An iterative tuning strategy for achieving Cramer Rao bound using Extended Kalman filter for a parameter estimation problem

by

Shyam Mohan M

(12104078)

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY KANPUR

June 2014

An iterative tuning strategy for achieving Cramer Rao bound using Extended Kalman filter for a parameter estimation problem

A Thesis Submitted

in Partial Fulfillment of the Requirements

for the Degree of

Master of Technology

by

Shyam Mohan M

(12104078)

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY KANPUR

June 2014

POST GRADUATE OFFICE
THESIS SUBMITTED
ON 12-06-2014
I. I. T., KANPUR - 16

CERTIFICATE

It is certified that the work contained in the thesis entitled "**An iterative tuning strategy for achieving Cramer Rao bound using Extended Kalman filter for a parameter estimation problem**" by **Shyam Mohan M** has been carried out under the supervision of Prof. M.R. Ananthasayanam (retired from the Department of Aerospace Engineering , Indian Institute of Science , Bangalore-560012) and myself . This work has not been submitted elsewhere for a degree.

12 June 2014

Dr. Naren Naik
Department of Electrical Engineering,
Indian Institute of Technology,
Kanpur - 208016.

ABSTRACT

The Kalman filter (KF) is a powerful tool widely used to estimate quantities in the presence of noise, be it in the process or the measurement. The catch in the approach is the need to know the process and measurement noise statistics as well as the initial estimate and covariance of the states and parameters. The setting of these typically apriori-unknown quantities is called tuning of the KF.

Tuning is typically done either manually or via adaptive estimation; however there is still a need for adaptive estimation of the unknown system parameters as well as the noise statistics, that yields accurate estimates and error covariances for the states as well as parameters.

In our work an iterative tuning procedure is developed not only to estimate the parameters but also to achieve the Cramer Rao bound (CRB), thus making it an efficient estimator. A smoothing technique is used to solve the initial conditions for the state, and a new virtual measurement concept is introduced to initialize the filter for unknown parameters. The measurement noise matrix is estimated by the Myers and Tapley covariance matching scheme.

A new method for estimating the process noise is introduced which uses the difference between the stochastic and deterministic trajectory. A tuning strategy is developed for both zero and non zero process noise case and many sensitivity studies are conducted on test problems to study the robustness of the algorithm, using a user friendly graphic user interface.

Dedicated

to

IIT Kanpur

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to my thesis supervisor Dr. Naren Naik for his valuable guidance and continuous support throughout my thesis work. His encouraging words and composed positive attitude were and will always be a source of inspiration to me. I am highly indebted to him for giving me the freedom to work on the area which interested me and taking a course on Kalman Filtering concepts. It has always been a pleasant experience to work with him.

I also had the opportunity to work with Prof. M. R. Ananthasayanam who is a retired professor from the department of Aerospace Engineering(IISc-Bangalore). I would always be grateful to him for making me understand the problem statement and continuously monitor my work so that its never off the track.

My greatest respect and gratitude to the faculty members of EE department for imparting their invaluable knowledge and their selfless dedicated efforts. Prof. Aditya K Jagannatham (EE,IITK) made us understand the basics of statistical signal processing which was very useful and interesting.

I thank all my classmates for giving me the strength for competency. I remember the quality time that I spent with my friend Rashid Ummer who would make me understand simple things in a simple way.

Friends are indeed an invaluable gift.I would like to thank my friends Rakhesh, Suraj, Harish, Karthik, Josh and Joshua for their moral support. My whole hearted thanks to my friend Sindhu for always wishing my happiness and success. I would also like to express my gratitude towards my colleagues especially Nidhi, Shivangi, Kaveri, Arun , Saurabh Shukla, Faisal, Sandeep Chandwani, Sidharth Khurana, Umesh Hasija, Sumit Bhugra, Shubham, Arpit, Vishwaneel Dubey, Chidambara Raj, Harinder Singh Basra, Anurag Khare, Harsh Wardhan, Ashutosh, Anish Jose, Amit Hingar, Asheesh Kumar, Ankit, Ashish Goel, Sukhpreet Singh, Snehdeep, Abhishek, Amit Sinha and Saurabh Shrivastav for their good wishes.

I would like to make a special mention of my younger brother Man Mohan M who have been the greatest source of support by all means. Last but not the least I thank my parents who gave me love and affection all through my life.

- Shyam Mohan M

Contents

Chapter 1

Introduction

Kalman proposed the standard five steps for filtering methods in his famous paper in year 1960 [1] which is well known as Kalman filter(KF). The solution was only formal since it assumed that the noise statistics (X_0, P_0, Q, R) and the parameters(θ) of the system model are known. The euphoria created after the introduction of Kalman filter was followed by a depression posing a big question on how to run the filter in an unknown noise statistics environment. Several research works have tackled the adaptive procedures which is still going on. Tuning is done manually even today since the adaptive filtering method had not matured to a great level. This manual approach may give satisfactory results after many trial and errors but will be inferior to the case when the noise statistics are estimated adaptively. Most of the literature assumes that the system parameters are known and attempts to solve the problem. The methods for estimating the unknown noise covariances in the Kalman Filter literature can be broadly classified into four categories namely -covariance matching, correlation techniques, Bayesian and maximum likelihood methods. The covariance matching technique [7] is an adaptive algorithm in which the estimates of the process noise covariance (Q) and measurement noise covariance (R) are computed at every sampling instant. A problem that arises with this technique is that it does not maintain the positive definiteness of Q and R. Moreover, Odelson et al. [12] have shown that the covariance matching method yields biased estimates of the covariances. Correlation techniques estimate the covariance matrices by making use of the sample auto-correlations between the innovations [15]. A least square estimate of Q and R is generated as proposed by Mehra [13] and Belanger [14]. Bayesian methods and in particular the maximum likelihood based methods formulated the covariance estimation problem as maximisation of the likelihood function associated with the innovations. The expectation-maximisation (EM) algorithm developed by Dempster et al. [16] is an iterative procedure that uses the complete data likelihood function to compute the maximum likelihood estimates of the model parameters. Shumway , Stoffer [17] and Raghavan et al. [18] provide a framework to identify state space models along with the state and measurement noise covariances for linear systems using the EM algorithm.Recently, Vinay A Bavdekar et al. [11] showed direct optimisation and EM algorithm based technique in estimating Q and R. These methods are not preferred as they are computationally demanding. A more practical monte carlo approach was suggested by Jaleel Valappil and Christos Georgakis [10] which uses parameter covariance matrix for finding Q which usually gives a lower estimate than expected. R.M.O Gemson and M.R. Ananthasayanam [5] highlighted the importance of tuning P_0. However none of the literature work spoke about achieving the CRB's of the parameter estimates.

Our work aims to solve the tuning problem using heuristic combined with intuitive approaches when both the parameters and the noise statistics are unknown in case of a nonlinear system. In order to obtain an insight into such a problem when the noise statistics are unknown one has to check for its convergence for multiple input values and many sensitivity studies has to be conducted. If one wishes good CRB's of the parameter estimates then adaptive tuning cannot be ignored especially the right choice of P_0 corresponding to the unknown parameters for an optimum filter operation. Only a well tuned filter can achieve the Cramer Rao bound criterion , making the EKF an efficient and consistent estimator. We now look into the insight of a problem using such a well tuned EKF to understand the numerical and statistical aspects of the parameter estimation . An adaptive tuning algorithm is developed and different case studies are carried out to prove the independence and robustness of the adaptive EKF algorithm for varying inputs.

Let us discuss how the chapters have been framed in the presented thesis work. Broadly the adaptive tuning problem tackled in the thesis can be divided into 3 sections-

- Adaptive EKF without process noise as discussed in chapter-4
- Adaptive EKF with low process noise(Q$<$0.1R) as discussed in chapter-6.1
- Adaptive EKF with high process noise(Q$\geq$0.1R) as discussed in chapter-6.2

The first chapter discussed about the various literature available in adaptive Kalman filtering technique along with their advantages and disadvantages. Chapter-2 introduces the reader to the basics of the Kalman filtering method and its extended version for non linear applications which is summarised at the end of the same chapter. Chapter-3 explains the maximum likelihood method used in estimation theory which can be solved using the well known Newton Raphson Optimisation procedure which forms our reference results. Chapter-4 is where the crux of our work is explained which arrives at an adaptive EKF solution that can achieve the CRB criterion which makes it an efficient estimator. The solution to the tuning problem is explained in an algorithmic form as well as using a flow chart. Chapter-5 lists out all the results obtained on conducting the case study as well as the sensitivity studies to check for the robustness and efficiency of the proposed adaptive EKF algorithm. Chapter-6 discuss the problem in presence of process noise which is divided into two sections , one that deals with low process noise case and the other with high process noise case. The chapter also introduces a new method of estimating the process noise. Appendix-A explains the creation of a Graphic User Interface (GUI) using MATLAB that will develop an user friendly environment to conduct sensitivity studies.

Chapter 2

Kalman Filtering method

In this chapter we will discuss the different Kalman filtering methods depending upon the type of system.

2.1 KF for Linear system

Consider the following linear discrete dynamic process represented by the linear difference equation

$$X(k) = F_{k,k-1}X(k-1) + B_k u_k + \omega(k) \tag{2.1}$$

The process noise samples $\omega(k) \sim$ N(0,Q) are assumed to be independent and identically distributed(iid)

$$E\left[\omega(k)\right] = 0 \ \& \ E\left[\omega(k)\omega(j)^T\right] = Q\delta(k-j)$$

where E[ ] is the expectation operator and δ is the Kronecker delta function.

$$\delta(k-j) = \begin{cases} 0 & \text{if } k \neq j; \\ 1 & \text{if } k = j. \end{cases}$$

The measurement equation in presence of additive noise at discrete time instant 'k' is given by

$$Z(k) = H_k X(k) + \upsilon(k) \tag{2.2}$$

The iid measurement noise samples $\upsilon(k) \sim$ N(0,R)and process noise samples $\omega(k) \sim$ N(0,Q) are assumed to be mutually independent

$$E\left[\omega(k)\upsilon(j)^T\right] = 0 \ \forall \ \text{k,j}$$

where

$$\begin{aligned}
X &: \text{State variable} \\
Z &: \text{measured data} \\
F &: \text{state transition matrix} \\
H &: \text{measurement matrix} \\
N &: \text{total no. of measurement samples} \\
k &: \text{discrete time instant and k=1,2...N} \\
B &: \text{Control input matrix} \\
u &: \text{Control input} \\
\upsilon &: \text{white Gaussian noise}, \mathcal{N}(0, R) \\
\omega &: \text{white Gaussian noise}, \mathcal{N}(0, Q)
\end{aligned}$$

The filtering solution to be above linear system was given by Kalman in his famous paper in the year 1960 which is well known as the Kalman filter divided into two steps (predict and update) consisting of five equations which are as follows

Predict :

$$\begin{aligned}
\hat{X}_{k|k-1} &= F_{k,k-1}\hat{X}_{k-1|k-1} + B_k u_k \\
P_{k|k-1} &= F_{k,k-1}P_{k-1|k-1}F_{k,k-1}^T + Q
\end{aligned}$$

Update :

$$\begin{aligned}
K_k &= P_{k|k-1}H_k^T(H_k P_{k|k-1}H_k^T + R)^{-1} \\
\hat{X}_{k|k} &= \hat{X}_{k|k-1} + K_k(Z(k) - H_k\hat{X}_{k|k-1}) \\
P_{k|k} &= (I - K_k H_k)P_{k|k-1}(I - K_k H_k)^T + K_k R K_k^T
\end{aligned}$$

The initial state estimate and the covariance matrix being

$$\hat{X}_0 = E[X_0]$$
$$P_0 = E[(X_0 - \hat{X}_0)(X_0 - \hat{X}_0)^T]$$

and

$$\hat{X}_{k|k-1} : \text{Prior state estimate}$$
$$\hat{X}_{k|k} : \text{Posterior state estimate}$$
$$P_{k|k-1} : \text{Prior state covariance matrix}$$
$$P_{k|k} : \text{Posterior state covariance matrix}$$
$$K_k : \text{Kalman Gain}$$

The KF needs the values of the noise statistics P_0 , Q and R. However in almost every application the noise statistics and the parameters governing the state trajectories are unknown and hence an adaptive procedure is required for filter convergence. If we want to make it an efficient and consistent estimator then tuning becomes very important in the optimal filter operation which is discussed in the later chapters. The next section deals with the non linearities in the system.

2.2 EKF for non-linear system

Kalman filer is optimal only for a linear system with additive independent white noise. However in practical scenario no system is linear and hence we need a filter that can handle non-linearities in the system and measurement equations. Extended Kalman Filter(EKF) is a non linear version of the standard Kalman Filter (KF) with the optimality being compromised. In this section we will see how one can use Taylor series approximation and convert a non linear system into its linear equivalent at a nominal point [2]. A discrete non-linear dynamic system with zero control input can be modelled as

$$X(k) = f(X(k-1)) + \omega(k) \tag{2.3}$$

The measurement equation is given by

$$Z(k) = h(X(k)) + \upsilon(k) \tag{2.4}$$

where

f : non linear function in the state space

h : non linear function in the measurement space

The characteristics of the measurement noise ,$\upsilon$and process noise ,ω remain the same and they are iid samples. By Taylor series expansion Eq-(2.3) reduces to

$$X(k) = f(X_n(k-1)) + f^{'}(X_n(k-1))\tilde{X}(k-1) + HOT + \omega(k)$$

Ignoring the higher order terms (HOT) , the 1st order Taylor series approximation is given by

$$X(k) = f(X_n(k-1)) + f^{'}(X_n(k-1))\tilde{X}(k-1) + \omega(k)$$

Let us denote $f^{'}(X_n(k-1))$ as $F_{k,k-1}$ and we get

$$X(k) = f(X_n(k-1)) + F_{k,k-1}\tilde{X}(k-1) + \omega(k) \tag{2.5}$$

where

$$F_{k,k-1} = f^{'}(X_n(k-1)) = \left[\frac{\partial f}{\partial X}\right]_{X=X_n(k-1)}$$
$$\tilde{X}(k-1) = X(k-1) - X_n(k-1)$$

$X_n(k)$ is the nominal point at discrete time instant 'k' and k=1,2...N where N is the total number of samples.

2.2.1 Predict step

Assuming that we have the apriori information of the initial estimate as

$$\hat{X}_0 = E[X_0]$$
$$\hat{P}_0 = E[(X_0 - \hat{X}_0)(X_0 - \hat{X}_0)^T]$$

We can predict the state at any discrete time instant 'k' given by

$$\hat{X}_{k|k-1} = E[X(k)|Z(k-1)]$$

Taking the nominal point as the previous state estimate in Eq-(2.5) we get

$$X(k) \approx f(\hat{X}_{k-1|k-1}) + F_{k,k-1}\tilde{X}(k-1) + \omega(k) \tag{2.6}$$

where $\tilde{X}(k-1) = X(k-1) - \hat{X}_{k-1|k-1}$ and the state Jacobian matrix in general is

$$F_{k,k-1} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_{n_s}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_{n_s}}{\partial x_1} & \frac{\partial f_{n_s}}{\partial x_2} & \cdots & \frac{\partial f_{n_s}}{\partial x_{n_s}} \end{bmatrix}_{|X=\hat{X}_{k-1|k-1}} \tag{2.7}$$

where n_s is the total number of states and the complete state vector X is of size $n_s \times 1$ and the individual states are augmented as

$$X = [x_1, x_2 ... x_{n_s}]^T$$

Thus we get

$$\hat{X}_{k|k-1} = E[f(\hat{X}_{k-1|k-1}) + F_{k,k-1}\tilde{X}(k-1) + \omega(k)|Z(k-1)]$$
$$\hat{X}_{k|k-1} = f(\hat{X}_{k-1|k-1}) + E[F_{k,k-1}\tilde{X}(k-1)|Z(k-1)] + E[\omega(k)|Z(k-1)]$$
$$\hat{X}_{k|k-1} = f(\hat{X}_{k-1|k-1}) \tag{2.8}$$

where

$$E[\omega(k)|Z(k-1)] = 0$$
$$E[\tilde{X}(k-1)|Z(k-1)] = 0$$

The corresponding covariance matrix is given by

$$P_{k|k-1} = E\left[\{X(k) - \hat{X}_{k|k-1}\}\{X(k) - \hat{X}_{k|k-1}\}^T\right] \tag{2.9}$$

Using Eq-(2.5) and Eq-(2.9)

$$\begin{aligned} P_{k|k-1} &= E\left[\{f(\hat{X}_{k-1|k-1}) + F_{k,k-1}\tilde{X}(k-1) + \omega(k) - f(\hat{X}_{k-1|k-1})\}\ \{\}^T\right] \\ &= E\left[\{F_{k,k-1}\tilde{X}(k-1) + \omega(k)\}\ \{\}^T\right] \end{aligned}$$

We know that

$$\begin{aligned} &E\left[\tilde{X}(k-1)\tilde{X}(k-1)^T\right] = P_{k-1|k-1} \\ &E\left[\tilde{X}(k-1)\omega(k)^T\right] = 0 \\ &E\left[\omega(k)\omega(k)^T\right] = Q \end{aligned}$$

The prior covariance propagation equation is given by

$$P_{k|k-1} = F_{k,k-1}P_{k-1|k-1}F_{k,k-1}^T + Q \tag{2.10}$$

where $F_{k,k-1}$ is now the state Jacobian matrix.

2.2.2 Update step

This step is also known as the data assimilation step or measurement update step since it fuses the prior state estimate and the measured data through the optimal Kalman gain to get the updated state estimate. An unbiased optimal estimate in a least square sense is given by

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k(Z(k) - E[h(X(k)|Z(k)]) \tag{2.11}$$

where K_k is the optimal Kalman gain which will be derived at a later stage of the chapter. From Taylor series expansion at the nominal point being the predicted state we get

$$h(X(k) = h(X_{k|k-1}) + H_k(X(k) - \hat{X}_{k|k-1}) + HOT$$

By ignoring the higher order terms(HOT) we get

$$h(X(k) \approx h(X_{k|k-1}) + H_k(X(k) - \hat{X}_{k|k-1}) \tag{2.12}$$

where H_k is the measurement Jacobian matrix which in general can be written as

$$H_k = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \cdots & \frac{\partial h_1}{\partial x_{n_s}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_{n_m}}{\partial x_1} & \frac{\partial h_{n_m}}{\partial x_2} & \cdots & \frac{\partial h_{n_m}}{\partial x_{n_s}} \end{bmatrix}_{|X=\hat{X}_{k|k-1}}$$

where n_m is the total number of measured states and the complete state vector X is of size $n_s \times 1$ and the individual states are augmented as

$$X = [x_1, x_2 ... x_{n_s}]^T$$

Consider the term $E[h(X(k)|Z(k)]$ which can now be approximated as

$$E[h(X(k)|Z(k)] \approx h(X_{k|k-1}) + E[H_k(X(k) - \hat{X}_{k|k-1})|Z(k)])$$
$$E[h(X(k)|Z(k)] \approx h(X_{k|k-1})$$

where $E[H_k(X(k) - \hat{X}_{k|k-1})|Z(k)]$=0. Using the above approximation in Eq-(2.11) we get

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k(Z(k) - h(X_{k|k-1})) \tag{2.13}$$

Let us now get the corresponding error covariance matrix. The state error is defined as

$$\begin{aligned}
\tilde{X}(k) &= X(k) - \hat{X}_{k|k} \\
&= f(X(k-1)) + \omega(k) - \hat{X}_{k|k-1} - K_k(Z(k) - h(X_{k|k-1}) \\
&= f(X(k-1)) - \hat{X}_{k|k-1} + \omega(k) - K_k(h(X(k) - h(X_{k|k-1} + \upsilon(k))
\end{aligned}$$

Using Eq-(2.6) and Eq-(2.12) we get

$$\begin{aligned}
\tilde{X}(k) &\approx F_{k,k-1}\tilde{X}(k-1) + \omega(k) - K_k(H_k(X(k) - \hat{X}_{k|k-1}) + \upsilon(k)) \\
&\approx F_{k,k-1}\tilde{X}(k-1) + \omega(k) - K_kH_k(F_{k,k-1}\tilde{X}_{k|k-1} + \omega(k)) - K_k\upsilon(k) \\
&\approx (I - K_kH_k)F_{k,k-1}\tilde{X}(k-1) + (I - K_kH_k)\omega(k) - K_k\upsilon(k)
\end{aligned}$$

The posterior covariance equation is given by

$$\begin{aligned}
P_{k|k} &= E[\tilde{X}(k)\tilde{X}(k)^T] \\
P_{k|k} &= (I - K_kH_k)F_{k,k-1}P_{k-1|k-1}F_{k,k-1}^T(I - K_kH_k)^T + (I - K_kH_k)Q(I - K_kH_k)^T + K_kRK_k^T
\end{aligned}$$

Using Eq-(2.10) we get

$$\begin{aligned}
P_{k|k} &= (I - K_kH_k)P_{k|k-1}(I - K_kH_k)^T + K_kRK_k^T \\
P_{k|k} &= P_{k|k-1} - K_kH_kP_{k|k-1} - P_{k|k-1}H_k^TK_k^T + K_kH_kP_{k|k-1H^TK_k^T} + K_kRK_k^T \quad (2.14)
\end{aligned}$$

An optimal Kalman gain K_k is obtained by minimizing the trace of the posterior covariance matrix $P_{k|k}$,i.e

$$\frac{\partial tr(P_{k|k})}{\partial K_k} = 0$$

Taking the partial derivatives of Eq-(2.14) with respect to the gain we get

$$\begin{aligned}
0 &= -(H_kP_{k|k-1})^T - P_{k|k-1}H_k^T + 2K_kH_kP_{k|k-1}H_k^T + 2K_kR \\
0 &= -2P_{k|k-1}H_k^T + 2K_k(H_kP_{k|k-1}H_k^T + R) \\
K_k &= P_{k|k-1}H_k^T(H_kP_{k|k-1}H_k^T + R)^{-1} \quad (2.15)
\end{aligned}$$

Let us now reduce the computational complexity of the posterior covariance by substituting Eq-(2.15) in Eq-(2.14)

$$\begin{aligned}
P_{k|k} &= (I - K_k H_k) P_{k|k-1} - (I - K_k H_k) P_{k|k-1} H_k^T K_k^T + K_k R K_k^T \\
P_{k|k} &= (I - K_k H_k) P_{k|k-1} - (P_{k|k-1} H_k^T - K_k H_k P_{k|k-1} H_k^T - K_k R) K_k^T \\
P_{k|k} &= (I - K_k H_k) P_{k|k-1} - (P_{k|k-1} H_k^T - K_k (H_k P_{k|k-1} H_k^T + R)) K_k^T \\
P_{k|k} &= (I - K_k H_k) P_{k|k-1} - (P_{k|k-1} H_k^T - P_{k|k-1} H_k^T) K_k^T \\
P_{k|k} &= (I - K_k H_k) P_{k|k-1}
\end{aligned} \tag{2.16}$$

We can summarise the steps involved in running the recursive EKF using Eq-(2.3), Eq-(2.4), Eq-(2.8), Eq-(2.10), Eq-(2.13), Eq-(2.14), Eq-(2.16) which is formulated in the next section.The idea of the extended Kalman filter (EKF) was originally proposed by Stanley Schmidt so that the Kalman filter could be applied to nonlinear spacecraft navigation and GPS problems. The linearized equations were derived based on first order Taylor series approximation where the higher order terms (HOT) are assumed to be negligible and ignored when there is considerably less non-linearity in the system. If there is too much non-linearity in the system then one has to go with higher order Taylor series approximation or more appropriately the particle filtering method based on Monte Carlo simulations which are computationally expensive.

