Optimal Nonlinear Adaptive Observers for State, Parameter and Fault Estimation

by

Amir Valibeysi

B.Sc., Amirkabir University of Technology, 2013

Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Applied Science

in the
School of Mechatronic Systems Engineering
Faculty of Applied Sciences

© Amir Valibeysi 2015
SIMON FRASER UNIVERSITY
Summer 2015

All rights reserved. However, in accordance with the Copyright Act of Canada, this work may be reproduced, without authorization, under the conditions for “Fair Dealing.” Therefore, limited reproduction of this work for the purposes of private study, research, criticism, review and news reporting is likely to be in accordance with the law, particularly if cited appropriately.
Approval

Name: Amir Valibeygi
Degree: Master of Applied Science
Title: Optimal Nonlinear Adaptive Observers for State, Parameter and Fault Estimation

Examinining Committee: Chair: Jason Wang
Assistant Professor

Krishna Vijayaraghavan
Senior Supervisor
Assistant Professor

Siamak Arzanpour
Supervisor
Associate Professor

Mehrdad Moallem
Examiner
Professor
School of Mechatronic Systems Engineering

Date Defended/Approved: July 30, 2015
Abstract

The demand for reliable, fast and robust techniques for detection and estimation of faults in real-world systems is constantly increasing. A nonlinear adaptive observer for state, parameter and fault estimation is developed in this work using an optimal approach. This observer is capable of estimating unknown parameters as well as sensor faults in the system. The proposed observer also accounts for existing noise and disturbance in the system. By defining appropriate cost functions for each problem, the observer is made to satisfy a performance bound. A systematic method of checking existence conditions and calculating observer gains in terms of Linear Matrix Inequalities (LMIs) is presented. Types of nonlinearities considered are fairly general and encompass sector bounded, Lipschitz and dissipative nonlinearities. The observer can identify time-varying unknown parameters, bias and gain sensor faults. Compared with the method of extended Kalman filter, the proposed observer is not computationally intensive and in its relaxed form, does not require online solution to the Ricatti equation. The observer is applied to representative state space models including a wind turbine mechanical power transmission mechanism. The considered system model is highly nonlinear and contains input disturbances as well. The results are compared with the results obtained from extended Kalman filtering and show satisfactory performance in the presence of noise and disturbances.

Keywords: Nonlinear adaptive observer; state and parameter estimation; fault estimation; optimal estimation; wind turbine
To my Parents
Acknowledgements

I would like to express my deep gratitude to my supervisor, Prof. Krishna Vijayaraghavan, for his help and support throughout this project. Without him it would have been impossible to do this work. Further, I am sincerely thankful to my co-supervisor Dr. Siamak Arzanpour for his support and assistance. Also, my deep gratitude goes to Dr. Mehrdad Moallem for examining this thesis.

Moreover, I would like to take this opportunity to appreciate the support and encouragement of all my professors who have helped me take my academic steps so far, special thanks to Dr. Farid Golnaraghi, Dr. Kevin Oldknow, Dr. Abdolreza Ohadi, Dr. Mojtaba Sedighi, Dr. Amir Abdullah, Dr. Mitra Danesh Pazhouh and Dr. Mohammad Zareinejad. It has been a great privilege to have the experience of working with them. Their behavior and advice have been the most influential and inspiring motivations in my academic life.

Last but not least, I sincerely thank my parents for their unstinting love and support throughout my academic and non-academic life.
# Table of Contents

Approval .................................................................................................................. ii  
Abstract ................................................................................................................... iii  
Dedication ............................................................................................................... iv  
Acknowledgements ................................................................................................ v  
List of Tables ....................................................................................................... vi  
List of Figures ...................................................................................................... viii  
List of Acronyms ................................................................................................ ix  
List of Symbols .................................................................................................. xii  

## Chapter 1. Introduction and Literature Review ............................................. 1

1.1. Nonlinear estimation .................................................................................. 2
    1.1.1. Nonlinear Observers ...................................................................... 2
    1.1.2. Nonlinear adaptive observers ....................................................... 5
    1.1.3. Kalman Filter .............................................................................. 7
1.2. Fault detection, estimation and fault tolerant control ................................ 10
    1.2.1. Fault detection and estimation ..................................................... 11
    1.2.2. Observer based FDI ................................................................... 13
1.3. Purpose and scope of this work ............................................................... 16

## Chapter 2. Optimal Nonlinear Adaptive Observer ....................................... 18

2.1. System model ............................................................................................ 19
2.2. Preliminary: Nonlinear observer vs. the Extended Kalman Filter (EKF) ... 20
2.3. Observer model ......................................................................................... 21
2.4. Optimal observer ....................................................................................... 23
    2.4.1. $H_\infty$ filter as an alternative to the Kalman Filter .................... 23
    2.4.2. Observer design .......................................................................... 24
        Theorem 2.1 ....................................................................................... 25
        Theorem 2.2 ....................................................................................... 34
    2.4.3. Extension to general sector-bounded nonlinearity ..................... 38
    2.4.4. Time-varying unknown parameter ............................................. 39
        Theorem 2.3 ....................................................................................... 40
2.5. Numerical example .................................................................................... 44
    2.5.1. Example 1. 4\textsuperscript{th} order nonlinear system ................. 44
        State and parameter estimation in Lipschitz nonlinear system using EKF 44
        State and parameter estimation in Lipschitz nonlinear system using algorithm proposed in this paper .................. 49
    2.5.2. Example 2. 3\textsuperscript{rd} order nonlinear system ....................... 57
3.1. System model.........................................................................................62
3.2. Observer model.......................................................................................64
3.3. Optimal observer and fault estimator......................................................65
   3.3.1. Observer design.................................................................................65
   Theorem 3.1..................................................................................................66
   Theorem 3.2..................................................................................................75
3.4. Numerical example .................................................................................78
   3.4.1. 4th order nonlinear system...............................................................78

Chapter 4. Observer and FDI Implementation: Wind Turbine Case Study ....84
4.1. Wind turbine components.......................................................................85
4.2. Wind turbine modelling..........................................................................87
   4.2.1. Aerodynamic model .........................................................................87
   4.2.2. Gear train model...............................................................................89
4.3. Wind turbine operation...........................................................................92
4.4. Wind turbine control strategies..............................................................93
4.5. Simulation results..................................................................................96
   4.5.1. Example 1. State and parameter estimation for the wind turbine model........................................................................97
   4.5.2. Sensor gain and bias faults estimation for the wind turbine model....101

Chapter 5. Conclusions and Future Work...................................................106

References ....................................................................................................108
List of Tables

Table 4-1 Symbols used in the 3rd order wind turbine model.................................90
Table 4-2 Turbine and wind parameters ......................................................................91
Table 4-3 Wind turbine mechanical parameters ...........................................................91
List of Figures

Figure 1.1. Simultaneous Fault Detection & Isolation (FDI) and active Fault Tolerant Control (FTC) ................................................................. 11
Figure 1.2. Residual based fault detection .......................................................... 13
Figure 2.1. State 1 estimation using EKF ............................................................... 46
Figure 2.2. State 2 estimation using EKF ............................................................... 47
Figure 2.3. State 3 estimation using EKF ............................................................... 47
Figure 2.4. State 4 estimation using EKF ............................................................... 48
Figure 2.5. State estimation errors in EKF method ............................................. 48
Figure 2.6. Parameter estimation by EKF ............................................................. 49
Figure 2.7. State 1 estimation using the proposed observer .................................. 51
Figure 2.8. State 2 estimation using the proposed observer .................................. 51
Figure 2.9. State 3 estimation using the proposed observer .................................. 52
Figure 2.10. State 4 estimation using the proposed observer ................................ 52
Figure 2.11. State estimation errors for the proposed observer ........................... 53
Figure 2.12. Parameter estimation by the proposed observer ............................... 53
Figure 2.13. Cost function .................................................................................. 54
Figure 2.14. Comparison between state 1 estimation errors of EKF and the proposed observer ................................................................. 55
Figure 2.15. Comparison between state 2 estimation errors of EKF and the proposed observer ................................................................. 55
Figure 2.16. Comparison between state 3 estimation errors of EKF and the proposed observer ................................................................. 56
Figure 2.17. Comparison between state 4 estimation errors of EKF and the proposed observer ................................................................. 56
Figure 2.18 State 1 estimation using the proposed observer .................................. 58
Figure 2.19 State 2 estimation using the proposed observer .................................. 59
Figure 2.20. State 3 estimation using the proposed observer ................................ 59
Figure 2.21. State estimation errors ...................................................................... 60
Figure 2.22. Unknown parameter estimation ...................................................... 60
Figure 2.23. Cost function .................................................................................. 61
List of Acronyms

ARE       Algebraic Ricatti Equation  
DFIG      Doubly Fed Induction Generator  
EKF       Extended Kalman Filter  
FDE       Fault Detection and Estimation  
FDI       Fault Detection and Isolation  
FTC       Fault Tolerant Control  
HAWT      Horizontal Axis Wind Turbine  
LMI       Linear Matrix Inequality  
MIMO      Multiple Input Multiple Output  
MPPT      Maximum Power Point Tracking  
NDFS      Noise and Disturbance Free Systems  
SISO      Single Input Single Output  
SPR       Strictly Positive Real  
UKF       Unscented Kalman Filter
**List of Symbols**

\[ A \] System matrix  
\[ x \] State vector  
\[ \mu \] Unknown parameter vector  
\[ B \] Direction matrix for the unknown parameter  
\[ D \] Direction matrix for the unknown parameter  
\[ E \] Direction vector for the nonlinearity  
\[ \Phi \] Nonlinearity  
\[ F \] Direction matrix for disturbance vector  
\[ \omega \] Disturbance vector  
\[ H \] Direction matrix for input vector  
\[ u \] Input vector  
\[ C \] Output matrix  
\[ y \] Output vector  
\[ v \] Disturbance vector  
\[ C_B \] Direction matrix for the bias fault  
\[ C_{G1} \] Direction matrix for the gain fault  
\[ C_{G2} \] Direction matrix for the gain fault  
\[ P_W \] Total wind power  
\[ C_P \] Power coefficient  
\[ P_A \] Captured aerodynamic power  
\[ c \] Damping coefficient  
\[ k \] Total shafts and gearbox stiffness  
\[ N \] Gear ratio  
\[ J_g \] High-speed shaft moment of inertia  
\[ J_r \] Low-speed shaft moment of inertia  
\[ T_g \] Generator torque  
\[ T_r \] Aerodynamic torque applied by the wind on the low-speed shaft  
\[ R \] Blade radius
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_{nom}$</td>
<td>Nominal wind speed</td>
</tr>
<tr>
<td>$\lambda_{opt}$</td>
<td>Optimal tip speed ratio</td>
</tr>
<tr>
<td>$C_{p_{max}}$</td>
<td>Maximum power coefficient</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Air density</td>
</tr>
</tbody>
</table>
Chapter 1.

Introduction and Literature Review

Wind turbines are complex dynamic systems, consisting of several sub-systems that need to be controlled and simultaneously monitored for potential faults. In control systems, states are used to represent internal dynamics of a system. States of complex dynamic systems are not always readily measurable. “Observers” are used to estimate the states by using measurements (sensor outputs) together with the knowledge of inputs to the system and also the system model. The theory of observers constitutes an important field in control systems in its own right. Some linear and nonlinear control techniques require knowledge of all system states [1], [2]. Furthermore, state values are utilized by many fault detection techniques [3]. Typically, complex real world systems such as wind turbines are nonlinear and hence their respective observers need to be nonlinear. Availability of accurate system models is vital to the performance and even stability of regular state observers. Nevertheless, in practice, complex system would likely have one or more parameters whose values may not be known accurately. Further, physical systems would be subject to unknown disturbances and sensor noise. This thesis aims to address several important challenges involved with complex systems by developing robust adaptive observers.

We will review the literature on nonlinear observers and nonlinear adaptive observers for unknown parameter or fault estimation in this chapter. Additionally, we will concisely cover the optimal filtering method of Kalman Filter and its extensions to nonlinear systems. Furthermore, we will briefly review different types of available fault detection and isolation techniques and present some background results. The last part of this chapter overviews the purpose and scope of this thesis.
1.1. Nonlinear estimation

1.1.1. Nonlinear Observers

The development of nonlinear observers has extended the scope of state estimation to fairly general classes of nonlinear systems. Nonlinearities are so commonly encountered in physical systems that we need to devise estimators for such nonlinear models. In this section, we will briefly review the literature on nonlinear observers. Further, adaptive observers have been designed for different types and different formulations of nonlinearities. We will elaborate on these observers in the next section.

Raghavan et al. [4] developed an observer for a Multiple Input Multiple Output (MIMO) Lipschitz nonlinear system where the nonlinearity can be a function of all the states of the system (i.e. \( \Phi = \Phi(t, u, x), |\Phi(t, u, x_1) - \Phi(t, u, x_2)| < \gamma |x_1 - x_2| \)). It was shown earlier by Thau [5] that for the nonlinear system

\[
\dot{x} = Ax + g(t, u, y) + \Phi(t, u, x) \\
y = Cx
\]

where 1. the first nonlinearity \( g(t, u, y) \) is only dependent on the inputs and outputs, 2. the second nonlinearity \( \Phi(t, u, x) \) is dependent on any of the states and 3. the observer is constructed as

\[
\dot{\hat{x}} = A\hat{x} + g(t, u, y) + \Phi(t, u, \hat{x}) + L(y - C\hat{x})
\]

If \( \exists P, Q, L \) such that \( \gamma < \sigma_{\text{min}}(Q)/2\sigma_{\text{max}}(P) \) where \( P = P^T > 0 \) and \( Q = Q^T > 0 \) satisfying the Lyapunov equation

\[
(A - LC)^T P + P(A - LC) = -Q
\]
Then \( \lim_{t \to \infty} \tilde{x} = 0 \) where \( \tilde{x} = x - \hat{x} \). To use this theorem, once \( L \) is selected, the conditions can be checked to ensure the stability of the proposed observer. It does not, however, provide a means of selecting an appropriate \( L \) to guarantee a stable observer. Finding a solution to the equation (1.3) does not guarantee the satisfaction of the previous inequality. In effect, the obtained solution should satisfy both the equation and the inequality simultaneously for which there is no straightforward procedure. To address the difficulty of finding \( L \), Raghavan et al. [4] proposed the following theorem which enables us to select an appropriate gain:

For the system (1.1) and the observer (1.2), if \( \exists \epsilon > 0 \) such that the following equation has a solution \( P = P^T > 0 \)

\[
AP + PA^T + P \left( \gamma^2 I - \frac{1}{\epsilon} C^T C \right) P + I + \epsilon I = 0
\]

(1.4)

then the gain \( L = \left( \frac{1}{2\epsilon} \right) PC^T \) stabilizes the error dynamics for all \( \Phi \) with Lipschitz constant \( \gamma \). The previous equation is known as Algebraic Riccati Equation (ARE). Although this design provides a systematic algorithm for calculating proper observer gains through solution of the ARE, it does not shed light on the conditions required by the observer gain and the nature of the observer matrix \( (A - LC) \) for stability.

A more insightful gain calculating technique is presented by Rajamani [6] which imposes direct conditions on the matrix \( (A - LC) \) for stability. He proves that for the system (1.1) and observer (1.2), if the gain \( L \) is chosen such that \( (A - LC) \) is stable and

\[
\min_{\omega \in \mathbb{R}^+} \sigma_{\min}(A - LC - j\omega I) > \gamma
\]

(1.5)

where \( \sigma \) shows the singular value of its respective matrix, then the observer is asymptotically stable. This theorem directly relates the properties of the conventional \( (A - LC) \) matrix to the Lipschitz constant of the nonlinearity.
Other results for design of stable $H_{\infty}$ observers with time-varying gains for Lipschitz nonlinear systems can be found in [7], [8]. The performance of these observers can surpass ordinary constant gain observers. Nevertheless, they do not account for noise and disturbance either. Vijayaraghavan [9] has developed an observer for dissipative Lipschitz nonlinear systems of the following form subject to noise

\begin{align}
\dot{x} &= Ax + Hu + E\Phi(u, x) + w \\
y &= Cx + v
\end{align}

(1.6)

The observer has the form

\begin{align}
\dot{\hat{x}} &= A\hat{x} + Hu + E\Phi(u, \hat{x}) + L(y - Cx)
\end{align}

(1.7)

and can guarantee the performance

\begin{align}
J &= \frac{\int_0^T \dot{\hat{x}}^T W \dot{\hat{x}} dt}{\hat{x}_0^T \bar{P}_0 \hat{x}_0 + \int_0^T v^T Q^{-1} v dt + \int_0^T w^T R^{-1} w dt} < \frac{1}{\alpha}
\end{align}

(1.8)

If the derivation steps are taken for matrix Lipschitz nonlinear systems, one will arrive at the following Riccati equation for selecting the observer gains

\begin{align}
-\dot{P} &= W' - C^T Q^{-1} C + A^T \bar{P} + \bar{P} A + \bar{P} E E^T \bar{P} + \bar{P} R \bar{P}
\end{align}

(1.9)

with the gain selected as $L = \bar{P}^{-1} C^T Q^{-1}$. In this equation, $W' = \alpha W + G^T G$ and $G$ is the matrix Lipschitz constant. This observer is the foundation on which we will build our adaptive observer in this work.
1.1.2. Nonlinear adaptive observers

Typically, it is assumed that an appropriate system model and its parameters are known while designing regular observers like the Luenberger observer. This observer is merely capable of estimating the states of linear deterministic systems with known model and model parameters. However, it is prevalent in the real world to have system models in which the value of some of the model parameters are unknown, uncertain or cannot be extracted by experiments. With the development of adaptive observers, we have gained the ability to estimate unknown parameters of the system, or alternatively system faults, as well as the states of the system. Luders et al. [10] have developed adaptive linear observers for simultaneous state and parameters estimation in linear systems. They assume the only available knowledge is the order of the system and its inputs & outputs. For noise-free linear time-invariant systems, a comprehensive analysis of adaptive observers can be found in Narendra et al. [11].

Starting in the 1980s, the first nonlinear adaptive observers were designed with restrictive limitations on the type of nonlinearity in the system. Bastin et al. [12] proposed an adaptive observer and parameter adaptation law for noise-free Single Input Single Output (SISO) nonlinear systems of the form $\dot{x} = Ax + f(t, u, y)\mu; \ y = Cx$, where the nonlinearity is independent of unmeasured system states and only dependent on the inputs and outputs. Getting the output-nonlinear formulation from the initial state-space model requires a transformation takes place. Marino [13] presents the conditions required for system to be transformable into this output-nonlinear form. Moreover, Marino et al. [14] have presented an adaptive observer with arbitrary exponential convergence rate for the same class of linearly parametrized nonlinear system introduced in [12]. This work has been an extension to the linear results of [15].

The result of Thau [5] has been extended to unknown parameter systems in Rajamani et al. [16]. Besançon [17] has shown that the conditions for observer existence in Bastin et al. [12] and in Rajamani et al. [16] are equivalent, however, [16] does not require the state transformation steps of Bastin et al. [12].
Again, the challenge to find proper values of gain fulfilling the conditions exists. However, a systematic approach in terms of LMI conditions for finding gains for the nonlinear adaptive observer is proposed in Cho et al. [18]. Consider the following uncertain Lipschitz nonlinear system [18]

\[ \dot{x} = Ax + \Phi(u, x) + bf(u, x)\theta \]
\[ y = Cx \]

where the Lipschitz constants for \( \Phi(u, x) \) and \( f(u, x) \) are \( \gamma_1 \) and \( \gamma_2 \) respectively. Also, the bound \( ||\theta||_2 \leq \gamma_3 \) is assumed known on the vector of unknown parameters \( \theta \). They suggest the following observer and adaptation law

\[ \dot{x} = A\hat{x} + \Phi(u, \hat{x}) + bf(u, \hat{x})\hat{\theta} + L[y - C\hat{x}] \]
\[ \dot{\theta} = \frac{f(u, \hat{x})^T C_1 \hat{x}}{\rho} \quad \rho > 0 \]  

Now, if there exist \( P = P^T > 0 \) and \( L \) such that

\[ (A - LC)^TP + P(A - LC) + (\gamma_1 + \gamma_2\gamma_3||b||_2)PP + (\gamma_1 + \gamma_2\gamma_3)I < 0 \]
\[ b^TPC^\perp = 0 \]  

Then the proposed adaptive observer is convergent, \( \bar{x} \to 0 \) and \( [bf(u, x)\theta - bf(u, \hat{x})\hat{\theta}] \to 0 \) as \( t \to \infty \). This was one of the first works to develop an adaptive nonlinear observer with the nonlinearity dependent on any of the states and to present a straightforward gain calculation procedure.

