Linear Prediction Approaches to Compensation of Missing Measurements in Kalman Filtering

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by

Naeem Khan MSc Eng.
Control and Instrumentation Research Group
Department of Engineering
University of Leicester

November 21, 2011.

Abstract

Linear Prediction Approaches to Compensation of Missing Measurements in Kalman Filtering

Naeem Khan

Kalman filter relies heavily on perfect knowledge of sensor readings, used to compute the minimum mean square error estimate of the system state. However in reality, unavailability of output data might occur due to factors including sensor faults and failures, confined memory spaces of buffer registers and congestion of communication channels. Therefore investigations on the effectiveness of Kalman filtering in the case of imperfect data have, since the last decade, been an interesting yet challenging research topic. The prevailed methodology employed in the state estimation for imperfect data is the open loop estimation wherein the measurement update step is skipped during data loss time. This method has several shortcomings such as high divergence rate, not regaining its steady states after the data is resumed, etc.

This thesis proposes a novel approach, which is found efficient for both stationary and non-stationary processes, for the above scenario, based on linear prediction schemes. Utilising the concept of linear prediction, the missing data (output signal) is reconstructed through modified linear prediction schemes. This signal is then employed in Kalman filtering at the measurement update step. To reduce the computational cost in the large matrix inversions, a modified Levinson-Durbin algorithm is employed. It is shown that the proposed scheme offers promising results in the event of loss of observations and exhibits the general properties of conventional Kalman filters. To demonstrate the effectiveness of the proposed scheme, a rigid body spacecraft case study subject to measurement loss has been considered.

Acknowledgments

To Almighty ALLAH who gave me the ability and opportunity to accomplish this work;
to my supervisor Professor Dawei Gu, who helped me from the core of his heart;
to my parents, without their prayers this work would not be possible;
to all my colleagues specially Dr. Halim Alwi, Dr. Sajjad Fekri, Rosli Omar, Rihan Ahmed, Dr. Jose Manuel and Dr. Georgios Kladis, Engr. Siddique Ali and Engr. Abdul Hamid;

and to all family members specially to my beloved ones, Mohammad Umer, Maimoonah Khan and Mahreen Khan.

This work is partially supported by Higher Education Commission (HEC) of Pakistan.

Contents

List of Figures

List of Tables

Nomenclature and Abbreviations

Nomenclature

$\|\cdot\|$	Euclidean norm (vectors) or induced spectral norm (matrices)
RV	Random Variable
WSS	Wide-Sense Stationary
AR	Auto-Regressive
ZOH	Zero Order Hold
FOH	First Order Hold
MSD	Mass Spring Damper
DCM	Direction Cosine Matrix
MRP	Modified Rodrigues Parameters
MMSE	Minimum Mean Square Error
LTI	Linear Time Invariant
NE	Normal Equations
LPFO	Linear Prediction Filter Order
LPC	Linear Prediction Coefficients
MLPC	Modified Linear Prediction Coefficients
ARMA	Auto-Regressive Moving Average
MA	Moving Average
ARE	Algebraic Riccati Equation
OLE	Open-Loop Estimation
KF	Kalman Filter

OLKF	Open-Loop Kalman Filter
LOOB	Loss Of OBservations
CKF	Conventional Kalman Filter
CCLKF	Compensated Closed-Loop Kalman Filter
FDI	Fault Detection and Isolation
MBFDI	Model-Based Fault Detection and Isolation
LDA	Levinson-Durbin Algorithm
CLDA	Constraint Levinson-Durbin Algorithm
LPF	Linear Prediction Filter
${}_{ic}P_k$	i^{th} effected error covariance matrix in case of measurement loss
σ	Modified Rodrigues Parameters
ω	Angular velocity
z	Measurement (or observation) vector
$\bar{z}$	Proposed observation vector in case of measurement loss
Q	Process noise covariance matrix
R	Measurement noise covariance matrix
μ	Mean value

Chapter 1

Introduction

1.1 Background and Motivation

In 1940s, Norbert Wiener [112] investigated the minimum variance estimation problem which was restricted to stationary scalar signals and noises. The solution obtained by his invention was not recursive and needed storing of the entire past observed data. In 1960, R. E. Kalman presented the landmark theory of Kalman filtering as an alternative option of formulating the minimum mean square error (MMSE) filtering problem using state space methods [18]. Kalman filter generalises Wiener filter in terms of; a) accommodating vector signals and noises which might not be stationary; b) the solution in general is recursive (it can be seen in later chapters), hence it eliminates the necessity of storing the entire past data [61]. Researchers from various fields were quick to investigate the application of Kalman filter specially in the field of navigation. It has found to be a very efficient practical solution to a number of problems which were previously intractable using conventional method like Wiener filter [9].

Generally speaking, state estimation of a dynamical system deals with recovering requested information from available noisy output data. It is one of the fundamental and significant problem in control and signal processing applications [15]. Kalman filter is one of the best known tools for the state estimation problems of linear time invariant (LTI) systems [21], [4]. The main idea behind Kalman filter algorithm for a dynamical model is to calculate the covariance and gain matrices of the filter. With the help of these matrices, the updated state of the dynamical system is recursively computed from the previous estimate and new input data signals.

Despite many advantages of Kalman filtering algorithm, it however heavily depends on the knowledge of plant dynamics, information of unmeasured stochastic inputs (noise signals), and measured data. In many practical applications, however, these assumptions could be violated. For instance, the intermittent measurement loss, temporary sensor faults and failures, limited bandwidth of communication channels, confined memory space, congestion of a network and many other factors may lead to imperfect measurement data and unwelcome situations. These unpredicted events could make a conventional Kalman filter prone to failure, in producing the "correct" (or acceptable) estimation of states.

The research reported in this thesis started when performing an experiment on 'Configurable Systems Engineering Research Tool' (ConSERT) robots at Loughborough University at Systems Engineering Innovation Centre (SEIC) laboratory. During the experiment, while implementation of a consensus algorithm took place on ConSERT robots [1] when the sensor of one robot got failed and it has to take help from it's neighbour robots, in order to update its own co-ordinates and adjust its position compared to the combined target. That scenario triggered a question on how much a robot's sensor can rely on its previous data to predict its current position.

A literature survey was showing that abundant work has been available where insufficient data is considered in the process of state estimation taking different names such as intermittent observations [10, 99], sensor fault [96], incomplete information [5, 72], loss of data packets [3, 32, 34, 42, 88, 89, 114, 119], loss of information [92], lossy network [20, 90], loss of observations [8, 68, 86, 102, 116]. The main theme mentioned in the above references, is that the process of *prediction* is termed as (state) *estimation* due to the unavailability of output data. This technique is known as Open-Loop Estimation (OLE) or Open-Loop Kalman filtering [88]. Despite being a fast algorithm and comparatively simpler structure in the event of loss of data, OLE suffers from a few of limitations, namely: 1) in the presence of an adequate loss of data (in the time index), OLE diverges at much faster rate; 2) after the data starts resuming back into the system, huge oscillations and sharp spike(s) can be observed in order to obtain the steady state value; and 3) not regaining its steady state (state and covariance) values after the loss is recovered. Open-Loop estimation takes theoretically infinite time to retrieve the steady state values [53].

Hence in this thesis a novel state estimation methodology is proposed in which linear prediction theory is utilised to reconstruct missing data, integration with the process of Kalman filtering. A number of approaches are introduced that can be employed when observations of interest are not available for certain amount of time. Emphasis is made on derivations of the proposed techniques to provide measurement update step for estimation when measurement is not available. The related characteristic properties of the proposed approaches are explored along with the comparison to the existing OLE approach. The proposed algorithm is implemented on two case studies, namely a mass-spring-damper system and a rigid body spacecraft model.

[1] The results have been published in [81] and [57].

1.2 Thesis Contributions and Structure

1.2.1 Contributions

The main contributions of the thesis are outlined as follows:

1. Reconstruction of missing data through constraint linear prediction methods. In the routine procedure of linear prediction, linear prediction coefficients (weights assigned to the previous data samples) are computed based on arbitrary number of previous data samples. However, in practice this conventional approach is found computationally expensive. To overcome this drawback, several constraint-based approaches are proposed (Chapter 4) in order to limit the previous data samples, required to reconstruct the missing data, [53, 55],

2. Introduction of the concept of sliding window (Chapter 4), which consequently generates optimal results in terms of minimum mean square errors, [53],

3. Employment of optimally reconstructed signal in the process of state estimation. In this way, the linear prediction theory is integrated with the well-known state estimator – the Kalman filter, which generates the compensated Kalman filtering schemes, (Chapter 4),

4. The performance of compensated Kalman filter is investigated in terms of various parameters (Chapter 5). These parameters are compared with 1) the existing method of state estimation meant for missing data and 2) the conventional Kalman filter in order to provide complete insight to the proposed compensated Kalman filter schemes, (Chapter 5), [54],

5. The conventional schemes of reproducing the missing data (auto-covariance and autocorrelation) take much longer due to various factors. This shortcoming has been overcome by modifying and implementing the Levinson-Durbin algorithm. Comparison analysis in terms of state estimation error and time computation is performed through simulations, (Chapter 5),

6. In order to explore various unforseen limitations, the proposed compensated Kalman filtering algorithm is implemented on a rigid body spacecraft system subjected to intermittent data loss, (Chapter 6) [52].

1.2.2 Thesis Structure

The thesis is organised as follows:

Chapter 2 is dedicated to the preliminaries, definitions and basic concepts of state estimation and linear prediction theory which are found useful in understanding the existing and proposed state estimation techniques in the event of loss of output data. In addition, the methodology of basic standard Kalman filtering technique along with associated characteristic properties will be discussed.

In Chapter 3, the problem of data loss in Kalman filtering algorithm, the most extensively used solution – the Open-Loop Kalman filter and a brief overview of a few other techniques are described.

The modified linear prediction coefficients strategy for reconstructing missing data is discussed in Chapter 4 on the ground of a straightforward derivation of the proposed compensated Kalman filtering techniques (given the name of Compensated Closed Loop Kalman filter (CCLKF) scheme). Various topologies of the proposed algorithm based on the window of data samples to reconstruct the modified signal, are also defined.

Theoretical characteristics of the proposed schemes are presented in Chapter 5 to include the implementation of CCLKF on a case study of a mass-spring-damper (MSD) system which is subjected to loss of observations. It presents a discussion on the characteristic properties of state estimation in a generalised way and are not restricted to the considered mass-spring-damper example. A comprehensive comparison in terms of quality of estimation, computational demand and error analysis between the existing Open-Loop estimation and the proposed CCLKF approaches is presented.

Chapters 6 is an application chapter which would start with the modelling of a rigid body spacecraft system in Modified Rodrigues Parameter (MRP) and is followed by the implementation of proposed algorithm. It also discusses the design of a Lyapunov stability based controller used for stabilisation purposes of the spacecraft system. The properties described in Chapter 5 are verified in simulation results, in terms of applicability and the performance of both, the existing and proposed approaches.

Chapter 7 concludes the work in this thesis, with proposals for the future perspectives.

A few appendices can be found towards the end of this thesis. They are briefly meant for:

- useful theories related to the existing Levinson-Durbin algorithm, which will assist the proposed modified version of constraint Levinson-Durbin algorithm, (Appendix A),

- linearisation of the rigid body spacecraft model using Modified Rodrigues Parameters (Appendix B),
- computing linear prediction coefficients through the auto-covariance method for a non-stationary process is presented in details (Appendix C).

A graphical view of the thesis plan is shown in Figure 1.1 below.

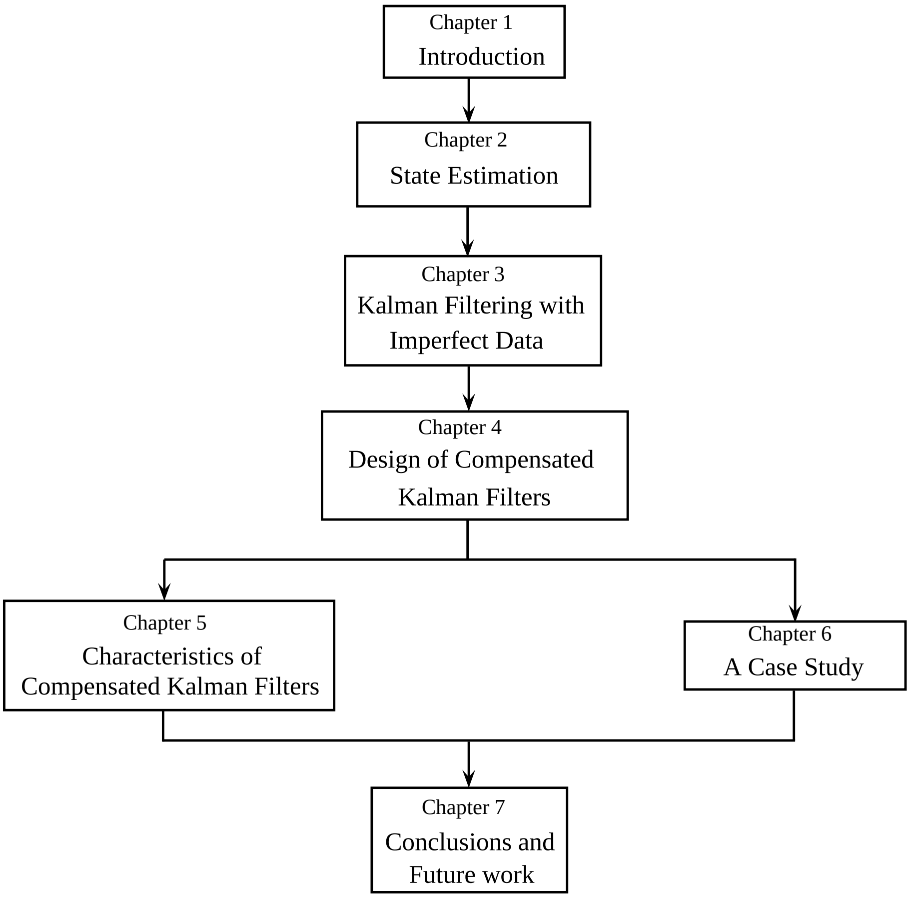

Figure 1.1: Thesis structure

Chapter 2

State Estimation

2.1 Introduction

Filtering is a technique to extract information from noise contaminated observations. If the signal and noise spectra are essentially non-overlapping, the design of a filter that allows the desired signal to pass while unwanted noise attenuating would be a possibility. The resulting filter would be either low pass, band pass/stop or high pass. However, when the noise and information signals are overlapped in spectrum, then the design of a filter to completely separate the two signals would not be possible. In such a situation the information is retrieved through estimations, smoothing or prediction. It will be useful to recall these procedures which are often used in retrieving of information. In simple words, filtering is the recovering of information e.g. $s(t)$ from noisy data say $z(t)$ at time t, using measurements up till time t [4]. Smoothing, on the other hand recovers information about $s(t)$ with the help of measurements later than time t, i.e. $z(t+\lambda)$ for some $\lambda > 0$. It can be called a delay in producing the information about $s(t)$ as compared to filtering operation. Another important concept is prediction, defined as the forecasting of information processing. In prediction, the information about $s(t+\lambda)$ for some $\lambda > 0$ is forecasted based on measurement up till time t. In a more concise way the three methods are summarised in the following statement [82]:

An estimate of a state $x(t)$ at time $[t+\tau]$, is $\hat{x}(t+\tau)$ using the observed data $\{z(s),\ s \in [0,t]\}$ where the time argument t may belong to discrete set $\{t_0 = 0, t_1, t_2, ...\}$. This estimate is required to be optimal with respect to least square estimation (LSE). For

- $\tau = 0$, *the process is called filtering;*
- $\tau < 0$, *the process is called smoothing;*
- $\tau > 0$, *the process is called prediction.*

State estimation has long been an active research subject in the control area because of its significance in understanding system behaviour and the design of control schemes. For example [98],

- state of an unstable system is estimated in order to implement a state feedback controller,
- winding current is estimated in order to control direction and speed of a DC motor,
- attitude is estimated to control the angular position and velocity of a satellite,
- sugar level is estimated to evaluate the health of a patient,
- economic growth of a product is estimated to prepare future strategy, and so on.

A few celebrated techniques for state estimation are Kalman filtering (and all its categories), particle filtering, H_∞ filtering [22], *etc.*

This chapter is organised as follows: Section 2.2 provides some basic definitions which help understanding of the proposed algorithms in the subsequent chapters. Based on this preliminary work, a simple structure of an unbiased estimator is presented in Section 2.3. A detailed description of standard discrete time Kalman filter algorithm is presented in Section 2.4 along with the list of areas where it is successfully implemented. The asymptotic stability related with Kalman filter is mentioned in Section 2.5.

2.2 Definitions

It is necessary to present some basic definitions as the introduction on random variables and their statistical properties such as mean, variance, correlation and distribution of random processes. The description is provided in a concise manner, the details can be found in [4, 14, 21, 32, 46, 98, 107].

Random Variable

In probability, a random variable is a variable whose value is not known. The possible values of a random variable might represent the possible outcomes of a yet-to-be-performed experiment [21]. A random variable (or a random signal) cannot be characterised by a simple, well-defined mathematical equation and its future values cannot be precisely predicted. However, the use of probability and statistical properties (e.g. mean and variance) is recommended to analyse the behavior of a random variable. If event $\mathcal{E}$ is a possible outcome of a random experiment then the probability of this event can be denoted by $p(\mathcal{E})$ and generally all possible outcomes of a random experiment represented by $\mathcal{E}_j$, (if finite) $\forall\ j = 1, 2, ...n$, follow

$$0 \leq p(\mathcal{E}_j) \leq 1 \qquad \& \qquad \sum_{j=1}^{n} p(\mathcal{E}_j) = 1 \tag{2.1}$$

The probability density function of a continuous random variable (RV) x, is a function which is summed to obtain the probability that the variable takes a value in a given interval. It is defined as

$$f(x) = \frac{dF(x)}{dx} \tag{2.2}$$

where $F(x)$ is the cumulative distribution function. [1] The probability density function satisfies

$$F(\infty) = \int_{-\infty}^{\infty} f(x)dx = 1 \tag{2.4}$$

Similarly, for a discrete random variable (RV), probability distribution (or probability function) is a list of probabilities associated with each of its possible value, i.e.

$$p(x_i) = P[X = x_i] \tag{2.5}$$

The probability distribution satisfies criteria defined in (2.1).

Mean

The expected or mean value of a random variable (RV) indicates the average or central value of that random variable. The mean value gives a general impression about an RV instead of complete details of its probability distribution (discrete RV) or probability density function (continuous RV). Two random variables with different distribution (or density function) might have the same mean values, hence only the mean value does not reveal complete information of a random variable. If X is a continuous RV with probability density function $f(x)$, the expected or mean value can be represented as

$$\mu = E[X] := \int_{-\infty}^{\infty} x f(x)dx \tag{2.6}$$

Similarly, if X is a discrete RV with possible values x_i where $i = \{1, 2, 3..., n\}$ and its probability

[1] The cumulative distribution function gives the probability that the random variable X satisfies the followings:

$$F(x) = p(X < x) \qquad \forall \quad -\infty < x < \infty \tag{2.3}$$

$p(x_i)$ is denoted as $P(X = x_i)$, then the mean or expected value can be defined as

$$\mu = E[X] := \sum x_i P(X = x_i) \tag{2.7}$$

where the elements are summed over all values of the random variable X.

Variance

The variance of a random variable shows how widely the values of that RV are likely to be. The larger the variance, the more scattered the observations about its mean value. In other words, variance shows the concentration of distribution about the mean value of that RV. Variance of continuous RV x, denoted by σ^2 or $V(x)$ is defined as

$$\sigma^2 = V(x) \;:=\; \int_{-\infty}^{\infty} (x - E(x))^2 f(x)dx \tag{2.8}$$

The square root of the variance of a random variable is known as standard deviation (σ).

If the random variable X is *discrete* then the variance can be defined as

$$\begin{aligned} Var[X] &= E[(X-\mu)^2] \\ &= E[X^2 - 2\mu X + \mu^2] \\ &= E[X^2] - 2\mu E[X] + \mu^2 \\ &= E[X^2] - \mu^2 \\ &= E[X^2] - E[X]^2 \end{aligned} \tag{2.9}$$

The variance and standard deviation of a random variable are always non-negative.

Correlation and Correlation Coefficient

In statistics, correlation shows how strongly two or more variables are related to one other. Linear correlation (also known as Pearson's correlation) determines the extend to which two variables are linearly related (or proportional) to each other.

Correlation coefficient denoted by ρ is a number. It indicates the degree to which two variables

X and Y are linearly related. The correlation coefficient can be expressed as

$$\begin{aligned} \rho_{X,Y} &= \frac{Cov(X,Y)}{\sigma_X \sigma_Y} \\ &= \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} \\ &= \frac{E[(X-E(X))(Y-E(Y))]}{\sqrt{E(X^2)-E^2(X)}\sqrt{E(Y^2)-E^2(Y)}} \\ &= \frac{E(XY)-E(X)E(Y)}{\sqrt{E(X^2)-E^2(X)}\sqrt{E(X^2)-E^2(Y)}} \end{aligned} \tag{2.10}$$

If the two variables are independent then $\rho = 0$ and if one variable (say Y) is a linear function of the other (X), then $\rho = [1, -1]$. An important note made here is this that correlation (or correlation coefficient) measures linear relationship only. However, it is possible that there exist a strong non-linear relationship between two variables while $\rho = 0$. A scattered diagram can give some information about any relationship between two variables. Also, the correlation coefficient does not show the influence of one variable on the other.

Autocorrelation

Autocorrelation is also some times called as lagged or serial correlation. It is the correlation of a signal with itself [79]. It indicates the similarity between observations as a function of time separation between them. In other words, autocorrelation of a random process describes the correlation between the values of the same process at different points in time. Therefore, if z is a process with mean μ and variance σ^2 then autocorrelation $\gamma_{(t,m)}$ at two different time instants t and m can be defined as

$$\gamma_{(t,m)} := \frac{E[(z_t - \mu_t)(z_m - \mu_m)]}{\sigma_t \sigma_m} \tag{2.11}$$

If the variance of this process has some value (not equal to zero) then γ_m is well defined between the range of 1 and -1. If z_t is a second order stationary process (sometimes called *wide-sense stationary* of WSS process), i.e. mean and variance are time independent, than autocorrelation only depends on the time-distance between the pair of values. Therefore, if

$$\tau := t - m \tag{2.12}$$

The last equation can be re-written as follows:

$$\gamma_\tau = \frac{E[(z_t - \mu)(z_{t+\tau} - \mu)]}{\sigma^2} = \gamma_{-\tau} \tag{2.13}$$

Autocorrelation helps in finding repeating patterns, such as periodic signals or fundamental frequency of a signal which cannot be determined due to some unwanted factors like noise. A positive autocorrelation value indicates some sort of tendency for a system to remain in the same state from one time instant to another. For example, the likelihood of tomorrow being rainy is higher if today is rainy than if today is dry [11]. Autocorrelation will be further discussed and utilised in Section 4.2.

Normal Distribution

The normal distribution (also known as Gaussian distribution) is a very important class of statistical distribution. This type of distribution is symmetric about the mean value and has a bell-shaped density curve with a single peak. Truly speaking, any normal distribution can be fully specified by two quantities: the mean μ (or m), where the peak of the density lies, and the standard deviation σ, which indicates the spreading distribution. In other words, if a real valued random variable X is normally distributed with mean m and variance $\sigma^2 \geq 0$, then it can be written as

$$X \sim N(m, \sigma^2) \tag{2.14}$$

i.e. the random variable X is normally distributed with mean m and variance σ^2. The normal probability density function of a scalar value can be shown analytically to be

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{[-\frac{(x-m)^2}{2\sigma^2}]} \tag{2.15}$$

The integral of the above function (which corresponds to the area under the curve) is unity. Sketch of a typical normal distribution is shown in Figure 2.1 for $\sigma = 4$ and $m = 13$.

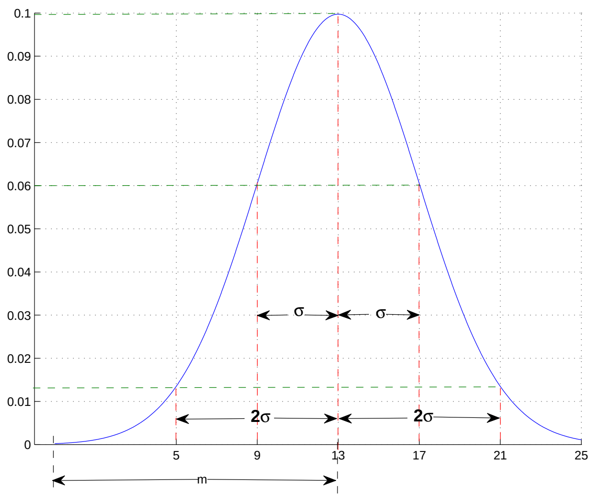

Figure 2.1: Normal probability density function (Figure adopted from [21]).

The probability that a normal distributed function lies outside $\pm 2\sigma$ is approximately 0.05 [21]. Standard normal distribution is defined with mean $m = 0$ and standard deviation $\sigma = 1$.

Therefore, from Equation (2.15), a standard normal distribution can be written as

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{[-\frac{x^2}{2}]} \tag{2.16}$$

The results of normal distribution (2.15) can be extended to two and more than two random variables (bivariate and multivariate normal distribution) as below:

$$f(x_1, x_2) = \frac{1}{\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{\left[-\frac{\frac{x_1^2}{\sigma_1^2} - 2\rho\frac{x_1x_2}{\sigma_1\sigma_2} + \frac{x_2^2}{\sigma_2^2}}{2(1-\rho^2)}\right]} \tag{2.17}$$

where ρ is the correlation between the two random variables x_1 and x_2. Similarly

$$f(x_1, x_2, ..., x_n) = \frac{1}{(2\pi)^{n/2}|P|^{1/2}} e^{[-\frac{(x-m)^T P^{-1}(x-m)}{2\sigma^2}]} \tag{2.18}$$

where $m := E[x]$ and $P := E[(x-m)(x-m)^T]$ are the mean and covariance of the vector x respectively.

Random Process

A random or stochastic process is a process whose behaviour is not completely predictable. A random process can be characterised by statistical properties. Due to randomness of random variables, the average values from a collection of signals rather than one individual signal, are usually studied [21]. Therefore, a random process $\{X(t), t \in T\}$ is a family of random variables where the index set T might be discrete ($T = \{0, 1, 2...\}$) or continuous ($T = [0, \infty)$). For a discrete time system, the word random sequence is the preferred terminology to use instead of random process. Daily life examples of random processes are stock index in Leicester Wholesale fruit market and hourly rainfall at city center *etc.*

Stationary Process

Stationarity is the quality in which statistical properties of a process do not change with time [106]. In other words, the probability density function (continuous case) or the probability distribution (discrete case) of a stationary random variable X_1 is independent of time t. Wherever the distribution is seen for some segment, the dynamics remain the same. This means that it does not matter when in time one observes the process.

Definition 1:

A process $\{X_t, t \in \mathbb{R}^n\}$ (where n is a positive integer) is said to be a weak-sense stationary (WSS) process if

- $E[X^2] < \infty \ \forall \ t \in \mathbb{R}^n$,

- $E[X_t] = \mu_t = \mu$, and
- $Cov(X_t, X_m) = E[(X_t - \mu_t)(X_m - \mu_m)] = \sigma^2(t, m) = \sigma^2(t - m)$

In other words one can say that a weak-sense stationary process must have finite variance, constant mean (first moment) and the second moment should only depend on $(t - m)$ and not on t or m [104].

Definition 2:
A process $\{X_t, t \in \mathbb{R}\}$ is strictly stationary process if all its higher-order moments are independent of time, i.e.

$$\begin{aligned} p(X_{t1} \leq x_1, X_{t2} \leq x_2, ..., X_{tn} \leq x_n) &= F(x_{t1}, x_{t2}, ..., x_{tn}) \\ &= F(x_{h+t1}, x_{h+t2}, ..., x_{h+tn}) \\ &= p(X_{h+t1} \leq x_1, X_{h+t2} \leq x_2, ..., X_{h+tn} \leq x_n) \end{aligned} \tag{2.19}$$

for any time shift t and h. Most statistical prediction methods are based on the assumption that the time series can be considered approximately stationary.

Gaussian Random Variable and Vector

A random variable X will be Gaussian variable (synonymously called normal random variable) if its probability density function $f_X(x)$ follows

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{[-\frac{(x-m)^2}{2\sigma^2}]} \tag{2.20}$$

where $m = \mu = E[X]$ and $\sigma = E[(X - m)^2] > 0$ are the mean and variance values of the random variable X, respectively [78]. For Gaussian random variables, a very common notation can be observed in literature, as

$$X \sim N(m, \sigma^2) \tag{2.21}$$

Similarly, if X is a vector of n-random variables and X has a nonsingular covariance matrix, then X will be a Gaussian (or normal) random vector if and only if its probability density is of the form

$$f_X(x) = \frac{1}{(2\pi)^{n/2}|P|^{1/2}} e^{[-\frac{(x-m)^T P^{-1}(x-m)}{2\sigma^2}]} \tag{2.22}$$

where $m = E[X]$ and $P = E[(X - m)(X - m)^T]$ are the mean and covariance of the vector X

respectively. The frequently encountered notation for a Gaussian random vector is

$$X \sim N(m, P) \tag{2.23}$$

A standard normal variable Z is defined for $\mu = 0$ and $\sigma = 1$, i.e.

$$Z \sim N(0, 1) \tag{2.24}$$

The probability density function for a standard normal variable implies

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(z)^2}{2}} \tag{2.25}$$

Gaussian or normal distribution is a frequently encountered assumption in many control and communication problems. This is because of its simple bell-shaped structure which requires less information (mean and variance) to store, receive and transmit.

White Noise

White noise is a random signal (or process) in which the power spectrum is distributed uniformly over all the frequency components in the full infinite range. This is purely a theoretical concept because if a signal has the same power at all components frequency, this is equivalent to a signal whose total power is infinite and therefore impossible to generate such a signal . However in practice, for finite frequency range "white" signal with a flat spectrum can be easily assumed.

Mathematically, a random signal v is a white signal if and only if it possesses the following two properties:

$$\begin{aligned} m_v &= E(v) = 0 \\ R_v &= E(vv^T) = \sigma^2 I \end{aligned} \tag{2.26}$$

A continuous white noise process w, holds the same properties as:

$$\begin{aligned} m_w(t) &= E(w(t)) = 0 \\ Q_w(t_1, t_2) &= E(w(t_1)w(t_2)^T) = \frac{S}{2}\delta(t_1 - t_2) \end{aligned} \tag{2.27}$$

i.e. it has zero mean and infinite power at zero time shift.

Often, when data (or measurement) is corrupted by some unwanted (noise) signal, the information inside the data is obtained by filtering or applying prediction (or estimation). A simple optimal estimator would be the one which bears the minimum prediction or estimation error, e.g. least square estimator.

Least Square Estimation

Linear least square estimation is an estimating method for unknown parameters in a linear regression model aiming to minimise the sum of the square of errors. Consider an unknown state vector x is related to the data or observation vector z as follows:

$$z = Hx + v \tag{2.28}$$

where v is the sensor noise. The variables are shown without any subscripts for simplicity with the aiming an estimate $\hat{x}$ that minimises the estimated measurement error $\varepsilon := z - H\hat{x}$. The estimation error can be characterised in terms of Euclidean vector norm $|\varepsilon|$, alternately – to minimise the scalar cost function J, where

$$J = (z - H\hat{x})^T(z - H\hat{x}) \tag{2.29}$$

To minimise this scalar cost function, standard minimisation procedure is carried out as:

$$\frac{\partial J}{\partial \hat{x}} = 0 \tag{2.30}$$

Substituting the value of J in Equation (2.30), and after a little algebra leads to

$$H^T H\hat{x} = H^T z \tag{2.31}$$

Equation (2.31) is sometimes called normal equation for the least square problem [32]. It can solved as

$$\hat{x} = (H^T H)^{-1} H^T z \tag{2.32}$$

provided the $(H^T H)$ is nonsingular. This product is known as Gramian matrix i.e. $\mathcal{G} = H^T H$

A few properties associated with random variables and processes namely mean, variance, correlation, correlation coefficient and distribution have been discussed in the preceding section. In the following section, the basis of the minimum variance estimator is uncovered.

2.3 A Classical Unbiased State Estimator

In statistics, bias of an estimator is the difference of estimated value and the actual value of the parameter (or state) being estimated. This difference is called estimation error. An unbiased estimator is one which yields zero estimation error while estimating the parameter or state of a system. It is difficult to design a completely unbiased estimator but to start with, it is assumed that the estimated value is equal to the actual value. This section is initialised with the standard state space realisation model of a discrete time linear system.

2.3.1 Discrete Case

Consider a discrete time LTI system

$$x_{k+1} = Ax_k + Bu_k + Gw_k \tag{2.33}$$

$$z_k = Cx_k + Du_k + v_k \tag{2.34}$$

Unbiased state estimator would be the one which gives zero mean error as the time step k approaches to infinity. Assuming the proposed structure of the discrete time estimator design is

$$\hat{x}_{k+1} = A_d\hat{x}_k + B_d u_k + K_d z_k, \tag{2.35}$$

where the subscript 'd' denotes discrete time system. K_d is the gain matrix of the designed estimator. The two noise sequences 'w_k' and 'v_k' are assumed to be zero mean, uncorrelated, Gaussian white noise signals. The corresponding error generated by the estimated state is

$$\begin{aligned} e_{k+1} &= x_{k+1} - \hat{x}_{k+1} \\ &= Ax_k + Bu_k + Gw_k - A_d\hat{x}_k - B_d u_k - K_d z_k \\ &= Ax_k + Bu_k + Gw_k - A_d\hat{x}_k - B_d u_k - K_d(Cx_k + Du_k + v_k) \\ &= (A - K_dC)x_k - A_d\hat{x}_k + (B - B_d - K_dD)u_k + Gw_k - K_dv_k \\ &= (A - K_dC)e_k + (A - K_dC - A_d)\hat{x}_k + (B - B_d - K_dD)u_k + Gw_k - K_dv_k \end{aligned} \tag{2.36}$$

The expected value of the error signal will be

$$E[e_{k+1}] = (A - K_dC)E[e_k] + \underbrace{(A - K_dC - A_d)}\,E[x_k] + \underbrace{(B - B_d - K_dD)}\,E[u_k] \tag{2.37}$$

For the above unbiased state estimator, the mean of the estimated error tends to be zero as the time step k tends to infinity,(i.e. $\lim_{k\to\infty} E[e_k] = 0 \ \forall \ u_k \ \forall \ E[\hat{x}_k]$ when $A_d = A - K_dC$, $B_d = B - K_dD$ and $(A - K_dC)$ is stable. Therefore, substituting the values of A_d and B_d in the proposed estimator (2.35) would result in,

$$\hat{x}_{k+1} = (A - K_dC)\hat{x}_k + (B - K_dD)u_k + K_dz_k \tag{2.38}$$

or

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + K_d(z_k - C\hat{x}_k - Du_k) \tag{2.39}$$

This is the discrete time general state estimator which results in zero error state estimation. Based on the concept of this unbiased state estimator, the theory of Kalman filter and its brief historical background is discussed in the following section.

