Synopsis Seminar Report

on

Frequency Estimation under Stationary and Non-stationary Conditions - A Case Study of Induction Motor Fault Diagnosis

by

Anik Kumar Samanta

(16AT91R11)

Under the supervision of

Prof. Aurobinda Routray

and

Prof. Swanand Ravindra Khare

Advanced Technology Development Centre

Indian Institute of Technology, Kharagpur

January 2021

1 Introduction

Signal parameters carry vital characteristic information about any physical process. This thesis attempts to retrieve the phase information of signals with time-varying and constant frequencies to gain insight into such processes. Conventional multi-component instantaneous frequency (IF) estimators decompose individual components by finding the maximum energy over the time-frequency plane before estimation [1]. However, the decomposition of closely-spaced and crossover units are not yet solved. Moreover, the Gabor-Heisenberg uncertainty principle limits the estimator's performance in terms of its time and frequency resolutions where one can only be gained at the expense of the other [2]. The uncertainty severely deteriorates the conventional methods' performance under higher chirp-rates and abrupt changes in frequency. The resolution of Fourier-based methods is dependent on the window length, and that of the model-based techniques are dependent on the number of parameters or the lags. Over-estimating the model order improves the resolution, but it can result in spurious ridges in the TF-plane. On the other hand, the sequential estimators mainly estimated a single component [3].

Classical stationary spectral estimators like discrete Fourier transform, periodogram, and modified-periodogram have low-resolution and weak accuracy. Alternatively, Subspace-based high-resolution estimators are computationally complex, require model-order information, and cannot give exact information about the amplitude. The model-based methods like autoregressive (AR)-spectrum require an inversion of the autocorrelation matrix [4]. The absence of regularization in AR-spectrum estimation leads to the inversion of an ill-conditioned matrix. Alternatively, regularized least square with Bayesian estimation involves the inversion of a well-conditioned matrix. Bayesian maximum a posterior estimation with Tikhonov prior [5] and variational Bayes with Bernoulli-Gaussian prior [6] are quite popular. In [7], the authors have considered a linear system's output for a given input to estimate the time-varying parameters.

Conventionally, stator current and vibrations signature are analysed for detecting faults in squirrel cage induction motors (SCIMs). The scope and advances of different approaches have been captured in various review articles [8,9]. The spectral fault features are numerically suppressed due to the presence of multiple high-magnitude frequency components. The non-adaptive selection of central frequency and pass-band of the notch filters are disadvantageous when the motors operate under variable frequency and low-load conditions. Teager-Kaiser energy operator [10] and oblique projection [11] have been mainly used for removing unwanted single components.

2 Motivations and Objective of the Thesis

The existing issues that motivate this thesis are as follows,

1. Majority of the IF-estimators indirectly estimate the frequency from time-frequency distributions.

2. High-resolution spectral estimators are computationally complex, non-sequential, and incapable of estimating the amplitude.

3. Conventional signal conditioners for SCIM fault detection are useful for removing only single component or components with harmonic ordering. Notch filters can suppress the required fault information under variable frequency and low-slip operations.

4. Conventional hypothesis tests primarily consider the presence or absence of faults. The generalized likelihood ratio test improves the accuracy under noisy conditions. However, it cannot incorporate the inherent fault information in the decision process.

5. Due to periodic maintenance, the availability of fault data is scarce. The use of generalized models for any motor is still an open area of research.

This thesis aims to detect weak faults in SCIMs, which require the detection and estimation of sinusoidal parameters under stationary and non-stationary conditions. The contributions of the thesis are summarized below,

1. Directly estimating IF of multiple sinusoidal frequency components using a linearized constrained Kalman filter. We extend the estimator for eliminating multiple dominant components for signal conditioning.

2. A Rayleigh-quotient-based spectral estimator which can estimate the frequency and amplitude accurately.

3. A minimum-distance hypothesis test that can incorporate the inherent fault component information for fault detection using only few cases of healthy-motor data.

4. A Bayesian maximum a-posterior sequential spectral estimator using a first-order Gauss-Markov process as the prior distribution.

5. Detection of weak SCIM faults using well-defined physics-based spectral signatures.

6. Two embedded and hardware realizations are proposed comprising of a simulink real-time based method for single motor and an IoT-enabled framework for multiple motor health monitoring.

The proposed methods have been tested and validated on a 22-kW SCIM laboratory setup. The use of vibration signal has also been validated using the publicly available Case Western Reserve University (CWRU) drive-end, 12 kHz bearing data [12]. Short descriptions of each contribution are presented in the subsequent sections.

3 Direct Estimation of Multiple Time-varying Frequencies of Non-stationary Signals

This section presents a generalized framework for real-time tracking of multiple time-varying sinusoidal frequencies of a non-stationary signal. The non-stationary signal is modeled as a time-varying autoregressive (TVAR) process. A non-linear state-space model is formed to truly represent the TVAR process, considering the frequencies as state variables. A signal with p components is defined as

$$x[n] = \sum_{i=1}^{p} \tilde{x}_i[n] + v[n], \tag{1}$$

where

$$\tilde{x}_i[n] = |A_i| e^{j(n\omega_i[n]+\phi_i)}. \tag{2}$$

Here $|A_i|$, $\omega_i[n]$, ϕ_i are the individual amplitude, instantaneous normalized frequency, and phase of the i^{th} component, respectively. $\omega_i[n] = 2\pi f_i[n]/F_s$ and $j = \sqrt{-1}$. $f_i[n]$ and F_s are the i^{th} instantaneous signal frequency (Hz) and sampling rate (samples/s), respectively. $v[n]$ is the additive white Gaussian noise with zero mean and variance $\boldsymbol{R} = \sigma_v^2$, i.e., $v[n] \sim \mathcal{N}(0, \sigma_v^2)$. The problem undertaken in the present work is the online estimation of $\omega_i[n]$ for every i at each instance of n of (2).