The results of Rajamani et al. [16] have been extended to tolerate perturbations on the observer gain and bounded input disturbance [19]. Ibrir [20] presents an observer for estimating unknown input coefficients for noise-free systems where the linear part of the
system has a controller canonical form with the output equal to the first state. Further, the nonlinearity affecting the \( i^{th} \) derivative \( \dot{x}_i \) is a function of only the states 1 to \( i \) (i.e. \( y = x_1; \dot{x}_i = x_{i+1} + \phi_i(x_1, x_2, ..., x_i), i < n; \dot{x}_n = \phi_i(x_1, x_2, ..., x_n) + \mu u \) allowing for a relatively simple observer construction. Vijayaraghavan [21] develops an adaptive observer for systems \( \dot{x} = Ax + \mu \beta x + \Phi(u, x); y = Cx; \bar{y} = \bar{C}x \) in which the relative degree from \( \bar{y} \) to \( \Phi \) (defined as \( r_\Phi, s.t. \bar{C}A^{i-1}\Phi = 0 \ \forall i < r_\Phi - 1 \)) is less than the relative degree from \( \bar{y} \) to \( \mu \) (defined as \( r_\mu, s.t. \bar{C}A^{i-1}\beta = 0 \ \forall i < r_\mu - 1 \)). It should be noted that many previous results \([16],[18]\) with the exception of Vijayaraghavan [21], require that the dynamics of the system from the unknown parameters to the output satisfies a strictly positive real-type condition. In order to ensure convergence, they require finding a Lyapunov matrix \( P > 0 \) for the observer error dynamics that satisfies \( b^TPC\perp = 0 \). On closer examination it is seen that the design is restricted to systems where \( b \) lies in the range space of \( C^T \) (since \( P \neq 0, b^T \subset \mathcal{N}(C) \) or equivalently \( b \subset \mathcal{R}(C^T) \), is a necessary condition for finding such a \( P \)). Even when \( b \) is in the range space of \( C^T \), the choice of \( P \) is restricted and an observer may not exist for every case. Vijayaraghavan [21] also presents observer design for the cases of time-varying parameters and multiple unknown parameters. However, the observer design relies on differentiating the sensor output and is limited to classes of Lipschitz nonlinear systems with almost no measurement noise. While the aforementioned works make great theoretical contributions, they do not consider measurement noise and are not suitable for practical observers.

1.1.3. **Kalman Filter**

The method of Kalman Filter (KF) is one of the most frequently used methods of filtering and estimation in noisy systems. Kalman Filter is an estimator for the linear-quadratic problem for estimating the states of a linear dynamic system perturbed by white noise [22]. It utilizes noisy measurement data for state estimation based on a recursive algorithm. The estimation process in KF is comprised of two steps; Prediction step and Update step. Kalman filter has been originally developed for discrete-time models;
however, continuous-time formulations of KF are also widely available in the literature. In the continuous time format, the two steps are combined in the form of a differential Ricatti equation [23].

Within the scope of noisy systems, there are traditionally three estimation problems that need to be addressed. These include state estimation, parameter estimation and dual (state and parameter) estimation. Kalman filter was originally designed for filtering and estimation in noisy linear systems, however, as it has grown into maturity, methods have been developed which extend the concept of Kalman filtering to nonlinear systems. Extended Kalman Filter (EKF) [24], [25] and Unscented Kalman Filter (UKF) [26], [27], [28] are two approaches developed in this respect. The extended Kalman filter linearizes nonlinear models so that the traditional Kalman filter can be applied. It is a first order linearization of the nonlinear system. This method requires online solution to the Ricatti equation (1.17) and is normally computationally intensive. Originally developed in the discrete-time form, its filtering process consists of prediction and update steps. For the nonlinear system

\[
x_k = f(x_{k-1}) + w_{k-1} \\
z_k = h(x_k) + v_k
\]  

where \( w_{k-1} \) and \( v_k \) are process and output noise both assumed to be zero-mean Gaussian with covariance \( R_k \) and \( Q_k \). The following predict and update steps are then applied

**Predict**

\[
\hat{x}_{k|k-1} = f(\hat{x}_{k-1|k-1}) \\
P_{k|k-1} = F_{k-1} P_{k-1|k-1} F_{k-1}^T + R_{k-1}
\]

\[
S_k = H_k P_{k|k-1} H_k^T + Q_k
\]

**Update**

\[
L_k = P_{k|k-1} H_k^T S_k^{-1} \\
\hat{x}_{k|k} = \hat{x}_{k|k-1} + L_k \left(z_k - h(\hat{x}_{k|k-1})\right)
\]
\[ P_{k|k} = (I - L_k H_k) P_{k|k-1} \]

where the following Jacobian definitions are used for \( F \) and \( H \)

\[
F_{k-1} = \frac{\partial f}{\partial x}|_{\hat{x}_{k-1|k-1}} \\
H_k = \frac{\partial h}{\partial x}|_{\hat{x}_{k|k-1}}
\]

In continuous time, the system model is formulated as

\[
\dot{x}(t) = f(x(t)) + w(t) \\
z(t) = h(x(t)) + v(t)
\]

and the prediction and update steps are combined resulting in

\[
\dot{\hat{x}} = f(\hat{x}) + L(y - h(\hat{x})) \\
\dot{\hat{P}} = FP + PF^T - PH^T Q^{-1} HP + R \\
L = PH^T Q^{-1}
\]

with \( F \) and \( H \) defined as

\[
F = \frac{\partial f}{\partial \hat{x}}|_{\hat{x}, u} \\
H = \frac{\partial h}{\partial x}|_{\hat{x}}
\]

In practice, however, there are challenges involved with implementation of Kalman filter-based methods. For one thing, having a suitable estimate of noise covariance matrices is a potential drawback. Consequently, research on extraction of noise covariance matrices from data has been the focus of extensive activities in this area. To use the EKF for parameter estimation, the unknown parameters should be added to the
state vector to form the augmented state vector. This also require virtual parameter disturbance be used in the process disturbance covariance matrix.

Higher order linearizations like second order and third order EKFs are also available in the literature [29]. The unscented Kalman filter outperforms EKF with approximately the same computational complexity. Kalman filter estimation is optimal only if some circumstances are met. These include the accuracy of the model and exact knowledge of noise covariance. It will be shown in the examples following the chapters that the method proposed in this work is computationally less intensive compared with the EKF. Also, the response is faster with the observer especially with high initial estimation errors.

1.2. Fault detection, estimation and fault tolerant control

The ever-increasing need for reliability in dynamic systems has propelled extensive research in fault detection. There are two reasons why online fault detection and accommodation is important. First, malfunction in parts of the system could lead to catastrophic failures and result in destructions that are not easily compensated. Our ability to identify the fault at its incipient stages can reduce the risk of imminent destructions. Second, pausing system operation for off-line fault inspection and then re-starting the system may not be economically feasible. We therefore need trustable, fast, sensitive fault accommodation schemes.

There are two major steps in fault accommodation: first, Fault Detection and Isolation (FDI) to detect and estimate the fault and then Fault Tolerant Control (FTC) to control the system in spite of the existing faults. FTC [30] aims to control the system to attain the pre-defined objectives despite the existing fault. FTC can be either passive or active. The former treats faults as uncertainties and applies robustness techniques against such uncertainties to achieve the goal of the system. Active FTC, on the other hand, relies on Fault Detection and Isolation (FDI) since knowledge of the fault data is necessary for
fault accommodation. Figure 1.1 illustrates a system comprising two fault accommodation modules. As observed in this figure, the FTC module uses information obtained from FDI module to modify the control rules when a fault occurs.

![Diagram of fault accommodation system](image)

**Figure 1.1. Simultaneous Fault Detection & Isolation (FDI) and active Fault Tolerant Control (FTC)**

1.2.1. **Fault detection and estimation**

Fault diagnostics is a key feature in advanced systems and processes due to destructive effects that potential faults can pose. FDI techniques should be able to detect the advent of faults as their first and foremost task. Next, they should be able to identify the location and severity of the faults. The most trivial approach to fault detection is using component redundancy. In this approach, the idea is to duplicate the operation of critical parts of the system. Further, a detection strategy based on non-similar operation of alternative system components is designed to infer existence, location and severity of the fault. This method is inefficient due to the cost imposed by using redundant hardware in the system. Other than the component redundancy approach, FDI approaches can be broadly classified into model based FDI [3], [31] and data-driven FDI [32]. Data-driven FDI techniques require knowledge of inputs and outputs of the system under operation. This eliminates the need for system models. Therefore this method becomes highly attractive when we do not know much about model of the system and whether it is linear or nonlinear. Methods like artificial neural networks are examples of data-driven fault
detection [33]. On the other hand, we mainly rely on the system model in model-based FDI and therefore require a proper model of the system in order for this to work.

Available models may not be too accurate as some uncertainty is associated with any model in the real world. Model based techniques should therefore be able to deal with such uncertainties. Within model based FDI, observer design and parity space method are two alternative strategies [34]. Observer based FDI can provide online fault detection for a wider class of systems. With advances in observer design techniques, increasing research focus is devoted to observer based FDI [3], [35]–[38]. Observer based FDI, in turn, can be sub-divided into residual based fault detection and Fault Detection and Estimation (FDE).

Sensors are one of the most prevalent sources of faults in dynamic systems. Sensors may experience various fault scenarios. One of the most common sensor faults is a constant deviation associated with the sensor readings called bias fault. Bias faults appear in many sensors as the sensor undergoes excessive hours of operation. It gradually becomes noticeable and affects operation of the system. Another type of sensor fault encountered in some sensors is sensor gain fault. Such a fault causes sensors to present the output as a factor of the actual output. A reliable fault detection scheme should be able to detect such faults before they start to affect the system operation severely or cause catastrophic damage to the system. In addition to the faults, there always exists some noise associated with any sensor in the system. Therefore, in order for the FDI algorithm to be practically applicable, noise should be accounted for.

Wind turbine mechanical power transmission mechanism is a highly nonlinear and uncertain system due to the nonlinear nature of its aerodynamic model and uncertain wind parameters. Estimation of existing sensor faults in such a system well in advance of catastrophic failures taking place is of great importance. This system is presented as a case study for utilizing our proposed fault estimation technique in this work.
1.2.2. Observer based FDI

This section focuses on the application of observers in FDI. Adaptive observers are first designed for estimating faults in the systems so that one can take advantage of Fault tolerant Control (FTC) approaches in the next steps. The application of adaptive observers for fault estimation has been the focus of numerous studies. Both time-invariant and time-varying faults could be re-constructed by utilizing adaptive observers. It is customary to have unknown immeasurable input disturbance and sensor noise in the real-world processes. Therefore, it is necessary to account for such noise and disturbance in the fault detection scenario as well. One of the main features of an active fault diagnostic scheme is maximum sensitivity to system faults and minimum sensitivity to noise and disturbances. In this work, we pursue our fault estimation technique to achieve this objective.

Using observers for detecting faults in the system is a useful yet complicated practice. States of the system are re-constructed by use of observers. Residual based FDI compares the estimated output(s) with the actual sensor reading(s) to generate a sensor residual. The observer is designed so that the direction of the residual can be used to identify faults [3]. Fault isolation is performed in this regard to arrive at a residual vector where each element of the vector indicates a particular fault scenario in the system [39].

![Residual based fault detection](image_url)
In some cases, a bank of observers is designed for FDI and the residuals generated from different observers are used for comparison and fault data extraction [40], [41]. Further, attempts are made to robustify residual-based FDI techniques against model uncertainties and noise by defining appropriate fault detection thresholds [42], [43] and using disturbance decoupling and unknown input observers [44], [45].

FDE, in contrast, relies on adding the fault to the list of system states to be estimated [35]–[37]. Adaptive estimation is closely related to FDE by modeling the fault as an unknown system parameter. We therefore develop adaptive nonlinear observers for both parameter estimation and fault detection.

Much of the literature on observer design and on FDI has focused on noise and disturbances free systems (NDFS). For such NDFS, Rajamani et al. [46] propose a residual-based FDI observer for additive sensor faults. Wang et al. [35], [36] build upon previous observer design results and propose an actuator and sensor fault diagnosis scheme for linear NDFS. However, this discussion does not provide a systematic procedure for calculating observer gains. These results are extended to nonlinear Lipschitz systems by Jiang et al. [30]. In their work, they consider the system

\[
\dot{x} = Ax + \Phi(x, t) + Bu + Ef(t) \\
y = Cx
\]

(1.18)

where \( \Phi(x, t) \) is the Lipschitz nonlinearity with constant \( \gamma \) and \( f(t) \) is the time-varying differentiable fault vector. They also assume bounds exist on both the fault and its derivative. Now, if there exists matrix \( K \) such that \( G(s) = C[sI - (A - KC)]^{-1}E \) is Strictly Positive Real (SPR): \( \text{Re}(G(j\omega)) > 0, \forall \omega > 0 \) and also \( \min(\sigma_{\min}(A - KC - j\omega I)) > \gamma \), then there exist matrices \( P > 0 \) and \( R \) such that

\[
PE = C^T R
\]

(1.19)
\[(A - KC)^T P + P(A - KC) + \varepsilon \gamma^2 I + \frac{P^2}{\varepsilon} + Q < 0\]

for any \(Q > 0\) and \(\varepsilon > 0\). The observer and fault estimator would have the form

\[
\begin{align*}
\dot{x} &= A\hat{x} + \Phi(\hat{x}, t) + Bu + E\hat{f}(t) + K[y - \hat{y}] \\
\hat{f}(t) &= \Gamma R^T[y - \hat{y}] - \sigma \Gamma \hat{f}(t) \quad (1.20) \\
\hat{y} &= C\hat{x}
\end{align*}
\]

Vijayaraghavan et al. [47] investigate actuator additive input fault estimation for Lipschitz nonlinear NDFS where the relative degree from the input \((r_u)\) and from the nonlinearity \((r_\phi)\) satisfy the condition \(r_\phi \geq r_u\). Zhang et al. [48] slightly modified the adaptive estimation law by including both proportional and integral terms and arrived at an observer for faster tracking of additive sensor and actuator faults. Unlike the majority of the works reviewed thus far that have focused on additive input and sensor faults, either constant (bias) or variable with time, Wang et al. [49] consider sensor gain faults for NDFS. They divide sensor faults into conditionally detectable faults and conditionally identifiable faults and investigate the number and location of sensor gain faults that can be identified simultaneously. According to them, for the system

\[
\begin{align*}
\dot{x} &= Ax + \Phi(u, x, t) + Bu \\
y &= Cx + Df_D(t) \\
x \in \mathbb{R}^n \\
\text{rank}(D) = q
\end{align*} \quad (1.21)
\]

if the fault matrix \(D\) can be devised in a way that \(\text{rank}\begin{bmatrix} \mu l - A \\ C \\ D \end{bmatrix} = n + q, \forall \mu \in \mathbb{C}^+\), then the fault vectors can be estimated using their proposed algorithm provided that all considered faults are identifiable. The gain faults considered in their work are modelled by time-varying bias faults.
Also, there are numerous research works investigating and developing sliding mode observers for fault detection and estimation [38], [50]–[52]. Such observers are appropriate in that they are robust against uncertainties.

Another approach would be to use the EKF with augmented state vector for faults estimation. As pointed out earlier, it requires specifying a virtual parameter disturbance covariance for simulation. Since such covariance does not exist in reality, one should assign a virtual value to this covariance. For instance, sometimes system disturbance covariance serves this purpose. Furthermore, the challenges involved with the EKF implementation discussed earlier uphold in the FDI scope as well.

1.3. Purpose and scope of this work

This thesis develops nonlinear adaptive observers for systems with existing input disturbance and measurement noise based on an optimal approach. The observer is employed for the twofold purpose of parameter estimation and fault estimation. Constant and time-varying unknown parameters as well as bias and gain sensor fault can be identified by the proposed observer.

In chapter 2, an adaptive observer is designed for uncertain nonlinear systems. It can achieve a desired cost performance defined by a cost function. The designed observer is capable of estimating all system states as well as time-varying unknown parameters. We first start with a Lipschitz nonlinearity and then extend the problem to the general class of sector-bounded nonlinear systems. The existence condition is formulated as a Linear Matrix Inequality (LMI) and a compact formulation for calculating the gains is provided. Also in this chapter, the performance of the proposed observer is compared with the method of Extended Kalman Filter (EKF) through an example.

In chapter 3, a nonlinear observer and fault estimator capable of identifying additive and gain sensor faults is presented. The considered system nonlinearity is fairly general and encompasses a variety of different physical nonlinearities. The minimum
performance is guaranteed by checking LMI conditions and LMI solutions are used for selecting observer gains. Also, noise and disturbance can be dealt with in this design.

Chapter 4 puts the FDI technique into practice employing illustrative wind turbine case studies. Wind turbine aerodynamic and torque transmission models are presented which involve some complexities due to their nonlinear and uncertain nature. Adaptive observers are designed for detection and estimation of the existing faults in the system.

Finally, the thesis is concluded in chapter 5 and recommendations for future extensions to this work are presented.
Chapter 2.

Optimal Nonlinear Adaptive Observer

In modelling some dynamic systems, while the basic dynamics of the system may be known, the parameters of the model would be harder to come by. For instance, in modelling a mechanical system, one may assume a certain nonlinear model for friction that is of sufficient accuracy, however, the value of the friction parameter used in that model is not known. If an observer is to be designed for such a system, it needs to estimate value of the unknown parameter simultaneous with estimating the states.

In this chapter, we intend to develop a nonlinear adaptive observer for such nonlinear systems with time-varying unknown parameters. The observer should be robust with respect to existing noise and disturbances in the system. We pursue an optimal methodology for gain calculation. First, we present a non-adaptive version of the observer and compare that with the Extended Kalman Filter. Next, an adaptive version of the observer is developed for dual estimation. We start with Lipschitz nonlinear systems with constant unknown parameters and then extend our theorem for more general nonlinearities and time-varying parameters.

A general framework is provided for the state-space models of this chapter with uncertainties and disturbances included in the model. Then, an observer model is proposed for the available system model. Next, the observer is developed by adopting an optimal approach and the observer existence conditions and the gain computation procedure is developed. Further, more general nonlinearities and time-varying parameters are presented and appropriate observers are designed. At the end of this chapter, two sample dynamical systems are studied using the methods of this work. The performance
and time-consumption of the proposed observer is investigated and compared with the Extended Kalman Filter.

2.1. System model

Consider the following Lipschitz nonlinear system

\[
\begin{align*}
\dot{x} &= Ax + E\Phi(u, x) + Fw + Hu \\
y &= Cx + v
\end{align*}
\]

(2.1)

where \(x \in \mathbb{R}^n\) is the state vector, \(A \in \mathbb{R}^{n \times n}\) is the system matrix, \(E \in \mathbb{R}^{n \times q}\), \(q \leq n\), \(\Phi(u, x) : \mathbb{R}^p \times \mathbb{R}^n \rightarrow \mathbb{R}^q\) is the nonlinearity, \(u\) is the vector of known inputs, \(w\) is a vector of input disturbances and \(v\) is measurement noise vector. \(F, H, C\) are matrices of appropriate dimension and \(y \in \mathbb{R}^m\) is the output vector. The nonlinearity \(\Phi(u, x)\) is matrix Lipschitz

\[
|\Phi(u, x) - \Phi(u, \hat{x})| \leq |G(x - \hat{x})| \\
\forall x, \hat{x} \in \mathbb{R}^n, u \in \mathbb{R}^p
\]

(2.2)

If the system model contains one or more unknown parameters, it can be rewritten as

\[
\begin{align*}
\dot{x} &= Ax + \sum \mu_i B_l D_l x + E\Phi(u, x) + Fw + Hu \\
y &= Cx + v
\end{align*}
\]

(2.3)

where \(\mu = [\mu_1 \mu_2 ... \mu_r]^T\) is the vector of unknown parameters, \(B_l \in \mathbb{R}^n\) and \(D_l \in \mathbb{R}^{1 \times n}\) represent the state-space direction in which the unknown parameters affect the system.

We assume that \((A, C)\) is observable and \(C(B_l D_l)\) is full rank which implies the detectability of the unknown parameter. At this point, we also assume that the unknown parameters are constant or vary sufficiently small with time that can be treated like constant. We will remove this assumption in section 2.4.3 and consider time-varying
unknown parameter as well. Also, the following bound is known on the unknown parameter

\[ |\mu_i| < \mu_{i\text{-max}} \]  \hspace{1cm} (2.4)

To illustrate modeling uncertainties and unknown parameters by equation (2.3), consider the second order system \( m\ddot{x} + c\dot{x} + kx = F \). This system can be represented in state space as

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} = \begin{bmatrix}
0 & 1 \\
-k/m & -c/m
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} + \begin{bmatrix}
0 \\
1/m
\end{bmatrix} F
\]

Assume that, for instance, the stiffness \( k \) is the uncertain parameter in the system. Then \( k \) can be replaced by \( k^* + \Delta k \) where \( k^* \) is the nominal value and \( \Delta k \) is the deviation of the actual stiffness from its nominal value. Therefore, the system can be rewritten in the following form with \( \Delta k \) being the unknown parameter.