2.4 Kalman Filtering

This section presents the basic, necessary and sufficient studies related to state estimation in the process of Kalman filtering to establish the proposed work of this thesis. The earlier sections were aimed at the foundation for Kalman filtering and the later chapters are written for innovative algorithms based on the concept presented in this chapter. It will be discussed that Kalman filter (KF) heavily depends on the output measurement data (in fact the true estimation is based on this information) which leads to unpleasant scenarios of the utilisation of KF if this information is not available.

In early 1940's Kolmogorov and Wiener discussed the problem of linear least square estimation for a stochastic process [46]. Kolmogorov formulated the filter design problem using statistical and frequency-domain characteristics. However their results are based while considering stationary processes with infinite or semi infinite observation interval. R.E. Kalman [47], presented the design of now well-known Kalman filter, which is presently in extensive use. Kalman filtering is an on-line recursive method adopted to estimate the state of a system with noise contaminated observations. Broadly speaking, in Kalman filtering there are two steps undertaken throughout the estimation process, one is the time update step (also known as *a priori* estimate or prediction) and the other is the measurement update (also known as *a posteriori* estimate or filtering update). KF's principle thus vigorously depends on measurement data.

KF is one the best tools available, providing optimal state estimation with minimum mean square error based on some information on system model and assumptions on uncertainties (noise). In simple words, a steady state constant gain KF is a recursive algorithm that is utilised to estimate the state variables of a linear time invariant (LTI) system, subject to stochastic noises, based on certain noise contaminated output variables [107]. Its simple structure and straightforward design methodology made it popular almost in every field of research, reflected in increasing application area. Thousands of research papers have been written about KF and its numerous applications such as navigation, tracking, power control, estimator and controller design in defence and industry *etc.* Table 2.4 gives a short summary of the application area in which KF has been excessively implemented.

No.	Application	References
1	Vehicle navigation	[40], [50], [1]
2	DC motors	[120], [91]
3	Aircraft system	[60], [117], [105]
4	Chemical process	[66], [13], [43]
5	Camera calibration	[71], [12]
6	Fault detection, isolation and accommodation	[44], [80], [39], [38]
7	Uncertain systems	[28], [118], [93]
8	Integrated fault detection	[124]
9	Spacecraft attitude estimation	[64], [56], [86]
10	Object visualisation	[7], [108], [115]
11	Human body tracking	[26], [121]
12	Embedded systems	[97]
13	Power generation and quality	[25], [30], [74]
14	Underwater vehicles	[63], [111]
15	Internet applications	[45], [90]
16	Chaotic synchronisation	[62]
17	Miscellaneous applications	[32] [21], [98], [15], [65]

Table 2.1: Kalman filter applications

Theoretically KF is an estimator for the instantaneous state of a linear dynamic system perturbed by white noise. With a few mathematical and stochastic process concepts, KF can be easily implemented in very sophisticated designs as KF is statistically '*optimal*' with respect to quadratic function of estimation error. Numerous applications can be found for KF in control of dynamic systems such as continuous manufacturing process, aircraft, ships or even attitude of a spacecraft system. It is quite often not possible or required to measure every variable that one wants to control and KF is found a better tool to infer such information from noisy measurements. In addition, KF is found a convenient and worthy solution for other areas beyond control, including prices of traded commodities and the flow of rivers during flood [32].

Rapid evolution in the digital control technology has made it possible to design very complicated controllers at a very low cost with less computational time. The increasing demand of digital computers has also altered the control system design options [107]. In general, digital control systems have many advantages over analog control systems including the following:

- low cost, weight and power consumption;
- high accuracy and reliability;
- ease of making software and design changes.

The majority of state estimation tools and control algorithms are implemented in digital electronics, and that is why researchers are normally concerned with eventual implementation [98].

These facts have brought us towards more emphasis on discrete time systems.

For this reason, this section is aimed at describing the basic Kalman filtering technique for discrete LTI systems. Attempts have been made to describe related features in the process of Kalman filtering such as estimated state, error covariance, Kalman filter gain, and residual vector. Asymptotic stability of the Algebraic Riccati Equation (ARE) associated with KF is also discussed in Section 2.5.

2.4.1 Design of A Discrete-time Kalman Filter

Assuming the discrete time system, for which state variables represented by the vector x, are governed by the equation

$$x_{k+1} = Ax_k + Bu_k + \xi_k \tag{2.40}$$

where $x \in \mathbf{R}^n$, $A_{n\times n}$ is the transition matrix, $u \in \mathbf{R}^m$ the deterministic input, $B_{n\times m}$ the control matrix and ξ the process noise which is assumed to be a Gaussian, zero mean white noise. The process noise ξ can be characterised as

$$\xi_k \sim N(0, Q_k) \tag{2.41}$$

in other words,

$$E[\xi_k] = 0 \qquad \& \qquad E[\xi_k \xi_k^T] = Q_k \tag{2.42}$$

The output observation equation is

$$z_k = Hx_k + \theta_k \tag{2.43}$$

in which $z \in \mathbf{R}^l$ is the measurement vector, $H \in \mathbf{R}^{l\times n}$ the output matrix and θ_k the measurement noise which is assumed as Gaussian, zero mean whit noise. i.e.

$$\theta_k \sim N(0, R_k) \tag{2.44}$$

or

$$E[\theta_k] = 0 \qquad \& \qquad E[\theta_k \theta_k^T] = R_k \tag{2.45}$$

The statistical properties of matrix R can be obtained from the statistical properties of the measurement devices and sensors. Suppose the measurement vector z constitutes l variables of interest, i.e. $z = (z_1, z_2, ...z_l)^T$. If $m := E[z]$ and $\sigma^2 := E[(z-m)(z-m)^T]$ are mean and variance of the measured signal, then the general form of the probability density of the measurement will be

$$p(Z) = \frac{1}{2\pi\sigma} exp[-\frac{1}{2}\frac{(z-m)(z-m)^T}{\sigma^2}] \tag{2.46}$$

or

$$p(z_1, z_2, ...z_n) = \frac{1}{(2\pi)^{n/2}\sqrt{\mid R \mid}} exp[-\frac{1}{2}(z-m)^T R^{-1}(z-m)] \tag{2.47}$$

m is the mean of the distribution. In the above equation R is the covariance matrix and can be found from

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1l} \\ r_{21} & r_{22} & \cdots & r_{2l} \\ \vdots & \vdots & \ddots & \vdots \\ r_{l1} & r_{l2} & \cdots & r_{ll} \end{bmatrix} \tag{2.48}$$

where $r_{ij} = E[(z_i - m_i)(z_j - m_j)] = \sigma_i \sigma_j$.

2.4.2 Optimal State Estimation

To better understand state estimation through KF, discrete time Kalman filtering can be broken down into two steps: *a priori* state estimation and *a posteriori* state estimation [123].

The *a priori* state estimation of the system (2.40) at time step k is represented by $x_{k+1|k}$ [2] and *a priori* error covariance of this estimate is denoted by $P_{k+1|k}$. The *a priori* or time update step is solely based on system model as it is calculated before incorporating measurements, i.e.

$$x_{k+1|k} = Ax_{k|k} + Bu_k \tag{2.49}$$

The *a priori* error will be

$$\begin{aligned} e_{k+1|k} &= x_{k+1|k} - x_{k+1} \\ &= Ax_{k|k} + Bu_k - Ax_k - Bu_k - \xi_k \\ &= Ae_{k|k} - \xi_k \end{aligned} \tag{2.50}$$

In a Gaussian process with zero mean and uncorrelated properties of the process noise, Equation (2.50) will lead to the *a priori* error covariance matrix defined as

$$\begin{aligned} P_{k+1|k} &:= E[e_{k+1|k} e_{k+1|k}^T] \\ &= AP_{k|k}A^T + Q_k \end{aligned} \tag{2.51}$$

KF is aimed at designing a recursive estimator in order to minimise the error in the state estimation [73]. Since the state and measurements are partly determined by stochastic process

[2] The subscript $\{l + m|m\}$ represents the prediction of $l + m$ entity at time step m.

(ξ_k and θ_k), the variables x and z are assumed jointly Gaussian, hence it is sufficient to seek the update of the *a priori* state estimation and covariance based on observation z_k [21] i.e.

$$x_{k+1|k+1} = \grave{K}_{k+1} x_{k+1|k} + K_{k+1} z_{k+1} \tag{2.52}$$

In the above equation the measurement is utilised to refine the *a priori* estimate. Thats why it is also called measurement update state estimation at time instant $k+1$.

2.4.3 Optimisation Problem

The two gain matrices $\grave{K}_{k+1}$ and K_{k+1} have to be determined optimally so that the estimate carries minimum error. The *a posteriori* estimate error is

$$e_{k+1|k+1} = x_{k+1|k+1} - x_{k+1} \tag{2.53}$$

From Equation (2.50), it can be deduced

$$x_{k+1|k} = e_{k+1|k} + x_{k+1} \tag{2.54}$$

Substituting Equation (2.52) in (2.53), the following result can be obtained;

$$\begin{aligned} e_{k+1|k+1} &= [\grave{K}_{k+1} x_{k+1|k} + K_{k+1} z_{k+1}] - x_{k+1} \\ &= \grave{K}_{k+1} x_{k+1|k} + K_k H x_{k+1} + K_{k+1} \theta_{k+1} - x_{k+1} \\ &= [\grave{K}_{k+1} (e_{k+1|k} + x_{k+1})] + K_{k+1} H x_{k+1} + K_{k+1} \theta_{k+1} - x_{k+1} \\ &= [\grave{K}_{k+1} + K_{k+1} H - I] x_{k+1} + K_{k+1} \theta_{k+1} + \grave{K}_{k+1} e_{k+1|k} \end{aligned} \tag{2.55}$$

With the assumptions of unbiased estimation and zero mean white noise sequences, it can be written that $E[e] = 0$ and $E[\theta_k] = 0$, therefore the expectation of (2.55) would result in

$$E[e_{k+1|k+1}] = E[(\grave{K}_{k+1} + K_{k+1} H - I) x_{k+1}] \tag{2.56}$$

The *a posteriori* state estimation error $E[e_{k+1|k+1}] = 0$ if

$$\grave{K}_{k+1} + K_{k+1} H - I = 0$$

or

$$\grave{K}_{k+1} = I - K_{k+1} H \tag{2.57}$$

Substituting this value in Equation (2.52) will generate

$$\begin{aligned} x_{k+1|k+1} &= [I - K_{k+1}H]x_{k+1|k} + K_{k+1}z_{k+1} \\ &= x_{k+1|k} + \underbrace{K_{k+1}[z_{k+1} - Hx_{k+1|k}]}_{correction-term} \end{aligned} \tag{2.58}$$

where K_{k+1} is the Kalman filter gain matrix, which provides weight to the correction term [21,32] in Equation (2.58). The correction term depends on the term $(z_{k+1} - Hx_{k+1|k})$ which is known as residual or innovation vector. It will be shown at later stages that K_{k+1} itself depends on residual covariance. It is worth mentioning that as the measurement error covariance R approaches zero, the actual measurement z_k is trusted more and more, and the predicted measurement (Hx_{k-1}) is credited less and less [109]. On the other hand, if the *a priori* (or predicted) estimate error covariance $(P_{k+1|k})$ approaches zero (in other words if measurement noise covariance matrix R increases), the belief in actual measurement (z_k) decreases while confidence on predicted measurement (Hx_{k-1}) increases.

The *a posteriori* error covariance matrix can be calculated as,

$$\begin{aligned} P_{k+1|k+1} &= E[e_{k+1|k+1}e_{k+1|k+1}^T] \\ &= E[(x_{k+1|k+1} - x_{k+1})(x_{k+1|k+1} - x_{k+1})^T] \\ &= E[\{x_{k+1|k} - x_{k+1} + K_{k+1}(Hx_{k+1} + \theta_{k+1} - Hx_{k+1|k})\} \\ &\quad \{x_{k+1|k} - x_{k+1} + K_{k+1}(Hx_{k+1} + \theta_{k+1} - Hx_{k+1|k})\}^T] \\ &= E[\{e_{k+1|k} + K_{k+1}(\theta_{k+1} - He_{k+1|k})\}\{e_{k+1|k} + K_{k+1}(\theta_{k+1} - He_{k+1|k})\}^T] \\ &= E[\{(I - K_{k+1}H)e_{k+1|k} + K_{k+1}\theta_{k+1}\}\{e_{k+1|k}^T(I - K_{k+1}H)^T + \theta_{k+1}^T K_{k+1}^T\}] \\ &= (I - K_{k+1}H)P_{k+1|k}(I - K_{k+1}H)^T + K_{k+1}R_{k+1}K_{k+1}^T \end{aligned} \tag{2.59}$$

This is the Riccati equation for the discrete time KF. For optimal value of Kalman filter gain K, which will provide minimum mean square error, define a cost function for the discrete time KF;

$$J = E[e_{k+1|k+1}^T e_{k+1|k+1}] \tag{2.60}$$

This is equivalent to say that

$$\begin{aligned} J &= trace\big(E[e_{k+1|k+1}e_{k+1|k+1}^T]\big) \\ &= trace(P_{k+1|k+1}) \\ &= trace[(I - K_{k+1}H)P_{k+1|k}(I - K_{k+1}H)^T + K_{k+1}R_{k+1}K_{k+1}^T] \end{aligned} \tag{2.61}$$

Minimising the cost function with respect to K_{k+1} as follows:

$$\frac{\partial J}{\partial K_{k+1}} = -2(I - K_{k+1}H)P_{k+1|k}H^T + 2K_{k+1}R_{k+1} = 0 \tag{2.62}$$

therefore

$$K_{k+1}R_{k+1} = P_{k+1|k}H^T - K_{k+1}HP_{k+1|k}H^T$$
$$K_{k+1}R_{k+1} + K_{k+1}HP_{k+1|k}H^T = P_{k+1|k}H^T$$
$$K_{k+1}(HP_{k+1|k}H^T + R_{k+1}) = P_{k+1|k}H^T$$
$$K_{k+1} = P_{k+1|k}H^T(HP_{k+1|k}H^T + R_{k+1})^{-1} \tag{2.63}$$

substitute this value in Equation (2.59) would result in

$$\begin{aligned} P_{k+1|k+1} &= (P_{k+1|k} - \underbrace{P_{k+1|k}H^T(HP_{k+1|k}H^T + R_{k+1})^{-1}}HP_{k+1|k}) \\ &= (I - K_{k+1}H)P_{k+1|k} \end{aligned} \tag{2.64}$$

In Equation (2.63), the factor $(HP_{k+1|k}H^T + R_{k+1})$ is known as covariance matrix of residual or innovation vector and is denoted by S. i.e.

$$S_{k+1} = HP_{k+1|k}H^T + R_{k+1} \tag{2.65}$$

The actual residual vector can be found out from equation

$$r_{k+1} = z_{k+1} - Hx_{k+1|k} \tag{2.66}$$

Equations (2.65) and (2.66) can be interpreted as: the elements of the r_{k+1} can be compared with the standard deviation obtained by taking square root of diagonal elements of S_{k+1} [4]. In order to achieve the correct solution for the problem, the computed residual should be increased to maximum of two standard deviation [85].

Equations (2.49),(2.51),(2.58), (2.63), and (2.64) completely describe the structure of standard discrete time Kalman filter [4,21,98]. In Figure 2.2 and Algorithm 1, the complete procedure of state estimation in Kalman filtering is summarised.

Note: There are various mathematically equivalent formats which are frequently encountered in the analysis of KF design. The detail of which can be found in [21], [4] and [98]. The following equation, appeared in numerous articles, is of great interest to the author from stability

Plant Dynamics Output Dynamics	$x_{k+1} = Ax_k + Bu_k + \xi_k$ $z_k = Hx_k + \theta_k$
Assumptions	$E[x_0] = \hat{x}_0$, $E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T] = P_0$ $\xi_k \sim N(0, Q_k)$, $\theta_k \sim N(0, R_k)$
State Prediction Error Covariance Prediction	$x_{k+1\|k} = Ax_{k\|k} + Bu_k$ $P_{k+1\|k} = AP_{k\|k}A^T + Q_k$
Residual Calculation	$r_{k+1} = z_{k+1} - Hx_{k+1\|k}$
Kalman filter Gain	$K_{k+1} = P_{k+1\|k}H^T(HP_{k+1\|k}H^T + R_{k+1})^{-1}$
Measurement updated state Measurement update error covariance	$x_{k+1\|k+1} = x_{k+1\|k} + K_{k+1}(r_{k+1})$ $P_{k+1\|k+1} = (I - K_{k+1}H)P_{k+1\|k}$

Figure 2.2: Recursive behaviour of conventional Kalman filter.

Algorithm 1 : Discrete-Time Kalman filtering

1: Initialise $x_{0|0}$, u_0, ξ_0, θ_0, $P_{0|0}$ and $k = 0$.
2: Prediction cycle:
$x_{k+1|k} = Ax_{k|k} + Bu_k$; State estimation
$P_{k+1|k} = AP_{k|k}A^T + Q_k$; Error covariance
3: Sensed measurements: $z_{k+1} = Hx_{k+1} + \theta_{k+1}$
4: Calculate the innovation (or residual) vector: $r_{k+1} = z_{k+1} - Hx_{k+1|k}$
5: Calculate the innovation covariance matrix: $S_{k+1} = HP_{k+1|k}H^T + R_{k+1}$
6: Calculate Kalman filter gain matrix: $K_{k+1} = P_{k+1|k}H^T S_{k+1}^{-1}$
7: Update cycle:
$x_{k+1|k+1} = x_{k+1|k} + K_{k+1}r_{k+1}$; State estimation
$P_{k+1|k+1} = (I - K_{k+1}H)P_{k+1|k}$; Error covariance
8: Time-step is updated $k = (k + 1)$.
9: Return to step (2).

perspectives. Substitute Equation (2.64) into (2.51) would result in

$$\begin{aligned} P_{k+1|k} &= A\big(P_{k|k-1} - P_{k|k-1}H^T(HP_{k|k-1}H^T + R_k)^{-1}HP_{k|k-1}\big)A^T + Q_k \\ &= AP_{k|k-1}A^T - AP_{k|k-1}H^T(HP_{k|k-1}H^T + R_k)^{-1}HP_{k|k-1}A^T + Q_k \end{aligned} \tag{2.67}$$

This equation is known as Discrete Algebraic Riccati Equation (DARE). It will be analysed for conventional Kalman filtering and for data loss case to explore the conditions requiring for asymptotic stability of the designed filter.

2.4.4 Numerical Example for A Discrete-time LTI System

In this subsection a simple discrete time LTI system adopted from [98] is simulated to estimate state and covariance using standard discrete time Kalman filtering scheme, discussed in the previous section. Several characteristics such as estimated error, Kalman filter gain and residual vector are investigated to show the performance of standard KF. A simple LTI system with

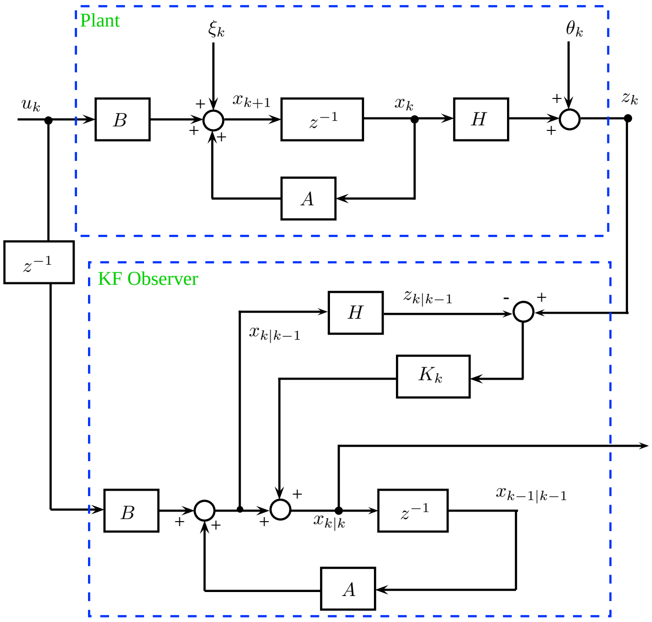

Figure 2.3: Structure of basic discrete time Kalman filter.

dynamics is as follows:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + \xi_k \\ z_k &= Hx_k + \theta_k \end{aligned} \tag{2.68}$$

where ξ_k and θ_k are assumed to be of zero mean, white noise and uncorrelated with any other signal or itself at other time step. The aim is to model a vehicle (UAV) running on a straight line. Two states are associated with this vehicle: its position p and velocity v. For simplicity, the input u of the system is assumed as the commanded acceleration and only one state is observed (measurement) at the output which is position p of the vehicle. If the acceleration is variable, then the velocity can be obtained by equation

$$\begin{aligned} u_k &= \frac{v_{k+1} - v_k}{dt} \\ v_{k+1} &= v_k + dt u_k \end{aligned} \tag{2.69}$$

Sampling frequency is assumed to be 10 Hz, hence the discrete time step $dt = 0.1$s. The complete state equation is constructed as

$$\begin{aligned} \begin{bmatrix} p_{k+1} \\ v_{k+1} \end{bmatrix} &= \begin{bmatrix} 1 & dt \\ 0 & 1 \end{bmatrix} \begin{bmatrix} p_k \\ v_k \end{bmatrix} + \begin{bmatrix} \frac{dt^2}{2} \\ dt \end{bmatrix} u_k + \begin{bmatrix} \tilde{p}_k \\ \tilde{v}_k \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0.1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} p_k \\ v_k \end{bmatrix} + \begin{bmatrix} 0.005 \\ 0.1 \end{bmatrix} u_k + \begin{bmatrix} \tilde{p}_k \\ \tilde{v}_k \end{bmatrix} \end{aligned} \tag{2.70}$$

where $x_k = \begin{bmatrix} p_k \\ v_k \end{bmatrix}$ is the state vector and $\xi_k = \begin{bmatrix} \tilde{p}_k \\ \tilde{v}_k \end{bmatrix}$ is the process noise vector. The output equation is described as

$$z_k = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} p_k \\ v_k \end{bmatrix} + \theta_k \tag{2.71}$$

with θ_k as sensor noise element associated with the single observed state p_k. The plant disturbance and sensor noise dynamics are characterised as

$$\begin{aligned} &E\left[\xi_k\right] = 0, \;\; E\left[\xi_k \xi_l^T\right] = Q\delta_{kl}, \;\; Q = I_{2\times 2} \\ &E\left[\theta_k\right] = 0, \;\; E\left[\theta_k \theta_l^T\right] = 100\delta_{kl} \end{aligned} \tag{2.72}$$

The output sensor noise level is assumed to be quite high relative to process noise. This is to highlight the efficiency of KF – the results obtained indicate the level of efficiency of KF in relation to how good a state estimation it is in comparison with noise contaminated/corrupted observations.

2.4.5 Simulation Results

The above mentioned system is simulated according to Algorithm 1. Some of the candidate results are shown and discussed below.

Figure 2.4 shows the performance of standard KF in the index of two estimated states – the position of the vehicle and its speed. It can be observed that, comparative to the actual noisy measurements, the estimated states (or filtered response) intelligently track the original states.

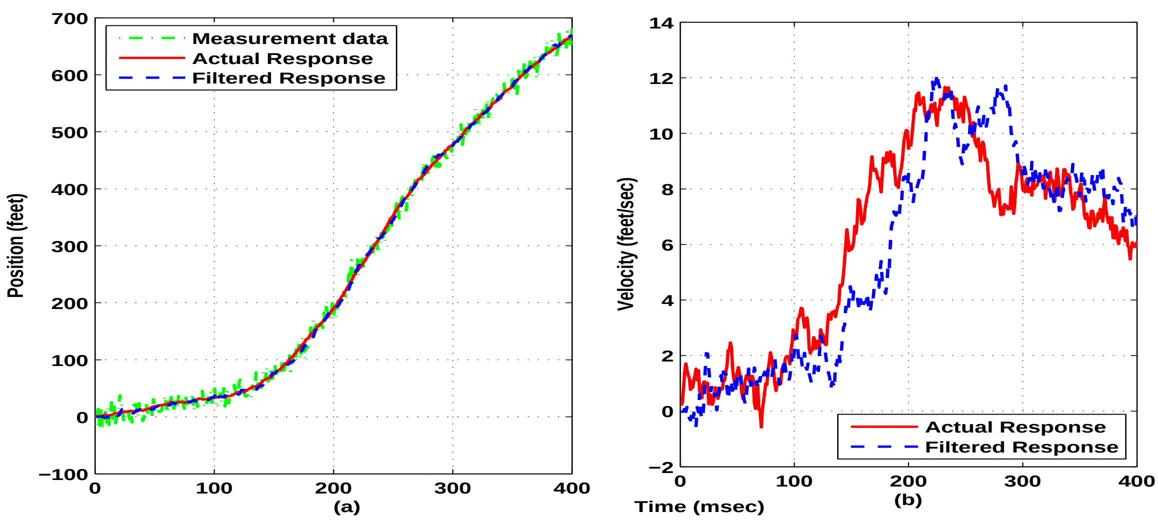

Figure 2.4: Estimations of (a) state 1 – $\hat{x}_1$ and (b) state 2 – $\hat{x}_2$.

This well-grounded realisation of KF results in errors in the two estimated states, which can be viewed in Figure 2.5.

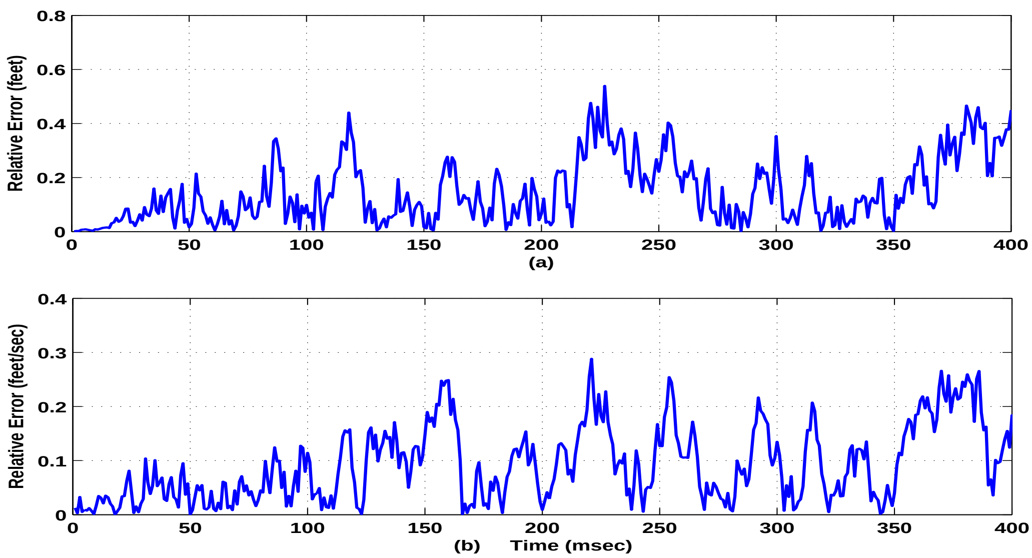

Figure 2.5: Absolute errors for the estimated (a) state 1 – $\hat{x}_1$ and (b) state 2 – $\hat{x}_2$.

Other indices which offer interesting results are residual vector (here a scalar) and output noise, which are shown in Figure 2.6. The more the two graphs resemble each other, the better estimation is, i.e. in a perfect state estimation, the residual measurement value would yield nothing but the output noise. Therefore, these parameters need to be explored in order to gain an insight to KF's performance with respect to the quality of state estimation.

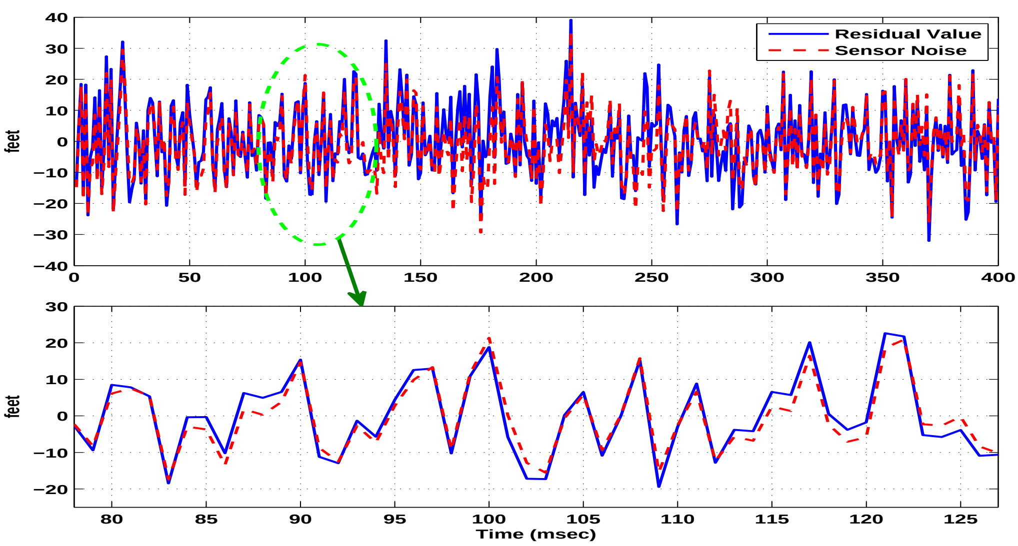

Figure 2.6: Residual observation signal and sensor noise.

In order to measure the time constraint for steady state value, the trajectory of Kalman filter gain elements and its predicted and updated error covariance matrices need to be analysed. From Equations (2.63) and (2.64), it can be concluded that the two parameters follow the same pattern as illustrated in Figure 2.7.

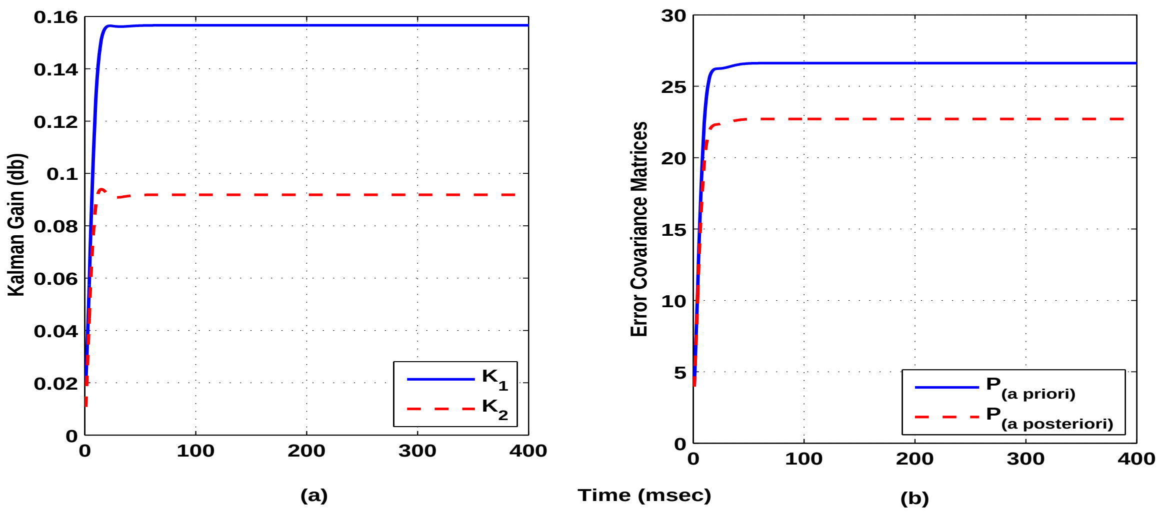

Figure 2.7: (a) Kalman filter gain components and (b) Traces of Predicted (*a priori*) & Updated (*a posteriori*) error covariance matrices.

2.5 Asymptotic Stability Associated with Kalman filtering

It is important to note that stability is not guaranteed for every optimal design of Kalman filter. *The design to be an optimal and the error covariance P to be bounded, the key requirements are complete observability, controllability and symmetric, bounded Q, R and bounded A matrices* [4]. From both practical and theocratical perspectives, KF design is required to be stable. Recalling the discrete ARE (2.67),

$$P_{k+1|k} = AP_{k|k-1}A^T - AP_{k|k-1}H^T(HP_{k|k-1}H^T + R_k)^{-1}HP_{k|k-1}A^T + Q_k \tag{2.73}$$

If ARE has a constant solution say $\bar{P}$ for constant Q and R matrices, then the designed filter is time invariant, i.e.

$$\bar{P} = A\bar{P}A^T - A\bar{P}H^T(H\bar{P}H^T + R)^{-1}H\bar{P}A^T + Q \tag{2.74}$$

with frozen Kalman filter gain as

$$K = \bar{P}H^T(H\bar{P}H^T + R)^{-1} \tag{2.75}$$

Equations (2.74) and (2.75) are sometimes known as steady state Riccati equation and steady state Kalman filter gain. Similarly, if a system under consideration is an unforced discrete time linear system, the following result could be achieved:

$$x_{k+1|k+1} = [A - KH]x_{k+1|k} + Kz_{k+1} \tag{2.76}$$

if $(A-KH)$ fulfill $|\lambda_j(A-KH)| < 1$ where λ_j are the eigenvalues of closed loop matrix $(A-KH)$, then the design filter is asymptotically stable [4].

Two reasons behind the drive of a time-invariant design and asymptotically stable filter are:

1. provided the designed filter is time-invariant and asymptotically stable, for any selection of initial condition on error covariance matrix (i.e. for any nonnegative symmetric initial condition P_0), the result $\lim_{k\to\infty} P_{k+1|k} = \bar{P}$ can be easily obtained [4].
2. with the conditions of time-invariant and asymptotic stability, the filter can be easily replaced by a steady state Kalman filter for a little performance compromise but reducing considerable computational efforts and memory requirements. The reason being, in steady state KF, the error covariance matrix or Kalman filter gain is not computed in real time, rather they are computed off-line [98].

