holds the balls. Fig(5.1) shows the various components of a rolling element bearing. Each of these elements is subject to intense stress when the motor is in operation. In this chapter we shall study the detection of bearing faults, specifically we shall study outer race faults. We shall also enhance the framework developed so far, by utilizing it to merge information from different types of sensors. In particular, we shall look at merging current and vibration information. We will also look at methods that determine other statistical aspects such as data sufficiency/stationarity and consistency of framework under information fusion.

Fig(5.2) shows the enhancement made to the framework. The biggest challenge in detecting bearing degradation is not instrumentation but signal processing. Typically the fault signature due to degradation is 50 to 80 dB smaller than the dominant components, be it in the vibration spectra or the current spectra.

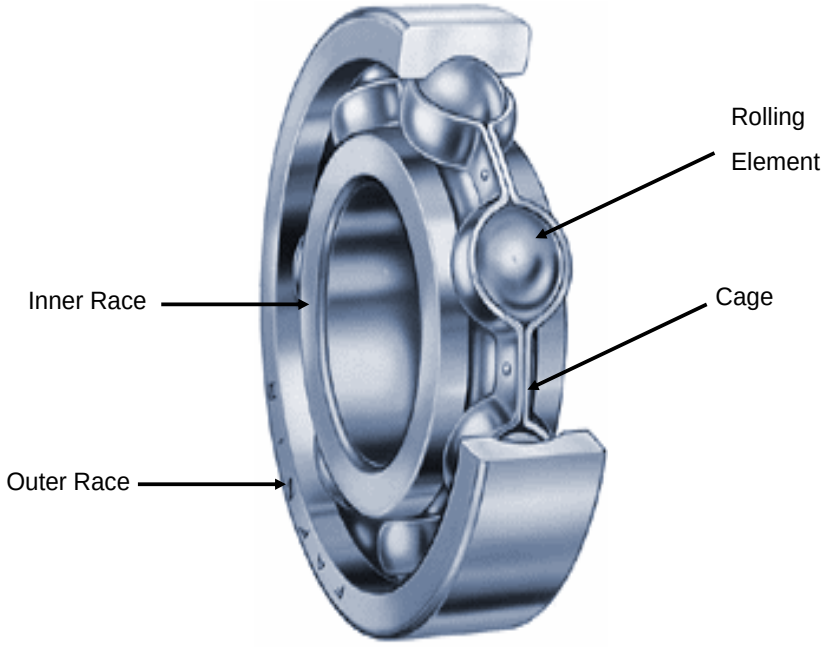

Figure 5.1. Cutaway view of a rolling element bearing

In the vibration spectra, the natural frequencies and forced frequencies of the apparatus have much larger amplitudes, while in the current spectra the pole-pass frequencies dominate [57]. In addition to these, the vibration generated by the test stand are also significant. These greatly interfere with the detection process.

In situations such as these, it is beneficial to utilize the information from all available sensors. The additions in the framework allow it to combine both the current and vibration sensors to obtain a more complete picture of the fault under study. This is achieved in a way that requires minimal additional processing. Apart from this each of these sensors can also give an independent fault measure, but combining the information makes the system less prone to false alarms. Unfortunately, combining information from sensors is a non-trivial task, especially since in this case the sensors are completely different in terms of operating characteris-

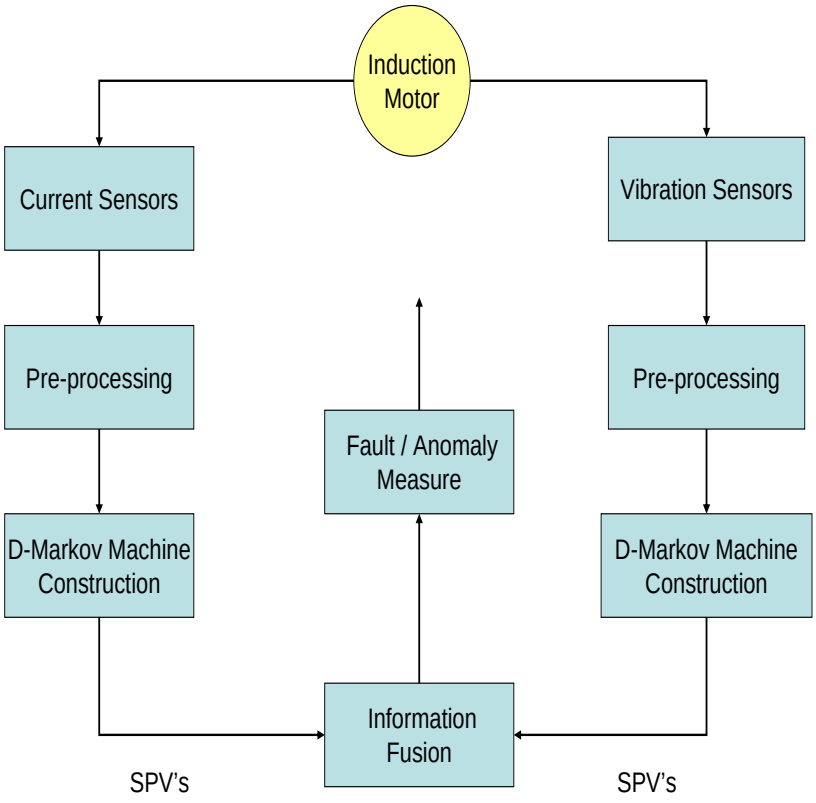

Figure 5.2. Bearing Fault Detection

tics like frequency response and accuracy. This is where the statistical approach used in this framework offers a significant advantage. Instead of merging raw information in the signal space, we perform sensor combination in a probability space. The data obtained from each sensor is processed independently using the methodology described in Chapter-3, beginning from signal acquisition all the way to generating the D-Markov machine. Subsequently, this information is merged after the stationary probability distribution is obtained from each of the sensors. For a faulted condition the merged probability distribution will be different from the healthy condition. All this is made possible due to the unique structure of the

D-Markov machine, and the fact that if we were to take a sufficiently long symbol string, the SPV's would converge to a stationary distribution. To understand the process of sensor fusion let us consider a Π or state transition matrix obtained from processing the information from one sensor. Let this matrix and the state probability vector associated with it be

$$\begin{bmatrix} 0.1 & 0.9 \\ 0.3 & 0.7 \end{bmatrix} \xrightarrow{SPV} \begin{bmatrix} 0.25 \\ 0.75 \end{bmatrix}$$

while another sensor could have a Π-matrix and SPV as shown below.

$$\begin{bmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{bmatrix} \xrightarrow{SPV} \begin{bmatrix} 0.333 \\ 0.667 \end{bmatrix}$$

When the information from both these sensors needs to be combined, we could either use their equally weighted average which would yield a new matrix and then obtain the SPV's as shown below:

$$\begin{bmatrix} 0.15 & 0.85 \\ 0.35 & 0.65 \end{bmatrix} \xrightarrow{SPV} \begin{bmatrix} 0.2917 \\ 0.7083 \end{bmatrix}$$

But since we are only interested in the SPV's we could combine them directly to obtain

$$\begin{bmatrix} 0.25 \\ 0.75 \end{bmatrix}, \begin{bmatrix} 0.333 \\ 0.667 \end{bmatrix} \rightarrow \begin{bmatrix} 0.2917 \\ 0.7083 \end{bmatrix}$$

Thus, merging information in probability space can be done both at the D-Markov machine level or after obtaining the stationary SPV's. It is only at this stage that the information from diverse sensors such as Hall-effect(current) and accelerometers(vibration) can be merged; due to the difference in the frequency spectra of the

fault signatures obtained from each of the sensors. However, after partitioning and representation of the symbol strings in a D-Markov machine, their combination is possible.

This entire process gives us three fault measures; one each from the current and the vibration, and an additional one from their combination. This, lowers the possibility of false alarm and only marginally increases the computational burden.

5.1 Bearing Fault Frequencies

To identify bearing faults, a good understanding of the geometry of the bearing is required; particularly, the characteristic vibration frequency for single point defects on each of the components of the bearing.

Bearing faults such as outer race, inner race, ball defect, and cage defects cause machine vibration at their respective characteristic frequency, f_v. The chances of the bearing failing in more than one point are extremely rare. The characteristic vibration frequencies due to bearing defects can be calculated for a given rotor speed and bearing dimensions. The characteristic frequencies for the four kinds of defects in bearings, namely the outer race, inner race, ball, and cage defects are given below.

$$f_{OD} = \frac{n}{2} f_{rm}(1 - \frac{BD}{OD} cos\phi) \tag{5.1}$$

$$f_{ID} = \frac{n}{2} f_{rm}(1 + \frac{BD}{OD} cos\phi) \tag{5.2}$$

$$f_{CD} = \frac{1}{2} f_{rm}(1 - \frac{BD}{OD} cos\phi) \tag{5.3}$$

$$f_{BD} = \frac{PD}{2BD} f_{rm}(1 - (\frac{BD}{OD})^2 cos^2\phi) \tag{5.4}$$

where f_{OD}-Outer race defect frequency, f_{ID}-Inner race defect frequency, f_{CD}-Cage

defect frequency, f_{BD}- Ball defect frequency, n-Number of balls in the bearing, f_{rm}-Speed of the rotor expressed in terms of frequency, BD-Ball diameter, OD-Outer Race diameter, PD-Pitch diameter and ϕ-Contact angle.

5.2 Experimental Procedure

Bearings usually consist of hardened steel and take a long time to fail, therefore for testing, failures must be artificially introduced. During this research the failures were introduced mechanically. Other researchers have used EDM (Electronic Discharge Machines) and shaft currents in addition to mechanical damage. In our case the motor under test was disassembled and the front bearing was pulled from the rotor shaft. It was subsequently damaged and hydraulically pressed back onto the shaft. In one of the bearings, a hole 1-mm in diameter was drilled in the outer race. In another bearing the outer-race was scratched with $1/16^{th}$ end mill to simulate a partially damaged bearing, similar to what happens due to brinelling. In all these experiments the stator, rotor and housing remained the same. The motors were tested after reassembly. Data was taken for different load conditions. This loading condition represents about 50% load. Hall-effect current sensors measure the current and a three-axis accelerometer, which is mounted on the front face plate directly above the bearing, measures vibration. Details of the experimental setup along with pictures are given in Appendix-A. Five runs each of the healthy, scratched race and 1-mm hole bearing were obtained. The first five runs were of the healthy motor, the next five were from the motor with the bearing with the scratched outer race and the last five were obtained from the motor with a 1-mm hole. The motors were loaded to 3 N-m and data was collected for 100 seconds. The radial acceleration picked up by the accelerometer is shown in Fig(5.3), which shows 6 seconds of data sampled at 1 kHz. It can be seen that in the time-domain,

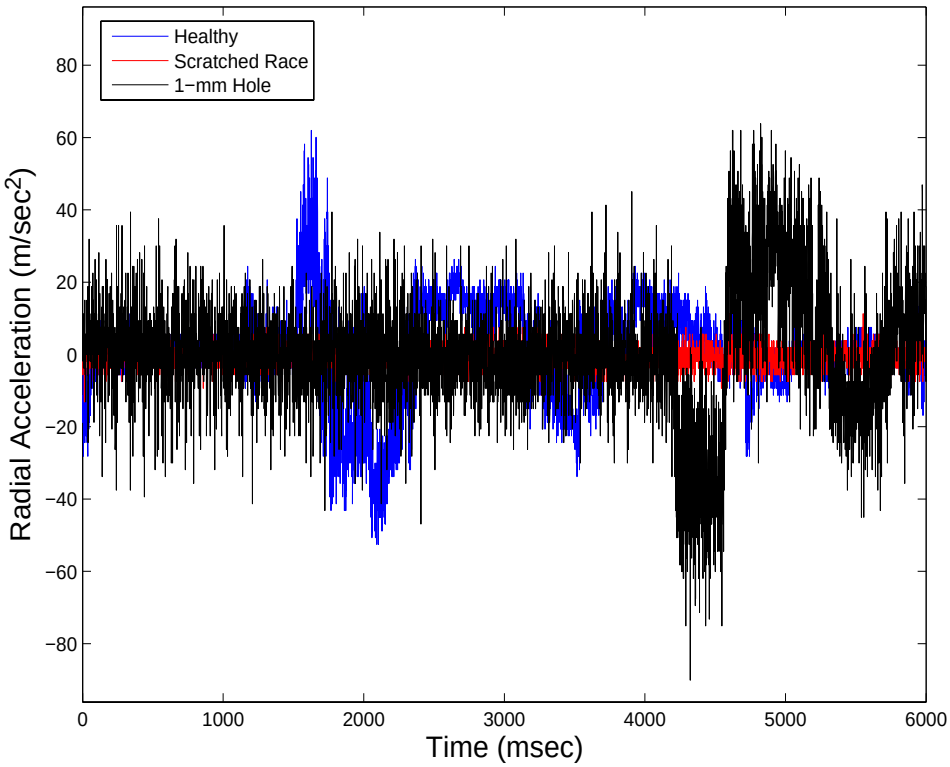

Figure 5.3. Radial Acceleration

it is extremely hard to separate the faulty bearing from the healthy one. There is a lot of noise in the vibration spectra caused by the various elements of the experimental setup, such as the test-stand. This makes the detection problem challenging, and hence, the need for a statistical fault detection framework that can accommodate the variations and yet extract the signal. The noise and the torque oscillations in the rotor, due to slight changes in the load, give rise to false peaks in the vibration frequency spectrum as well, as is shown in Fig (5.4). The acceleration is measured in all three directions, radial, axial and tangential; although only the tangential and radial components contain the fault information. Every time any of the balls in the bearing pass over the fault, they cause an acceleration in both the tangential and radial directions. This is due to the pitting of the ball, which causes a deflection in the radial direction, thus causing acceleration. The