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} = \begin{bmatrix}
0 & 1 \\
-k^*/m & -c/m
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} + \Delta k \begin{bmatrix}
0 \\
1/m
\end{bmatrix} \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} + \begin{bmatrix}
0 \\
1/m
\end{bmatrix} F
\]

This representation indicates \( B \) and \( D \) matrices. In this example, \( B = [0 \ 1/m]^T \) and \( D = [1 \ 0] \).

2.2. Preliminary: Nonlinear observer vs. the Extended Kalman Filter (EKF)

In this section, we consider the system (2.1) and study the optimal observer designed in [9] for this system and compare that with the EKF. We will show that the Riccati equation that emerges while designing the EKF (equation (1.17)) is similar to the one arrived at in the development of the optimal nonlinear observer (equation (1.9)) with some slight variations. These variations, however, fuel the differences observed in the performance of these alternative estimation methods. Equation (1.9) is re-written below
\[-\dot{P} = W' - C^T Q^{-1} C + A^T P + PA + PEE^T P + PRP \]  \hspace{1cm} (2.5)

Defining the change of variable \( P = \bar{P}^{-1} \) in this equation, it results that

\[ P\bar{P} = I \Rightarrow P\dot{\bar{P}} + \dot{P}\bar{P} = 0 \Rightarrow \dot{\bar{P}} = -P^{-1}\dot{P}P^{-1} \]  \hspace{1cm} (2.6)

Therefore, using the relation (2.6), equation (2.5) is re-written as

\[ \dot{P} = AP + PA^T - P(C^T Q^{-1} C - W') P + R + \frac{1}{\eta} EE^T \]  \hspace{1cm} (2.7)

Also, we can adopt \( F = A + EJ \) where \( J = \frac{d\Phi(u,X)}{dx} \) and re-write (1.17) as

\[ \dot{P} = AP + PA^T + EJP + PJ^T E^T - PC^T Q^{-1} CP + R \]  \hspace{1cm} (2.8)

Comparing equations (2.7) and (2.8), we can observer similarities and differences between these two Riccati equations. Defining a high bound on the nonlinearity in the observer design has added a constant term to the equation. In contrast, using the Jacobian and linearization in the EKF has added a new part to the first degree term which can affect stability of the system. This term \((EJP + PJ^T E^T)\) can be time-dependent.

Also, in the EKF formulation, the only constant term is the covariance matrix while in the observer formulation two distinct constant terms exist. This signifies the reason why uncertainty of the covariance matrix in the EKF can adversely affect the optimal performance of the filter.

### 2.3. Observer model

In this section, we intend to develop an adaptive version of the preceding observer for unknown parameter systems. For the system (2.3), the observer model is proposed as
\[
\dot{x} = A\hat{x} + \sum \mu_i B_i D_i \hat{x} + E\Phi(u, \hat{x}) + Hu + L(y - C\hat{x}) \\
\dot{\mu} = L\mu(y - C\hat{x}) = L\mu C\hat{x} + L\mu \nu
\]  

(2.9)

where \(\hat{x}\) is vector of estimated states and \(\hat{\mu}_i\) is estimated unknown parameter. The first of equations 2.4 is state reconstruction equation and the second is parameter adaptation law. These two relations are coupled as an estimation of the unknown parameter is being used with state reconstruction. \(L\) and \(L\mu\) are the observer gains. Dynamics of the estimation error for this system is written as

\[
\ddot{x} = \dot{x} - \dot{x} \\
= [Ax + \sum \mu_i B_i D_i x + E\Phi(u, x) + Fw + Hu] \\
- [A\hat{x} + \sum \hat{\mu}_i B_i D_i \hat{x} + E\Phi(u, \hat{x}) + Hu + L(y - C\hat{x})] \\
= (A - LC)\ddot{x} + E\ddot{\Phi} + B\ddot{D}_x \dot{\mu} + B\ddot{D}_x \dot{\mu} - L\nu + Fw
\]

(2.10)

and

\[
\ddot{\mu} = \dot{\mu} - \dot{\mu} \\
= -L\mu C\ddot{x} - L\mu \nu
\]

(2.11)

where the following notations have been used

\(\ddot{x} = x - \hat{x}, \quad \ddot{\mu} = \mu - \hat{\mu}\)

\(\ddot{\Phi} = \ddot{\Phi}(u, \ddot{x}, \hat{x}) = \Phi(u, x) - \Phi(u, \hat{x})\)

\(B = [B_1 | B_2 | ... | B_r]\)

\(D_x = diag(D_1 \ddot{x}, D_2 \ddot{x}, ..., D_r \ddot{x})\)

\(D_{x_i} = diag(D_1 \ddot{x}, D_2 \ddot{x}, ..., D_r \ddot{x})\)

Notice that
\[ \sum \mu_i B_i D_i x = B D_x \mu \]
\[ \sum \hat{\mu}_i B_i D_i \hat{x} = B D_{\hat{x}} \hat{\mu} \]
\[ \sum \mu_i B_i D_i x - \sum \hat{\mu}_i B_i D_i \hat{x} = B D_x \mu - B D_{\hat{x}} \hat{\mu} \]
\[ = B D_x \mu - B D_{\hat{x}} \mu + B D_{\hat{x}} \hat{\mu} - B D_{\hat{x}} \hat{\mu} \]
\[ = B D_{\hat{x}} \mu + B D_{\hat{x}} \hat{\mu} \]

As seen in (2.9), each of the observer and its adaptation law require one gain matrix be calculated to guarantee their convergence. In the following sections we would obtain a procedure for calculating appropriate gains for stabilizing the observer.

2.4. **Optimal observer**

### 2.4.1. \( H_\infty \) filter as an alternative to the Kalman Filter

In this section, we intend to develop a \( H_\infty \) nonlinear filter and estimator as an alternative to the renowned Kalman Filter. The \( H_\infty \) filter is an optimal filter and estimator that aims at estimating states/parameters in the presence of noisy data.

\( H_\infty \) filtering and estimation is set to minimize the worst-case estimation error of the space-state model. In contrast, Kalman filter is based upon minimizing the expected value of the variance of the estimation error. Kalman Filter is therefore a minimum variance estimator if the noise is Gaussian and linear minimum variance estimator if the noise is not Gaussian. If other cost functions like worst-case estimation error is intended to be minimized, KF cannot perform the task [53].

Further, there is no assumption required on the statistics of the system in \( H_\infty \) observers which makes the design more convenient and flexible [53]. As discussed earlier, one of the major challenges involved with designing accurate Kalman Filters is the problem of finding noise covariance matrices. In practice, some approximation of the
covariance matrices can be extracted from experimental data. However, the accuracy of covariance matrices affects the filter performance significantly. Therefore, in applications involving Kalman Filter and its extended forms, substantial time and cost should be devoted to understanding statistical nature of the process and the noise. In $H_\infty$ filters, in contrast, some weighting matrices are used instead of covariance matrices for the noise and disturbance attenuation purposes. These filters aim to show robustness against noise uncertainty. Despite these differences, both filters are optimal in essence and their design procedure accounts for estimating values of states/parameters from noisy data. A comparison between the nonlinear forms of these two filters without parameter uncertainty was conducted in section 2.2. It was shown that solving a differential Riccati equation is the core element of both designs, however, the derived equations differ to some extent. Despite the Kalman Filter which is based on the online linearization of the system for gain calculation, our proposed observer will adopt a single high-bound limit on the nonlinearity and will accordingly design a filter without the Riccati equation being dependent on the previously estimated states.

2.4.2. Observer design

In this section we adopt a game theory approach to $H_\infty$ filtering. Consider the following cost function

$$J = \frac{\int_0^T \tilde{x}^T W \tilde{x} dt + \int_0^T \tilde{\mu}^T W_\mu \tilde{\mu} dt}{\tilde{x}_0^T P_0 \tilde{x}_0 + \tilde{\mu}_0^T P_\mu \tilde{\mu}_0 + \int_0^T \nu^T Q^{-1} \nu dt + \int_0^T \nu^T R^{-1} \nu dt} \quad (2.12)$$

where $W, W_\mu, P_0, P_\mu, Q$ and $R$ are appropriate weight matrices selected for each problem. This cost is a ratio of quadratic functions of estimation errors over initial errors and disturbances. We wish to minimize this cost function. Both state estimation errors and parameter estimation errors are taken into account in this cost function. Also, since the observer convergence should be independent of initial observer values and robust against
noise and disturbances, they have been included in the denominator of the cost function as well.

We would formulate the filtering problem as a constrained dynamic optimization problem by utilizing Lagrange multipliers. The objective is to minimize the previously defined cost function subject to the constraint of state equations.

For the observer (2.9), provided that proper values of gains $L$ and $L_{\mu}$ have been selected, the cost function $J$ grows smaller with time. The following theorem presents necessary conditions for stability of the observer and a method for finding the gains.

**Theorem 2.1.**

For the system (2.3) the observer (2.9) can guarantee a performance

\[
J < \frac{1}{\alpha}
\]  
(2.13)

after a sufficiently large time $T$ if $\mathcal{D}_{\hat{x}} \neq 0$, and $\exists \epsilon = \text{diag}(\epsilon_1, \epsilon_2, ..., \epsilon_r) > 0$, $\eta > 0$ such that

\[
\bar{P} = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{bmatrix} > 0
\]  
(2.14)

when driven by the differential equation

\[
\begin{aligned}
\frac{d}{dt} \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^T & P_{22} \end{bmatrix} &= -\begin{bmatrix} \bar{W} & P_{11}BD_{\hat{x}} + A^T P_{12} \\ * & P_{12}^TBD_{\hat{x}} + D_{\hat{x}}^T B^T P_{12} + W_{\mu}' \end{bmatrix} \\
&\quad - \bar{P} \begin{bmatrix} 1/\eta EE^T + B \epsilon^{-1} B^T + FRF^T & 0 \\ * & 0 \end{bmatrix} \bar{P}
\end{aligned}
\]
(2.15)

where
\[
\bar{W} = \alpha W + P_{11} A + A^T P_{11} + \sum \varepsilon_i \mu_{i_{\text{max}}}^2 D_i^T D_i + \eta G^T G - C^T Q^{-1} C
\]

and observer gains are chosen as

\[
L = [P_{11} - P_{12} P_{22}^{-1} P_{12}^T]^{-1} C^T Q^{-1}
\]

\[
L_\mu = -P_{22}^{-1} P_{12}^T [P_{11} - P_{12} P_{22}^{-1} P_{12}^T]^{-1} C^T Q^{-1}
\]

(2.16)

\(\alpha\) is a user-specified performance bound.

Before starting the proof, we present the following lemma:

S-Procedure Lemma [54]: If \(V_1: \mathbb{R}^r \to \mathbb{R}\) and \(V_2: \mathbb{R}^s \to \mathbb{R}\) be such that \(V_2 \leq 0\), then \(V_1 < 0\) iff \(\exists \varepsilon > 0\) such that

\[
V_1 - \varepsilon V_2 < 0
\]

(2.17)

Proof: The cost function inequality (2.13) can be rewritten as

\[
J_1 = -\bar{x}_0^T P_0 \bar{x}_0 - \bar{\mu}_0^T P_0 \bar{\mu}_0
\]

\[+
\int_0^T \left[ \alpha \bar{x}^T W \bar{x} + \alpha \bar{\mu}^T W \bar{\mu} - \nu^T Q^{-1} \nu - w^T R^{-1} w \right] dt < 0
\]

(2.18)

Define

\[
V_i = \mu_{i_{\text{max}}}^2 D_i^T D_i \bar{x} - \mu_{i_{\text{max}}}^2 \bar{x}^T D_i^T D_i \bar{x} \leq 0
\]

(2.19)

\[
V_\Phi = \Phi^T \Phi - \bar{x}^T G^T G \bar{x} \leq 0
\]

(2.20)
From S-Procedure lemma, \( J_1 < 0 \) iff \( \exists \varepsilon_i, \eta > 0 \) such that

\[
J_2 := J_1 - \int_0^T (\sum \varepsilon_i V_i + \eta V_\phi) dt < 0
\]  
(2.21)

Since we have the constraints (2.9) in our optimization problem, we define the Hamiltonian or the augmented cost function as

\[
\bar{J} = J_2 + 2 \int_0^T \lambda^T [(A - LC)\ddot{x} + E\ddot{\Phi} + B\mathcal{D}_x \ddot{\mu} + B\mathcal{D}_x \ddot{\mu} - L\nu + Fw - \ddot{x}] dt
\]  
\[+ 2 \int_0^T \lambda_{\mu}^T [-L\mu C\ddot{x} - L\mu \nu - \ddot{\mu}] dt
\]  
(2.22)

Equation (2.22) can be rewritten as

\[
\bar{J} = -\ddot{x}_0^T P_0 \ddot{x}_0 - \ddot{\mu}_0^T P_0 \ddot{\mu}_0 + \int_0^T (\mathcal{H} - 2\lambda_{\mu}^T \ddot{\mu} - 2\lambda^T \ddot{x}) dt
\]  
(2.23)

where

\[
\mathcal{H} = \alpha \ddot{x}^T W \ddot{x} + \alpha \ddot{\mu}^T W_{\mu} \ddot{\mu} - \nu^T Q^{-1} \nu - w^T R^{-1} w
\]  
\[+ 2\lambda^T [(A - LC)\ddot{x} + E\ddot{\Phi} + B\mathcal{D}_x \ddot{\mu} + B\mathcal{D}_x \ddot{\mu} - L\nu + Fw]
\]  
\[- 2\lambda_{\mu}^T [L\mu C\ddot{x} + L\mu \nu] - \sum \varepsilon_i [(\mu_i^2 - \mu_{i-max}) \ddot{x}^T D_i^T D_i \ddot{x}]
\]  
\[-\eta[\ddot{\Phi}^T \ddot{\Phi} - \ddot{x}^T G^T G \ddot{x}]
\]

This can be reformulated as

\[
\mathcal{H} = \ddot{x}^T W' \ddot{x} + \ddot{\mu}^T W_{\mu}' \ddot{\mu} - \nu^T Q^{-1} \nu - w^T R^{-1} w
\]  
\[+ 2\lambda^T [(A - LC)\ddot{x} + E\ddot{\Phi} + B\mathcal{D}_x \ddot{\mu} + B\mathcal{D}_x \ddot{\mu} - L\nu + Fw]
\]  
\[- 2\lambda_{\mu}^T [L\mu C\ddot{x} + L\mu \nu] - \sum \varepsilon_i [\mu_i^2 \ddot{x}^T D_i^T D_i \ddot{x} - \eta \ddot{\Phi}^T \ddot{\Phi}]
\]  
(2.24)
where

\[ W' = \alpha W + \sum \varepsilon_i \mu_{i_{\text{max}}}^2 D_i^T D_i + \eta G^T G \]

\[ W'_\mu = \alpha W_\mu \]

Integrating \( \lambda^T \dot{x} \) and \( \lambda_{\mu}^T \dot{\mu} \) by parts, we have

\[
\bar{J} = -\dot{x}_0^T P_0 x_0 - \dot{\mu}_0^T P_0 \bar{\mu}_0 - 2[\lambda^T(T)\dot{x}(T) - \lambda^T(0)\dot{x}_0] \\
-2[\lambda_{\mu}^T(T)\dot{\mu}(T) - \lambda_{\mu}^T(0)\dot{\mu}_0] + \int_0^T (\mathcal{H} + 2\bar{\lambda}^T \bar{x} + 2\bar{\lambda}_{\mu}^T \bar{\mu}) \, dt
\]

(2.25)

To solve the dynamic constrained optimization problem, we wish to find stationary points of \( \bar{J} \) with respect to its variables. We find these stationary points, first with respect to initial state and parameter estimates in what follows. Setting \( (\partial \bar{J} / \partial x_0)^T = 0, \)

\[ \lambda_0 = P_0 \ddot{x}_0 \]  
(2.26)

Setting \( (\partial \bar{J} / \partial \mu_0)^T = 0, \)

\[ \lambda_{\mu 0} = P_{\mu 0} \ddot{\mu}_0 \]  
(2.27)

We now assume

\[ \bar{\lambda}: = \begin{bmatrix} \lambda \\ \lambda_{\mu} \end{bmatrix} = \bar{P} \begin{bmatrix} \ddot{x} \\ \ddot{\mu} \end{bmatrix} \]  
(2.28)

Hence from (2.26) and (2.27) it is concluded that
\[
\bar{P}(0) = \begin{bmatrix} P_0 & 0 \\ 0 & P_{\mu 0} \end{bmatrix}
\]  

(2.29)

The observer is also expected to show robustness against noise and disturbances present in the system. Therefore, it is necessary to find its stationary points with respect to existing disturbances and noise. To this end, we set \(- (\partial \bar{J}/\partial w)^T = 0,\)

\[
F^T \lambda = R^{-1} w
\]  

(2.30)

Setting \(- (\partial \bar{J}/\partial \nu)^T = 0,\)

\[
Q^{-1} \nu + L^T \lambda + L_{\mu}^T \lambda_{\mu} = 0
\]  

(2.31)

Defining the observer gain vector as

\[
\bar{L}^T = [L^T \ L_{\mu}^T]
\]  

(2.32)

We conclude from (2.31) and (2.32)

\[
Q^{-1} \nu + \bar{L}^T \bar{\lambda} = 0
\]  

(2.33)

Setting \((\partial \bar{J}/\partial L)^T = 0,\)

\[
C \bar{x} + \nu = 0
\]  

(2.34)

Defining \(\bar{C} = [C \ 0]\) and using (2.34) and (2.28) in (2.33),

\[
-Q^{-1} \bar{C} P^{-1} \bar{\lambda} + \bar{L}^T \bar{\lambda} = 0
\]  

(2.35)
Since the above equation needs to hold for all $\bar{\lambda}$,

$$\bar{L} = \bar{P}^{-1}\bar{C}^{T}Q^{-1} \quad (2.36)$$

For optimality with respect to the nonlinearity, set $\left(\partial \bar{J}/\partial \Phi\right)^{T} = 0$,

$$E^{T} \lambda = \eta \bar{\Phi} \quad (2.37)$$

Setting $\left(\partial \bar{J}/\partial \mu_{i}\right)^{T} = 0$,

$$\bar{x}^{T}D_{i}^{T}B_{i}^{T}\lambda - \epsilon_{i}\mu_{i}\bar{x}^{T}D_{i}^{T}D_{i}\bar{x} = 0 \quad (2.38)$$

Since $\bar{x}^{T}D_{i}^{T} = D_{i}\bar{x}$ is a scalar and the above equation needs to hold for all $\bar{x}$,

$$B_{i}^{T}\lambda - \epsilon_{i}\mu_{i}D_{i}\bar{x} = 0 \quad (2.39)$$

Setting $\left(\partial \bar{J}/\partial \bar{x}\right)^{T} = 0$,

$$W'\bar{x} + (A - LC)^{T}\lambda + \sum \mu_{i}D_{i}^{T}B_{i}^{T}\lambda - C^{T}L_{\mu}^{T}\lambda_{\mu} - \sum \epsilon_{i}\mu_{i}^{2}D_{i}^{T}D_{i}\bar{x} + \dot{\lambda} = 0 \quad (2.40)$$

From (2.39) it can be concluded that

$$W'\bar{x} + (A - LC)^{T}\lambda - C^{T}L_{\mu}^{T}\lambda_{\mu} + \dot{\lambda} = 0 \quad (2.41)$$

or

$$W'[I \ 0]\bar{P}^{-1}\bar{\lambda} + (A - LC)^{T}[I \ 0]\bar{\lambda} - C^{T}L_{\mu}^{T}[0 \ I]\bar{\lambda} + [I \ 0]\dot{\lambda} = 0 \quad (2.42)$$

Setting $\left(\partial \bar{J}/\partial \bar{\mu}_{i}\right)^{T} = 0$
\[ W_{\mu} \hat{\mu}_{i} + \hat{x}^T D_i^T B_i^T \lambda + \dot{\lambda}_{\mu-i} = 0 \]  

(2.43)

Stacking the above equation for different i-s,

\[ W_{\mu} \hat{\mu} + D_\hat{x} B^T \lambda + \dot{\lambda}_{\mu} = 0 \]  

(2.44)

or

\[ W_{\mu} [0 \ I] \bar{P}^{-1} \bar{\lambda} + D_\hat{x} B^T [I \ 0] \bar{\lambda} + [0 \ I] \dot{\lambda}_\mu = 0 \]  

(2.45)