The diagonal elements of the covariance matrix should indicated the uncertainty in the state estimates if right values of P_0,Q,R is chosen. Even though EKF is suboptimal due to linearisation errors , it is the de facto standard in the theory of non-linear state estimation and can be made as an efficient and consistent estimator if Cramer Rao criterion is achieved by the diagonal elements of the covariance matrix. This can be achieved by adaptive tuning of the filter which is discussed in the later chapters and forms the crux of our work.

2.2.3 EKF Summary

The system equations and recursive EKF can be summarised as follows

System & Observation Model :

$$X(k) = f(X(k-1)) + \omega(k)$$
$$Z(k) = h(X(k)) + \upsilon(k)$$

Initialisation :

$$\hat{X}_0 = E[X_0]$$
$$P_0 = E[(X_0 - \hat{X}_0)(X_0 - \hat{X}_0)^T]$$

Predict step :

$$\hat{X}_{k|k-1} = f(\hat{X}_{k-1|k-1})$$
$$P_{k|k-1} = \mathbf{F}_{k,k-1} P_{k-1|k-1} \mathbf{F}^T_{k,k-1} + Q$$

Update step :

$$K_k = P_{k|k-1} \mathbf{H}_k^T (\mathbf{H}_k P_{k|k-1} \mathbf{H}_k^T + R)^{-1}$$
$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + K_k (Z(k) - h(\hat{X}_{k|k-1}))$$
$$P_{k|k} = (I - K_k \mathbf{H}_k) P_{k|k-1}$$

where

$\hat{X}_{k|k-1}$: Prior state estimate

$\hat{X}_{k|k}$: Posterior state estimate

$P_{k|k-1}$: Prior state covariance matrix

$P_{k|k}$: Posterior state covariance matrix

K_k : Optimal Kalman Gain

$\mathbf{F}_{k,k-1}$: State Jacobian matrix

$\mathbf{H}_k$: Measurement Jacobian matrix

Chapter 3

Maximum likelihood estimation

In this chapter we will discuss the method of maximum likelihood estimation(MMLE) using the well-known Newton Raphson optimisation scheme which forms the reference results to check for efficiency of the adaptive EKF. We will also discuss about the Cramer Rao bound (CRB) criterion used in estimation theory (ET) and statistics. Any estimator is said to be efficient and MVU(Minimum variance and unbiased) which satisfies the Cramer Rao bound. It also sets a bench mark to evaluate the performance of the given estimator.

3.1 Newton Raphson Optimisation

Method of maximum likelihood estimation (MMLE) is one of the fundamental method in estimation theory which was introduced by Fisher in the year 1922. The likelihood function(L) of the measurement set(Z) for a given parameter θ can be written as

$$L(Z|\theta) = p(Z(1), Z(2), ..Z(N)|\theta)$$

Assuming that the measurement noise is white and Gaussian the likelihood function of a multivariate RV is given by

$$\begin{aligned} L(Z|\theta) = & \frac{1}{\sqrt{(2\pi)^N |R|}} exp(-1/2) \times \\ & \sum_{k=1}^{N} (Z_\theta(k) - Z(k))^T R^{-1} (Z_\theta(k) - Z(k)) \end{aligned}$$

where Z_θ is the predicted measurement for a given parameter θ.
The log-likelihood function is given by

$$log(L) = log\left[\frac{1}{\sqrt{(2\pi)^N|R|}}\right] - \frac{1}{2}\sum_{k=1}^{N}(Z_\theta(k) - Z(k))^T R^{-1}(Z_\theta(k) - Z(k))$$

We can maximize the above log-likelihood function using

$$\frac{\partial log(L)}{\partial \theta} = 0$$

Thus the problem reduces to a weighted least squares where the cost function(J) is given by

$$J(\theta) = \sum_{k=1}^{N}(Z_{\hat{\theta}}(k) - Z(k))^T R^{-1}(Z_{\hat{\theta}}(k) - Z(k))$$

The above function is minimised using the well known Newton Raphson iterative procedure [6] as follows

$$\hat{\theta}^{i+1} = \hat{\theta}^i - (D^2J)^{-1}DJ \tag{3.1}$$

where 'i' is the iteration number , D and D^2 represent the first and second order gradients of the function $J(\theta)$ with respect to θ given by

$$DJ = \sum_{k=1}^{N}(Z_{\hat{\theta}}(k) - Z(k))R^{-1}D(Z_{\hat{\theta}}(k))^T$$
$$D^2J \approx \sum_{k=1}^{N}D(Z_{\hat{\theta}}(k))R^{-1}D(Z_{\hat{\theta}}(k))^T$$

The second gradient being only an approximation approaches the correct value as the optimum value of θ is reached. The unknown R can be estimated as

$$\hat{R} = \frac{1}{N}\sum_{k=1}^{N}(Z_{\hat{\theta}}(k) - Z(k))(Z_{\hat{\theta}}(k) - Z(k))^T \tag{3.2}$$

The CRB or uncertainty in the parameter is given by

$$P_{\hat{\theta}} = diag[(D^2J)^{-1}] \tag{3.3}$$

The correlation co-efficient matrix element of i^{th} row and j^{th} column is given by

$$\rho_{ij} = \frac{d_{ij}}{\sqrt{d_{ii}d_{jj}}} \tag{3.4}$$

where d_{ij} is the i^{th} row and j^{th} column element of the second gradient matrix (D^2J).The above results obtained from the Newton Raphson optimisation procedure is used as reference results to compare with the EKF results.A brief procedure of the NR optimisation scheme is summarised below

- Start with guess values of R and θ
- Update θ using Eq-(3.1)
- Estimate R using Eq-(3.2)
- Repeat the above two steps till equillibrium or maximum iteration is reached.
- Save the values of $\hat{\theta}$, CRB , R and ρ for comparison with EKF results.

3.2 Simulated Reference Results

The MMLE using NR optimisation procedure was tested on a constant system. The system modelling and equations are framed in chapter-5. The simulated results are as follows

	theta	**%CRB***	$\hat{R}$
NR	0.9999	0.0038	0.0484

$* \%CRB = \frac{\sqrt{P_{\hat{\theta}}}}{\theta} \times 100$

The estimated parameters and histogram results of $\hat{\theta}$ and $\hat{R}$ with true values being θ=1 and R=0.05 for 100 ensembles and 10 iterations are displayed below

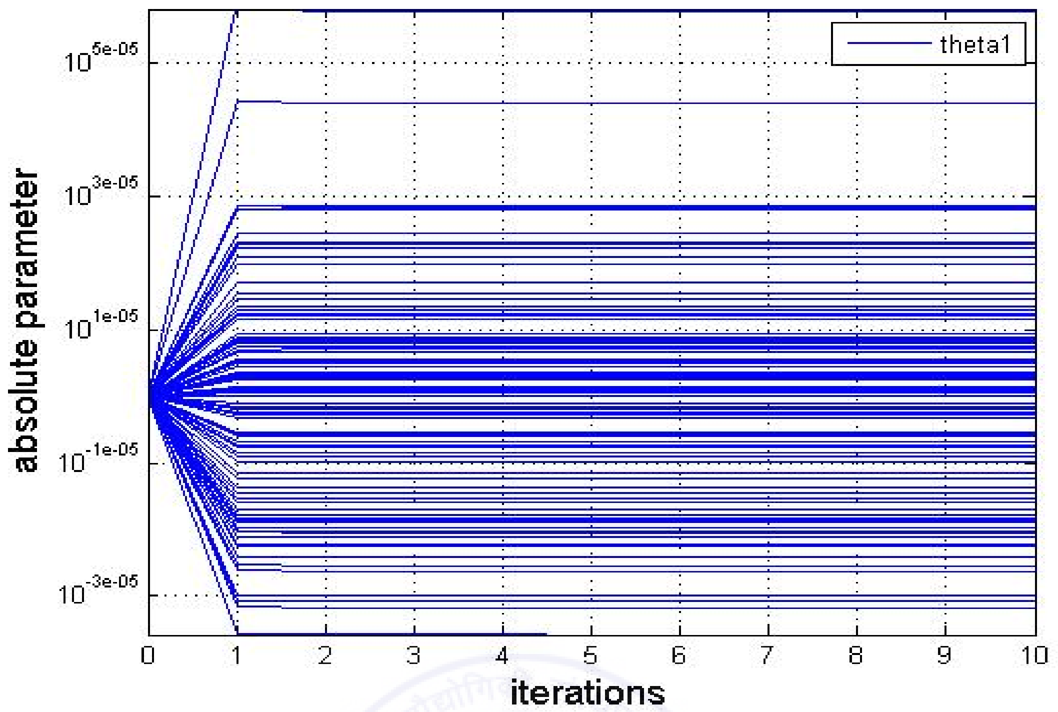


Figure 3.1: NR-Constant system, $\hat{\theta}$ vs iterations for 100 ensembles

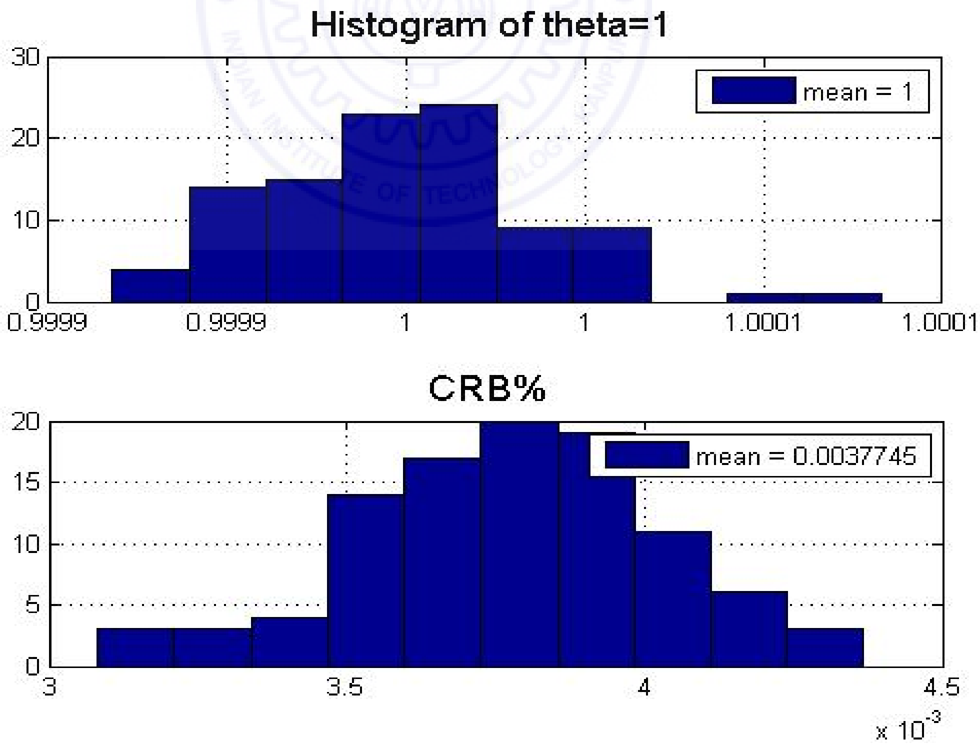


Figure 3.2: NR-Constant system, Histogram of θ=1 and its CRB

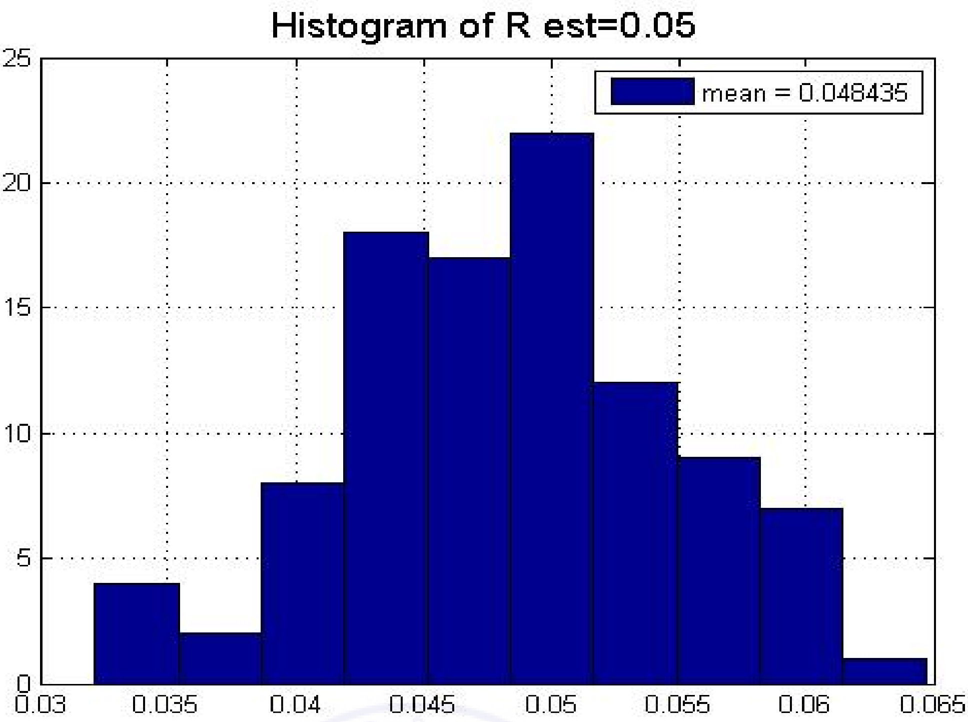


Figure 3.3: NR-Constant system, Histogram of R

The MMLE using NR optimisation procedure was also tested on a Spring Mass Damper(SMD) system with weak non-linearity. The system modelling and equations are framed in chapter-5. The simulated results are as follows

Table 3.1

	theta	**%CRB***	$\hat{R}$
NR	3.9969 , 0.4006 , 0.6074	1.8779 , 3.2293 , 33.2294	0.0102 , 0.0443

$$* \ \%CRB = \frac{\sqrt{P_{\hat{\theta}}}}{\theta} \times 100$$

The correlation coefficient matrix of the estimated parameters is

$$\rho = \begin{bmatrix} 1.0000 & -0.2041 & -0.9063 \\ -0.2041 & 1.0000 & 0.4011 \\ -0.9063 & 0.4011 & 1.0000 \end{bmatrix}$$

The histogram results of $\hat{\theta}$ and $\hat{R}$ with true values being θ=[4,0.4,0.6] and R=diag[0.01,0.04] for 100 ensembles and 10 iterations are displayed below

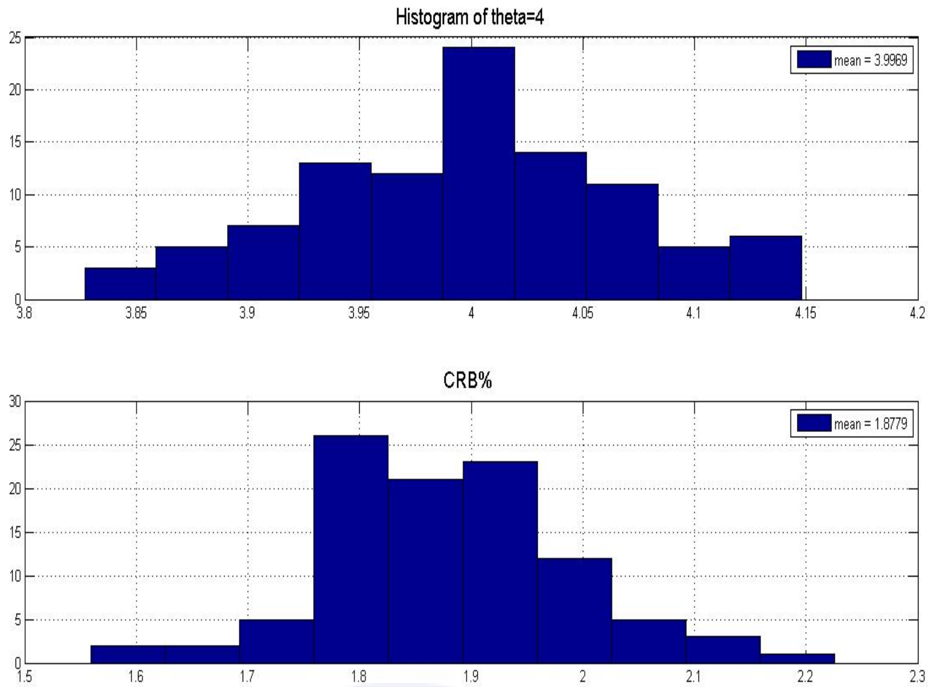


Figure 3.4: NR-SMD system, Histogram of θ=4 and its CRB

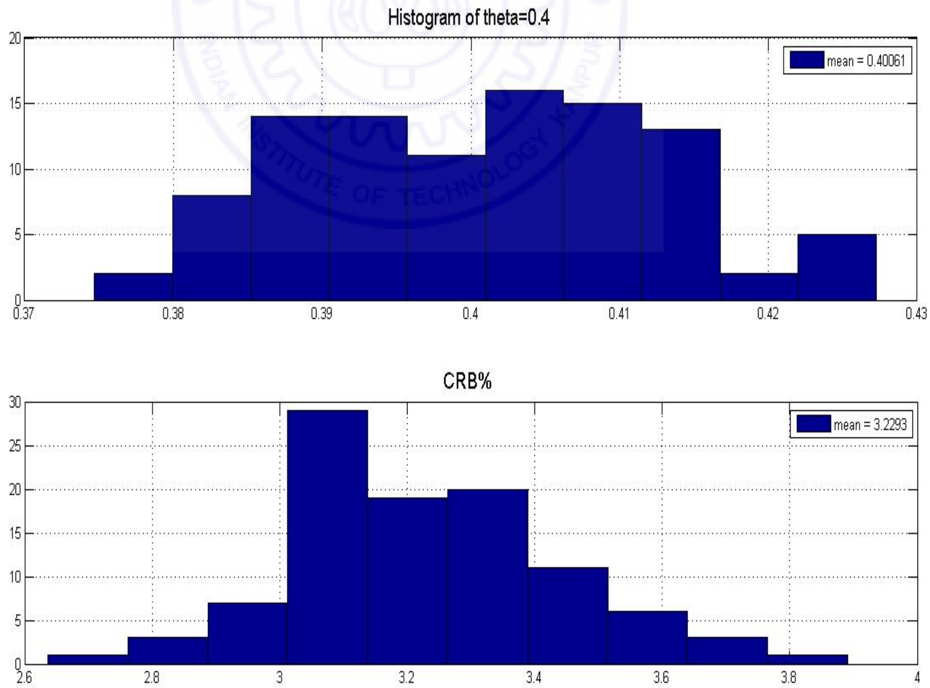


Figure 3.5: NR-SMD system, Histogram of θ=0.4 and its CRB

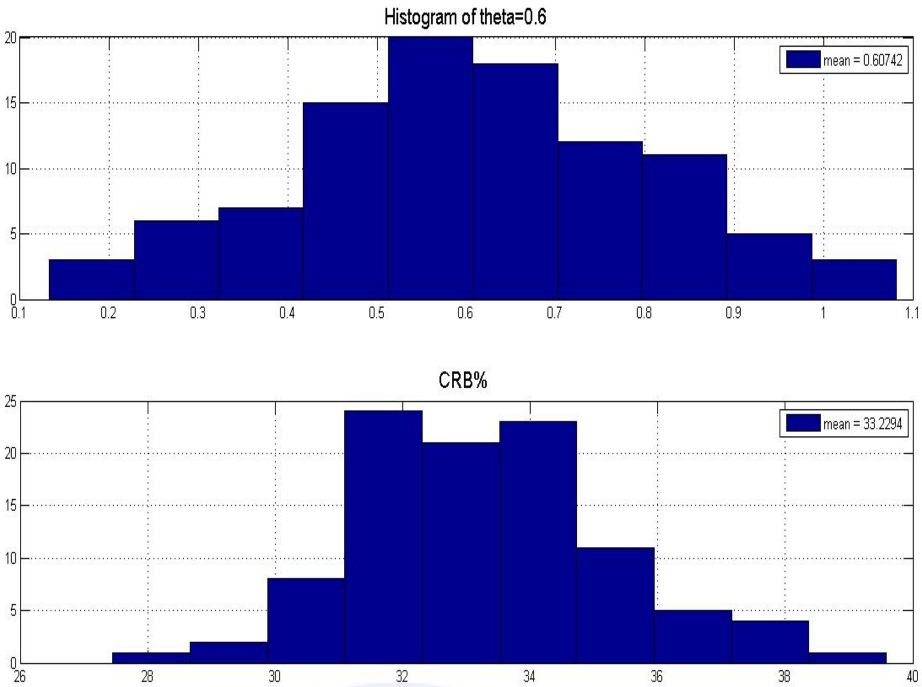


Figure 3.6: NR-SMD system, Histogram of θ=0.6 and its CRB

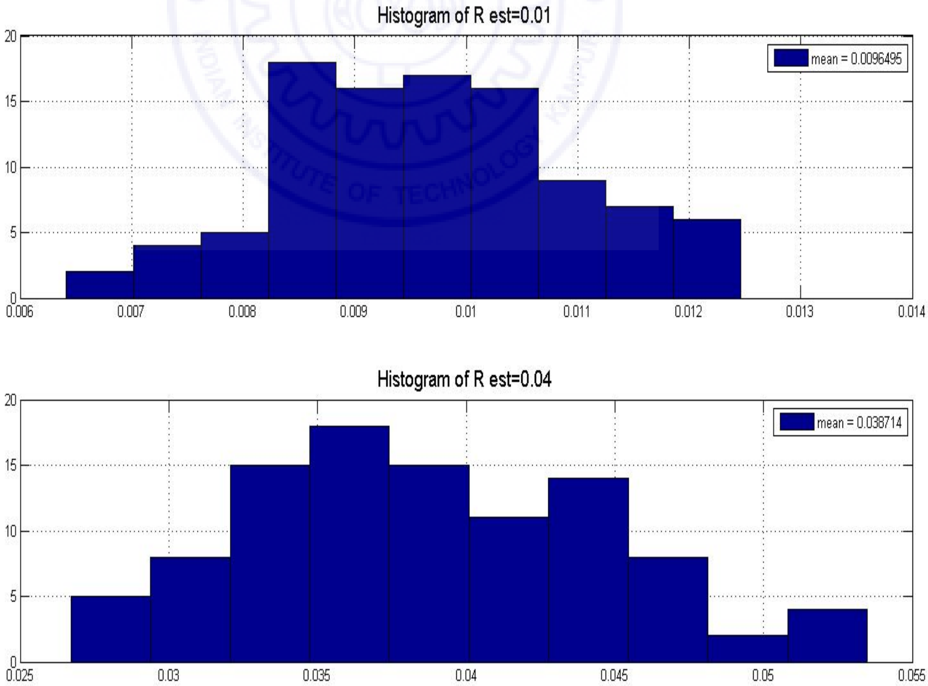


Figure 3.7: NR-SMD system, Histogram of R

Chapter 4

EKF tuning without process noise

What is filter tuning ? If the filter has to give a stable and consistent output then one need to know the process and measurement noise statistics as well as the initial estimate and covariance of the states and parameters. The setting of these typically apriori-unknown quantities is called filter tuning. In this chapter we will discuss the adaptive tuning aspects of EKF with zero process noise for achieving Cramer Rao bounds in a parameter estimation problem which is the main contribution of our work. Rigorous time consuming optimisation techniques and ad-hoc procedures used in adaptive filtering have always lead to improper and inadequate results.The best way to solve the tuning problem is to use heuristic procedures which gives encouraging results.