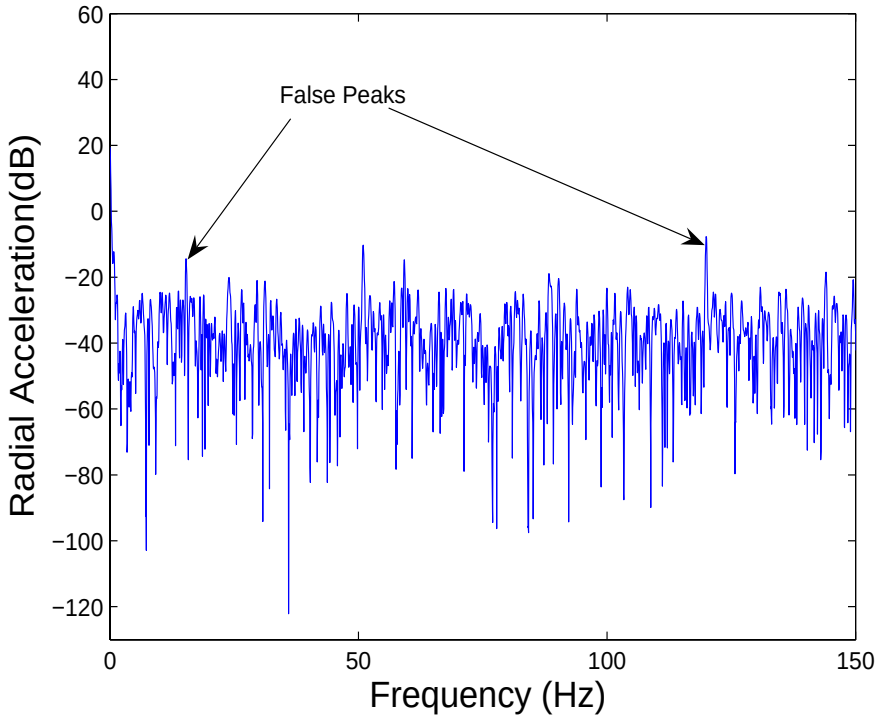

Figure 5.4. Radial Acceleration

same pitting action also causes the ball to lose its pure rolling condition, causing it to slip and produce a tangential acceleration [58]. The term acceleration in this context means a deviation from the normal operating speed; this could be a positive or a negative deviation. The mechanical vibration caused by the bearing defect results in air gap eccentricity. Oscillations in air gap length, in turn, cause variations in the flux distribution. The variations in flux distribution are coupled loosely with the supply (stator winding) current. These variations produce harmonics in the stator current. The mechanical fault frequencies are modulated by the electrical frequency of the stator current. This is the effect that shows up, with an extremely small amplitude in the stator current [59].

$$f_{CF} = |f_e \pm m f_v| \tag{5.5}$$

where f_{CF} is the fault frequency in the stator currents. f_e is the electrical frequency of the stator currents, f_v is the mechanical fault frequency derived earlier for a particular fault, and $m \in \mathbb{N}$.

The stator current helps in the early detection of bearing faults, since there is less contamination due to noise and vibrations from the test stand and other factors. However, it suffers from the drawback that it is not strongly coupled to the bearing and so it is not very reliable. But neither is the vibration measurement in our experiment, since the accelerometer is placed on the casing above the faulted bearing instead of directly on the bearing [60]. In practical applications this will be the nearest location to the bearing the accelerometer can be placed to get a correct reading. The next section discusses the results obtained from the current and vibration signals independently.

5.3 Experimental Results

The analysis follows the steps outlined in Chapter-3. For the analysis 40 seconds of steady-state operating data was used. The wavelet used was 'db4' at scales 8.0 and 8.2 for both the tangential and radial acceleration. The implementation in this case was using the CWT, since for the mechanical vibrations there exist no harmonics. These scales correspond to the fault frequencies of 89.29 and 87.11 Hz, which are in the estimated range of fault frequencies calculated from the bearing dimensions. The scales used in the analysis using the current signals are 4.15 and 4.2; these scales correspond to frequencies 172.12 and 170.07 Hz, respectively. Once the wavelet coefficients are obtained for each of the scales for every time shift they are stacked (scale-series stacking). Table(5.1) shows the current signals from the motor that has no bearing damage. If we were to compare this particular state transition matrix (Π) matrix to the ones in the previous chapters for healthy

motors, we will find that the transition probabilities differ. This is due to the pre-processing stage. Each particular fault has a different kind of pre-processing. In this particular case we used the current signal from a single line, instead of computing the Parks vector transform.

Table 5.1. Π-Matrix with Uniform Partitioning for Healthy Bearing for Stator Current Signals

	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7
q_0	0.6667	0.3333	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
q_1	0.0002	0.6960	0.2310	0.0725	0.0001	0.0001	0.0000	0.0000
q_2	0.0000	0.2538	0.3941	0.2201	0.1320	0.0000	0.0000	0.0000
q_3	0.0000	0.0555	0.0098	0.5315	0.3172	0.0827	0.0033	0.0000
q_4	0.0000	0.0031	0.0579	0.2357	0.5342	0.1130	0.0562	0.0000
q_5	0.0000	0.0000	0.0035	0.1128	0.3116	0.4641	0.0680	0.0399
q_6	0.0000	0.0000	0.0000	0.0000	0.0938	0.2048	0.4643	0.2371
q_7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0591	0.5020	0.4389

Table 5.2. Π-Matrix with Uniform Partitioning with a 1-mm hole in the Bearing for Current Signals

	q_0	q_1	q_2	q_3	q_4	q_5	q_6	q_7
q_0	0.5000	0.0000	0.5000	0.0000	0.0000	0.0000	0.0000	0.0000
q_1	0.0000	0.6677	0.2493	0.0830	0.0000	0.0000	0.0001	0.0000
q_2	0.0000	0.2443	0.3656	0.2025	0.1876	0.0000	0.0000	0.0000
q_3	0.0000	0.0547	0.0466	0.5126	0.2959	0.0767	0.0134	0.0000
q_4	0.0000	0.0028	0.0612	0.2471	0.5119	0.1270	0.0500	0.0000
q_5	0.0000	0.0000	0.0022	0.1484	0.2693	0.4526	0.0567	0.0709
q_6	0.0000	0.0000	0.0000	0.0000	0.1424	0.1839	0.4055	0.2682
q_7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0834	0.4242	0.4924

5.3.1 Problems with Vibration Signals

There is a well known phenomenon that makes the detection of bearing damage harder than most other problems. That is the effect of machine speed and loading,

particularly for machines operating under load, the detection of bearing failure is harder. This phenomenon is reported in [57] and is shown in Fig(5.5). Tables

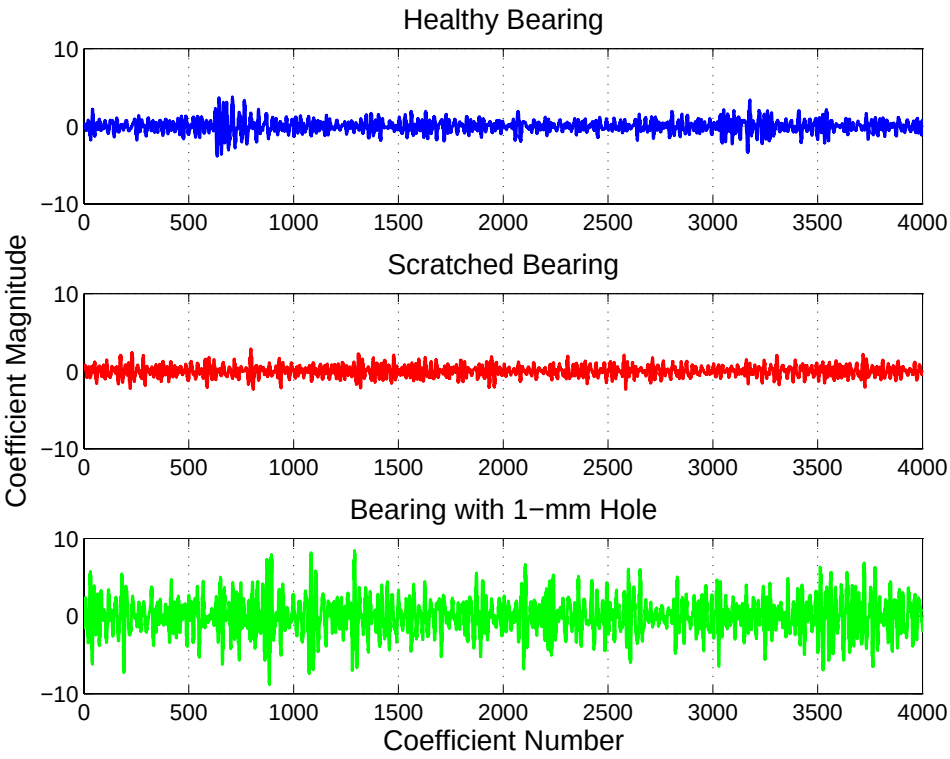

Figure 5.5. Wavelet Coefficients of Radial Acceleration at Fault Frequencies

(5.3) and (5.4) give the SPV's for all three bearings under test. Intuitively, as the fault progresses we should see states peripheral states like q_0 and q_7 see an increased probability of occurrence. Instead we observe the opposite effect in both the radial and tangential acceleration, when comparing the healthy bearing with the scratched race bearing. This effect reverses itself subsequently for a larger amount of damage, i.e. the 1-mm hole. If we were to determine the health of the machine by simply looking at the SPV's we would get the impression that the faulted condition is better than the nominal. However, the metrics (1-norm, standard deviation and Kullback distance) used to quantify the fault determine only how much the state probability vectors have changed. The radial acceleration

Table 5.3. p-Vectors Radial Acceleration

	Healthy	Scratched Race	1-mm hole
q_0	0.0003	0.0000	0.0010
q_1	0.0016	0.0000	0.0103
q_2	0.0148	0.0000	0.1150
q_3	0.6412	0.6774	0.4303
q_4	0.3320	0.3226	0.3632
q_5	0.0089	0.0000	0.0736
q_6	0.0009	0.0000	0.0060
q_7	0.0003	0.0000	0.0006

Table 5.4. p-Vectors for Tangential Acceleration with Uniform Partitioning

	Healthy	Scratched Race	1-mm hole
q_0	0.0003	0.0000	0.0002
q_1	0.0015	0.0000	0.0020
q_2	0.0110	0.0000	0.0203
q_3	0.4439	0.4459	0.4512
q_4	0.5278	0.5541	0.4971
q_5	0.0136	0.0000	0.0268
q_6	0.0017	0.0000	0.0022
q_7	0.0003	0.0000	0.0003

is usually the most accurate measure of the fault, as compared to the tangential acceleration and current sensors, which are coupled by the electromagnetic field in the induction motor. This is probably the reason that a lot of researchers have used this signal alone in their analysis [61] [62].