Combining (2.42) and (2.45), it follows that

\[
\begin{bmatrix}
W' & 0 \\
0 & W_{\mu}'
\end{bmatrix} \bar{P}^{-1} \bar{\lambda} + \begin{bmatrix}
(A - LC)^T \\
D_\hat{x} B^T
\end{bmatrix} - C^T L_{\mu}^T \\
0
\end{bmatrix} \bar{\lambda} + \dot{\bar{\lambda}} = 0
\]  

(2.46)

Substituting for \( \dot{\lambda} \) from (2.28), we arrive at

\[
\begin{bmatrix}
W' & 0 \\
0 & W_{\mu}'
\end{bmatrix} \bar{P}^{-1} \bar{\lambda} + \begin{bmatrix}
(A - LC)^T \\
D_\hat{x} B^T
\end{bmatrix} - C^T L_{\mu}^T \\
0
\end{bmatrix} \bar{\lambda} + \dot{\bar{\lambda}} \bar{P}^{-1} \bar{\lambda} \\
+ \bar{P} \left[ (A - LC) \bar{x} + E \Phi + B D_\hat{x} \bar{\mu} + B D_\hat{x} \bar{\bar{\mu}} - L \nu + F_w \right] = 0
\]  

(2.47)

Using (2.30), (2.37) and (2.39), we find

\[
\begin{bmatrix}
W' & 0 \\
0 & W_{\mu}'
\end{bmatrix} \bar{P}^{-1} \bar{\lambda} + \begin{bmatrix}
(A - LC)^T \\
D_\hat{x} B^T
\end{bmatrix} - C^T L_{\mu}^T \\
0
\end{bmatrix} \bar{\lambda} + \dot{\bar{\lambda}} \bar{P}^{-1} \bar{\lambda} \\
+ \bar{P} \left[ (A - LC) \bar{x} + \frac{1}{\eta} E E^T \lambda + \sum \frac{1}{\epsilon_i} B_i B_i^T \lambda + B D_\hat{x} \bar{\mu} - L \nu + F R F^T \lambda \right] = 0
\]  

(2.48)

after some manipulation
\[
\begin{bmatrix}
   W' & 0 \\
   0 & W'_\mu
\end{bmatrix}
\bar{p}^{-1}\bar{\lambda} + \left[
\begin{array}{cc}
   (A - LC)^T & -C^T L_{\mu}^T \\
   D_x B^T & 0
\end{array}
\right] \bar{\lambda} + \bar{P} \bar{p}^{-1} \bar{\lambda} \\
+ \bar{p} \left[
\begin{array}{cc}
   A & B D_x \\
   0 & 0
\end{array}
\right] \bar{p}^{-1} + \bar{p} \left[
\begin{array}{cc}
   \frac{1}{\eta} EE^T + \frac{1}{\epsilon_i} B_i B_i^T + FRF^T & 0 \\
   0 & 0
\end{array}
\right] \bar{\lambda} = 0
\]

(2.49)

Since the above equation needs to hold for all \( \bar{\lambda} \),

\[
\begin{bmatrix}
   W' & 0 \\
   0 & W'_\mu
\end{bmatrix}
\bar{p}^{-1} + \left[
\begin{array}{cc}
   (A - LC)^T & -C^T L_{\mu}^T \\
   D_x B^T & 0
\end{array}
\right] + \bar{p} \left[
\begin{array}{cc}
   \frac{1}{\eta} EE^T + \frac{1}{\epsilon_i} B_i B_i^T + FRF^T & 0 \\
   0 & 0
\end{array}
\right] = 0
\]

(2.50)

Multiplying by \( \bar{P} \) to the right

\[
\begin{bmatrix}
   W' & 0 \\
   0 & W'_\mu
\end{bmatrix} + \left[
\begin{array}{cc}
   (A - LC)^T & -C^T L_{\mu}^T \\
   D_x B^T & 0
\end{array}
\right] + \bar{P} \left[
\begin{array}{cc}
   \frac{1}{\eta} EE^T + \frac{1}{\epsilon_i} B_i B_i^T + FRF^T & 0 \\
   0 & 0
\end{array}
\right] = 0
\]

(2.51)

Using (2.36),

\[
\begin{bmatrix}
   W' & 0 \\
   0 & W'_\mu
\end{bmatrix} - \left[C^T \right] Q^{-1} [C 
0] + \left[
\begin{array}{cc}
   A & 0 \\
   D_x B^T & 0
\end{array}
\right] \bar{P} + \bar{p} \\
+ \bar{P} \left[
\begin{array}{cc}
   \frac{1}{\eta} EE^T + \frac{1}{\epsilon_i} B_i B_i^T + FRF^T & 0 \\
   0 & 0
\end{array}
\right] = 0
\]

(2.52)

Now

32
Therefore we have

\[
\begin{bmatrix}
A & 0 \\
B\mathcal{D}_\hat{x} & 0
\end{bmatrix}
\begin{bmatrix}
\bar{p} \\
0
\end{bmatrix}
+ \begin{bmatrix}
A & B\mathcal{D}_\hat{x} \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
\bar{p} \\
0
\end{bmatrix}
= \begin{bmatrix}
P_{11}A & P_{11}B\mathcal{D}_\hat{x} \\
P_{12}^T A & P_{12}^T B\mathcal{D}_\hat{x}
\end{bmatrix}
\begin{bmatrix}
P_{11} & P_{12} \\
P_{12}^T & P_{12}^T
\end{bmatrix}
\] (2.54)

Using (2.54) in (2.52), we arrive at

\[
\frac{d}{dt} \bar{P} = -\begin{bmatrix}
\bar{W} & P_{11}B\mathcal{D}_\hat{x} + A^T P_{12} \\
0 & P_{12}^T B\mathcal{D}_\hat{x} + \mathcal{D}_\hat{x} B^T P_{12} + W'_\mu
\end{bmatrix}
\begin{bmatrix}
\bar{p} \\
0
\end{bmatrix}
\begin{bmatrix}
1 & 0 \\
\eta & 0
\end{bmatrix}
\begin{bmatrix}
EE^T + B\epsilon^{-1}B^T + FRF^T & 0 \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
\bar{p} \\
0
\end{bmatrix}
\] (2.55)

where

\[
\bar{W} = aW + P_{11}A + A^T P_{11} + \sum \epsilon_i \mu_i^2 \mathcal{D}_i D_i + \eta G^T G - C^T Q^{-1}C
\]

This is equivalent to equation (2.15) and proves the first part of the theorem.

Now if \( \bar{P}' = \bar{P}^{-1} \),

\[
\bar{L}' = Q^{-1}C
\begin{bmatrix}
P_{11}' & P_{12}' \\
P_{22}' & 0
\end{bmatrix}
= \begin{bmatrix}
Q^{-1}CP_{11}' & Q^{-1}CP_{12}'
\end{bmatrix}
\] (2.56)

since
We find

\[ P_{11}^I = [P_{11} - P_{12}P_{22}^{-1}P_{12}^T]^{-1} \]
\[ P_{21}^I = -P_{22}^{-1}P_{12}^T[P_{11} - P_{12}P_{22}^{-1}P_{12}^T]^{-1} \]

Equation (2.55) is a matrix Riccati differential equation. The observer is stable if and only if \( P \) remains positive definite for all \( t \in [0 \ T] \). Solution to the equation (2.55) gives stable observer gain vector at each time provided that \( P \) remains positive definite. That equation is dependent on the state estimation \( \hat{x} \). Therefore, the above theorem requires online solution to the matrix differential Riccati equation which is computationally intensive. One way to avoid solving this equation in real-time is to find a stable steady state solution. Finding such a solution seems impossible since the adaptation law requires time-varying matrix \( P \). We can, however, modify our cost function in a way to obtain a possible steady state solution. The following theorem presents this new design methodology.

**Theorem 2.2**

*For the system (2.3) if \( \gamma_{2i} \geq |D_i\hat{x}| \geq \gamma_{1i}, \max(|(D_i\hat{x})(D_i\hat{x})^{-1}|) \leq \rho_{i-max}, \) the observer (2.9) can guarantee a performance

\[
J = \frac{\int_0^T \hat{x}^TW\hat{x}dt + \int_0^T \hat{\mu}^TD_\hat{x}W_\mu D_\hat{x}\hat{\mu}dt}{\hat{x}_0^TP_0\hat{x}_0 + \hat{\mu}_0^TP_\mu\hat{\mu}_0 + \int_0^T v^TQ^{-1}vdt + \int_0^T w^TR^{-1}wdt} < \frac{1}{\alpha} \quad (2.59)
\]

after a sufficiently large time \( T \), if \( \exists \epsilon = \text{diag} \left( \epsilon_1, \epsilon_2, \ldots, \epsilon_r \right) > 0, \eta > 0, P_{11} > 0, P_{21}' \) such that \( \forall \rho_i \leq \rho_{i-max} \) and
\[
\begin{bmatrix}
M_{11} & M_{12} & P_{11}B & P_{11}E & P_{11}F \\
* & M_{22} & P_{21}'B & P_{21}'E & P_{21}'F \\
* & * & -\varepsilon & 0 & 0 \\
* & * & * & -\eta I & 0 \\
* & * & * & * & -R^{-1}
\end{bmatrix} < 0 \tag{2.60}
\]

where

\[
M_{11} = \alpha W + \sum \varepsilon_i \mu_i^2 D_i^T D_i + \eta G^T G - C^T Q^{-1} C + P_{11} A + A^T P_{11}
\]

\[
M_{12} = P_{11} B + \rho P_{12}' + A^T P_{12}'
\]

\[
M_{22} = P_{12}'^T B + B^T P_{12}' + \alpha W_\mu
\]

\[
\rho = \text{diag}(\rho_i)
\]

and observer gains are chosen as

\[
L = [P_{11} - P_{12}' D_\hat{x} P_{22}^{-1} D_\hat{x}^T P_{12}']^{-1} C^T Q^{-1}
\]

\[
L_\mu = -P_{22}^{-1} D_\hat{x} P_{12}' [P_{11} - P_{12}' D_\hat{x} P_{22}^{-1} D_\hat{x}^T P_{12}']^{-1} C^T Q^{-1}
\]

\[
P_{22} \text{ is chosen such that } P_{22} > \gamma_2 P_{12}'^T P_{11}^{-1} P_{12}' \gamma_2, \quad \gamma_2 = \text{diag}(\gamma_{2i}).
\]

Proof: Similar to Theorem 2.1., the cost function inequality (2.13) can be rewritten as

\[
J_1 = -\bar{x}_0^T P_0 \bar{x}_0 - \bar{\mu}_0^T P_0 \bar{\mu}_0
\]

\[
+ \int_0^T [\alpha \bar{x}^T W \bar{x} + \alpha \bar{\mu}^T D_\hat{x} W_\mu D_\hat{x} \bar{\mu} - \nu^T Q^{-1} \nu - \omega^T R^{-1} \omega] dt < 0 \tag{2.62}
\]

Define
\[ V_i = \mu_i^2 \ddot{x}_i^T D_i^T D_i \ddot{x}_i - \mu_{i-max}^2 \ddot{x}_i^T D_i^T D_i \ddot{x}_i \leq 0 \quad (2.63) \]

\[ V_{\phi} = \Phi^T \ddot{\Phi} - \ddot{x}_i^T G^T G \ddot{x}_i \leq 0 \quad (2.64) \]

From S-Procedure lemma, \( J_1 < 0 \) iff \( \exists \varepsilon_i, \eta > 0, \Xi = \Xi^T > 0 \) such that

\[ J_2 := J_1 + \int_0^T \left( -\sum \varepsilon_i V_i - \eta V_{\phi} + [\ddot{x}_i^T \ddot{\mu}^T] \Xi [\ddot{x}_i^T \ddot{\mu}^T]^T \right) dt < 0 \quad (2.65) \]

It should be noted that (2.65) differs from (2.21) by \( V_\Xi = \int_0^T ([\ddot{x}_i^T \ddot{\mu}^T] \Xi [\ddot{x}_i^T \ddot{\mu}^T]^T) dt \). Following the same steps as in theorem 1 and finding the stationary points of \( \bar{J} \) with respect to its variables, we arrive at a slightly different differential Riccati equation

\[
\begin{align*}
\frac{d}{dt} \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} = & - \begin{bmatrix} \bar{W} & P_{11} B \mathcal{D} \ddot{x} + A^T P_{12} \\ P_{12}^T B^T \mathcal{D} \ddot{x} + \mathcal{D} \ddot{x} B^T P_{12} + \mathcal{D} \ddot{x} W_{\mu}' \mathcal{D} \ddot{x} \end{bmatrix} \\
& - P \begin{bmatrix} \frac{1}{\eta} EE^T + \sum \frac{1}{\varepsilon_i} B_i B_i^T + FRF^T & \ast \\ \ast & 0 \end{bmatrix} P - \Xi
\end{align*}
\]

where

\[
\bar{W} = \alpha W + P_{11} A + A^T P_{11} + \sum \epsilon_i \mu_{i-max}^2 D_i^T D_i + \eta G^T G - C^T Q^{-1} C \]

\[ W_{\mu}' = \alpha W_{\mu} \]

We intend to find a steady state solution so that online solution to the Riccati equation is not required. Suppose \( \exists P_{21}' \) s.t. \( \forall \rho_i \leq \rho_{i-max} \) and \( \rho = diag(\rho_i) \) such that
\[
\begin{bmatrix}
\bar{W} & P_{11}B + A^T P_{12} + \rho P'_{12} \\
* & P'_{12}^T B + B^T P'_{12} + W'_{\mu}
\end{bmatrix} + \\
\begin{bmatrix}
P_{11} & P'_{12} \\
P'_{21} & P_{22}
\end{bmatrix}
\begin{bmatrix}
\frac{1}{\eta} EE^T + \sum \frac{1}{\epsilon_i} B_i B_i^T + FR F^T & 0 \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
P_{11} & P'_{12} \\
P'_{21} & P_{22}
\end{bmatrix} < 0
\] (2.67)

then at time \( t_0 \) for a choice \( P_{12}(t_0) = D_x^T P_{12}', \rho = (\dot{D}_x)(D_x)^{-1} \) and an appropriate \( \Xi \)

\[
\begin{bmatrix}
\bar{W} & P_{11}B'D_x + A^T P_{12} + P'_{12}\dot{D}_x \\
* & P'_{12}^T B'D_x + D_xB^T P_{12} + D_xW'_{\mu}D_x
\end{bmatrix} + \bar{p}
\begin{bmatrix}
\frac{1}{\eta} EE^T + \sum \frac{1}{\epsilon_i} B_i B_i^T + FR F^T & 0 \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
P_{11} & P'_{12} \\
P'_{21} & P_{22}
\end{bmatrix} = 0
\] (2.68)

Substituting (2.68) into (2.66) we find

\[
\frac{d}{dt} \begin{bmatrix} P_{11} & P_{12} \\ P'_{21} & P_{22} \end{bmatrix} = \begin{bmatrix} 0 & P'_{12}\dot{D}_x \\ * & 0 \end{bmatrix}
\] (2.69)

Hence \( P_{12}(t > t_0) = D_x^T P_{12}' = P_{21}^T \).

Taking the Schur’s complement, (2.67) can be written as the LMI (2.60). The inequalities (2.67) and (2.60) are independent of \( \hat{x}(t) \). This significantly simplifies the design procedure since the existence can be checked by solving the pertinent LMI. Once checked, one can use the relations (2.16) to calculate the gains. Also, this is computationally more efficient because differential Riccati equation needs not be solved online.

37
2.4.3. **Extension to general sector-bounded nonlinearity**

For the class of state-space models of equation (2.3), assume that the nonlinearity $\Phi(u, x)$ satisfies a generalized sector bounded condition proposed in Vijayaraghavan [55], i.e.

$$
\tilde{\Phi}^T \tilde{\Phi} + (\gamma_1 - \gamma_2) \tilde{\Phi}^T G \tilde{x} - \gamma_1 \gamma_2 \tilde{x}^T G^T G \tilde{x} \leq 0
$$

(2.70)

This type of nonlinearity is highly general. It encompasses a variety of different nonlinearities including Lipschitz, one-sided Lipschitz, bounded Jacobian, monotonically increasing and dissipative nonlinearities. The rest of the system model considered here is the same as in section 2.1. Here we follow the same procedure as in previous sections to extract conditions for the adaptive observer applicable to the sector-bounded nonlinear systems.

Similar to theorem 1, the cost function inequality (2.13) can be rewritten as

$$
J_1 = -\tilde{x}_0^T P_0 \tilde{x}_0 - \tilde{\mu}_0^T P_0 \tilde{\mu}_0 + \int_0^T [\alpha \tilde{x}^T W \tilde{x} + \alpha \tilde{\mu}^T D_\chi W \tilde{\mu} D_\chi \tilde{\mu} - \nu^T Q^{-1} \nu - w^T R^{-1} w] dt < 0
$$

(2.71)

Define

$$
V_i = \mu_i^2 \tilde{x}^T D_i^T D_i \tilde{x} - \mu_{i_{-\text{max}}}^2 \tilde{x}^T D_i^T D_i \tilde{x} \leq 0
$$

(2.72)

$$
V_\phi = \tilde{\Phi}^T \tilde{\Phi} + (\gamma_1 - \gamma_2) \tilde{\Phi}^T G \tilde{x} - \gamma_1 \gamma_2 \tilde{x}^T G^T G \tilde{x} \leq 0
$$

(2.73)

From S-Procedure lemma, $J_1 < 0$ if and only if $\exists \varepsilon, \eta > 0$ such that

$$
J_2 := J_1 - \int_0^T (\varepsilon V_\mu + \eta V_\phi) dt < 0
$$

(2.74)
Following similar steps as in Theorem 2.1., we need to find stationary points of the augmented cost function with respect to the nonlinearity. Setting \( \left( \frac{\partial \bar{J}}{\partial \bar{\Phi}} \right)^T = 0 \)

\[
E^T \lambda = \eta \bar{\Phi} + \frac{\eta}{2} (\gamma_1 - \gamma_2) G \bar{x}
\] (2.75)

Following the same procedure, we finally arrive at

\[
\frac{d}{dt} \bar{P} = - \begin{bmatrix}
\bar{W} & P_{11} B \mathcal{D} \bar{x} + A^T P_{12} \\
* & P_{12}^T B \mathcal{D} \bar{x} + \mathcal{D} \mathcal{S} B^T P_{12} + W' \\
\end{bmatrix}
\]

\[
- \bar{P} \begin{bmatrix}
\frac{1}{\eta} EE^T & \frac{1}{\epsilon} BB^T + FR \mathcal{F}^T \\
* & 0
\end{bmatrix}
\bar{P}
\] (2.76)

with \( A' \) and \( \bar{W} \) defined as

\[
A' = A - \frac{1}{2} (\gamma_1 - \gamma_2) EG
\]

\[
\bar{W} = \alpha W + P_{11} A' + A'^T P_{11} + \epsilon \mu_{\text{max}}^2 D^T D + \eta G^T G - C^T Q^{-1} C + \frac{\eta}{4} (\gamma_1 - \gamma_2)^2 G^T G
\]

This equation is similar to equation (2.55) except for the definitions of \( A' \) and \( \bar{W} \). Applying the methods of Theorem 2.2, equation (2.60) is also applicable here with the recently defined \( A' \) and \( \bar{W} \).