4.1 Augmented System

The unknown parameters of the system can be augmented as additional states [3] using a constant system model as

System model

$$X(k) = \begin{bmatrix} x(k) \\ \theta(k) \end{bmatrix} = \begin{bmatrix} f(x(k-1), \theta) \\ \theta(k-1) \end{bmatrix} \tag{4.1}$$

Observation model

$$Z(k) = h(x(k), \theta) + v(k) \tag{4.2}$$

where

x(k) is the state vector of size $n_s \times 1$

$\theta(k)$ is the augmented state vector or unknown parameter vector of size $n_p \times 1$

X(k) is the complete state vector of size $n_s + n_p \times 1$

Z(k) is the measurement vector of size $n_m \times 1$

4.2 Estimation of measurement noise

The measurement covariance matrix R is considered as one of the easiest noise statistic that can be estimated. A simple heuristic way of estimating R was suggested by Myers and Tapley in the mid 1970's based on covariance matching technique as discussed below. Consider the intuitive estimates of measurement noise samples $\hat{r}(k) = \mathcal{N} \sim (\bar{r}, C_r)$ given by

$$\hat{r}(k) = Z(k) - H_k \hat{X}_{k|k-1} \tag{4.3}$$

Sample mean of these noise samples is given by

$$\bar{r} = \frac{1}{N} \sum_{k=1}^{N} \hat{r}(k)$$

Sample covariance of these noise samples is given by

$$C_r = \frac{1}{N-1} \sum_{k=1}^{N} \{\hat{r}(k) - \bar{r}\}\{\hat{r}(k) - \bar{r}\}^T \tag{4.4}$$

Using Eq-(4.3)

$$\begin{aligned}
cov\{\hat{r}(k)\} &= cov\{Z(k) - H_k \hat{X}_{k|k-1}\} \\
&= cov\{H_k X(k) + \upsilon(k) - H_k \hat{X}_{k|k-1}\} \\
&= cov\{H_k (X(k) - \hat{X}_{k|k-1}) + \upsilon(k)\} \\
&= cov\{H_k (X(k) - \hat{X}_{k|k-1})\} + cov\{\upsilon(k)\}
\end{aligned}$$

where

$$cov\{(X(k) - \hat{X}_{k|k-1})\upsilon(k)\} = 0$$

cov is the covariance operator given by

$$cov(X) = E\left[(X - E[X])(X - E[X])^T\right]$$

Thus we get

$$
\begin{aligned}
cov\{\hat{r}(k)\} &= H_k cov\{(X(k) - \hat{X}_{k|k-1})\} H_k^T + cov\{\upsilon(k)\} \\
cov\{\hat{r}(k)\} &= H_k P_{k|k-1} H_k^T + R \\
E[cov\{\hat{r}(k)\}] &= \frac{1}{N}\sum_{k=1}^{N} H_k P_{k|k-1} H_k^T + R \qquad (4.5)
\end{aligned}
$$

Equating Eq-(4.4) and Eq-(4.5) we get an estimate of R given by

$$
\begin{aligned}
\frac{1}{N}\sum_{k=1}^{N} H_k P_{k|k-1} H_k^T + \hat{R} &= \frac{1}{N-1}\sum_{k=1}^{N}\{\hat{r}(k) - \bar{r}\}\{\hat{r}(k) - \bar{r}\}^T \\
\hat{R} &= \frac{1}{N-1}\sum_{k=1}^{N}\left\{(r_k - \bar{r})(r_k - \bar{r})^T - \frac{N-1}{N} H_k P_k H_k^T\right\} \qquad (4.6)
\end{aligned}
$$

where r(k) are the measurement noise samples and $\bar{r}$ is its sample mean.The estimated measurement noise samples $r(k)$ and $P(k)$ can be any of the three different values mentioned below

- Using innovations where $r(k) = Z(k) - h(\hat{X}_{k|k-1})$ and $P_k = P_{k|k-1}.\hat{X}_{k|k-1}$ and $P_{k|k-1}$ is the prior estimate and prior covariance respectively.
- Using residue where $r(k) = Z(k) - h(\hat{X}_{k|k})$ and $P_k = P_{k|k}.\hat{X}_{k|k}$ and $P_{k|k}$ is the posterior estimate and posterior covariance respectively.
- Using smoothed data where $r(k) = Z(k) - h(\hat{X}_{k|N})$ and $P_k = P_{k|N}$ where $\hat{X}_{k|N}$ and $P_{k|N}$ is the smoothed estimate and smoothed covariance respectively.

4.3 Estimation of X_0 and P_0 for states

In order to get the tuning process right the initial state and the corresponding covariance matrix needs to be estimated correctly. The importance of the initial state covariance matrix in achieving Cramer-Rao bound was first shown by Gemson and M R Ananthasayanam [5]. However achieving the CRB's of the parameter estimates was still pending. Hence we go for smoothing technique which can be categorized as

- Fixed point smoothing which estimates the state at a fixed point of past time instant.
- Fixed lag smoothing which estimates the state at a fixed delay in the past time instant.
- Fixed interval smoothing which assumes that all measurement sets are readily available on the same time interval.

Let us now find out the basic difference between Prediction , filtering and smoothing methods.

- Prediction at a time instant 'k' involves estimating the apriori state with the given measurements upto and not including 'k^{th}' time instant which can be represented as $\hat{X}_{k|k-1} = E[X_k|Z(1), Z(2)...Z(k-1)]$.
- A Filter estimates the posterior state based on all measurements upto and including time instant 'k' which is denoted as $\hat{X}_{k|k} = E[X_k|Z(1), Z(2)...Z(k)]$.
- A smoother estimates the state based on the past as well as future measurements say upto time instant 'N' which can be represented as $\hat{X}_{k|N} = E[X_k|Z(1), Z(2)...Z(N)]$.

In the next sections we will discuss two different fixed interval optimal smoothing techniques, RTS smoothing which was suggested by Rauch, Tung and Striebel(RTS) and Modified Bryson Frazier(MBF) smoothing which was suggested by Bryson and Frazier.

4.3.1 RTS Smoothing

RTS smoother [8] is an optimal two pass smoothing technique which runs backwards in time for k=N-1,N-2...0. The recursive RTS smoothing algorithm is explained below

$$
\begin{aligned}
K_s(k) &= P_{k|k} F_{k,k-1} P_{k+1|k}^{-1} \\
\hat{X}_{k|N} &= \hat{X}_{k|k} + K_s(k)[\hat{X}_{k+1|N} - \hat{X}_{k+1|k}] \\
P_{k|N} &= P_{k|k} + K_s(k)[P_{k+1|N} - P_{k+1|k}]K_s(k)^T
\end{aligned}
$$

where

$F_{k,k-1}$ is the state Jacobian matrix

K_s is the smoothed gain

$\hat{X}_{k|N}$ is the smoothed state estimate

$P_{k|N}$ is the smoothed covariance matrix

4.3.2 MBF Smoothing

The RTS smoothing can be sometimes computationally complex because of the covariance matrix inverse used in the smoothing equations. A computationally less complex smoothing algorithm was suggested by Bryson and Frazier [9] which uses the Kalman Gains from the forward filter. The MBF smoother is initialised as follows

$$
\begin{aligned}
\hat{\Lambda}_N &= 0 \\
\hat{\lambda}_N &= 0
\end{aligned}
$$

The recursive MBF smoothing technique which runs backwards in time for k=N-1,N-2...0 is explained below

$$
\begin{aligned}
\tilde{\Lambda}_k &= H_k^T S_k^{-1} H_k + (I - K_k H_k)^T \hat{\Lambda}_k (I - K_k H_k) \\
\hat{\Lambda}_{k-1} &= F_{k,k-1} \tilde{\Lambda}_k F_{k-1,k}^T \\
\tilde{\lambda}_k &= -H_k^T S_k^{-1} I_k + (I - K_k H_k)^T \hat{\lambda}_k \\
\hat{\lambda}_{k-1} &= F_{k,k-1}^T \tilde{\lambda}_k
\end{aligned}
$$

$$P_{k|N} = P_{k|k} - P_{k|k}\hat{\Lambda}_k P_{k|k}$$
$$\hat{X}_{k|N} = \hat{X}_{k|k} - P_{k|k}\hat{\lambda}_k$$
or
$$P_{k|N} = P_{k|k-1} - P_{k|k-1}\tilde{\Lambda}_k P_{k|k-1}$$
$$\hat{X}_{k|N} = \hat{X}_{k|k-1} - P_{k|k-1}\tilde{\lambda}_k$$

where

$\hat{X}_{k|N}$ is the smoothed state estimate

$P_{k|N}$ is the smoothed covariance matrix

$I_k = Z(k) - H_k\hat{X}_{k|k-1}$ is the innovation sequence

$S_k = H_k P_{k|k-1} H_k^T + R$ is the innovation covariance matrix

$F_{k,k-1}$ is the state Jacobian matrix

K_k is the forward Kalman gain

Thus we get

$$\left.\begin{aligned} \hat{x}_0 &= \hat{x}_{0|N} \\ P_{x_0} &= P_{x_0|N} \end{aligned}\right\} \tag{4.7}$$

where $\hat{x}_{0|N}$ and $P_{x_0|N}$ are the initial smoothed state (RTS or MBF).

4.4 Estimation of X_0 and P_0 for parameters

The initial state covariance matrix(P_0) for states was estimated by smoothing technique.In this section let us estimate the initial state covariance matrix(P_0) for the parameters using virtual measurement concept. Till today the tuning of the parameter covariance marix is done manually by trial methods and generally it is taken as unity without any mathematical basis. The unknown parameters(θ)are augmented as states using the constant signal model $\dot{\theta} = \theta$ or $\theta(k) = \theta(k-1)$ as seen in Eq-(4.1). Our aim is to see what would have happened if we could observe or measure the unknown parameters of the system and what can we do if we do not have them. Since parameters cannot be measured let us assume a virtual measurement(Z_v) of the parameters represented by a virtual observation model given by

$$Z_v(k) = \theta + \vartheta_v(k) \tag{4.8}$$

where the subscript 'v' indicates a virtual noise. The vectors Z_v , θ and ϑ_v are vector of size $n_p \times 1$ where n_p is the number of unknown parameters. Virtual noise samples ϑ_v(k)$\sim$N(0,R_v) for all time instants k=1,2,...N are assumed to be zero mean white Gaussian noise described as

$$E\left[\vartheta_v(k)\right] = 0$$
$$E\left[\vartheta_v(k)\vartheta_v(j)^T\right] = R_v\delta(k-j)$$

where δ is the Kronecker delta function.

$$\delta(k-j) = \begin{cases} 0 & \text{if } k \neq j; \\ 1 & \text{if } k = j. \end{cases}$$

Least Square estimate of the unknown parameter θ is given by the sample mean of the virtual measurements

$$\hat{\theta} = \frac{1}{N}\sum_{k=1}^{N} Z_v(k)$$

The corresponding covariance of the estimate [4] is given by

$$\begin{aligned} P_{\hat{\theta}} &= E\left[\{\hat{\theta}-\theta\}\{\hat{\theta}-\theta\}^T\right] \\ &= E\left[\left[\frac{1}{N}\sum_{k=1}^{N} Z_v(k) - \theta\right]\left[\frac{1}{N}\sum_{k=1}^{N} Z_v(k) - \theta\right]^T\right] \\ &= E\left[\left[\frac{1}{N}\sum_{k=1}^{N}\{Z_v(k) - \theta\}\right]\left[\frac{1}{N}\sum_{k=1}^{N}\{Z_v(k) - \theta\}\right]^T\right] \end{aligned}$$

$$
\begin{aligned}
P_{\hat{\theta}} &= E\left[\left[\frac{1}{N}\sum_{k=1}^{N}\{\vartheta_v(k)\}\right]\left[\frac{1}{N}\sum_{k=1}^{N}\{\vartheta_v(k)\right]^T\right] \\
&= \frac{1}{N^2}\left[\sum_{k=1}^{N} E\{\vartheta_v(k)\vartheta_v(k)^T\}\right] \\
&= \frac{1}{N^2}\left[\sum_{k=1}^{N} R_v\right] \\
&= \frac{R_v}{N}
\end{aligned}
$$

Thus we get

$$
R_v = N \times P_{\hat{\theta}} \tag{4.9}
$$

The best possible initial estimates for such a system are given by

$$
\begin{aligned}
\hat{\theta}_0 &= Z_v(k) \text{ for any k=1,2,..N} \\
P_{\theta_0} &= R_v
\end{aligned} \tag{4.10}
$$

Since we do not have the value of R_v which is virtual , an estimate of the initial state covariance matrix is given by

$$
P_{\theta_0} = N \times P_{\hat{\theta}} \text{ ; from Eq-(4.9),(4.10)}
$$

Hence we can scale up the covariance estimated by the Kalman filter at the end of the iteration by a factor of N and use it as the initial state covariance matrix for the next iteration.The strategy is to use the available information after a pass or iteration for estimating the parameters and noise statistics and use the estimated values in the next pass iteration. Since we do not have the virtual measurement(Z_v) , the initial state estimate($\hat{\theta}_0$) is taken as the parameter estimated by the Kalman filter at the end of the previous iteration ($\hat{\theta}_{N|N}$). Thus at any given iteration 'i'

$$
\left.
\begin{aligned}
\hat{\theta}_0^i &= \hat{\theta}_{N|N}^{i-1} \\
P_{\theta_0}^i &= N \times P_{\hat{\theta}}^{i-1}
\end{aligned}
\right\} \tag{4.11}
$$

Combining the results from the above two sections for our augmented model, at any iteration 'i' the estimated X_0 and P_0 is given by

$$\left.\begin{aligned} \hat{X}_0^i &= [\hat{x}_0^i, \hat{\theta}_0^i] \\ P_0^i &= \mathrm{diag}[P_{x_0}^i, P_{\theta_0}^i] \end{aligned}\right\} \tag{4.12}$$

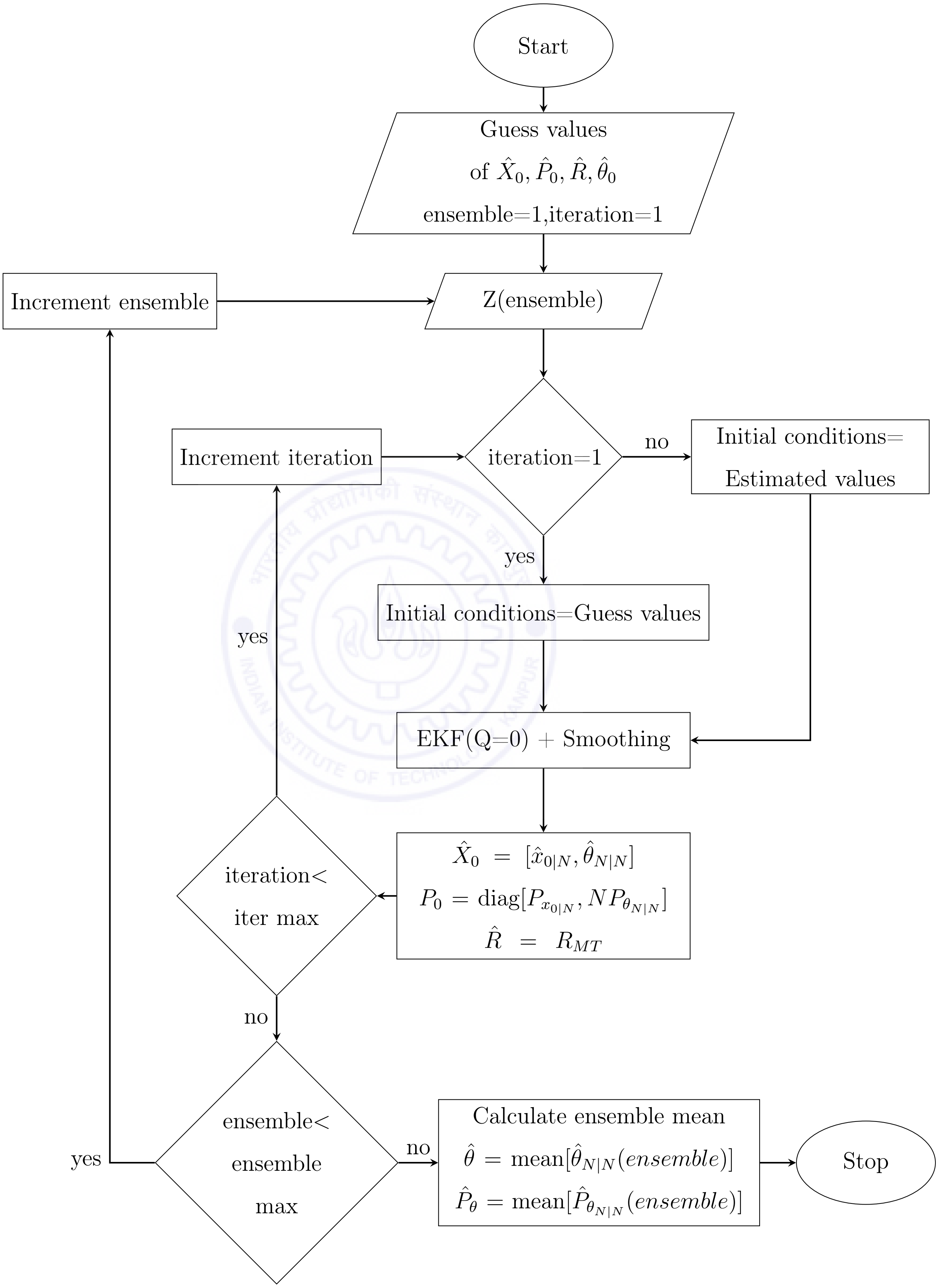
The above iterative tuning strategy would require data processing through a scouting pass or iteration where the first iteration is carried out with guess values of X_0 and P_0.The next iterations can be ran with the estimated values at the end of the previous iteration as per the Eq-(4.12).If we know the true initial states then we can use $\hat{x}_0 = X_0$ and P_{x_0}=0 in which case the best results are obtained.

4.5 EKF tuning algorithm

In this section an adaptive tuning algorithm for parameter estimation problem is discussed assuming that there is no process model mismatch (Q=0) and the unknowns are X_0, P_0, R and θ.

1. Given the model and the measurements the EKF scouting pass can be run with guess values of X_0, P_0, R and θ
2. The measurement noise covariance matrix (R) can be estimated by a simple heuristic procedure suggested by Myers and Tapley (MT)
3. The initial state and its covariance matrix for the actual states can be estimated by Modified Bryson Frazier (MBF) or Rauch Tung Striebel (RTS) smoothing technique Eq-(4.7).
4. The initial state and covariance matrix for the augmented states (unknown parameters) can be estimated using the virtual measurement concept as per Eq-(4.11).
5. The above steps(1-4) is be repeated for a pre-defined maximum number of iteration till statistical equilibrium is reached when the results do not change.
6. Ensemble runs using different measurement data is carried out by repeating the above steps(1-5) as seen in the below flow chart.

Adaptive EKF flowchart for Q=0



Chapter 5

Results for zero process noise case

In this chapter different case studies are conducted using the proposed algorithm and its robustness is tested with different sensitivity studies. The case study include a simple constant signal system and involved spring mass and damper system.