The tangential acceleration will also be affected, but more by damage along the surface. So instead of the depth of the damage, the tangential acceleration is affected more by the length along the surface. The current also follows the tangential acceleration to a large extent; both these signals are looking for a deviation from pure rolling. They may not be able to differentiate the depth-wise severity of the fault. They will, however, be able to determine if deterioration has occurred at all, and differentiate in terms of the surface length of the deterioration. Thus, tangential vibration and stator current signals provide a good early warning signal

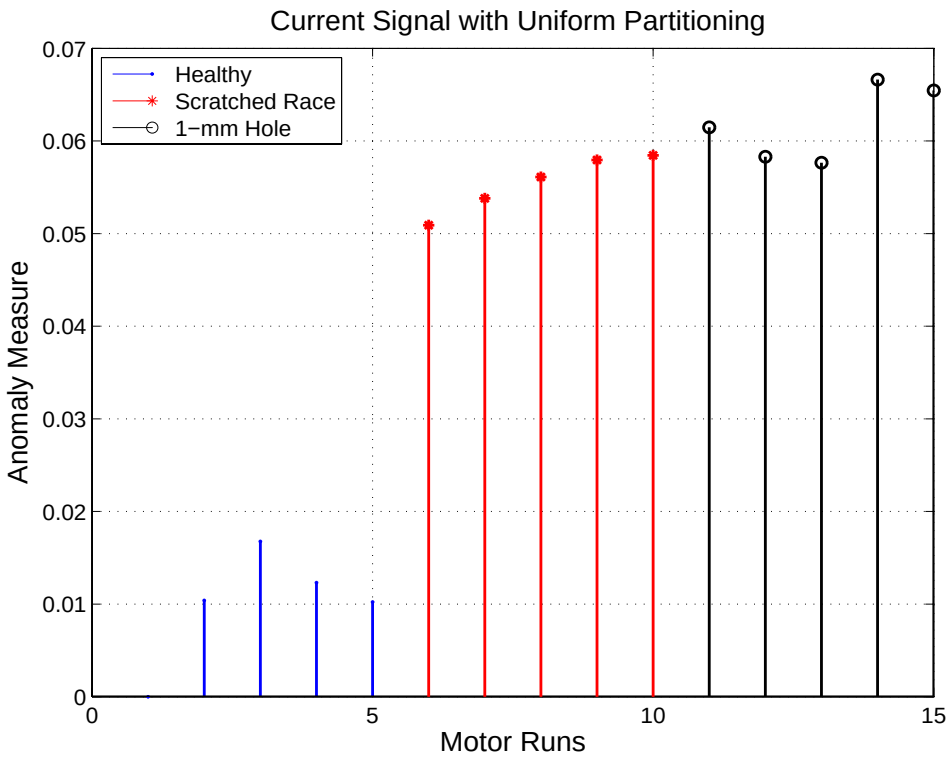

Figure 5.6. Current signal

for damage.

5.3.2 Sensor Fusion by Way of Mixed Multinomials

In this section we shall extend the fault detection framework so far to incorporate the information from different sensors.

Figs(5.6), (5.7) and (5.8) show the fault measures from the current signals, radial acceleration and tangential acceleration respectively.

As discussed in the previous section, the radial acceleration provides us with a better sense of the depth of the fault. Whereas, the tangential and current signals provide a better indication of the surface area of damage. By combining the radial acceleration along with either the tangential acceleration or the current signals, we can obtain a better fault scheme that gives us the severity of the damage both

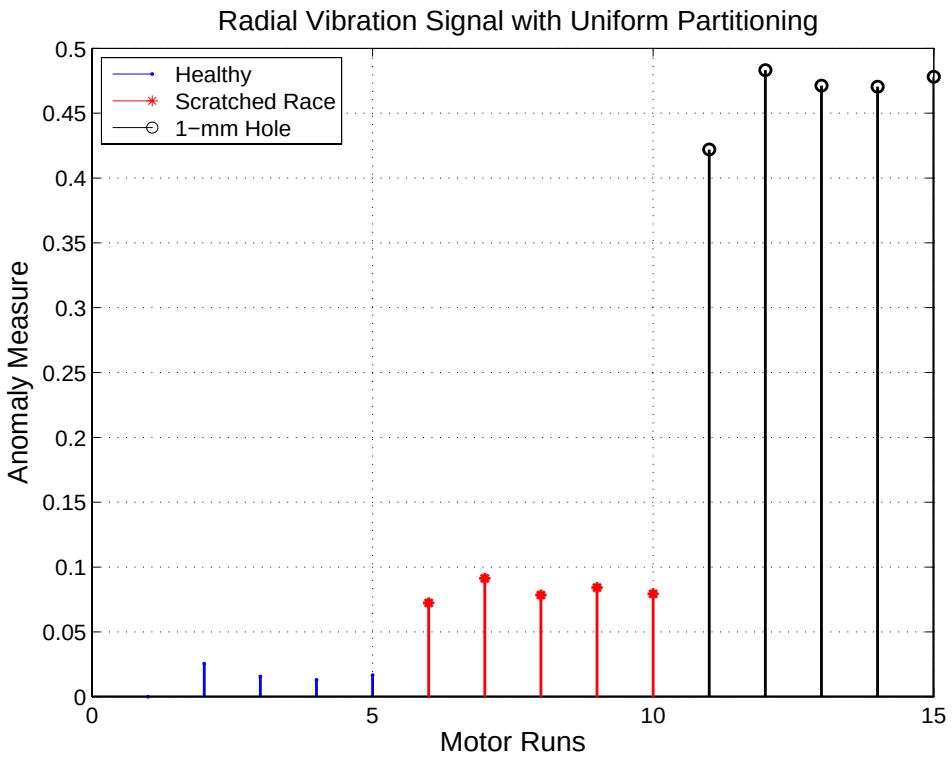

Figure 5.7. Radial Acceleration signal

depth-wise and surface area wise. Since all the signals form multinomial distributions, refer Section(2.4.1); it is possible to combine information from both of them, although the sensors used are completely different in operating characteristics, and so are the fault signatures. The process for combining two multinomial distributions is given in Section(2.4.3) and is repeated here for readability. Given two multinomial distributions $\mathbf{X} \sim M(n_1; p_1, p_2, \ldots, p_k)$ and $\mathbf{Y} \sim M(n_2; \tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_k)$, the combined distribution is defined as [45]

$$\mathbf{Z} \sim (n_1 + n_2; \frac{n_1 p_1 + n_2 \tilde{p}_1}{n_1 + n_2}, \frac{n_1 p_2 + n_2 \tilde{p}_2}{n_1 + n_2}, \ldots, \frac{n_1 p_k + n_2 \tilde{p}_k}{n_1 + n_2}) \tag{5.6}$$

This combined multinomial distribution can be used to give a more robust indication of bearing failure as is shown in Fig(5.9). As a metric the 1-norm between the nominal and the faulted condition is used for all the anomaly measure

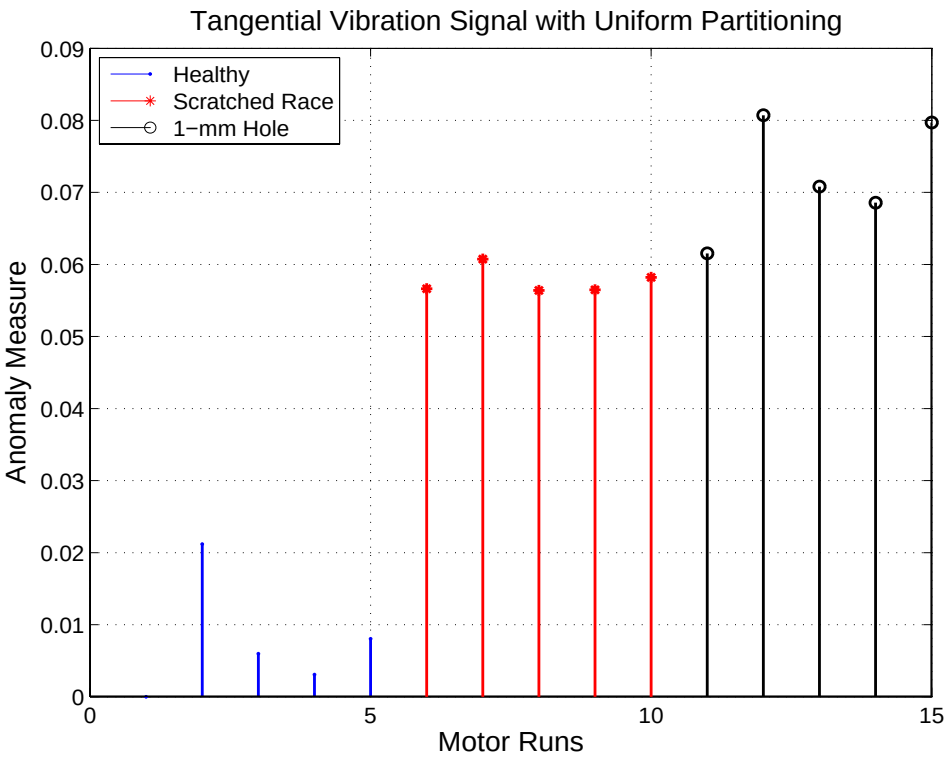

Figure 5.8. Tangential Acceleration signal

plots.

Fig(5.10) illustrates this concept. In this case, the tangential acceleration is combined using the multinomial combination with the stator current signals. It is also possible to combine signals that are very much alike, which is the case with the radial and tangential accelerations, as we shall see in the next section. The framework allows for three measures to be computed, without significant loss of information. Thus, providing three fault metrics where there would have been one.

This fusion of information can also be used to combine the information from sensors that are not reliable on their own. This concept is not explored in this thesis.

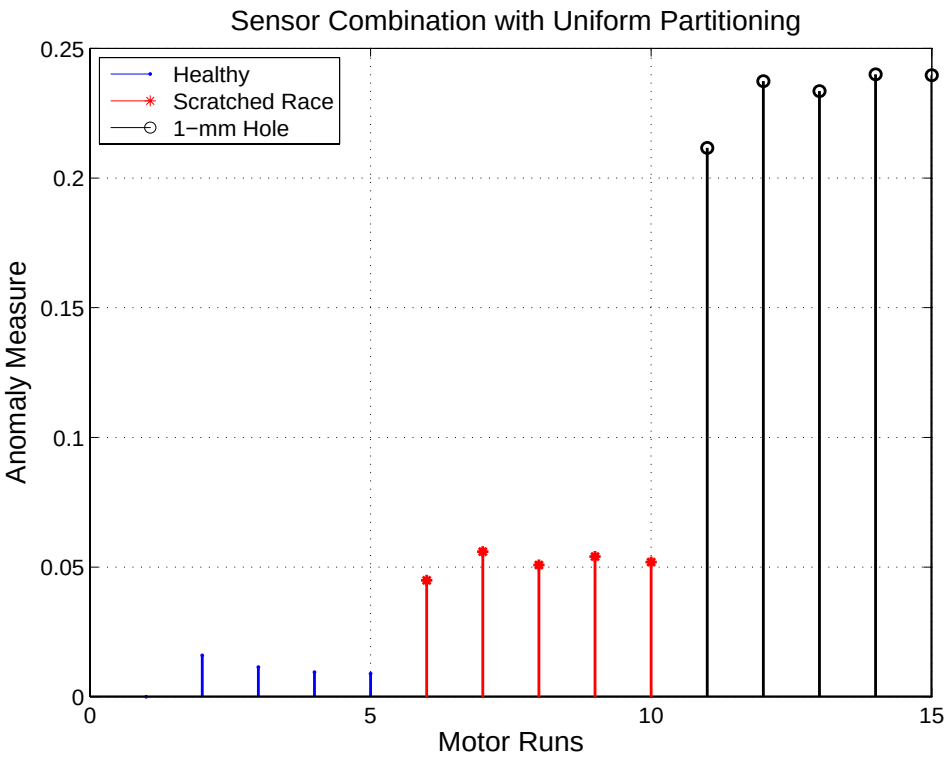

Figure 5.9. Combined Radial Acceleration and Current Signals

5.3.3 Validation using Sensor Combination using Vibration Signals

In the previous section the framework was extended to allow for sensor information fusion. It is important to experimentally validate this sensor information fusion. The results can be validated from the vibration signals. We use the tangential and radial acceleration signals for this purpose. Combining the radial and tangential acceleration signals by computing their modulus, we get the net acceleration in the tangential-radial plane. This is signal is analyzed using the CWT. This anomaly measure can then be compared with that obtained from combining the radial and tangential acceleration information using mixed multinomials. Fig(5.11) shows the sensor combination by way of mixed multinomials. Whereas, Fig(5.12) shows the anomaly measures for the case where the modulus of the radial and tangential