### 2.4.4. **Time-varying unknown parameter**

Suppose the unknown parameter in system (2.3) is time varying. Then the dynamics of the estimation error for this observer can be written as
\[\dot{x} = (A - LC)x + E\ddot{\phi} + B'D\ddot{x}\mu + B'D\ddot{x}\tilde{\mu} - Lv + Fw\]
\[\dot{\mu} = \dot{\mu} - L_\mu Cx - L_\mu v\]

where \(\dot{\mu}\) is a vector of \(\dot{\mu}_i\) values. We will assume that in addition to the bound on the parameter variation, a bound on its rate of change is also known \textit{a priori}

\[|\dot{\mu}_i| < \dot{\mu}_{i\text{-max}}\]  

\textbf{Theorem 2.3}

For the system (2.3) with a time-varying unknown parameter satisfying (2.78), suppose \(\gamma_2 \geq |D_\dot{x}| \geq \gamma_1\), \(\max\left(|(D_\dot{x})(D_\ddot{x})^{-1}|\right) = \rho_{i\text{-max}}\), the observer (2.9) can guarantee a performance

\[J = \frac{J_{Nr}}{J_{Dr}} < \frac{1}{\alpha}\]

with

\[J_{Nr} = \int_0^T \ddot{x}^TW\ddot{x}dt + \int_0^T \ddot{\mu}^TD_\mu W_\mu D_\ddot{x}\ddot{\mu}dt\]

\[J_{Dr} = \ddot{x}_0^TP_0\ddot{x}_0 + \ddot{\mu}_0^TP_\mu\ddot{\mu}_0 + \int_0^T v^TQ^{-1}vdT + \int_0^T w^TR^{-1}wdT + \int_0^T \mu_{max}^T D_\dot{x} S^{-1} D_\ddot{x}\ddot{\mu}_{max}dt\]

after a sufficiently large time \(T\), if \(\exists \epsilon = \text{diag}(\epsilon_1, \epsilon_2, ..., \epsilon_r) > 0\), \(\eta > 0\), \(P_{11} > 0\), \(P_{22} > 0\) and \(P'_{21}\) such that \(\forall |\rho_i| \leq \rho_{i\text{-max}}\)
\[
\begin{bmatrix}
M_{11} & M_{12} & P_{11}B & P_{11}E & P_{11}F & P_{12}^T \\
M_{12}^T & M_{22} & P_{21}^T & P_{21}^T & P_{21}^T & P_{22}^T \\
B^T P_{11} & B^T P_{12}^T & -\epsilon & 0 & 0 & 0 \\
E^T P_{11} & E^T P_{12}^T & 0 & -\eta I & 0 & 0 \\
F^T P_{11} & F^T P_{12}^T & 0 & 0 & -R^{-1} & 0 \\
P_{12}^T & P_{22}^T & 0 & 0 & 0 & -S
\end{bmatrix} < 0
\] (2.80)

where

\[
M_{11} = \alpha W + \sum_{i} \epsilon_i \mu_{i_{\text{max}}} D_i^T D_i + \eta G^T G - C^T Q^{-1} C + P_{11} A + A^T P_{11}
\]

\[
M_{12} = P_{11} B + \rho P_{21}^T + A^T P_{21}^T
\]

\[
M_{22} = P_{21}^T B + B^T (P_{21}^T) + \alpha W\mu
\]

\[
\rho = \text{diag}(\rho_i)
\]

and observer gains are chosen as

\[
L = \left[P_{11} - P_{12}^T D\hat{x} P_{22}^{-1} D\hat{x} P_{12}^T\right]^{-1} C^T Q^{-1}
\] (2.81)

\[
L_\mu = -P_{22}^{-1} D\hat{x} P_{12}^T \left[P_{11} - P_{12}^T D\hat{x} P_{22}^{-1} D\hat{x} P_{12}^T\right]^{-1} C^T Q^{-1}
\]

Proof: Similar to theorem 1, the cost function inequality (2.79) can be rewritten as

\[
\begin{align*}
J_1 &= -\tilde{x}_0^T P_0 \tilde{x}_0 - \tilde{\mu}_0^T P_0 \tilde{\mu}_0 \\
&\quad + \int_0^T \left[\alpha \tilde{x}^T W \tilde{x} + \alpha \tilde{\mu}^T D\hat{x} W_\mu D\hat{x} \tilde{\mu} - v^T Q^{-1} v - w^T R^{-1} w \\
&\quad - \mu_{\text{max}}^T D\hat{x} S^{-1} D\hat{x} \mu_{\text{max}}\right] dt < 0
\end{align*}
\]
Define

\[ V_i = \mu_i^2 \bar{x}^T D_i^T D_i \bar{x} - \mu_i \max \bar{x}^T D_i^T D_i \bar{x} \leq 0 \]  

(2.83)

and

\[ V_\Phi = \Phi^T \Phi - \bar{x}^T G^T G \bar{x} \leq 0 \]  

(2.84)

and

\[ V_{\mu} = \mu^T Z \mu - \mu_{\max}^T Z \mu_{\max} \leq 0 \]  

(2.85)

From S-Procedure lemma, \( J_1 < 0 \) iff \( \exists \varepsilon_i, \eta > 0, \Xi = \Xi^T > 0 \) such that

\[ J_2 = J_1 + \int_0^T \left( -\sum \varepsilon_i V_i - \eta V_\Phi - V_{\mu} + [\bar{x}^T \ \bar{\mu}^T] \Xi [\bar{x}^T \ \bar{\mu}^T]^T \right) dt < 0 \]  

(2.86)

The augmented cost function can be written as

\[ \bar{J} = -\bar{x}_0^T P_0 \bar{x}_0 - \bar{\mu}_0^T P_0 \bar{\mu}_0 + \int_0^T (\mathcal{H} - 2\lambda_{\mu}^T \bar{\mu} - 2\lambda^T \bar{x}) dt \]  

(2.87)

where
\[ H = \alpha \ddot{x}^T W \ddot{x} + \alpha \dot{\mu}^T W \mu - \nu^T Q^{-1}\nu - w^T R^{-1}w - \dot{\mu}_{\text{max}}^T (D_{\dot{x}} S D_{\dot{x}})^{-1} \mu_{\text{max}} + 2\lambda^T [(A - LC) \ddot{x} + E \Phi + B'D_{\dot{x}} \mu + B'D_{\dot{x}} \mu - L \nu + Fw] \]
\[-2\lambda^T [L \mu C \ddot{x} + L \nu \nu - \dot{\mu}] - \sum \epsilon_i [(\mu_i^2 - \mu_i^{\text{max}}) \dddot{x}^T D_i^T D_i \dddot{x}] \]
\[-\mu^T Z \dddot{x} + \mu_{\text{max}}^T Z \mu_{\text{max}} - \eta [\Phi^T \Phi - \dddot{x}^T G^T G \dddot{x}] + [\dddot{x}^T \dddot{\mu}] \Xi [\dddot{x}^T \dddot{\mu}]^T \]

To ensure \( \mu_{\text{max}}^T Z \mu_{\text{max}} - \mu_{\text{max}}^T D_{\dot{x}} S^{-1} D_{\dot{x}} \mu_{\text{max}} = 0 \) we choose \( Z = D_{\dot{x}} S^{-1} D_{\dot{x}} \). Following the same steps as in theorem 1, we set to zero all the previously mentioned partial derivatives as well as the derivative with respect to the new defined variable \( \dddot{\mu} \). To find the stationary points of \( \dddot{J} \) with respect to \( \dddot{\mu} \), we set \( (\partial \dddot{J} / \partial \dddot{\mu})^T = 0 \) which gives

\[ \dddot{\mu} = Z^{-1} \lambda_{\mu} = D_{\dot{x}}^{-1} S D_{\dot{x}}^{-1} \lambda_{\mu} \] (2.88)

Following the steps of theorem 1, we find

\[ \frac{d}{dt} \dddot{p} = - \begin{bmatrix} \dddot{W} & P_{11} B D_{\dot{x}} + A^T P_{21}^T \\ D_{\dddot{x}} B^T P_{11} + P_{21} A & P_{21} B D_{\dot{x}} + D_{\dddot{x}} B^T P_{21}^T + D_{\dddot{x}} W_{\mu}^T D_{\dddot{x}} \end{bmatrix} \]
\[-\dddot{p} \begin{bmatrix} 1 & EE^T + B(e^{-1}B)^T + FRF^T \\ 0 & D_{\dddot{x}}^{-1} S D_{\dddot{x}}^{-1} \end{bmatrix} \dddot{p} = Z \] (2.89)

where

\[ \dddot{W} = \alpha W + P_{11} A + A^T P_{11} + \sum \epsilon_i \mu_i^{\text{max}} D_i^T D_i + \eta G^T G - C^T Q^{-1} C \]

\[ W_{\mu}' = \alpha W_{\mu} \]

Following the proof of theorem 2, we arrive at LMI (2.80) and equation (2.81) for calculating the observer gains.
2.5. Numerical example

In this section, two models of nonlinear systems are employed to illustrate the application of the proposed observer. The models include uncertainties in the system parameters, as well as input disturbances and measurement noise.

2.5.1. Example 1. 4th order nonlinear system

Consider the following 4th order nonlinear system in the framework (2.3)

\[
\dot{x} = Ax + \sum \mu_i B_i D_i x + E\Phi(u, x) + Fw + Hu
\]

\[
y = Cx + \nu
\]

\[
A = \begin{bmatrix}
0 & 1 & 0 & 0 \\
-25 & -1 & 15 & 0.3 \\
0 & 0 & 0 & 1 \\
15 & 0.3 & -15 & -0.3
\end{bmatrix}
\]

\[
\mu_1 = 1, \quad \mu_{1-max} = 2.5,
\]

\[
B_1 = [0 \ 0 \ 0 \ 1]^T, \quad D_1 = [-0.25 \ -0.0033 \ 0.25 \ 0.0033]^T
\]

\[
E = [0 \ 1 \ 0 \ 0]^T, \quad \Phi = \sin(x_2), \quad G = [0 \ 1 \ 0]^T
\]

\[
F = [0 \ 0 \ 0 \ 1]^T, \quad \omega \sim (0, 10^{-4})
\]

\[
H = [0 \ 0 \ 0 \ 1]^T, \quad u = 35\sin(2\pi t)
\]

\[
C = [1 \ 0 \ 0 \ 0], \quad \nu \sim \left(0, \begin{bmatrix}10^{-4} & 0 & 0 & 0 \end{bmatrix}ight)
\]

The model is representative of a two-mass spring system with damping. The considered unknown parameter models some uncertainty in one of the masses. The objective is to estimate the states of the system and the unknown parameter simultaneously. We intend to apply both methods of extended Kalman filtering and the method proposed by this work to compare their performance and time-efficiency.

State and parameter estimation in Lipschitz nonlinear system using EKF

In discrete time systems, the unknown parameter EKF is constructed by augmenting the unknown parameter to the states [24], [25], [56]. We will apply a similar
approach to continuous time systems by defining the augmented state vector $\tilde{x} = [x^T \mu]^T$.

It is customary to write the EKF system equation in the form

$$\begin{align*}
\dot{x} &= f(\tilde{x}, u) + w \\
y &= h(\tilde{x}) + v
\end{align*}$$

The EKF is then constructed as

$$\hat{x} = f(\hat{x}, u) + L(y - h(\tilde{x}))$$

where

$$L = PH^TQ^{-1}$$

$$\dot{P} = FP + PF^T - LHP + \bar{R}$$

$$F = \left. \frac{\partial f}{\partial \tilde{x}} \right|_{\tilde{x}, u}, H = \left. \frac{\partial h}{\partial \tilde{x}} \right|_{\tilde{x}}$$

In this example,

$$F = \left. \frac{\partial f}{\partial \tilde{x}} \right|_{\tilde{x}, u} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ -25 & -1 + \cos(\tilde{x}_2) & 15 & 0.3 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 15 - 0.25 \mu_1 & 0.3 - 0.0033 \mu_1 & -15 + 0.25 \mu_1 & -0.3 + 0.0033 \mu_1 & S \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$S = -0.25 \tilde{x}_1 - 0.0033 \tilde{x}_2 + 0.25 \tilde{x}_3 + 0.0033 \tilde{x}_4$$

$$H = \left. \frac{\partial h}{\partial \tilde{x}} \right|_{\tilde{x}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

As discussed earlier, EKF requires virtual parameter disturbance (VPD) in order to estimate the unknown parameter. Since there is no methodology for selecting this VPD-covariance, we will use VPD-covariance equal to the covariance of the disturbance $w$. Hence

$$\bar{R} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & R & 0 \\ 0 & 0 & 0 & 0 & R \end{bmatrix}$$
The simulations were carried out in Matlab® R2012b using Simulink® on an Intel(R) Core(TM) i5-4570 CPU @ 3.20 GHZ computer. The simulation for the combined system-observer took approximately 3.42 seconds. The overhead of simulating just the system was found to be 1.40 seconds. Figures 2-1, 2-2, 2-3 and 2-4 show the evolution of actual state values and their estimations while 2-6 shows the evolution of the parameter estimate (note the different time scale for parameter estimation). Also figure 2-5 indicates the state estimation errors. As observed in these figures, the filter shows satisfactory estimation of the states and the parameter.

![Figure 2.1. State 1 estimation using EKF](image-url)
Figure 2.2. State 2 estimation using EKF

Figure 2.3. State 3 estimation using EKF
Figure 2.4. State 4 estimation using EKF

Figure 2.5. State estimation errors in EKF method
Figure 2.6. Parameter estimation by EKF

State and parameter estimation in Lipschitz nonlinear system using algorithm proposed in this paper

Now we intend to apply the methods of section 2.4 to estimate the states and parameters of the system. Notice that in this example, \( n = 4 \), \( q = 1 \) and \( r = 1 \) and system matrices are

\[
A = \begin{bmatrix}
0 & 1 & 0 & 0 \\
-25 & -1 & 15 & 0.3 \\
0 & 0 & 0 & 1 \\
15 & 0.3 & -15 & -0.3
\end{bmatrix},
B_1 = \begin{bmatrix}
0 \\
0 \\
0 \\
1
\end{bmatrix},
D_1 = \begin{bmatrix}
-0.25 \\
-0.0033 \\
0.25 \\
0.0033
\end{bmatrix}^T,
\mu_1 = 1, \mu_{1-max} = 2.5
\]

\[
E = \begin{bmatrix}
1 \\
0 \\
0
\end{bmatrix}, \Phi = \sin(x_1)
\]

\[
F = \begin{bmatrix}
1 \\
0 \\
0
\end{bmatrix}, H = \begin{bmatrix}
1 \\
0 \\
0
\end{bmatrix}, C = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0
\end{bmatrix}
\]
\[ G = [0 \ 1 \ 0 \ 0]^T, R = 10^{-4}, Q = \begin{bmatrix} 10^{-4} & 0 \\ 0 & 10^{-4} \end{bmatrix} \]

We will assume that \( \max \left( |(D_i \hat{x})(D_i \hat{x})^{-1}| \right) = \rho_{i_{\text{max}}} = 0.2 \). Solving the LMI, we find \( \eta = 1.227 \times 10^3 \), \( \epsilon_1 = 1.269 \times 10^4 \) and

\[
P_{11} = 10^4 \times \begin{bmatrix} 3.450 & -0.023 & -2.077 & -0.038 \\ -0.023 & 0.132 & -0.047 & -0.008 \\ -2.077 & -0.047 & 2.003 & -0.051 \\ -0.038 & -0.008 & -0.051 & 0.125 \end{bmatrix}
\]

\[
P_{21} = [198.8 \ -124.3 \ -1231.3 \ -182.5]
\]

This simulation was carried out with \( x_0 = [20,40,40,80]^T, \hat{x}_0 = [0,0,0,0]^T \) and \( \hat{\mu}(0) = 2 \) using the same system as before. The simulation took approximately 0.018 seconds for solving the LMI and 2.70 seconds for the combined system-observer. Subtracting the overhead of simulating the system, we find that the present algorithm would results in 35% improvement in simulation time. Results of constructing the adaptive observer for this system are presented in the following figures. Figures 2-7, 2-8, 2-9 and 2-10 show the evolution of the actual state values and their estimations, while figure 2-12 shows the evolution of the parameter estimation. Furthermore, figure 2-11 indicates state estimation errors of the observer and figure 2-13 shows the cost function evolution with time. Compared to the EKF, the proposed observer is about 35% faster and is also able to estimate both the states and the unknown parameter with smaller convergence time. Also, since the initial guess includes significant error, the observer convergence to the actual values is faster than the EKF.
Figure 2.7. State 1 estimation using the proposed observer

Figure 2.8. State 2 estimation using the proposed observer
Figure 2.9. State 3 estimation using the proposed observer

Figure 2.10. State 4 estimation using the proposed observer
Figure 2.11. State estimation errors for the proposed observer

Figure 2.12. Parameter estimation by the proposed observer
Comparison between EKF and the proposed observer

EKF requires that the initial state estimates be close to the actual states. Hence the proposed observer will outperform the EKF for large error in initial state estimate. To illustrate, the results of the two simulations are compared for $x = [20, 40, 40, 80]^T$ and $\hat{x} = [0, 0, 0, 0]^T$ and the comparison of estimation errors are plotted below.
Figure 2.14. Comparison between state 1 estimation errors of EKF and the proposed observer

Figure 2.15. Comparison between state 2 estimation errors of EKF and the proposed observer
Figure 2.16. Comparison between state 3 estimation errors of EKF and the proposed observer

Figure 2.17. Comparison between state 4 estimation errors of EKF and the proposed observer
It is evident in the Figures 2-14 to 2-17 that the proposed algorithm outperforms EKF in terms of convergence speed for large errors in the initial state estimates. However, the extended Kalman filter outperforms the observer in noise attenuation. As stated earlier, this is due to the optimality of the Kalman filter and is dependent on the availability of accurate noise covariance matrix. In other words, if the covariance matrix for the noise was not readily available, or was not accurate enough, the noise attenuation level in the EKF would be less accurate.

2.5.2. Example 2. 3rd order nonlinear system

Consider the following dissipative nonlinear model of two rotating masses connected together via a gear train with a nonlinear viscous drag affecting the second rotating mass and linear damping affecting both masses and the gear train. This system IS formulated as a dissipative system with $n = 3$, $q = 1$, $r = 1$ and the following system matrices.

$$
A = \begin{bmatrix} -1.1 & 2 & -5 \\ 2 & -2 & 10 \\ 1 & -2 & 0 \end{bmatrix}, B_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, D_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}^T, \mu_1 = 0.15
$$

$$
E = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \Phi = -sgn(x_2) \times x_2^2
$$

$$
F = \begin{bmatrix} 0.1 \\ 0 \\ 0 \end{bmatrix}, H = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, u = 0.1\sin(2\pi t), C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}
$$

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

We will set $max(|(D_i\dot{x})(D_i\dot{x})^{-1}|) = \rho_{i-max} = 5$.

Solving the LMI, we find $\eta = 34.24$, $\epsilon_1 = 93.31$ and

$$
P_{11} = \begin{bmatrix} 49.23 & 0 & 42.97 \\ 0 & 34.24 & 0 \\ 42.97 & 0 & 165.33 \end{bmatrix}
$$
\[ P_{12}' = [-26.34 \quad 0 \quad -54.66]^T \]

Results of simulating the adaptive observer for the dissipative system are indicated in the following figures.

![Figure 2.18 State 1 estimation using the proposed observer](image-url)
Figure 2.19 State 2 estimation using the proposed observer

Figure 2.20 State 3 estimation using the proposed observer
Figure 2.21. State estimation errors

Figure 2.22. Unknown parameter estimation
Both the states and unknown parameter estimates are convergent to their true values and the cost function’s decay with time is also evident. Using the results of Theorem 2.2 has enabled us to solve LMIs in advance and eliminate the need for solving the Riccati differential equation online.

**Figure 2.23. Cost function**
Chapter 3.

Nonlinear Observer for Sensor Fault Estimation

Sensors are amongst the most vulnerable components in a complex dynamic system. The sensitivity of sensors to environmental factors and their deterioration with aging are two major causes of sensor faults. Some types of sensor faults may not cause catastrophic failures at their early stages; however, if not diagnosed and isolated, they can lead to disastrous destructions resulting from untrue sensor signals leading to untrue control signals. Therefore, periodic sensor calibration and off-line inspection is performed to check the healthy operation of sensors in many plants. Online fault detection and estimation, on the other hand, has gained increasing interest due to its ability to detect the onset of faults without requiring stopping system operation.

In this chapter, we will develop a nonlinear adaptive observer for nonlinear systems in order to estimate bias and gain sensor faults. The observer should be robust with respect to the noise and disturbances present in the system. A similar procedure can be followed as in chapter 2, to start with Lipschitz nonlinearities and then extend to sector-bounded nonlinearities. However, we skip treating Lipschitz nonlinearities separately and go directly to sector-bounded nonlinearities considering that Lipschitz nonlinearities are a sub-set of this general class. Simulation results at the end of this chapter illustrate the application of the fault estimation observer to a sample system.

3.1. System model

Consider the following sector-bounded nonlinear system
\[
\dot{x} = Ax + E\Phi(u, x) + Fw + Hu
\]
\[
y = Cx + C_B\theta_B + \theta_G C_{G1} C_{G2} \hat{x} + v
\]

where \(x \in \mathbb{R}^n\) is the state vector, \(A \in \mathbb{R}^{n \times n}\) is the system matrix, \(E \in \mathbb{R}^{n \times q}\), \(q \leq n\), \(\Phi(u, x): \mathbb{R}^p \times \mathbb{R}^n \rightarrow \mathbb{R}^q\) is the nonlinearity, \(u\) is the vector of known inputs, \(w\) is a vector of unknown input disturbances, \(y \in \mathbb{R}^m\) is the output vector and \(v\) is measurement noise vector. \(F, H, C\) are matrices of appropriate dimension. The pair \((A, C)\) is assumed to be observable. The nonlinearity \(\Phi(u, x)\) is assumed to satisfy a generalized sector bounded condition

\[
\Phi^T \Phi + (\gamma_1 - \gamma_2)\Phi^T G \bar{x} - \gamma_1 \gamma_2 \bar{x}^T G^T G \bar{x} \leq 0
\]

where \(\bar{x} = x - \hat{x}\) and \(\bar{\Phi} = \Phi(u, \bar{x}, \hat{x}) = \Phi(u, x) - \Phi(u, \hat{x})\). The type of nonlinearity considered is highly general and encompasses a variety of different nonlinearities including Lipschitz, one-sided Lipschitz, bounded Jacobian, monotonically increasing and dissipative nonlinearities. The definition implies that the nonlinearity lies within two bounds in a more general sense than the Lipschitz nonlinearity.