5.1 Case Study

5.1.1 Constant System Model

Consider a constant signal system represented by the discrete system

$$x_1(k) = \theta x_1(k-1)$$

where $x_1(n_s = 1)$ at any instant of time is a constant state=10 and the parameter $\theta = 1(n_p = 1)$. If the parameter is unknown it can be augmented as unknown state x_2 and the above state space equation becomes

$$\begin{aligned} x_1(k) &= x_2(k-1)x_1(k-1) \\ x_2(k) &= x_2(k-1) \end{aligned}$$

The above nonlinear equation is in the form $X(k) = f(X(k-1))$ where $X(k) = [x_1(k), x_2(k)]^T$ and f being the nonlinear function whose Jacobian matrix (2×2) is calculated as per E-(2.7) given by

$$F_{k,k-1} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix}_{|X=\hat{X}_{k-1|k-1}}$$

$$F_{k,k-1} = \begin{bmatrix} x_2 & x_1 \\ 0 & 1 \end{bmatrix}_{|X=\hat{X}_{k-1|k-1}}$$

The linear measurement model is described by the following equation

$$Z(k) = HX(k) + \upsilon(k)$$

where $\upsilon(k) \sim \mathcal{N}(0, R)$ and the measurement matrix for $n_m = 1$ is H=[1 0].
n_s is the number of original states.
n_p is the number of augmented states.
$n_s + n_p$=Total unknown states.
n_m is the number of measured states.

The simulated results are tabulated in table-5.1 and the plots are shown below for a constant system(constant state=10) with R=0.05 for 100 ensembles each with 20 iterations.X_0 was chosen as [10.1,1.1] with P_0=10 and initial R=0.1. MBF smoothing was used for estimating initial states and its covariance and R was estimated using smoothed innovations.

Table 5.1: Typical SMD system results

	theta	**%CRB***	$\hat{R}$
NR	1.0000	0.0038	0.0484
EKF	1.0000	0.0040	0.0476

$* \%CRB = \frac{\sqrt{P_{\hat{\theta}}}}{\theta} \times 100$, crb ratio=1.0526

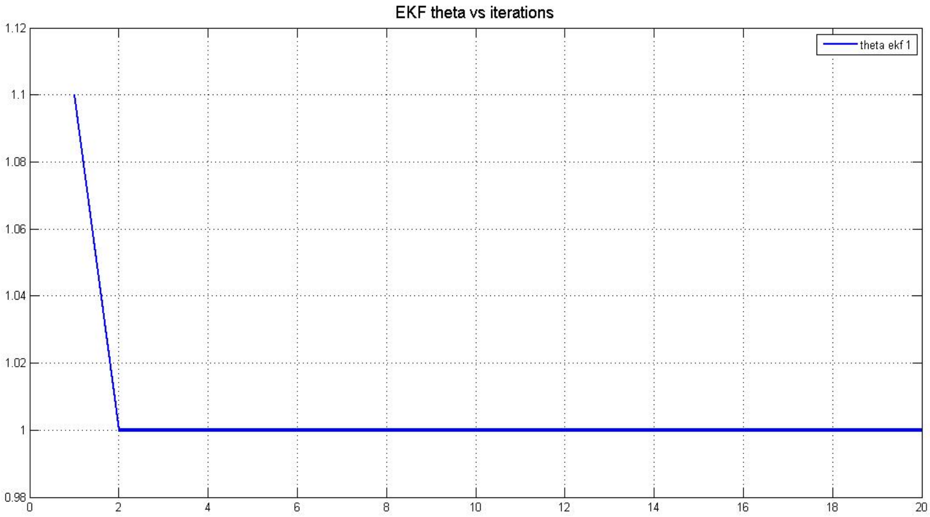


Figure 5.1: EKF-Constant system, theta vs iteration

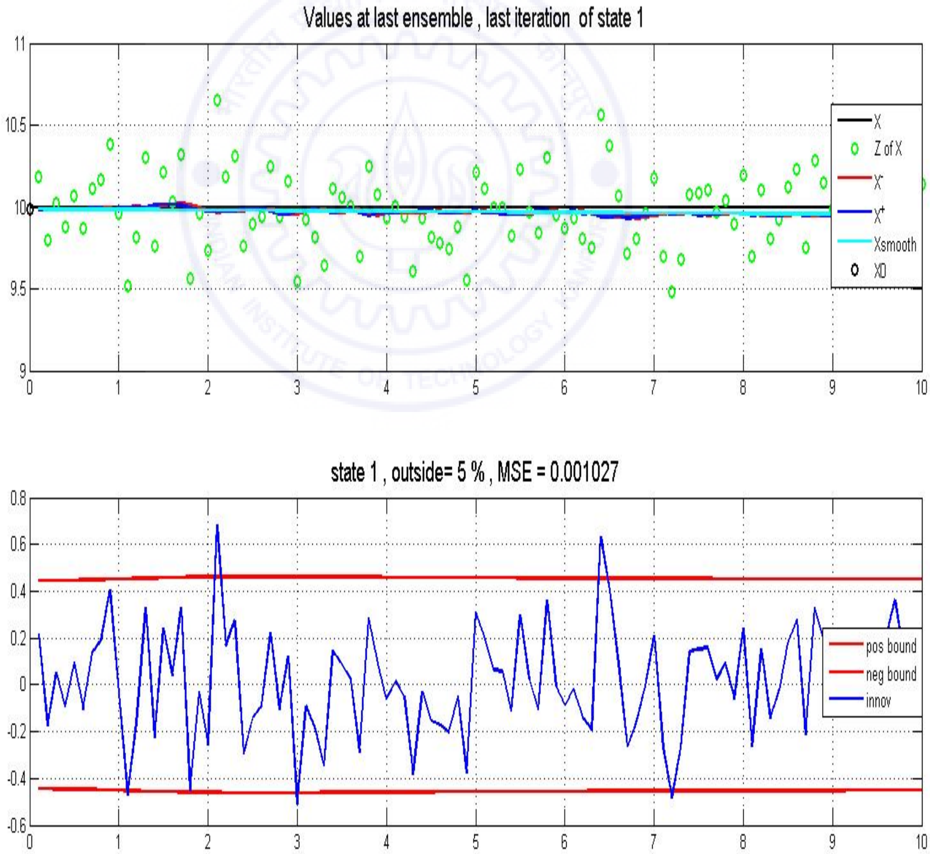


Figure 5.2: EKF-Constant system, Last iteration values of constant state=10

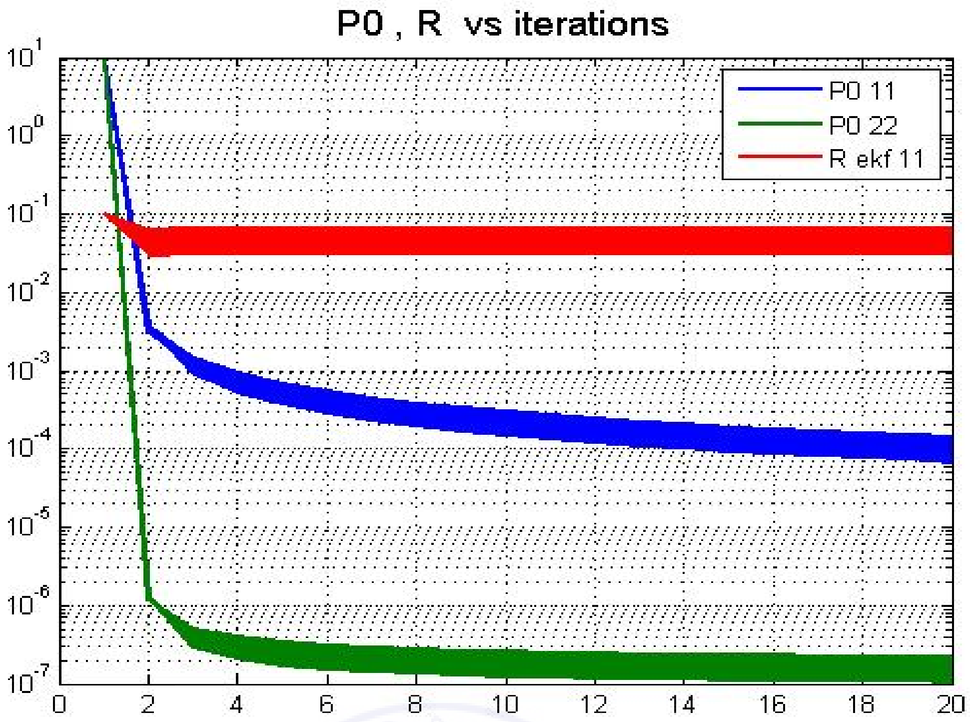


Figure 5.3: EKF-Constant system, P0 and R vs iteration

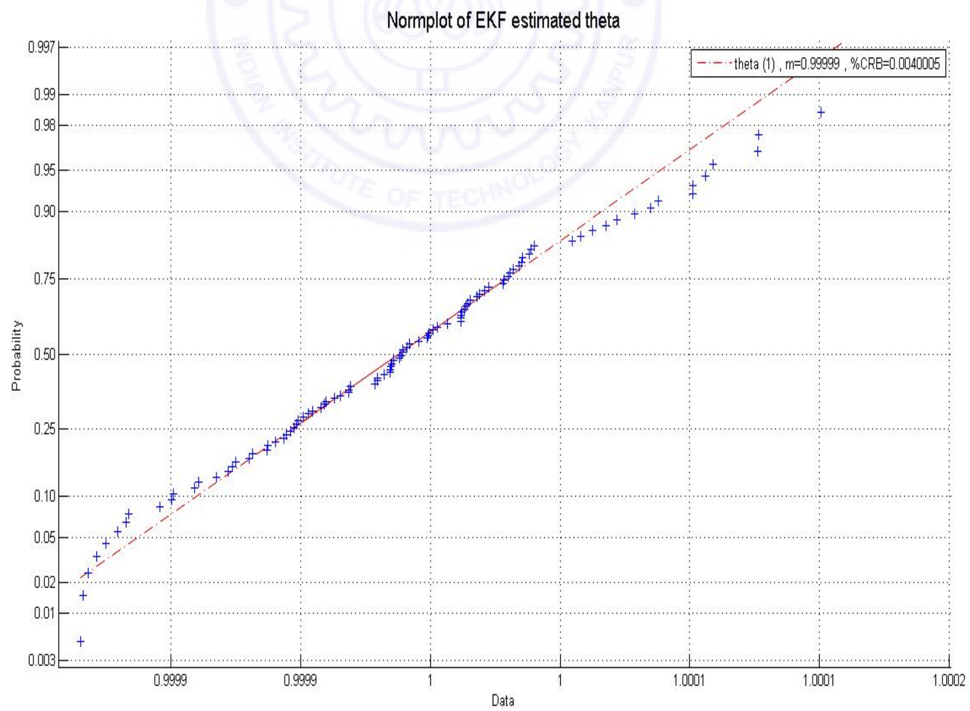


Figure 5.4: EKF-Constant system, Normal Probability plot of $\hat{\theta}$

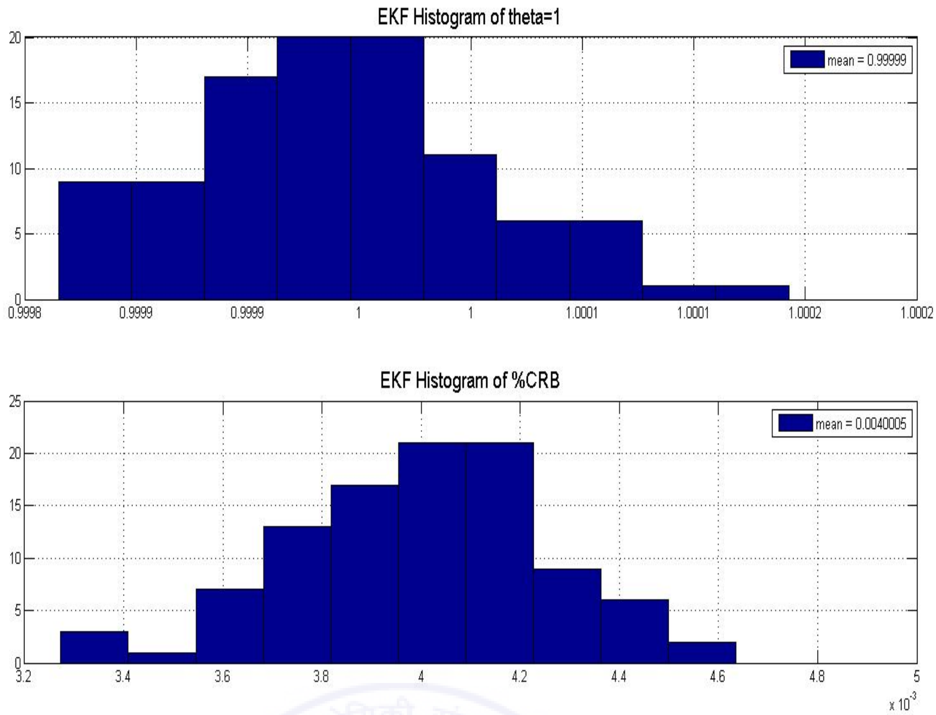


Figure 5.5: EKF-Constant system, Histogram of $\hat{\theta}$

5.1.2 SMD System Model

Consider the SMD system with weak nonlinear parameter defined by the second order differential equation [6] given by

$$m\ddot{x}_1 + c\dot{x}_1 + k_1 x_1 + k_2 x_1^3 = U$$

The state space form with zero control input is

$$\left.\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\theta_1 x_1 - \theta_2 x_2 - \theta_3 x_1^3 \end{aligned}\right\} \tag{5.1}$$

where

x_1 : displacement state with the initial condition $x_1(0) = 1$

x_2 : velocity state with the initial condition $x_2(0) = 0$

m : mass constant

c : viscous damping co-efficient

k : spring constant

θ_1 : parameter(k_1/m)

θ_2 : parameter(c/m)

θ_3 : parameter(k_2/m)

The displacement(x_1) and velocity(x_2) states constitute the original states(n_s=2) which are unknown and needs to be estimated or in other words tracked.The unknown parameters(n_p=3) $\theta_1 = 4$,$\theta_2 = 0.4$ and $\theta_3 = 0.6$ are augmented as unknown states x_3,x_4 and x_5 respectively using constant signal model $\dot{\theta} = 0$ or $\theta_k = \theta_{k-1}$.The parameter θ_3 is considered as a weak parameter since its value do not affect the system dynamics to a great extent and estimation of such parameter is a major challenge. The state space equation in Eq-(5.1) now becomes

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = -x_3x_1 - x_4x_2 - x_5x_1^3$$

The linearized model is represented as

$$\Delta\dot{X} = A\Delta X$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ -x_3 - 3x_1^2x_5 & -x_4 & -x_1 & -x_2 & -x_1^3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The state Jacobian with δt as the sampling time period is given by

$$F_{k,k-1} = e^{A\delta t}|_{X(k)=\hat{X}(k-1)}$$

The linear measurement model is described by the following equation

$$Z(k) = HX(k) + \upsilon(k)$$

where $\upsilon(k) \sim \mathcal{N}(0, R)$ & the complete state vector with a total of 2+3=5 states is

$$X(k) = \Big[x_1(k), \quad x_2(k), \quad x_3(k), \quad x_4(k), \quad x_5(k)\Big]^T$$

The measurement matrix for $n_m = 2$ is

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

The results for 50 ensemble runs each with 10 iterations are tabulated in table-5.2.The system is assumed to be free of process noise and the measurement noise m_noise=[0.01,0.04]. The initial X_0 is chosen to be $[1.0, 0.0, 5, 0.5, 0.5]^T$ with $P0\_power = R0\_power = 0$. MBF smoothing technique was used and the measurement noise matrix R was estimated using smoothed innovations. Total no. of measurement samples are N=100 with a sampling time of dt=0.1s and true parameter values being $\theta_1 = 4$, $\theta_2 = 0.4$ and $\theta_3 = 0.6$.

Table 5.2: Typical SMD system results

	theta	**%CRB***	$\hat{R}$
NR	3.9969 , 0.4006 , 0.6074	1.8779 , 3.2293 , 33.2294	0.0096 , 0.0387
EKF	3.9836 , 0.4013 , 0.6775	1.8605 , 3.1759 , 33.8951	0.0094 , 0.0378

$* \ \%CRB = \frac{\sqrt{P_{\hat{\theta}}}}{\theta} \times 100$, crb ratio= [0.9907 , 0.9835 , 1.0200]