Theorem 1 *[4]*
Consider a time invariant asymptotically stable designed KF for the system described in (2.40) and (2.43) with initial conditions $m := E[x_0]$ and $P_0 := [(x_0 - x_{0|0})(x_0 - x_{0|0})^T]$, then the estimated state $\hat{x} \to 0$ and algebraic Riccati equation (2.67) converges to a constant solution $\bar{P}$ as time step $k \to \infty$.

Proof

Since the *a priori* state estimate at time step $k-1$ is

$$x_{k|k-1} = AE[x_k] \tag{2.77}$$

Using chain rule it can be written as

$$x_{k|k-1} = A^k E[x_0] = A^k m \tag{2.78}$$

provided $|\lambda_i(A)| < 1$ – due to asymptotic stability, $\hat{x}_k = x_{k|k-1} \to 0$ as time step $k \to \infty$. The second part of the theorem can be deduced in the similar way. The *a priori* error covariance at time step $k-1$ would result in

$$\begin{aligned} P_{k|k-1} &= E[e_{k|k-1}e_{k|k-1}^T] \\ &= A^k P_0 (A^k)^T + \sum_{j=0}^{k-1} A^k Q (A^k)^T \end{aligned} \tag{2.79}$$

Hence with the condition $k \to \infty$, $P_k \to \bar{P} = \sum_{j=0}^{k-1} A^{k-1} Q (A^{k-1})^T$, which completes the proof. The details of such analysis can be found in several text books e.g. [4,21,98] *etc.* Similarly, the unique solution, associated with the ARE in Kalman filtering can be found through the following theorem.

Theorem 2 *The ARE has unique positive semidefinite solution P_∞ if and only if the following conditions are held:*

1. *(A,H) is detectable*
2. *(A-KH,G) is stabilisable* [3]

Proof :
The proof can be seen in the first part of Theorem 23, [98].

2.6 Summary

State estimation plays an important role in most of the areas of engineering and science and, without doubt, is vital in many physical systems where direct measurement of a system's state is either not available or available but corrupted by some random variable such as noise. For this reason, it was considered useful to present a few characteristics of random variable. This chapter starts with standard definitions related to a random variable. Thereafter, the basic notion of an unbiased state estimator for a discrete time system has been discussed. Based on these definitions and the general unbiased state estimator, a thorough description of Kalman filtering is presented. Attempts have been made to present all related theory associated with Kalman filter. Attention has been paid to describe the necessary conditions, required for the asymptotic stability of the ARE related to Kalman filter. The effectiveness of Kalman filter is shown by simulating a simple discrete time system.

During the description of Kalman filter, it is seen that Kalman filter depends heavily on certain parameters including the plant output data or measurements. In the events where the output data is subjected to random loss, the effectiveness of Kalman filter needs to be investigated. The

[3] $GG^T = Q$

next chapter is aimed at describing the performance of Kalman filter is such situations. Various diagnosing techniques, existing remedies to handle loss of observation in Kalman filtering and their associated properties are intended to be discussed in the following chapter.

Chapter 3

Kalman Filtering with Imperfect Data

3.1 Problem of Imperfect Data

This chapter is aimed at the analysis of state estimation through KF algorithm in distributed control system settings where different components of a control system communicate over a wireless network. A common phenomenon can be seen in Fig. 3.1, where plant, sensors and controllers are physically located at different locations and require a communication network to exchange (or transfer) critical information for system control [88]. Due to its fully mobile operation, fast deployment and flexible installation, wireless sensor networks (WSN) plays a vital role in distributed control applications. However, its use is restricted due to inevitable dilemma of data loss caused by limited spectrum of a channel, time varying channel gains, interference, congestions, limited bandwidth of buffer registers, collision, transmission errors, and many other deficits.

As mentioned in the previous chapter, Kalman filter (KF) depends on system dynamics, information of the unknown noise signals and timely received measurement signals. Hence any absence of the above may lead to adverse conditions and risking failure to the estimation tech-

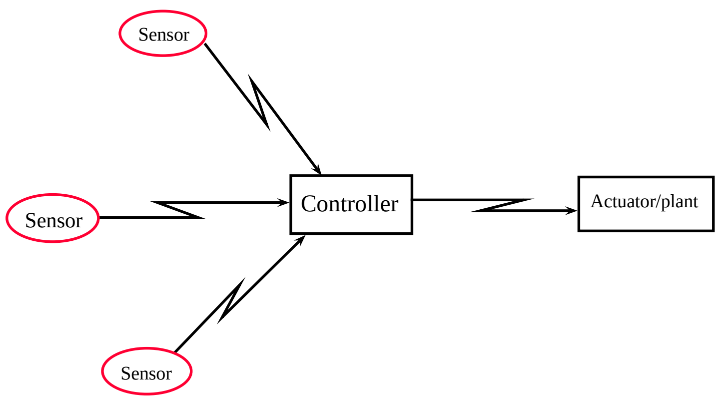

Figure 3.1: Wireless sensor network (Figure adopted from [88]).

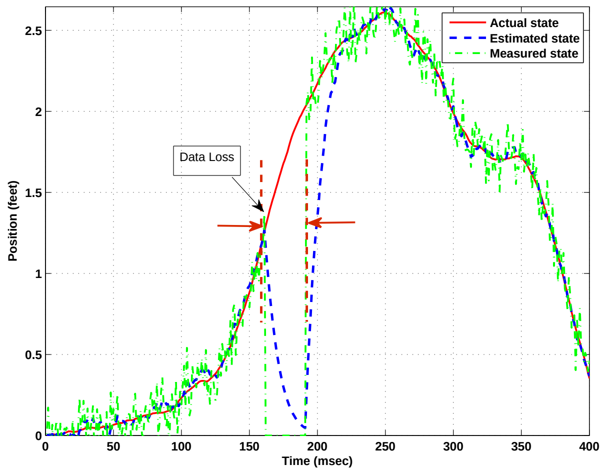

Figure 3.2: State – 1 estimation through conventional Kalman filter with data loss between (160-190) ms.

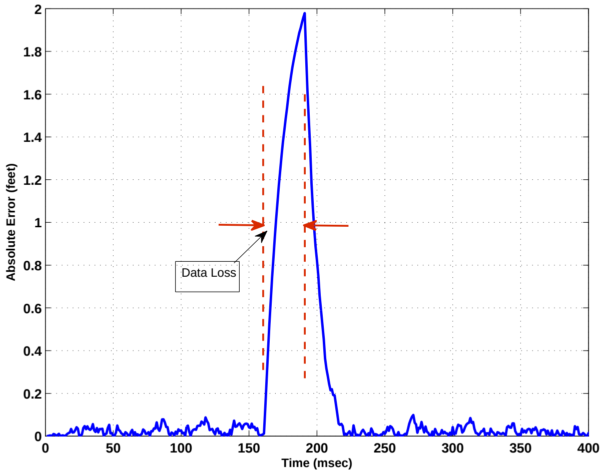

Figure 3.3: Absolute error in estimated state – 1 using conventional Kalman filtering with data loss between (160-190) ms.

nique. Among these, KF heavily relies on output measurement data – in fact it can be called an integral requirement for the complete and successful estimation. Figures 3.2 and 3.3 respectively show the state estimation and its corresponding error for the system discussed in Chapter 2, Section 2.4, subjected to a loss of observation of 30 ms. It is evident that estimation without any measurement will diverge swiftly if no cure is performed. Consequence of this could be severe if estimated states are to be used in a feedback control loop. Hence, emphasising KF's vulnerability to provide state estimation under such circumstances requires careful and non-trivial analysis.

One of the major contributors causing loss of data is the inefficient utilisation of communication bandwidth channel. For this reason, Goodwin [31] and his coworkers presented a generalised predictive control in order to minimise the bandwidth utilisation of a communication channel, which otherwise would lead to output data loss (packet drop). A communication constraint has been imposed to restrict all data transmitted in both links (from sensor to controller and from controller to actuator). In simple words, data is categorised in several quantised levels and one level of data is transmitted at a time. However, due to quantisation, the associated model becomes nonlinear, for which a suboptimal scheme based on nonlinear moving horizon optimisation is proposed.

3.2 Data Loss Detection

As mentioned previously, there are several factors which cause the output data loss. One of the major reasons behind the loss of observations is sensor faults and failures. Therefore, it should be interesting to state a few fault diagnosing techniques. In this regards, several researchers have contributed including [76, 77, 113]. Willsky *et al.* [113] have presented a thorough survey for abrupt changes (mainly failures) in stochastic systems such as failure-sensitive filters, voting systems, jump process formulations, innovations-based detection systems which may cause loss of observations. Another approach to detect data loss through fault diagnostic analysis is a *generalised observer scheme.*

3.2.1 Generalised Observer Scheme

A wide range of fault detection logics are available for data loss detection [77], [76]. For instant, Instrument Fault Detection (IFD) proposed in [16] and [17], sensor fault detection, Actuator Fault Detection (AFD) and Component Fault Detection (CFD) schemes are thoroughly discussed by Patton *et al.* [77]. The basic notion of an observer with output detection logic circuit is shown in Figures 3.4 and 3.5 [1].

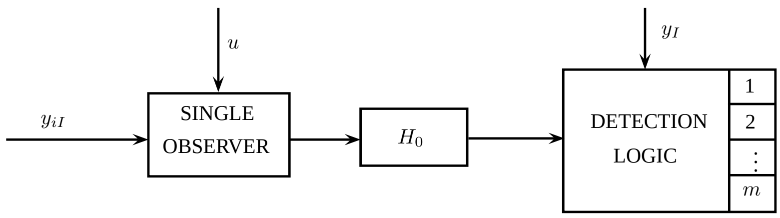

Figure 3.4: A simple observers scheme (Figure adopted from [77]).

In the IFD subsystems (Figure 3.5), there are m observers, one for each of the instruments to be monitored. Such designs are called Dedicated Observer Schemes (DOS) because each observer is specially designed to operate on a single instrument signal. Dedicated scheme is an important form of Generalised Observer Schemes (GOS) based on analytical redundancy technique. Such basic and simple design offers versatile features for generalisation. Importantly saying, the performance of DOS is limited by several factors such as [77]

[1] In the figures, UIFDO stands for Unknown Input Fault Detection Observer.

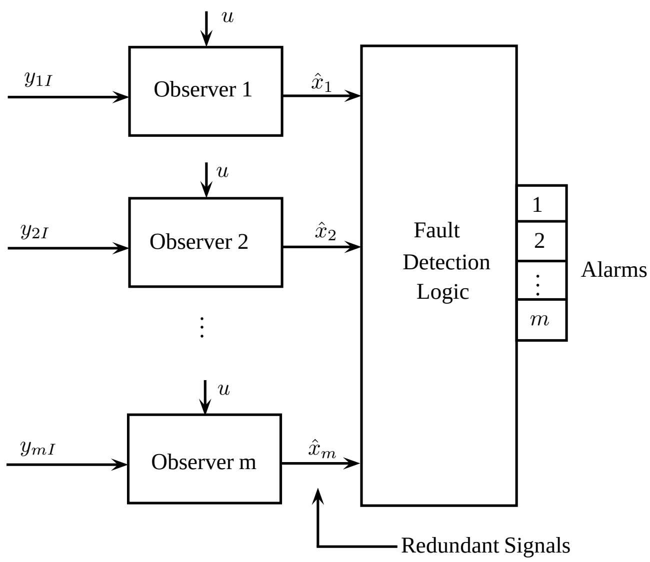

Figure 3.5: A dedicated observer scheme (Figure adopted from [77]).

1. uncertainty in the plant dynamics,
2. plant might have input commands other than simple step functions or harmonic oscillations. In such case, the plant may manoeuvre outside the limits for which the IFD system has been designed.
3. IFD schemes are based on theoretical assumptions which may be violated when implemented practically.

The Generalised Observer Scheme (GOS) has been successfully implemented for AFD, IFD and CFD [77]. For example, a single fault at a time, in one of the m sensors of a system is to be detected with the help of Unknown Input Observer Scheme (UIOS) [77]. This is achieved through Generalised Observer Schemes proposed by [27] and shown in Figures 3.6 and 3.7. In this scheme the i^{th} observer ($i = 1, 2, ...m$) is driven by all the measurement data except the i^{th} measurement. This is because, y_i is assumed to be imperfect due to sensor fault or failure and hence does not contain any information.

In some sensitive applications such as nuclear power plants, approaches to FDI based on the concept of analytical redundancy are limited for two reasons [58]:

- analytical redundancy based approaches are less familiar to engineers than traditional manual methods,
- analytical redundancy based approaches involve considerable expenditure in time and money (due to requiring many diagnostic and remedical actions) together with possible loss of power generation. In addition, false alarms and shutdowns would lead to costly consequences.

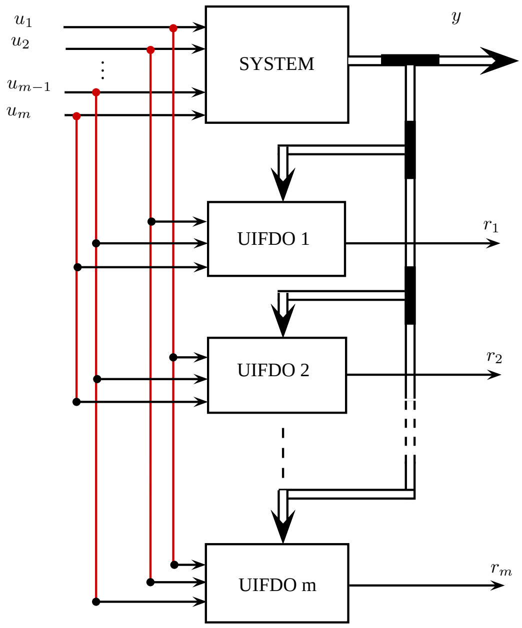

Figure 3.6: GOS for AFD (Figure adopted from [77]).

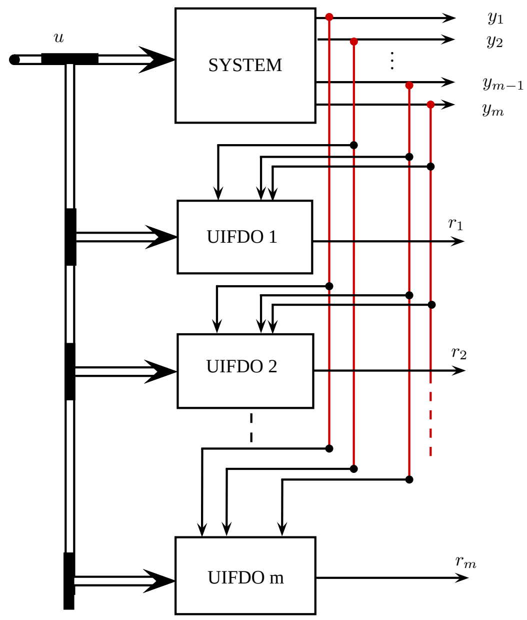

Figure 3.7: GOS for IFD (Figure adopted from [77]).

To avoid analytical redundancy techniques including GOS (e.g. a bank of Kalman filters [35], [60], [59], [117]), a single estimator (e.g. Kalman filter) is employed for output data loss (it is deliberately called as Loss Of OBservations or LOOB in this research) in the process of state estimation. State estimation under data loss scenarios has remained an interesting and challenging research topic during the last three decades, see [3, 6, 22, 42, 49, 88, 94, 99, 102, 116]. Researchers are taking keen interest in investigating the alternative consequences caused by LOOB in the process of Kalman filtering.

However, most of the researchers are interested in running Kalman filter in an Open-Loop fashion in the event of LOOB, as can be viewed from the above cited articles and references therein. KF predicts the states of the system only and no update is performed whenever output data is unavailable (packet loss time instant) due to any reason, i.e. sensitivity matrix $H = 0$, [32]. In other words, in the Open-Loop based estimation algorithm, prediction is referred to as *estimation*. More specifically, prediction is based on system model and processed as a state estimation without being updated by observations due to the unavailability of the observed data. Nonetheless, this approach may diverge in practice in the presence of sufficiently long data loss duration.

In order to grasp a clear idea, Open-Loop Kalman filter methodology, its associated properties and the drawbacks of employing Open-Loop Kalman filtering in the state estimation are presented in the following section. In addition, a brief literature review related to Open-Loop estimation and another technique known as Zero Order Hold are presented towards the end of this chapter.

3.3 Existing Solutions for Data Loss in Kalman Filtering

It was mentioned before that blind estimation through Kalman filter eventually leads to divergence in the event of data loss. Several approaches to handle LOOB have been proposed in various literature including this section. Discussions on the methodology, properties and associated drawbacks of the most effective technique known as the Open-Loop Estimation algorithm are to follow.

3.3.1 Open-Loop Kalman Filter

In literature, Open-Loop Kalman filtering scheme, or simply Open-Loop Estimation (OLE) is an estimation technique in which the measurement residual is ignored by forcing the Kalman filter gain matrix K_k to a zero matrix, i.e. no update step is performed. In other words in OLE algorithm, a prediction problem is identical to a filtering problem. In practice, OLE has been shown to be a straightforward and rapid approach to accommodate system missing data. A simple graphical representation of Open-Loop Kalman filtering is shown Fig 3.8. Whenever the output data (measurement or observation) is diagnosed to be lost due to any reason, the two switches (S_1 and S_2) behave open. In other words, no data arrives at filter, no Kalman filter gain is calculated and hence no update for state and covariance is performed. Hence only prediction is performed as it is not possible to obtain observational update, due to loss of measurement.

In OLE scheme, the five fundamental equations borrowed from the conventional Kalman filtering procedure, discussed in Section 2.4 are as follows:

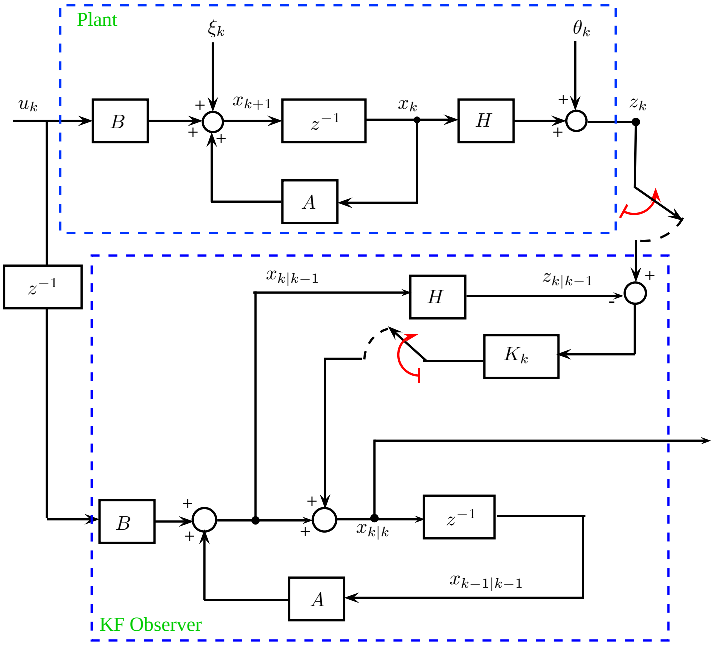

Figure 3.8: Open-Loop Kalman filter

The *a priori* step

$$_ox_{k+1|k} = A_o x_{k|k} + Bu_k \tag{3.1}$$

$$_oP_{k+1|k} = A_o P_{k|k} A^T + Q_k \tag{3.2}$$

The leading subscript 'o' denotes the Open-Loop Estimation scheme. It is important to note that predicted state and covariance matrix are not affected directly by the LOOB, rather they are influenced in the time update step.

The residual calculation step

The "compensated measurement" is represented as [2]

$$_oz_{k+1} = H_o x_{k+1|k}$$

Kalman filter gain

This will cause the residual vector to be zero (2-norm) and consequently [3]

$$_oK_{k+1} = 0 \tag{3.3}$$

[2]The actual observation vector is modelled as $z_{k+1} = Hx_{k+1} + v_{k+1}$
[3]i.e. Kalman filter gain matrix is a zero matrix.

The *a posteriori* step

It means no correction can be made to the prediction step quantities (state and error covariance). From Equations (2.58) and (2.64), it can be concluded that

$$ {}_o x_{k+1|k+1} = {}_o x_{k+1|k} \tag{3.4}$$

$$ {}_o P_{k+1|k+1} = {}_o P_{k+1|k} \tag{3.5}$$

Clearly, in the OLKF scheme, the *a posteriori* state estimation and error covariance strictly follow the *a priori* state estimation and error covariance, respectively.

3.3.2 Properties of the Open-Loop Kalman Filter

In this section, two important features related to the Open-Loop Kalman filtering technique are discussed.

First, Kalman filter gain (K_k) indicates the weight, which is integrated into the predicted state estimation and covariance matrix required at time step k. The more the predicted observation vector ($\hat{z}_k = H\hat{x}_{k|k-1}$) agrees with that of the actual observation vector ($z_k = Hx_k + v_k$), the more credence should be attributed to the predicted step (both predicted state estimation and predicted state covariance). In contrary, the less prediction step conforms to the actual state, a little recommendation should be assigned to the predicted step. Towards this end, in OLKF approach, since the predicted observations are virtually taken as actual measurements, no correction is made and hence the Kalman filter gain is kept zero during the loss of measurement period.

The second feature is that, if the sensor noise covariance R, is large enough, this would cause the Kalman filter gain to be smaller to emphasise that Kalman filter could rely less on the measurements. Therefore, in those situations where a huge measurement noise is imposed into the system, this approach could be a suitable option for measurement update step. These features can be observed from Equations (2.58),(2.63), and (2.64). Therefore the update step quantities follow the predicted step quantities;

$$\begin{aligned} {}_o x_{k+1|k+1} &\leftarrow {}_o x_{k+1|k} \\ {}_o P_{k+1|k+1} &\leftarrow {}_o P_{k+1|k} \end{aligned} \tag{3.6}$$

Therefore, it could be concluded that in cases where (a) estimated observation vector is almost equal to actual measurements or (b) sensor noise intensity is sufficiently large,

$$ {}_o K_k \approx 0 \tag{3.7}$$

In simple words, Kalman filter gain calculation step is ignored. Hence

$$
\begin{aligned}
{}_o x_{k+1|k+1} &= {}_o x_{k+1|k} \\
{}_o P_{k+1|k+1} &= {}_o P_{k+1|k}
\end{aligned}
\tag{3.8}
$$

In addition, there are numerous properties associated with OLKF technique which have been explored by many researchers which are briefly discussed in the following subsection.

3.3.3 A Literature Review for Open-Loop Kalman Filter

It is neither possible nor desirable to comprehend all the related works with the Open-Loop estimation used for loss of observations in control, communication and wireless networked systems. Therefore, in this section a very brief summary associated with data loss in the Kalman filtering is presented. Many researchers have employed and investigated this technique in various fashions since four decades and the research is still in progress.

Perhaps, Nahi [73] was the first who explored and discussed the insufficient data in the process of filtering in late 60's by elaborating a very common scenario. For a generalised state estimation baed on minimum mean error, he discussed two cases; the first one can be called as randomly single sample data loss and the other as a complete sequence of data loss. i.e.

- Case 1:

$$z_k = \gamma_k H x_k + \theta_k \tag{3.9}$$

  where 'k' is a single time step, and

- Case 2:

$$z_k = \gamma_k H x_k + \theta_k \qquad \forall \qquad k = 1, 2, ... \tag{3.10}$$

  where 'k' is a time sequence. The random variable γ_k is defined as

$$\gamma_k := \begin{cases} 1; & \text{with probability } p_k \\ 0; & \text{with probability } 1 - p_k \end{cases} \tag{3.11}$$

For both of these cases, minimum mean square estimators are computed.

Observations from different sensors are treated collectively in [99] and [92], in contrast to Liu *et al.* [68], where different sensor readings are treated individually. In simple words observations may be fully received, partially received or fully lost. In both of these works, lower and upper bounds of the threshold values with respect to the loss, have been provided. Figure 3.9 and subsequent equations easily describe the problem considered in [68].

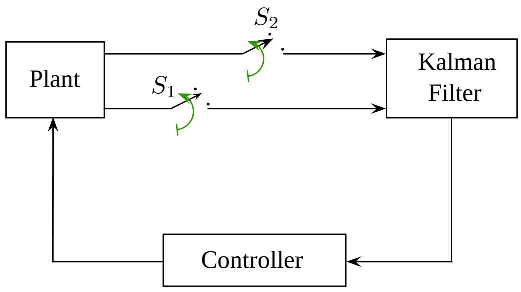

Figure 3.9: Partial observation loss

$$z_{i,k} = \gamma_{i,k} H_{i,k} x_k + \theta_{i,k} \qquad \forall \qquad k = 1, 2, ... \tag{3.12}$$

where the sensor noise is characterised as

$$p(\theta_{i,k}) \sim \begin{cases} \mathcal{N}(0, R_{ii}) & \text{if } \gamma_{i,k}{=}1 \\ \mathcal{N}(0, \sigma_i^2 I) & \text{if } \gamma_{i,k}{=}0 \end{cases} \tag{3.13}$$

with $\sigma \to \infty$. The work in [68] is similar to [99], where various threshold values are theoretically determined for different cases.

Another interesting concept revealed by some researchers is the so-called sojourn times (denoted by α^* and β^*) in the event of data loss e.g. [42], [41], [116], [114] *etc.* Sojourn times are the consecutive time sequences during which the data is available and then is not. They are described as follows:

$$\begin{aligned} \alpha_1 &= inf\{k : k \geq 1, \gamma_k = 0\} & \beta_1 &= inf\{k : k \geq \alpha_1, \gamma_k = 1\}, \\ \alpha_2 &= inf\{k : k \geq \beta_1, \gamma_k = 0\} & \beta_2 &= inf\{k : k \geq \alpha_2, \gamma_k = 1\}, \\ &\vdots \\ \alpha_k &= inf\{k : k \geq \beta_{k-1}, \gamma_k = 0\} & \beta_k &= inf\{k : k \geq \alpha_k, \gamma_k = 1\} \end{aligned} \tag{3.14}$$

The sojourn times are defined as

$$\alpha_k^* = \alpha_k - \beta_{k-1} \qquad \& \qquad \beta_k^* = \beta_k - \alpha_k \tag{3.15}$$

The concept of sojourn times is shown in Figure 3.10 where $l_1 \geq 1$ and $l_2 \geq 1$. In [41] and [42], upper bound for the peak error covariance matrix is investigated in terms of these two time indices for an unstable scalar model.

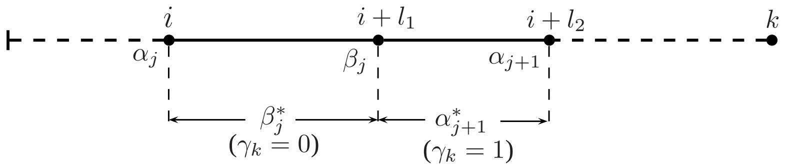

Figure 3.10: Sojourn times

Micheli in [70] has considered a time delay in the data arrival which may also be translated as lost or inaccurate measured data. In addition, he has consider random sampling for the data loss which is a resembling technique to the ZOH technique discussed below. In [88] and [89], the system is assumed to be subjected to both LOOB and delay of observation at the same time. All these works have suggested switching to an Open-Loop estimator when there is a LOOB and back to the Closed-Loop estimator when observation is arrived at the measurement channels. This consequently aims at designing an estimator which is strongly time-varying and stochastic in nature.

In [99], the authors have discussed the convergence of ARE in the event of loss of data for Open-Loop Kalman filtering. The authors have discussed stability effects of the state estimation and have shown a threshold limit for data loss, above which the expected value of error covariance becomes unbounded as the time goes to infinity. They have considered the arrival of observation data as time independent i.e. for the whose simulation period γ_k is independent of time. On the other hand, authors in [92] have considered somewhat similar study but from different prospectives; by defining an information gain $\pi_g = \frac{No.\ \ of\ \ data\ \ packets\ \ recieved}{No.\ \ of\ \ step}$. A Kalman filter estimator has been designed which has been found stable theoretically, for any positive value of π_g. Some other research work will be discussed shortly, after highlighting a few notable shortcomings of the OLKF scheme.

Besides these characteristic properties, there are a few drawbacks associated with Open-Loop Kalman filtering technique which are mentioned in the following subsection.

3.3.4 Shortcomings of the Open-Loop Kalman Filter

Despite of fast estimation technique and simpler structure, there are a few drawbacks associated with the OLKF approach in the event of loss of observation which are briefly discussed below.

1. In practice, Open-Loop Kalman filtering (or OLE) scheme may diverge in the presence of adequate data loss. In other words, the error covariance is likely to exceed the error limits

very quickly, if such upper and lower bounds of error covariance are provided [68].[4]

2. Another shortcoming of the Open-Loop estimation technique is that when the system measurement (observation) is resumed after the loss period, sharp spikes phenomenon and/or oscillation in the estimated parameters can be observed. This is because Open-Loop Kalman filter gain is set to zero when data loss occurs. But, when observation is repossessed (after the recovery of fault/failure, for example), the filter gain surges to a high value (–quite higher than nominal value), oscillates and then attains the steady state value in order to compensate the impact of data loss [54]. This, consequently resulting in sudden peaks to reach the normal trajectories of the estimated parameters, which is an undesirable behaviour in the state estimation process.

3. Another important shortcoming associated with OLKF approach is that its steady state values (state and covariance) after the loss is recovered are not regained. It takes theoretically infinite time to retrieve the steady state values. The reason behind this deficit is the recursive behaviour of KF, i.e. the predicted quantities are employed at the filtering step and the filtering quantities are adopted at the prediction step. Therefore, once at any step efficiency of estimation is decreased, its effect will be transferred at the next step and so on.

3.3.5 Zero Order Hold Technique

Due to several shortcomings, most recently, some researchers [19,22,24,94,95] (and the references therein) have attempted to avoid the Open-Loop estimation by some techniques. One of the major differences of such recent techniques is that, the researchers have considered the loss of observations as well as loss of inputs. This is similar to a combined sensor and actuator faults. In [22], [94] and [24] the authors have reconstructed the loss input and output signal by auto-regressive models. The main idea behind this technique is shown in Figure 3.11. However there are certain limitation associated with this technique too, i.e.

- the most recent data sample of the signal, needs to be stored during the whole estimation process[5],
- constantly employing one measurement sample at measurement update may not provide an optimal solution for an adequate data (input and output) loss,
- this technique requires strict correlation among the signal data samples [95], [24],
- this technique results in random sampling [70],

[4]For the critical limits of the LOOB in OLE scheme, which will provide bounded error covariance matrices, authors in [99] have provided a detailed discussion.

[5]The reason this point is declared as a drawback is that, in Kalman filtering, once any data sample is utilised at measurement update step, it is discarded. No storing of data is required in the process of Kalman filtering.

- if the input signal changes during the data loss time, this may further invalidate the scheme.

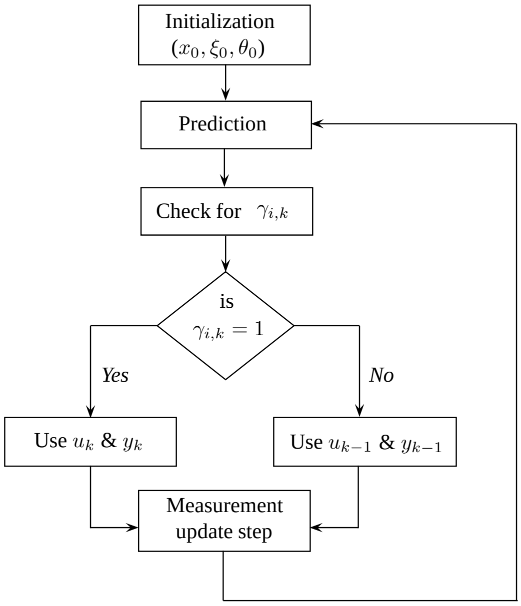

Figure 3.11: Zero Order Hold estimation

The method described in the above references is quite simple, and is known as Zero Order Hold (ZOH) technique. In this technique, the last data sample (both for input and output) is required to be stored and updated throughout the operation. This is summarised in the flow-chart diagram 3.11. In the diagram, the random variables $\gamma_{i,k} := \{\gamma_{1,k}, \gamma_{2,k}\}$ are for input data sample and for output data sample.

In the process of compensated optimal state estimation, however, the ZOH technique is briefly introduced in [55] besides other simple techniques in the event of loss of observation data. It is necessary to mention that, employing ZOH technique drives the time varying Kalman filter to a steady state Kalman filter during the data loss time. This is because, using the last sensor readings would result in the the last residual, and hence the last covariance and Kalman filter gain matrices at measurement update step. In order to avoid random sampling and stochastic behaviour of the designed KF, involved in ZOH technique as discussed in [70], a number of approaches have been proposed in [55] to compensate loss of observations in the process of state estimation. These approaches are indirectly linked to linear prediction theory, for which a detailed discussion will be provided in the following chapter.

3.4 Summary

In this chapter, the problem of missing measurements has been considered in the process of Kalman filtering. Several factors causing the adverse condition of data loss are reported in the literature. Among these, sensor faults/failures and limited bandwidth of a communication channel are notable and therefore briefly discussed in this chapter. Generalised observer scheme based on analytical redundancy technique, usually employed to detect instrument fault is cited for interesting readers.

The findings being that Kalman filter heavily depends on measurement data, which has led it to an improper solution for state estimation in situations of missing or imperfect data. Researchers have attempted to handle state estimation in such events of loss of observations by proposing a few techniques. The most distinguished method, known as Open-Loop Kalman filtering, predicts during the loss of observation time period and processes the predicted parameters without any update. Though this technique is quite simple and fast but is found an unhealthy solution for a reasonable data loss duration. Bearing in mind the design structure and related shortcomings of Open-Loop Kalman filtering algorithm, another solution named Zero Order Hold (ZOH) has been offered by some researchers to deal with the LOOB. Though ZOH technique has overcome majority of the limitations associated with OLKF technique, the method suffers from other drawbacks. Therefore, it is an essential requirement for more accurate and reliable estimation algorithms in the event of LOOB.

How much of the missing measurements can be reproduced through some techniques and how much Kalman filter can sustain with this reproduced data is a very challenging problem. In the following chapter, a novel state estimation technique based on Kalman filtering is presented, where observations are not available for measurement update step. The missing data is thereby generated through linear prediction theory. In other words, in doing so the two concepts of Kalman filtering and linear prediction for the intermittent observations will be combined.