With state vector $\boldsymbol{\theta}[n]$ defined as $\boldsymbol{\theta}[n] = [\omega_1[n], \omega_2[n], \cdots, \omega_{\tilde{q}}[n]]^T$, $0 \leq \omega_i[n] < 2\pi$, and the non-linear observation as $x[n] = h(\boldsymbol{\theta}[n]) + v[n]$, we propose Algorithm 1 for instantaneous estimation of multiple frequency components. A constrained optimization estimates the optimal Kalman gain so that the states are restricted within 0 to 2π as

$$\begin{aligned} \hat{\boldsymbol{K}} &= \underset{K}{\operatorname{argmin}} \operatorname{Trace}\left[(\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H}[n])\,\boldsymbol{P}[n|n-1](\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H}[n])^H + \boldsymbol{K}\boldsymbol{R}\boldsymbol{K}^H\right] \\ &\text{subject to}: \; 0 < \hat{\theta}_i[n|n-1] + K_i\left(x[n] - h(\hat{\boldsymbol{\theta}}[n|n-1])\right) < 2\pi, \quad \forall i \end{aligned} \tag{3}$$

in which $\boldsymbol{K}$, $\boldsymbol{R}$, and $\boldsymbol{P}$ are the Kalman gain, observation error covariance, and state error covariance, respectively. K_i is the Kalman gain for the state θ_i, and $i = 1, 2, \ldots, p$. $\boldsymbol{I}$ is an $p \times p$ identity matrix and $(\cdot)^H$ denotes the conjugate transpose operation.

Algorithm 1 Sequential estimation of multiple frequencies

Input: $\hat{\boldsymbol{\theta}}[n-1]$, $x[n]$, $\boldsymbol{P}[n-1]$, $\boldsymbol{Q}$, $\boldsymbol{R}$.
Output: $\hat{\boldsymbol{\theta}}[n]$, $\boldsymbol{P}[n]$.

1: **for all** n such that $n > p$ **do**
2: *Prediction of state:* $\hat{\boldsymbol{\theta}}[n|n-1] = \hat{\boldsymbol{\theta}}[n-1]$
3: *Prediction of Minimum Mean Square Error:* $\boldsymbol{P}[n|n-1] = \boldsymbol{P}[n-1] + \boldsymbol{Q}$
4: *Kalman gain:* $\boldsymbol{K} = \boldsymbol{P}[n|n-1]\boldsymbol{H}^T[n]\left\{\boldsymbol{H}[n]\boldsymbol{P}[n|n-1]\boldsymbol{H}^T[n] + \boldsymbol{R}\right\}^{-1}$
5: *Correction:* $\hat{\boldsymbol{\theta}}[n] = \hat{\boldsymbol{\theta}}[n|n-1] + \boldsymbol{K}\left\{x[n] - h(\hat{\boldsymbol{\theta}}[n|n-1])\right\}$
6: **If** $(\hat{\theta}_i[n] < 0)$ or $(\hat{\theta}_i[n] > 2\pi)$
7: Re-evaluate the Kalman gain with (3) and go to step 5.
8: **End If**
9: *Minimum Mean Square Error*: $\boldsymbol{P}[n] = \boldsymbol{I} - \boldsymbol{K}\boldsymbol{H}[n]\boldsymbol{P}[n|n-1]$
10: **end for**

3.1 Statistical Evaluation of the Instantaneous Frequency Estimates

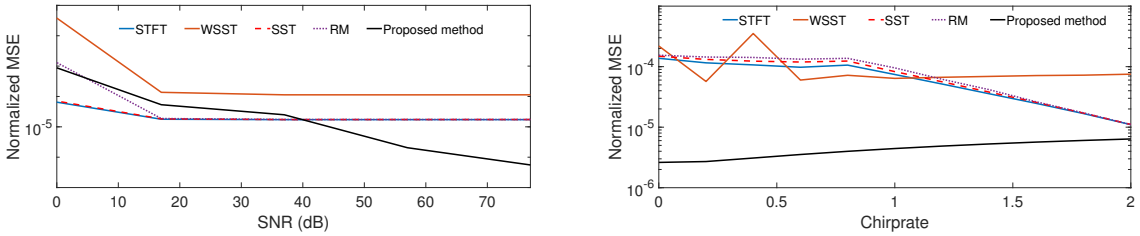
Various statistical measures are adopted to compare the efficacy of the proposed method with four established methods for IF estimation. The compared techniques are short time Fourier transform (STFT) with Gaussian window, analytic Morlet-based wavelet synchrosqueezed transform (WSST) [13], Fourier synchrosqueezed transform (SST) [14], and reassignment method (RM) [14]. Normalized mean-squared error (NMSE) is used to compare the methods in terms of SNR, chirp-rate, resolution, and abrupt change in frequency. We have defined the NMSE as

$$\text{NMSE} = \frac{1}{ML}\sum_{m=1}^{M}\sum_{n=0}^{L-1}\left(\frac{\omega[n] - \hat{\omega}[n]}{\omega[n]}\right)^2, \tag{4}$$

where $\omega[n]$ and $\hat{\omega}[n]$ are the true and estimated frequencies at n^{th} instant, L is the length of the sequence, and M is the total number of trials. 1000 Monte Carlo simulations (MCS) for different noise realizations are used to evaluate the NMSE.

Figure 1a compares the performance of all the methods under different SNR levels. The NMSE of the proposed method for estimating a single component is high when the noise is high. However, with increasing SNR, the NMSE of the proposed method decreases drastically compared to the other methods. It is observed that the NMSE of the other methods converges asymptotically to particular error variances. When the SNR increases, the proposed estimator starts giving more priority to the data than the model. As a result, the NMSE decreases with increasing SNR compared to the other methods.

The next simulation investigates the NMSE for different chirp-rates. As observed in Fig. 1b, the proposed method has an NMSE, which is relatively less for a constant component. The NMSE increases slightly with increasing chirp-rate. However, it is lower than the other methods. For the simulation, we have considered a single frequency component with different chirp-rates.