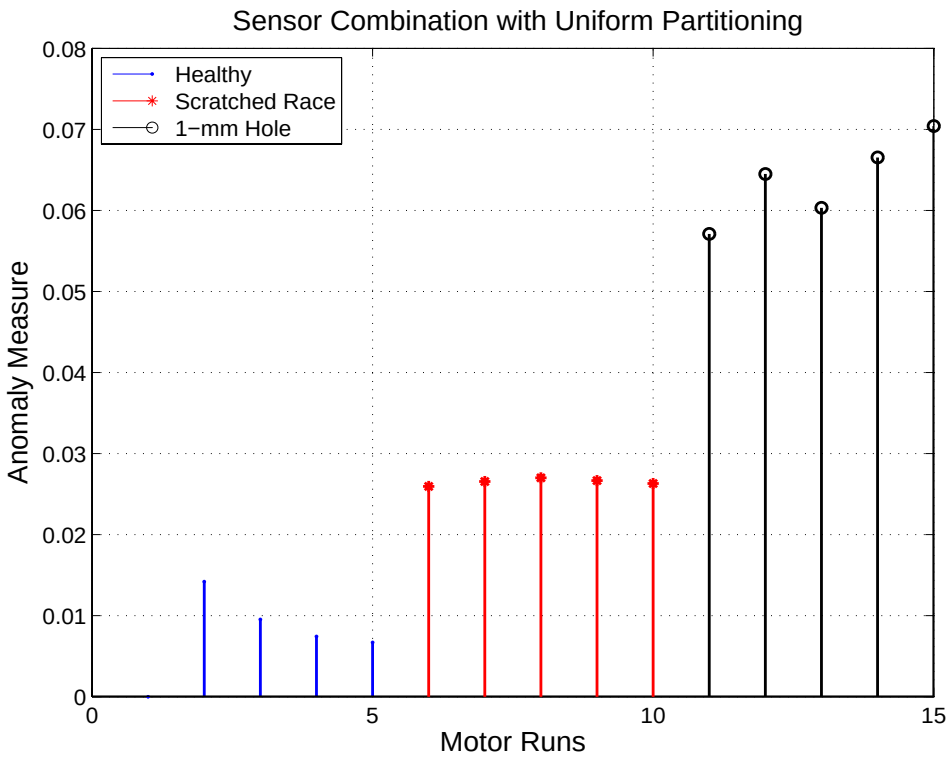

Figure 5.10. Combined Tangential Acceleration and Stator Current Signals

acceleration is analyzed. Both the anomaly measures are normalized for comparison. For the most ideal sensor combination technique, the anomaly measures from the two cases should be identical. However, in this particular framework, there is loss of information at every stage. This will cause the anomaly measures to be slightly different. However, for fault detection purposes either of the methods can be used. In fact, instead of computing the anomaly measure directly from the modulus of acceleration, we can obtain it from the individual tangential and radial acceleration, thereby reducing the computation involved.

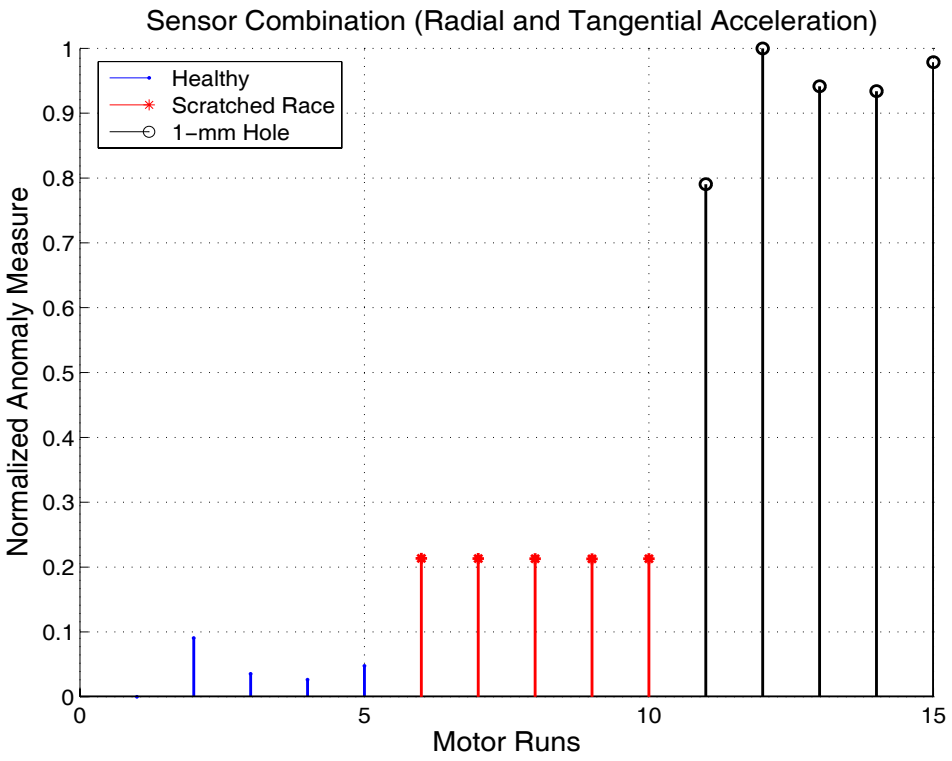

Figure 5.11. Combination Occurs after Tangential and Radial Acceleration are processed independently

5.4 Data Sufficiency or Stopping Rule

For any statistical method to be effective, there must be a minimum amount of observed data. This is to ensure that the estimate represents the underlying process as best possible. Such a test has been defined in [63]. It utilizes the SPV (**p**-vector) to arrive at the minimum length required for convergence. This rule makes use of the Perron-Frobenius theorem, given in Appendix-B, which relates the SPV's with the Π matrix, which is an irreducible matrix by construction to the unique unit eigenvalue. By imposing a condition on both the SPV's and the Π matrices for a finite data length k, we can obtain the data sufficiency condition.

$$\| \mathbf{p}(k)(I - \Pi(k) \|_{\infty} \leq \frac{1}{k} \rightarrow 0, k \rightarrow \infty \tag{5.7}$$

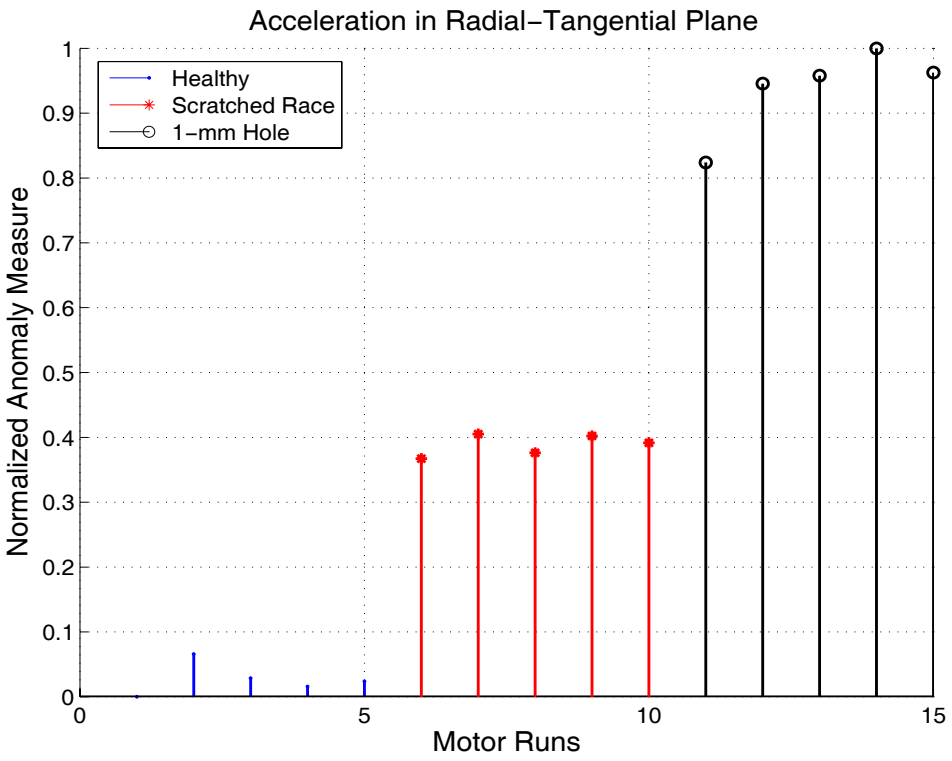

Figure 5.12. Acceleration Modulus is the Data-Stream to the Framework

which is equivalent to

$$\| \mathbf{p}(k) - \mathbf{p}(k+1) \|_\infty \leq \frac{1}{k} \rightarrow 0, k \rightarrow \infty \tag{5.8}$$

Assuming an identical distribution, the expected value of $\| \mathbf{p}(k) \|_\infty$ to be $1/n$ where n is the number of states. A lower bound on the amount of data required would be

$$k_{stop} = Integer(\frac{n}{\eta}) \tag{5.9}$$

where $0 < \eta \ll 1$ is a constant that is decided a-priori depending on the accuracy desired. For a given D-Markov machine with unit depth and $|\Sigma| = 8$, let us assume the acceptable limit on η is 0.001. This makes k_{stop} to be 8000 data points.

Table(5.5) shows the state probability vectors for the data analysis from 2000

data points which corresponds to one second of data at a 1 kHz sampling rate, for 2 scales. This is compared with state probability vectors for 80000 data points which represent 40 seconds of data. The SPV's for both the cases are given in the table.

Table 5.5. **p**-Vectors

40 seconds		1 second	
q_0	0.0001	q_0	0.0035
q_1	0.1071	q_1	0.1065
q_2	0.0735	q_2	0.0720
q_3	0.2331	q_3	0.2375
q_4	0.2941	q_4	0.2855
q_5	0.1415	q_5	0.1440
q_6	0.0988	q_6	0.0945
q_7	0.0518	q_7	0.0565

Thus, if the bound was fixed at $\eta = 0.001$; 2000 data points or one second of data is insufficient for the stationarity condition to be satisfied. Thus, depending on the bound, the Markov property decides the amount of data required for the distribution to converge. On the flip side; the amount of time taken to process $40,000$ data points as opposed to 1000 data points is at least 40 times more, assuming the entire computation is linear. The best data length for this process would be 8000 for the desired accuracy and the fixed D-Markov machine structure. This stopping rule computes the minimum amount of data length required for convergence. However, convergence is not a necessary criteria for detection. Convergence proves data sufficiency, which means any additional data will not change the SPV's or the Π matrix for the given structure.

5.5 Summary

The framework is used to detect outer race faults in bearings. The geometry of the bearing is used to calculate the regions of fault location. Problems with the vibration sensors were explored and, the metric is shown as a useful tool in tackling this problem. Information from both the current and vibration sensors, which are mounted on the exterior, is used. The framework is extended to allow for a mixed multinomial method for sensor fusion. This fusion is compared and shown to be similar to the anomaly measure derived from signals that are merged in their raw form. A data sufficiency check is devised using the stationary Markov property, this ensures accuracy of the framework.

The following chapter summarizes the research work and provides future direction.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

This thesis describes a statistical framework for detection of failures in various components of induction motors. The framework has been validated experimentally for stator voltage imbalance, rotor bar breakage, and bearing damage on a laboratory test bed comprised of a motor-dynamometer test bed that can be both operated in the line-fed as well as inverter-fed mode.

The technique comprises of both off-line and on-line analysis. The off-line analysis is used to characterize and isolate the fault, while the on-line analysis is used to provide a real-time, embedded solution for detecting both electrical and mechanical faults. The off-line analysis uses established techniques such as power spectral analysis and Parks vector transforms, while the on-line analysis is based on a symbolic dynamics approach. The symbolic dynamics method consists of fault signature extraction using wavelet analysis followed by discretization and symbolic pattern representation. The stochastic representation of the symbolic information, i.e. in the D-Markovian framework, ensures that the fault detection improves as more data is acquired.