Chapter 4

Design of Compensated Kalman Filters

4.1 Introduction

Kalman filter (KF) is an important tool of retrieving information (state) from noisy measurements when direct access to the system's information is either not possible or accessible but of no use. KF predicts the states of a system (*a priori* step) and thereafter updates those states with the help of noisy measurements (*a posteriori* step). This is called one complete iteration. However, in the absence of noisy measurements, it would not be possible to generate an optimal estimation of the system's states. To overcome this issue, the prevailed scheme known as Open-Loop Kalman filtering predicts the system's state and processes it without any update to the next iteration. However this scheme cannot produce a bounded estimation error when the observations are lost for an adequate time period. A few other flaws associated with Open-Loop Kalman filtering have been briefly discussed in the previous chapter. Consequently urgent calls are required to propose for some compensated estimation approaches with the aim of producing acceptable state estimation with bounded error in the presence of reasonable data loss [3].

This chapter plans to propose those estimation approaches to tackle loss of observations efficiently. Keeping in mind the basic structure of discrete time KF, a straightforward procedure has been taken place to design compensated Kalman filtering approaches. In the event of loss of observations (LOOB), observation signal is reconstructed through linear prediction schemes. Therefore, it is considered important and necessary to discuss linear prediction theory.

This chapter is organised as follows: Section 4.2 is dedicated to a brief discussion on linear prediction (LP) theory, its related categories, and various possible modes involved in the linear prediction process. The routine methods of linear prediction could not decide the number of previous data samples. For this reason, the current methods of LP are modified by providing an upper bound on the number of data samples employed in Section 4.3. In this section, the lossy observation signal is reconstructed through external linear prediction technique. Two straightforward algorithms are proposed to decide the optimal value of order of the linear prediction filter. Based on different reconstructed signals, compensated Kalman filtering schemes are designed in Section 4.4. The complete structure of Kalman filter in the event of LOOB and the compensated signal at measurement update step are analysed at the time of occurrence of LOOB and after the observation is resumed. Some features related to those combinations are presented too in order to highlight their suitability.

4.2 Linear Prediction Theory

Linear prediction (LP) is in fact a system identification process where signal is reconstructed from its previous signal samples [69]. In other words, linear prediction is a mathematical and optimisation tool for estimating the future values of a signal based on previous values (and sometimes inputs as well). The theory of linear prediction has been widely used in a variety of engineering applications [103]. Its diverse area of applications can be found in data forecasting, speech coding, video coding, speech recognition, signal restoration, model-based spectral analysis, model-based interpolation and impulse/step input detection [106]. The future values of a signal can be reconstructed by adopting the external linear prediction approach based on the following assumptions [14]:

1. the signal has correlation between its samples and
2. the signal's statistical properties vary slowly with time.

In the following subsection an overview of linear prediction theory, its key topologies and the models wherein LP is implemented will be briefly discussed.

4.2.1 Topologies of Linear Prediction schemes

Linear Prediction itself can be termed as a system identification problem, where the parameters of an auto-regressive series are estimated within the series itself [36], [84], i.e.

$$\bar{z}_k = \sum_{j=1}^{p} \alpha_j z_{k-j} + \sum_{i=0}^{N} b_i u_{k-i} \tag{4.1}$$

where $\bar{z}_k$ is the signal to be predicted through linear prediction scheme, α_j are the weights assigned to the previous observations (known as *Linear Prediction Coefficients* or LPC), b_j are the weights assigned to previous input signal values u_{k-j}. Linear prediction (LP) method can be done in forward and backward directions, sometimes called as External and Internal linear prediction, respectively. External and internal linear prediction techniques are shown in Figures 4.1 and 4.2.

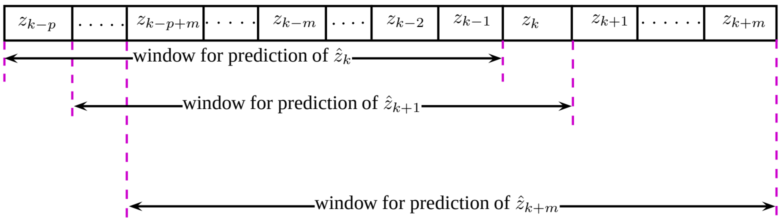

Figure 4.1: Selection of window for External Linear Prediction (for the case $m < p$).

In Figure 4.1, the data samples z_{k+l} where $l = \{1, 2, ...m\}$ are predicted through a window, consists of p previous data samples. In this case, the data is truly predicted in the future.

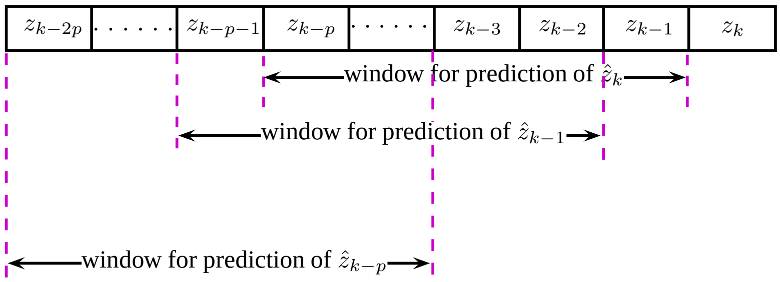

Figure 4.2: Selection of window for Internal Linear Prediction (for the case $m = p$).

In Figure 4.2, the data samples z_{k-l} where $l = \{1, 2, ...p\}$ are predicted using previous p data samples. This methods is in fact model identification problem [14]. Linear prediction can be performed through least square error (in case of deterministic process), mean square error and maximum a posteriori (MAP) (for random processes) methods [36, 106].

Figure 4.3 describes all possible methods related to linear prediction schemes.

Computing the weights or linear prediction coefficients through linear prediction (LP) schemes can be performed using various models [83]. These models can be derived based on different combinations of previous input and output data samples.

- **Auto-regressive (AR) Model**

  If $N = 0 \Rightarrow b_i = b_0$ then (4.1) reduces to

$$\bar{z}_k = \sum_{j=1}^{p} \alpha_j z_{k-j} + b_0 u_k \tag{4.2}$$

  This is called an All-Pole filter because its transfer function in z-plane would be

$$H(z) = \frac{b_0}{1 + \sum_{j=1}^{p} \alpha_j z^{-j}} \tag{4.3}$$

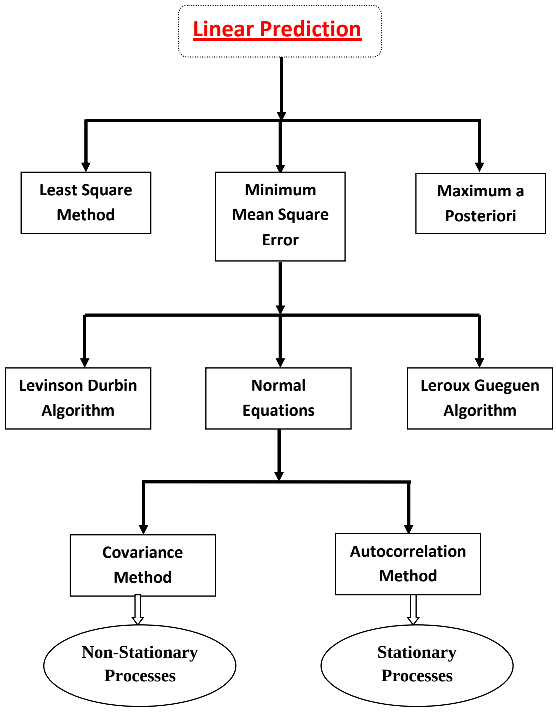

Figure 4.3: Linear prediction techniques.

- **Moving Average (MA) Model**
  When the prediction solely depends on input and no previous observation is utilised i.e. $p = 0 \Rightarrow \alpha_j = 0$, hence (4.1) reduces to

$$\bar{z}_k = \sum_{j=1}^{p} b_j u_{k-j} \tag{4.4}$$

  This is known as finite impulse response (FIR) model.

- **Auto-Regressive Moving Average (ARMA) Model**
  The ARMA model is formed where both input and observations are involved in the prediction process i.e. when both N and p have some real positive values as per illustrated below:

$$\bar{z}_k = \sum_{j=1}^{p} \alpha_j z_{k-j} + \sum_{j=1}^{N} b_j u_{k-j} \tag{4.5}$$

The AR model has received more attention than other models due to its

- simple structure,
- independency from input signal [75],
- convenience because some physical applications like speech processing (an acoustic tube model for speech production) can be easily modeled [36],

- speed and efficiency embedded in its algorithm in calculating linear prediction coefficients (LPC) [36],

For these reasons, AR model is considered in this thesis. Though there are certain limitations while employing AR model in linear prediction which will be highlighted in the subsequent sections. The necessary and sufficient details of computing LPC in AR-mode, would be discussed in following subsection.

4.2.2 Autoregressive Model

Linear prediction coefficients (also known as linear predictive coding) of a random process requires a model called autoregressive (AR) model [104]. A random process is said to be autoregressive process if it can be generated using the following recursive equation:

$$z_t = c + \sum_{j=1}^{p} \alpha_j z_{t-j} + v_t \tag{4.6}$$

where $\alpha_j = [\alpha_1, \alpha_2, ...\alpha_p]$ are the parameters of the AR model, c is a constant, v_t is a noise signal and p is the window size – also known as Linear Prediction filter (LPF) order. If input to the system is a white noise signal, than the model is termed as an all-pole Infinite Impulse Response (IIR) filter. The constant c can be omitted for simplicity.

The optimal values of α_j are computed through auto-correlation function using Yule-Walker equation, which is

$$\gamma_m = \sum_{j=1}^{p} \alpha_j \gamma_{m-j} + \sigma^2 \delta_m \tag{4.7}$$

where δ_m is kroncker delta. It is characterised as follows:

$$\delta_m = \begin{cases} 1 & \text{if m} = 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.8}$$

Equation (4.7) is obtained by multiplying both sides of (4.6) by z_{t-m} and then taking the expectation. i.e.

$$E[z_t z_{t-m}] = E\left[\sum_{j=1}^{p} \alpha_j z_{t-j} z_{t-m}\right] + E\big[v_t z_{t-m}\big] \tag{4.9}$$

Equation (4.9) is in fact a set of p linear equations with p unknowns $\alpha_j = \{\alpha_1, \alpha_2, ...\alpha_p\}$ and is known as *Normal Equations*. Throughout this thesis solving normal equations by routine methodology would mean LP through autocorrelation method.

Conventionally, various ways have been adopted to solve the system of *Normal Equations*, in common use are the *covariance* formulation (details in Appendix C) which is an appropriate solution for a non-stationary process and the *autocorrelation* formulation which is used for a stationary process.

4.2.3 Linear Prediction Coefficients through Autocorrelation Method

Consider a random signal $z(k)$ is to be predicted, given its p past values; it is done as

$$\bar{z}(k) = \sum_{j=1}^{p} \alpha_j z(k-j) \tag{4.10}$$

The following measures are taken to find the optimal values of α_j. The error generated by this compensated observation vector would be

$$e(k) = z(k) - \bar{z}(k) \tag{4.11}$$

where $z(k)$ is the actual observation vector. The cost function consists of mean square prediction error as;

$$J = E[e^2(k)] = E\big[\{z(k) - \sum_{j=1}^{p} \alpha_j z(k-j)\}^2\big] \tag{4.12}$$

The optimal values of linear prediction coefficients (LPC) can be calculated by equating the partial derivatives of the cost function J with respect to α_i to zero, i.e.

$$\frac{\partial J}{\partial \alpha_i} = 2E\Big\{\big[z(k) - \sum_{j=1}^{p} \alpha_i z(k-j)\big] z(k-i)\Big\} = 0 \tag{4.13}$$

or

$$\begin{aligned}
E[z(k)z(k-i)] - E\big[z(k-i)\sum_{j=1}^{p}\alpha_i z(k-j)\big] &= 0 \\
E[z(k)z(k-i)] - \sum_{j=1}^{p}\alpha_i E\{z(k-i)z(k-j)\} &= 0 \\
E[z(k)z(k-i)] - \sum_{j=1}^{p}\alpha_i E\{z[(k+j)-i]z[(k+j)-j]\} &= 0 \qquad (\because k = k+j) \\
E[z(k)z(k-i)] - \sum_{j=1}^{p}\alpha_i E\{z[k-(i-j)]z(k)\} &= 0 \\
E[z(k)z(k-i)] - \sum_{j=1}^{p}\alpha_i E\big\{z(k)z[k-(i-j)]\big\} &= 0 \\
r_\gamma(i) - \sum_{j=1}^{p}\alpha_i R_z[i,j] &= 0 \\
\sum_{j=1}^{p}\alpha_i R_z[i,j] &= r_\gamma(i) \qquad (4.14)
\end{aligned}$$

or

$$R_z \cdot A_\alpha = r_\gamma \qquad (4.15)$$

The optimal values of the linear prediction coefficients can be calculated as

$$A_\alpha = R_z^{-1} r_\gamma \qquad (4.16)$$

where R_z is $p \times p$ Hermitian matrix of autocorrelations, and is composed of

$$R_z = \begin{bmatrix} R[0] & R[1] & \cdots & R[p-1] \\ R[1] & R[0] & \cdots & R[p-2] \\ \vdots & \vdots & \ddots & \cdots \\ R[p-1] & R[p-2] & \cdots & R[0] \end{bmatrix}, \qquad (4.17)$$

r_γ is the column vector as

$$r_\gamma = \begin{bmatrix} R[1] \\ R[2] \\ \vdots \\ R[p] \end{bmatrix} \qquad (4.18)$$

and

$$A_\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix} \tag{4.19}$$

is the LPC array. Generally fixed window is employed for calculating LPC in the linear prediction process. However sliding window is introduced in order to utilise more information in the signal reconstruction. By sliding window it means that, for every data sample to be reconstructed through LPC schemes, an independent window composed of p samples is chosen. This window is drifted in forward (External LP) or backward (Internal LP) directions as the number of data to be predicted, increases. With regards to the sliding window concept (shown in Figure 4.4) for the reconstruction of p data samples, a signal through the internal linear prediction scheme requires window size of $n = 2p$.

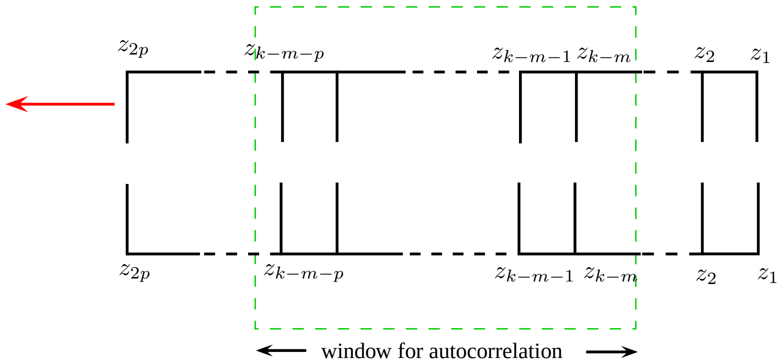

Figure 4.4: Sliding window concept for calculating m^{th} data sample through autocorrelation method.

4.2.4 Linear Prediction Coefficients through Auto-covariance Method

Stationary process is quite simple and less computational due to its time independency property. In real world applications, however non-stationary processes can be frequently encountered for which the statistical properties such as mean and covariance are time-dependent. In such cases, the linear prediction coefficients should be calculated through auto-covariance method instead of the usual autocorrelation method. In this work, main emphasis has been considered on stationary processes and therefore the details theory, concepts and calculations for non-stationary processes through auto-covariance method, are avoided in order to focus on the main course of the work. A brief summary of linear prediction coefficients through auto-covariance method is, however, included in Appendix C for the interested readers. The detailed description can be found in many text books like [14, 83, 84]. The final shape of the end results are as follows:

$$\underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix}}_{A_\alpha} = \underbrace{\begin{bmatrix} R[1,1] & R[1,2] & \cdots & R[1,p] \\ R[2,1] & R[2,2] & \cdots & R[2,p] \\ \vdots & \vdots & \ddots & \cdots \\ R[p,1] & R[p,2] & \cdots & R[p,p] \end{bmatrix}^{-1}}_{R_z} \underbrace{\begin{bmatrix} R[0,1] \\ R[0,2] \\ \vdots \\ R[0,p] \end{bmatrix}}_{r_\gamma} \tag{4.20}$$

where $R[m,n] := E\left[z_{k-m}^T z_{k-n}\right]$, A_α, R_z and r_γ in Equation 4.20, using auto-covariance method are derived in Appendix C.

Autocorrelation (or auto-covariance) matrix R_z has to be inverted at every step of calculation while computing LPC through Normal Equations. Computing LPC through autocorrelation (or auto-covariance) involves $O(p^3)$ multiplies and additions and inversion of $p \times p$ matrix [103]. Consequently, more computational efforts need to be undertaken for an adequate value of p (linear prediction filter order). In order to overcome this issue, Levinson-Durbin, Leorux-Gueguen and Schur algorithms are adopted resulting in much lesser computational time. Due to simplicity, Levinson-Durbin algorithm is selected to test the computational time for the same prediction quality. Utilising Levinson-Durbin method for calculating linear prediction coefficients reduces the computational effort from $O(p^3)$ to $O(p^2)$, [103].

4.2.5 Levinson-Durbin Algorithm

Levinson-Durbin algorithm (LDA) reduces the computational time by exploiting the Toeplitz symmetry property inherent in the autocorrelation matrix [83]. Due to its simple structure and straightforward computations, LDA is implemented in this thesis to reconstruct the lost information (missing measurements) and is integrated with the process of Kalman filtering. The conventional LDA is summarised below.

Levinson-Durbin algorithm is an alternative for matrix inversion found in Yule-Walker Equations (4.7). Similar to the *Normal Equations*, LDA is a recursive prediction method where parameters of an autoregressive series are predicted with less computational efforts. The basic idea of this algorithm is to calculate the solution of $(i+1)^{th}$ predictor from the i^{th} order predictor's solution. The process is repeated until some assigned upper limit has been reached. Each recursion solves the i^{th} pole-problem, finding the solution that minimises the mean-square error for each order predictor, by updating the lower order solution [103]. The LDA algorithm is described in the following steps:

1. Initialising the order of linear predictor filter as

$$\begin{aligned} \alpha_0^0 &= 0; \\ J^0 &= R[0] \end{aligned} \tag{4.21}$$

i.e. initial coefficients computed for zero-order of linear prediction filter are zero and the mean square prediction error is the first element of autocorrelation window i.e. R[0].

2. Hypothetical coefficients, also known as Reflection-Coefficients are computed as

$$\kappa_i = \frac{1}{J^{i-1}}\left[r(i) + \sum_{j=1}^{i-1} \alpha_j^{i-1} R[i-j]\right] \tag{4.22}$$

3. Using reflection coefficients, the updated linear prediction coefficients are calculated as

$$\begin{aligned} \alpha_i^i &= \kappa_i \\ \alpha_j^i &= \alpha_j^{i-1} - \kappa_i \alpha_{i-j}^{i-1} \end{aligned} \tag{4.23}$$

where $1 \leq j \leq i-1$, and the most recent reflection coefficient by i^{th} order predictor (κ_i) is utilised to update the coefficients computed through $(i-1)th$ order filter coefficients.

4. The minimum mean square error for the i^{th} predictor is updated as

$$J^i = (1 - \kappa_i^2) J^{i-1} \tag{4.24}$$

This step is exercised to monitor the accuracy of the prediction by calculating the mean square error.

5. Step-2 to step-4 are repeated for $i = 1, 2, ..., p$. At p^{th} step, the final version of the linear prediction coefficients is shown as follows:

$$\mathbf{a}^T = \alpha_j^p, \;\; \forall \;\; 1 \leq j \leq p \tag{4.25}$$

The transfer function based on these LPC is given by

$$H(z) = \frac{A}{1 - \sum_{j=1}^{p} \alpha_j z^{-1}} \tag{4.26}$$

The detailed theory associated with Levinson-Durbin algorithm has been presented in Appendix A. There are two important features associated with the Levinson-Durbin algorithm [103] which are stated below:

1. The mean square prediction error is always greater than zero due to the inability to compute for a perfect prediction i.e.

$$\begin{aligned} & J^i = (1-\kappa_i^2)J^{i-1} > 0 \\ \Rightarrow \quad & (1-\kappa_i^2) = J^i/J^{-1} > 0 \\ \Rightarrow \quad & \kappa_i^2 < 1 \\ \Rightarrow \quad & |\kappa_i| < 1 \end{aligned} \tag{4.27}$$

   i.e. the magnitude of reflection coefficients is less than unity [14].

2. From the first feature, ($|\kappa_i| < 1$), the poles of the transfer function (4.26) are inside the unit circle. Therefore, the roots of the p^{th} polynomial lie inside the unit circle and thus is minimum-phase [103].

4.3 Modified Linear Prediction Coefficients

In the routine methodologies of computing linear prediction coefficients or weights (wether auto-correlation or auto-covariance), there is no clear or distinct way to decide the number of previous measurement samples [14, 83]. These weights are emphasised due to the fact that they define the contributions of previous observations made for reconstructing the missing data – which can be viewed from the data autocorrelation standpoint [84]. From the fundamental structure of computing LPC, this is computationally costly issue and non-optimal. For this reason, certain constraints are introduced in this section, in order to compute LPC optimally by defining Linear Prediction filter Order (LPFO). The process of computing linear prediction coefficients (LPC) is summarised as follows:

The error generated by the proposed observation vector is

$$e_z(k) = \bar{z}_k - z_k \tag{4.28}$$

The residual based cost function is defined as

$$J_k := E\big[e_z(k)^T e_z(k)\big] \tag{4.29}$$

The optimal values of the LPC, α_j are computed by minimising the above cost function, J_k, i.e.

$$\begin{aligned}
\frac{\partial J_k}{\partial \alpha_i} &= \frac{\partial}{\partial \alpha_i} E\big[e_z(k)^T e_z(k)\big] \\
&= E\Big[\frac{\partial\big(e_z(k)^T e_z(k)\big)}{\partial \alpha_i}\Big] \\
&= E\Big[\frac{\partial J_k}{\partial \bar{z}_k}\frac{\partial \bar{z}_k}{\partial \alpha_i}\Big] \qquad (4.30)
\end{aligned}$$

Taking the derivatives of Equations (4.10) and (4.29) and substituting in Equation (4.30). i.e.

$$\frac{\partial J_k}{\partial \bar{z}_k} = 2E\big[e_z(k)^T\big] = 2E\big[\bar{z}_k - z_k\big]^T \qquad (4.31)$$

Differentiating $\bar{z}_k$ with respect to α_i will result in

$$\frac{\partial \bar{z}_k}{\partial \alpha_i} = z_{k-i} \qquad (4.32)$$

In order to minimise the cost function J_k, from Equation (4.30),

$$\frac{\partial J_k}{\partial \alpha_i} = 0 \qquad (4.33)$$

From Equation (4.30) through Equation (4.32),

$$\begin{aligned}
2E\Big[\big(\bar{z}_k - z_k\big)^T\Big] z_{k-i} &= 0 \\
E\Big[\bar{z}_k^T z_{k-i}\Big] - E\Big[z_k^T z_{k-i}\Big] &= 0
\end{aligned}$$

Substituting the value of $\bar{z}_k$ from Equation (4.10), will result in

$$\sum_{j=1}^{p} \alpha_j E\big[z_{k-j}^T z_{k-i}\big] = E[z_k^T z_{k-i}] \qquad (4.34)$$

which can be simplified as

$$R_z A_\alpha = r_\gamma \qquad (4.35)$$

or

$$A_\alpha = R_z^{-1} \cdot r_\gamma \qquad (4.36)$$

where

$$R_z = \begin{bmatrix} R[0] & R[1] & R[2] & \cdots & R[p-1] \\ R[1] & R[0] & R[1] & \cdots & R[p-2] \\ R[2] & R[1] & R[0] & \cdots & R[p-3] \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R[p-1] & R[p-2] & R[p-3] & \cdots & R[0] \end{bmatrix} \tag{4.37}$$

is the autocorrelation matrix,

$$A_\alpha = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_p \end{bmatrix}^T \tag{4.38}$$

is the linear prediction coefficients array and

$$r_\gamma = \begin{bmatrix} R[1] & R[2] & R[3] & \cdots & R[p] \end{bmatrix}^T \tag{4.39}$$

is the modified autocorrelation array (vector) with

$$E\left[z_{k-j}^T z_{k-i}\right] = \begin{cases} R[0], & \text{if j = i} \\ R\left[|j-i|\right], & \text{if j} \neq \text{i} \end{cases} \tag{4.40}$$

The optimal values of LPC are obtained by solving Equation (4.36). It is worth noticing that the complexity of linear prediction filter (LPF) is directly related to Equation (4.36) which depends on the size of matrix R_z, in other words the order of LPF. The higher order of an LPF does not necessarily reflect the optimal reconstruction of a signal [14], therefore has led to the discussion on the 'optimal order' of the LPF below.

4.3.1 Order of Linear Prediction Filter

The selection process of the order of an LPF is important to achieve the optimal reconstruction, and hence optimal estimation of the signals. In order to obtain optimal values for LPF orders, the applications of straightforward approaches are found better than the others. The following assumption needs to be considered for that purpose:

Assumption: $p \leq T_{st}/T_{sm}$ where T_{sm} is the sampling time, T_{st} is the starting time of loss of observations (LOOB), and p is the order of LPF.

This assumption in fact indicates that the number of previous observation samples, utilised in the reconstruction of missing measurements is limited by the starting time of LOOB and the sampling frequency. This assumption indicates an additional importance for the selection of an appropriate LPF's order (see Algorithm 3).

Assuming the starting time instant of LOOB to be at time $t = k$. Having the measurement updated step known at time instant $k - 1$, two methods are proposed which are aimed at yielding an optimal order of the LPF. The first method is summarised in Algorithm 2.

Algorithm 2 : Selection of the LP filter order (first method)

1: **Compute** $e_m(k-1) = \max\ (e_i)$, where $e_i = x_i - \hat{x}_i \quad \forall \quad i = \{1, 2, \dots k-1\}$.
2: **Initialisation** $j = 1$, Compute R_z and r_γ;
3: **Recursion**: $j = 2 \dots p$
4: **Obtain** $\bar{z}$ through Equation (4.10)
  Calculate measurement updated state estimation ${}_c\hat{x}_k$ based on this compensated observations
  Compute $e_j(k) = x_k - {}_c\hat{x}_k$.
5: **Check** Is $e_j(k) \leq e_m(k-1)$
  Yes $n \longleftarrow j$: Order of the LP filter
  Else $j \longleftarrow j + 1$
6: **Repeat** Step 4

Algorithm 2 reconstructs the data signal which will not produce error higher than the maximum error limit in the normal procedure of without any data loss. On successful implementation, Algorithm 2 can be called an ideal proposition as it does not permit any influence of LOOB to the process of state estimation. Nevertheless, alternative approaches might be proposed to tackle the problem of order of the selection process.

In real time processing, however Algorithms 2 may be found saturated i.e. the constraint mentioned in the algorithm may not be fulfilled during the whole recursion process. In other words, non of the combination of previous samples may generate error, less than the maximum error generated by the normal case (without any data loss). Therefore, to avoid this dilemma, Algorithm 3 assures results in all circumstances. To summarise, Algorithms 2 is a better choice from efficiency point of view – it provides lesser error with smaller computational time but its implementation is uncertain. On the other hand, Algorithm 3 provides inevitable results but is computationally expensive.

4.4 Compensated State Estimation Algorithms

In the event of LOOB, conventional KF (CKF) and OLKF schemes may be found inappropriate solutions to use for optimal state estimation problem. The reason being is that false measurement update step is achieved in CKF as opposed to OLKF, whereby no measurement update is performed during the data loss period. This issue has been discussed in Chapter 3. Therefore in this section, the observation vector reproduced through modified external linear prediction scheme (Section 4.3), is utilised in the process of Kalman filtering for the systems subjected to LOOB. Based on various combinations, a number of modified linear prediction coefficient

Algorithm 3 : Selection of the LP filter order (second method)

Initialisation $j = 1$, Compute R_z and r_γ through Equations (4.40)

Recursion $j = 2, 3, \ldots M \leq T_{st}/T_{sm}$ (LPFO)

2.1 Calculate LPC

2.2 Calculate compensated observation signal $\bar{z}_k|_{\alpha_\gamma^j}$

2.3 Calculate compensated state estimation ${}_c\hat{x}_k^j$ based on this signal

2.4 Calculate compensated state estimation error $\bar{e}_j = x_k - {}_c\hat{x}_k^j$

Trace $\epsilon_{th} = \min(\bar{e}_j)$, whereas $\bar{e}_j \in \{\bar{e}_2, \bar{e}_3, \ldots \bar{e}_M\}$

Select $\bar{z}_j$ which results in ϵ_{th}

Decide $p \leftarrow j$ i.e. LPFO.

(MLPC) techniques are introduced, which are discussed below.

Approach I: Zero-Order Hold (ZOH)

Perhaps the easiest way to perform measurement update step is to hold the last sensor reading all the time. Whenever, LOOB is diagnosed (Section 3.2), this holding observation value can be employed to obtain updated state and covariance values. If the LPF parameters are provided as $p = 1$ and $\alpha = 1$, hence the compensated observation vector is retranslated as

$$\begin{aligned} \bar{z}_k &= z_{k-1} = Hx_{k-1} + \theta_{k-1} \\ &= H\bar{x}_k + \bar{\theta}_k \end{aligned} \tag{4.41}$$

In literature, the above scheme is known as the Zero-Order Hold (ZOH) technique [67], [86], [53] and [22] due to the fact that the last observation is held to provide the measurement update step for the predicted state and error covariance. The ZOH method may be considered an effective approach for systems with sufficiently slow measurement samples, which may be considered stationary for a few sampling times and hence the last sensor reading can be locked (such as ship dynamics and commercial airplanes). However, this scheme might fail in reconstruction of the given signal if the sampling frequency of the sensors' measurements is high. In other words, if some nonlinearities are involved in the output data then the last previous data sample may not bear fruitful results for an adequate data loss (such as jet fighters, UAVs and spacecraft systems).

Approach II: First-Order Hold (FOH)

If the order of LPF $p = 2$ and $\alpha_i = [\alpha_1, \alpha_2]$, this would result in

$$\bar{z}_k = \alpha_1 z_{k-1} + \alpha_2 z_{k-2} \tag{4.42}$$

The scheme is known as First-Order Hold (FOH) approach which is divided into two subclasses as discussed below.

- *Mean (or Average) Based Approach (MBA)*
  If $\alpha_1 = \alpha_2 = 0.5$ i.e. the two weights are set equal, therefore

$$\begin{aligned} \bar{z}_k &= \frac{z_{k-1} + z_{k-2}}{2} \\ &= H\bar{x}_k + \bar{\theta}_k \end{aligned} \tag{4.43}$$

  where $\bar{x}_k = \frac{x_{k-1}+x_{k-2}}{2}$ and $\bar{\theta}_k = \frac{\theta_{k-1}+\theta_{k-2}}{2}$. It can be seen in the above approach, computation of the modified linear prediction coefficients (MLPC) is not required in relation to the linear prediction system. On the other hand, an MLPC for the two sensor readings are achieved as a vital consideration to obtain performance boost as discussed in the subsequent subclass.

- *MLPC Type-I: Diverse Weighted Observations*
  In the Mean Based Approach (MBA), two observations are weighed identically which may not always indicate a favourable solution. Alternatively, a different weighing scheme assigned to the two sensor readings (z_{k-1} and z_{k-2}) can be considered which is referred to as MLPC Type-I method. [1]

From the autocorrelation standpoint of view [84], the far most data can impart less in the linear prediction as compared to closer data, i.e. the more one depart from signal at time step k, the less correlation can be found, hence nearer sensor readings should be credited more than comparatively far measurements, i.e.

$$\begin{aligned} \bar{z}_k &= \alpha_1 z_{k-1} + \alpha_2 z_{k-2} \\ &= H\bar{x}_k + \bar{\theta}_k \end{aligned} \tag{4.44}$$

where $\bar{x}_k = \alpha_1 x_{k-1} + \alpha_2 x_{k-2}$ and $\bar{\theta}_k = \alpha_1 \theta_{k-1} + \alpha_2 \theta_{k-2}$. In this approach, the theory of linear prediction (LP) is integrated with the weights imposed to the previous measurements.

[1] The reason this type is separated from MBA approach and given the name MLPC Type-I is that, in this approach Linear Prediction algorithms 2 or 3 can be implemented to compute α_1 and α_2. Contrary to MBA approach where $\alpha_1 = \alpha_2 = 0.5$ i.e. equal weights are assigned to the two previous sensor readings.

Approach III: Trend-Based Algorithm

In some case studies, it might be noticed that observation vector contains a characteristic components like the trend, the periodic or seasonal component etc., in the received data [48]. In such cases, the estimation would not be affected by LOOB to a high extent. This is because, by finding the trend, the lost information can be easily recovered. These trends may vary from time to time and from a system to another [48]. Consider a simple trend between the data samples as

$$\bar{z}_k = z_{k-1} + [z_{k-1} - z_{k-2}] \tag{4.45}$$

or

$$\begin{aligned} \bar{z}_k &= Hx_{k-1} + \theta_{k-1} + [Hx_{k-1} + \theta_{k-1} - Hx_{k-2} - \theta_{k-2}] \\ &= Hx_{k-1} + \theta_{k-1} + [H\Delta_{x_{k-1}} + \Delta_{\theta_{k-1}}] \\ &= H(x_{k-1} + \Delta_{x_{k-1}}) + (\theta_{k-1} + \Delta_{\theta_{k-1}}) \\ &= H\bar{x}_k + \bar{\theta}_k \end{aligned} \tag{4.46}$$

where $\Delta_{x_{k-1}} = x_{k-1} - x_{k-2}$ and $\Delta_{\theta_{k-1}} = \theta_{k-1} - \theta_{k-2}$, $\bar{x}_k = x_{k-1} + \Delta_{x_{k-1}}$ and $\bar{\theta}_k = \theta_{k-1} + \Delta_{\theta_{k-1}}$. Equation (4.45) can be deduced by substituting the LPFO $p = 2$, $\alpha_1 = 2$ and $\alpha_2 = -1$ in Equation (4.10). Hence, the theory of linear prediction is not applicable to this approach and therefore is a trivial technique.

Approach IV: Moving Average Approach

For $p \geq 2$, the approach is termed as Moving Average (MA) approach in which every measurement is associated with an equal weight. Therefore, measurement equation is written as

$$\begin{aligned} \bar{z}_k &= \frac{1}{n}\sum_{j=1}^{p} z_{k-j} \\ &= H\frac{1}{n}\sum_{j=1}^{p} x_{k-j} + \frac{1}{n}\sum_{j=1}^{p} \theta_{k-j} \\ &= H\bar{x}_k + \bar{\theta}_k \end{aligned} \tag{4.47}$$

where $\bar{x}_k = \dfrac{1}{n}\displaystyle\sum_{j=1}^{p} x_{k-j}$ and $\bar{\theta}_k = \frac{1}{n}\displaystyle\sum_{j=1}^{p} \theta_{k-j}$.