(a) NMSE versus SNR levels.

(b) Frequency estimate error for different chirp-rates levels, $N = 599$, SNR $= 60$ dB.

Figure 1: NMSE for frequency estimation with 1000 MCS.

4 The Rayleigh-quotient Spectral Estimator

We propose a low-complexity and accurate spectral estimator based on the theory of Rayleigh-quotient [15]. We have used the spectral estimator for analyzing current and vibration signals for detecting the SCIM faults. For any symmetric matrix $\boldsymbol{A}$, with known eigenvector $\boldsymbol{v}$, we can approximate the eigenvalue (λ) by [15]

$$\lambda = \frac{\boldsymbol{v}^H \boldsymbol{A} \boldsymbol{v}}{\boldsymbol{v}^H \boldsymbol{v}}. \tag{5}$$

For a symmetric auto-correlation matrix ($\boldsymbol{R}_x$) of size L, the eigenvalues can be found by

$$\hat{h}(\omega) = \frac{\boldsymbol{w}(\omega)^H \hat{\boldsymbol{R}}_x \boldsymbol{w}(\omega)}{\boldsymbol{w}(\omega)^H \boldsymbol{w}(\omega)}. \tag{6}$$

Given a sinusoidal frequency ω to be present ([16], Pg. 452) in the signal, $\boldsymbol{w}(\omega)$ as in (7) becomes an eigenvector of $\hat{\boldsymbol{R}}_x$.

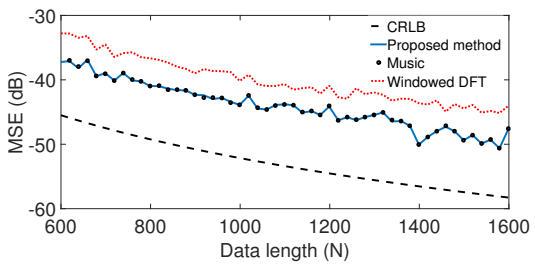
$$\boldsymbol{w}(\omega) = \begin{bmatrix} e^{j\omega \cdot 0} & e^{j\omega} & \cdots & e^{j\omega \cdot (L-1)} \end{bmatrix}^H, \tag{7}$$

The value of the peak in $\hat{h}(\omega)$ is an eigenvalue corresponding to the eigenvector $\boldsymbol{w}(\omega)$ and is proportional to the amplitude of the sinusoid present in the signal. The location of the peak gives the frequency. The relation between the eigenvalue and the sinusoidal amplitude is derived as

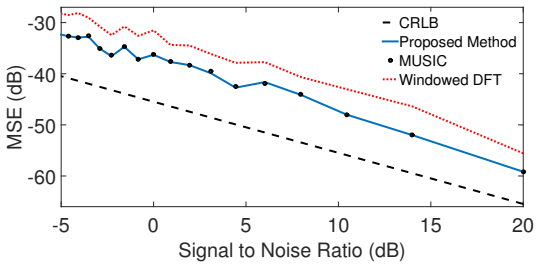
$$\hat{A}_k = \sqrt{\hat{h}(\omega_k) \Big/ L}. \tag{8}$$

4.1 Statistical Evaluation of the Rayleigh Quotient Spectral Estimator

The accuracy and robustness of the Rayleig spectrum is compared with discrete Fourier transform (DFT) and multiple signal classifier (MUSIC). In Fig. 2a, mean square error (MSE) between the input sinusoidal frequency and the location of the peak obtained from the respective spectral estimator are evaluated by 100 trials for data length N. It is inferred that MUSIC and the proposed spectral estimator have slightly higher accuracy for similar data length than DFT with Chebyshev window. In Fig. 2b, the MSE is evaluated with different signal-to-noise ratio (SNR) for 100 trials each. It is found that the performance of the proposed spectral estimator in a noisy environment is quite robust and is equivalent to that of MUSIC and is marginally better than DFT with windowed data. A single sinusoid of 50.11 Hz with unity amplitude has been used for both experiments and $F_s = 200$ Hz. The proposed method is most suitable for fault detection as it is fast and can estimate the value of fault frequency. Unlike MUSIC, it can also estimate the amplitude accurately and doesn't require any information about the number of sinusoids.



(a) Frequency estimate error for different values of N.



(b) Frequency estimate error for different SNR levels, $N = 599$.

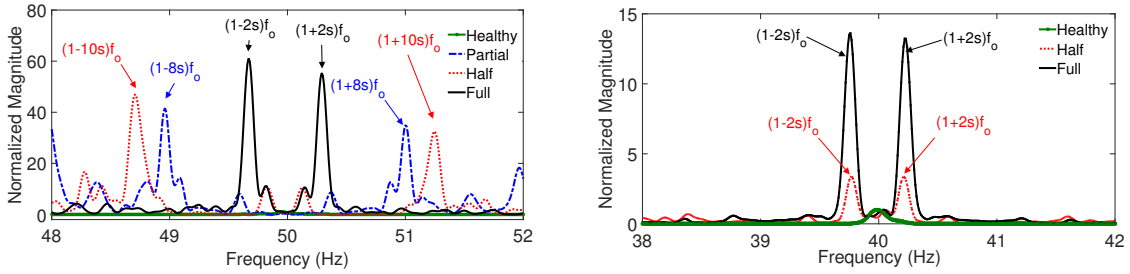
Figure 2: Evaluation of frequency estimate error for different data length and SNR levels

4.2 Stator Current Analysis to Detect Partially a Broken Rotor Bar

We analyze the effect of different broken rotor bar (BRB) fault levels on the stator current spectra, with similar loading in all the cases. An EKF-based signal conditioner is used for the fundamental component suppression and the spectrum is normalized with the magnitude of the attenuated component. It is observed in Fig. 3a that, with increasing BRB severity, the fault component's magnitude also increases. The fault frequency peaks for different cases don't exactly match due to slight variations in each case's supply frequency. For partial and half BRB cases, the fourth and fifth fault harmonic components are exceptionally prominent for the 50 Hz supply. Fig. 3b shows the spectrum of the stator current obtained by the proposed spectral estimator for 40 Hz supply frequency and 0.33% slip for all the cases.