The system (3.1) is assumed to experience sensor bias fault and sensor gain fault. A bias fault in sensor ‘i’ can be written as \(y_i = C_i x + \theta_B\). More generally, a single bias fault can be written as the product of unknown magnitude \(\theta_B \in \mathbb{R}\), acting in a known output direction \(C_B \in \mathbb{R}^m\) giving \(y = Cx + C_B \theta_B\). A gain fault in sensor ‘i’ can be written as \(y_i = C_i x + \theta_G C_j x\), where \(i \neq j\) corresponds to a fault from cross-talk. More generally, a single gain fault will be written as \(y = Cx + C_{G1} \theta_G C_{G2} x\), where \(\theta_G \in \mathbb{R}\) is the unknown magnitude of the gain, \(C_{G2} \in \mathbb{R}^{1 \times n}\) is the direction of the fault in the state-space and \(C_{G1} \in \mathbb{R}^{m \times 1}\) is the direction of the output affected by the fault. An upper bound on the gain fault magnitude is assumed to be known, i.e.

\[
|\theta_G| < \theta_{G max}
\]
For example in a system with four states \((n = 4)\) and two sensors \((m = 2)\) with a bias fault in the first sensor and gain fault in the second sensor, one can write

\[
\begin{bmatrix}
  y_1 \\
  y_2
\end{bmatrix} = \begin{bmatrix}
  x_1 + \theta_B \\
  x_3(1 + \theta_G)
\end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \theta_B + \theta_G \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}
\]

Therefore, with \(C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\), a bias fault in the first sensor can be modeled with \(\mathcal{C}_B = \begin{bmatrix} 1 & 0 \end{bmatrix}\), while a gain fault in the second sensor can be modeled with \(\mathcal{C}_G_1 = \begin{bmatrix} 0 & 1 \end{bmatrix}^T, \mathcal{C}_G_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}\).

Note 1. If multiple bias faults are to be investigated, \(\mathcal{C}_B \theta_B\) should be replaced by \(\sum \mathcal{C}_B_i \theta_{B_i}\) and the observer design procedure follows similar steps. Similarly, if multiple gain faults are to be investigated, \(\sum \theta_{G_i} \mathcal{C}_{G1-i} \mathcal{C}_{G2-i} x\) should be replaced and the observer design procedure follows the same steps.

### 3.2. Observer model

For the system (3.1), the observer model is constructed as

\[
\begin{align*}
\dot{x} &= Ax + E\Phi(u, \hat{x}) + Hu + L(y - C\hat{x} - \mathcal{C}_B \hat{\theta}_B - \mathcal{C}_G \hat{\theta}_G C_G \hat{x}) \\
\dot{\hat{\theta}}_B &= L_{\theta B}(y - C\hat{x} - \mathcal{C}_B \hat{\theta}_B - \mathcal{C}_G \hat{\theta}_G C_G \hat{x}) \\
\dot{\hat{\theta}}_G &= L_{\theta G}(y - C\hat{x} - \mathcal{C}_B \hat{\theta}_B - \mathcal{C}_G \hat{\theta}_G C_G \hat{x})
\end{align*}
\]  

(3.4)

where \(\hat{x}\) is vector of estimated states and \(\hat{\theta}_B\) is the magnitude of estimated bias fault and \(\hat{\theta}_G\) is the magnitude of estimated gain fault. \(L, L_{\theta B}, \text{ and } L_{\theta G}\) are the observer gains. Dynamics of the estimation error for this system is written as

\[
\begin{align*}
\dot{\hat{x}} &= (A - LC)\hat{x} + E\hat{\Phi} - LC_B \hat{\theta}_B - L_{\theta B} C_G \hat{\theta}_B \hat{x} - L_{\theta G} C_G C_G \hat{x} - Lv + Fw \\
\dot{\hat{\theta}} &= -L_{\theta} C \hat{x} - L_{\theta} C_B \hat{\theta}_B - L_{\theta} \theta G C_G \hat{x} - L_{\theta} \theta G C_G C_G \hat{x} - L_{\theta} v
\end{align*}
\]  

(3.5)
Where

\[ \theta = [\theta_B \ \theta_G]^T, \quad L_\theta = [L_{\theta B} \ L_{\theta G}]^T \]

\[ \tilde{x} = x - \hat{x}, \quad \tilde{\theta}_B = \theta_B - \hat{\theta}_B, \quad \tilde{\theta}_G = \theta_G - \hat{\theta}_G \]

\[ \Phi = \Phi(u, \tilde{x}, \hat{x}) = \Phi(u, x) - \Phi(u, \hat{x}) \]

To obtaining the above relation, note that

\[ y - C\tilde{x} - C_1\hat{\theta}_1 - \hat{\theta}_G C G_1 C G_2 \hat{x} \]
\[ = Cx - C\hat{x} + C_B \theta_B - C_B \hat{\theta}_B + \theta_G C G_1 C G_2 x - \hat{\theta}_G C G_1 C G_2 \hat{x} + \nu \]  
(3.6)

\[ = C\tilde{x} + C_B \tilde{\theta}_B + \theta_G C G_1 C G_2 \tilde{x} + \hat{\theta}_G C G_1 C G_2 \hat{x} + \nu \]

3.3. **Optimal observer and fault estimator**

In this section, an optimal observer and estimator is developed for estimating the states and existing faults in the system (3.1). The observer is similar to the adaptive observer of chapter 2 in that it is based on defining a cost function of errors over disturbances and uses the method of Lagrange multipliers for solving the problem.

3.3.1. **Observer design**

Consider the following cost function

\[ J = \frac{\int_0^T \tilde{x}^T W \tilde{x} dt + \int_0^T \tilde{\theta}^T W_\theta \tilde{\theta} dt}{\tilde{x}_0^T P_0 \tilde{x}_0 + \tilde{\theta}_0^T P_{\theta_0} \tilde{\theta}_0 + \int_0^T v^T Q^{-1} v dt + \int_0^T w^T R^{-1} w dt} \]  
(3.7)

where \( W, W_\theta, P_0, P_{\theta_0}, Q \) and \( R \) are appropriate matrices selected for each problem and \( \tilde{\theta} = [\tilde{\theta}_B \ \tilde{\theta}_G]^T \). This cost function accounts for all estimation errors in its numerator. As
in previous chapter, we would formulate the problem as a constrained dynamic optimization problem. The objective is to minimize this cost function subject to the constraint of the state equations.

The following theorem investigates the existence and selection of gain matrices to stabilize the observer.

**Theorem 3.1**

*For the system (3.1) the observer (3.4) can guarantee a performance

\[
J < \frac{1}{\alpha}
\]

after a sufficiently large time \(T\), if \(\exists \eta > 0, \xi > 0\) such that

\[
\bar{P} = \bar{P}^T = \begin{bmatrix} P_{11} & P_{12} & P_{13} \\ * & P_{22} & P_{23} \\ * & * & P_{33} \end{bmatrix} > 0
\]

when driven by the differential equation

\[
\frac{d}{dt} \bar{P} = -\begin{bmatrix} W & A^TP_{21} - C^TQ^{-1}C_B & A^TP_{31} - C^TQ^{-1}C_GC_G\hat{x} \\ * & W'_{\theta_B} - C_B^TQ^{-1}C_B & -C_B^TQ^{-1}C_GC_G\hat{x} \\ * & * & W'_{\theta_g} - (C_GC_G\hat{x})^TQ^{-1}C_GC_G\hat{x} \end{bmatrix} \bar{P}
\]

\[
-\bar{P} \begin{bmatrix} \frac{1}{\eta} EE^T + FRF^T & 0 & 0 \\ * & 0 & 0 \\ * & * & 0 \end{bmatrix} \bar{P}
\]

where
\[
\tilde{W} = \alpha W + \eta \gamma_1 \gamma_2 G^T G + \xi \theta_{\text{max}}^2 C_2 C_2^T + P_{11} A' + A'^T P_{11} + \frac{\eta}{4} (\gamma_1 - \gamma_2)^2 G^T G - C^T Q'^{-1} C
\]

\[
A' = A - \frac{1}{2} (\gamma_1 - \gamma_2) E G
\]

and observer gains are chosen as

\[
\bar{L} = \bar{P}^{-1} \tilde{C}^T Q'^{-1}
\]

\[
\bar{L} = \begin{bmatrix}
L^T & L_{\theta_B}^T & L_{\theta_G}^T
\end{bmatrix}^T
\]

\[
\tilde{C} = \begin{bmatrix}
C & C_B & C_{G1} C_{G2} \bar{x}
\end{bmatrix}
\]

\[
Q' = \left( \frac{1}{\xi} C_{G1} C_{G1}^T + Q \right)
\]

and $\alpha$ is a user-specified performance bound.

Proof: The sub-optimal cost function (3.7) can be rewritten as

\[
J_1 = -\bar{x}_0^T P_0 \bar{x}_0 - \bar{\theta}_0^T P_0 \bar{\theta}_0 + \int_0^T \left[ \alpha \bar{x}^T W \bar{x} + \alpha \bar{\theta}^T W_0 \bar{\theta} - v^T Q^{-1} v - w^T R^{-1} w \right] dt
\]

Define

\[
V_\phi = \bar{\Phi}^T \bar{\Phi} + (\gamma_1 - \gamma_2) \tilde{\Phi}^T G \bar{x} - \gamma_1 \gamma_2 \bar{x}^T G^T G \bar{x} \leq 0
\]
\[ V_\theta = \theta_G^2 \tilde{x}^T C_{G2}^T C_{G2} \tilde{x} - \theta_{\theta_{\max}}^2 \tilde{x}^T C_{G2}^T C_{G2} \tilde{x} \leq 0 \]  

(3.14)

From S-Procedure Lemma, \( J_1 < 0 \) if and only if \( \exists \eta, \xi > 0 \) such that

\[ J_2' = J_1 - \int_0^T (\eta V_\phi) dt - \int_0^T (\xi V_\theta) dt < 0 \]  

(3.15)

Define the Hamiltonian or the augmented cost function as

\[
\bar{J} = J_2 + 2 \int_0^T \lambda^T [(A - LC) \tilde{x} + E \tilde{\Phi} - L C_B \tilde{\theta}_B - L \theta_G C_G C_{G2} \tilde{x} - L \tilde{\theta}_G C_G C_{G2} \tilde{x} - L \nu + F w - \dot{x}] dt \\
+ 2 \int_0^T \lambda_{\theta}^T \left[ -L_{\theta} C \tilde{x} - L_{\theta} C_B \tilde{\theta}_B - L_{\theta} \theta_G C_G C_{G2} \tilde{x} - L_{\theta} \tilde{\theta}_G C_G C_{G2} \tilde{x} - L_{\theta} \nu - \dot{\theta} \right] dt
\]  

(3.16)

After some manipulation,

\[ \bar{J} = -\tilde{x}_0^T P_0 \tilde{x}_0 - \tilde{\theta}_0^T P_0 \tilde{\theta}_0 + \int_0^T \left( \mathcal{H} - 2 \lambda^T \dot{x} - 2 \lambda_{\theta}^T \dot{\theta} \right) dt \]  

(3.17)

Where

\[
\mathcal{H} = \tilde{x}^T W' \tilde{x} + \tilde{\theta}^T W_{\theta}' \tilde{\theta} - \nu^T Q^{-1} \nu - w^T R^{-1} w \\
+ 2 \lambda^T [(A - LC) \tilde{x} + E \tilde{\Phi} - L C_B \tilde{\theta}_B - L \theta_G C_G C_{G2} \tilde{x} - L \tilde{\theta}_G C_G C_{G2} \tilde{x} - L \nu + F w] \\
+ 2 \lambda_{\theta}^T \left[ -L_{\theta} C \tilde{x} - L_{\theta} C_B \tilde{\theta}_B - L_{\theta} \theta_G C_G C_{G2} \tilde{x} - L_{\theta} \tilde{\theta}_G C_G C_{G2} \tilde{x} - L_{\theta} \nu \right] \\
- \eta \tilde{\Phi}^T \tilde{\Phi} - \eta (\gamma_1 - \gamma_2) \tilde{\Phi}^T G \tilde{x} - \xi \theta_G^2 \tilde{x}^T C_{G2}^T C_{G2} \tilde{x}
\]

and
\[ W' = aW + \eta y_1 y_2 G^T G + \xi \theta_{z_{\max}}^2 C_{G2}^T C_{G2} \]

\[ W'_\theta = \alpha W_\theta \]

Integrating \( \lambda^T \dot{x} \) and \( \lambda_\theta^T \dot{\theta} \) by parts we have

\[
\tilde{J} = -\bar{x}_0^T P_0 \bar{x}_0 - \bar{\theta}_0^T P_\theta \bar{\theta}_0 - 2[\lambda^T (T) \bar{x}(T) - \lambda^T (0) \bar{x}_0] \\
- 2[\lambda_\theta^T (T) \bar{\theta}(T) - \lambda_\theta^T (0) \bar{\theta}_0] + \int_0^T (\mathcal{H} + 2\lambda^T \ddot{x} + 2\dot{\lambda}_\theta^T \ddot{\theta}) dt
\]

We intend to find the stationary points of \( \tilde{J} \) with respect to the variables. Setting \( \left( \frac{\partial \tilde{J}}{\partial \bar{x}_0} \right)^T = 0 \) and \( \left( \frac{\partial \tilde{J}}{\partial \bar{\theta}_0} \right)^T = 0 \), we arrive at \( \lambda_0 = P_0 \bar{x}_0 \) and \( \lambda_{\theta_0} = P_{\theta_0} \bar{\theta}_0 \) or

\[
\tilde{\lambda}_0 := \begin{bmatrix} \lambda_0 \\ \lambda_{\theta_0} \end{bmatrix} = \begin{bmatrix} P_0 & 0 \\ 0 & P_{\theta_0} \end{bmatrix} \begin{bmatrix} \bar{x}_0 \\ \bar{\theta}_0 \end{bmatrix}
\]

(3.18)

Inspired by these relations, we assume

\[
\bar{x} := \begin{bmatrix} \lambda \\ \lambda_{\theta} \end{bmatrix} = \bar{P} \begin{bmatrix} \bar{x} \\ \bar{\theta} \end{bmatrix}
\]

(3.19)

Setting \( \left( \frac{\partial \tilde{J}}{\partial w} \right)^T = 0 \) and \( \left( \frac{\partial \tilde{J}}{\partial v} \right)^T = 0 \) yields

\[
F^T \lambda = R^{-1} w
\]

(3.20)

and

\[
Q^{-1} v + L^T \lambda + L_{\theta}^T \lambda_{\theta} = 0
\]

(3.21)

Defining the gain matrix \( \bar{L} \) as stated in the theorem statement, (3.22) can be rewritten as
\[
Q^{-1}v = -L^T \lambda \tag{3.23}
\]

Using (3.20) and rearranging the equation,

\[
v = -QL^T \bar{P} \begin{bmatrix} \bar{x} \\ \tilde{\theta} \end{bmatrix} \tag{3.24}
\]

Setting derivative with respect to the Sector bounded nonlinearity \((\partial \bar{f}/\partial \bar{\Phi})^T = 0,\)

\[
E^T \lambda = \eta \bar{\Phi} + \frac{\eta}{2}(\gamma_1 - \gamma_2)G\bar{x} \tag{3.25}
\]

Setting \((\partial \bar{f}/\partial \theta_C)^T = 0,\)

\[
C_{G1}^T (L^T \lambda + L_\theta^T \lambda_\theta) + \xi \theta_C C_{G2} \bar{x} = 0 \tag{3.26}
\]

Using (3.20) and simplifying,

\[
\theta_C C_{G2} \bar{x} = -\frac{1}{\xi} C_{G1}^T L^T \bar{P} \begin{bmatrix} \bar{x} \\ \tilde{\theta} \end{bmatrix} \tag{3.27}
\]

Setting \((\partial \bar{f}/\partial L)^T = 0\) yields

\[
C\ddot{x} + C_B\ddot{\theta}_B + \theta_C C_{G1} C_{G2} \ddot{x} + \theta_C C_{G1} C_{G2} \dot{\bar{x}} + v = 0 \tag{3.28}
\]

Reformulating into a matrix form and substituting for \(v\) from (3.24) we have

\[
\begin{bmatrix} C & C_B & C_{G1} C_{G2} \hat{x} \end{bmatrix} \begin{bmatrix} \ddot{x} \\ \hat{\theta}_B \\ \hat{\theta}_G \end{bmatrix} + \theta_C C_{G1} C_{G2} \ddot{x} = Q L^T \bar{P} \begin{bmatrix} \ddot{x} \\ \hat{\theta}_B \\ \hat{\theta}_2 \end{bmatrix} \tag{3.29}
\]

Using the definition for \(\bar{C}\) from the theorem statement and by using (3.27),

70
\[
\tilde{C} [\tilde{x}] - \frac{1}{\xi} C_{G1} C_{G1}^T \tilde{P} [\tilde{x}] = Q \tilde{P} [\tilde{x}]
\]

(3.30)

The preceding equation needs to hold for every \([\tilde{x}^T \quad \tilde{\theta}^T]^T\). Hence after simplification and by using the definition of \(Q'\) from the theorem statement we find

\[
\bar{L} = \tilde{P}^{-1} C^T Q'^{-1}
\]

(3.31)

Setting \((\partial \bar{L} / \partial \bar{x})^T = 0\) yields

\[
W' \tilde{x} + (A - LC)^T \lambda - \theta_2 (LC_{G1} C_{G2})^T \lambda - (L_{\theta} C)^T \lambda_\theta - \theta_G (L_{\theta} C_{G1} C_{G2})^T \lambda_\theta - \frac{\xi}{2} \theta_G^2 C_{G2}^T C_{G2} \tilde{x} - \frac{\eta}{2} (y_1 - y_2) G^T \Phi + \dot{\lambda} = 0
\]

(3.32)

Using (3.25),

\[
W' \tilde{x} + (A - LC)^T \lambda - \theta_2 (LC_{G1} C_{G2})^T \lambda - (L_{\theta} C)^T \lambda_\theta - \theta_2 (L_{\theta} C_{G1} C_{G2})^T \lambda_\theta - \frac{\xi}{2} \theta_G^2 C_{G2}^T C_{G2} \tilde{x} - \frac{1}{2} (y_1 - y_2) G^T E^T \lambda + \frac{\eta}{4} (y_1 - y_2)^2 G^T G \tilde{x} + \dot{\lambda} = 0
\]

(3.33)

Using (3.27),

\[
W' \tilde{x} + (A - LC)^T \lambda - (L_{\theta} C)^T \lambda_\theta - \frac{1}{2} (y_1 - y_2) G^T E^T \lambda + \frac{\eta}{4} (y_1 - y_2)^2 G^T G \tilde{x} + \dot{\lambda} = 0
\]

(3.34)

or

\[
\left( W' + \frac{\eta}{4} (y_1 - y_2)^2 G^T G \right) [I \ 0] \tilde{\lambda} - (L_{\theta} C)^T [0 \ I] \tilde{\lambda} + [I \ 0] \dot{\lambda} = 0
\]

(3.35)

or

\[
\left( W' + \frac{\eta}{4} (y_1 - y_2)^2 G^T G \right) [I \ 0] \tilde{\lambda} - (L_{\theta} C)^T [0 \ I] \tilde{\lambda} + [I \ 0] \dot{\lambda} = 0
\]

(3.35)
Setting \((\partial \tilde{J}/\partial \tilde{\theta}_1)^T = 0\),

\[
W'_{\theta_B} \tilde{\theta}_B - C_B^T L^T \lambda - C_B^T L_0^T \lambda_\theta + \lambda_{\theta_B} = 0
\]  \(\text{(3.36)}\)

or

\[
W'_{\theta_B}[0 \ 0 \ I \ 0] \bar{P}^{-1} \bar{\lambda} - C_B^T L^T [I \ 0] \bar{\lambda} - C_B^T L_0^T [0 \ I] \bar{\lambda} + [0 \ I \ 0] \bar{\lambda} = 0
\]  \(\text{(3.37)}\)