Approach V: MLPC Type-II

For the moving average approach, however, the theory of MLPC cannot be implemented, likewise with that of the MBA approach. This scheme has been found efficient for simple first order systems [54,55]. However, all previous data should not be treated identically if order of the system increases. In other words, different weights should be assigned based upon their contributions on the predicted (future) estimations. The proposed observation signal is rewritten as

$$\bar{z}_k = \sum_{j=1}^{p} \alpha_j z_{k-j} \tag{4.48}$$

where the values of α_j are computed through Equation (4.36) and p is order of the LP filter.

Equations (4.41) to (4.47) except Equation (4.45), are written in a unique way, i.e.

$$\bar{z}_k = H\bar{x}_k + \bar{\theta}_k,$$

which will assist in deriving the compensated Kalman filter scheme in a generalised framework. This particular form of observation vector would generate the following residual covariance matrix:

$$\bar{S}_k := E[\bar{e}_{zk}\bar{e}_{zk}^T] = H \ {}_{3c}P_k H^T + \bar{R}_k, \tag{4.49}$$

where $\bar{e}_{zk} = \bar{z}_k - \hat{z}_k$, $\bar{R}_k = E[\bar{\theta}_k\bar{\theta}_k^T]$ and ${}_{3c}P_{k|k-1} \stackrel{\text{def}}{=} E\big[{}_ce_{k|k-1}({}_ce_{k|k-1})^T\big] = E\big[(\bar{x}_k - x_{k|k-1})(\bar{x}_k - x_{k|k-1})^T\big]$ are respectively residual error vector, residual error covariance matrix and affected state error covariance matrix in case of LOOB. By employing the compensated observation vector in Kalman filter during LOOB, it should be interesting to observe the behaviour of various properties of the standard Kalman filter. It is necessary to inspect the attributes of the proposed Compensated Closed Loop KF and compare them with those of standard KF without any data loss. In the following section, the above proposed observation vector is employed at measurement update step when a system is subjected to intermittent observations. Compensated Kalman filter gain and the corresponding error covariance matrices are derived in an apparent approach.

4.4.1 Compensated Closed Loop Kalman Filter Gain and Error Covariance Matrices

In this section a straightforward procedure is performed to address the Compensated Closed Loop Kalman filtering (CCLKF) algorithm in order to obtain an optimal gain matrix in the event of loss of observations. Attempts have been made to show that the CCLKF guarantees a minimum error covariance. Consider the predicted state vector as

$$x_{k|k-1} = Ax_{k-1|k-1} + Bu_{k-1} \tag{4.50}$$

with a corresponding state error vector as

$$e_{k|k-1} = Ae_{k-1|k-1} + \xi_{k-1} \tag{4.51}$$

and the error covariance matrix as

$$P_{k|k-1} = E[e_{k|k-1}e_{k|k-1}^T] = AP_{k-1|k-1}A^T + Q_{k-1} \tag{4.52}$$

The unavailability of observation will in fact lead to consider a measurement reconstruction approach based on Equation (4.10). The results are intended to obtain in a generalised framework, therefore

$$\bar{z}_k = H\bar{x}_k + \bar{\theta}_k \tag{4.53}$$

This would result in *a posteriori* state estimate as

$$_cx_{k|k} = x_{k|k-1} + {}_cK_k(\bar{z}_k - Hx_{k|k-1}) \tag{4.54}$$

where $\bar{z}_k$ depends upon the approach employed and $_cK_k$ is the compensated Kalman filter gain matrix. The *a posteriori* estimation error is as follows:

$$\begin{aligned} _ce_{k|k} &= x_k - {}_cx_{k|k} \\ &= x_k - x_{k|k-1} - {}_cK_k[\bar{z}_k - Hx_{k|k-1}] \\ &= e_{k|k-1} - {}_cK_k[H\bar{x}_k + \bar{\theta}_k - Hx_{k|k-1}] \\ &= e_{k|k-1} - {}_cK_kH_ce_{k|k-1} + {}_cK_k\bar{\theta}_k \end{aligned} \tag{4.55}$$

It is important to note that the error signals $e_{k|k-1}$ and $_ce_{k|k-1}$ are different but are assumed to be correlated. The *a posteriori* error covariance can be obtained as

$$_cP_{k|k} = E[({}_ce_{k|k})({}_ce_{k|k})^T] \tag{4.56}$$

The following result can be easily obtained by substituting Equation (4.55) in Equation (4.56) and taking the expectation operation;

$$_cP_{k|k} = {}_{1c}P_{k|k-1} - {}_{2c}P_{k|k-1}H^T{}_cK_k^T - {}_cK_kH_{2c}P_{k|k-1} + {}_cK_k(H_{3c}P_{k|k-1}H^T + \bar{R}_k){}_cK_k^T \tag{4.57}$$

where

$$
\begin{aligned}
{}_{1c}P_{k|k-1} &\stackrel{\text{def}}{=} E\big[e_{k|k-1}e_{k|k-1}^T\big] \\
{}_{2c}P_{k|k-1} &\stackrel{\text{def}}{=} E\big[e_{k|k-1}({}_ce_{k|k-1})^T\big] \\
{}_{3c}P_{k|k-1} &\stackrel{\text{def}}{=} E\big[{}_ce_{k|k-1}({}_ce_{k|k-1})^T\big]
\end{aligned} \tag{4.58}
$$

are different *a priori* error covariance matrices caused by LOOB. No matter how long the plant output data is lost, *a posteriori* error covariance will constitute these three types of error covariance matrices. To obtain an optimal Kalman filter gain, define a cost function as

$$J_k := E[({}_ce_{k|k})^T {}_ce_{k|k}] \tag{4.59}$$

which is equivalent to

$$
\begin{aligned}
J_k &= trace\big\{E[{}_ce_{k|k}({}_ce_{k|k})^T]\big\} \\
&= trace({}_cP_{k|k})
\end{aligned} \tag{4.60}
$$

Substituting Equation (4.57), Equation (4.60) can be minimised with respect to ${}_cK_k$,

$$\frac{\partial J_k}{\partial_c K_k} = 0 - 2_{2c}P_{k|k-1}H^T + 2_cK_kH_{3c}P_{k|k-1}H^T + 2_cK_k\bar{R}_k = 0 \tag{4.61}$$

which results in

$${}_cK_k = {}_{2c}P_{k|k-1}H^T(H_{3c}P_{k|k-1}H^T + \bar{R}_k)^{-1} \tag{4.62}$$

Substituting this value in Equation (4.57), the final update error covariance matrix adopts the following structure

$${}_cP_{k|k} = {}_{1c}P_{k|k-1} - {}_{2c}P_{k|k-1}H^T(H_{3c}P_{k|k-1}H^T + \bar{R}_k)^{-1}H_{2c}P_{k|k-1} \tag{4.63}$$

These are the optimal Kalman filter gain and its corresponding error covariance matrix for the first sampling time of data loss occurrence. If LOOB is continued, the general structure of the compensated Kalman filter gain and updated error covariance would be as follows:

$${}_cK_{k+l} = {}_{2c}P_{k+l|k+l-1}H^T(H_{3c}P_{k+l|k+l-1}H^T + \bar{R}_{k+l})^{-1} \tag{4.64}$$

and

$${}_cP_{k+l|k+l} = {}_{1c}P_{k+l|k+l-1} - {}_{2c}P_{k+l|k+l-1}H^T(H_{3c}P_{k+l|k+l-1}H^T + \bar{R}_{k+l})^{-1}H_{2c}P_{k+l|k+l-1} \tag{4.65}$$

where $l = \{1, 2, 3....m\}$ corresponds to the data loss steps.

It is necessary to specify that due to data loss, the above equations of the gain and error covariance matrices cannot be further simplified, contrary to conventional Kalman filter, (Equations (2.63) and (2.64)) . In other words, these equations will always have three different types of error covariance matrices within the LOOB duration as

$$
\begin{aligned}
E\left[e_{k+l|k+l-1}e^T_{k+l|k+l-1}\right] &= {}_{1c}P_{k+l|k+l-1} \\
E\left[e_{k+l|k+l-1}({}_ce_{k+l|k+l-1})^T\right] &= {}_{2c}P_{k+l|k+l-1} \\
E\left[({}_ce_{k+l|k+l-1})({}_ce_{k+l|k+l-1})^T\right] &= {}_{3c}P_{k+l|k+l-1}
\end{aligned}
$$

It is also worthwhile to discuss the behaviour of Kalman filter after the observation is resumed. It is assumed that the loss of observation has occurred between time steps k and $k+m-1$, for some positive value of m. Therefore, the predicted state estimation at time step $k+m-1$ is calculated as

$$
{}_cx_{k+m|k+m-1} = A_cx_{k+m-1|k+m-1} + Bu_{k+m-1} \tag{4.66}
$$

The state estimation error generated from the predicted state will be

$$
{}_ce_{k+m|k+m-1} = A_ce_{k+m-1|k+m-1} + \xi_{k+m-1}, \tag{4.67}
$$

and predicted error covariance for such state prediction is

$$
{}_cP_{k+m|k+m-1} = A_cP_{k+m-1|k+m-1}A^T + Q_{k+m-1} \tag{4.68}
$$

Due to the availability of actual data, the measurement update state will be [2]

$$
{}_cx_{k+m|k+m} = {}_cx_{k+m|k+m-1} + {}_cK_{k+m}(z_{k+m} - H_c x_{k+m|k+m-1}), \tag{4.69}
$$

with the predicted state error as

$$
\begin{aligned}
{}_ce_{k+m|k+m} &= x_{k+m} - {}_cx_{k+m|k+m-1} - {}_cK_{k+m}(z_{k+m} - H_c x_{k+m|k+m-1}) \\
&= {}_ce_{k+m|k+m-1} - {}_cK_{k+m}(H_c e_{k+m|k+m-1} + \theta_{k+m}).
\end{aligned} \tag{4.70}
$$

This shows that only one error element appears in the measurement updated step. Therefore, the proposed filtering scheme makes the limits of conventional KF once the observations are available. If a cost function is emanated from the same course, an optimal value for the Kalman

[2] The reason why the leading subscript ‘c’ still appears while observations are now available is, because theoretically it takes infinite time to recover the effect of data loss. However, in practice it may take a few samples to bring the estimation back to the standard estimation trajectory. To avoid confusion, index ‘c’ can be omitted from the state and covariance equations.

filter gain, which guarantees the minimum estimation error, will be

$$ {}_cK_{k+m} = {}_cP_{k+m|k+m-1}H^T\left(H_cP_{k+m|k+m-1}H^T + R_{k+m}\right)^{-1}, \tag{4.71} $$

with the consequent error covariance matrix of

$$ \begin{aligned} {}_cP_{k+m|k+m} &= {}_cP_{k+m|k+m-1} - {}_cK_{k+m}H_cP_{k+m|k+m-1} \\ &= \left(I - {}_cK_{k+m}H\right)_cP_{k+m|k+m-1}, \end{aligned} \tag{4.72} $$

where R_{k+m} is the actual measurement noise covariance matrix.

Hence, the proposed compensated observation scheme bears similar trends for the Kalman filter gain and error covariance matrices in the event of LOOB. It will then bring the above two matrices to the same structures once the sensed measurements are available again. For a discrete time LTI system, subjected to a random data loss, the proposed approach is summarised in a concise manner in Algorithm 4.

In the following subsection, a few related characteristic properties of the proposed approaches are summarised.

4.4.2 Properties of the Proposed Compensated Estimation Approaches

It will be interesting to highlight some of the features associated with the proposed compensated estimation approaches, discussed in the previous subsection.

In Approach I (ZOH), a sensor reading is utilised repeatedly to extract information whenever LOOB occurs. This approach has been successfully implemented in the event of loss of data e.g. [22, 37, 54, 55]. Using ZOH approach, the last measurement sample and hence the same Kalman filter gain and error covariance matrices are held throughout the data loss period. In simple words, *using this scheme the Kalman filter performs like steady state Kalman filter.* The only requirement of this method is that, the most recent sensor reading must be stored all the times. It should be noted that conventional Kalman filter does not require any observation to store after utilising it at measurement update step – as can be viewed in Algorithm 1, Chapter 3.

The reason why two measurement readings would be useful in estimating the state of a system (Approaches II and III) is because it is not necessary that the output data signal z_k exhibits *"markov property"*, therefore z_{k-1} and z_{k-2} may convey more information jointly about z_k than z_{k-1} alone [4]. In Approach III, previous two observations are considered according to their

Algorithm 4 : Closed Loop Estimation algorithm using MLPC

1: **I**t is assumed that loss of observation occurs at time step k. In other words z_k is not available and everything is as normal upto time step $k-1$. Therefore, the predicted state estimation at time step $k-1$ is calculated as

$$x_{k|k-1} = Ax_{k-1|k-1} + Bu_{k-1},$$

2: **T**he state estimation error generated from the predicted state will be

$$e_{k|k-1} = Ae_{k-1|k-1} + \xi_{k-1},$$

and predicted error covariance for such state prediction is

$$P_{k|k-1} = AP_{k-1|k-1}A^T + Q_{k-1}$$

3: **I**nitialisation of the compensated case be,

$$\begin{aligned}
{}_cx_{k|k-1} &= x_{k|k-1} \\
{}_ce_{k|k-1} &= e_{k|k-1} \\
{}_cP_{k|k-1} &= P_{k|k-1} \\
{}_cK_{k-1} &= K_{k-1}
\end{aligned} \tag{4.73}$$

4: **D**ue to the unavailability of actual data, the measurement update state will be

$${}_cx_{k|k} = {}_cx_{k|k-1} + {}_cK_k(\bar{z}_k - H_cx_{k|k-1}),$$

where

$${}_cK_k = {}_cP_{k|k-1}H^T\big(H_cP_{k|k-1}H^T + R_k\big)^{-1},$$

with the consequent error covariance matrix of

$$\begin{aligned}
{}_cP_{k|k} &= {}_cP_{k|k-1} - {}_cK_kH_cP_{k|k-1} \\
&= \big(I - {}_cK_kH\big)_cP_{k|k-1},
\end{aligned}$$

where R_k is the actual measurement noise covariance matrix.

5: **T**ime-step is updated.

6: **Return** to Prediction step with $x_{k|k} = {}_cx_{k|k}$ and $P_{k|k} = {}_cP_{k|k}$ and definitions of (4.73).

relative distances in correlation frame of reference, i.e. nearer measurements are represented by a higher correlation as compared to those of far measurements. Employing Approaches II and III for compensation of loss-of-observations will require the last (recent) two sensor readings to be stored all the times.

In Approaches IV and V, the span of utilising previous observations is extended to benefit the situations where data is lost for a longer duration of time. However, this approach is more ex-

Algorithm 5 : Proposed CCLKF Algorithm

1: **Recursion:** $k = \{0, 1, 2, ...\}$
2: **C**ompute Predicted Quantities: $x_{k+1|k}$ and $P_{k+1|k}$
3: **C**heck: Status of η_{k+1}
  if $\eta_{k+1} = 1$
    Run Conventional Kalman filter (Algorithm 1; Chapter 2) to obtain Filtered Quantities ($x_{k+1|k+1}$ and $P_{k+1|k+1}$).
  Else
    Jump to Closed Loop Estimator (Algorithm 4) to obtain Compensated Filtered Quantities (${}_{c}x_{k+1|k+1}$ and ${}_{c}P_{k+1|k+1}$).
  end
4: **T**ime-step is updated
5: **R**eturn to Prediction step;

pensive as it requires sufficiently large storage for holding the previous observations. In addition, employing more measurement data in order to compute linear prediction coefficients causes more computational efforts. However, in short-period of loss of data scenarios, Approach IV and V are found effective [55].

It is important to mention that all these approaches are based on the assumption that noise signals are Gaussian in nature.

4.5 Summary

In this chapter, an innovative solution for a common but important issue of state estimation under data loss is presented. Conventional Kalman filter does not offer optimal and bounded-error state estimation due to its profound dependency on the measurements. Researchers have proposed schemes such as Open-Loop Estimation to tackle such problems. However, under adequate period of loss of observations, such schemes fail to produce bounded errors and hence optimal state estimation. In this chapter, a novel method is presented, in which the lost observation is reproduced through linear prediction techniques, given the name of modified linear prediction coefficients or MLPC algorithms. The conventional methods of linear prediction theory do not offer a distinct way to decide the number of data samples used in computing linear prediction coefficients (LPC). For this reason, some constraint based approaches are presented which optimally decide the number of data samples employed in calculating LPC.

Through a straightforward methodology, the compensated Kalman filter gain and the associated *a posteriori* error covariance matrices are derived. It has been found that the Compensated Closed-Loop KF (CCLKF) approach employed, for state estimation in the event of LOOB yields some extra elements in the optimal Kalman filter gain and error covariance matrices. However, these extra elements vanish once the observation is resumed. The chapter is ended up by some useful discussion on the characteristic properties of the proposed approaches.

To conclude, mathematical description of an innovative Kalman filter is addressed in this chapter when a compensated observation signal is used at measurement update step in the event when actual observation is not available. In the next chapter, the proposed compensated estimation scheme (CCLKF) is tested by simulating a simple case study example of mass-spring-damper (MSD) system. It is intended to compare the performances of the two schemes (the existing Open-Loop Kalman filter and the proposed CCLKF) with respect to conventional Kalman filter (CKF) algorithm without any data loss. All the performance revealing parameters such as estimated state, Kalman filter gain, error analysis, residual vector and error covariance matrices shall be presented. Computational analysis for the two state estimators will also be explored with respect to CKF.

Chapter 5

Characteristics of the Compensated Kalman Filter

5.1 Introduction

It has been discussed that loss of output data is a non-trivial issue in the majority of control and communication systems. More attention has been made to the estimation methods, where measurement plays an integral role in retrieving the system information (states). However those systems may suffer from data loss due to several factors discussed in the previous chapters. Existing compensated estimation tools such as Open-Loop Kalman filtering produce poorer results when data is not available for a reasonable time. To generate improved state estimation results, compensated Kalman filtering techniques have been presented thoroughly in the previous chapter based on linear prediction theory. The conventional linear prediction schemes have certain limitations which were overcome by imposing a few constraints on the number of previous data sample employed. Certain characteristic properties related to the proposed approaches are described which assist in selecting a proposed approach. Mathematical formulations for the proposed design have been derived for both the commence of data loss and after the data is resumed.

In order to explore the behaviour of these approaches in the process of state estimation, a Mass-Spring-Damper (MSD) system studied in [23], is first tested in terms of various parameters for the conventional KF. The proposed Compensated Closed Loop Kalman filtering (CCLKF) approach is then analysed in terms of a list of properties. The main emphasis in this chapter is on the comparison of the two approaches – Open-Loop Kalman filtering (OLKF) and the proposed CCLKF in the event of loss of data.

This chapter is organised as follows: In Section 5.2, a simple second order mass-spring-damper (MSD) system is described. The proposed CCLKF algorithm is implemented on MSD system to explore various characteristics in Section 5.3. Also included is a thorough comparison between the existing OLKF approach and the proposed CCLKF scheme in terms of update state and covariance estimations, Kalman filter gains, error analysis and computational analysis. The performance of the proposed CCLKF is explored for various orders of linear prediction filter. The associated shortcomings of the proposed CCLKF scheme are briefly mentioned in Section 5.4.

5.2 Mass-Spring-Damper System

In this section the description of a simple second order mechanical system namely a mass-spring-damper (MSD) system is demonstrated. Such systems are common control experimental devices frequently encountered in many technical laboratories.

5.2.1 System Dynamics

The *two-degree-of-freedom (2DOF)* mass-spring-damper system is depicted in Figure 5.1. The

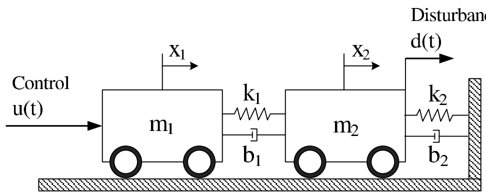

Figure 5.1: Two cart mass-spring-damper system

dynamics of such a system can be described by two 2nd-order differential equations, by Newton's Second Law,

$$m_i\ddot{x}_i + b_i\dot{x}_i + k_i x_i = u_i \qquad \forall \quad i = \{1, 2\} \tag{5.1}$$

where x_i are the displacements of the two masses from the equilibrium points and $u_i = F_i$ are the forces acting on the masses, with m_i being the masses, b_i the damper constants and k_i the spring constants. An equivalent block diagram in terms of electrical network is shown in Figure 5.2. Applying Kirchhoff's circuit laws (in other words superposition theorem) would result in the following equations. For simplicity the time-dependency (subscript – t) is omitted. [1]

$$\begin{aligned} u - m_1\ddot{x}_1 - k_1(x_1 - x_2) - b_1(\dot{x}_1 - \dot{x}_2) &= 0 \\ m_2\ddot{x}_2 - k_1x_1 - b_1\dot{x}_1 + (k_1 + k_2)x_2 + (b_1 + b_2)x_2 - \xi &= 0 \end{aligned} \tag{5.2}$$

where ξ is the plant disturbance associated with the speed of mass m_2. The two equations can be further simplified as ————

[1] $u_1 = u$ is the only applied force, while $u_2 = \xi$ is the noise source.

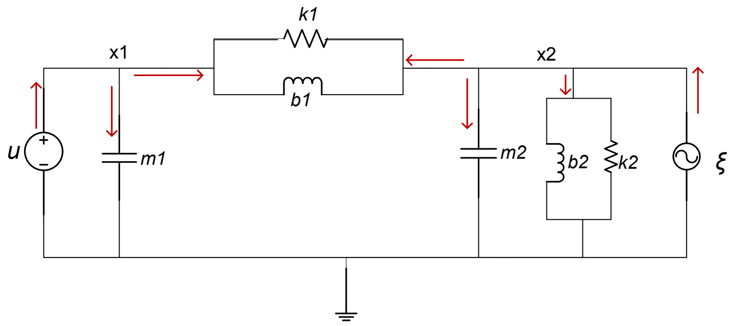

Figure 5.2: Equivalent electrical circuit diagram

$$
\begin{aligned}
m_1\ddot{x}_1 &= -k_1x_1 + k_1x_2 - b_1\dot{x}_1 + b_1\dot{x}_2 + u \\
m_2\ddot{x}_2 &= k_1x_1 - (k_1+k_2)x_2 + b_1\dot{x}_1 - (b_1+b_2)\dot{x}_2 + \xi
\end{aligned}
\tag{5.3}
$$

or

$$
\begin{aligned}
\ddot{x}_1 &= -\frac{k_1}{m_1}x_1 + \frac{k_1}{m_1}x_2 - \frac{b_1}{m_1}\dot{x}_1 + \frac{b_1}{m_1}\dot{x}_2 + \frac{u}{m_1} \\
\ddot{x}_2 &= \frac{k_1}{m_2}x_1 - \frac{(k_1+k_2)}{m_2}x_2 + \frac{b_1}{m_2}\dot{x}_1 - \frac{(b_1+b_2)}{m_2}\dot{x}_2 + \frac{\xi}{m_2}
\end{aligned}
\tag{5.4}
$$

The state space model of the system would be as follows:

$$
\dot{x}(t) = Ax(t) + Bu(t) + L\xi(t) \tag{5.5}
$$

where the state vector is defined as

$$
x^T(t) = [\ x_1(t) \quad x_2(t) \quad \dot{x}_1(t) \quad \dot{x}_2(t)\ ] \tag{5.6}
$$

with

$$
A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{k_1}{m_1} & \frac{k_1}{m_1} & -\frac{b1}{m1} & \frac{b1}{m1} \\ \frac{k_1}{m_2} & -\frac{k_1+k_2}{m_1} & \frac{b_1}{m_2} & -\frac{b_1+b_2}{m_2} \end{bmatrix} \tag{5.7}
$$

$$
B^T = [\ 0 \quad 0 \quad \tfrac{1}{m_1} \quad 0\ ] \tag{5.8}
$$

and

$$
L^T = [\ 0 \quad 0 \quad 0 \quad \tfrac{1}{m_2}\ ] \tag{5.9}
$$

The output dynamics can be described as

$$z(t) = Hx(t) + \theta(t) \tag{5.10}$$

Only one state is measured i.e. the displacement of mass m_2. Therefore, the output matrix can be described as

$$H = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} \tag{5.11}$$

5.2.2 Application of Mass-Spring-Damper System

The mass-spring-damper is a simple but practical application for state estimation, such as a robot arm (m_1) when lifting an object (m_2). This application is considered in this work because it is a simple "physical" system yet intuitively provides reasonable meaning of a real application [33]. For example, in a 2-D robot arm, the position of mass m_2 is a critical information which is normally read by a laser or sonic sensor. However, it may happen that an obstacle stays in front of the sensor and blocks the measurements temporarily. Similarly in automotive manufacturing processes a short-period of data loss is easy to imagine.

5.3 Simulation Results

For simulation purpose, the known parameters are selected as $m_1 = m_2 = 1$, $k_1 = 1$, $k_2 = 0.15$ and $b_1 = b_2 = 0.1$ and sampling frequency $T_s = 1ms$, [23]. Substituting these known values, the matrices will be as follows:

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 1 & -0.1 & 0.1 \\ 0.1 & -1.15 & 0.1 & -0.2 \end{bmatrix} \tag{5.12}$$

and

$$B^T = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} \tag{5.13}$$

$$L^T = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} \tag{5.14}$$

The two noise signals considered are characterised as

$$E\big[\xi(t)\big] = 0, \qquad E\big[\xi(t)\xi(\tau)\big] = \Xi\delta(t-\tau), \qquad \Xi = 1 \tag{5.15}$$

$$E\big[\theta(t)\big] = 0, \qquad E\big[\theta(t)\theta(\tau)\big] = 10^{-6}\delta(t-\tau) \tag{5.16}$$

For the purpose of this study, continuous-time dynamics of the MSD system are transformed to an appropriate discrete-time model using the sampling time of $T_s = 0.001$ s. In the simulation results, all four states are highlighted to show the performance of CCLKF and OLKF during LOOB period.

5.3.1 Updated State Estimation

In the first set of simulation results, performance of the conventional Kalman filter (CKF) is tested in normal operation. This has been achieved by assuming that there is no observation loss. The results of the normal operation are shown in Figures 5.3 and 5.4. Figure 5.3 shows the single noisy output measurement (z_k, position of mass m_2) along with its actual state, x_2, and the associated state estimation, $\hat{x}_2$. As expected in the failure-free case (normal operation), the conventional discrete-time Kalman filter is able to successfully estimate the position of mass m_2 within the provided simulation range.

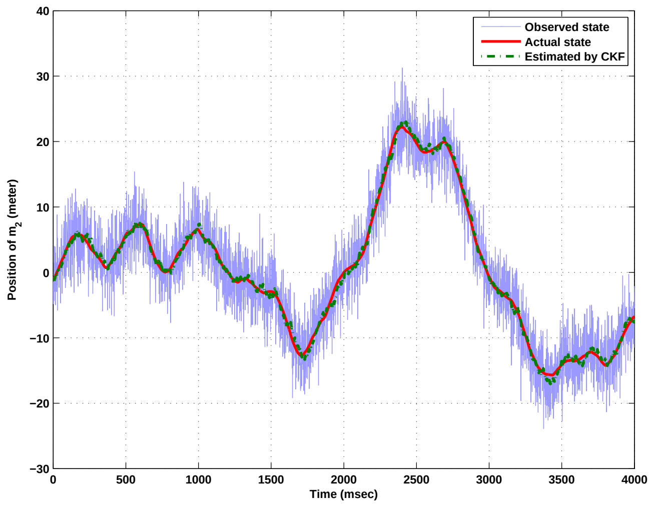

Figure 5.3: Estimation of state 2 without any LOOB.

In Figure 5.4 the three remaining states (namely, x_1, position of mass m_1; x_3, velocity of mass m_1; and x_4, velocity of mass m_2) are illustrated together with their associated state estimations. It can be seen from Figure 5.4 that, in normal operational mode, CKF is working perfectly, as all other states (which were not observed at the output of the system) are estimated reasonably well. It is worth noticing that, as expected, the effect of the plant disturbance is more severe on the velocity of mass m_2 than in other states - this can be seen by some sort of juggling in the results of state 4, x_4, as depicted in Figure 5.4.

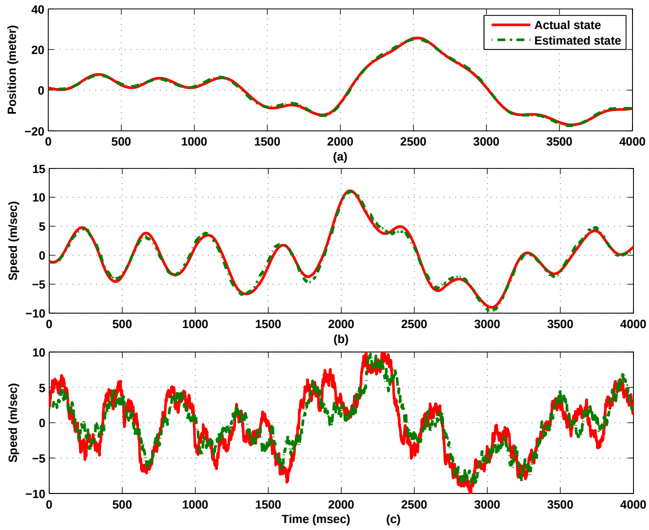

Figure 5.4: Estimations of (a) state 1, x_1 (b) state 3, x_3 and (c) state 4, x_4 – without any LOOB.

Secondly a typical set of simulation results is obtained for MSD system subject to the fault-induced measurements. The results of the proposed method (CCLKF) are compared with those of Open-Loop Kalman filtering technique. The loss of observation caused by any abnormal factor (such as temporary sensor fault or transmission channel failure) is assumed to occur between $1.30 - 1.85$ sec. The maximum allowable order of the LP filter p is 40 in Algorithm 3, Chapter 4.

Figures 5.5 – 5.8 show the state estimation results. Each figure has two parts; the first part shows the whole simulation period while in the second part of the figure, the loss of observation interval is highlighted in order to depict clearly the performance of the proposed CCLKF scheme over the existing OLKF approach. These figures express that the proposed CCLKF methodology has the ability of estimating the desirable outcomes (all four states $(x_1(t) - x_4(t))$ of the system accurately in the event of measurement loss.

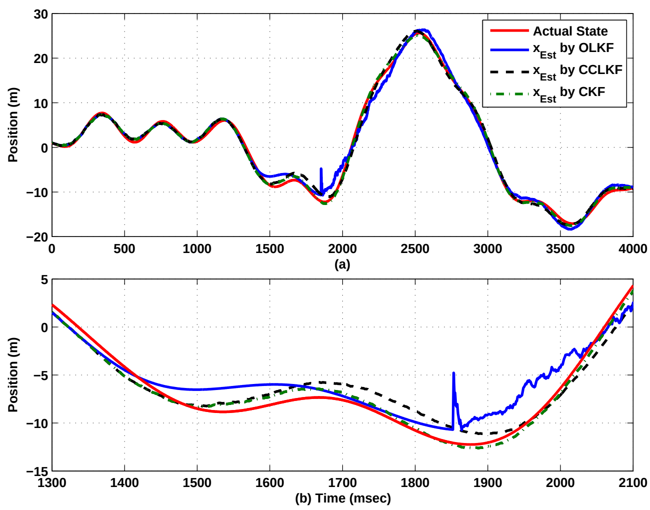

Figure 5.5: Estimation of state x_1 by OLKF, CCLKF and CKF (without any data loss).

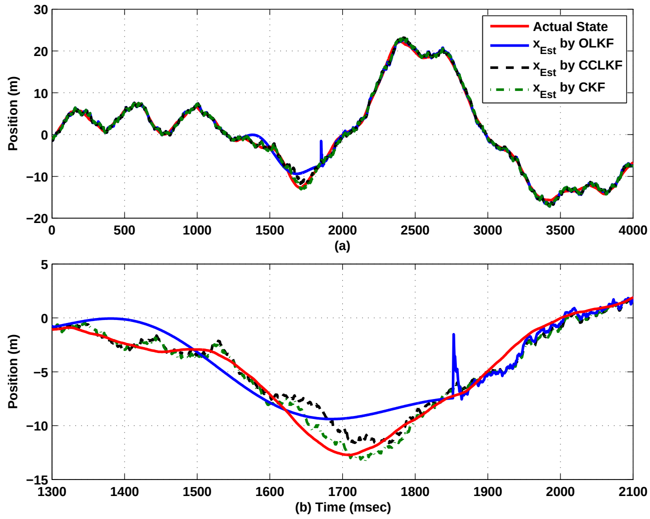

Figure 5.6: Estimation of state x_2 by OLKF, CCLKF and CKF (without any data loss).

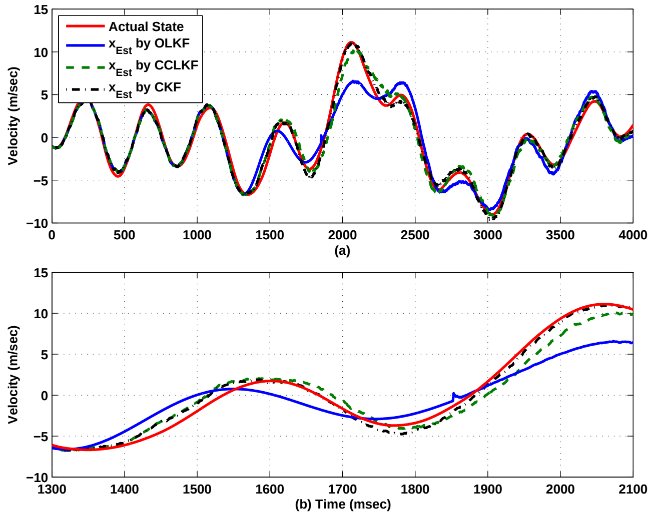

Figure 5.7: Estimation of state x_3 by OLKF, CCLKF and CKF (without any data loss).

5.3.2 State Estimation Error

Figures 5.9 and 5.10 illustrate the superiority of the CCLKF over the OLKF by manifesting the error analysis for the measured and unmeasured states. The error analysis is mainly highlighted for the measured state. It can be seen that during the LOOB, the state estimation error using the OLKF approach exceeds abruptly. On the other hand, the error generated by the proposed CCLKF approach is significantly small and hence it is less influenced. In fact this is one of the major achievements which can be experienced when applying CCLKF scheme.