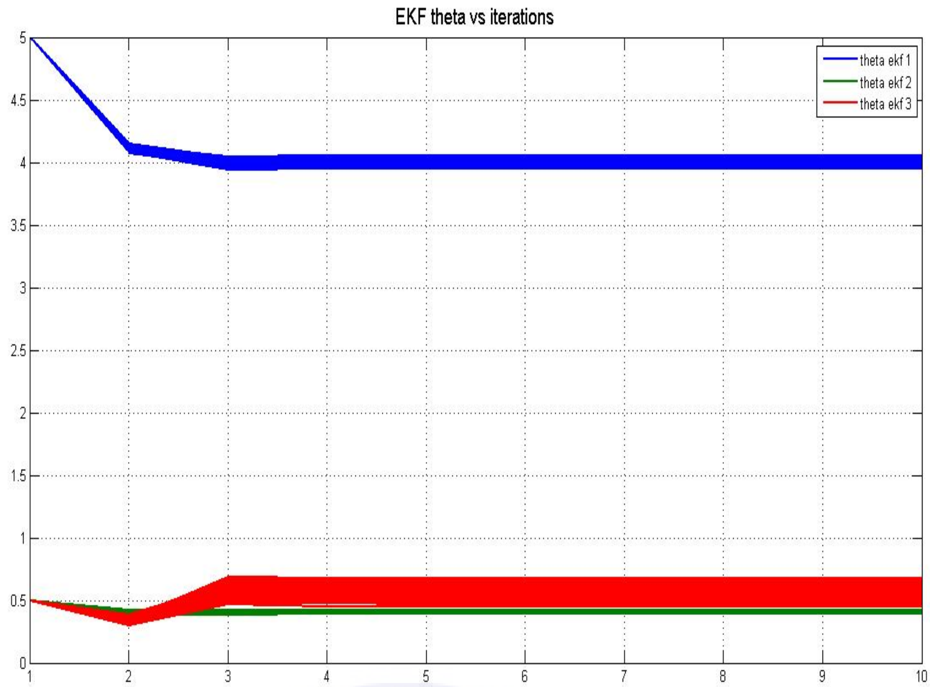


Figure 5.6: EKF-SMD system, $\hat{\theta}$ vs iterations

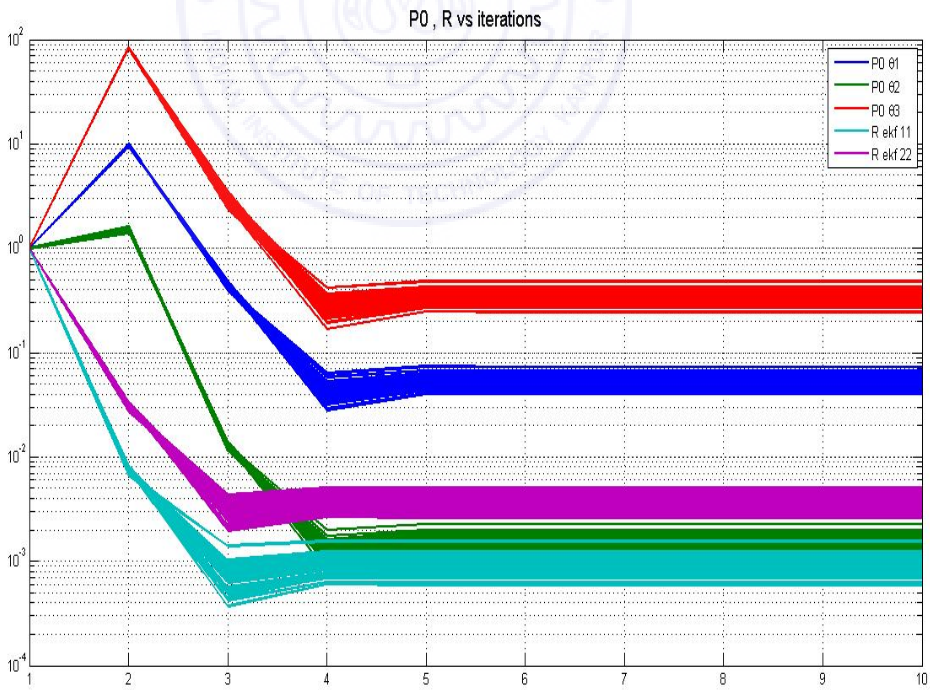


Figure 5.7: EKF-SMD system, P0 and R vs iteration

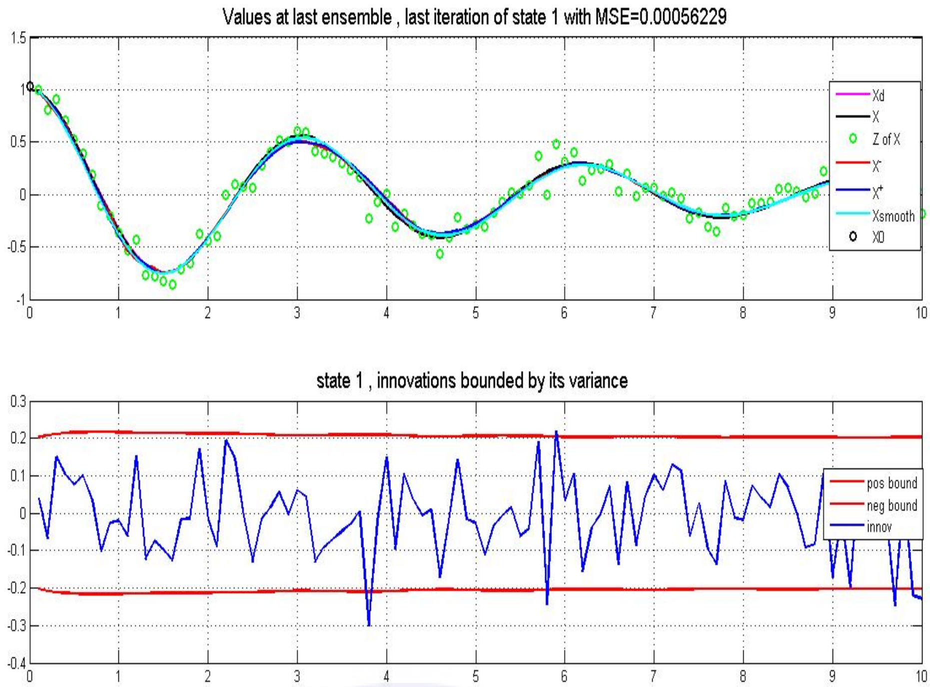


Figure 5.8: EKF-SMD system,Last iteration values of state=1

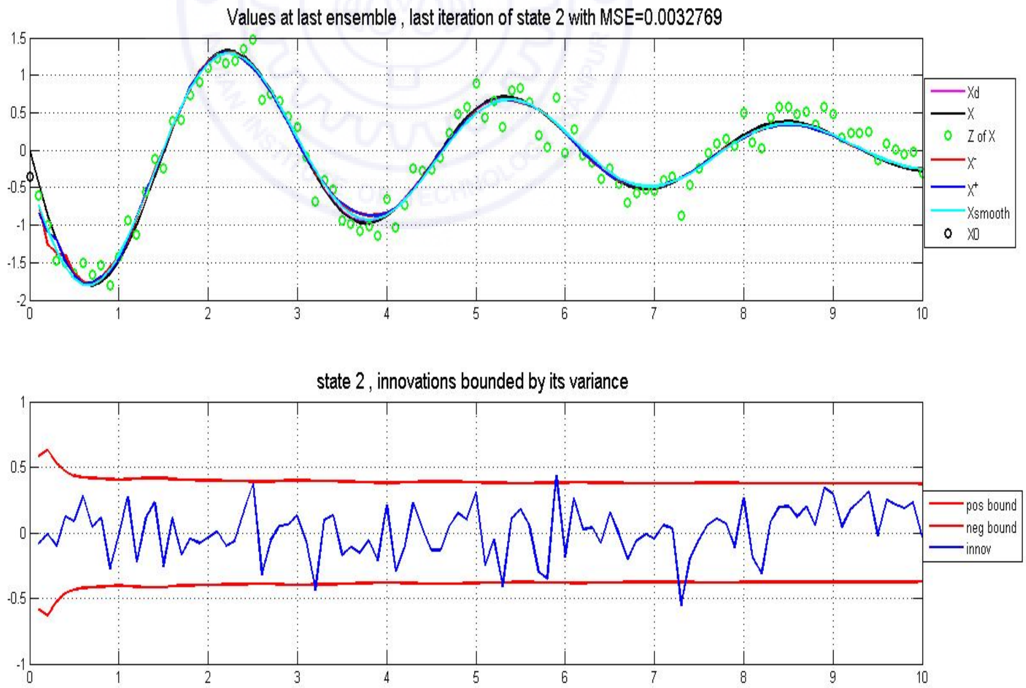


Figure 5.9: EKF-SMD system,Last iteration values of state=2

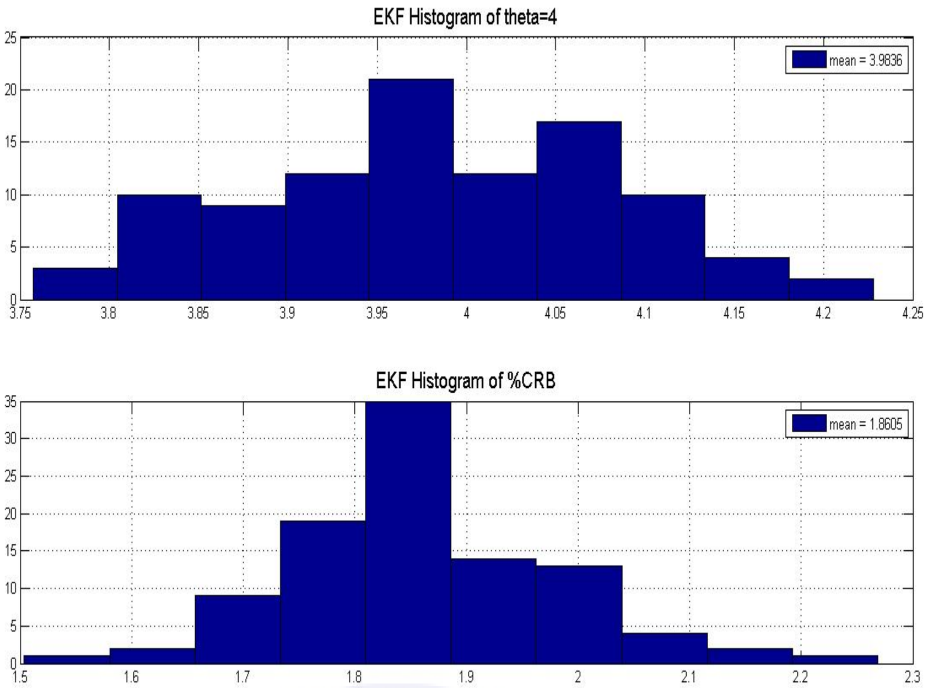


Figure 5.10: EKF-SMD system, Histogram of $\hat{\theta}_1$

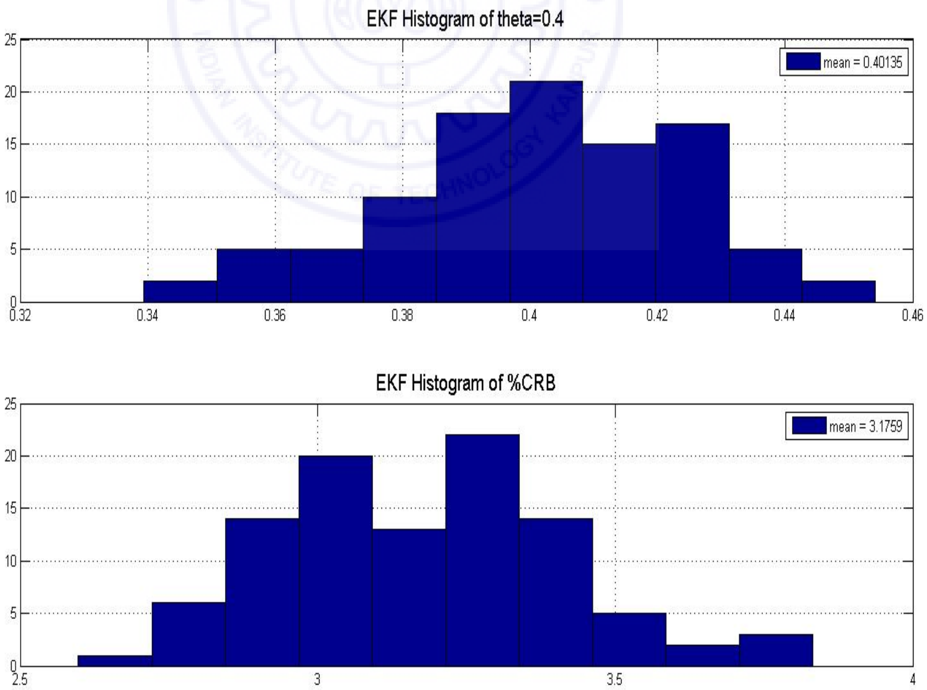


Figure 5.11: EKF-SMD system, Histogram of $\hat{\theta}_2$

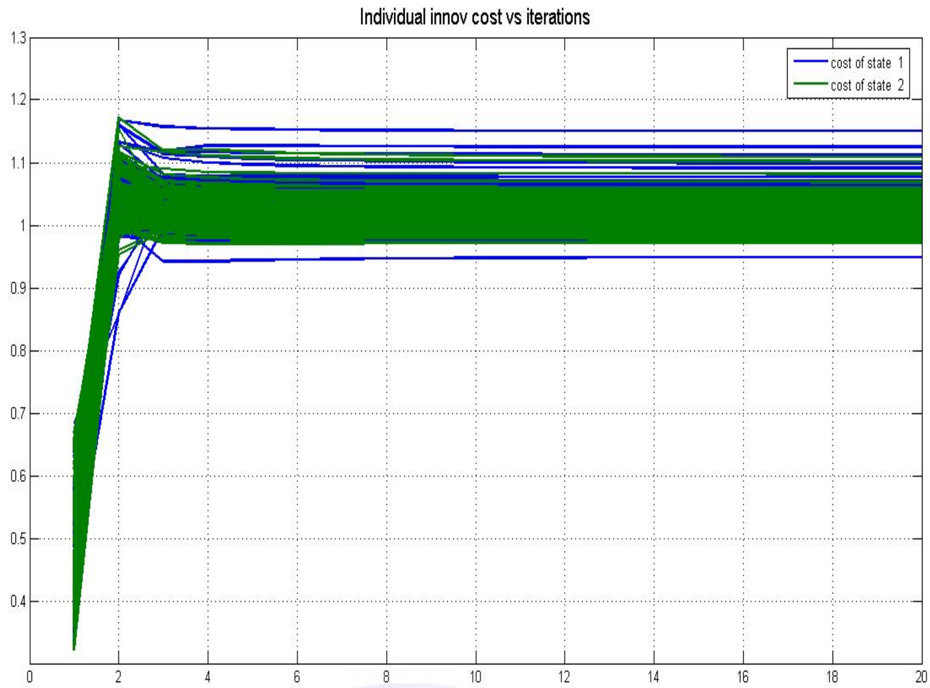


Figure 5.12: EKF-SMD system, innovation cost (J1) vs iteration

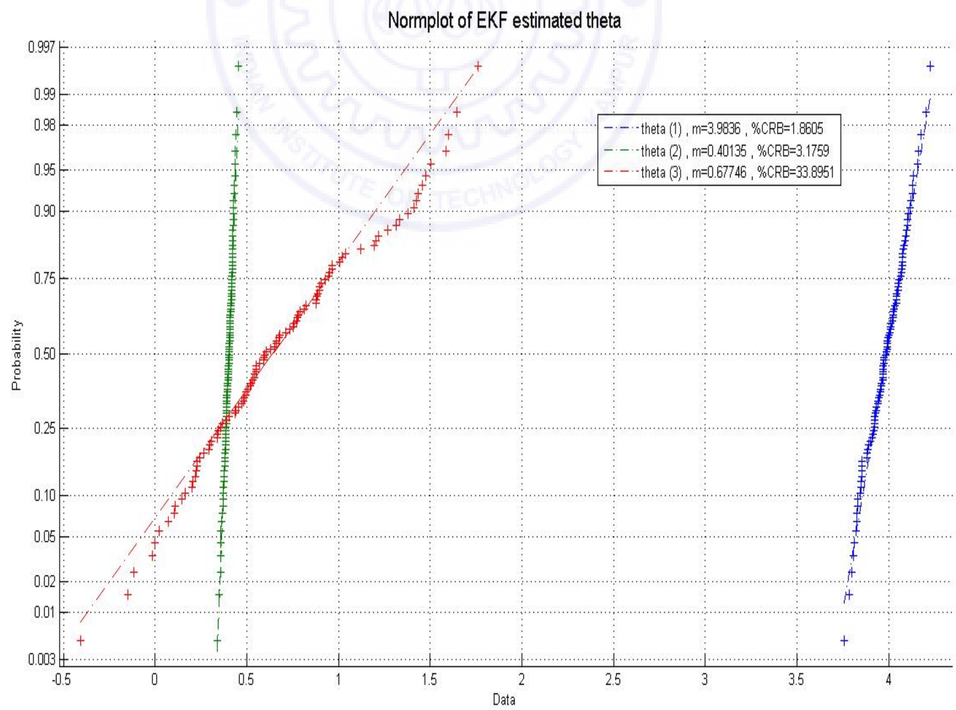


Figure 5.13: EKF-SMD system, Normal Probability plot of $\hat{\theta}$

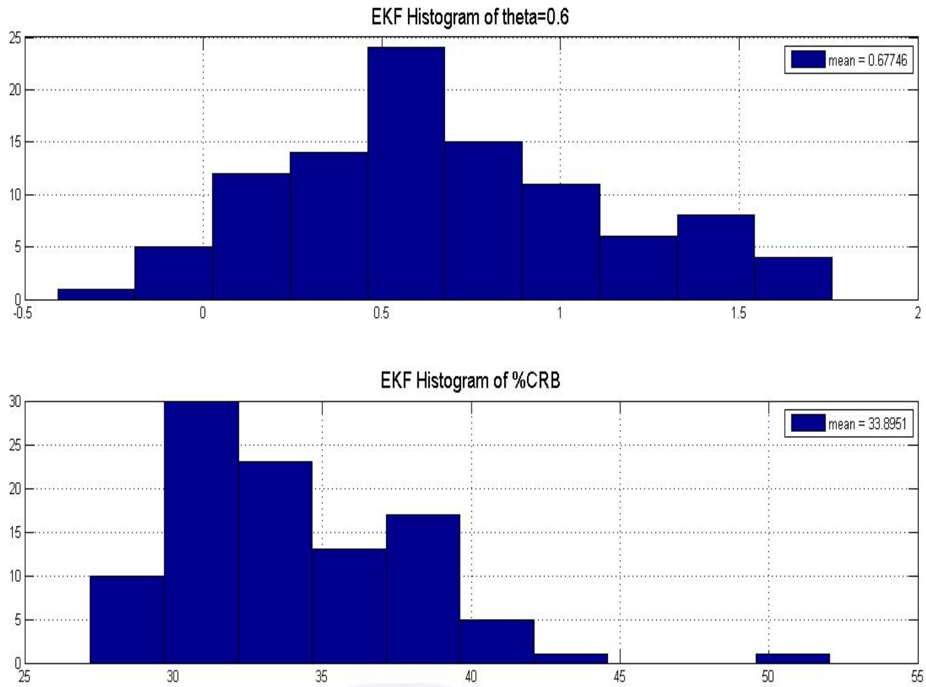


Figure 5.14: EKF-SMD system, Histogram of $\hat{\theta}_3$

Observations :

- The ratio of the variance of the parameter estimated by EKF to that of the CRB estimated by NR is very close to unity as seen in table- 5.2.Thus the proposed adaptive EKF is an efficient estimator.

- The innovations are bounded by the confidence levels determined by its standard deviation $\pm\sqrt{H_k P_{k|k-1} H_k^T + R}$ as seen in Fig- 5.8 and 5.9.

- The normal probability plot is a straight line and thus the estimated parameters by EKF are distributed normally.

- The cost function (J1) is defined as

$$J1 = \frac{1}{N}\sum_{k=1}^{k=N}(Z - H_k\hat{X}_{k|k-1})^T(H_k P_{k|k-1} H_k + R)^{-1}(Z - H_k\hat{X}_{k|k-1}) \qquad (5.2)$$

The cost values of each state tend towards unity as expected.

5.2 Sensitivity Study

In this section let us conduct different sensitivity studies on the SMD system to test the robustness of the proposed tuning algorithm and show that its results are independent of the varying input. The typical values of R chosen is R=diag(0.001,0.004), N=100 and discrete time t_k=0,0.1..10 with dt=0.1 seconds. Its observed that in case of the non-linear SMD system equilibrium is achieved within 10 iterations and CRB's are estimated after 10 ensemble runs. The different sensitivity studies are-

- Sensitivity Study 1 : R is estimated using the three different options as explained in 4.2.
- Sensitivity Study 2 : R is estimated using the three different options but with only the last half of the 'r' samples.
- Sensitivity Study 3 : To show that the results are unique and independent for a wide rage of initial P_0 ranging from 10^{-5} to 10^5.
- Sensitivity Study 4 : To show that the results are unique and independent for a wide rage of initial R from R_0power=-4 to 1.
- Sensitivity Study 5 : Parameter covariance Scaling up is varied from 0.1*N to 10*N to check for its effect.
- Sensitivity Study 6 : Extreme initial values of X_0 is tested for the robustness of the algorithm.
- Sensitivity Study 7 : Smoothing options are varied.
- Sensitivity Study 8 : m_noise is increased to check for the convergence of the $\hat{R}$, θ and its CRB.
- Sensitivity Study 9 : Experiments varying the combination of N and dt with $Scale\_up$=N.
- Sensitivity Study 10 : Experiments varying the ensemble runs to show that the CRB ratio tend towards unity.

where some of the variables used in the MATLAB program are described as below

N : number of measurement samples

dt : sampling time interval

m_noise : measurement noise $R = diag(m\_noise)$

p_noise : process noise $Q = diag(p\_noise)$

$ensemble\_max$: maximum number of ensembles

$iter\_max$: maximum number of iterations

r_option : $\hat{R}$ using 1-innovations,2-residue,3-smoothed residue

$smooth\_option$: 1-RTS smoothing ,2-MBF smoothing

$last\_n$: The number of last few samples used for $\hat{R}$

$Scale\_up$: The factor N used for initial parameter covariance

$P0\_power$: Initial $P_0 = 10^{P0\_power}$

$R0\_power$: Initial $R_0 = 10^{R0\_power}$

$x0\_initial$: Guess value of initial state vector(X_0)

$x0\_est$: EKF Esimated X_0

$P0\_est$: EKF Estimated P_0

R_nr : Ensemble mean of NR estimated R

R_ekf : Ensemble mean of EKF estimated R

Q_MT : Estimated Q by MT method

Q_DSDT : Estimated Q by DSDT method

Q_ekf : Ensemble mean of EKF estimated Q

$theta\_ekf$: Ensemble mean of EKF estimated θ

$theta\_nr$: Ensemble mean of NR estimated θ

crb_ekf : Ensemble mean of EKF estimated %CRB

crb_nr : Ensemble mean of NR estimated %CRB

crb_ratio : ratio of crb_ekf to crb_nr

$theta\_corrcoef\_nr$: Correlation coefficient of the NR estimated parameters

$theta\_corrcoef\_ekf$: Correlation coefficient of the EKF estimated parameters.

Sensitivity Study-1

Exp No.	1	2	3
N	100	100	100
dt	0.1	0.1	0.1
ensemble max	100	100	100
iter max	10	10	10
r_option	1	2	3
smooth option	2	2	2
last_n	100	100	100
Scale_up	100	100	100
P0 power	0	0	0
R0 power	0	0	0
initial x0 1	1.1	1.1	1.1
initial x0 2	0.1	0.1	0.1
initial x0 3	5	5	5
initial x0 4	0.5	0.5	0.5
initial x0 5	0.5	0	0
x0 est 1	0.998145	0.997853	0.998042125
x0 est 2	0.005361	0.006592	0.005430013
x0 est 3	3.995628	3.99479	3.995711195
x0 est 4	0.399995	0.399989	0.399963053
x0 est 5	0.618569	0.622596	0.618501276
P0 est 1	0.00000902	0.00000737	8.64E-06
P0 est 2	0.000133	0.000109	0.000129213
P0 est 3	0.056337	0.048949	0.055847607
P0 est 4	0.001745	0.001508	0.001724831
P0 est 5	0.445607	0.383873	0.440701248
R nr 11	0.000965	0.000965	0.000965026
R nr 22	0.003871	0.003871	0.003871353
R ekf 11	0.000957	0.000856	0.000932194
R ekf 22	0.003677	0.003145	0.003707745
theta nr 1	3.998956	3.998956	3.998956033
theta nr 2	0.400244	0.400244	0.400244064
theta nr 3	0.602411	0.602411	0.602410622
theta ekf 1	3.995628	3.99479	3.995711195
theta ekf 2	0.399995	0.399989	0.399963053
theta ekf 3	0.618569	0.622596	0.618501276
crb ratio 1	0.998554	0.930824	0.994294194
crb ratio 2	1.02204	0.950214	1.016140173
crb ratio 3	1.057078	0.981226	1.051430216

Sensitivity Study-2

Exp No.	1	2	3
N	100	100	100
dt	0.1	0.1	0.1
ensemble max	100	100	100
iter max	10	10	10
r_option	1	2	3
smooth option	2	2	2
last_n	50	50	50
Scale_up	100	100	100
P0 power	0	0	0
R0 power	0	0	0
initial x0 1	1.1	1.1	1.1
initial x0 2	0.1	0.1	0.1
initial x0 3	5	5	5
initial x0 4	0.5	0.5	0.5
initial x0 5	0	0	0
x0 est 1	0.998067	0.998031	0.998078
x0 est 2	0.006057	0.00622	0.00592
x0 est 3	3.995365	3.995261	3.99553
x0 est 4	0.399936	0.399934	0.399935
x0 est 5	0.619628	0.620097	0.619065
P0 est 1	8.85E-06	8.33E-06	8.63E-06
P0 est 2	0.000132	0.000124	0.000129
P0 est 3	0.057151	0.053856	0.055631
P0 est 4	0.001766	0.001663	0.001719
P0 est 5	0.450508	0.424506	0.438548
R nr 11	0.000965	0.000965	0.000965
R nr 22	0.003871	0.003871	0.003871
R ekf 11	0.000972	0.000917	0.000948
R ekf 22	0.003773	0.003567	0.003673
theta nr 1	3.998956	3.998956	3.998956
theta nr 2	0.400244	0.400244	0.400244
theta nr 3	0.602411	0.602411	0.602411
theta ekf 1	3.995365	3.995261	3.99553
theta ekf 2	0.399936	0.399934	0.399935
theta ekf 3	0.619628	0.620097	0.619065
crb ratio 1	1.004737	0.975309	0.991276
crb ratio 2	1.026909	0.996737	1.013206
crb ratio 3	1.061901	1.030766	1.047709