5 Targeted Removal of Multiple Dominant Components and Minimum Distance Detection

A dominating i^{th} component without the noise is represented by $\tilde{x}_i[n]$ in (2). We will estimate $\omega_i[n]$ and $\tilde{x}_i[n]$ for multiple dominant components for each instance of n, and then use the estimated $\tilde{x}_i[n]$ for its elimination from the input signal $x[n]$. We define $\tilde{q}$ as the number of



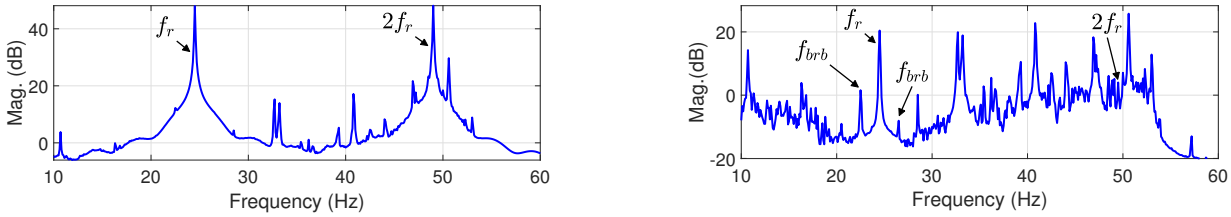
(a) 50 Hz supply (0.26% slip) (b) 40 Hz supply (0.33% slip)

Figure 3: Spectrum of stator current with different levels of BRB fault for 50 Hz and 40 Hz supply frequency

dominating components, assumed to be a known quantity. We presume that the magnitudes of the dominating $\tilde{q}$ components are significantly higher than the other $(q - \tilde{q})$ components, i.e., $|A_i| \gg |A_j|$, $\forall \{(i, j) : i \in [1, \tilde{q}], j \in [\tilde{q} + 1, q]\}$. With complex amplitude representation in (2), we have $\tilde{x}_i[n] = A_i e^{jn\omega_i[n]}$ with $A_i = |A_i| e^{j\phi_i}$. From the nonlinear observation $x[n] = h(\boldsymbol{\theta}[n]) + v[n]$, we obtain the conditioned signal $\hat{x}_0[n]$ by eliminating the estimated $h(\hat{\boldsymbol{\theta}}[n])$ from the input $x[n]$ using Algorithm 1 as

$$\hat{x}_0[n] = x[n] - h(\hat{\boldsymbol{\theta}}[n]). \tag{9}$$

An example test case to compare the spectrum of the raw and conditioned vibration signal for a faulty motor is shown in Fig. 4. We observe in Fig. 4(a) that the spectral leakage from the dominant rotational frequency and its harmonics has obscured the fault components. The proposed signal conditioning has attenuated the dominant components and reduced the spectral leakage. As a result, the fault component peaks have conspicuously emerged in Fig. 4(b). The Rayleigh quotient spectrum has been used to generate the plot of Fig. 4.



(a) (b)

Figure 4: Spectrum analysis of motor vibration (a) before and (b) after signal conditioning.

5.1 The Minimum Distance Detector

To design the test, we use the real part of (2) as

$$x(n) = A\cos(\omega n + \phi) + v(n). \tag{10}$$

The fault frequency (ω) is assumed to be time-invariant, and the hypothesis testing (11) is formulated under the assumption that $v(n) \sim \mathcal{N}(0, \sigma^2)$ as

$$\begin{aligned} \mathcal{H}_0 :&\quad x(n) = A\cos(\omega n + \phi) + v(n); \quad A < \gamma, \\ \mathcal{H}_1 :&\quad x(n) = A\cos(\omega n + \phi) + v(n); \quad A \geq \gamma, \end{aligned} \tag{11}$$

where γ is the tunable threshold that can be learned from the healthy data. The fault component along with noise is represented as

$$x(n) = \alpha_1 \cos(\omega n) + \alpha_2 \sin(\omega n) + v(n),$$

where $A = \sqrt{\alpha_1^2 + \alpha_2^2}$, and $\phi = \arctan\left(-\alpha_2/\alpha_1\right)$. Collating N data points, the observation $\boldsymbol{x}$ is represented as

$$\boldsymbol{x} = \boldsymbol{M}\boldsymbol{\alpha} + \boldsymbol{v},$$

where $\boldsymbol{M} = \begin{bmatrix} 1 & \cdots & \cos(N-1)\omega \\ 0 & \cdots & \sin(N-1)\omega \end{bmatrix}^T$, and $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$. The minimum distance detector chooses the hypothesis that has the higher conditional likelihood $p\left(\boldsymbol{x}; \hat{\boldsymbol{\alpha}}_i, \hat{\sigma}_i^2 | \mathcal{H}_i\right)$ [17],

$$p\left(\boldsymbol{x}; \hat{\boldsymbol{\alpha}}_i, \hat{\sigma}_i^2 | \mathcal{H}_i\right) = C \exp\left[-\frac{\left(\boldsymbol{x} - \boldsymbol{M}\hat{\boldsymbol{\alpha}}_i\right)^T \left(\boldsymbol{x} - \boldsymbol{M}\hat{\boldsymbol{\alpha}}_i\right)}{2\hat{\sigma}_i^2}\right].$$