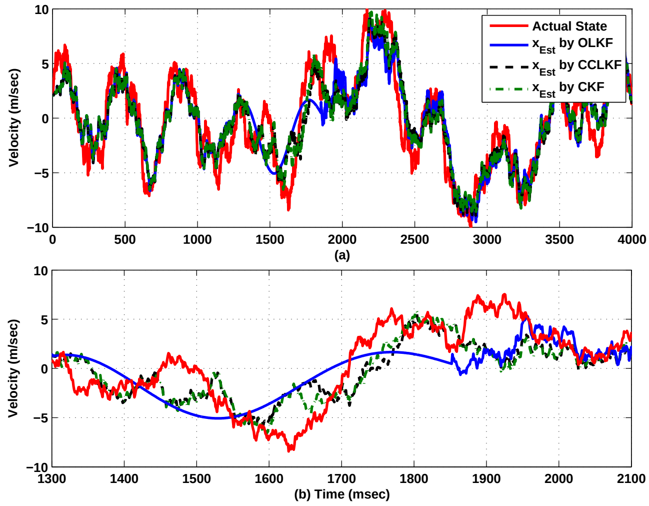

Figure 5.8: Estimation of state x_4 by OLKF, CCLKF and CKF (without any data loss).

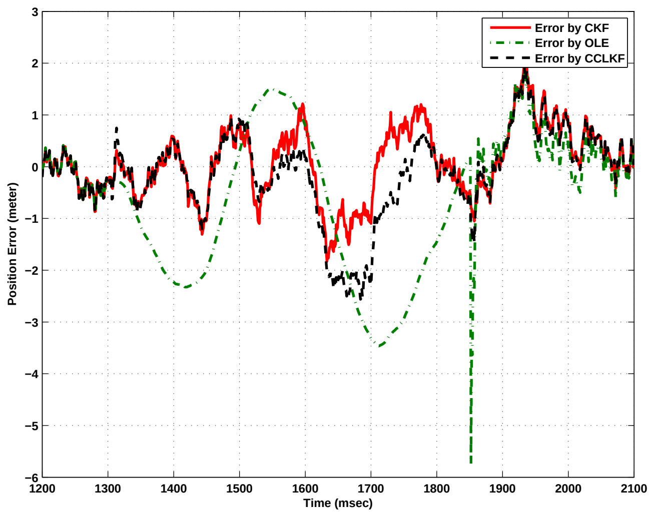

Figure 5.9: Estimation error of state x_2.

5.3.3 Kalman Filter Gain

The simulation results for the Kalman filter gain computed based on the OLKF and the proposed CCLKF approaches are shown in Figure 5.11. In the event of LOOB, the gain's elements are forced to zero until LOOB is recovered in OLKF approach. It is also worthwhile to stress that at the time of switching from Open-Loop Estimator back to normal (conventional) Kalman filtering, there are undesirable excessive oscillations in the gain's elements to help hold the system's state swiftly. Kalman filter gains using the proposed method are fairly smooth on comparison. Due to the gains being calculated online, it can be seen that when the LOOB occurs, the compensated Kalman filter gains are very close to those of original system. This will turn out that the proposed design based on CCLKF could successfully predict the state estimation of the original system even if an adequate period of loss of observation has occurred. Clearly, such an optimal compensated filter will provide minimum error for the estimation vector.

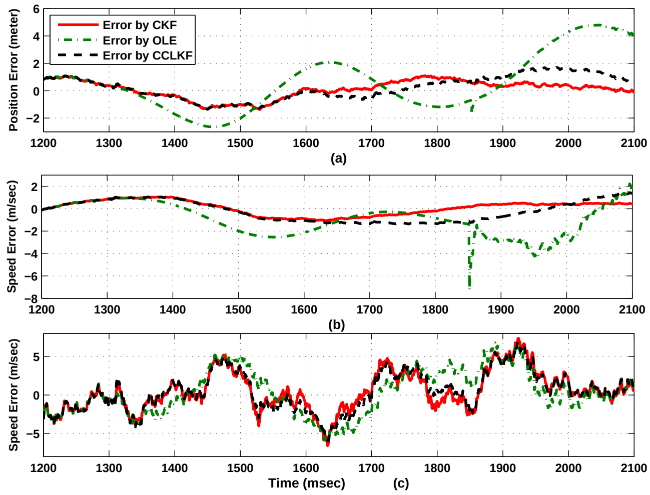

Figure 5.10: Estimation error of states x_1, x_3 and x_4

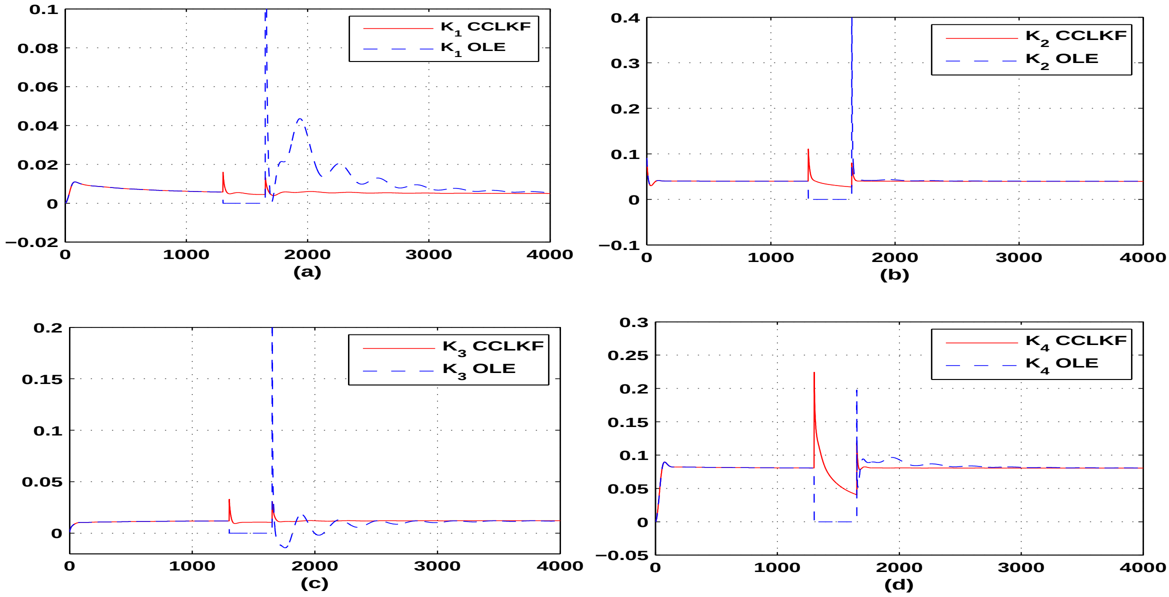

Figure 5.11: Comparison of Kalman filter gain's elements

5.3.4 Error Covariance

In this section error covariance analysis is explored for the two approaches (CCLKF and OLKF), as shown in Figure 5.12. As no update step is performed in OLKF approach, the update error covariance diverges abruptly from its normal trajectory. It is likely to happen that the error covariance would exceed the error bounds swiftly. On the other hand, the proposed CCLKF technique keeps error covariance bounded for a much longer time and hence keeps the design filter stable for longer in the event of loss of observation.

This feature is highlighted by providing error covariance for various orders of linear predictor filter (LPFO) in CCLKF approach. For a low LPFO, the divergence (deviation from the normal trajectory) rate is normally higher as compared to large LPFO. However for any LPFO, the results are much better than those of OLE scheme. It is important to mention at this point that after a certain limit of LPFO any further decrease in the expected value of the error (and hence error covariance) is not feasible. Some theoretical discussion on the convergence of the error covariance in the event of LOOB has been provided in the subsequent section.

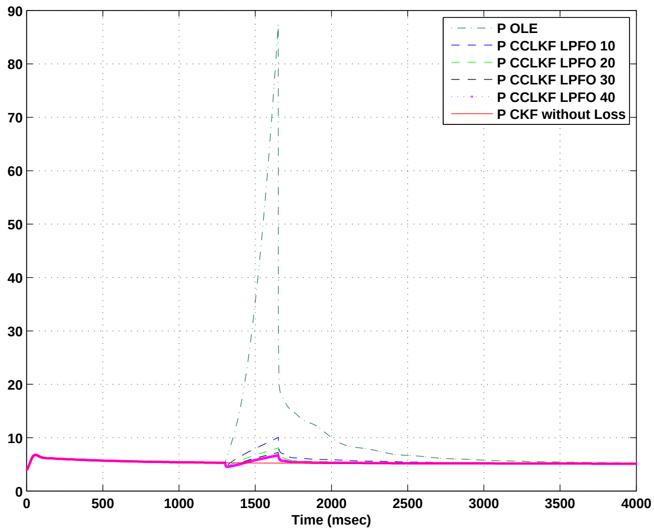

Figure 5.12: Traces of error covariance matrices of OLE and CCLKF

5.3.5 Computation Analysis

There are certain disadvantages for the proposed estimation techniques. The major issue associated with the proposed CCLKF scheme is its computational time. It is fairly easy to deduce that as the number of observation samples increases in the external linear prediction scheme in order to reconstruct the missing data, the computational time for its corresponding state estimation increases.

An overview of computational time analysis for various LPFO has been shown in Table 5.1. These results are obtained using a desktop computer running MATLAB, under Microsoft Windows Vista with 2.0GHz Intel Core 2 Dual processor with 2GB RAM. It was mentioned that a higher order of Linear Predictor filter order does not mean an optimal reconstruction of lost signal [14], and hence does not guarantee an optimal state estimation. This statement can be verified from Table 5.1, which is obtained for 0.55 s data-loss. In fact, for a particular LPFO, the error gets saturated and any further increase in the LPFO does not bear fruitful results in the estimation error. Some of the graphical representative results are shown in Figure 5.14.

With 0.55 s data-loss duration, the OLKF approach takes 2.3436 s in order to compute the four state estimations. On the other hand, the conventional Kalman filter (CKF) takes 2.5330 s in

No.	LPFO	Error 1	Error 2	Error 3	Error 4	Comp. Time (sec)
CKF Error	→	0.4297	0.4520	0.4150	1.6061	3.7366
OLE Error	→	1.1157	0.7002	1.1597	1.8181	2.9984
1	10	1.7405	0.8043	1.5178	1.7721	5.1498
2	20	0.6678	0.4843	0.6965	1.6256	7.4886
3	30	0.6582	0.4752	0.6063	1.6225	10.7856
4	40	0.5489	0.4630	0.5091	1.6183	16.5492
5	50	0.5257	0.4611	0.4717	1.6186	23.2230
6	60	0.4920	0.4585	0.4794	1.5984	33.5825
7	70	0.4282	0.4553	0.4180	1.5946	44.4685
8	80	0.4319	0.4518	0.4236	1.6057	61.1164

Table 5.1: Computation analysis for conventional Kalman filter (without any data loss), Open-Loop Kalman filter and proposed CCLKF

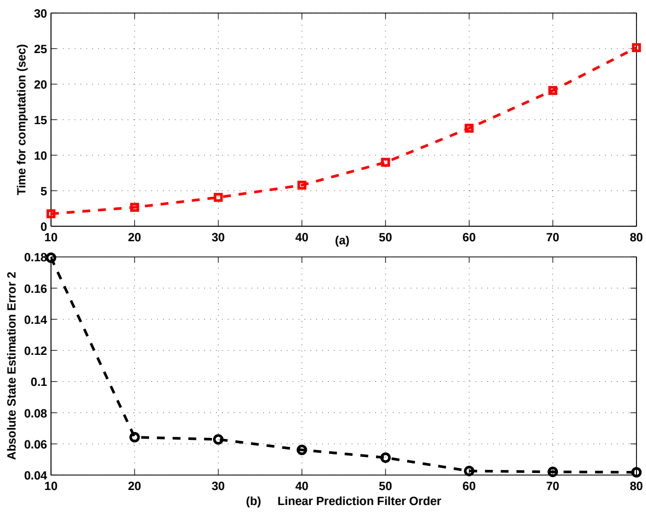

Figure 5.13: Analysis of (a) computational time and (b) corresponding error for various LPFO

normal operation (without any data loss) for the same simulation period. The reason OLKF's computational time is less than CKF approach is because during LOOB time, OLKF scheme only computes predicted quantities (state and error covariance) and no update is performed.

5.3.6 Levinson-Durbin Algorithm Results

Computing LPC through autocorrelation or auto-covariance involves $O(p^3)$ multiplications and additions and an inversion of a $p \times p$ matrix [103]. Consequently, more computational efforts need to be undertaken for an adequate LPF order. In Chapter 4, it has been mentioned that the computational effects due to the inversion of large autocorrelation matrices in the routine processes can be reduced by a few techniques. Out of these, the Levinson-Durbin algorithm (LDA) was briefly studied. In the conventional LDA, there is no distinct way to decide the upper limit of the linear prediction filter's order, which in turn is not very efficient in terms of the reduction of computation time. Therefore, a constraint has been introduced to limit the order of LPF in processing the existing Levinson-Durbin algorithm. Of course, this is an open problem and quite a few constraints can be introduced to limit/decide the order of LPF. A simple Constraint LDA (CLDA) is summarised below. [2]

Algorithm 6 : Constraint Levinson Durbin Algorithm (CLDA)

1: Initialisation $l = 0$; $E^0 = R_\gamma[0]$

2: Set a threshold estimated error value of $e_{th} \stackrel{\text{def}}{=} \bar{e}$, where $\bar{e} = \max(e_i) \in \{e_1, e_2, \ldots, e_{k-1}\}$

3: Recursion $l = \{1, 2, 3....p\}$ LPF order

3.1 Compute Reflection Coefficients as $\kappa_l = \frac{1}{E^{l-1}}\left[R_\gamma[l] + \sum_{j=1}^{p-1} a_j^{l-1} R_\gamma[l-j]\right]$

3.2 Calculate LPCs for l^{th} order predictor as
$a_l^l = -\kappa_l$
$a_j^l = a_j^{l-1} - \kappa_l a_{l-j}^{l-1} \quad \forall \;\; j = \{1, 2, \ldots, l-1\}$

3.3 a) Calculate $\bar{z}_l(k)$ based on the available LPC.
b) Calculate the $\hat{x}_l$ based on $\bar{z}_l(k)$
c) Compute error signal e_l for this $\hat{x}_l$
d) **I**s $e_l \leq e_{th}$,
Yes; stop the process and $p \leftarrow l$
The resultant linear prediction coefficients are $a_j = a_j^{(p)}$,
where $j = \{1, 2, \ldots, p\}$
Else Compute minimum mean square prediction error associated with the last p^{th} predictor by $E^l = E^{l-1}(1 - \kappa_l^2)$,
Update $l \leftarrow l + 1$ and
Repeat Step **3.1**

4: Return

This time data loss is introduced at another point to observe the flexibility and hence the performance of the proposed algorithm. Missing data has occurred between 2.3 - 2.8 s (for 0.5 seconds). The time consumed by OLKF is 2.4128 seconds to accommodate such a data loss. On the other hand, conventional Kalman filter (without any data loss) takes 2.4512 seconds to complete the whole simulation process. The computational time for various orders of LPF in CCLKF approach while employing the constraint Levinson-Durbin algorithm (CLDA) is shown in Table 5.2. Performance of the CLDA is more prominent for higher order of Linear Prediction filter (LPFO). This is because, the routine procedure to solve the *Normal Equations* involves large autocorrelation matrices inversions (of the order LPFO $\times$ LPFO), which is computationally cumbersome specially for higher orders of LPF.

[2]It is assumed that this algorithm is implemented at time step k

LPFO	Compt. time (NE)	Compt. time (CLDA)
10	5.1498	3.9012
20	7.4886	4.4582
30	10.7856	6.0426
40	16.5492	7.9714
50	23.2230	10.1721
60	33.5825	12.0896
70	44.4685	15.0436
80	61.1164	18.5693

Table 5.2: Computation analysis in seconds for Normal Equation (Autocorrelation) and Constraint Levinson-Durbin algorithm

The performance of the proposed CCLKF approach based on CLDA is tested by employing the mass-spring-damper (MSD) system, for a 0.5 s data loss (from 2.3 - 2.8 seconds). The CLDA has been tested for various orders of linear prediction filter ranging from $10-80$ and the corresponding results are shown in Figures 5.14 – 5.16. Among these, the best results in terms of least square estimation error, are observed for LPFO = 40. Thereafter by increasing the order of LPF, the error is found saturated.

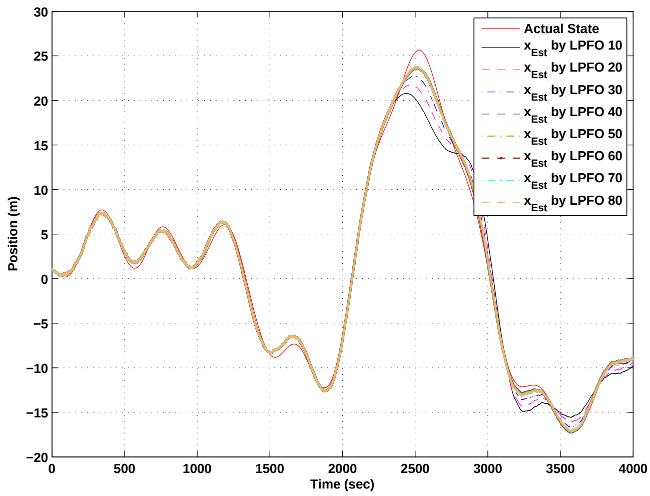

Figure 5.14: Estimation of state x_1 of MSD system using CLDA for various LPFO.

In Figures 5.17 and 5.18, the performance of CCLKF using CLDA is compared with that of CCLKF based on *Normal Equations* scheme. From the two figures and Table 5.2, it can be observed that almost for the same quality of estimation, the computational time can be reasonably reduced by using constraint Levinson-Durbin algorithm.

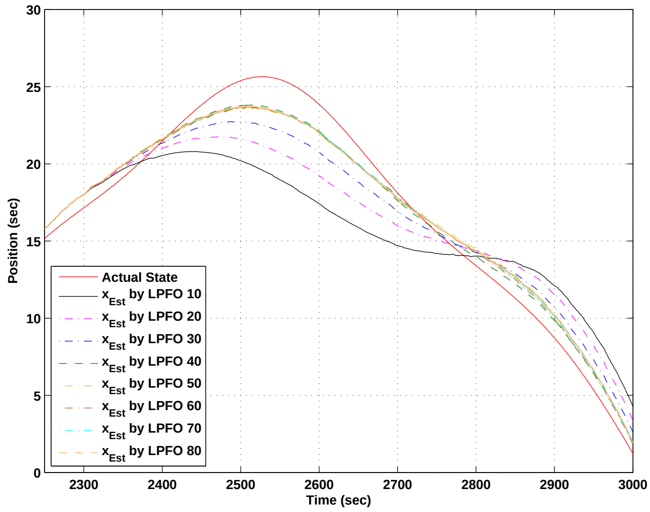

Figure 5.15: Enlarge view of the Fig 5.14.

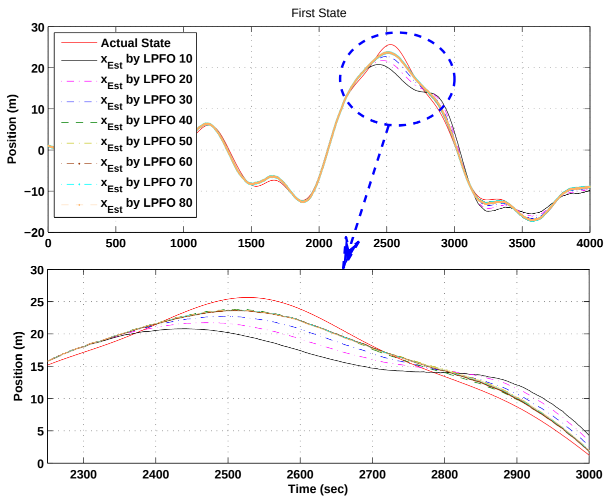

Figure 5.16: Combined picture the Figures 5.14 & 5.15

5.4 Shortcomings of the Proposed CCLKF

The proposed CCLKF approach could be considered a useful estimation algorithm for many challenging engineering applications suffering from loss of measurements. However,

- The main drawback of CCLKF is its computational burden which requires a very fast processor even if the order of LPF (p) is adequate. Achieving the desired performance of the CCLKF demands a considerable time in order to compute the compensated observational signal. Depending on system dynamics, order of LP filter and various other factors, computational time varies. However, having the hope that new technologies will bring sufficiently fast processors in future, this drawback could pale into insignificance when using CCLKF.

- Second drawback associated with CCLKF approach, which in fact is related to the above

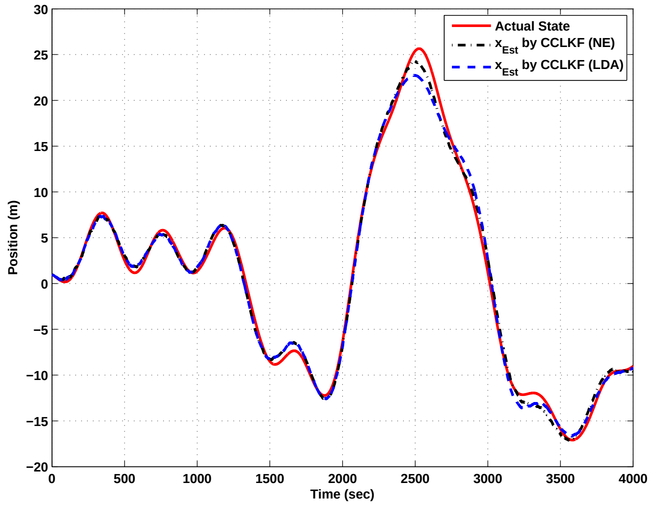

Figure 5.17: Estimation of state x_1 by NE and CLDA for LPFO = 30

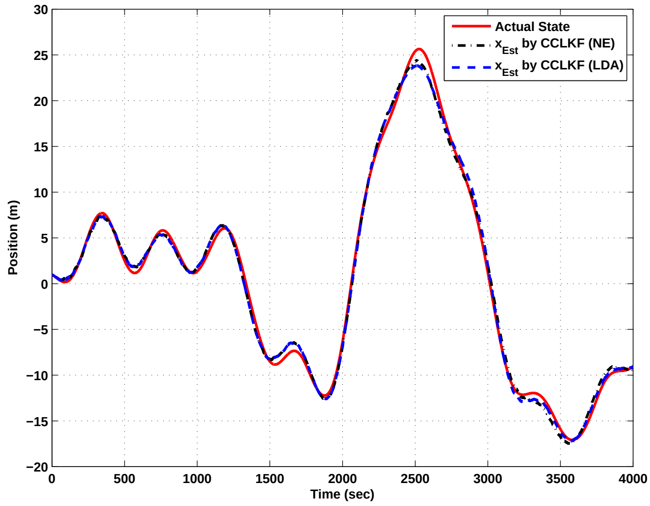

Figure 5.18: Estimation of state x_1 by NE and CLDA for LPFO = 40

issue, is hardware limitation due to large requirement to store previous observations (especially it is true for multi-input multi-output plants). At every data loss time instant, p-number of observations need to be retreated at the computation of compensated observation, which leads to the need of a buffer register that could store and update these p data samples.

- Third drawback associated with the proposed approach is that the repossessing of the steady state values (Subsection 3.3.4) are not fully attained in the CCLKF. However, this shortcoming has been improved by reducing the difference between the steady state and estimated values. Deviation from steady state covariance is related with convergence and asymptotic stability issues which will be thoroughly studied in the subsequent section.

- The compensated Kalman filtering algorithms are proposed based on Auto-regressive model where only output data (measurements) are considered in calculating the linear prediction coefficients. It does not take into account the control input (which ultimately would lead to ARMA model instead of autoregressive model) in order to simplify the matter. However, by doing so might weaken the proposed algorithms if the control signal has significant changes during the data loss period. This feature has explicit relationship with the assumption of slowly time varying characteristics of the process, made in Chapter 4.

5.5 Summary

In order to tackle loss of observation in state estimation, compensated Kalman filtering schemes have been proposed in Chapter 4. It was necessary to investigate various features resulting from implementation of the proposed approach to an appropriate example such as mass-spring-damper subjected to loss of observations. In this chapter, the performance of the proposed CCLKF scheme has been analysed in terms of various parameters e.g. estimated state and covariance, corresponding state estimation error, Kalman filter gain, *etc.* Attempts have been made to show these properties in a generalised framework beyond the considered mass-spring-damper system. As expected, the conventional autocorrelation method of solving *Normal Equations* has been found computationally expensive due to the inversion of large-dimension matrices at every instant of LOOB. This deficit has been overcome and the required computation efforts are reduced by modifying the existing Levinson-Durbin algorithm. The modified Levinson-Durbin algorithm has been found more effective for higher orders of linear prediction filter.

In the next chapter, the two compensated estimation schemes are implemented to estimate the attitude of a rigid body spacecraft system, subjected to loss of measurements. This case study example has been considered to explore the practical limitations of the two schemes.

Chapter 6

Compensated Estimation and Control of a Rigid Body Spacecraft

6.1 Introduction

Previous chapters have thoroughly presented an important problem of control and communication systems where the observations are subjected to random loss. This issue is possibly caused by various factors including intermittent sensor faults, temporary channel failures, congestion of network systems and limited memory of buffer registers. It has been found that, if problems arise in the process of state estimation where the output data plays an important role in retrieving system's information (states), the results would be catastrophic. In order to cope with such critical environment, compensated state estimation techniques have been proposed, based on modified linear prediction theory in Chapter 4. A simple mass-spring-damper (MSD) example was simulated to show some of the candidate results with comparison of the existing approach (Open-Loop Estimation) in Chapter 5 along with the discussion on the proposed CCLKF algorithm's characteristics.

In this chapter, the span of utilisation of the proposed algorithm is extended to a rigid body spacecraft system. Spacecraft technology is one of the emerging research field, which depends on the ground based data as one of the integral forms of communication. For a successful completion of a spacecraft mission, it relies on smooth arrival of sensor data. Any interruption in the arrival of data would effect the performance of the spacecraft attitude control system. Hence, in situations of intermittent observations, the proposed compensated close loop Kalman filtering algorithm is implemented to a rigid body spacecraft system in this chapter.

This chapter is organised as follows: Section 6.2 provides discussion on the nonlinear plant and output dynamics of a rigid body spacecraft model. Due to wide range of rotation, the plant and output dynamics are represented in modified Rodrigues parameterisations. The nonlinear plant

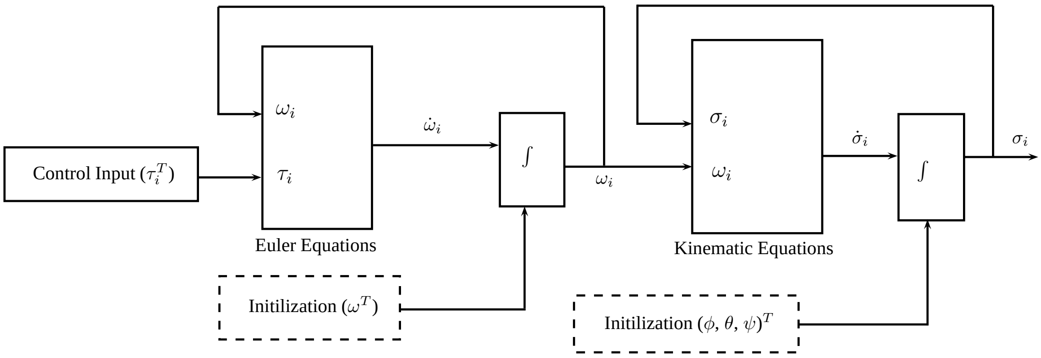

Figure 6.1: Nonlinear spacecraft attitude model

and output dynamics are linearised on nonzero operating points which are computed through Levenberg-Marquardt iterative method. The original continuous-time model has been discretised for simulation purposes. The original plant is unstable, therefore for stabilisation purposes, a Lyapunov based controller is presented in Section 6.3. In Section 6.4, the proposed CCLKF approach is recapitulated in a concise manner. Simulation studies based on various performance indices of a rigid spacecraft model subject to observation loss are illustrated through a numerical example in Section 6.5.

6.2 Spacecraft Dynamics

In literature, development of spacecraft model using quaternion representations has been driven by several factors such as

- linear treatment to prediction equations [51],
- avoidance of gimbal lock situation [64], and
- avoidance of involvement of any trigonometric function unlike Euler angles analysis.

However quaternion parameterisations suffer from the unit norm constraint [101]. To avoid that, Modified Roderigues Parameters (MRP) representation is employed to test the performance of CCLKF approach.

The plant dynamics of a spacecraft system are normally described while considering *Kinematic* equations only, (see [29, 56, 64, 122]). However in this work, the spacecraft system is modelled as a rigid body and its attitude model is described by two set of equations namely “Euler equations of rotational dynamics” and “Kinematic equations”, as follows:

6.2.1 Plant Dynamics using Modified Rodrigues Parameterisations

In this section, the nonlinear spacecraft model is presented in Modified Rodrigues Parameters (MRP) representations. The reason for carrying out this slightly complex parametrisation technique (MRP) is, to avoid the shortcomings associated with the quaternion parameters. Both *kinematic* and *dynamics* equations are employed to construct the nonlinear model of the rigid body spacecraft system.

The *Kinematic* equations in terms of MRP are defined as

$$\dot{\sigma} \stackrel{\text{def}}{=} T(\sigma)\overline{\omega}, \tag{6.1}$$

where $\sigma_{3\times 1}$ is the modified Rodrigues parameter vector,

$$T(\sigma) = \frac{1}{2}\left[\left(\frac{1-\sigma^T\sigma}{2}\right) I_{3\times 3} + S(\sigma) + \sigma\sigma^T\right], \tag{6.2}$$

and $\overline{\omega}_{3\times 1}$ is the noisy angular velocity vector which is defined as

$$\overline{\omega} \stackrel{\text{def}}{=} \begin{bmatrix} \overline{\omega}_1 \\ \overline{\omega}_2 \\ \overline{\omega}_3 \end{bmatrix} = \begin{bmatrix} \omega_1 - \nu_1 \\ \omega_2 - \nu_2 \\ \omega_3 - \nu_3 \end{bmatrix}, \tag{6.3}$$

where ν_i are bias elements caused by the rate gyros. The bias elements are assumed to possess Gaussian property with zero mean, i.e.

$$\nu_i \sim N(0, \Upsilon) \tag{6.4}$$

where Υ is the bias variance, and $S(\sigma) := \sigma \times \sigma^T$. Euler equations of *rotational dynamics* are shown as

$$J\dot{\overline{\omega}} = -S(\overline{\omega})J\overline{\omega} + \tau \tag{6.5}$$

The inertia matrix $J_{3\times 3}$ is defined as

$$J \stackrel{\text{def}}{=} \begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \\ J_{31} & J_{32} & J_{33} \end{bmatrix} \quad \Rightarrow \quad J^{-1} \equiv \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \tag{6.6}$$

ie. J is considered to be invertible. In Equation (6.5), $\tau_{3\times 1}$ is the control input and $S(\overline{\omega}) = \overline{\omega}\times\overline{\omega}^T$ is the skew symmetric matrix.

Linearisation

The linearised plant dynamics for the rigid body spacecraft system using MRP representations are obtained using Equations (6.1) and (6.5), which result in

$$\dot{x} = Ax + Bu + G\xi \tag{6.7}$$

where $x = [\sigma \;\; \omega]^T$, $u = \tau$ and ξ are state vector, control input vector and noise vector respectively. The linearised Jacobian matrices A,B and G are computed in Appendix B.

6.2.2 Spacecraft Output Dynamics

The spacecraft attitude parameters can be measured using a combination of reference sensors such as a sun sensor, star sensor or earth sensor, and inertial reference systems [110]. The latter measure the rates of rotation about each axis, and the integration of the rate of rotation would give the attitude, because they are working on a numerical integration. Such sensors however drift over time, hence the need for the reference sensors to act as a calibration arises. At least two reference sensors are required to give the three axis position. Some sensors give better accuracy than others and one needs to select based on the requirements of the mission and payload. Having Euler angles at the input model and MRP parameters at the output model, a precise relationship is required to be established between the two distinct representations.

There are two approaches to achieve output linearised model in terms of MRP; the first indirect approach which derives the desire relationship from Euler angles to quaternion and thereafter from quaternion to MRP, and the second is direct approach where the output is directly obtained by equating the two Direction Cosine Matrices (DCM) of Euler angles and MRP. The former approach requires massive calculations to derive the desired model plus the loss of performance during the two conversions and therefore the output linearised model is derived using the direct approach. For the same operating points, both direct and indirect models are expected to generate nearly similar results.

The two DCM in terms of MRP and Euler angles representations, of sequence 3-2-1 are shown in the following equations, [87].

$$\begin{aligned} DCM_\sigma \;=\; & \frac{1}{(1+\sigma^2)^2} \\ & \begin{bmatrix} 4(\sigma_1^2 - \sigma_2^2 - \sigma_3^2) + (1-\sigma^2)^2 & 8\sigma_1\sigma_2 + 4\sigma_3(1-\sigma^2) & 8\sigma_1\sigma_3 - 4\sigma_2(1-\sigma^2) \\ 8\sigma_2\sigma_1 - 4\sigma_3(1-\sigma^2) & 4(-\sigma_1^2 + \sigma_2^2 - \sigma_3^2) + (1-\sigma^2)^2 & 8\sigma_2\sigma_3 + 4\sigma_1(1-\sigma^2) \\ 8\sigma_3\sigma_1 + 4\sigma_2(1-\sigma^2) & 8\sigma_3\sigma_2 - 4\sigma_1(1-\sigma^2) & 4(-\sigma_1^2 - \sigma_2^2 + \sigma_3^2) + (1-\sigma^2)^2 \end{bmatrix} \end{aligned} \tag{6.8}$$

$$DCM_e = \begin{bmatrix} \cos\theta\cos\psi & \cos\theta\sin\psi & -\sin\theta \\ \sin\phi\sin\theta\cos\psi - \cos\phi\sin\psi & \sin\phi\sin\theta\sin\psi + \cos\phi\cos\psi & \sin\phi\cos\theta \\ \cos\phi\sin\theta\cos\psi + \sin\phi\sin\psi & \cos\phi\sin\theta\sin\psi - \sin\phi\cos\psi & \cos\phi\cos\theta \end{bmatrix} \tag{6.9}$$

In this case, the conversion deals with the nonlinear trigonometric functions. In general format, one can write as:

$$z(t) = h(x(t)) + v(t) \tag{6.10}$$

In order to obtain the linearised model, the output Jacobian matrix is obtained as [1]

$$z = Cx + v = \begin{bmatrix} \frac{\partial\phi}{\partial\sigma_1} & \frac{\partial\phi}{\partial\sigma_2} & \frac{\partial\phi}{\partial\sigma_3} & \\ \frac{\partial\theta}{\partial\sigma_1} & \frac{\partial\theta}{\partial\sigma_2} & \frac{\partial\theta}{\partial\sigma_3} & \mathbf{0}_{3\times3} \\ \frac{\partial\psi}{\partial\sigma_1} & \frac{\partial\psi}{\partial\sigma_2} & \frac{\partial\psi}{\partial\sigma_3} & \end{bmatrix} \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix} + v \tag{6.11}$$

In order to obtain the output linearised model, the following procedure has been taken place.