The approach uses all the possible a-priori fault information for the diagnosis of the particular kind of fault. The extensive one-time off-line analysis, makes the on-line analysis, which is the monitoring process when the machine is in operation, computationally efficient.

Several variations in the framework modules were also explored. Beginning with the effect of different wavelets, and justification for choosing a particular kind of wavelet and also looking at different implementations of wavelets, i.e. CWT and the filter-bank implementation of the DWT. In addition to this the effect of the alphabet size and the depth of the D-Markov machine were also explored. Properties such as data sufficiency and stationarity were explored, and criteria were developed to test them. The methodology was extended further to incorporate information from multiple sensors. This was achieved by abstracting the D-Markov machine to a stationary multinomial distribution and the combination with another multinomial distribution obtained from another sensors.

6.2 Future Work

- *Formulation of alternative partitioning techniques:* These include partitioning to obtain other probability distributions. For example, the Poisson distribution, which has the advantage of being independent of the data length. Thus, dependent only on one parameter, the mean.

- *Optimization of effects of the parameters (e.g., D and $\mid \Sigma \mid$), length of the data set on the performance of anomaly detection:* Since the fault processes are modeled as distributions. Maximum-Likelihood Estimation (MLE) could be used to determine parameters such as the length and the number of symbols required. In addition, concepts such as stochastic complexity can be

used to determine the depth D of the D-Markov machine. In addition to this, concepts such as the entropy rate can also be used to determine the optimum depth. Some of these methods have been applied to simpler systems. Additionally, an optimization approach to a the problem of finding the optimal D, $\mid \Sigma \mid$ while using the minimum length of data. Thus, incorporating the appropriate amount of memory, and choosing the correct discretization interval so as not to loose accuracy.

- *Application of the framework to life cycle testing:* This framework can also be extended to include the life-cycle testing of faults such as bearing failures. Where data from a class of bearings is obtained from healthy to a failed condition. If sufficient life-cycle data is obtained then we can determine the remaining useful life of a sample for a given anomaly measure.

Appendix A

Experimentation Details

A.1 Experimental Setup

The experimental setup consists of motor-dynamometer test rig. The motors under test were 3-phase, 145 frame 2-HP Y-connected squirrel cage induction motors, with a floating neutral made by Baldor Motors Inc. The bearings on the motor were generic type 6205. The dynamometer is an air-cooled hysteresis brake, Model HD-805 N made by Magtrol Inc. The motors are coupled using a flexible Lovejoy coupling to the dynamometer. The sensor board comprised of hall effect current and voltage sensors, LA-55P and LV-25P by LEM. The voltage and current signals are sensed then passed through an anti-aliasing filter and fed into a dSPACE DS-1103 PPC card, with on board 16-bit ADC's.

The accelerometers used in this experiment are made by PCB Piezotronics, model 356A25 triaxial, general purpose, ceramic shear ICP accelerometers, with a conversion ratio of 25 mV/g ($2.6mV/(m/s^2)$), and bandwidth of 1 to 10000 Hz. The measurement range is ±200 g pk ($\pm 1960 m/s^2$ pk) with a resolution of 0.0002 g rms ($0.002 m/s^2$ rms). This gives the accelerometers a range of 1-5000 Hz, with an accuracy of 10%. The signals obtained from the accelerometers are signal

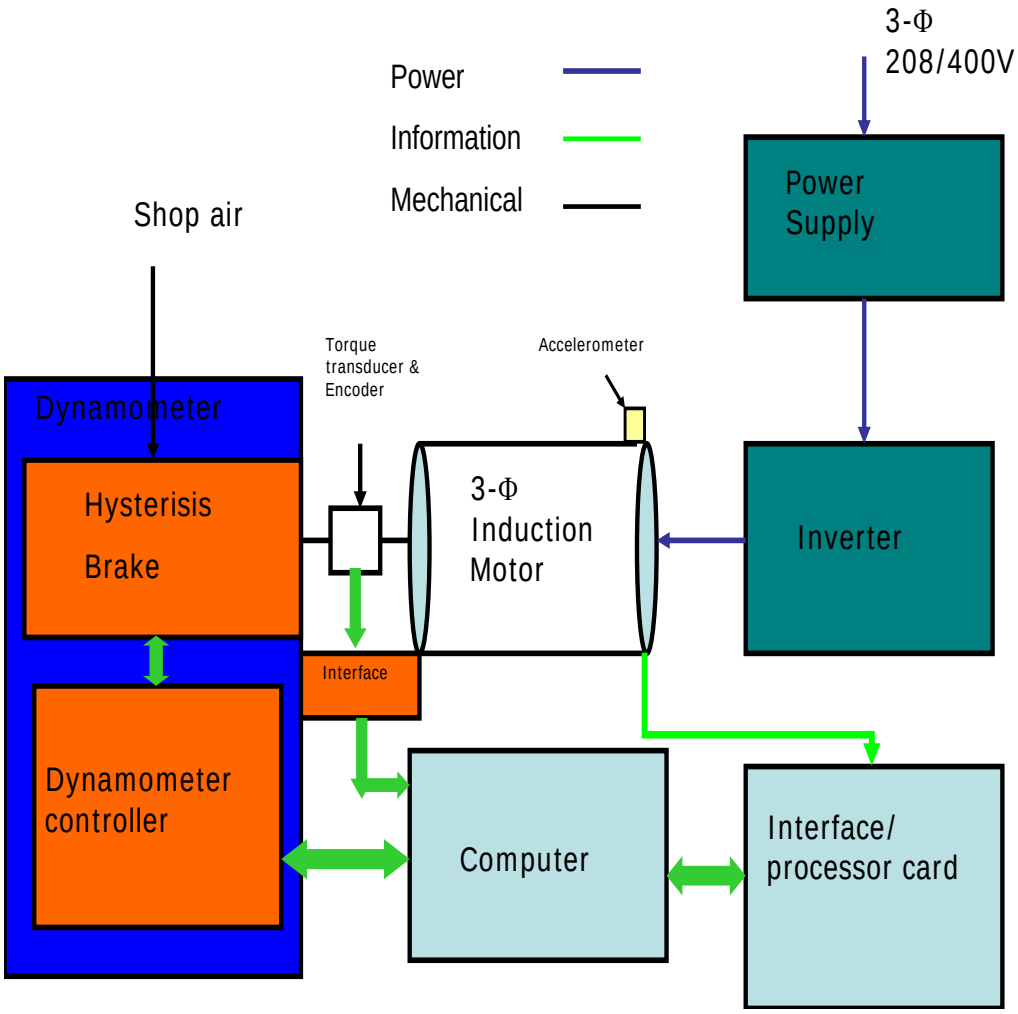

Figure A.1. Test Apparatus

conditioned using a servo amplifier with unity gain (Model 841 A).

A commercial grade UVC inverter made by TB Woods is used to drive the motor. The inverter can be used in both open loop and closed loop speed control modes. In the experiments discussed in this paper, it is operated in open loop V/Hz mode.