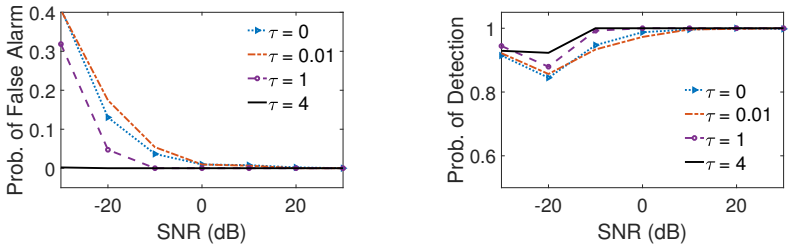
Alternatively, the hypothesis $\mathcal{H}_i$ is chosen, if the $D_i^2(x) = \left(\boldsymbol{x} - \boldsymbol{M}\hat{\boldsymbol{\alpha}}_i\right)^T \left(\boldsymbol{x} - \boldsymbol{M}\hat{\boldsymbol{\alpha}}_i\right)$ under that hypothesis is minimum. Neglecting the $\boldsymbol{x}^T\boldsymbol{x}$ term as it is same under both the hypothesis, we have

$$D_i^2(x) = -2\boldsymbol{x}^T\boldsymbol{M}\hat{\boldsymbol{\alpha}}_i + \hat{\boldsymbol{\alpha}}_i^T\boldsymbol{M}^T\boldsymbol{M}\hat{\boldsymbol{\alpha}}_i. \tag{12}$$

Rearranging the terms of (12), we define the test statistic for $\mathcal{H}_i$ as

$$T_i(x) = \boldsymbol{x}^T\boldsymbol{M}\hat{\boldsymbol{\alpha}}_i - \frac{1}{2}\hat{\boldsymbol{\alpha}}_i^T\boldsymbol{M}^T\boldsymbol{M}\hat{\boldsymbol{\alpha}}_i. \tag{13}$$

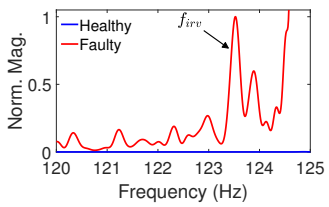
A MCS is conducted to determine the proposed detector's performance using a synthetic signal. The parameters of the input signal (10) used for the MCS are $\omega \sim \mathcal{U}(0.01\pi, 0.5\pi)$, $\phi \sim \mathcal{U}(0, \pi)$, $p(A|\mathcal{H}_0) \sim \mathcal{U}(0, \gamma - \tau)$, $p(A|\mathcal{H}_1) \sim \mathcal{U}(\gamma + \tau, 10)$, and $\gamma = 5$, where $\mathcal{U}(c, d)$ denotes the uniform distribution within the range c and d. The tolerance τ decides the separation between the two hypotheses from which A is drawn. Figure 5 shows the effect of the input SNR on the detector's performance for different tolerance values. We observe that with increasing SNR, the detector performance improves as the probability of detection increases, and the likelihood of false alarm decreases. Also, the performance of the detector improves as τ increases. We have carried out 1000 trials for each SNR value.



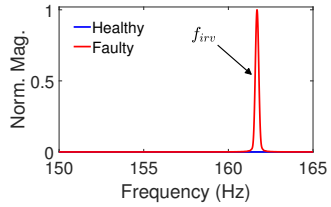
(a) Effect on the probability of false alarm

(b) Effect on the probability of detection

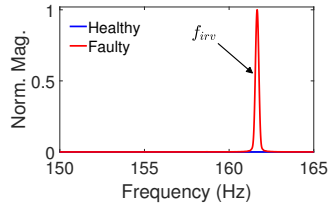
Figure 5: Effect of SNR for different values of tolerances (τ) on the probability of false alarm and probability of detection for the designed detector.



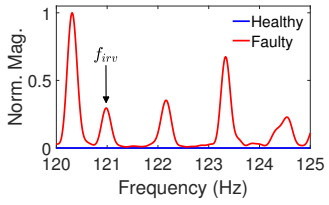
(a) Lab setup, low load, 0.2% slip, $f_r = 24.95$.



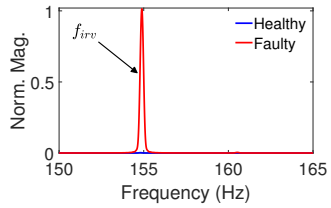
(b) CWRU, 0.007-inch fault, 0 HP load, $f_r = 29.95$.



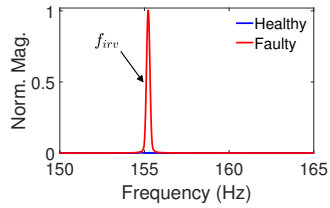
(c) CWRU, 0.028-inch fault, 0 HP load, $f_r = 29.95$.



(d) Lab setup, 67% load, 1.93% slip, $f_r = 24.5175$.



(e) CWRU, 0.007-inch fault, 3 HP load, $f_r = 28.83$.



(f) CWRU, 0.028-inch fault, 3 HP load, $f_r = 28.83$.

Figure 6: Spectrum of the conditioned vibration signal obtained from the lab setup and CWRU motor with inner-race fault for different load levels and severity.

5.2 Vibration Analysis to Detect Inner Raceway Defect

We have used a damaged inner raceway for multiple experiments to compare the faulty motor signature with a healthy motor data. Figures 6a and 6d shows the spectrum of the conditioned vibration signal from the single axis (driving end) sensor with two different load conditions for the lab setup. With increased loading, identifying the inner-raceway fault component has improved considerably as the number of spurious peaks has reduced. The fault component for the CWRU data is quite prominent for all the cases. Figures 6b and 6e gives the spectrum of the motor with a 0.007-inch (0.1778 mm) fault for no-load and 3-HP load, respectively. The vibration spectrum for increased severity with the 0.028-inch (0.7112 mm) fault in the inner-raceway for two load conditions are shown in Fig. 6c and Fig. 6f, respectively. For the CWRU data, the fault components are considerably prominent for all the cases, and they match with the theoretical fault component.