6.2.3 Linearised Output Model

Considering the direct method, the two Direction Cosine matrices (DCM) associated with Euler parameters and MRP are compared element-wise, in order to compute the linearised output model of the rigid body spacecraft system.

ϕ-Differentiations

To obtain the derivatives of ϕ with respect to σ_i, the same procedure is followed: Comparing the last element in the two DCMs

$$\cos\phi\cos\theta = DCM[3,3] = \frac{4(-\sigma_1^2 - \sigma_2^2 + \sigma_3^2) + (1-\sigma^2)^2}{(1+\sigma^2)^2} \tag{6.12}$$

Differentiating it with respect to σ_1, would result in

$$-\sin\phi\cos\theta\frac{\partial\phi}{\partial\sigma_1} = \underbrace{\frac{\partial}{\partial\sigma_1}\Big[\frac{4(-\sigma_1^2 - \sigma_2^2 + \sigma_3^2) + (1-\sigma^2)^2}{(1+\sigma^2)^2}\Big]}_{N_1} \tag{6.13}$$

[1] The subscript ‘t’ is dropped for simplicity.

or

$$\frac{\partial\phi}{\partial\sigma_1} = \frac{N_1}{-\sin\phi\cos\theta} = \frac{N_1}{-DCM[2,3]}$$
$$= \frac{-(1+\sigma^2)^2}{8\sigma_2\sigma_3 + 4\sigma_1(1-\sigma^2)}N_1 \tag{6.14}$$

where

$$N_1 = \frac{\partial}{\partial\sigma_1}\Big[\frac{4(-\sigma_1^2-\sigma_2^2+\sigma_3^2)+(1-\sigma^2)^2}{(1+\sigma^2)^2}\Big] \tag{6.15}$$

Apply quotient rule to compute the right hand side of Equation 6.15, i.e.

$$N_1 = \frac{v_1\dot{u}_1 - u_1\dot{v}_1}{v_1^2}$$

where

$$\begin{aligned} u_1 &= 4(-\sigma_1^2-\sigma_2^2+\sigma_3^2)+(1-\sigma^2)^2, \\ v_1 &= (1+\sigma^2)^2, \\ \dot{u}_1 &= \frac{\partial u_1}{\partial\sigma_1} = -8\sigma_1 + 2(1-\sigma^2)(-2\sigma_1) = -4\sigma_1(3-\sigma^2) \\ \dot{v}_1 &= 2(1+\sigma^2)2\sigma_1 = 4\sigma_1(1+\sigma^2) \end{aligned}$$

therefore

$$\begin{aligned} N_1 &= \frac{(1+\sigma^2)^2[-4\sigma_1(3-\sigma^2)] - [4(-\sigma_1^2-\sigma_2^2+\sigma_3^2)+(1-\sigma^2)^2]4\sigma_1(1+\sigma^2)}{(1+\sigma^2)^4} \\ &= \frac{4\sigma_1(1+\sigma^2)}{(1+\sigma^2)^4}\Big[-(1+\sigma^2)(3-\sigma^2) - 4(-\sigma_1^2-\sigma_2^2+\sigma_3^2) - (1-\sigma^2)^2\Big] \\ &= \frac{-4\sigma_1}{(1+\sigma^2)^3}\Big[(1+\sigma^2)(3-\sigma^2) + 4(-\sigma_1^2-\sigma_2^2+\sigma_3^2) + (1-\sigma^2)^2\Big] \end{aligned} \tag{6.16}$$

Substitute this value in (6.14) will result in

$$\begin{aligned} \frac{\partial\phi}{\partial\sigma_1} &= \frac{-(1+\sigma^2)^2}{8\sigma_2\sigma_3+4\sigma_1(1-\sigma^2)}\frac{-4\sigma_1}{(1+\sigma^2)^3}[(1+\sigma^2)(3-\sigma^2)+4(-\sigma_1^2-\sigma_2^2+\sigma_3^2)+(1-\sigma^2)^2] \\ &= \frac{4\sigma_1\big[(1+\sigma^2)(3-\sigma^2)+4(-\sigma_1^2-\sigma_2^2+\sigma_3^2)+(1-\sigma^2)^2\big]}{\big[8\sigma_2\sigma_3+4\sigma_1(1-\sigma^2)\big](1+\sigma^2)} \\ &= \frac{4\sigma_1\big[4-4\pi\big]}{\big[8\sigma_2\sigma_3+4\sigma_1(1-\sigma^2)\big](1+\sigma^2)} \\ &= H_1\sigma_1\big[1-\pi\big] \end{aligned} \tag{6.17}$$

where $\pi = -(\sigma_1^2+\sigma_2^2-\sigma_3^2)$ and $H_1 = \frac{16}{\big[8\sigma_2\sigma_3+4\sigma_1(1-\sigma^2)\big](1+\sigma^2)}$

In the similar way, the remaining two ϕ-differentiation are computed.

$$\frac{\partial\phi}{\partial\sigma_2} = H_1\big[\sigma_2\{1-\pi\}\big] \tag{6.18}$$

$$\frac{\partial\phi}{\partial\sigma_3} = H_1\big[-\sigma_3\{\sigma^2+\pi\}\big] \tag{6.19}$$

θ-Differentiations

Comparing the element-[1,3] of the both DCMs

$$-\sin\theta = DCM[1,3] = \frac{8\sigma_1\sigma_3 - 4\sigma_2(1-\sigma^2)}{(1+\sigma^2)^2} \tag{6.20}$$

differentiating with respect to θ

$$\begin{aligned} -\cos\theta\frac{\partial\theta}{\partial\sigma_1} &= \underbrace{\frac{\partial}{\partial\sigma_1}\Big[\frac{8\sigma_1\sigma_3 - 4\sigma_2(1-\sigma^2)}{(1+\sigma^2)^2}\Big]}_{N_2} \\ \frac{\partial\theta}{\partial\sigma_1} &= \frac{N_2}{-\cos\theta} \end{aligned} \tag{6.21}$$

N_2 is computed as follows:

$$N_2 = \frac{\partial}{\partial\sigma_1}\Big[\frac{8\sigma_1\sigma_3 - 4\sigma_2(1-\sigma^2)}{(1+\sigma^2)^2}\Big] \tag{6.22}$$

Applying the quotient rule as before: $N_2 = \frac{v_2\dot{u}_2 - u_2\dot{v}_2}{v_2^2}$, where

$$\begin{aligned} u_2 &= 8\sigma_1\sigma_3 - 4\sigma_2(1-\sigma^2) \\ v_2 &= v_1 = (1+\sigma^2)^2 \\ \dot{u}_2 &= 8\sigma_3 + 8\sigma_1\sigma_2 \\ \dot{v}_2 &= \dot{v}_1 = 4\sigma_1(1+\sigma^2) \end{aligned}$$

$$\begin{aligned} N_2 &= \frac{(1+\sigma^2)^2(8\sigma_1\sigma_2 + 8\sigma_3) - [8\sigma_1\sigma_3 - 4\sigma_2(1-\sigma^2)]4\sigma_1(1+\sigma^2)}{(1+\sigma^2)^4} \\ &= \frac{8}{(1+\sigma^2)^3}\big[(1+\sigma^2)(\sigma_1\sigma_2+\sigma_3) - 2\sigma_1[2\sigma_1\sigma_3 - \sigma_2(1-\sigma^2)]\big] \end{aligned} \tag{6.23}$$

Also from the DCM of Euler angle, it can be found easily that

$$\begin{aligned}
\cos\theta &= \sqrt{DCM[2,3]^2 + DCM[3,3]^2} \\
&= \sqrt{\Big[\frac{8\sigma_2\sigma_3 + 4\sigma_1(1-\sigma^2)}{(1+\sigma^2)^2}\Big]^2 + \Big[\frac{4(\overbrace{-\sigma_1^2 - \sigma_2^2 + \sigma_3^2}^{-\pi}) + (1-\sigma^2)^2}{(1+\sigma^2)^2}\Big]^2} \\
&= \frac{1}{(1+\sigma^2)^2}\sqrt{\big[8\sigma_2\sigma_3 + 4\sigma_1(1-\sigma^2)\big]^2 + \big[-4\pi + (1-\sigma^2)^2\big]^2} \qquad (6.24)
\end{aligned}$$

Substituting the values of N_2 from (6.23) and $\cos\theta$ from (6.24) into (6.21), will result in

$$\begin{aligned}
\frac{\partial\theta}{\partial\sigma_1} &= \frac{\frac{8}{(1+\sigma^2)^3}\big[(1+\sigma^2)(\sigma_1\sigma_2+\sigma_3) - 2\sigma_1[2\sigma_1\sigma_3 - \sigma_2(1-\sigma^2)]\big]}{\frac{1}{(1+\sigma^2)^2}\sqrt{\big[8\sigma_2\sigma_3 + 4\sigma_1(1-\sigma^2)\big]^2 + \big[-4\pi + (1-\sigma^2)^2\big]^2}} \\
&= \frac{8\big[(1+\sigma^2)(\sigma_1\sigma_2+\sigma_3) - 2\sigma_1[2\sigma_1\sigma_3 - \sigma_2(1-\sigma^2)]\big]}{(1+\sigma^2)\sqrt{\big[8\sigma_2\sigma_3 + 4\sigma_1(1-\sigma^2)\big]^2 + \big[-4\pi + (1-\sigma^2)^2\big]^2}} \\
&= \frac{8\sigma_3\big[(3-\sigma^2)\sigma_1 + (1-4\sigma_1^2)\big]}{(1+\sigma^2)\sqrt{\big[8\sigma_2\sigma_3 + 4\sigma_1(1-\sigma^2)\big]^2 + \big[-4\pi + (1-\sigma^2)^2\big]^2}} \\
&= H_2\big[2\sigma_3\big[(3-\sigma^2)\sigma_1 + (1-4\sigma_1^2)\big]\big] \qquad (6.25)
\end{aligned}$$

where

$$H_2 = \frac{4}{(1+\sigma^2)\sqrt{\big[8\sigma_2\sigma_3 + 4\sigma_1(1-\sigma^2)\big]^2 + \big[-4\pi + (1-\sigma^2)^2\big]^2}} \qquad (6.26)$$

In the same way, the other two elements of θ are computed.

$$\frac{\partial\theta}{\partial\sigma_2} = H_2\big[2\{\sigma_2\sigma_3(3-\sigma^2) + 2\sigma_1(1+\sigma^2-4\sigma_3^2)\}\big] \qquad (6.27)$$

$$\frac{\partial\theta}{\partial\sigma_3} = H_2\big[\sigma^4 + 6\sigma^2 - 2\sigma^2\sigma_2^2 - 8\sigma_1\sigma_2\sigma_3 - 1\big] \qquad (6.28)$$

ψ-Differentiations

To obtain the ψ differentiating elements, the element DCM[1,1] of both Euler and MRP parameterisations is selected and differentiated with respect to three σ elements, as follows:

$$\cos\theta\cos\psi = \frac{4(\sigma_1^2 - \sigma_2^2 - \sigma_3^2) + (1-\sigma^2)^2}{(1+\sigma^2)^2} \qquad (6.29)$$

Differentiating with respect to σ_1,

$$\begin{aligned}
-\cos\theta\sin\psi\frac{\partial\psi}{\partial\sigma_1} &= \underbrace{\frac{\partial}{\partial\sigma_1}\Big(\frac{4(\sigma_1^2-\sigma_2^2-\sigma_3^2)+(1-\sigma^2)^2}{(1+\sigma^2)^2}\Big)}_{N_3} \\
\frac{\partial\psi}{\partial\sigma_1} &= \frac{N_3}{-\cos\theta\sin\psi} \\
&= \frac{-N_3}{DCM[1,2]}
\end{aligned} \tag{6.30}$$

To compute N_3 using quotient rule, this time

$$\begin{aligned}
u_3 &= 4(\sigma_1^2-\sigma_2^2-\sigma_3^2)+(1-\sigma^2)^2 \\
v_3 &= (1+\sigma^2)^2 \\
\dot{u}_3 &= 4\sigma_1(1+\sigma^2) \\
\dot{v}_3 &= 4\sigma_1(1+\sigma^2)
\end{aligned}$$

therefore

$$\begin{aligned}
N_3 &= \frac{(1+\sigma^2)^2 4\sigma_1(1+\sigma^2)-[4(\sigma_1^2-\sigma_2^2-\sigma_3^2)+(1-\sigma^2)^2]4\sigma_1(1+\sigma^2)}{(1+\sigma^2)^4} \\
&= \frac{4\sigma_1}{(1+\sigma^2)^3}\big[(1+\sigma^2)^2-\{4(\sigma_1^2-\sigma_2^2-\sigma_3^2)+(1-\sigma^2)^2\}\big]
\end{aligned} \tag{6.31}$$

Substituting the above value of N_3 and DCM[1,2] in (6.30) would generate

$$\begin{aligned}
\frac{\partial\psi}{\partial\sigma_1} &= -\frac{\frac{4\sigma_1}{(1+\sigma^2)^3}\big[(1+\sigma^2)^2-\{4(\sigma_1^2-\sigma_2^2-\sigma_3^2)+(1-\sigma^2)^2\}\big]}{\frac{1}{(1+\sigma^2)^2}[8\sigma_1\sigma_2+4\sigma_3(1-\sigma^2)]} \\
&= \frac{4\sigma_1\big[(1+\sigma^2)^2-\{4(\sigma_1^2-\sigma_2^2-\sigma_3^2)+(1-\sigma^2)^2\}\big]}{(1+\sigma^2)[8\sigma_1\sigma_2+4\sigma_3(1-\sigma^2)]} \\
&= H_3\Big[2\sigma_1\big\{(\sigma^4+1)+2(\sigma^2-2\sigma_1^2)\big\}\Big]
\end{aligned} \tag{6.32}$$

where

$$H_3 = \frac{4}{(1+\sigma^2)[8\sigma_1\sigma_2+4\sigma_3(1-\sigma^2)} \tag{6.33}$$

In the same the other two elements of the ψ are computed,

$$\frac{\partial\psi}{\partial\sigma_2} = H_3\Big(-2\sigma_2\big[1+2(2\sigma_1^2-\sigma^2)\big]\Big) \tag{6.34}$$

$$\frac{\partial\psi}{\partial\sigma_3} = H_3\Big(-2\sigma_3\big[1+2(2\sigma_1^2-\sigma^2)\big]\Big) \tag{6.35}$$

With the help of these nine partial derivatives elements ($\frac{\partial\phi}{\partial\sigma_i}$, $\frac{\partial\theta}{\partial\sigma_i}$, $\frac{\partial\psi}{\partial\sigma_i}$), the output linearised model can be constructed.

6.3 Control Scheme Design

It is worthwhile to mention that in this work more emphasis has been made to obtain bounded error state (attitude) estimation of the spacecraft system in the event of measurement loss and not the control problem. However, for the sake of being able to obtain estimation results, a control system design is provided.

In the conventional design methods for the spacecraft systems, controllers use two parameters namely, angular velocity and attitude parameter – see e.g. [100] and [101]. However, an output feedback control law proposed in [2], has been considered here to stabilise the plant where it is assumed that $\dot{\sigma}$ and ω are not accessible to measure. The employed control scheme consists of two loops, an inner loop and an outer loop. The inner loop has a transfer function and the outer loop has a unity feedback path. The control law is summarised as follows:

$$\tau = T(\sigma)^T[S_p\tilde{\sigma}(t) - \sigma^*(t)], \tag{6.36}$$

where $T(\sigma)$ has been defined in Equation (6.1), τ is the control input and

$$\begin{aligned} &S_p = diag(s_{p1}, s_{p2}, s_{p3}),\\ &\tilde{\sigma} = \sigma_d(t) - \hat{\sigma}(t),\\ &\sigma^*(t) = N\sigma(t), \quad with\\ &N = diag\Big(s_{d1}\frac{\alpha_1 s}{s+\alpha_1}, s_{d2}\frac{\alpha_2 s}{s+\alpha_2}, s_{d1}\frac{\alpha_3 s}{s+\alpha_3}\Big) \end{aligned} \tag{6.37}$$

The positive definite matrices (S_p, N), in the control scheme are the design parameters. The candidate Lyapunov function is

$$V(\sigma^*, \dot{\sigma}, \tilde{\sigma}) = \frac{1}{2}\big(\dot{\sigma}^T H^* \dot{\sigma} + \tilde{\sigma}^T S_p \tilde{\sigma} + \sigma^{*T}\{\alpha S_d\}^{-1}\sigma^*\big) \tag{6.38}$$

where H^* and S_d are defined as

$$\begin{aligned} H^* &:= (T(\sigma)^{-1})^T J T(\sigma)^{-1} \quad \& \\ S_d &:= diag(s_{d1}, s_{d2}, s_{d3}) \end{aligned} \tag{6.39}$$

The time derivative of this Lyapunov function is found as

$$\begin{aligned}\dot{V} &= -\dot{\sigma}^T \sigma^* + \sigma^{*T}\{\alpha S_d\}^{-1}(S_d \alpha \dot{\sigma} - \alpha \sigma^*) \\ &= -\sigma^{*T} S_d^{-1} \sigma^* \end{aligned} \tag{6.40}$$

The selection of gain elements (s_{pi}, s_{di} and α_i) to stabilise the nonlinear plant could be any positive values as far as the convergence is concerned. The detailed discussion on the asymptotic stability associated with the proposed control law can be found in [2]. In the subsequent section, the two algorithms – the Open-Loop Kalman filtering (OLKF) and the Compensated Closed Loop Kalman filtering (CCLKF) techniques are applied along with the conventional Kalman filtering (without any data loss) scheme.

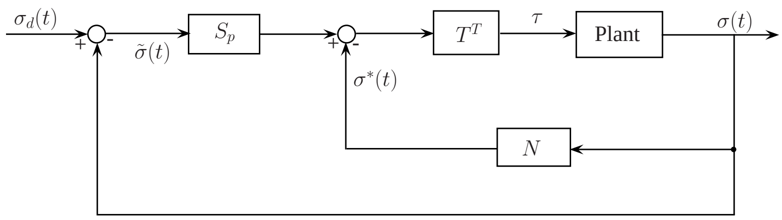

Figure 6.2: Attitude stabilisation using output feedback control

6.4 Compensated Kalman Filter - CCLKF

In this section the CCLKF approach, proposed in Chapter 4 is revised in a concise manner. Assuming the LOOB is detected at time step $= t$. The proposed CCLKF scheme employed, is summarised in the following way.

- **P**rediction cycle: The predicted state and covariance matrix at time step $(t-1)$ will be,

$$\begin{aligned} x_{t+1|t} &= Ax_{t|t} + Bu_t, \\ P_{t+1|t} &= AP_{t|t}A^T + Q_t, \end{aligned} \tag{6.41}$$

  where $Q_t := \mathbb{E}[\xi_t \xi_t^T]$ is the process error covariance matrix.

- **O**bservation vector

$$z_t = \gamma_t\big(Hx_t\big) + \theta_t, \tag{6.42}$$

  with γ_t characterised as follows:

$$\gamma_t = \begin{cases} 1; & \text{if No LOOB} \\ 0; & \text{otherwise} \end{cases} \tag{6.43}$$

- **C**heck for an observation loss:
  if $\gamma_t = 1 \rightarrow$ No LOOB has occurred.
  $\Rightarrow$ Run a conventional Kalman filter [21],
  if $\gamma_t = 0 \rightarrow$ An abnormal condition is detected (LOOB case)
  $\Rightarrow$ The actual observation is not available to which the prediction step is updated.
  $\Rightarrow$ Run the compensated Kalman filter, which is summarised below.

- **O**btain the LPFO i.e. p through constraint-based algorithms (Chapter 4)

- **C**ompute the autocorrelation matrix R_γ as

$$R_\gamma = \begin{bmatrix} R_\gamma[0] & R_\gamma[1] & R_\gamma[2] & \cdots & R_\gamma[n-1] \\ R_\gamma[1] & R_\gamma[0] & R_\gamma[1] & \cdots & R_\gamma[n-2] \\ R_\gamma[2] & R_\gamma[1] & R_\gamma[0] & \cdots & R_\gamma[n-3] \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R_\gamma[n-1] & R_\gamma[n-2] & R_\gamma[n-3] & \cdots & R_\gamma[0] \end{bmatrix}, \tag{6.44}$$

  and the modified autocorrelation array r_γ as

$$r_\gamma = \begin{bmatrix} r_\gamma[1] & r_\gamma[2] & r_\gamma[3] & \cdots & r_\gamma[n] \end{bmatrix}^T \tag{6.45}$$

  where

$$\begin{aligned} E\big[z(t-i)^T z(t-j)\big] &= \begin{cases} R_\gamma[0], & \text{if i = j} \\ R_\gamma\big[|i-j|\big], & \text{if i} \neq \text{j} \end{cases} \\ r_\gamma[j] &= E\big[z(t)^T z(t-j)\big] \end{aligned} \tag{6.46}$$

- **C**ompute the Linear Prediction Coefficients (LPC) as

$$A_\alpha = [\alpha_j]^T = R_\gamma^{-1} \cdot r_\gamma \tag{6.47}$$

- **C**alculate the compensated measurement vector as

$$\bar{z}_t = \sum_{j=1}^{p} \alpha_j z(t-j) \equiv H\bar{x}_t + \bar{\theta}_t \tag{6.48}$$

- **O**btain the compensated residual vector as $\bar{z}_t - \hat{z}_t$.

- **C**alculate the compensated Kalman filter gain ${}_cK_t = {}_cP_{t|t-1}H^T\big(H_cP_{t|t-1}H^T + R_t\big)^{-1}$.

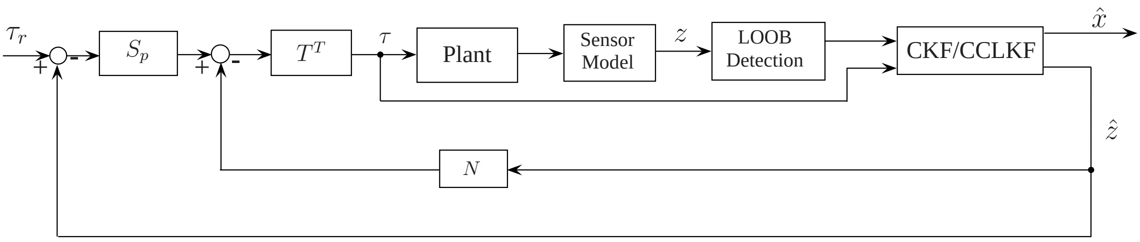

Figure 6.3: Compensated attitude estimation block diagram

- **M**easurement update step will proceed as:

$$
\begin{aligned}
{}_c x_{t|t} &= {}_c x_{t|t-1} + {}_c K_t \left[\bar{z}_t - H_c x_{t|t-1} \right] \\
{}_c P_{t|t} &= {}_c P_{t|t-1} - {}_c K_t H_c P_{t|t-1}
\end{aligned} \tag{6.49}
$$

- **R**eturn to Step 1 (prediction cycle).

2

6.5 Numerical Example

In this section, the two compensated state estimation techniques (OLKF and CCLKF) are applied to the rigid spacecraft model which is subjected to loss of measurements. The drawbacks of the OLKF approach along with the added advantages of the CCLKF scheme are illustrated through extensive simulations. Several typical simulation results are shown in this section including attitude estimation, angular velocity estimation, control effort and absolute error for the two techniques.

6.5.1 Spacecraft Model

The continuous spacecraft model is discretized according to Nyquist–Shannon sampling theorem. The non-zero equilibrium points in Equations (6.1) and (6.5) are computed off-line through Levenberg-Marguardt iterative least-square scheme. The plant model is linearised using Jacobian linearisation at these operating points to conclude the state space model – Appendix B provides more details on this matter.

[2] As mentioned in Chapter 4, at the very first moment of LOOB, the following inilisations are taken into account after the prediction step

$$
\begin{aligned}
{}_c x_{k|k-1} &= x_{k|k-1} \\
{}_c e_{k|k-1} &= e_{k|k-1} \\
{}_c P_{k|k-1} &= P_{k|k-1} \\
{}_c K_{k-1} &= K_{k-1}
\end{aligned}
$$

.

The mathematical description of the linearised spacecraft attitude model is described by the following state space model:

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + G\xi_t \\ z_t &= Hx_t + Du_t + \theta_t \end{aligned} \tag{6.50}$$

Jacobian Matrices for MRP representations :

where

$$A = \begin{bmatrix} -0.0160 & 0.0621 & 0.3567 & 0.2151 & 0.2087 & -0.0133 \\ -0.0621 & -0.0160 & 0.1462 & -0.2010 & 0.2010 & -0.0962 \\ -0.3567 & -0.1462 & -0.0160 & -0.0580 & 0.0779 & 0.2247 \\ 0 & 0 & 0 & 0.0465 & 0.0661 & 0.0234 \\ 0 & 0 & 0 & -0.0435 & 0.0163 & -0.1275 \\ 0 & 0 & 0 & -0.1921 & 0.0274 & -0.0628 \end{bmatrix}$$

$$B^T = \begin{bmatrix} 0 & 0 & 0 & 0.0503 & -0.0033 & -0.0027 \\ 0 & 0 & 0 & -0.0033 & 0.0595 & -0.0054 \\ 0 & 0 & 0 & -0.0027 & -0.0054 & 0.0673 \end{bmatrix}$$

$$G^T = \begin{bmatrix} -0.2151 & 0.2010 & 0.0580 & -0.0465 & 0.0435 & 0.1921 \\ -0.2087 & -0.2010 & -0.0779 & -0.0661 & -0.0163 & -0.0274 \\ 0.0133 & 0.0962 & -0.2247 & -0.0234 & 0.1275 & 0.0628 \end{bmatrix}$$

$$H = \begin{bmatrix} 2.7069 & -2.7731 & 0.3026 & 0 & 0 & 0 \\ 2.6196 & 2.5571 & -1.1132 & 0 & 0 & 0 \\ 1.1156 & -0.2864 & 2.6252 & 0 & 0 & 0 \end{bmatrix}; \; D = \mathbf{0}_{3\times 3};$$

The initial state vector is assumed as $x_0 = \begin{bmatrix} 0.2 & 0.2 & 0.2 & -0.3 & -0.4 & 0.2 \end{bmatrix}^T$. The operating points computed through Levenberg-Marguardt method are found at $u = [0.1284 \;\; 0.5767 \;\; 0.9365]^T$; $\sigma = [0.1741 \;\; 0.0447 \;\; -0.4097]^T$; $\omega = [0.4779 \;\; -0.4779 \;\; 0.2287]^T$.
The gain parameters are selected through simulink optimisation tool as follows: S_p=80; $S_d = 95$; $\alpha = 45$;

The plant noise covariance matrix and measurement noise covariance matrix are respectively given as

$$\begin{aligned} Q &= diag\{0.01,\ 0.01,\ 0.01,\ 0.01,\ 0.01,\ 0.01\} \\ R &= diag\{0.05,\ 0.05,\ 0.05\} \end{aligned}$$

Some representative simulation results using MRP representations are shown for the above linearised spacecraft model in the following subsection.

6.5.2 Simulation Results

In this section, a typical set of simulation results is provided for the rigid body spacecraft model discussed in the earlier sections subject to an induced loss of measurements. Simulation results are achieved in Matlab/Simulink environment using zero-order hold scheme with a sampling time of $T_s = 10$ ms.

The results for the proposed method (CCLKF) and existing Open-Loop approach (OLKF) are compared with those of the conventional Kalman filter (CKF) without any data loss. The loss of observation has been assumed to occur due to the sensor fault (or transmission channel failure) and remains for 15 s. After such a failure, the measurement reading is resumed. It is worthwhile to emphasis that duration of loss of observation (LOOB) is small enough to assure that the CCLKF is capable of reconstructing such data loss accurately. Nonetheless, such data losses could occur intermittently as it is the case in real-world applications.

Results of Modified Rodrigues Attitude Parameters

Figures 6.4 – 6.6 show the nominal state estimation results and illustrate the performance of the two approaches in terms of attitude parameters estimation ($\sigma_1 - \sigma_3$) with loss of observation (LOOB) occurred at all three channels. In this set of simulations, the LOOB is assumed to occur at time $t = 30$ s. From Figures 6.4 – 6.6, it can be seen that results produced by the OLKF diverge immediately from the nominal steady-state values during the loss of observation. In contrary, the proposed CCLKF approach is comparatively stable and does not deviate significantly in the event of data loss from the failure-free dynamic model. After the data is resumed, effectively less oscillations can be found in the simulation results for CCLKF approach than OLKF approach. From the three figures, it can be easily observed that the proposed CCLKF scheme significantly outperforms the OLKF approach.

It is also worthwhile to mention that deviation from the steady-state (or nominal value) is dependant on the length of the data loss period (failure duration).

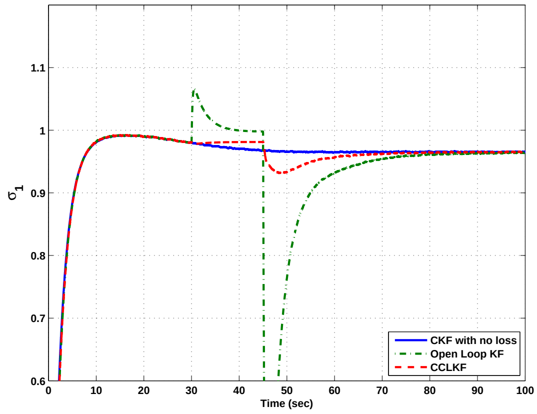

Figure 6.4: Simulation results of $\sigma_1(t)$. Loss of output data has been considered from time $t = 30$ s to $t = 45$ s. Thereafter, the observation is resumed.

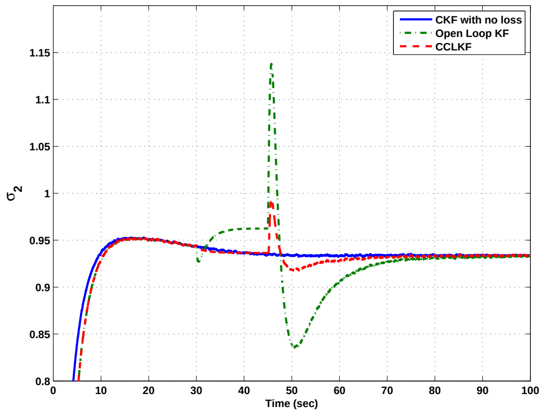

Figure 6.5: Simulation results of $\sigma_2(t)$. Loss of output data has been considered from $t = 30$ to $t = 45$ s. After that the observation is resumed.

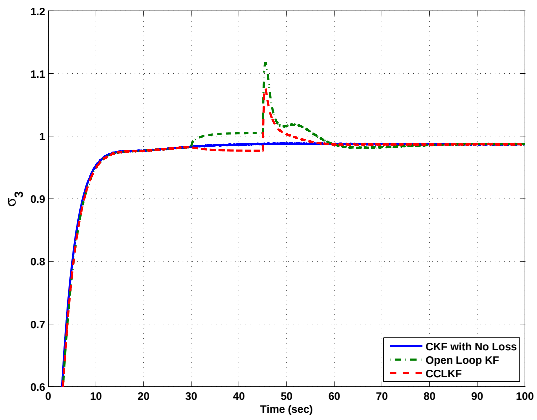

Figure 6.6: Simulation results of $\sigma_3(t)$. Loss of output data has been considered from $t = 30$ to $t = 45$ s. After that the observation is resumed.

Angular Velocity Results

In addition to attitude parameters, there are three states of angular velocities associated with the rigid body spacecraft model under investigation. Figures 6.7 – 6.9 show the distinctions of the two approaches in the event of LOOB as an index of angular velocity. It can be seen from these figures that, the unavailability of observation causes OLKF approach to be a poor solution for state estimation in the event of data loss. It is important to realise that there are abrupt changes and chattering, in the angular velocity estimation results of the OLKF approach. Nonetheless, the compensation-based measurement update cycle in the CCLKF approach provides considerable improvement in the angular velocity estimation results.

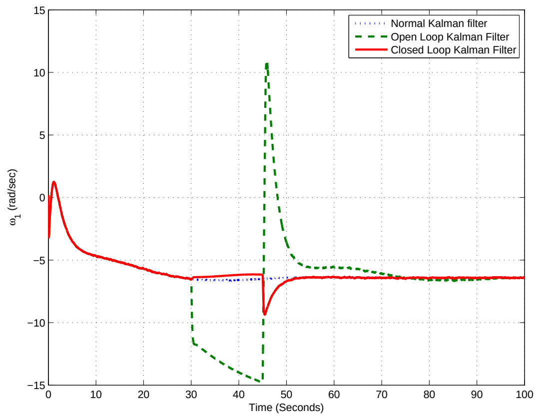

Figure 6.7: Simulation results of angular velocity $\omega_1(t)$. Loss of output data has been considered from $t = 30$ to $t = 45$ s only.

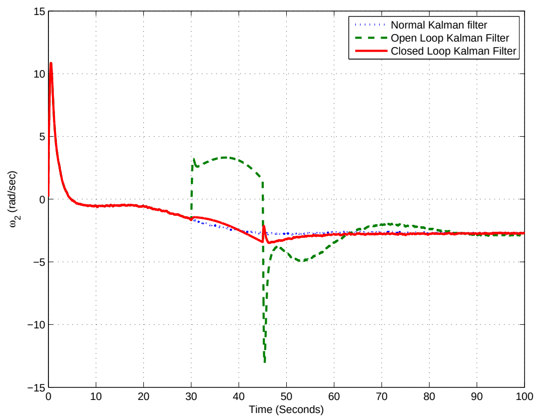

Figure 6.8: Simulation results of angular velocity $\omega_2(t)$. Loss of output data has been considered from $t = 30$ to $t = 45$ s only.