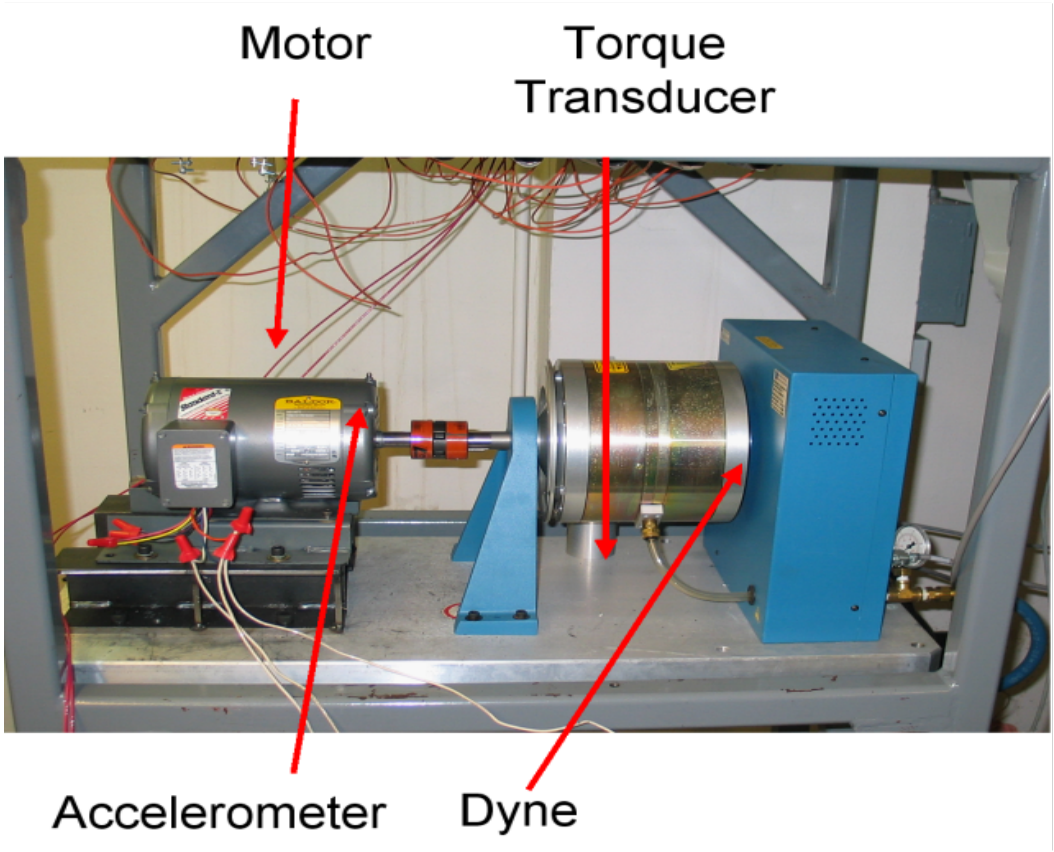

Figure A.2. Motor Test Bed

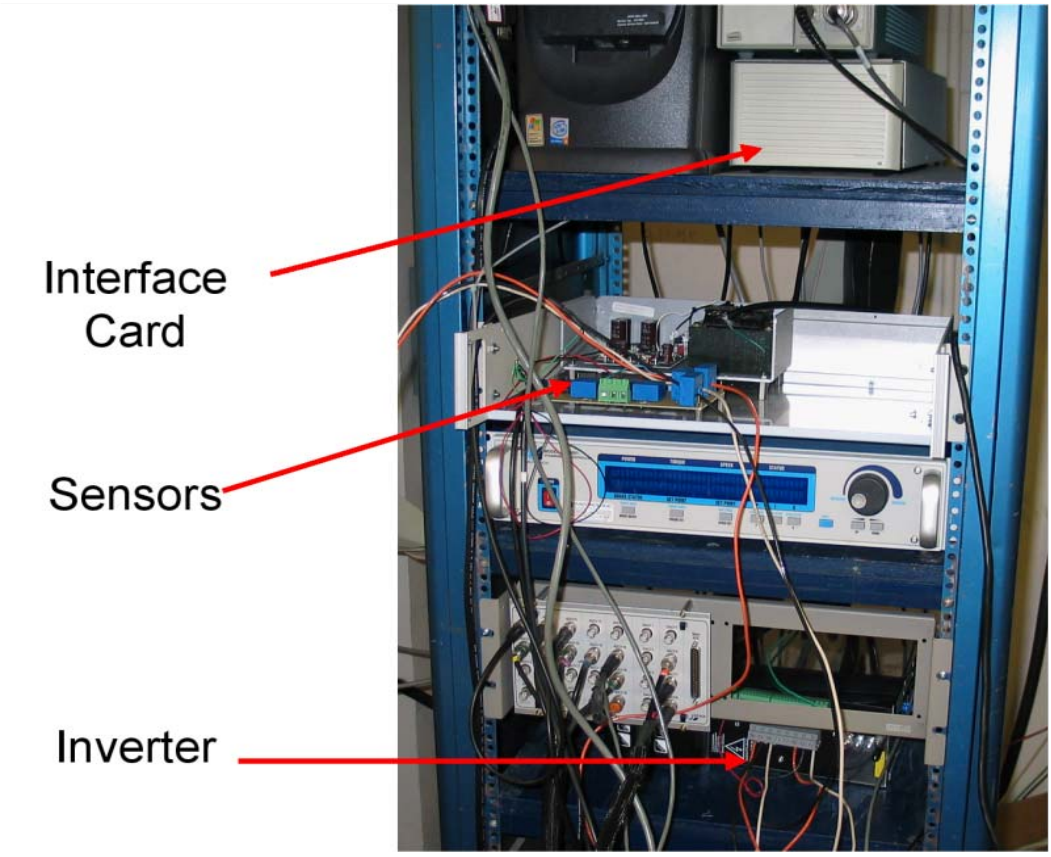

Figure A.3. Sensor Board

Figure A.4. Damaged Rotor

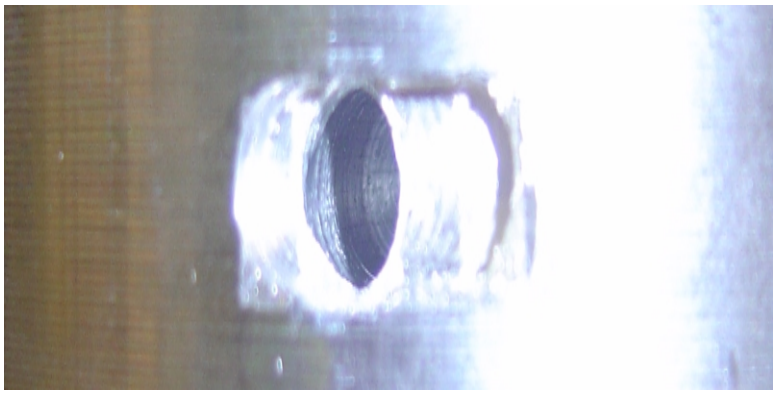

Figure A.5. Zoomed-in View of the Broken Rotor Bar

Figure A.6. Bearing with 1-mm Hole

Figure A.7. Stator Windings

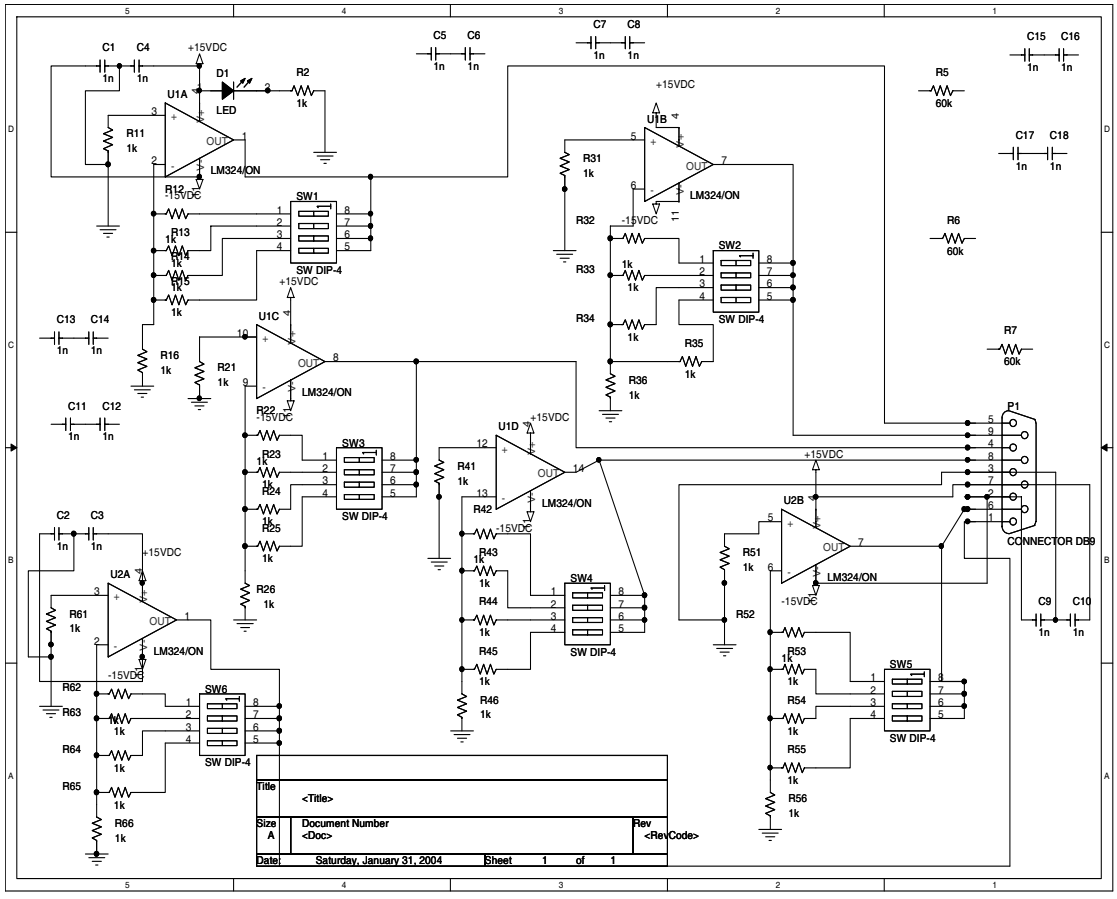

Figure A.8. Voltage and Current Sensor Board

Appendix B

Additional Theorems

Definition 1. *A stochastic symbolic stationary process* $\boldsymbol{S} = ...S_{-2}S_{-1}S_0S_1S_2...$ *is called* D^{th} *order Markov process if the probability of the next symbol depends only on the previous D symbols, i.e. the following condition holds:*

$$P(S_i/S_{i-1}S_{i-2}.....S_{i-D}S_{i-D-1}....) = P(S_i/S_{i-1}S_{i-2}.....S_{i-D}) \tag{B.1}$$

Theorem 1. *Perron-Frobenius: Let A be a nonnegative matrix. Denote its spectrum by* $\sigma(A)$*. Then the spectral radius* $\rho(A)$ *is an eigenvalue, that is,* $\rho(A) \in \sigma(A)$*, and is associated to a nonnegative eigenvector.*

If, in addition, A is an irreducible matrix, then $|\rho(A)| \geq |\lambda|$*, for all* $\lambda \in \sigma(A)$*,* $\lambda \neq \rho(A)$*, and* $\rho(A)$ *is a simple eigenvalue associated to a positive eigenvector.*

If, in addition, A is a primitive matrix, then $\rho(A) > |\lambda|$ *for all* $\lambda \in \sigma(A)$*,* $\lambda \neq \rho(A)$*.*

Appendix C

Other Pattern Recognition Techniques used for Fault Detection

This appendix briefly describes the three pattern recognition techniques which use time series data as inputs, for comparison with the symbolic-time-series-based anomaly detection method:

- Principal Component Analysis (PCA)
- Multilayer Perceptron Neural Network (MLPNN)
- Radial Basis Function Neural Network (RBFNN)

C.1 Principal Component Analysis (PCA) for anomaly detection

Feature extraction methods in statistical pattern recognition determine an appropriate subspace of dimensionality $q \in \mathbb{N}$, where $\mathbb{N}$ is the set of positive integers, using either linear or nonlinear methods in the original feature space of dimensionality n ($q \leq n$). The best known linear feature extractor relies on the Principal

Component Analysis (PCA) [26]. The eigenvectors of the $(n \times n)$ (positive semi-definite) covariance matrix of the time series data, corresponding to the q largest eigenvalues, form the n-dimensional patterns. The linear transformation is defined as

$$Y = HX \tag{C.1}$$

where X is the given $(n \times d)$ pattern matrix, made of n row vectors; H is the $q \times n$ linear transformation matrix whose rows represent q feature vectors of dimension n; and Y is the derived $d \times q$ pattern matrix. Since the PCA method uses the most expressive features (e.g., eigenvectors with the largest eigenvalues), it effectively approximates the data by a linear subspace using the mean squared error criterion.

To detect growth in anomaly from time series data, principal component analysis is performed for dimensionality reduction. If the time response of an appropriate process variable $y(t)$ is sampled to generate a time series sequence y_k, then data samples of large enough length ($\ell = d\ n$) can be used to capture the dynamical characteristics of the observed process. The length ℓ of time series data is partitioned into d subsections, each being of length $n = \frac{\ell}{d}$, where $d > n$. The resulting $(d \times n)$ data matrix is processed to generate the $(n \times n)$ covariance matrix that is positive-definite or positive-semidefinite real-symmetric. The next step is to compute the orthonormal eigenvectors $\mathbf{v}^1, \mathbf{v}^2, ..., \mathbf{v}^n$ and the corresponding eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ that are arranged in decreasing orders of magnitude. The eigenvectors associated with the first (i.e., largest) q eigenvalues are chosen as the feature vectors such that

$$\frac{\sum_{i=q+1}^{n} \lambda_i}{\sum_{i=1}^{n} \lambda_i} < \eta \tag{C.2}$$

where the threshold $\eta \ll 1$ is a positive real close to 0. The resulting pattern is

the matrix, consisting of the feature vectors as columns,

$$\widetilde{M} = \begin{pmatrix} \sqrt{\lambda_1}\mathbf{v}_1 & \dots & \sqrt{\lambda_q}\mathbf{v}_q \end{pmatrix} \tag{C.3}$$

The above steps are executed for time series data under the nominal (stationary) condition to obtain $\widetilde{M}_0$. The metric below uses any of the metrics (usually 1-norm) to calculate the anomaly measure.

$$\hat{\mathcal{M}}_k \equiv d\left(\widetilde{M}_k, \widetilde{M}_0\right)$$

C.2 Multi-Layer Perceptron Neural Network (MLPNN) for anomaly detection

The multilayer perceptron neural network (MLPNN) is the most commonly used family of feed-forward neural networks for pattern classification tasks [26]. The MLPNN is a collection of connected processing elements, called nodes or neurons [64] [65]. Its structure is fixed by choosing the number of layers as well as the (possibly different) number of neurons in each layer. The MLPNN is trained based on the information contained in a given set of inputs and target outputs. The training phase includes modelling of the input-output system architecture and identification of the synapsis weights. A set of inputs is passed forward through the network yielding trial outputs which are then compared to the target outputs to obtain the error (i.e., the deviation of the trial output from the target output). The network parameters (i.e., synapsis weights and biases) are adjusted until the error is within specified limits. If the specified bound is exceeded, the error is passed backwards through the net and the training algorithm adjusts the synapsis weights.

The back-propagation algorithm has been used in this dissertation. The simplest implementation of back-propagation learning updates the network weights and biases in the direction in which the performance function decreases most rapidly. The mean square error criterion is adopted in the recursive algorithm to update the weight vectors $\{w_k\}$ as follows:

$$\mathbf{w}_{n+\mathbf{1}} = \mathbf{w}_n - \alpha_n \mathbf{g}_n \tag{C.4}$$

where $\mathbf{g}_n$ is the gradient and α_n is the learning rate.

Different layers in MLP neural networks may contain different numbers of neurons. Time series signals enter into the input layer nodes, progress forward through the hidden layers, and finally emerge from the output layer. Each node i at a given layer k receives a signal from all nodes j in its preceding layer $(k-1)$ through a synapsis of weight w_{ij}^k and the process is carried onto the nodes in the following layer $(k+1)$. The weighted sum of signals x_j^{k-1} from all nodes j of the layer $(k-1)$ together with a bias w_{i0}^k produces the excitation z_i^k that, in turn, is passed through a nonlinear activation function f to generate the output x_i^k from the node i at the layer k. This is mathematically expressed as

$$z_i^k = \sum_j w_{ij}^k x_j^{k-1} + w_{i0}^k \tag{C.5}$$

$$x_i^k = f\left(z_i^k\right) \tag{C.6}$$

Various choices for the activation function f are possible; the hyperbolic tangent function $f(x) = tanh(x)$ has been adopted in this thesis.

For anomaly detection, the MLPNN is trained by setting a set of N input vectors, each of dimension ℓ, and a specified target output vector τ of dimension q. This implies that the input layer has ℓ neurons and the output layer has q neurons.