Sensitivity Study-3

Exp No.	1	2	3	4	5
N	100	100	100	100	100
dt	0.1	0.1	0.1	0.1	0.1
ensemble max	100	100	100	100	100
iter max	10	10	10	10	10
r_option	3	3	3	3	3
smooth option	2	2	2	2	2
last_n	100	100	100	100	100
Scale_up	100	100	100	100	100
P0 power	-5	-3	0	3	5
R0 power	0	0	0	0	0
initial x0 1	1.1	1.1	1.1	1.1	1.1
initial x0 2	0.1	0.1	0.1	0.1	0.1
initial x0 3	5	5	5	5	5
initial x0 4	0.5	0.5	0.5	0.5	0.5
initial x0 5	0	0	0	0	0
x0 est 1	1.051282	0.998617	0.998042	0.997856	1.016379
x0 est 2	0.103577	0.01272	0.00543	0.007002	-0.09516
x0 est 3	3.993991	3.992756	3.995711	3.994863	4.032467
x0 est 4	0.435435	0.401107	0.399963	0.400069	0.411357
x0 est 5	0.718094	0.636632	0.618501	0.62319	0.458482
P0 est 1	5.96E-06	9E-06	8.64E-06	8.77E-06	1.31E-05
P0 est 2	9.44E-06	0.000113	0.000129	0.000131	0.000259
P0 est 3	0.070083	0.055637	0.055848	0.055933	0.086467
P0 est 4	0.002737	0.00174	0.001725	0.001729	0.00412
P0 est 5	0.415814	0.430728	0.440701	0.442227	0.564035
R nr 11	0.000965	0.000965	0.000965	0.000965	0.000965
R nr 22	0.003871	0.003871	0.003871	0.003871	0.003871
R ekf 11	0.001252	0.000933	0.000932	0.000932	0.001395
R ekf 22	0.004841	0.00371	0.003708	0.003707	0.006301
theta nr 1	3.998956	3.998956	3.998956	3.998956	3.998956
theta nr 2	0.400244	0.400244	0.400244	0.400244	0.400244
theta nr 3	0.602411	0.602411	0.602411	0.602411	0.602411
theta ekf 1	3.993991	3.992756	3.995711	3.994863	4.032467
theta ekf 2	0.435435	0.401107	0.399963	0.400069	0.411357
theta ekf 3	0.718094	0.636632	0.618501	0.62319	0.458482
crb ratio 1	1.113539	0.992421	0.994294	0.995054	1.11115
crb ratio 2	1.279174	1.020581	1.01614	1.017355	1.216937
crb ratio 3	1.021042	1.039574	1.05143	1.053235	1.139764

Sensitivity Study-4

Exp No.	1	2	3
N	100	100	100
dt	0.1	0.1	0.1
ensemble max	100	100	100
iter max	10	10	10
r_option	3	3	3
smooth option	2	2	2
last_n	100	100	100
Scale_up	100	100	100
P0 power	0	0	0
R0 power	-4	-1	1
p noise 1	0	0	0
p noise 2	0	0	0
initial x0 1	1.1	1.1	1.1
initial x0 2	0.1	0.1	0.1
initial x0 3	5	5	5
initial x0 4	0.5	0.5	0.5
initial x0 5	0	0	0
x0 est 1	1.01648	1.002979	1.000004
x0 est 2	0.036095	0.00139	0.01099
x0 est 3	3.995383	4.000925	3.994555
x0 est 4	0.412167	0.401919	0.401609
x0 est 5	0.651196	0.597181	0.629007
P0 est 1	1.24E-07	5.02E-06	6.39E-06
P0 est 2	1.88E-06	7.7E-05	9.39E-05
P0 est 3	0.055899	0.05435	0.05525
P0 est 4	0.001761	0.001661	0.001694
P0 est 5	0.362288	0.400575	0.417325
R nr 11	0.000965	0.000965	0.000965
R nr 22	0.003871	0.003871	0.003871
R ekf 11	0.000996	0.000934	0.000938
R ekf 22	0.003948	0.003719	0.003716
theta nr 1	3.998956	3.998956	3.998956
theta nr 2	0.400244	0.400244	0.400244
theta nr 3	0.602411	0.602411	0.602411
theta ekf 1	3.995383	4.000925	3.994555
theta ekf 2	0.412167	0.401919	0.401609
theta ekf 3	0.651196	0.597181	0.629007
crb ratio 1	0.993875	0.980908	0.988963
crb ratio 2	1.022906	0.997162	1.006763
crb ratio 3	0.953178	1.002594	1.023354

Sensitivity Study-5

Exp No.	1	2	3
N	100	100	100
dt	0.1	0.1	0.1
ensemble max	100	100	100
iter max	10	10	10
r_option	3	2	3
smooth option	2	2	2
last_n	100	100	100
Scale_up	10	100	1000
P0 power	0	0	0
R0 power	0	0	0
initial x0 1	1.1	1.1	1.1
initial x0 2	0.1	0.1	0.1
initial x0 3	5	5	5
initial x0 4	0.5	0.5	0.5
initial x0 5	0	0	0
x0 est 1	0.999397338	0.99785	0.998951
x0 est 2	-0.001025025	0.00659	0.001954
x0 est 3	3.99933089	3.99479	3.998165
x0 est 4	0.400039353	0.39999	0.399768
x0 est 5	0.600295816	0.6226	0.605885
P0 est 1	8.03095E-06	7.4E-06	8.72E-06
P0 est 2	0.000111211	0.00011	0.000131
P0 est 3	0.004626735	0.04895	0.565968
P0 est 4	0.000155845	0.00151	0.017388
P0 est 5	0.035801024	0.38387	4.47212
R nr 11	0.000965026	0.00097	0.000965
R nr 22	0.003871353	0.00387	0.003871
R ekf 11	0.000933061	0.00086	0.000932
R ekf 22	0.003721107	0.00315	0.003704
theta nr 1	3.998956033	3.99896	3.998956
theta nr 2	0.400244064	0.40024	0.400244
theta nr 3	0.602410622	0.60241	0.602411
theta ekf 1	3.99933089	3.99479	3.998165
theta ekf 2	0.400039353	0.39999	0.399768
theta ekf 3	0.600295816	0.6226	0.605885
crb ratio 1	0.905018699	0.93082	1.000922
crb ratio 2	0.965906451	0.95021	1.020256
crb ratio 3	0.947805382	0.98123	1.059048

Sensitivity Study-6

Exp No.	1	2	3
N	100	100	100
dt	0.1	0.1	0.1
ensemble max	100	100	100
iter max	20	20	20
r_option	3	3	3
smooth option	2	2	2
last_n	100	100	100
Scale_up	100	100	100
P0 power	0	0	0
R0 power	0	0	0
initial x0 1	1.1	2	1.5
initial x0 2	0.1	-1	-0.1
initial x0 3	100	100	100
initial x0 4	50	50	-50
initial x0 5	10	10	-50
x0 est 1	0.997049	0.997798	1.004513
x0 est 2	0.008715	0.004462	-0.02179
x0 est 3	3.993559	3.996034	4.012831
x0 est 4	0.399876	0.399715	0.400002
x0 est 5	0.629414	0.616067	0.531486
P0 est 1	1.18E-05	1.47E-05	5.41E-06
P0 est 2	0.000183	0.000222	8.12E-05
P0 est 3	0.057199	0.057966	0.054077
P0 est 4	0.0018	0.001864	0.001654
P0 est 5	0.476341	0.500251	0.400471
R nr 11	0.000965	0.000965	0.000965
R nr 22	0.003871	0.003871	0.003871
R ekf 11	0.000933	0.000931	0.000935
R ekf 22	0.003703	0.003707	0.00373
theta nr 1	3.998956	3.998956	3.998956
theta nr 2	0.400244	0.400244	0.400244
theta nr 3	0.602411	0.602411	0.602411
theta ekf 1	3.993559	3.996034	4.012831
theta ekf 2	0.399876	0.399715	0.400002
theta ekf 3	0.629414	0.616067	0.531486
crb ratio 1	1.006182	1.012978	0.978419
crb ratio 2	1.038087	1.056294	0.995011
crb ratio 3	1.092865	1.120217	1.002338

Sensitivity Study-7

Exp No.	1	2	3
N	100	100	100
dt	0.1	0.1	0.1
ensemble max	100	100	100
iter max	20	20	20
r_option	3	3	3
smooth option	2	1	1
last_n	100	100	100
Scale_up	100	100	100
P0 power	0	0	0
R0 power	0	0	0
initial x0 1	1.5	1.5	1.5
initial x0 2	-0.1	-0.1	-0.1
initial x0 3	10	10	10
initial x0 4	-10	-10	-10
initial x0 5	-10	-10	-10
x0 est 1	1.004513	0.973772	0.997504
x0 est 2	-0.02179	0.113361	0.007303
x0 est 3	4.012831	3.934023	3.995108
x0 est 4	0.400002	0.402314	0.399889
x0 est 5	0.531486	0.95657	0.622327
P0 est 1	5.41E-06	4.22E-06	0.000928
P0 est 2	8.12E-05	6.83E-05	0.00368
P0 est 3	0.054077	0.061585	0.08012
P0 est 4	0.001654	0.001703	0.00292
P0 est 5	0.400471	0.477475	1.106238
R nr 11	0.000965	0.000965	0.000965
R nr 22	0.003871	0.003871	0.003871
R ekf 11	0.000935	0.000972	0.000928
R ekf 22	0.00373	0.003926	0.00368
theta nr 1	3.998956	3.998956	3.998956
theta nr 2	0.400244	0.400244	0.400244
theta nr 3	0.602411	0.602411	0.602411
theta ekf 1	4.012831	3.934023	3.995108
theta ekf 2	0.400002	0.402314	0.399889
theta ekf 3	0.531486	0.95657	0.622327
crb ratio 1	0.978419	1.043734	1.190893
crb ratio 2	0.995011	1.009502	1.322075
crb ratio 3	1.002338	1.093835	1.665408

Sensitivity Study-8

Exp No.	1	2	3
N	100	100	100
dt	0.1	0.1	0.1
ensemble max	100	100	100
iter max	20	20	20
m_noise 1	0.001	0.01	0.1
m_noise 2	0.004	0.04	0.4
r_option	3	3	3
last_n	100	100	100
Scale_up	100	100	100
P0 power	0	0	0
R0 power	0	0	0
initial x0 1	1.5	1.5	1.5
initial x0 2	-0.1	-0.1	-0.1
initial x0 3	10	10	10
initial x0 4	-10	-10	-10
initial x0 5	-10	-10	-10
x0 est 1	1.004513	0.961731	0.931173
x0 est 2	-0.02179	0.152609	0.169753
x0 est 3	4.012831	3.912254	3.947159
x0 est 4	0.400002	0.404769	0.429035
x0 est 5	0.531486	1.085786	1.044261
P0 est 1	5.41E-06	3.57E-05	0.000375
P0 est 2	8.12E-05	0.000581	0.005628
P0 est 3	0.054077	0.613267	7.482487
P0 est 4	0.001654	0.016301	0.17873
P0 est 5	0.400471	4.918385	77.77094
R nr 11	0.000965	0.00965	0.096463
R nr 22	0.003871	0.038714	0.387125
R ekf 11	0.000935	0.0094	0.093335
R ekf 22	0.00373	0.037481	0.37374
theta nr 1	3.998956	3.996888	3.99042
theta nr 2	0.400244	0.400606	0.400395
theta nr 3	0.602411	0.607416	0.625187
theta ekf 1	4.012831	3.912254	3.947159
theta ekf 2	0.400002	0.404769	0.429035
theta ekf 3	0.531486	1.085786	1.044261
crb ratio 1	0.978419	1.040175	1.125638
crb ratio 2	0.995011	0.985742	1.017697
crb ratio 3	1.002338	1.104137	1.293075

Sensitivity Study-9

Exp No.	1	2	3
N	20	50	100
dt	0.5	0.25	0.1
ensemble max	100	100	100
iter max	20	20	20
r_option	3	3	3
smooth option	1	1	1
last_n	20	50	100
Scale_up	20	50	100
P0 power	0	0	0
R0 power	0	0	0
initial x0 1	1.01	1.01	1.01
initial x0 2	-0.01	-0.01	-0.01
initial x0 3	5	5	5
initial x0 4	0.5	0.5	0.5
initial x0 5	0.5	0.5	0.5
x0 est 1	1.011498801	1.003012	0.997814
x0 est 2	-0.049403091	-0.02	0.005803
x0 est 3	4.01158435	4.005734	3.995643
x0 est 4	0.398880862	0.400452	0.39986
x0 est 5	0.506924405	0.55947	0.619436
P0 est 1	2.16946E-05	9.98E-06	4.05E-06
P0 est 2	0.000212243	0.000133	6.01E-05
P0 est 3	0.030258728	0.040522	0.054376
P0 est 4	0.001158253	0.001568	0.001611
P0 est 5	0.154136829	0.283611	0.398554
R nr 11	0.000886974	0.000952	0.000965
R nr 22	0.003756566	0.003747	0.003871
R ekf 11	0.000736738	0.000898	0.000933
R ekf 22	0.00325852	0.003433	0.003712
theta nr 1	3.988822441	3.997484	3.998956
theta nr 2	0.400437181	0.400993	0.400244
theta nr 3	0.626105285	0.608772	0.602411
theta ekf 1	4.01158435	4.005734	3.995643
theta ekf 2	0.398880862	0.400452	0.39986
theta ekf 3	0.506924405	0.55947	0.619436
crb ratio 1	0.911229107	0.95725	0.981104
crb ratio 2	0.8655972	0.914827	0.982084
crb ratio 3	0.8364888	0.923103	0.999913

Sensitivity Study-10

Table 5.3: Sensitivity results for 10 ensembles

	theta	**%CRB**	$\hat{R}$
NR	3.9924 , 0.4029 , 0.6270	0.5982 , 1.0309 , 10.5952	0.0010 , 0.0039
EKF	3.9927 , 0.4031 , 0.6241	0.5813 , 0.9838 , 10.0520	0.0009 , 0.0037

$$theta\_corrcoef\_nr = \begin{bmatrix} 1.0000 & -0.4408 & -0.9744 \\ -0.4408 & 1.0000 & 0.5007 \\ -0.9744 & 0.5007 & 1.0000 \end{bmatrix}$$

$$theta\_corrcoef\_ekf = \begin{bmatrix} 1.0000 & -0.3003 & -0.9724 \\ -0.3003 & 1.0000 & 0.3324 \\ -0.9724 & 0.3324 & 1.0000 \end{bmatrix}$$

Table 5.4: Sensitivity results for 100 ensembles

	theta	**%CRB**	$\hat{R}$
NR	3.9990 0.4002 0.6024	0.5933 1.0201 10.5026	0.0010 , 0.0039
EKF	4.0001 0.4003 0.5979	0.5780 0.9779 10.0342	0.0009 , 0.0038

$$theta\_corrcoef\_nr = \begin{bmatrix} 1.0000 & -0.2041 & -0.9072 \\ -0.2041 & 1.0000 & 0.3993 \\ -0.9072 & 0.3993 & 1.0000 \end{bmatrix}$$

$$theta\_corrcoef\_ekf = \begin{bmatrix} 1.0000 & -0.1819 & -0.9066 \\ -0.1819 & 1.0000 & 0.3716 \\ -0.9066 & 0.3716 & 1.0000 \end{bmatrix}$$

Observation : As the number of ensemble runs are increased the correlation coefficient of the EKF estimated parameters and that of the NR method matches with each other.

Chapter 6

EKF tuning with process noise

In this chapter we will see an extended version of the tuning algorithm for a non zero process noise case. The process noise has been regarded as the most notorious unknown statistic in a Kalman filtering problem. In most of the literature the process noise covariance matrix, Q is estimated assuming that the parameters (θ) of the model is known. A simple heuristic method of estimating Q is the Myers and Tapley(MT) algorithm [7]. In a non-linear system with unknown parameters (θ) the MT algorithm gives a very high and oscillating estimates of the process noise covariance matrix with the diagonal elements being forced to positive value whenever it becomes negative. Another method of estimating the process noise was suggested by Vallapil and Georgakis [10] which gives a very low estimate than expected. Hence we need a different approach to solve the process noise problem. The low and high process noise case are treated separately in the following sections.

6.1 MT method for low process noise

In this section we will try to solve the problem for a low process noise case(Q<0.1R) in a heuristic intuitive way. Let us first refresh the basic covariance propagation equation as explained in section 2.2.1. The state propagation and the corresponding prior covariance propagation equation is given by

$$\hat{X}_{k|k-1} = f(\hat{X}_{k-1|k-1})$$
$$\hat{P}_{k|k-1} = F_{k,k-1}\hat{P}_{k-1|k-1}F^T_{k,k-1} + Q$$

where $F_{k,k-1}$ is the state Jacobian matrix and Q is the process noise covariance matrix. The following sections discuss how the prior covariance propagation equation vary with balanced state propagation equation.

6.1.1 Ideal case with true process

Let us assume an ideal case with known process $\omega(k)$ for k=1,2...N. The balanced state propagation equation can be written as

$$\hat{X}_{k|k-1} = f(\hat{X}_{k-1|k-1}) + \omega(k) \tag{6.1}$$

Following the same steps as in section 2.2.1 we get the equation given by

$$\hat{P}_{k|k-1} = E\left[\{f(\hat{X}_{k-1|k-1}) + F_{k,k-1}\tilde{X}(k-1) + \omega(k) - f(\hat{X}_{k-1|k-1}) - \omega(k)\}\ \{\}^T\right]$$
$$\hat{P}_{k|k-1} = E\left[\{F_{k,k-1}\tilde{X}(k-1)\}\ \{F_{k,k-1}\tilde{X}(k-1)\}^T\right]$$
$$\hat{P}_{k|k-1} = F_{k,k-1}\hat{P}_{k-1|k-1}F^T_{k,k-1} + 0$$

The above equation suggests that the EKF can be run with Q=0 even if there is a process model mismatch simply by balancing the state propagation equation as per Eq-(6.1). Unfortunately we do not have the process noise samples and a more practical case is discussed in the next section.

6.1.2 Practical case with estimated process

Now let us balance the state propagation equation by estimates of ω (first order moments) as suggested by Myers and Tapley which is

$$\hat{\omega}_{MT}(k) = \hat{X}_{k|k} - \hat{X}_{k-1|k}$$

By balancing the state propagation equation we get

$$\hat{X}_{k|k-1} = f(\hat{X}_{k-1|k-1}) + \hat{\omega}_{MT}(k) \tag{6.2}$$

Let us discard the subscript 'MT' and denote $\hat{\omega}_{MT}$ as $\hat{\omega}$. The corresponding covariance equation is given by

$$\hat{P}_{k|k-1} = E\left[\{f(\hat{X}_{k-1|k-1}) + F_{k,k-1}\tilde{X}(k-1) + \omega(k) - f(\hat{X}_{k-1|k-1} - \hat{\omega}(k))\}\ \{\}^T\right]$$
$$\hat{P}_{k|k-1} = E\left[\{F_{k,k-1}\tilde{X}(k-1) + \omega(k) - \hat{\omega}(k)\}\ \{F_{k,k-1}\tilde{X}(k-1) + \omega(k) - \hat{\omega}(k)\}^T\right]$$
$$\hat{P}_{k|k-1} = E\left[\{F_{k,k-1}\tilde{X}(k-1) + \tilde{\omega}(k)\}\ \{F_{k,k-1}\tilde{X}(k-1) + \tilde{\omega}(k)\}^T\right]$$

where $\tilde{\omega}(k) = \omega(k) - \hat{\omega}(k)$ and also we know that

$$E\left[\tilde{X}(k-1)\tilde{X}(k-1)^T\right] = \hat{P}_{k-1|k-1}$$
$$E\left[\tilde{X}(k-1)\tilde{\omega}(k)^T\right] = 0$$

Thus we get

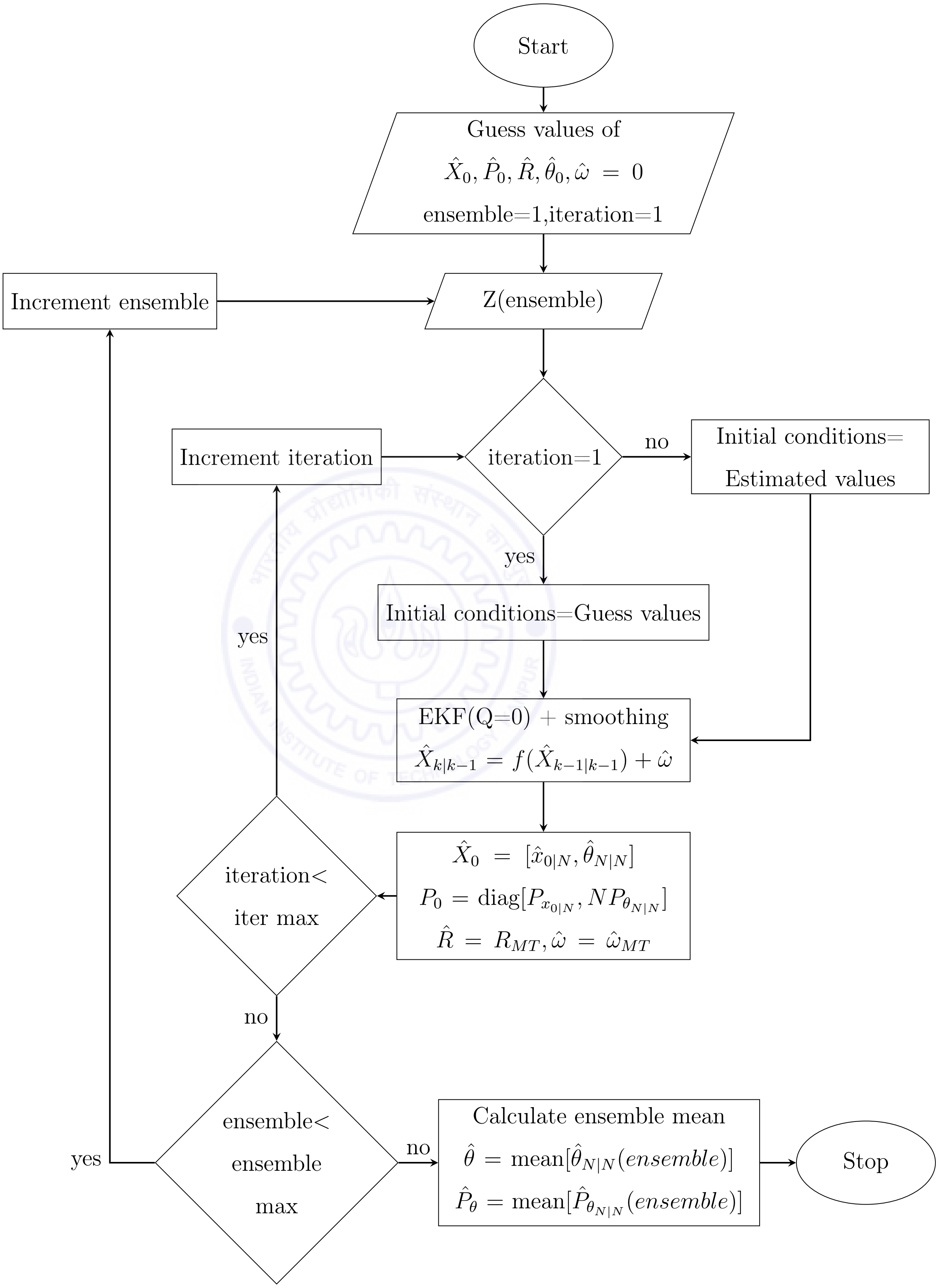
$$P_{k|k-1} = F_{k,k-1}\hat{P}_{k-1|k-1}F_{k,k-1}^T + Q_k$$
$$Q_k = E\left[\tilde{\omega}(k)\tilde{\omega}(k)^T\right]$$

Assuming that the estimates of the process noise samples $\hat{\omega}(k)$ are very close to the true process or the actual process ie $\tilde{\omega}(k) \rightarrow 0$ as $\hat{\omega}(k) \rightarrow \omega(k)$ we can still run the EKF with Q≈0 as run in the ideal case earlier. What we have achieved by balancing the state propagation equation is the fact that we can supress the more unstable second order moments (Q) with the help of the estimates of process noise samples ($\hat{\omega}$). The EKF predict step equations now becomes

$$\left.\begin{aligned} \hat{X}_{k|k-1} &= f(\hat{X}_{k-1|k-1}) + \hat{\omega}_{MT}(k) \\ P_{k|k-1} &= F_{k,k-1}\hat{P}_{k-1|k-1}F_{k,k-1}^T + 0 \end{aligned}\right\} \tag{6.3}$$

A flow chart describing the above said tuning procedure is shown below

Adaptive EKF flowchart for Q<0.1R



6.1.3 Results using MT method

The above tuning algorithm was studied on a SMD system described in section-5.1.2. For 50 ensembles each with 10 iterations, R=diag(0.001,0.004) and Q=0.001× R, the initial conditions were chosen as $X_0 = [1.0, 0.0, 5, 0.5.0.5]^T, P0\_power = 0$ and $R0\_power$=0. Innovations were used in estimating R with $last\_n$=N/2 and MBF smoothing was used for which we get the results as tabulated below

Table 6.1: SMD system results for very low process noise

	$\hat{\theta}$	$\%CRB$	$\hat{R}$
REF	4.0, 0.4, 0.6	1.0217, 1.5357, 14.0774	0.001,0.004
EKF	4.005, 0.4003, 0.586	0.5799, 0.9842, 10.0811	0.001,0.0038

Note : For Q>0 case the reference results is no more from the Newton Raphson method which cannot differentiate between the measurement noise and process noise but the true values of the parameter and the injected measurement noise, (m_noise)(variable as used in GUI in section-A.2) will act as the reference results(REF). When Q≠0 then the reference $\%CRB$ is evaluated as follows

$$\%CRB = \frac{\sigma_{\hat{\theta}}}{\theta} \times 100$$

where $\sigma_{\hat{\theta}}$ is the standard deviation of the estimated parameters over many ensembles.