Setting \((\partial \tilde{J}/\partial \tilde{\theta}_G)^T = 0\),

\[
W'_{\theta_G} \tilde{\theta}_G - \hat{x}^T C_{g_2}^T C_{g_1}^T L^T \lambda - \hat{x}^T C_{g_2}^T C_{g_1}^T L_0^T \lambda_\theta + \lambda_{\theta_G} = 0
\]  \(\text{(3.38)}\)

or

\[
W'_{\theta_G}[0 \ 0 \ I \ 0] \bar{P}^{-1} \bar{\lambda} - \hat{x}^T C_{g_2}^T C_{g_1}^T L^T [I \ 0] \bar{\lambda} - \hat{x}^T C_{g_2}^T C_{g_1}^T L_0^T [0 \ I] \bar{\lambda}
\]

\[
+ [0 \ 0 \ I] \bar{\lambda} = 0
\]  \(\text{(3.39)}\)

Combining (3.35), (3.37) and (3.39) into a compact matrix form,

\[
\begin{bmatrix}
W'' & 0 & 0 \\
0 & W'_{\theta_B} & 0 \\
0 & 0 & W'_{\theta_G}
\end{bmatrix}
\bar{P}^{-1} \bar{\lambda} +
\begin{bmatrix}
(A - LC - \frac{1}{2}(\gamma_1 - \gamma_2)EG)^T \\
-(LC_B)^T \\
-(LC_{g_1}C_{g_2}\hat{x})^T
\end{bmatrix}
\bar{\lambda} + \hat{\lambda}
\]

\[
= 0
\]  \(\text{(3.40)}\)

where

\[
W'' = W' + \frac{\eta}{4}(\gamma_1 - \gamma_2)^2 G^T G
\]

Substituting for \(\hat{\lambda}\),

72
\[
\begin{bmatrix}
W'' & 0 & 0 \\
0 & W_{\theta_B}' & 0 \\
0 & 0 & W_{\theta_g}'
\end{bmatrix}
\tilde{\lambda} + \begin{bmatrix}
(A - LC - \frac{1}{2}(\gamma_1 - \gamma_2)EG)^T & -(L_0C)^T \\
-(LC_B)^T & -(L_0C_B)^T \\
-(LC_{G1}C_{G2}\hat{x})^T & -(L_0C_{G1}C_{G2}\hat{x})^T
\end{bmatrix}
\tilde{\lambda} \\
+ \tilde{p}\tilde{\lambda} + \bar{p} \begin{bmatrix}
(A - LC)\bar{x} + E\bar{\Phi} - LC_B\bar{\theta}_B - L\theta_C C_{G1}C_{G2}\bar{x} - L\bar{\theta}_G C_{G1}C_{G2}\hat{x} - L\nu + Fw \\
-L\bar{\theta}_C \bar{x} - L\theta_C B\bar{\theta}_B - L\theta_G C_{G1}C_{G2}\bar{x} - L\bar{\theta}_G C_{G1}C_{G2}\hat{x} - L\nu
\end{bmatrix} = 0
\]

(note that \(\tilde{\lambda} = [L_{\theta_1} \quad L_{\theta_2}]^T\), hence all block matrices in the above equation have the same dimension). Simplifying using (3.21), (3.25) and (3.28),

\[
\begin{bmatrix}
W'' & 0 & 0 \\
0 & W_{\theta_B}' & 0 \\
0 & 0 & W_{\theta_g}'
\end{bmatrix}
\tilde{\lambda} + \begin{bmatrix}
(A - LC - \frac{1}{2}(\gamma_1 - \gamma_2)EG)^T & -(L_0C)^T \\
-(LC_B)^T & -(L_0C_B)^T \\
-(LC_{G1}C_{G2}\hat{x})^T & -(L_0C_{G1}C_{G2}\hat{x})^T
\end{bmatrix}
\tilde{\lambda} \\
+ \tilde{p}\tilde{\lambda} + \bar{p} \begin{bmatrix}
A\bar{x} + \frac{1}{\eta}EE^T\lambda - \frac{1}{2}(\gamma_1 - \gamma_2)EG\bar{x} + FRF^T\lambda \\
0 \\
0
\end{bmatrix} = 0
\]

(3.42)

After some basic matrix calculations and since the above equation needs to hold for all \(\tilde{\lambda}\),

\[
\begin{bmatrix}
W'' & 0 & 0 \\
0 & W_{\theta_B}' & 0 \\
0 & 0 & W_{\theta_g}'
\end{bmatrix}
- \begin{bmatrix}
(LC)^T & (L_0C)^T \\
(LC_B)^T & (L_0C_B)^T \\
(LC_{G1}C_{G2}\hat{x})^T & (L_0C_{G1}C_{G2}\hat{x})^T
\end{bmatrix} \bar{\lambda} + \tilde{p} + \bar{p} \begin{bmatrix}
\frac{1}{\eta}EE^T + FRF^T & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}
\tilde{p}
\]

\[
+ \begin{bmatrix}
(A - \frac{1}{2}(\gamma_1 - \gamma_2)EG)^T & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix} \tilde{p} + \bar{p} \begin{bmatrix}
(A - \frac{1}{2}(\gamma_1 - \gamma_2)EG & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix} = 0
\]

(3.43)

Replacing from (3.31),
\[
\begin{bmatrix}
W'' & 0 & 0 \\
0 & W_{\theta B}' & 0 \\
0 & 0 & W_{\theta g}'
\end{bmatrix} + \begin{bmatrix}
C^T & 0 \\
C_B^T & 0 \\
(C_G1C_G2\hat{x})^T & (C_G1C_G2\hat{x})^T
\end{bmatrix} Q'^{-1} \begin{bmatrix}
C^T & 0 \\
C_B^T & 0 \\
(C_G1C_G2\hat{x})^T & (C_G1C_G2\hat{x})^T
\end{bmatrix}^T + \bar{P} \begin{bmatrix}
1 EE^T + FRF^T & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}.
\]

Defining \( A' = A - \frac{1}{2} (\gamma_1 - \gamma_2)EG \) as in the theorem statement and noting that

\[
\begin{bmatrix}
A' & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix} + \bar{P} \begin{bmatrix}
A' & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix} = \begin{bmatrix}
P_{11}A' + A'^T P_{11} & A'^T P_{12} & A'^T P_{13} \\
P_{12}A' & 0 & 0 \\
P_{13}A' & 0 & 0
\end{bmatrix}
\]

The equation (3.44) can be written as

\[
\begin{bmatrix}
\bar{W} & A'^T P_{12} - C^T Q'^{-1} C_B & A'^T P_{13} - C^T Q'^{-1} C_G1C_G2\hat{x} \\
p_{12}' A' - C_B^T Q'^{-1} C_B & W_{\theta 1}' - C_B^T Q'^{-1} C_B & -C_B^T Q'^{-1} C_G1C_G2\hat{x} \\
p_{13}' A' - (C_G1C_G2\hat{x})^T Q'^{-1} C_B & -(C_G1C_G2\hat{x})^T Q'^{-1} C_B & W_{\theta g}' - C_G2\hat{x}(C_G1)^T Q'^{-1} C_G1C_G2\hat{x}
\end{bmatrix}

\begin{bmatrix}
1 EE^T + FRF^T & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix} + \bar{P} \begin{bmatrix}
1 EE^T + FRF^T & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix} \bar{P} = 0
\]

(3.46)

The above theorem would require Riccati differential equation be solved online in order to compute \( \bar{P} \) and \( \bar{L} \). For many physical systems with sufficiently slow dynamics, a numerical solution to the Riccati differential equation can be calculated and a practical observer is feasible with the advent of microprocessors with high computational power.
Nonetheless, it would be beneficial to avoid solving the Riccati differential equation online.

Choosing a time-varying matrix $\tilde{P}$ (or equivalently gain $L$) to satisfy the preceding inequality at all times is not a trivial task. Instead, we can try to push equation (3.46) to have a steady-state solution for which $\eta > 0, \xi > 0$ and $\tilde{P} > 0$ exist. This is not feasible at first glance, since the adaptive estimation law requires the matrix $\tilde{P}$ to be time-varying. However, the following theorem attempts to modify the structure of $\tilde{P}$ by defining a bound on the variable terms. To extract existence conditions for stable solutions, we use a slightly modified cost function.

**Theorem 3.2**

For the system (3.1), if $\max(|E G_2 \hat{x} | / |E G_2 \hat{x}|) < \rho_{max}$, the observer (3.4) can guarantee a performance

$$J = \frac{\int_0^T \hat{x}^T W \hat{x} dt + \int_0^T \hat{x}^T C_2 W_\theta \hat{x} dt + \int_0^T \hat{x}^T C_2 \xi W_{\theta_0} \hat{x} \xi \hat{x} dt}{\hat{x}^T P_0 \hat{x} + \hat{x}^T P_{\theta_0} \hat{x} + \int_0^T \nu^T Q^{-1} \nu dt + \int_0^T \nu^T R^{-1} \nu dt} \leq \frac{1}{\alpha} \tag{3.47}$$

after a sufficiently large time $T$, if $\exists \tilde{P}', \eta > 0, \xi > 0$ such that

$$\tilde{P}' = \tilde{P}'^T = \begin{bmatrix} P_{11} & P_{12} & P_{13}' \\ * & P_{22} & P_{23}' \\ * & * & P_{33} \end{bmatrix} > 0$$

Satisfies the following LMI for $\forall \rho \leq \rho_{max}$

$$\begin{bmatrix} M_{11} & M_{12} & M_{13} & P_{11}E & P_{11}F \\ * & M_{22} & M_{23} & P_{21}E & P_{21}F \\ * & * & M_{33} & P_{31}E & P_{31}F \\ * & * & * & -\eta I & 0 \\ * & * & * & * & -R^{-1} \end{bmatrix} < 0 \tag{3.48}$$
with

\[ 
M_{11} = \alpha W + \xi \theta_{z_{\max}}^2 \frac{C_G^T C_G}{\xi G^T G + P_{11}A' + A'TP_{11} + \frac{\eta}{4}(y_1 - y_2)^2 G^T G - C^T Q^{-1}C} 
\]

\[ 
M_{12} = -C^T Q^{-1}C_B + A'TP_{12} 
\]

\[ 
M_{13} = -C^T Q^{-1}C_G + A'TP_{13}' + \rho P_{13}' 
\]

\[ 
M_{22} = -C_B^T Q^{-1}C_B + \alpha W_{\theta_B} 
\]

\[ 
M_{23} = -C_B^T Q^{-1}C_G + \rho P_{23}' 
\]

\[ 
M_{33} = -\rho^2 C_B^T Q^{-1}C_G + \rho^2 (\alpha W_{\theta_2}) 
\]

\[ 
A' = A - \frac{1}{2}(y_1 - y_2)EG 
\]

The gains should be choosen as \( \bar{L} = \bar{P}^{-1}C^T Q^{-1} \) where \( \bar{L} = [L^T \quad L_{\theta_B}^T \quad L_{\theta_G}^T]^T \), \( \bar{C} = 
\]

\[
[C \quad C_B \quad C_G C_G^\hat{x}] \quad \text{and} \quad Q' = \left( \frac{1}{\xi} C_G C_G^T + Q \right). \bar{P} = \begin{bmatrix} P_{11} & P_{12} & P_{13}' C_G \hat{x} \\ * & P_{22} & P_{23}' C_G \hat{x} \\ * & * & P_{33} \end{bmatrix}.
\]

Proof: Similar to Theorem 3.1, the cost function inequality (3.8) can be rewritten as

\[ 
J_1 = -\bar{x}_0^T P_0 \bar{\tilde{x}}_0 - \bar{\tilde{\theta}}_0^T P_0 \bar{\tilde{\theta}}_0 
\]

\[ 
+ \int_0^T \left[ \alpha \bar{x}^T W \bar{x} + \alpha \bar{\theta}_B^T W_{\theta_B} \bar{\theta}_B + \alpha \bar{\theta}_G^T C_G \hat{x} W_{\theta_G} C_G \hat{x} \bar{\theta}_G 
- \nu^T Q^{-1} \nu - w^T R^{-1} w \right] dt 
\]

(3.49)

with \( W_{\theta} \) defined as

76
\[ W_\theta = \text{diag}(W_{\theta_B} C_{G2} \hat{x} W_{\theta_C} C_{G2} \hat{x}) \]  

(3.50)

Define

\[ V_\theta = \theta_C^2 \bar{x}^T C_{G2}^T C_{G2} \bar{x} - \theta_C^2 \bar{x}^T C_{G2}^T C_{G2} \bar{x} \leq 0 \]  

(3.51)

\[ V_\phi = \bar{\phi}^T \bar{\phi} + (\gamma_1 - \gamma_2) \bar{\phi}^T G \bar{x} - \gamma_1 \gamma_2 \bar{x}^T G^T G \bar{x} \leq 0 \]  

(3.52)

From S-Procedure lemma, \( J_1 < 0 \) if and only if \( \exists \eta > 0, \xi > 0 \) and \( \Xi = \Xi^T > 0 \) such that

\[
J_2 := J_1 + \int_0^T \left( -\eta V_\phi - \xi V_\theta + [\bar{x}^T \bar{\theta}^T] \Xi [\bar{x}^T \bar{\theta}^T]^T \right) dt < 0
\]

(3.53)

It should be noted that (2.65) differs from (2.21) by \( V_\Xi = \int_0^T \left( [\bar{x}^T \bar{\theta}^T] \Xi [\bar{x}^T \bar{\theta}^T]^T \right) dt \). Following the same steps as in Theorem 3.1 and finding the stationary points of \( \bar{J} \) with respect to its variables, we arrive at a slightly different differential Riccati equation

\[
\dot{\bar{P}} + \begin{bmatrix}
\bar{W} & A^T P_{12} - C^T Q^{\prime -1} C_B \\
* & W_{\theta_B} - C_B^T Q^{\prime -1} C_B
\end{bmatrix}
\begin{bmatrix}
A^T P_{13} - C^T Q^{\prime -1} C_{G1} C_{G2} \hat{x} \\
-C_B^T Q^{\prime -1} C_{G1} C_{G2} \hat{x}
\end{bmatrix}
+ \begin{bmatrix}
\frac{1}{\eta} E E^T & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix} \bar{P} + \Xi = 0
\]

(3.54)

Suppose \( \exists P'_{13}, P'_{23} \) such that \( \forall \rho \leq \rho_{\text{max}} \)
\[
\begin{bmatrix}
\bar{W} & A^T P_{12} - C^T Q'^{-1} C_B + \rho P'_{12} & A^T P'_{13} - C^T Q'^{-1} C_{G1} + \rho P'_{13} \\
* & W_{\theta_B}' - C_B^T Q'^{-1} C_B & -C_B^T Q'^{-1} C_{G1} \\
* & * & W_{\theta_G}' - (C_{G1})^T Q'^{-1} C_{G1}
\end{bmatrix} < 0
\]

(3.55)

then at time \( t_0 \), for a choice of \( P_{13}(t_0) = P'_{13} C_{G2} \hat{x}(t_0) \) and \( P_{23}(t_0) = P'_{23} C_{G2} \hat{x}(t_0) \), and appropriate \( \Xi > 0 \), (3.54) can be satisfied by choosing \( \rho = (C_{G2} \hat{x})(C_{G2} \hat{x})^{-1} \) and

\[
\hat{P} = \begin{bmatrix}
0 & 0 & P'_{13} C_{3} \hat{x} \\
* & 0 & P'_{23} C_{3} \hat{x} \\
* & * & 0
\end{bmatrix}
\]

(3.56)

Hence \( P_{13}(t > t_0) = P'_{13} C_{G2} \hat{x}(t) \) and \( P_{23}(t > t_0) = P'_{23} C_{3} \hat{x}(t) \). Taking the Schurs’s complement, (3.55) can be written as the LMI (3.48). The inequalities (3.55) and (3.48) are independent of \( \hat{x}(t) \) which results in a more convenient design procedure using the LMI and relations (3.11) for the gains. Also, this is computationally more efficient as the differential Riccati equation needs not be solved online.

### 3.4. Numerical example

3.4.1. **4th order nonlinear system**

Consider the following 4th order nonlinear system in the framework (3.1)

\[
\dot{x} = Ax + E\Phi(u, x) + Fw + Hu \\
y = Cx + C_B \theta_B + \theta_G C_{G1} C_{G2} x + \nu
\]
\[ A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -25 & -0.2 & 15 & 0.2 \\ 0 & 0 & 0 & 1 \\ 15 & 0.2 & -15 & -0.2 \end{bmatrix} \]

\[ E = [0 \ 2 \ 0 \ 0]^T, \quad \Phi = -sgn(x_2)x_2^2, \]

\[ F = [0 \ 0 \ 0 \ 1]^T, \quad \omega \sim (0, 10^{-4}) \]

\[ H = [0 \ 0 \ 0 \ 1]^T, \quad u = \sin(2\pi t) \]

\[ C = [1 \ 0 \ 0 \ 0], \quad \nu \sim (0, [10^{-4} \ 0 \ 0]) \]

The model is representative of a two-mass spring systems with nonlinear dissipative damping. We want to observe the existence and severity of bias and gain sensor fault in the first sensor. We choose the bias fault magnitude \( \theta_B = 0.04 \). Next we define the following fault matrix \( C_B = [1 \ 0]^T \). Also, for the gain fault with magnitude \( \theta_G = 0.1 \), we have \( C_{G1} = [1 \ 0]^T, C_{G2} = [1 \ 0 \ 0 \ 0] \). For this dissipative system, \( \gamma_1 \to \infty, \gamma_2 = 0, G = [0 \ 1 \ 0 \ 0] \).

Selecting an appropriate value for \( \alpha \) is a compromise between the observer performance and solvability of the linear matrix inequalities. A reasonable value can often be found after some try and error. The LMI solution for \( \alpha = 1, \ \rho = 1 \) yeilds

\[ \xi = 2197.3, \eta = 20913, P_{23} = 5.11 \times 10^3, P_{22} = 2.194 \times 10^4, P_{33} = 2.2033 \times 10^4 \]

\[ P_{11} = 10^4 \begin{bmatrix} 5.528 & 0 & -3.279 & -0.033 \\ 0 & 0.219 & 0 & 0 \\ -3.279 & 0 & 3.339 & -0.009 \\ -0.033 & 0 & -0.009 & 0.223 \end{bmatrix} \]

\[ P_{12} = [274.0 \ 0 \ 381.8 \ 152.9]^T \]

\[ P_{13} = [592.9 \ 0 \ 389.2 \ 161.7]^T \]

The following figures indicate the estimation of the states and the bias fault as well as the cost function evolution with time. The results show satisfactory state and fault estimation for the 4th order system.
Figure 3.1 State 1 estimation

Figure 3.2 State 2 estimation
Figure 3.3 State 3 estimation

Figure 3.4 State 4 estimation
Figure 3.5 Bias fault estimation

Figure 3.6 Gain fault estimation
Figure 3.7 Cost function

As observed in Figure 3.5 and Figure 3.6 both faults are estimated by the observer, but in a relatively larger time span compared with state estimations. Additionally, Figure 3.7 indicates the evolution of the cost function with time. It is evident that after some initial growth, the cost function becomes smaller with time which signifies the expected optimization process.
Chapter 4.

Observer and FDI Implementation: Wind Turbine Case Study

Energy generation through renewable sources is gaining increasing popularity in the world, especially due to environmental concerns. This increase in popularity has led to larger investment in this area and increasing number of renewable plants installed each day. Wind power, one of the most important sources of renewable energy, is being exploited by installing wind turbines in many countries. It is expected that 20% of the total electrical energy generation in the United States would be from wind farms by 2030 [57].

Wind turbines are generally electro-mechanical systems capable of converting wind’s mechanical energy into electrical energy. Typically, they are complex machines that require incessant monitoring due to their potential failure risks and damage costs. Wind turbines require effective fault detection techniques due to several reasons. First, many turbines are installed offshore or in remote areas where persistent personal inspection is not feasible. Additionally, proper FDI can prevent catastrophic failures by detecting faults at their early stages. If an online technique is employed for detecting and estimating the extent of small faults or deviations from normal system operation, the risk of system failures would be significantly reduced. The fault data can then be used by control algorithms to operate safe in spite of present faults. Further, it highly reduces maintenance costs of the system. Consequently it has become an essential part of different wind turbine technologies around the globe.
A survey of different wind turbine FDI and FTC approaches available in the literature can be found in [64]. Different components of the wind turbine system including sensors, actuators and power transmission components may become faulty during turbine operation. Among various fault scenarios that may occur in a wind turbine, sensor faults are being investigated in this work. As will be explained later, the estimation of faults in the sensors measuring angular velocities of turbine shafts is essential to safe turbine operation.

In this chapter, we intend to implement the parameter and fault estimation methods proposed in the previous chapters on a wind turbine power transmission mechanism. Our observer is a good candidate for FDI in this problem since different nonlinearities, disturbances and uncertainties are associated with the wind turbine model. We start with a brief overview of wind turbine components, modelling and control algorithms and then discuss the steps towards FDI implementation.