If the time series data are obtained from an ergodic process, then a data set of length $N.\ell$ can be segmented into N vectors of length ℓ to construct the input and target pattern matrices, $\mathcal{P}$. The input pattern matrix $\mathcal{P} \in \mathbb{R}^{\ell \times N}$ is obtained from the N input vectors as

$$\mathcal{P} \equiv \left[\mathbf{p}^1 \ \mathbf{p}^2 \ \cdots \ \mathbf{p}^N\right] \tag{C.7}$$

where $\mathbf{p}^k \equiv \left[y_{(k-1)\ell+1} \ y_{(k-1)\ell+2} \ \cdots \ y_{k\ell}\right]^T$ and each y_k is a sample from the ensemble of the time series data. The corresponding output matrix $\mathcal{O}$ is the output of the trained MLPNN under the input pattern $\mathcal{P}$.

$$\mathcal{O} \equiv \left[\mathbf{o}^1 \ \mathbf{o}^2 \ \cdots \ \mathbf{o}^N\right] \tag{C.8}$$

where $\mathbf{o}^i \in \mathbb{R}^q$ is the output of the trained MLPNN under the input $\mathbf{p}^k \in \mathbb{R}^\ell$. The performance vector $\mathbf{u} \in \mathbb{R}^q$ is obtained as the average of the N outputs.

$$\mathbf{u} \equiv \frac{1}{N} \sum_{k=1}^{N} \mathbf{o}^k \tag{C.9}$$

The time series data under the nominal condition generates the input pattern matrix $\mathcal{P}_0$ that, in turn, is used to train the MLPNN with respect to a target output vector τ. The resulting output of the trained MLPNN with $\mathcal{P}_0$ as the input is $\mathcal{O}_0$ and the performance vector is $\mathbf{u}_0$.

$$\hat{\mathcal{M}}_k \equiv d\left(\mathbf{u}_k, \mathbf{u}_0\right)$$

The equation above gives a measure of the change in the performance vector, which is an indication of fault.

C.3 Radial Basis Function Neural Network (RBFNN) for anomaly detection

The radial basis function neural network (RBFNN) is a commonly used tool for pattern identification [64], where the activation of a hidden unit is determined by the distance between the input vector and the prototype vector; the RBFNN is essentially a nearest neighbor type of classifier. A radial basis function has the following structure:

$$f(y,\alpha) = exp\left(-\frac{\sum_k |y_k - \mu|^\alpha}{N\ \theta_\alpha}\right) \tag{C.10}$$

where the exponent parameter $\alpha \in (0,\infty)$; and μ and θ_α are the center and α^{th} central moment of the data set, respectively. For $\alpha = 2$, $f(\bullet)$ becomes Gaussian, which is the typical radial basis function used in the neural network literature. To perform anomaly detection, the first task is to obtain the sampled time series data when the dynamical system is in the nominal condition and then the mean μ and the central moment θ_α are calculated as

$$\mu = \frac{1}{N}\sum_{k=1}^{N} y_k; \qquad \text{and} \qquad \theta_\alpha = \frac{1}{N}\sum_{k=1}^{N} |y_k - \mu|^\alpha \tag{C.11}$$

The distance between any vector y and the center μ is obtained as $d(y,\mu) \equiv (\sum_n |y(n) - \mu|^\alpha)^{\frac{1}{\alpha}}$. Following Eq. (C.10), the radial basis function at the nominal condition is: $f_0 = f(y)$. Under all conditions including anomalous ones, the parameters μ and θ are kept fixed.

$$\hat{\mathcal{M}}_k = d(f_0, f_k)$$

where the $d(\bullet,\bullet)$ is an appropriately defined distance function.

Appendix D

Magnetic Equivalent Circuit of Induction Machine

The magnetic equivalent circuit of the induction machine is shown Fig(D.1).

Algebraic equation is developed based on node type equation. Scalar magnetic potential is assigned for every node.

The algebraic equation can be established based on Gauss's law that the net flux flowing into every node must be equal to zero as expressed.

$$A_{11}u_1 + A_{22}u_2 + A_{22}u_3 = 0 \quad \text{(D.1)}$$

$$A_{32}u_2 + A_{33}u_3 + \phi rt = 0 \quad \text{(D.2)}$$

$$u_2 - u_1 + F_{st} + R_{st}\phi_{st} = 0 \quad \text{(D.3)}$$

$$A_{44}u_4 - \phi_{rt} = 0 \quad \text{(D.4)}$$

In addition to basic algebraic equations given above. Permeance matrices are defined next. It is assumed that the induction machine k stator teeth and l rotor teeth.

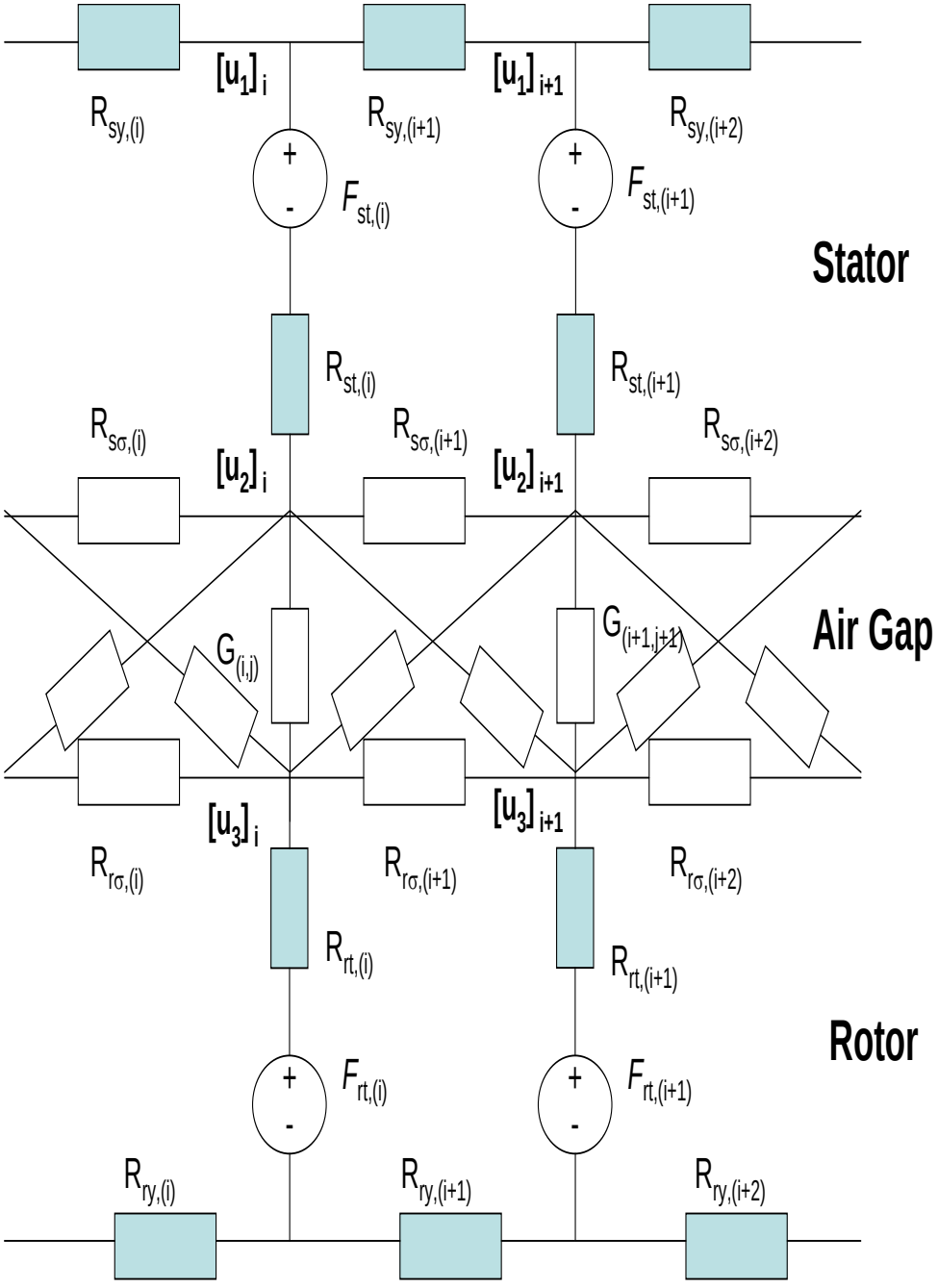

Figure D.1. Magnetic equivalent circuit of the induction machine

Table D.1. Induction machine data

Specification	Rate
Power	10 Hp
Speed	1500 rpm
Efficiency	0.9
Power Factor	0.85
Input Voltage	381 V
Frequency	50 Hz
Phases	3
Poles	4
Stator/Rotor slot	36/28

$$A_{11} = \begin{bmatrix} (G_{sy,1}+G_{sy,2}) & -G_{sy,2} & 0 & \dots & 0 & 0 & -G_{sy,1} \\ -G_{sy,1} & (G_{sy,1}+G_{sy,2}) & -G_{sy,3} & \dots & 0 & 0 & 0 \\ & & & \dots & & & \\ -G_{sy,1} & 0 & 0 & \dots & 0 & -G_{sy,k} & (G_{sy,1}+G_{sy,2}) \end{bmatrix} \tag{D.5}$$

$$A_{22} = \begin{bmatrix} (2G_{s\sigma}+\sum_{j=1}^{l} G_{1,j}) & -G_{s\sigma} & 0 & \dots & 0 & 0 & -G_{s\sigma} \\ -G_{s\sigma} & (2G_{s\sigma}+\sum_{j=1}^{l} G_{2,j}) & -G_{s\sigma} & \dots & 0 & 0 & 0 \\ & & & \dots & & & \\ -G_{s\sigma} & 0 & 0 & \dots & 0 & -G_{s\sigma} & (2G_{s\sigma}+\sum_{j=1}^{l} G_{k,j}) \end{bmatrix} \tag{D.6}$$

$$A_{23} = \begin{bmatrix} G_{1,1} & G_{1,2} & \dots & G_{1,l} \\ G_{2,1} & G_{2,2} & \dots & G_{2,l} \\ & & \dots & \\ G_{k,1} & G_{k,2} & \dots & G_{k,l} \end{bmatrix} \tag{D.7}$$

$$A_{32} = A_{23}^{T} \tag{D.8}$$

$$A_{33} = \begin{bmatrix} (G_{r\sigma,1} + G_{r\sigma,2} + \sum_{j=1}^{k} G_{j,1}) & -G_{r\sigma,2} & 0 & \dots & 0 & 0 & -G_{r\sigma,l} \\ -G_{r\sigma,2} & (G_{r\sigma,2} + G_{r\sigma,3} + \sum_{j=1}^{k} Gj,1) & -G_{r\sigma,3} & \dots & 0 & 0 & 0 \\ & & & \dots & & & \\ -G_{r\sigma,l} & 0 & 0 & \dots & 0 & -G_{r\sigma,l-1} & (G_{r\sigma,k-1} + G_{r\sigma,k} + \sum_{j=1}^{k} G_{j,l}) \end{bmatrix} \tag{D.9}$$

Bibliography

[1] A. H. Bonnett and G. C. Soukup, *Cause and Analysis of Stator and Rotor Failures in Three-Phase Squirrel-Cage Induction Motors.* IEEE Transactions on Industry Applications, vol. 28, No. 4, July/August 1992.

[2] P. Fernandes, *Contact fatigue in rolling-element bearings.* Engineering Failure Analysis,, v 4, n 2, Jun, 1997, p 155-160.

[3] M. Cash, T. Habetler, and G. Kliman, *Insulation failure prediction in AC machines using line-neutral voltages.* Industry Applications, IEEE Transactions on , Volume: 34 , Issue: 6, Nov.-Dec. 1998.

[4] M. Eltabach, A. Charara, and I. Zein, *A comparison of external and internal methods of signal spectral analysis for broken rotor bars detection in induction motors.* Industrial Electronics, IEEE Transactions on , Volume: 51 , Issue: 1, Feb. 2004.

[5] M. S. N. Said, M. E. H. Benbouzid, and A. Benchaib, *Detection of Broken Bars in Induction Motors Using an Extended Kalman Filter for Rotor Resistance Sensorless Estimation.* IEEE Transactions On Energy Conversion, VOL. 15, NO. 1,, March 2000.

[6] J. M. Spanjaard, P. J. Cherman, L. B. Chite, and S. Lau, *Periodic autoregressive time-frequency analysis for monitoring of rotating machinery with variable period.* Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, 1996, p 465-468.

[7] P. Marino, V. Mungiguerra, F. Russo, and F. Vasca, *Parameter and state estimation for induction motors via interlaced least squares algorithm and Kalman filter.* Power Electronics Specialists Conference, 1996. PESC '96 Record., 27th Annual IEEE , Volume: 2 , 23-27, June 1996.