6 Sequential Bayes Maximum a-Posterior Spectral Estimation

A q-component sinusoidal signal $x(n)$ with amplitude A_i, normalized frequency ω_i, and phase ϕ_i given by $x(n) = \sum_{i=1}^{q} A_i \cos(n\omega_i + \phi_i)$ can be represented by a TVAR model with parameters a_i as

$$x(n) = \sum_{i=1}^{p} a_i(n)x(n-i) + v(n), \tag{14}$$

where $v(n)$ is the additive white Gaussian noise with zero mean and covariance σ_v^2, i.e., $v(n) \sim \mathcal{N}(0, \sigma_v^2)$, and $p = 2q$. Defining $\boldsymbol{x}(n) = \begin{bmatrix} x(n) & x(n-1) & \cdots & x(n-p+1) \end{bmatrix}^T$, the

joint distribution of the observation using the chain rule is given as

$$p\left(\boldsymbol{x}(n)|\boldsymbol{x}(n-1),\boldsymbol{a}(n),\sigma_v^2\right)=\prod_{i=p}^{n}p\left(x(i)|x(i-1)\cdots x(i-p+1),\boldsymbol{a},\sigma_v^2\right). \tag{15}$$

Hence,

$$p\left(\boldsymbol{x}(n)|\boldsymbol{x}(n-1),\boldsymbol{a}(n),\sigma_v^2\right)=k_1\exp\left\{-\sum_{i=p}^{n}\frac{\left[x(i)-\boldsymbol{a}^T(n)\boldsymbol{x}(i-1)\right]^2}{2\sigma_v^2}\right\}, \tag{16}$$

where $k_1=(2\pi\sigma_v^2)^{-(n-p+1)/2}$. To introduce uncertainty in the model parameters, we define the parameter $\boldsymbol{a}(n)$ as a Gauss-Markov random walk. The random-walk enables the sequential estimation paradigm and is given as

$$\boldsymbol{a}(n)=\boldsymbol{a}(n-1)+\boldsymbol{u}(n), \tag{17}$$

where $\boldsymbol{u}(n)$ is Gaussian distribution denoted as $\boldsymbol{u}(n)\sim\mathcal{N}\left(\boldsymbol{0},\sigma_u^2\boldsymbol{I}\right)$, and $\boldsymbol{0}$ is a vector with p zeros. The prior distribution of the parameter denoted by $p(\boldsymbol{a}(n))$ from (17) is given as

$$p\left(\boldsymbol{a}(n)\right)=k_2\exp\left\{-\tfrac{1}{2\sigma_u^2}\left[\boldsymbol{a}(n)-\boldsymbol{a}(n-1)\right]^T\left[\boldsymbol{a}(n)-\boldsymbol{a}(n-1)\right]\right\}, \tag{18}$$

where $k_2=(2\pi\sigma_u^2)^{-p/2}$. The posterior probability of the parameter $\boldsymbol{a}(n)$ conditioned on the observation is given by

$$p\left(\boldsymbol{a}(n)|\boldsymbol{x}(n)\right)=\frac{p\left(\boldsymbol{x}(n)|\boldsymbol{a}(n)\right)p\left(\boldsymbol{a}(n)\right)}{p\left(\boldsymbol{x}(n)\right)}. \tag{19}$$

Therefore, putting (16) and (18) in (19), we have the posterior probability as

$$p\left(\boldsymbol{a}(n)|\boldsymbol{x}(n)\right)\propto k_1k_2\exp\left\{-\sum_{i=p}^{n}\tfrac{\left[x(i)-\boldsymbol{a}^T(n)\boldsymbol{x}(i-1)\right]^2}{2\sigma_v^2}\right\}\exp\left\{-\tfrac{\left[\boldsymbol{a}(n)-\boldsymbol{a}(n-1)\right]^T\left[\boldsymbol{a}(n)-\boldsymbol{a}(n-1)\right]}{2\sigma_u^2}\right\}, \tag{20}$$

Maximizing (20) with respect to the unknown parameter vector, and using Algorithm 2, we obtain the spectrum.

6.1 Statistical Evaluation of the Bayes Spectral Estimator

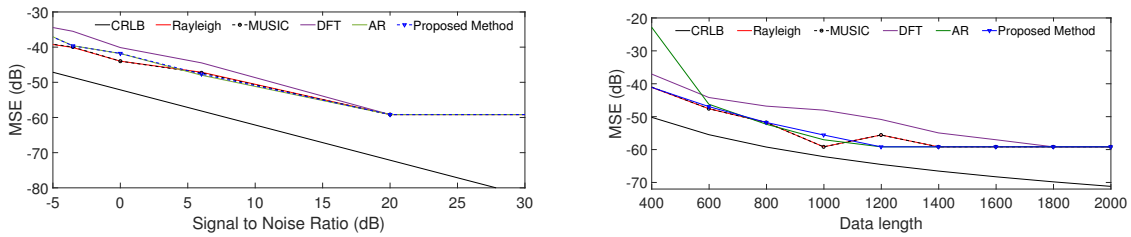
We evaluate the peak-MSE of the spectrum and how it is affected by the input SNR as shown in Fig. 7a. For this simulation, 100 trials are performed for each SNR. The noise is generated from a Gaussian distribution with zero mean and appropriate variance for each trial. The signal's true frequency was kept constant at 50.11 Hz, and amplitude was fixed at 1 unit. The performance of all the spectral estimators follows each other quite closely. If observed minutely, the proposed estimator performs marginally better when the SNR is high. However, its performance is comparable to the AR-spectrum.

The MSE with different values of N is plotted in Fig. 7b. For this simulation, a single component with a frequency of 51.11 Hz was taken. The SNR was fixed at 6.98 dB. It is observed that the MSE decreases with increasing data length for all the estimators. The MSE of all the estimators asymptotically converges to a constant value, after which N does not affect the MSE. It is also observed that the asymptotic value is achieved fastest by both the AR-spectrum and the proposed method. However, the proposed method performs better than the AR-spectrum when N is low.