4.1. Wind turbine components

Horizontal Axis Wind Turbines (HAWT) are generally divided into two major categories; fixed speed and variable speed. There are some disadvantages associated with using fixed speed turbines including energy inefficiency and lower control freedom. Such drawbacks are moving the current commercial trend in wind turbine industry towards variable speed wind turbines. Variable speed wind turbines are more flexible with regards to variable wind speed and can provide more efficient power.

Figure 4.1 indicates different components of a horizontal axis wind turbine. The major components include turbine blades, low speed shaft, gearbox, high speed shaft and the generator. Turbine blades are aerodynamically designed components for exploiting maximum possible wind energy. The blades rotate a low speed shaft which transfers the power to a high speed shaft through a gearbox. The high speed shaft is connected to the generator. Also, the type of the generator used is different depending on the complexity
of the turbine [58]. Nowadays, many turbines utilize Doubly Fed Induction Generators (DFIG). The stator of a DFIG is directly connected to the grid. The rotor is connected to the grid via a back-to-back converter (view Figure 4.2). The role of this converter is to provide appropriate rotor currents based on its control logic in a way to achieve the general turbine control objectives [60].

Figure 4.1 Major wind turbine components[59]
4.2. Wind turbine modelling

In this section, we present a mathematical model for a variable speed wind turbine. The model comprises aerodynamic and gear train sub-models. Modelling the generator and converter is beyond the scope of this work.

As we will see in the next sections, the system comprises a highly nonlinear model due to the nature of the problem.

4.2.1. Aerodynamic model

Total power contained in the wind ($P_w$) crossing the turbine blades at speed $v$ can be expressed as

$$P_w = \frac{1}{2} \rho A v^3$$  \hspace{1cm} (4.1)
where $\rho$ is air density and $A$ is the total blade swept area equal to $\pi R^2$ ($R$ is blade radius). Only a part of the wind energy is transferred from the wind to the blades and the rotor of the wind turbine. Therefore, the captured aerodynamic power is a portion of the total wind power [61]

$$P_A = C_P P_W$$

(4.2)

where $P_A$ is the captured aerodynamic power and $C_P$ is power coefficient of the wind turbine indicating what portion of the wind energy is being captured by the turbine. $C_P$ is a function of the blade’s geometry (including blade profile and pitch angle) and also the ratio $\frac{\omega_r}{v}$ where $\omega_r$ is rotor rotational speed and $v$ is the wind speed. Therefore the aerodynamic power can be expressed as

$$P_A = \frac{1}{2} C_P \rho A v^3$$

(4.3)

From this relation, the rotor torque can be obtained as

$$T_r = \frac{P_A}{\omega_r} = \frac{C_P \rho A v^3}{2 \omega_r}$$

(4.4)

To express the dependency of $C_P$ on other parameters, we first define the parameter tip speed ratio which is proportional to the ratio of turbine angular velocity over the wind speed

$$\lambda = \frac{\omega_r R}{v}$$

(4.5)

$R$ is turbine blade radius. $C_P$ is dependent on both the blade pitch angle and tip speed ratio. The nonlinear dependency of $C_P$ on $\lambda$ and $\theta$ (blade pitch angle) can be expressed through diagrams, lookup tables or some semi-analytical relations. The relations for $C_P$
are dependent on the turbine and the manufacturer. One of the available models for $C_P$ is [62]:

$$C_P(\lambda, \theta) = 0.22 \left( \frac{116}{\lambda_i} - 0.4\theta - 5 \right) e^{-\frac{12.5}{\lambda_i}}$$

$$\frac{1}{\lambda_i} = \frac{1}{\lambda + 0.08\theta} - \frac{0.035}{1 + \theta^3}$$  \hspace{1cm} (4.6)

Using this relation, the dependency of $C_P$ on $\lambda$ is illustrated in Figure 4.3 for five different values of $\theta$.

![Figure 4.3 $C_P - \lambda$ plot for different values of $\theta$](image)

4.2.2. **Gear train model**

The power captured by the turbine is usually transferred to the generator through a gear train. The gear train plays the role of connecting the low-speed rotor shaft to the high-speed generator shaft. The shafts and the gear train experience some torsion due to the high torques they endure. For the purpose of our modelling, we adopt a two-mass model for the gearbox. In Figure 4.4, $r$ subscripts denote rotor quantities and $g$ subscript
denote generator quantities. $N$ is the gear ratio, and $c$ and $k$ are combined masses and gearbox damping and stiffness coefficients.

The following 3rd order state-space model is used for the gear train

$$
\begin{bmatrix}
\dot{\omega}_g \\
\dot{\omega}_r \\
\dot{\Delta \theta}
\end{bmatrix} =
\begin{bmatrix}
-\frac{c}{N^2 J_g} & \frac{c}{N J_g} & -k \\
\frac{c}{N J_r} & -\frac{k}{J_r} & \frac{1}{N} \\
1 & -1 & 0
\end{bmatrix}
\begin{bmatrix}
\omega_g \\
\omega_r \\
\Delta \theta
\end{bmatrix} +
\begin{bmatrix}
1 & 0 \\
0 & 1 \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
T_g \\
T_r
\end{bmatrix}
\tag{4.7}
$$

where the following model parameters and variables are employed

<table>
<thead>
<tr>
<th>Table 4-1 Symbols used in the 3rd order wind turbine model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symbol</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>$c$</td>
</tr>
<tr>
<td>$k$</td>
</tr>
<tr>
<td>$N$</td>
</tr>
<tr>
<td>$J_g$</td>
</tr>
</tbody>
</table>
\( J_r \)  Low-speed shaft moment of inertia

\( T_g \)  Generator torque

\( T_r \)  Aerodynamic torque by wind on the low-speed shaft

Generator torque is the control input through which different turbine control strategies can be implemented. \( T_r \) is the aerodynamic torque applied to the blades. The following table summarizes sample nominal values for the turbine and wind parameters.

**Table 4-2 Turbine and wind parameters**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Parameter</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R )</td>
<td>Blade radius</td>
<td>70</td>
<td>( m )</td>
</tr>
<tr>
<td>( V_{nom} )</td>
<td>Nominal wind speed</td>
<td>12.5</td>
<td>( m/s )</td>
</tr>
<tr>
<td>( \lambda_{opt} )</td>
<td>Optimal tip speed ratio</td>
<td>8.42</td>
<td>-</td>
</tr>
<tr>
<td>( C_{p_{max}} )</td>
<td>Maximum power coefficient</td>
<td>0.4345</td>
<td>-</td>
</tr>
<tr>
<td>( \rho )</td>
<td>Air density</td>
<td>1.1225</td>
<td>( kg/m^3 )</td>
</tr>
</tbody>
</table>

**Table 4-3 Wind turbine mechanical parameters**

The following table summarizes nominal values for the turbine’s mechanical power transmission system.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Parameter</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>( J_g )</td>
<td>High-speed shaft moment of inertia</td>
<td>400</td>
<td>( kg.m^2 )</td>
</tr>
<tr>
<td>( J_r )</td>
<td>Low-speed shaft moment of inertia</td>
<td>( 5.5 \times 10^7 )</td>
<td>( kg.m^2 )</td>
</tr>
<tr>
<td>( c )</td>
<td>Damping coefficient</td>
<td>600</td>
<td>( kg.m^2/s )</td>
</tr>
<tr>
<td>( k )</td>
<td>Total shafts and gearbox stiffness</td>
<td>( 3 \times 10^9 )</td>
<td>( kg.m^2/s^2 )</td>
</tr>
</tbody>
</table>
4.3. Wind turbine operation

The operation of a variable speed wind turbine is highly dependent on the wind speed. At very low wind speeds, the turbines are not let free to operate. Generally, turbines are installed in regions where the average wind speed is relatively high and the turbines are active for a great part of their life. The wind speed at which the turbine begins operation is called cut-in wind speed. This speed is dependent on the turbine manufacturer and varies between turbines, but typically falls within 3 m/s and 5 m/s.

![Wind turbine operation range](image)

**Figure 4.5 Wind turbine operation range**

When wind speed reaches this level the turbine starts operation. For some interval of wind speed variation, i.e. from cut-in speed up to the rated speed, the turbine will be controlled in Maximum Power Point Tracking (MPPT) mode. This mode of control maximizes the extracted power from the wind and increases turbine efficiency. Once the
rated wind speed is achieved (~ 12m/s), the turbine reaches its maximum (nominal) power point. Higher speeds are deemed dangerous for the turbine and should be avoided. Therefore, at wind speeds above the nominal wind speed, a new control strategy called blade pitch control will take over the operation aiming at continuing exploiting constant power without damaging the turbine. This is achieved by employing blade pitch control, which controls the blade pitch angle. Finally, in case the wind speed is higher than a pre-defined threshold, the turbine operation will fully stop in order to prevent damage to the turbine. These four operation regions are indicated graphically in Figure 4.5. In this figure, the blue lines indicate exploited power by the wind turbine at each wind speed under the explained control strategies. As observed in this figure, within the MPPT region, the power increases with increase in wind speed. However, in wind speeds above the rated speed, the power remains constant by utilizing the pitch control mechanism to prevent damage [60].

### 4.4. Wind turbine control strategies

As mentioned before, in region 2, Maximum Power Point Tracking (MPPT) is employed. For MPPT, we should control the DFIG torque in a way that the speed of rotor varies proportional to wind speed and optimal tip speed ratio is retained. MPPT should provide the reference value for the active generator power. When the turbine is operating at the maximum power point, the following optimal parameters are defined

\[ \lambda_{opt} = \frac{\omega_r R}{v} \quad (4.8) \]

\[ C_p = C_{p_{max}} \quad (4.9) \]

Now, we use one of the available relations for the reference value for active generator torque which is
\[ T_g = \frac{1}{2} \rho \pi R^3 \frac{R^2 \omega_r^2 C_{p\text{max}}}{\lambda_{opt}^2} = K_{opt} \omega_r^2 \]

\[ K_{opt} = \frac{1}{2} \rho \pi \frac{R^5}{\lambda_{opt}^3} C_{p\text{max}} \]

This reference value is the output of the high-level control unit fed into the converter. Also, the reference value for reactive generator power is zero. Active and reactive power, independently, provide the reference value for rotor current control loop. Current control loop aims to achieve the final objective of maximum power extraction by means of the converter connected to the rotor of the generator [60].

Figure 4.6 Variable speed wind turbine general control strategy
On the other hand, in high wind speeds, pitch regulation is employed for keeping output power at its rated value so that the turbine is kept safe from damage.

The other system input, i.e. the aerodynamic torque is not directly measurable. There are two approaches to deal with this problem. Some works consider designing unknown input observers in which the knowledge of the direction of the input is used to decouple state estimations from the unknown input [45], [63]. The other technique that we have adopted here relies on using the existing relations for the input torque despite the uncertainties associated with its variables. Following this method, we would be able to approximately reconstruct the input and feed it into the observer. Then, we also add the disturbance term to our state equations to account for uncertainties associated with this input calculation in the direction of the aerodynamics torque input. The uncertainties show up in two different parts of the $T_r$ relation. First, wind speed fluctuations and its variation across the turbine blades is a source of uncertainty dissemination. We cannot measure the effective wind speed affecting the blades accurately. Moreover, the uncertainty associated with the available relations for computing $C_p$ could be significant. The following block diagram indicates the spread of uncertainties within the system model.

Also, the system is highly nonlinear due to the relation used for $C_p$ and therefore a nonlinear observer should be designed.
4.5. Simulation results

In this section, we have conducted simulations to implement the proposed state estimation and fault estimation technique on the wind turbine model. We only consider the mechanical power transmission model and aerodynamic torque model described in section 4.2. The wind turbine model and parameters explained in sections 4.2 and 4.3 are used for our simulations. It is assumed that the generator and converter have effectively provided the optimal generator torque. Therefore the following matrices are derived for our state-space model

\[
A = \begin{bmatrix}
-\frac{c}{N^2J_g} & \frac{c}{NJ_g} & -\frac{k}{NJ_g} \\
\frac{c}{NJ_r} & -\frac{c}{J_r} & \frac{k}{J_r} \\
\frac{1}{N} & -1 & 0
\end{bmatrix}
\]

\[
F = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T \quad H = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T \quad \omega \sim (0, 10^{-4})
\]

The nonlinearity is modelled as \( E = [0 \ 1 \ 0]^T, \Phi = \frac{1}{2J_r\omega_r} C_p\rho A v^3 \). As explained before, there is some uncertainty associated with this nonlinearity. First, available
relations and data for $C_p$ are approximate and do not provide exact values for all available turbines. Further, wind speed $\nu$ cannot be measured precisely. Wind speed fluctuations are significant across the blade swept area and an effective wind speed value is not readily measurable. Therefore we include disturbance term in the state equation.

It is intended to use the methods of Chapter 2 and Chapter 3 to estimate an uncertain parameter and existing sensor faults. In the first example, we intend to use the observer to estimate the value of the stiffness assuming our knowledge of it is uncertain. In the second example we aim to estimate bias fault in the second sensor and gain fault in the first one.

4.5.1. Example 1. State and parameter estimation for the wind turbine model

It is intended to estimate the value of the stiffness $k$ assuming that we know an approximate value for it. To obtain the matrices $B$ and $D$, we should investigate the effect of the uncertainty of $k$ on the system matrix $A$. If we show $k = k^* + \Delta k$, then the uncertain matrix $A$ becomes

$$
\begin{bmatrix}
-\frac{c}{N^2 J_g} & \frac{c}{N J_g} & -\frac{k^* + \Delta k}{J_g} \\
\frac{c}{N J_g} & \frac{k^* + \Delta k}{J_g} & -\frac{k^*}{J_g} \\
\frac{1}{N} & -1 & 0
\end{bmatrix}
= \begin{bmatrix}
-\frac{c}{N^2 J_g} & \frac{c}{N J_g} & -\frac{k^*}{J_g} \\
\frac{c}{N J_g} & \frac{k^*}{J_g} & -\frac{k^*}{J_g} \\
\frac{1}{N} & -1 & 0
\end{bmatrix}
+ \begin{bmatrix}
0 & 0 & -\Delta k \\
0 & 0 & \frac{\Delta k}{J_r} \\
0 & 0 & 0
\end{bmatrix}
$$

where $\mu = -5 \times 10^{-12} \Delta k$, $B = 2 \times 10^9 \times \left[ \frac{1}{N J_g} \ -\frac{1}{J_r} \ 0 \right]^T$, $D = 10^2 \times [0 \ 0 \ 1]$. We now use the model parameter values of previous sections to simulate the turbine and design the observer. We use the approximate value $k = 3 \times 10^9$ for the stiffness in our simulations while the exact value is $3.014 \times 10^9$. Therefore, $\Delta k = 0.014 \times 10^9$. We
now solve the LMIs of section 2.4 to obtain the observer gains. The considered nonlinearity satisfies the sector-bounded condition with

\[ \gamma_1 = \gamma_2 = 1, \quad G = [0 \ 10 \ 0] \]

Selecting an appropriate value for \( \alpha \) is a compromise between the observer performance and solvability of the linear matrix inequalities. A reasonable value can often be found after some try and error.

Solving the LMIs for \( \alpha = 0.1, \ \rho = 1 \) yeilds

\[ \xi = 2.814 \times 10^6, \eta = 2.743 \times 10^2, P_{22} = 1.84 \times 10^6 \]

\[
P_{11} = 10^3 \begin{bmatrix} 0.0005 & -0.010 & 0.353 \\ -0.010 & 19.655 & 21.561 \\ 0.353 & 21.561 & 325.945 \end{bmatrix}, \quad P_{12} = -10^5 \begin{bmatrix} 0.0002 \\ 0.1308 \\ 2.2036 \end{bmatrix}
\]

The simulation results follow below.

![Figure 4.8 State 1 estimation](image)
Figure 4.9 State 2 estimation

Figure 4.10 State 3 estimation
Figure 4.11 State estimation errors

Figure 4.12 Unknown parameter estimation
Both the states and the parameter estimates converge to the true values in spite of the uncertainties and disturbances.

4.5.2. **Sensor gain and bias faults estimation for the wind turbine model**

In this example, we define $C_B$, $C_{G1}$ and $C_{G2}$ to model bias fault in the second sensor and gain fault in the first sensor.

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \nu \sim \left( 0, \begin{bmatrix} 10^{-3} \\ 0 \\ 0 \\ 10^{-3} \end{bmatrix} \right)$$

Since bias fault occurs in the second sensor and gain fault occurs in the first sensor,

$$C_B = [0 \ 1]^T \quad \theta_B = 0.4 \quad C_{G1} = [1 \ 0]^T \quad C_{G2} = [1 \ 0 \ 0] \quad \theta_G = 0.5 \quad \theta_{G_{max}} = 4$$

The considered nonlinearity satisfies the sector-bounded condition with

$$\gamma_1 = \gamma_2 = 1, \quad G = [0 \ 10 \ 0]$$

An observer design for $\alpha = 0.001, \ \rho = 2$ yeilds

$$\xi = 23.41, \ \eta = 2942.5, \ P_{23}' = -0.017, P_{22} = 1, P_{33} = 1$$

$$P_{11} = 10^3 \begin{bmatrix} 0.0013 & -0.103 & 0.371 \\ -0.103 & 15.90 & 229.62 \\ 0.371 & 229.62 & 10182.8 \end{bmatrix}, \quad P_{12} = \begin{bmatrix} 0.0083 \\ -2.887 \\ -134.2 \end{bmatrix}, \quad P_{13}' = \begin{bmatrix} 0.321 \\ -6.977 \\ -174.62 \end{bmatrix}$$

Also, we find the gain using $\bar{L} = \begin{bmatrix} L^T & L_{\theta_B}^T & L_{\theta_G}^T \end{bmatrix}^T = \bar{P}^{-1}CQ^{-1}$
Figure 4.13 State 1 estimation

Figure 4.14 State 2 estimation
Figure 4.15 State 3 estimation

Figure 4.16 Bias fault estimation
Figure 4.17 Gain fault estimation

Figure 4.18 Cost function
It is evident in these figures that state estimations are convergent to the actual values and bias and gain faults are also identified simultaneously. The last figure indicates the evolution of the cost function with time which verifies the operation of the observer.
Chapter 5.

Conclusions and Future Work

Designing adaptive observers for nonlinear systems can be beneficial in addressing issues related to control and fault estimation of dynamic systems. This work has presented an optimal nonlinear observer for unknown parameter and fault estimation in nonlinear systems. The observer also presents robustness against noise and disturbances in the system.

The estimation problem is formulated as a dynamic constrained optimization problem using Lagrange multipliers and appropriate cost functions are defined to guarantee minimization of estimation errors. Towards obtaining stationary points of the cost functions with respect to each of the variables, partial derivatives are calculated and a Riccati differential equation formulation is achieved. It is shown that the Riccati equation for the simple case of system without uncertainty is similar to the Riccati equation available for the Extended Kalman Filter with some variations. These variations stem from the conceptual difference among these two methods; especially the way nonlinearity is dealt with in each approach. In the adaptive case, a small modification is made to the cost function to arrive at a more convenient gain calculation procedure and obviate the need for online solution to the Riccati equation. Observer existence conditions and the gain selection procedure are then formulated in terms of Linear Matrix Inequalities that can be conveniently solved using the available solvers. To use the observer in its relaxed form, however, a condition is imposed on a scalar quantity denoting the ratio of derivative of state estimates over state estimations. This ratio should remain smaller than a bound to achieve stability when using the LMI results. The methods are extended to estimate time-varying parameters as well. Finally, the observer
is successfully utilized for developing a robust parameter and sensor fault estimation method for wind turbine mechanical power transmission system. The described system is highly nonlinear and effect of noise and disturbances on the system model is significant.

The results of the observer simulations show satisfactory behaviour in terms of both performance and time-consumption. It also enables us to estimate two types of sensor faults in a vast class of nonlinear dynamical systems. The generality of the nonlinearities accounted for in the observer design allows us to apply this observer to a vast class of systems. The developed observer is computationally less intensive in its relaxed form compared with the alternative method of Extended Kalman Filtering. Furthermore, the observer design procedure does not require knowledge of noise covariance matrices which is a major advantage over the EKF. The steady state noise elimination characteristics are more accurate in the EKF provided that accurate noise covariance matrices are available. The observer, on the other hand, is able to exhibit faster transient response convergence especially for the cases where large initial estimation errors exist.

For future research, the following extensions can be made to this work:

- Including parameter convergence rate in the cost function for improving parameter and fault estimation speed
- Extending the theorems to account for actuator faults
- Developing Fault Tolerant Control (FTC) methods and utilizing the estimated faults for use in FTC
- More advanced modeling of wind turbine including generator and converter models and extending the observer design to estimate faults in the generator and converter
References