[8] T.-S. Park, S.-H. Kim, N.-J. Kim, J.-Y. Yoo, and G.-T. Park, *Speed-sensorless vector control of an induction motor using recursive least square algorithm.*

Electric Machines and Drives Conference Record, 1997, IEEE International , 18-21, May 1997.

[9] M. Cirrincione, M. Pucci, G. Cirrincione, and G.-A. Capolino, *A new experimental application of least-squares techniques for the estimation of the induction motor parameters.* Industry Applications, IEEE Transactions on , Volume: 39 , Issue: 5, Sept.-Oct. 2003.

[10] D. Baillie and J. Mathew, *Comparison of autoregressive modeling techniques for fault diagnosis of rolling element bearings.* Mechanical Systems and Signal Processing, v 10, n 1, Jan, 1996, p 1-17.

[11] J. L. Kohler, J. Sottile, and F. C. Trutt, *Condition Monitoring of Stator Windings in Induction Motors: Part I- Experimental Investigation of the Negative-Sequence Impedance Detector.* IEEE Transactions on Industry Applications, vol. 38, No. 5, September/October 2002.

[12] F. C. Trutt, J. Sottile, and J. L. Kohler, *Online Condition Monitoring of Induction Motors.* IEEE Transactions on Industry Applications, Volume: 38 , No. 6, Nov-Dec 2002.

[13] J. Sottile, F. C. Trutt, and J. L. Kohler, *Condition Monitoring of Stator Windings in Induction Motors: Part II- Experimental Investigation of Voltage Mismatch Detector.* IEEE Transactions on Industry Applications, vol. 38, No. 5, September/October 2002.

[14] J. Kueck, J.D.and Criscoe and N. Burstein, *Assessment of valve actuator motor rotor degradation by Fourier analysis of current waveform.* Energy Conversion, IEEE Transactions on , Volume: 7 , Issue: 3, Sept 1992.

[15] S.-H. Lee, S. Kim, J. M. Kim, and M. H. Lee, *Fourier and Wavelet transformations for the fault detection of induction motor with stator current.* IECON Proceedings (Industrial Electronics Conference), v 1, IECON 2004 - 30th Annual Conference of IEEE Industrial Electronics Society, 2004, p 383-388.

[16] L. Eren and M. J. Devaney, *Bearing Damage Detection via Wavelet Packet Decomposition of the Stator Current.* IEEE Transactions on Instrumentation and Measurement, v 53, n 2, April, 2004, p 431-436.

[17] S.-H. Kim, T.-S. Park, J.-Y. Yoo, and G.-T. Park, *Speed-sensorless vector control of an induction motor using neural network speed estimation.* Industrial Electronics, IEEE Transactions on , Volume: 48 , Issue: 3, June 2001.

[18] R. Patton, J. Chen, and C. Lopez-Toribio, *Fuzzy observers for nonlinear dynamic systems fault diagnosis.* Decision and Control, 1998. Proceedings of the 37th IEEE Conference on , Volume: 1 , 16-18, Dec. 1998.

[19] M.-Y. Chow, R. Sharpe, and J. Hung, *On the application and design of artificial neural networks for motor fault detection. I and II.* Industrial Electronics, IEEE Transactions on , Volume: 40 , Issue: 2, April 1993.

[20] P. Goode and M. yuen Chow, *Using a neural/fuzzy system to extract heuristic knowledge of incipient faults in induction motors. Part I-Methodology.* Industrial Electronics, IEEE Transactions on , Volume: 42 , Issue: 2, April 1995.

[21] K. Kim and A. G. Parlos, *Induction motor fault diagnosis based on neuropredictors and wavelet signal processing.* IEEE/ASME Transactions on Mechatronics, v 7, n 2, June, 2002, p 201-219.

[22] H. Nejjari and M. Benbouzid, *Monitoring and diagnosis of induction motors electrical faults using a current Park's vector pattern learning approach.* Electric Machines and Drives, 1999. International Conference IEMD '99 , 9-12 May, 1999.

[23] S. M. Cruz, H. A. Toliyat, and A. M. Cardoso, *DSP implementation of the multiple reference frames theory for the diagnosis of stator faults in a DTC induction motor drive.* IEEE Transactions on Energy Conversion, v 20, n 2, June, 2005, p 329-335.

[24] D. Lind and M. Marcus, *A Introduction to Symbolic Dynamics and Coding.* Cambridge University Press, United Kingdom, 1995.

[25] H.E. Hopcroft, R. Motwani and J. Ullman, *Introduction to Automata Theory, Languages, and Computation, 2nd ed.* Addison Wesley, Boston, 2001.

[26] R.Duda, P.Hart, and D.Stork, *Pattern classification.* John Wiley and Sons Inc, 2001.

[27] R. Badii and A. Politi, *Complexity hierarchical structures and scaling in physics.* Cambridge University Press, United Kingdom, 1997.

[28] R. Davidchack, Y. Lai, E. Bolt, and H. Dhamala, *Estimating generating partitions of chaotic systems by unstable periodic orbits.* Physical Review, 2000, no. 75.

[29] J. P. Crutchfield and K. Young, *Inferring Statistical Complexity.* Physical Review Letters 63, pp. 105-108, 1989.

[30] Shalizi, C. R., K.L. Shalizi and J. Crutchfield, *An Algorithm for Pattern Discovery in Time Series.* SFI Working Paper 02-10-060, 2002.

[31] A. Ray, *Symbolic Dynamic Analysis of Complex Systems for Anomaly Detection.* Signal Processing, Paper 1640, 2004.

[32] C. Daw, C. Finney, and E. Tracy, "A review of symbolic analysis of experimental data," *Review of Scientific Instruments*, vol. 74, no. 2, pp. 915–930, 2003.

[33] M. Kennel and M. Buhl, *Estimating good discerte partitions form observed data: symbolic false nearest neighbors.* http://arxiv.org/PS cache/nlin/pdf/0304/0304054.pdf, 2003.

[34] L. Rabiner, "A tutorial on hidden markov models and selected applications in speech processing," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[35] M. Benbouzid and G. Kliman, *What stator current processing-based technique to use for induction motor rotor faults diagnosis?* Energy Conversion, IEEE Transactions on , Volume: 18 , Issue: 2 pp.238 - 244, June 2003.

[36] S. G. Mallat;, *A Wavelet Tour of Signal Processing, 2/e.* Academic Press, 1998.

[37] I. Daubechies, *Ten lectures on wavelets.* Society for Industrial and Applied Mathematics,, 1992.

[38] P. Abry, *Ondelettes et turbulence. Multirsolutions, algorithmes de dcomposition, invariance d'chelles.* Diderot Editeur, Paris, 1997.

[39] P. P. Vaidyanathan, *Multirate systems and filter banks.* PTR Prentice Hall, 1993.

[40] V. Rajagopalan and A.Ray, "Symbolic time series analysis via wavelet based partitioning." Seville, Spain: IEEE Conference on Decision and Control, Dec 2005, pp. 5245–5250.

[41] R. Samsi, V. Rajagopalan, J. Mayer, and A. Ray, *Early detection of voltage imbalances in three-phase induction motors.* American Control Conference, 2005. Proceedings of the 2005, June 8-10, 2005 Page(s):478 - 483.

[42] R. Samsi, V. Rajagopalan, and A. Ray, *Wavelet-Based Symbolic Analysis for Detection of Broken Rotor Bars in Inverter-Fed Induction Motors.* To appear in the American Control Conference, 2006. Proceedings of the 2006, 2006.

[43] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis, 2nd ed.* Cambridge University Press, 2003.

[44] E. W. Weisstein, *Discrete Distribution.* MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/DiscreteDistribution.html.

[45] N. Balakrishnan and V. B. Nevzorov, *A primer on statistical distributions.* Wiley, c2003.

[46] A. Surana, A. Ray, and S. Chin, *Anomaly Detection in Complex Systems.* 5th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Process, Washington, D.C., 2003.

[47] P. C. Krause, O. Wasynczuk, and S. D. Sudhoff, *Analysis of electric machinery.* Piscataway, NJ : IEEE Press ; New York, NY : Institute of Electrical and Electronics Engineers, Inc, 1995.

[48] V. Ostovi*c, *Dynamics of saturated electric machines.* New York, N.Y. : Springer-Verlag, 1989.

[49] C. Jeraputra, *Analysis of electromagnetically induced acoustic noise in electrical machines.* Thesis (M.S.)–Pennsylvania State University, 1999.

[50] J. S. Bay, *Fundamentals of linear state space systems.* WCB/McGraw-Hill,ISBN: 0256246394, 1999.

[51] R. Casimir, E. Boutleux, G. Clerc, and F. Chappuis, *Broken bars detection in an induction motor by pattern recognition.* Power Tech Conference Proceedings, 2003 IEEE Bologna, Volume 2, 23-26 June 2003 Page(s):7 pp. Vol.2.

[52] Z. Ye and B. Wu;, *Online rotor bar breakage detection of three phase induction motors by wavelet packet decomposition and artificial neural network.* Power Electronics Specialists Conference, 2001. PESC. 2001 IEEE 32nd Annual, Volume 4, 17-21 June 2001 Page(s):2209 - 2216 vol. 4.

[53] F. Bellini, A.and Filippetti, G. Franceschini, C. Tassoni, and G. Kliman, *Detection of broken rotor bars in induction motors using state and parameter estimation: Quantitative evaluation of induction motor broken bars by means of electrical signature analysis.* Industry Applications, IEEE Transactions on , Volume: 37 , Issue: 5, Sept.-Oct. 2001.

[54] K. Cho, J. Lang, and S. Umans, *Detection of broken rotor bars in induction motors using state and parameter estimation.* Industry Applications, IEEE Transactions on , Volume: 28 , Issue: 3, May-June 1992.

[55] M. E. H. Benbouzid, *A review of induction motors signature analysis as a medium for faults detection.* IEEE Trans. Ind. Electron., vol. 47, pp. 984993, Oct. 2000.

[56] D. Siyambalapitiya and P. McLaren, *Reliability improvement and economic benefits of online monitoring systems for large induction machines.* Industry Applications, IEEE Transactions on Volume 26, Issue 6, Page(s):1018 - 1025, Nov.-Dec. 1990.

[57] J. Stack, T. Habetler, and R. Harley, *Effects of machine speed on the development and detection of rolling element bearing faults.* Power Electronics Letters, IEEE, Volume 1, Issue 1, March 2003 Page(s):19 - 21.

[58] T. A. Harris, *Rolling bearing analysis.* John Wiley, 2001.

[59] B. Yazici and G. Kliman, *An adaptive statistical time-frequency method for detection of broken bars and bearing faults in motors using stator current.* IEEE Transactions on Industry Applications, Volume 35, Issue 2, March-April 1999 Page(s):442 - 452.

[60] A. Prabhakar, S.and Mohanty and A. Sekhar, *Application of discrete wavelet transform for detection of ball bearing race faults.* Tribology International, v 35, n 12, December, 2002, p 793-800.

[61] S. McInerny and Y. Dai, *Basic vibration signal processing for bearing fault detection.* Education, IEEE Transactions on, Volume 46, Issue 1, Feb. 2003 Page(s):149 - 156.

[62] Y. Shao and K. Nezu, *Extracting symptoms of bearing faults in the wavelet domain.* Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering, v 218, n 1, February, 2004, p 39-51.

[63] A. Ray, *Signed Real Measure of Regular languages for Discrete Event Supervisory Control.* International Journal of Control,, Vol. 78, No. 12, August 2005, pp. 949-967.

[64] C. M. Bishop, *Neural Networks for Pattern Recognition.* Oxford University Press Inc., New York, 1995.

[65] M. Markou and S. Singh, "Novelty detection: a review – parts 1 and 2," *Signal Processing*, vol. 83, pp. 2481–2521, 2003.

Vita

Rohan Samsi

Rohan Samsi was born in Mumbai, India. After completing his undergraduate studies in Electrical Engineering from Manipal Institute of Technology, he joined Pennsylvania State University in August 2001. His research interests include fault detection, electrical drives and power electronics. He also holds a Masters degree in Mechanical Engineering. Currently he works for Osram-Sylvania in their Electronic Control Systems division. His interests also include playing soccer and cricket.