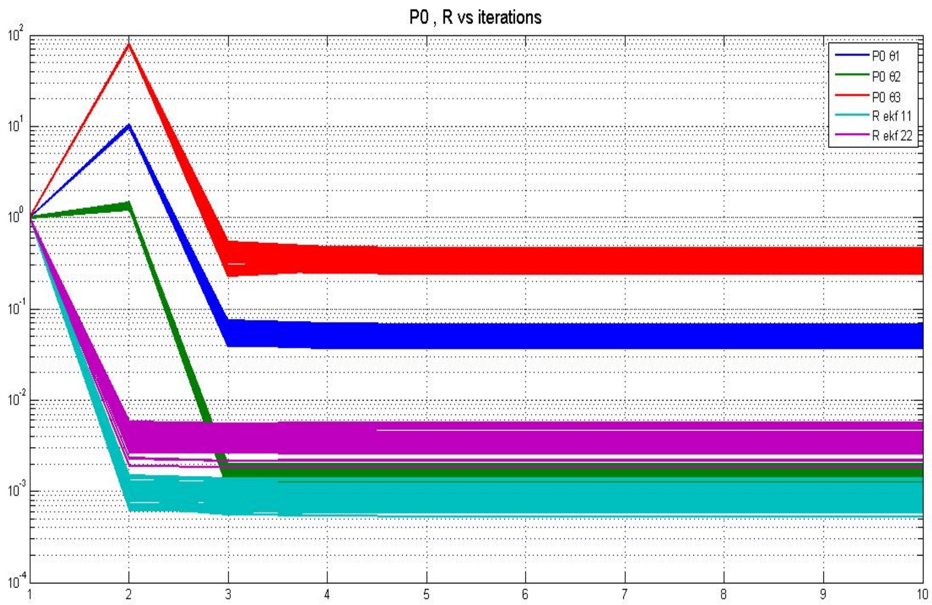


Figure 6.1: EKF-Low process noise , Noise Statistics vs iterations

6.1.4 Is process noise really a noise ?

It is quite strange to note that the above algorithm estimates the process noise and add them back into the EKF to get filter convergence when the sole purpose of a filter is to remove noise and not add them.A question that remains is whether the process noise is really a noise ? Generally 'noise' is a term associated with unwanted signal and 'filter' is a term associated with something that removes unwanted signal. In the context of Kalman filter the term 'filter' is used since it reduces the unwanted measurement noise and not process noise. If Kalman filter has to reduce process noise then it would have tracked the deterministic trajectory X_d and not the stochastic trajectory X.The term 'noise' is used in conjecture with the word process since its statistics are often unknown which can be represented as a Gaussian random process $\omega(k) \sim \mathcal{N}(0, Q)$. The process noise is not an unwanted signal and infact its something inherent in the system whose statistical information is needed for the Kalman filter to converge in the form of first or second order moments. In almost all literature the first order moments $\omega(k)$ are assumed to be zero since $E\left[\omega(k)\right]$=0 and the second order moments $E\left[\omega(k)\omega(k)^T\right]$=Q is estimated using different techniques where as in the above method the other way round.

6.2 DSDT method for high process noise

The following section explains a new method of estimating process noise matrix, Q using the difference between the stochastic and the deterministic trajectory as a measure of the process noise. We call this scheme as DSDT method which is the abbreviation for 'Difference between the Stochastic and Deterministic Trajectory'.

Generation of deterministic trajectory : A non-linear system with zero process noise can be modelled as

$$X_d(k) = f(X_d(k-1))$$

By 1st order Taylor series approximation we get

$$X_d(k) = f(X_n(k-1)) + f^{'}(X_n(k-1))\tilde{X}_d(k-1)$$

X_n is the nominal point and let us denote $f^{'}(X_n(k-1))$ as $F_{k,k-1}$ to get

$$X_d(k) = f(X_n(k-1)) + F_{k,k-1}\tilde{X}_d(k-1) \tag{6.4}$$

where the subscript 'd' indicates the deterministic trajectory of the process,

$$F_{k,k-1} = f^{'}(X_n(k-1)) = \left[\frac{\partial f}{\partial X_d}\right]_{X_d=X_n(k-1)}$$
$$\tilde{X}_d(k-1) = X_d(k-1) - X_n(k-1)$$

Generation of Stochastic trajectory : A non-linear system with process noise can be modelled as

$$X(k) = f(X(k-1)) + \omega(k)$$

It is assumed that the nominal point X_n taken at any discrete time instant 'k' for both the deterministic and stochastic trajectory is same.By 1st order Taylor series approximation we get

$$X(k) = f(X_n(k-1)) + f^{'}(X_n(k-1))\tilde{X}(k-1) + \omega(k)$$
$$X(k) = f(X_n(k-1)) + F_{k,k-1}\tilde{X}(k-1) + \omega(k) \tag{6.5}$$

where $\tilde{X}(k-1) = X(k-1) - X_n(k-1)$.The stochastic trajectory can take any direction depending upon the process noise samples $\omega(k)$ ~N(0,Q).

6.2.1 Process noise sample reconstruction

Given the non linear function f,the stochastic trajectory(X) and the deterministic trajectory(X_d),the process noise samples at any discrete time instant 'k' can be reconstructed assuming $\omega(0) = 0$ and $X_d(0) = X(0) = X_0$ which is given by

$$\omega(k) = X(k) - X_d(k) - \sum_{j=1}^{k-1} F_{k,j}\omega(j) \tag{6.6}$$

where $F_{k,j} = f^{'}(X_n(k-1)) \times f^{'}(X_n(k-2)) \times .....f^{'}(X_n(j))$

6.2.2 Proof by Mathematical Induction

In this section let us prove Eq-(6.6) by the method of Mathematical Induction.

When k=1

To prove that Eq-(6.6) is true for k=1.

$\omega(1) = X(1) - X_d(1)$; from Eq-(6.4)

$\omega(1) = f(X_n(0)) + F_{1,0}\tilde{X}(0) + \omega(1) - [f(X_n(0)) + F_{1,0}\tilde{X}_d(0)]$; from Eq-(6.6),(6.5)

$\omega(1) = \omega(1)$; $X_d(0) = X(0)$ and $\tilde{X}_d(0) = \tilde{X}(0)$

When k=n

Assume that Eq-(6.6) is true for k=n where 1≤n≤ N and we get

$$\omega(n) = X(n) - X_d(n) - \sum_{j=1}^{n-1} F_{n,j}q(j) \tag{6.7}$$

When k=n+1

To prove that Eq-(6.6) is true for k=n+1 assuming its true for k=n

$$X(n) = X_d(n) + q(n) + \sum_{j=1}^{n-1} F_{n,j}\omega(j) \text{ ; from Eq-(6.7)}$$

Subtracting $X_n(n)$ and pre-multiplying both sides with $F_{n+1,n}$ we get

$$F_{n+1,n}[X(n) - X_n(n)] = F_{n+1,n}[X_d(n) - X_n(n)] + F_{n+1,n}\omega(n) + \sum_{j=1}^{n-1} F_{n+1,n}F_{n,j}\omega(j)$$

from Eq-(6.4),(6.5) we get

$$X(n+1) - \omega(n+1) - f(X_n(n)) = X_d(n+1) - f(X_n(n) + \sum_{j=1}^{n-1} F_{n+1,j}\omega(j) + F_{n+1,n}\omega(n)$$

$$X(n+1) - q(n+1) = X_d(n+1) + \sum_{j=1}^{n} F_{n+1,j}\omega(j)$$

$$\omega(n+1) = X(n+1) - X_d(n+1) - \sum_{j=1}^{n} F_{n+1,j}\omega(j)$$

Thus Eq-(6.6) is true for k=n+1 and hence proved by mathematical induction.

6.2.3 Recursive equation for faster processing

Consider the third term in Eq-(6.6) and if there are N=100 samples , to reconstruct q(100) we would need 100-1=99 summation terms and each j^{th} summation term needs 100-j multiplications $F_{100,j} = F_{100,99}F_{99,98}...F_{j+1,j}$ making the estimation process highly computational. A solution to this is to frame a recursive equation of the third term. The third term can be written as

$$q(k) = \sum_{j=1}^{k-1} F_{k,j}\omega(j) \tag{6.8}$$

$$q(k) = \left[\sum_{j=1}^{k-2} F_{k,j}\omega(j)\right] + F_{k,k-1}\omega(k-1)$$

$$q(k) = \left[\sum_{j=1}^{k-2} F_{k,k-1}F_{k-1,k-2}...F_{j+1,j}\omega(j)\right] + F_{k,k-1}\omega(k-1)$$

$$q(k) = F_{k,k-1}\left[\sum_{j=1}^{k-2} F_{k-1,j}\omega(j) + \omega(k-1)\right]$$

$$\text{wkt } q(k-1) = \sum_{j=1}^{k-2} F_{k-1,j}\omega(j)$$

$$q(k) = F_{k,k-1}\left[q(k-1) + \omega(k-1)\right] \tag{6.9}$$

Thus at any instant of time 'k' the third term has only two summation terms each with just one multiplication term as per Eq-(6.9) which otherwise would be k-1 and k-j respectively as per Eq-(6.8). Thus we get almost exact reconstruction of q from Eq-(6.6) using the recursive Eq-(6.9). Even though the expectation of the third term E$\{q(k)\}$ is zero these terms cannot be avoided in the process noise estimation since the stochastic state at a time instant 'k' depends on the process noise samples $\omega(j)$ for all time instants 1≤j≤k-1 where both $\omega(0)$ and q(0) is zero.

6.2.4 Q estimation by DSDT method

Our problem is to track X using the model that generated X_d by injecting appropriate Q into the EKF when the parameters of the system model $F_{k,k-1}$ associated with X_d are unknown. The new method of estimation of process noise samples as per Eq-(6.6) have the following unknowns and assumptions

- X is unknown which can be substituted by the posterior estimate ($\hat{X}_{k|k}$) and its assumed that $\hat{X}_{k|k} \to X(k)$.
- X_d is unknown since initial state $X_{0|0}$ is unknown.The deterministic trajectory is generated with the initial nominal point $X_n(0) = \hat{X}_{0|0}$ and its assumed that $\hat{X}_{0|0} \to X(0)$.
- The parameters of the state transition matrix $F_{k,k-1}$ is unknown and hence an estimate of the deterministic state transition matrix $\hat{F}_{k,k-1}$ based on the latest parameter estimated during the iterative procedure is used for the estimation of Q assuming that $\hat{F}_{k,k-1} \to F_{k,k-1}$ as $\hat{\theta} \to \theta$.

Thus an estimate of process noise samples is given by

$$\hat{\omega}(k) = \hat{X}_{k|k} - \hat{X}_d(k) - \sum_{j=1}^{k-1} F_{k,j}\hat{\omega}(j) \tag{6.10}$$

The above three assumptions are valid since we make many iterative runs along with smoothing. This lead us to the approximation $\omega(k) \approx \hat{\omega}$(k) and we get

$$\hat{Q}_{DSDT} = \frac{1}{N-1}\sum_{k=1}^{N}\left[\{\hat{\omega}(k)-\bar{\omega}\}\{\hat{\omega}(k)-\bar{\omega}\}^{T}\right] \tag{6.11}$$
$$\bar{\omega} = \frac{1}{N}\sum_{k=1}^{N}\hat{\omega}(k)$$

Unlike Myers and Tapley (MT) method which measures independent estimates of the process noise sample, our new estimate of the process noise samples at discrete time 'k' denoted as $\hat{\omega}(k)$ depends on all the previous estimates of the process noise samples $\hat{\omega}(l)$ where 1≤l≤k-1. Hence it will be too cumbersome to estimate the process noise covariance matrix Q by the covariance matching technique as suggested by Myers and Tapley which is now estimated as sample covariance as per EQ-(6.11). Due to the three assumptions, the DSDT method works fine only for high process noise case , $Q \geq 0.1R$. Also the results are good when only last few samples, typically last 50% of the samples are considered in estimating Q as a sample covariance of $\hat{\omega}$. **NOTE:** If Q=0 then both the stochastic (X) and the deterministic (X_d) trajectory are same and thus R can be estimated using the sample covariance of (z-HX_d).

6.2.5 Results using DSDT method

Tuning of X_0, P_0,R and Q is done as per the below flow chart. Since the simultaneous estimation of Q and R is difficult we ignore the second terms in the MT equation which otherwise would make the diagonal values of R as negative. The estimated R now becomes higher than expected and is lowered by using a factor(ρ_f) in the interval (0,1) which is robust. The optimum value of this factor is obtained my minimizing the cost function(J_{ρ_f})[19] given by

$$J_{\rho_f} = \frac{1}{N}\sum_{k=1}^{N} r(k)' S_k^{-1} r(k) + ln(|S_k|)$$

where r is the innovation samples and S is the innovation covariance. Fig-6.2 shows the variation of J_{ρ_f} with the ρ_f or R_factor(as used in GUI in section-A.2) and the subsequent figures show the plots obtained for a typical case of Q=R. Many sensitivity studies are conducted to check the robustness of the algorithm and its observed that we get consistent result with the estimated parameter covariance(P_θ) and the numerically evaluated variance of the estimated parameters. The default input values chosen for the sensitivity study on the SMD system are given below and the results are tabulated in 6.2.

$$
\begin{aligned}
N &: 100 \\
dt &: 0.1 \\
m\_noise &: [0.001, 0.004] \\
p\_noise &: [0.001, 0.004] \\
ensemble\_max &: 50 \\
iter\_max &: 20 \\
r\_option &: 1 - \text{using innovations} \\
q\_option &: 2 - \text{using posterior estimates} \\
smooth\_option &: 2 - \text{MBF smoothing} \\
last\_n &: N/2 \\
Scale\_up &: N \\
P0\_power &: 0 \\
R0\_power &: 0 \\
Q0\_power &: 0 \\
x0\_initial &: [1, 0, 5, 0.5, 0.5]
\end{aligned}
$$

The input variations for the sensitivity analysis are as follows

$$
\begin{array}{lll}
Study-1 & X0\_initial & : [1, 0, 10, 1, 1], [1, 0, 10, -1, -1], [1, 0, -10, 1, 1] \\
Study-2 & P0\_power & : -5, 1, 5 \\
Study-3 & Scale\_up & : N/10, 10 * N, 100 \times N \\
Study-4 & R0\_power & : -2, 1, 2 \\
Study-5 & p\_noise & : 0.1, 1, 10 \times m\_noise \\
Study-6 & N & : 50, 100, 1000
\end{array}
$$

Table 6.2: Sensitivity Analysis
N-negligible$\leq$ 5% error , M-medium$\leq$ 15% error , L-Large$>$ 15% error

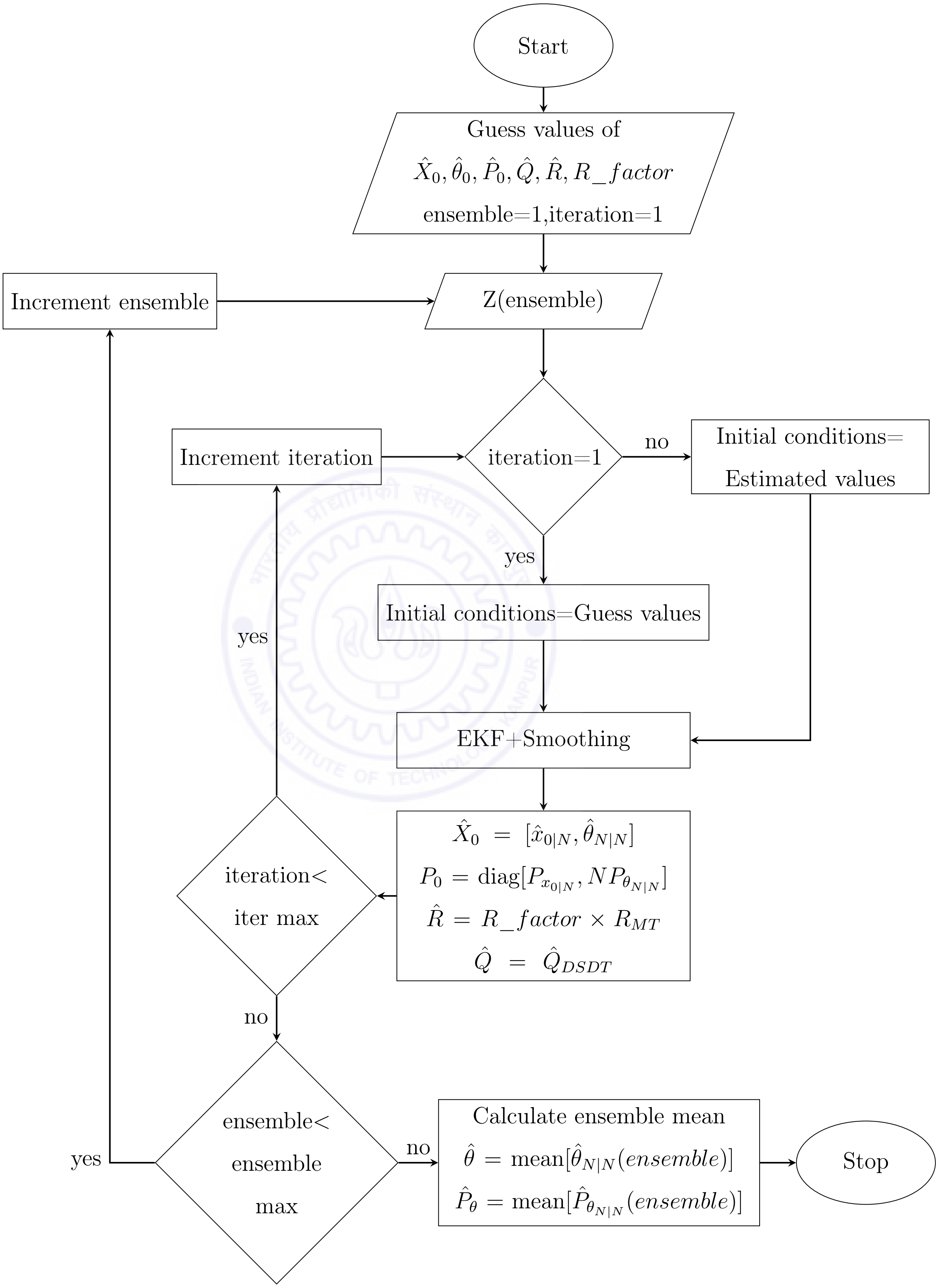
Study no.	% error in $\hat{\theta}$			% error in P_θ		
	θ_1	θ_2	θ_3	CRB_1	CRB_2	CRB_3
1	N,N,N	N,N,N	M,M,M	L,L,L	N,N,N	N,N,N
2	N,N,N	N,N,N	M,M,M	M,M,M	M,M,M	N,N,N
3	N,N,N	M,M,M	N,N,N	M,M,M	M,M,L	N,N,N
4	N,N,N	M,M,M	M,M,L	N,N,N	M,M,M	N,N,N
5	N,N,N	N,N,N	M,N,M	N,N,M	N,N,M	N,N,N
6	N,N,N	M,M,N	M,M,N	M,N,N	L,M,N	L,N,N

Sensitivity Analysis contd.
N-negligible$\leq$ 5% error , M-medium$\leq$ 15% error , L-Large$>$ 15% error

Study no.	% error in $\hat{R}$		% error in $\hat{Q}$	
	R_{11}	R_{22}	Q_{11}	Q_{22}
1	N,N,N	N,N,N	N,N,N	M,M,M
2	N,N,N	N,N,N	N,N,N	N,N,N
3	N,N,N	N,N,M	N,N,N	N,N,M
4	N,N,N	N,N,N	N,N,N	N,N,N
5	N,N,N	M,N,M	M,N,N	N,N,N
6	N,N,N	N,N,N	M,N,N	M,N,N

Note : The scalar factor ρ_f or R_factor change with the values of the actual injected process noise(p_noise). The optimum values of ρ_f obtained for $p\_noise = 0.1, 1, 10 \times m\_noise$ is 0.68,0.38 and 0.08 respectively.