Algorithm 2 Sequential spectral estimation

Input: $\hat{\boldsymbol{a}}(n-1)$, $x(n)$, $\boldsymbol{R}(n-1)$, $\boldsymbol{\rho}(n-1)$, $\hat{\sigma}_u^2(n-1)$, $\hat{\sigma}_v^2(n-1)$.
Output: $\hat{\boldsymbol{a}}(n)$, $P(\omega)$.

1: **for all** n such that $n > p$ **do**
2: $\quad \boldsymbol{R}_{\boldsymbol{xx}}(n) = \boldsymbol{R}_{\boldsymbol{xx}}(n-1) + \boldsymbol{x}(n-1)\boldsymbol{x}^T(n-1)$
3: $\quad \boldsymbol{\rho}_{\boldsymbol{x}}(n) = \boldsymbol{\rho}_{\boldsymbol{x}}(n-1) + \boldsymbol{x}(n-1)x(n)$
4: $\quad \hat{\sigma}_v^2(n) = \frac{1}{n-p+1}\left[\hat{\sigma}_v^2(n) + \left\{x(n) - \hat{\boldsymbol{a}}^T(n)\boldsymbol{x}(n-1)\right\}^2\right]$
5: $\quad \hat{\sigma}_u^2(n) = \frac{1}{p}\left\|\hat{\boldsymbol{a}}(n) - \hat{\boldsymbol{a}}(n-1)\right\|_2^2$
6: $\quad \kappa(n) = \sigma_v^2(n)/\sigma_u^2(n)$
7: $\quad \hat{\boldsymbol{a}}(n) = \left[\boldsymbol{R}_{\boldsymbol{xx}}(n) + \kappa\boldsymbol{I}\right]^{-1}\left[\boldsymbol{\rho}_{\boldsymbol{x}}(n) + \kappa\boldsymbol{a}(n-1)\right]$
8: $\quad \boldsymbol{x}(n) = \left[x(n), \boldsymbol{x}^T(n-1)\right]^T$
9: **end for**
10: $P(\omega, n) = \hat{\sigma}_v^2(n) \Big/ \left|\sum_{i=0}^{p} \hat{a}_i[n]e^{-j\omega}\right|^2$,

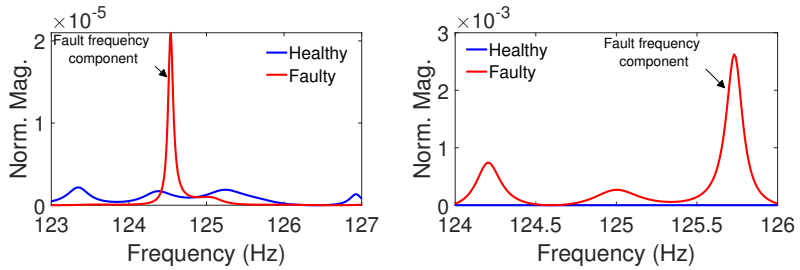


(a) MSE variation with the SNR. $N = 1000$ (b) MSE variation with data length. SNR = 6.98 dB.

Figure 7: Mean squared error evaluation

6.2 Stator Current Analysis to Detect Outer Raceway Defect

Figure 8 shows the spectrum of the stator current for a motor with outer raceway fault when the motor is operated under low-load and high-load conditions. The fault specific components can be easily distinguished when compared to a healthy motor spectrum.



(a) Very low load, 0.2% slip. (b) 67%load, 1.93% slip

Figure 8: Outer-race fault spectrum extracted from the stator current.

7 Embedded System Development for Online Fault Detection

Two schemes for online implementations of the fault detector are presented in this thesis. First, we use the Rayleigh-quotient spectrum for detecting fault components. Simulink Real-time was used for implementing the fault detection algorithm using an Intel embedded hardware. A single-phase stator current was recorded using NI-PCI 6024 E analog interface.

We have used the sequential Bayes spectrum of conditioned stator current for fault diagnosis in the second implementation. This framework is useful for monitoring multiple motors with IoT-enabled modules that can record and process stator current and transmit the decision to a central server for storing. We use local processing units to reduce large scale data transmission. Two ESP modules are configured as master-slave units to function as the data-handler and fault detector, respectively. A sampling rate of 400 Hz is suitable for acquiring the data without aliasing for both the implementations, considering lower execution time.

8 Conclusions

This thesis represents the research for the detection and estimation of frequency-based features under adverse situations when the low-amplitude signals of interest are buried under noise and masked by the presence of high-magnitude elements. We have adopted a model-based approach for simultaneous estimation of a multicomponent signal's instantaneous frequency using a linearized-observation constrained-Kalman approach. The phase abstraction property of the method was instrumental in extracting the dominant modes of the signal. The simulation results have shown the estimator to be suitable for time-varying closely spaced and sharply changing elements. The technique outperformed other methods for signal having higher chirp-rate. However, the performance under low-SNR needs improvement. The removal of dominant modes condition the signal and reduces the spectral leakage. The provision of using a precise model can further improve the accuracy of the method.

We have proposed two spectral estimators in the thesis for stationary conditions when the frequency is constant in the observation window. The Rayleigh-quotient-based method has high frequency and amplitude accuracy and requires low computational resources. The approach is data-driven and doesn't require any information about the underlying model. Alternatively, Bayes MAP spectral estimator is model-based and can incorporate knowledge of the underlying model. It is also sequential and has higher accuracy than the Rayleigh-quotient spectrum. The use of an accurate model can further improve the precision of the approach.

The algorithms presented in this thesis have been validated using the case study of detecting the low-amplitude frequency component of weak SCIM faults. A minimum distance-based hypothesis test has been proposed for incorporating the inherent fault information. Two embedded platforms are also presented in this thesis for hardware realization of the suggested algorithms. The Simulink Real-time-based hardware is appropriate for detecting faults in a single motor. However, the hardware is costly, but its flexibility for initial feasibility studies is advantageous. On the other hand, the IoT-based system has been tailor-made for dedicated fault detection in a multiple-motor scenario. Nevertheless, further investigation into the memory, speed, and power requirement of the methods are required for product development.

9 Organization of the Thesis

The thesis is organized as follows:

Chapter 1: This chapter provides the background, motivation, and objective of this thesis to detect and estimate low-amplitude frequency components and their utility in diagnosing SCIM faults.

Chapter 2: Proposes a generalized framework for real-time tracking of multiple time-varying sinusoidal frequencies of a non-stationary signal.

Chapter 3: Introduces the Rayleigh-quotient-based spectral estimator and uses it for detecting partially broken rotor bar faults using stator current.

Chapter 4: This chapter enhanced the IF-estimator to eliminate multiple dominant components of a vibration signal. A minimum distance detector is proposed for generalized fault detection.

Chapter 5: Proposes a Bayesian MAP-based sequential spectral estimator for detecting SCIM faults using a single-phase stator current input.

Chapter 6: This chapter discusses a standalone and an IoT-based embedded system for online SCIM fault detection.

Chapter 7: This chapter concludes the thesis and describes future research directions.

10 Publications from the Thesis

Patent Filed:

1. A. Routray, A. Naha, **A. K. Samanta**, Amey Pawar, & Chandrasekhar Sakpal, "A system for assessment of multiple faults in induction motors", WO2019167086A1, 2019.

Journals:

1. **A. K. Samanta**, A. Routray, S.R. Khare, & A. Naha, "Minimum Distance-based Detection of IncipientInduction Motor Faults using Rayleigh Quotient Spectrum of Conditioned Vibration Signal [Accepted]", *IEEE Trans. Instrum. Meas.*

2. **A. K. Samanta**, A. Routray, S.R. Khare, & A. Naha, "Direct Estimation of Multiple Time-varying Frequencies of Non-stationary Signals", *Signal Process.*, vol.169, 2020.

3. **A. K. Samanta**, A. Naha, A. Routray, & A. K. Deb "Fast and accurate spectral estimation for online detection of partial broken bar in induction motors", *Mech. Sys. Signal Process.*, vol. 98, pp. 63-77, 2018.

References

[1] N. A. Khan and B. Boashash, "Instantaneous frequency estimation of multicomponent nonstationary signals using multiview time-frequency distributions based on the adaptive fractional spectrogram," *IEEE Signal Process. Lett.*, vol. 20, no. 2, pp. 157–160, 2013.

[2] M. Vetterli, J. Kovačević, and V. K. Goyal, *Foundations of signal processing.* Cambridge University Press, 2014.

[3] A. Routray, A. K. Pradhan, and K. P. Rao, "A novel Kalman filter for frequency estimation of distorted signals in power systems," *IEEE Trans. Instrum. Meas.*, vol. 51, no. 3, pp. 469–479, jun 2002.

[4] T. L. Hansen, B. H. Fleury, and B. D. Rao, "Superfast Line Spectral Estimation," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2511–2526, 2018.

[5] H. Fu and P. Y. Kam, "MAP/ML Estimation of the Frequency and Phase of a Single Sinusoid in Noise," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 834–845, 2007.

[6] M.-A. Badiu, T. L. Hansen, and B. H. Fleury, "Variational Bayesian Inference of Line Spectra," *IEEE Trans. Signal Process.*, vol. 65, no. 9, pp. 2247–2261, 2017.

[7] Y. Chu and C. M. Mak, "A New Parametric Adaptive Nonstationarity Detector and Application," *IEEE Trans. Signal Process.*, vol. 65, no. 19, pp. 5203–5214, 2017.

[8] S. Nandi, H. A. Toliyat, and X. Li, "Condition Monitoring and Fault Diagnosis of Electrical Motors — A Review," *IEEE Trans. Energy Convers.*, vol. 20, no. 4, pp. 719–729, dec 2005.

[9] Y. Lei, B. Yang, X. Jiang, F. Jia, N. Li, and A. K. Nandi, "Applications of Machine Learning to Machine Fault Diagnosis: A review and Roadmap," *Mech. Syst. Signal Process.*, vol. 138, p. 106587, 2020.

[10] M. Pineda-sanchez, J. Perez-cruz, J. Pons-llinares, V. Climente-alarcon, and J. A. Antonino-daviu, "Application of the Teager – Kaiser Energy Operator to the Fault Diagnosis of Induction Motors," *IEEE Trans. Energy Convers.*, vol. 28, no. 4, pp. 1036–1044, 2013.

[11] E. Elbouchikhi, V. Choqueuse, F. Auger, and M. E. H. Benbouzid, "Motor Current Signal Analysis based on a Matched Subspace Detector," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 12, pp. 3260–3270, 2017.

[12] "Case Western Reserve University Bearing Data Center." [Online]. Available: http://csegroups.case.edu/bearingdatacenter/home

[13] G. Thakur, E. Brevdo, N. S. Fučkar, and H.-T. Wu, "The synchrosqueezing algorithm for time-varying spectral analysis: Robustness properties and new paleoclimate applications," *Signal Process.*, vol. 93, no. 5, pp. 1079–1094, 2013.

[14] F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H.-T. Wu, "Time-frequency reassignment and synchrosqueezing: An overview," *IEEE Signal Process. Mag.*, vol. 30, no. 6, pp. 32–41, 2013.

[15] A. K. Samanta, A. Naha, A. Routray, and A. K. Deb, "Fast and accurate spectral estimation for online detection of partial broken bar in induction motors," *Mech. Syst. Signal Process.*, vol. 98, pp. 63–77, 2018.

[16] M. H. Hayes, *Statistical Digital Signal Process. and Modeling.* John Wiley & Sons, 2010.

[17] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory.* Prentice-Hall (Upper Saddle River, NJ [ua]), 1998.