Adaptive EKF flowchart for Q≥0.1R



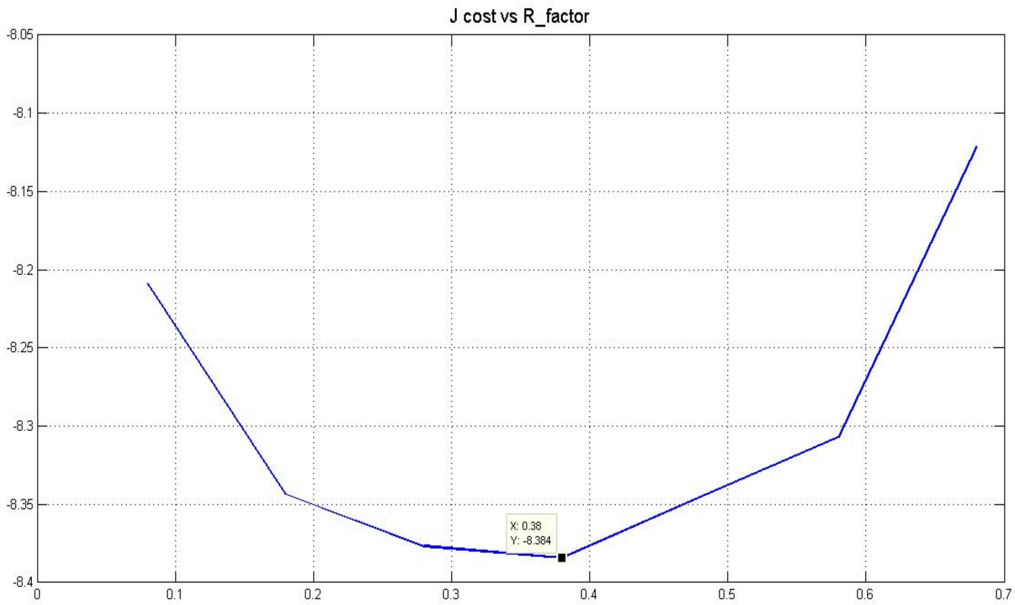


Figure 6.2: EKF-High process noise(Q=R) , Cost J_{ρ_f} vs R_factor

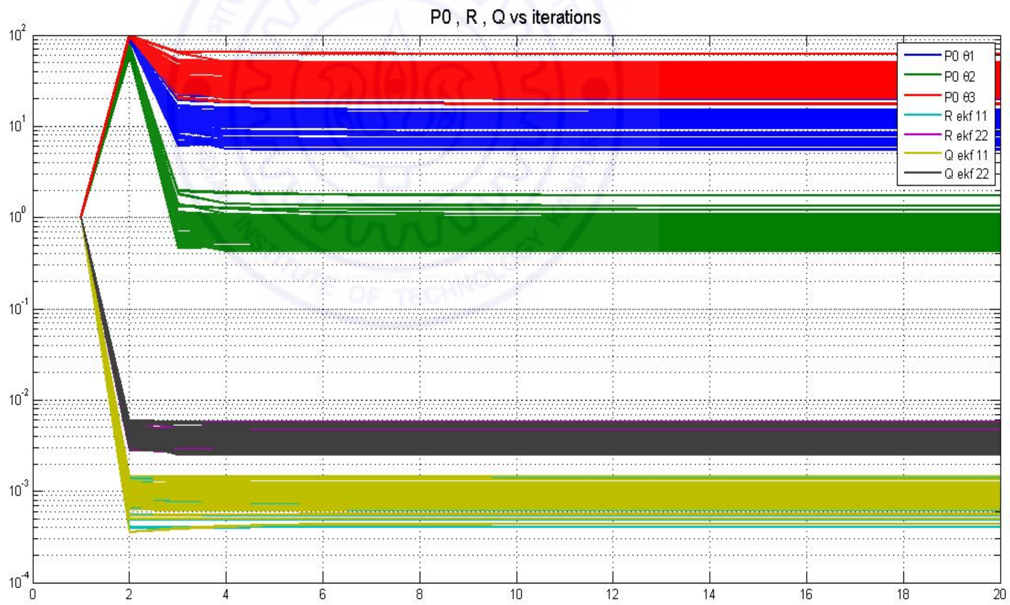


Figure 6.3: EKF-High process noise(Q=R) , Noise statistics vs iteration

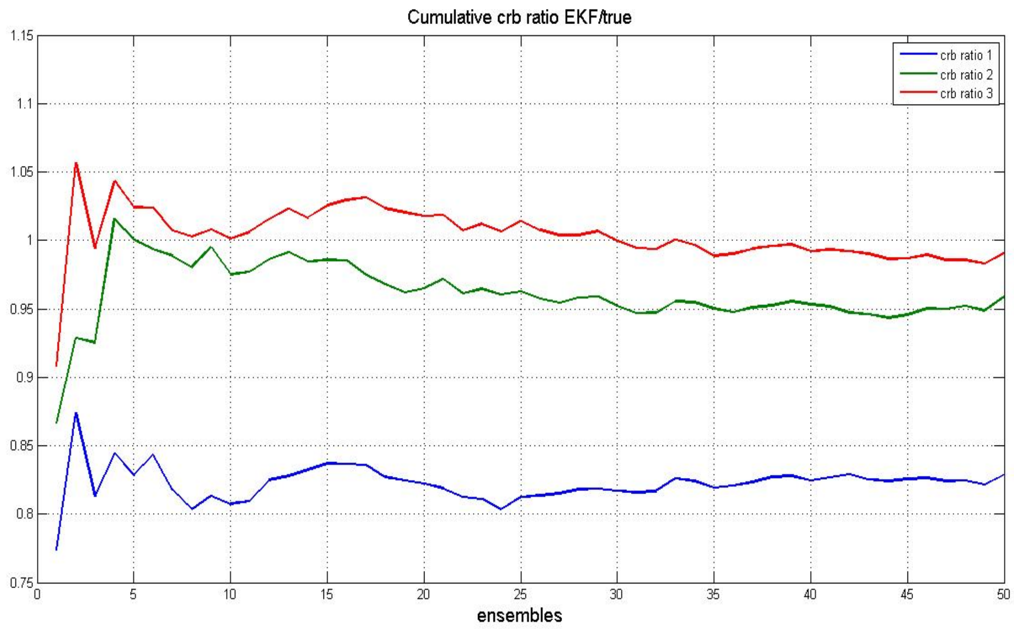


Figure 6.4: EKF-High process noise(Q=R) , Cumulative CRB ratio

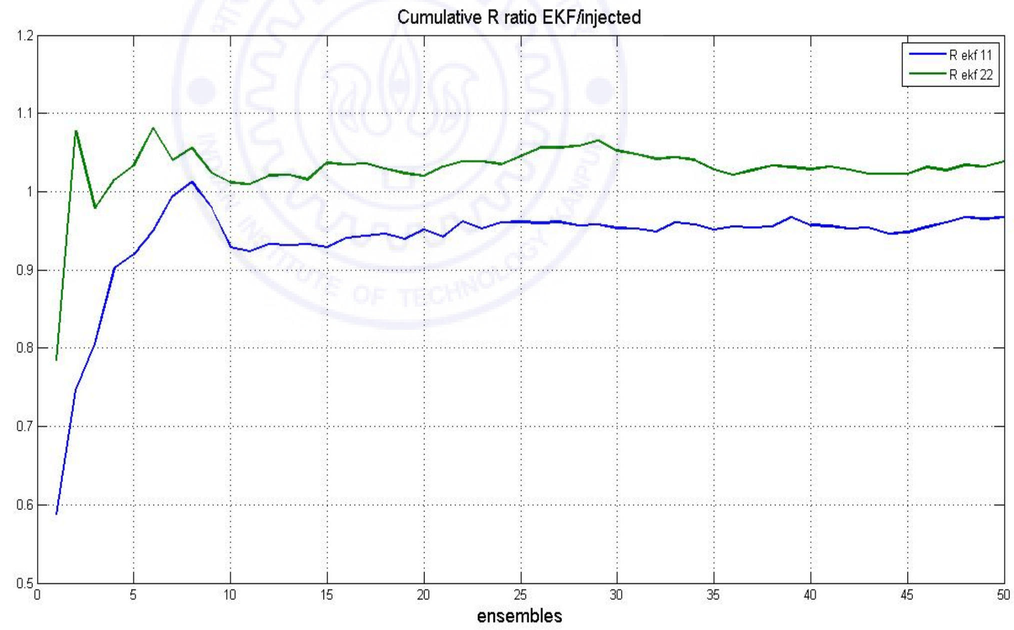


Figure 6.5: EKF-High process noise(Q=R) , Cumulative R ratio

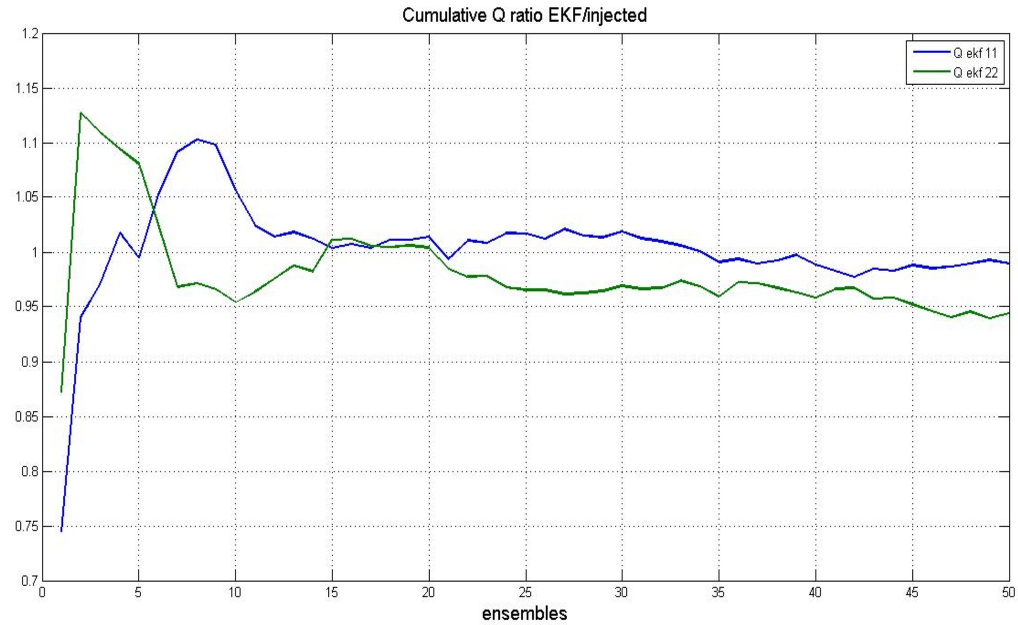


Figure 6.6: EKF-High process noise(Q=R) , Cumulative Q ratio

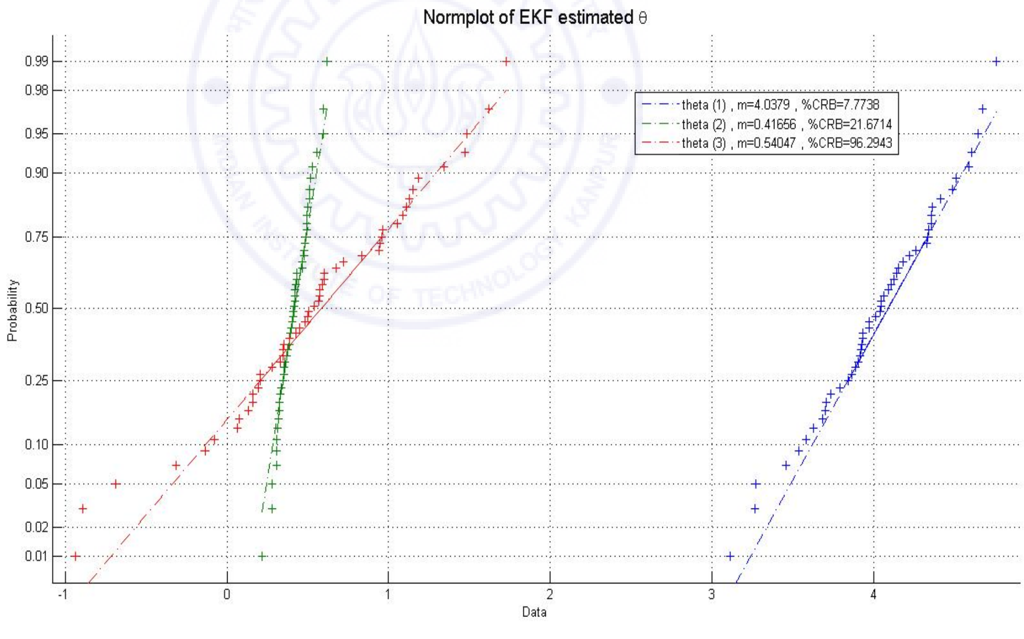


Figure 6.7: EKF-High process noise(Q=R) , Normplot of $\hat{\theta}$

Conclusions and Future Work

An iterative scheme for the parameter estimation problem using adaptive EKF is proposed which can achieve Cramer Rao Bound criteria in presence of unknown noise statistics. The parameter covariance which has a direct impact on the CRB's are tuned using virtual measurement concept. The histograms show that the estimated parameters are distributed about the true value and its CRB's are comparable to the MMLE using Newton Raphson optimisation scheme as shown in table- 5.1 and 5.2. Different sensitivity studies are conducted to conclude that the results are unique and independent of the initial guess values of X_0,P_0 and R. The sensitivity study showed that the initial guess value of the parameter can go upto $\pm500\%$ of the true value. Its observed that the results do not change over a wide range of P_0 chosen between 10^{-5} to 10^{5}. As the iterations proceed statistical equilibrium is achieved as seen in Fig-5.3 and 5.7. Monte Carlo runs of several ensembles of measured data yied high histogram results of the parameter estimates about the true value as seen in Fig-5.10,5.11 and 5.14. An user-friendly MATLAB Graphic User Interface(GUI) software is developed which can be used to estimate parameter of any unknown system and automatically create an excel report of the sensitivity study conducted. The high process noise problem was solved using the a new DSDT method and sensitivity studies were conducted which showed negligible error of the statistics and θ in most cases. Consistent results were obtained with the estimated parameter covariance(P_θ) and the numerically evaluated variance of the estimated parameters. As a future work the process noise can be augmented as additional unknown states using the samples from DSDT method which can be a new technique in solving process noise problem.

Using Eq-(6.9) in Eq-(6.10) we get

$$\hat{\omega}(k) = \hat{X}_{k|k} - \hat{X}_d(k) - F_{k,k-1}\left[\hat{q}(k-1) + \hat{\omega}(k-1)\right] \tag{6.12}$$

$$\hat{q}(k-1) = \sum_{j=1}^{k-2} F_{k-1,j}\hat{\omega}(j)$$

The above non-linear model Eq-(6.12) can be used for the state propagation of process noise samples if we are augmenting the process noise samples as unknown states in which case the state Jacobian needs to be worked out.

Appendices

INDIAN INSTITUTE OF TECHNOLOGY KANPUR

Appendix A

MATLAB Graphic User Interface

In this section we will discuss the creation of a graphical user interface (GUI) of the various inputs that be used to test the proposed algorithm for its efficiency and consistency. The MATLAB GUI creates a user friendly environment for running the program with different control inputs as well create an automatically drafted report of the experiment conducted. The excel report can quickly take the user through the details of the input and output of the algorithm. Basically GUI [20] is simply a figure window with user defined components for control over the program. The control components can be built using GUIDE (GUI Development Environment) layout editor tool. GUIDE enables the user to create GUI without writing scripts or commands and it creates an automatic MATLAB program that is linked with the GUI figure. This automatic program when run opens the GUI figure for further tasks. It can also be modified to incorporate several other routines or tasks defined by the user. The GUI components has one or more callback functions and by virtue of its name it calls back MATLAB for performing the instruction defined in the function. This type of programming is event driven and asynchronous and event external to the software triggers the callback execution.

A.1 GUI Components

The different components used in building the GUI has many properties which can be edited in the property inspector.Some of them are explained below-

- Uipanel :- Its a panel showing a group of buttons that are aligned together.In other words they are containers that arrange GUI components into a group. For ex: 1. in the GUI layout figure is a uipanel that has four different options of selecting the case study or system on which the algorithm has to be run.

- Radio button :- It can be considered as a toggle switch used to select or deselect different options with only one option being selected at once. For ex: 2. in the GUI layout figure is a radio button to conduct the case study on a nonlinear SMD system.

- Static text :- It is used to display text that cannot be changed by the user.They are typically used for naming GUI components. For ex: 3. in the GUI layout figure is a static text to set the number of ensembles and iterations to be used. It also has a counter below displaying the current ensemble and iteration being executed by the program which are reset to zero in the beginning.

- Edit text :- It is a text box to give input to the program using characters that are typed by the end user.For ex: 4. in the GUI layout figure is an edit text box to input the diagonal values of the measurement noise separated by commas.

- Pop-Up menu :- When clicked , it opens up a list of items stored in its handle 'string' to be chosen by the user. For ex: 5. in the GUI layout figure is a popup menu to choose what should be used in estimating the measurement noise matrix R , whether innovations or residue or smoothed data.

- Push button :- Its is used to execute a set of intructions defined in its callback function when the user clicks on it. For ex: 6. in the GUI layout figure is a push button to create an automatically drafted excel report after the sensitivity study is conducted.

- Uitable :- Its a simple table with pre-defined number of rows and columns which is used generally to display outputs. For ex: 7. in the GUI layout figure is a Uitable that shows the estimated values as columns and the rows shows the results separately for EKF and NR method.

A.2 GUI Layout Figure

Bibliography

[1] Kalman R.E , "A New Approach to Linear Filtering and Prediction Problems". *Journal of Basic Engineering* , 1960

[2] Dan Simon , "Optimal State estimation - Kalman,H_∞ and non linear approaches". *Wiley Interscience (pg-395-425)* , 2006

[3] Crassidis J.L and Junkins J.L ,"Optimal estimation of dynamic systems". *CRC Press,* 2^{nd} *Edition (pg-575-600)* ,2012.

[4] Steven M. Kay , "Fundamentals of Statistical Signal Processing : Estimation Theory". *Prentice Hall Signal Processing Series Volume 1 (pg-27-35)* , 2010

[5] R.M.O. Gemson and M.R. Ananthasayanam, "Importance of initial state covariance matrix for the parameter estimation using adaptive extended Kalman filter", *AIAA JOURNAL,A98-37215* , 1998.

[6] M.R. Ananthasayanam , H.S Suresh and M.R Muralidharan ,"GUI based software for teaching parameter estimation technique using MMLE". *Report 2001 FM 1,Project-JATP/MRA/053,IISc-Bangalore* , 2001

[7] Kenneth A Myers and Bryon D Tapley ,"Adaptive Sequential estimation with unknown noise statistics", *IEEE transactions on automatic control*, 1976.

[8] H. E. Rauch,F. Tung and C. T. Striebel, "Maximum Likelihood Estimates of Linear Dynamic Systems", *AIAA Journal,VOL. 3, NO.-8* , August 1965.

[9] Bierman G.J , "A new computationally efficient fixed-interval, discrete-time smoother". *Decision and Control including the Symposium on Adaptive Processes, 1981 20th IEEE Conference* ,1981.

[10] Valappil J and Georgakis C , "Systematic Estimation of State Noise Statistics for Extended Kalman Filters." *AIChe Journal Vol 46, No. 2* , February 2000.

[11] Vinay A Bavdekar , Anjali P Deshpande , Sachin C Patwardhan , "Identification of process and measurement noise covariance for state and parameter estimation using extended Kalman filter". *Journal of Process control* , 2011

[12] B.J. Odelson, A. Lutz, J.B. Rawlings, "The autocovariance-least squares method for estimating covariances: application to model-based control of chemical reactors".*IEEE Trans. Control Syst. Technol. 14* , 2006.

[13] R.K. Mehra, "On the identification of variances and adaptive Kalman filtering, *IEEE Trans. Automat. Control*, 1970.

[14] P.R. Belanger, "Estimation of noise covariances for a linear time-varying stochastic process". *Automatica 10* , 1974.

[15] B.J. Odelson, M.R. Rajamani, J.B. Rawlings, "A new autocovariance least-squares method for estimating noise covariances", *Automatica 42*, 2006.

[16] A.P. Dempster, N.M. Laird, D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm". *J. R. Stat. Soc. B 3*, 1977.

[17] R.H. Shumway, D.S. Stoffer, "Time Series Analysis and its Applications". *Springer-Verlag New York Inc., New York* , 2000.

[18] H. Raghavan, A.K. Tangirala, R.B. Gopaluni and S.L. Shah ,"Identification of chemical processes with irregular output sampling".
Control Eng. Pract. 14 , 2006.

[19] Maj Ashwin Yadav , "Innovations cost function based tuned and constant gain Kalman Filter approaches to target tracking in wireless sensor networks". *Thesis submiited to EE, IITK (pg-15)* , June 2012

[20] www.mathworks.in/help/pdf_doc/matlab/buildgui.pdf, "Creating Graphical User Interface".*Revised for Matlab 8.3 , Release-2014a* , 2014