Control Effort Results

Another important concept considered is the control signal input which poses significant performance impacts. Due to the recursive behaviour of Kalman filter and output feedback control system, the effects of loss of observation transverses to the input parameters too as of the MRP entities and angular velocities. During the failure period, large overshoots were always observed when OLKF technique is applied in comparison to the CCLKF approach as evidenced in Figures 6.10 – 6.12.

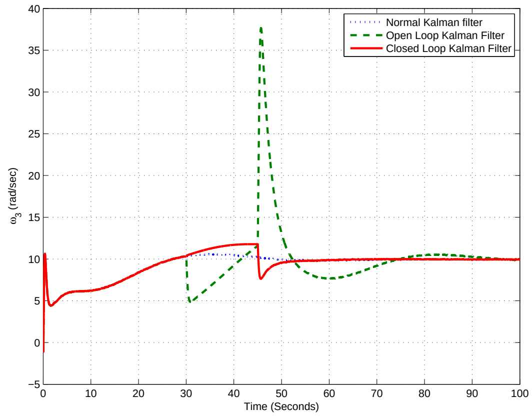

Figure 6.9: Simulation results of angular velocity $\omega_3(t)$. Loss of output data has been considered from $t = 30$ to $t = 45$ s only.

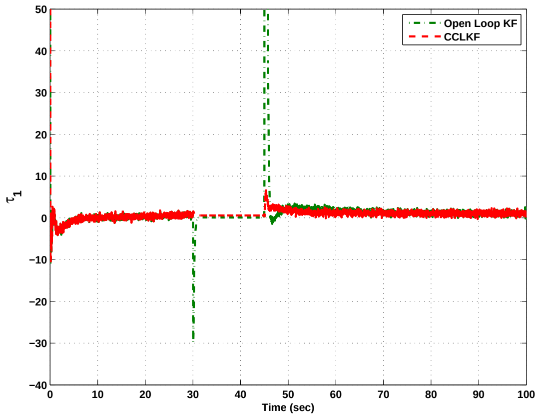

Figure 6.10: Control signal of τ_1 associated with output feedback controller.

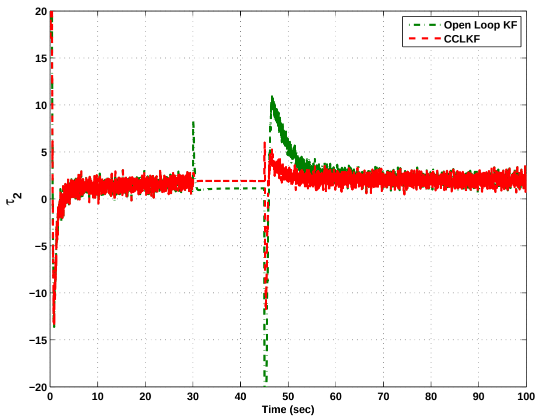

Figure 6.11: Control signal of τ_2 associated with output feedback controller.

Error Analysis Results

Another important characteristic property showing the efficiency of the proposed closed loop Kalman filtering approach compared to that of the OLKF approach is the integrated error generated, i.e. *A reference input - plant filtered output*. Most importantly, less chattering and

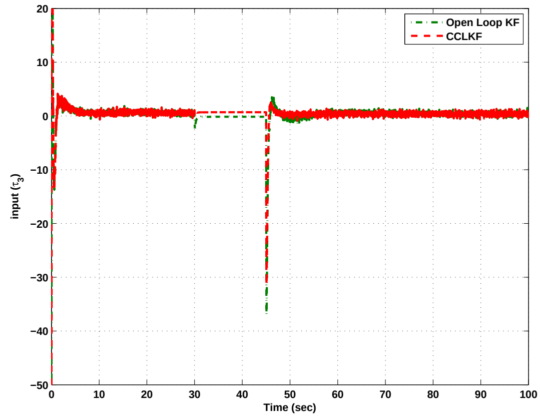

Figure 6.12: Control signal of τ_3 associated with output feedback controller.

disruption can be tracked in case of using CCLKF approach compared to those of the OLKF technique as shown in Figures 6.13 – 6.15.

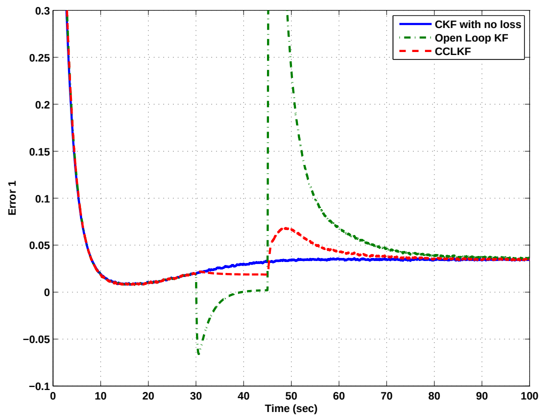

Figure 6.13: Error in output channel 1.

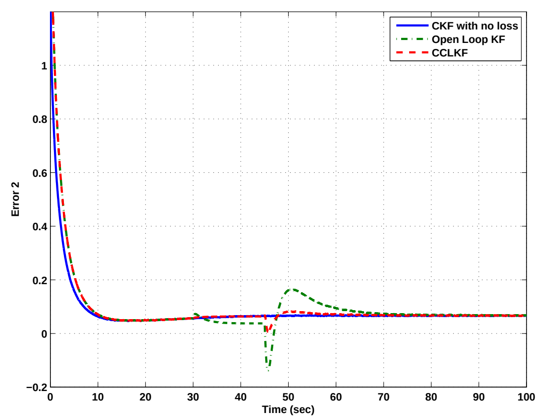

Figure 6.14: Error in output channel 2.

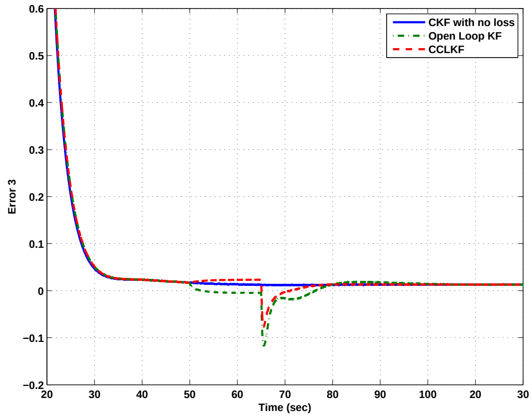

Figure 6.15: Error in output channel 3.

6.6 Summary

State estimation plays an important role in the attitude control of spacecraft systems. That's why, in this chapter, attitude estimation and control are considered for a rigid spacecraft model which is subjected to output measurement loss. The plant dynamics of the rigid body spacecraft system are obtained while utilising both kinematic and dynamics models. The output model is achieved in the modified Rodrigues parameterisations.

Conventional Kalman filter could fail in providing bounded attitude estimation error in the event of loss of measurement because of the unavailability of measurements at the *a posteriori* step. In order to handle this issue, two compensated estimation techniques known as Open-Loop Kalman filtering (OLKF) and compensated close loop Kalman filtering (CCLKF) schemes are implemented to the rigid spacecraft model, subjected to output data loss. The proposed (CCLKF) approach was successfully applied to the linearised model of the rigid body spacecraft to generate some promising results in such adverse conditions. The simulation results were compared with those of available Open-Loop Kalman filter estimation approach along with the conventional KF without any data loss. Simulation results based on MRP representations are shown for various performance indices. Various properties were illustrated and discussed through a numerical example. These results give comprehensive comparison between the two OLKF and CCLKF approaches from which the performances can be easily analysed.

Chapter 7

Conclusion and Future Work

7.1 Conclusion

In this thesis, innovative Kalman filtering techniques are presented, which are found very efficient and vigorous towards loss of output data. The problem of loss of output data or observation can be frequently encountered in many control and communication systems. Such issue may be caused by many undesirable factors such as temporary faults and failures, congestion of communication channels, limited space of buffer registers, transmission error, limited spectrum, time varying channel gains, interference. The application of conventional Kalman filter requires the knowledge of plant dynamics, information of unmeasured stochastic inputs, and measured data. For example, in Kalman filtering technique, the state of a system is predicted through system's dynamics information and input signal and thereby updated through measurement (output) data. Hence, in the absence of any of these information, conventional Kalman filter would not be able to generate correct (bounded error) state estimation.

The prevailed remedy for such problem has remained the Open-Loop Kalman filtering (OLKF) approach. In OLKF technique, the state of a system is predicted only during the data loss time interval and is processed to the next time instant without any update. In other words, in OLKF scheme, the *prediction* phenomenon is considered as *filtering*. OLKF is a fast and simple way of handling loss of data in the state estimation. In practice, this technique takes even lesser time for computation than any other solution. However, a number of displeasing factors have made OLKF technique an improper solution for loss of data in the Kalman filtering. Some of the factors are listed as follows:

- Inability of OLKF scheme to produce optimal estimation. The estimation error and hence error covariance could easily exceed the upper bounds provided for a system. In other words, this technique is found very sensitive to the loss of data in terms of divergence. It is important to mention that the performance of OLKF is much better than conventional KF subjected to data loss which does not take into account any remedy towards data loss.

The reason being is that CKF considers false residual ($r_k = 0 - H\hat{x}_k$) calculation in the event of data loss instead of actual residual ($r_k = z_k - H\hat{x}_k$) due to the unavailability of the measurement data, consequently results in a huge estimation error in a very short interval.

- The severity of oscillations and chattering. This problem can be observed at the start and/or end of data loss times. The possible reason behind this problem is the sudden cease of the update-cycle (when the data loss starts) and immediate start of the update-cycle (when the data is resumed after the loss). Therefore, depending upon the system dynamics, large oscillations can be observed at the start and/or end times of data loss.

- KF is a recursive state estimation method i.e. the update step is performed on the predicted quantities (state and covariance) and the prediction step is based on the updated quantities. Hence, any deficit at one step would affect the state and covariance at the other step. The recursive behaviour of the Kalman filter is shown in Figure 7.1. Due to this property, OLKF technique could not attain the steady state values when the observation is resumed after data loss. In fact, theoretically it takes infinite time to achieve steady state (stable) values.

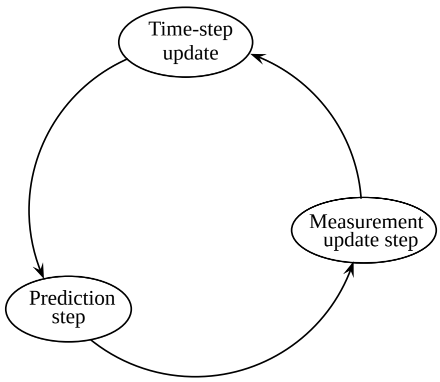

Figure 7.1: Recursive behaviour of Kalman filter

Open-Loop Kalman filtering scheme, also known as Open-Loop estimation is thoroughly studied in Chapter 3. A few associated features with OLKF technique are also summarised to aid with the selection of this approach in the event of data loss. Another way to perform the measurement update step in Kalman filtering in the event of data loss is to record the last sensor reading all the time. The technique is known as Zero Order Hold. Chapter 3 is ended by an overview for Zero Order Hold technique.

The primary objective of this thesis is to overcome the shortcomings associated with OLKF technique and provide efficient estimation performance in the event of data loss. Towards this end, linear prediction techniques are employed to construct a signal from its previous data samples. Two types of linear prediction techniques are in common use; internal (or backward) linear prediction and external (or forward) linear prediction techniques. Internal linear prediction method is a system identification problem whereby given signal coefficients are computed through autocorrelation theory. On the other hand, external linear prediction theory is purely a prediction concept. Therefore, loss of observation can be tackled with external linear prediction technique with some suitable amendments.

In Chapter 4, the data loss has been reconstructed through the external linear prediction technique. Conventionally, there is no distinct method to decide the number of previous data samples which could optimally reproduce the lost observations. For this purpose, several effective algorithms are presented. In addition solving linear prediction schemes in the routine methodology involves inversion of huge matrices which dimensions depend on the order of linear prediction filter. The higher the order of linear prediction filter, the more computational time it would take. In order to handle this problem, the Levinson-Durbin algorithm, as opposed to Leorux-Gueguen and Schur algorithms, has been adopted due to its simpler structure and distinct methodology in the thesis.

The compensated observation signal has been utilised to reformulate the conventional Kalman filter in Chapter 4. Straightforward procedure has been adopted to obtain the optimal Kalman filter gain and error covariance matrices using compensated observation signal. Different topologies are briefly introduced on the ground of various compensated signal's contributions. During the data loss time period, gain and error covariance matrices have extra terms caused by loss of observations. However, once the actual observation signal is resumed, the structures of these matrices are reduced to the normal ones as is the case of conventional KF. Related discussion including features and shortcomings with the proposed schemes have been described.

Having an adequate derivations of Kalman filter design while employing compensated observational signal at measurement update step, it was necessary to test the properties of this compensated KF (CCLKF) design in the situation of data loss. For this purpose, Chapter 5 has been approached with the theoretical implementation of the CCLKF algorithm on a mass-spring-damper case study. Through this test, majority of the possible characteristic properties associated with KF are explored in terms of CCLKF approach such as update state estimation and error covariance, Kalman filter gain and residual analysis. In order to investigate the efficiency, the properties explored for CCLKF approach are compared with those of Open-Loop KF approach. Computational analysis for the CCLKF has been carried out for two scenario; first when computing linear prediction coefficients through routine methodology using autocorrelation matrix technique and secondly employing Levinson-Durbin algorithm. The former is found very expensive in terms of computational time while the latter is computationally efficient particularly for higher orders of linear prediction filter.

One of the widely used application area of Kalman filtering is spacecraft technology. Attitude estimation in spacecraft system involves sensors such as rate gyros and accelerometers which upon any fault or failure would cause unavailability of output data. This might lead to adverse situations in attitude estimation and control. For this reason, attitude estimation for a rigid body spacecraft system in the event of output data loss in Modified Rodrigues Parameters representations has been considered in Chapter 6. The plant dynamics are obtained through Euler equations of rotational dynamics and Kinematic equations. The output dynamic model is obtained through sensor in terms of Euler angles. Due to a number of associated shortcomings, Euler parameters are converted into Modified Rodrigues Parameterisations (MRP). The conversion of output model from Euler angles to MRP representations has been achieved by equating the two DCM associated with Euler angles and MRP representations.

The nonlinear rigid body spacecraft system is linearised on nonzero operating points. The original system is unstable, therefore for stabilisation, a Lyapunov based control law has been analysed. Simulation results are described for various indeces such as attitude parameters, angular velocity, control input and error signal. These results have shown the proposed compensated KF approach in the event of data loss significantly outperforms the existing Open-Loop KF approach.

There are some drawbacks associated with the proposed compensated close loop Kalman filtering approaches, which are mentioned below:

- The main drawback allied with the proposed scheme is its computational time. This is because, during data loss time, previously stored data is recalled and optimal linear prediction coefficients are calculated for these data samples through some techniques (e.g. autocorrelation method), which in turn requires extra time while computing state estimations. The computational time is directly related to the number of data samples recalled (order of linear prediction filter) for reconstructing this compensated observations signal. However, by employing some techniques such as Levinson-Durbin algorithm, this issue has been tackled to some extent.

- The second problem of the CCLKF approach is its hardware requirement to store previous observations (especially for multi-input multi-output systems). At every data loss time instant, p-number of data samples need to be retreated in order to calculate the compensated observation, which leads to a need of buffer register that could store and update these p data samples. 'p' is the order of linear prediction filter.

- The third drawback associated with the proposed approaches is that the repossessing of steady state values (estimated state and error covariance) is not fully attained. However, this shortcoming is much improved in comparison with the Open-Loop KF approach.

- The proposed schemes consider auto-regressive model (i.e. previous observation data) when computing the linear prediction coefficients (LPC), i.e.

$$\bar{z}_k = \sum_{j=1}^{p} \alpha_j z_{k-j} \tag{7.1}$$

It does not take into account the control input (i.e. ARMA process) in order to simplify the matter.

$$\bar{z}_k = \sum_{j=1}^{p} \alpha_j z_{k-j} + \sum_{i=0}^{N} b_i u_{k-i} \tag{7.2}$$

However, this might invalidate the proposed approaches if the control signals have abrupt changes during the data loss period. This feature can be explicitly related to the assumption of slowly time varying characteristics of the process made in Chapter 4 in the derivation of LPC.

There will always be a trade-off between the listed shortcomings and the related advantages of the proposed CCLKF approach depend upon various factors such as amount of data loss, the nature of system under considerations and the value of computational efforts. For example, the above mentioned additional requirements can be compromised for a complex and bulky spacecraft system in order to achieve a successful completion of a mission in the event of loss of measurements, caused by any undesirable conditions.

7.2 Future Work

Five tasks have been proposed for future work of this thesis. The brief discussion of these tasks is as follows:

Linear state estimation tools have been adopted to analyse the linear time invariant systems. In future, nonlinear state estimation techniques such as Extended-Kalman filter (EKF) and Unscented Kalman filter (UKF) are intended to be applied, with the latter being the latest addition to the techniques and would best suit nonlinear systems. Through some more mathematical expressions, the compensated Kalman filtering schemes are intended to be derived in terms of EKF and UKF structures. The implementation of EKF is expected to be an easy task due to its great resemblance with the conventional Kalman filter, implemented in Chapters 5 and 6.

Insufficient data on nonlinear models of a rigid body spacecraft and Uninhabited Autonomous Vehicles (UAV) systems are planned to be looked into. In a networked system, where various UAV agents perform to a combined task, loss of data will be considered and nonlinear state estimation techniques will be employed.

Linear parameter varying (LPV) systems are emerging research area in many practical applications. It is also an intention to extend the work of this thesis to LPV systems. In order to achieve this goal, initially a straightforward Kalman filtering scheme will be designed for a simple LPV system with bounded parameter variations. Thereafter, the proposed compensated estimation technique will be modified to cope an LPV application with an induced fault or failure.

Likewise the successful implementation of Constraint Levinson-Durbin algorithm to reduce computational burdensome of the proposed CCLKF approach, other existing methods of Leroux-Geuguen [14] and Sch*ü*r algorithms [83] are required to be amended and updated accordingly to verify and improve the efficiency of the proposed algorithms. This task is expected to give justification of employing any scheme to reduce the computational efforts.

It is also intended to perform some quantitative analysis for the stability of the CCLKF algorithm with given loss period (in time index) and estimation (upper/lower) error bounds. It should be interesting to establish some relationship between the amount of data loss and a given error bounds.

By all means, various unforseen and practical implications related to implementation of the proposed algorithms need to be explored.

Appendix A

Theory of Levinson-Durbin Algorithm

This appendix is meant for providing sufficient details of Levinson-Durbin Algorithm (LDA). Theory for LDA can be found in plenty of text books such as [14, 83, 84].

A simple autoregressive (AR) process can be represented as:

$$z_t = \sum_{j=1}^{p} \alpha_j z_{t-j} + v_t \tag{A-1}$$

where $\alpha_j = [\alpha_1, \alpha_2, ...\alpha_p]$ are the parameters of the AR model, v_t is a noise signal and p is the window size. The optimal values of α_i are computed through auto-correlation function using Yule-Walker equation, which is

$$R_m = \sum_{j=1}^{p} \alpha_j R_{m-j} + \sigma^2 \delta_m \tag{A-2}$$

where δ_m is kroncker delta. It is characterised as follows:

$$\delta_m = \begin{cases} 1 & \text{if m = 0} \\ 0 & \text{otherwise} \end{cases} \tag{A-3}$$

Equation (A-2) is obtained by multiplying both sides of (A-1) by z_{t-m} and then taking the expectation. i.e.

$$E[z_t z_{t-m}] = E\Big[\sum_{j=1}^{p} \alpha_j z_{t-j} z_{t-m}\Big] + E\big[v_t z_{t-m}\big] \tag{A-4}$$

Equation (A-4) is in-fact a set of p linear equations with p unknowns $\alpha_j = \{\alpha_1, \alpha_2, ...\alpha_p\}$ and is known as *Normal Equations*. Equation (A-2) for $m = 0$ can be written as

$$
\begin{aligned}
J_m &= R[0] - \sum_{j=1}^{p} \alpha_j R[j] \\
&= R[0] + \mathbf{a}\mathbf{R}_\gamma
\end{aligned} \tag{A-5}
$$

where $J_m = \sigma^2$, $\mathbf{a} = -\{\alpha_1, \alpha_2, \ldots, \alpha_p\}$ and $\mathbf{R}_\gamma^T = \big\{R[1], R[2], \ldots, R[p]\big\}$.
From the theory of linear prediction, to compute linear prediction coefficients through *Normal Equations*, differentiating Equation (A-4) with respect to α_j the following equation can be deduced (Equation-4.15):

$$
\begin{aligned}
\mathbf{r}_\gamma &= -\mathbf{a}\mathbf{R}_z \\
\mathbf{r}_\gamma + \mathbf{a}\mathbf{R}_z &= 0
\end{aligned} \tag{A-6}
$$

where

$$
R_z = \begin{bmatrix} R[0] & R[1] & \cdots & R[p-1] \\ R[1] & R[0] & \cdots & R[p-2] \\ \vdots & \vdots & \ddots & \cdots \\ R[p-1] & R[p-2] & \cdots & R[0] \end{bmatrix} \tag{A-7}
$$

is $p \times p$ autocorrelation matrix, and

$$
r_\gamma = E[z_k^T z_{k-i}] \tag{A-8}
$$

is autocorrelation array for $i = \{1, 2, \ldots, p\}$.

Augmenting Equations (A-5) and (A-6) would result in

$$
\begin{bmatrix} R[0] & \mathbf{R}_\gamma^T \\ \mathbf{r}_\gamma & \mathbf{R}_z \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{a} \end{bmatrix} = \begin{bmatrix} J_m \\ \mathbf{0} \end{bmatrix} \tag{A-9}
$$

The extended version of Equation (A-9) without any subscript would be

$$
\begin{bmatrix} R[0] & R[1] & \cdots & R[p] \\ R[1] & R[0] & \cdots & R[p-1] \\ \vdots & \vdots & \ddots & \vdots \\ R[p] & R[p-1] & \cdots & R[0] \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} J_m \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{A-10}
$$

Keeping in mind that the aim is to compute the optimal values of the LPC, provided the autocorrelation values of $R[l]$ with $l = \{0, 1, 2, \ldots, p\}$ and J is the minimum mean square error. Levinson-Durbin algorithm (LDA), finds the solution to the p^{th} order predictor from the $(p-1)^{th}$ order predictor. In simple words, LDA is an iterative-recursive algorithm where initially, solution of zero-order predictor is computed, which is then utilised to find the solution of the first order predictor. This procedure is performed as follows:

Zero Order Predictor (LPFO = 0)

Initialising

$$R[0] = J_0 \tag{A-11}$$

This equation can be achieved from (A-5) with zero lag which in fact is the variance of the signal it self. Equation (A-11) can be extended with some modifications as

$$\begin{bmatrix} R[0] & R[1] \\ R[1] & R[0] \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} J_0 \\ \Delta_0 \end{bmatrix} \tag{A-12}$$

Equation (A-12) can be obtained from equation (A-10) by substituting $a_1 = 0$. However, this substitution violates the optimal condition of computing LPC, hence to compensate the effect, an additional term Δ_0 is introduced. The following result can be inferred from Equation (A-12);

$$\Delta_0 = R[1] \tag{A-13}$$

From the property of Toeplitz matrix, the above equation can be written as

$$\begin{bmatrix} R[0] & R[1] \\ R[1] & R[0] \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \Delta_0 \\ J_0 \end{bmatrix} \tag{A-14}$$

First Order Predictor (LPFO = 1)

The next step of LDA takes the values, computed in the zero order predictor as follows:

$$\underbrace{\begin{bmatrix} R[0] & R[1] \\ R[1] & R[0] \end{bmatrix}}_{R_z[1]} \begin{bmatrix} 1 \\ a_1^{(1)} \end{bmatrix} = \begin{bmatrix} J_1 \\ 0 \end{bmatrix} \tag{A-15}$$

where $a_1^{(1)}$ is the LPC computed through first order LDA, and J_1 represents the minimum mean square error for first order predictor. To calculate these two unknowns ($a_1^{(1)}$ and J_1), considering, the solution of the form

$$\begin{bmatrix} 1 \\ a_1^{(1)} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \kappa_1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{A-16}$$

where κ_1 is a constant and known as reflection coefficient. Multiplying both sides with the correlation matrix $R_z[1]$ would result in

$$\begin{bmatrix} R[0] & R[1] \\ R[1] & R[0] \end{bmatrix} \begin{bmatrix} 1 \\ a_1^{(1)} \end{bmatrix} = \begin{bmatrix} R[0] & R[1] \\ R[1] & R[0] \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \kappa_1 \begin{bmatrix} R[0] & R[1] \\ R[1] & R[0] \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{A-17}$$

From Equations (A-12), (A-14) and (A-15)

$$\begin{bmatrix} J_1 \\ 0 \end{bmatrix} = \begin{bmatrix} J_0 \\ \Delta_0 \end{bmatrix} - \kappa_1 \begin{bmatrix} \Delta_0 \\ J_0 \end{bmatrix} \tag{A-18}$$

which implies

$$0 = \Delta_0 - \kappa_1 J_0 \tag{A-19}$$

or

$$\kappa_1 = \frac{\Delta_0}{J_0} = \frac{R[1]}{J_0} = \frac{R[1]}{R[0]} \tag{A-20}$$

From Equation (A-16), the first LPC can be achieved as:

$$a_1^{(1)} = -\kappa_1 \tag{A-21}$$

The cost function for the first order of the LPF is updated from Equation (A-18) as:

$$\begin{aligned} J_1 &= J_0 - \kappa_1 \Delta_0 \\ &= J_0 - \kappa_1(\kappa_1 J_0) \\ &= J_0 - \kappa_1^2 J_0 \\ &= J_0(1 - \kappa_1^2) \end{aligned} \tag{A-22}$$

Equation A-12 is extended as:

$$\begin{bmatrix} R[0] & R[1] & R[2] \\ R[1] & R[0] & R[1] \\ R[2] & R[1] & R[0] \end{bmatrix} \begin{bmatrix} 1 \\ a_1^{(1)} \\ 0 \end{bmatrix} = \begin{bmatrix} J_1 \\ 0 \\ \Delta_1 \end{bmatrix} \tag{A-23}$$

where Δ_1 is included for compensation purpose, which can be found as

$$\Delta_1 = R[2] + a_1^{(1)} R[1] \tag{A-24}$$

Through the property of Toeplitz matrix, Equation (A-23) can be written as

$$\begin{bmatrix} R[0] & R[1] & R[2] \\ R[1] & R[0] & R[1] \\ R[2] & R[1] & R[0] \end{bmatrix} \begin{bmatrix} 0 \\ a_1^{(1)} \\ 1 \end{bmatrix} = \begin{bmatrix} \Delta_1 \\ 0 \\ J_1 \end{bmatrix} \tag{A-25}$$

Second Order Predictor (LPFO = 2)

Following the previous footsteps, the extended version of first order predictor will be

$$\underbrace{\begin{bmatrix} R[0] & R[1] & R[2] \\ R[1] & R[0] & R[1] \\ R[2] & R[1] & R[0] \end{bmatrix}}_{R_z[2]} \begin{bmatrix} 1 \\ a_1^{(2)} \\ a_2^{(2)} \end{bmatrix} = \begin{bmatrix} J_2 \\ 0 \\ 0 \end{bmatrix} \tag{A-26}$$

$a_1^{(2)}$, $a_2^{(2)}$ and J_2 are unknowns in the above equation. Assuming, a proposed solution of the form

$$\begin{bmatrix} 1 \\ a_1^{(2)} \\ a_2^{(2)} \end{bmatrix} = \begin{bmatrix} 1 \\ a_1^{(1)} \\ 0 \end{bmatrix} - \kappa_2 \begin{bmatrix} 0 \\ a_1^{(1)} \\ 1 \end{bmatrix} \tag{A-27}$$

where κ_2 is the reflection coefficient for 2^{nd} order LPF. Multiplying the above equation with $R_z[2]$ and the simplifying would result in

$$\begin{bmatrix} J_2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} J_1 \\ 0 \\ \Delta_1 \end{bmatrix} - \kappa_2 \begin{bmatrix} \Delta_1 \\ 0 \\ J_1 \end{bmatrix} \tag{A-28}$$

From Equations (A-24) and (A-28), the second reflection coefficient can be deduced as

$$\kappa_2 = \frac{1}{J_1}(R[2] + a_1^{(1)} R[1]) \qquad\qquad (\kappa_2 = \frac{\Delta_1}{J_1}) \tag{A-29}$$

From Equation (A-27)

$$\begin{aligned} a_2^{(2)} &= -\kappa_2 \\ a_1^{(2)} &= a_1^{(1)} - \kappa_2 a_1^{(1)} \end{aligned} \tag{A-30}$$

The cost function is obtained for the next step as (from Equations (A-28) and (A-29))

$$J_2 = J_1(1 - \kappa_2^2) \tag{A-31}$$

Third Order Predictor (LPFO = 3)

For the purpose of summarising the results in a concise manner, only the final results are shown for the 3^{rd} order of LPF as the derivation is straightforward.

$$\kappa_3 = \frac{1}{J_2}(R[3] + a_1^{(2)}R[2] + a_2^{(2)}R[1]) \tag{A-32}$$

with

$$\begin{aligned} a_3^{(3)} &= -\kappa_3 \\ a_2^{(3)} &= a_2^{(2)} - \kappa_3 a_1^{(2)} \\ a_1^{(3)} &= a_1^{(2)} - \kappa_3 a_2^{(2)} \end{aligned} \tag{A-33}$$

and

$$J_3 = J_2(1 - \kappa_3^2) \tag{A-34}$$

In conventional Levinson-Durbin method, this process is repeated till the final assigned order for Linear prediction filter is reached. The summary of this approach is described below in Algorithm 7.

Algorithm 7 : Levinson Durbin Algorithm (LDA)

1: Initialisation $l = 0$; $J_0 = R[0]$

2: Recursion $l = \{1, 2, 3....p\}$ order of LPF

3.1 Compute Reflection Coefficients as $\kappa_l = \frac{1}{J_{l-1}}\left[R[l] + \sum_{j=1}^{p-1} a_j^{l-1} R[l-j]\right]$

3.2 Calculate LPCs for l^{th} order predictor as
$a_l^l = -\kappa_l$
$a_j^l = a_j^{l-1} - \kappa_l a_{l-j}^{l-1}\ \forall\ j = \{1, 2, \ldots, l-1\}$

3.3 **I**s $l = p$,
Yes; Stop Computation of LPC
Process $a_l^l\ \forall\ l = \{1, 2, ...p\}$
No; Compute cost function associated with the last l^{th} predictor by $J_l = J_{l-1}(1 - \kappa_l^2)$
Update $l \leftarrow l + 1$ repeat Step 3.1

3: The resultant linear prediction coefficients are $a_j = a_j^{(p)}$ where $j = \{1, 2, \ldots, p\}$

Appendix B

Linearisation of A Rigid Body Spacecraft Model in Modified Rodrigues Parameterisations

The kinematic equations in Modified Rodrigues parameterisations are as follows:

$$\dot{\sigma} \overset{\text{def}}{=} T(\sigma)\overline{\omega} \tag{B-1}$$

where $\sigma_{3\times 1}$ is the modified rodrigous parameter vector,

$$T(\sigma) = \frac{1}{2}\left[\left(\frac{1-\sigma^T\sigma}{2}\right) I_{3\times 3} + S(\sigma) + \sigma\sigma^T\right] \tag{B-2}$$

The dynamic equation is as follows:

$$J\dot{\overline{\omega}} = -S(\overline{\omega})J\overline{\omega} + \tau \tag{B-3}$$

Expanding Equations (B-1) and (B-3) will result in

$$f_1 \overset{\text{def}}{=} \dot{\sigma}_1 = \frac{1}{4}\left[\overline{\omega}_1(1+\sigma_1^2-\sigma_2^2-\sigma_3^2)\right] + \frac{1}{2}\left[\overline{\omega}_2(\sigma_1\sigma_2-\sigma_3)\right] + \frac{1}{2}\left[\overline{\omega}_3(\sigma_1\sigma_3+\sigma_2)\right] \tag{B-4}$$

$$f_2 \overset{\text{def}}{=} \dot{\sigma}_2 = \frac{1}{2}\left[\overline{\omega}_1(\sigma_1\sigma_2+\sigma_3)\right] + \frac{1}{4}\left[\overline{\omega}_2(1-\sigma_1^2+\sigma_2^2-\sigma_3^2)\right] + \frac{1}{2}\left[\overline{\omega}_3(\sigma_2\sigma_3-\sigma_1)\right] \tag{B-5}$$

$$f_3 \overset{\text{def}}{=} \dot{\sigma}_3 = \frac{1}{2}\left[\overline{\omega}_1(\sigma_1\sigma_3-\sigma_2)\right] + \frac{1}{2}\left[\overline{\omega}_2(\sigma_2\sigma_3+\sigma_1)\right] + \frac{1}{4}\left[\overline{\omega}_3(1-\sigma_1^2-\sigma_2^2+\sigma_3^2)\right] \tag{B-6}$$

$$
\begin{aligned}
f_4 \stackrel{\text{def}}{=} \dot{\overline{\omega}}_1 \quad = \quad & a[\overline{\omega}_3(J_{21}\overline{\omega}_1 + J_{22}\overline{\omega}_2 + J_{23}\overline{\omega}_3) - \overline{\omega}_2(J_{31}\overline{\omega}_1 + J_{32}\overline{\omega}_2 + J_{33}\overline{\omega}_3)] \\
& -b[\overline{\omega}_3(J_{11}\overline{\omega}_1 + J_{12}\overline{\omega}_2 + J_{13}\overline{\omega}_3) + \overline{\omega}_1(J_{31}\overline{\omega}_1 + J_{32}\overline{\omega}_2 + J_{33}\overline{\omega}_3)] \\
& +c[\overline{\omega}_2(J_{11}\overline{\omega}_1 + J_{12}\overline{\omega}_2 + J_{13}\overline{\omega}_3) - \overline{\omega}_1(J_{21}\overline{\omega}_1 + J_{22}\overline{\omega}_2 + J_{23}\overline{\omega}_3)] \\
& +a\tau_1 + b\tau_2 + c\tau_3
\end{aligned} \tag{B-7}
$$

$$
\begin{aligned}
f_5 \stackrel{\text{def}}{=} \dot{\overline{\omega}}_2 \quad = \quad & d[\overline{\omega}_3(J_{21}\overline{\omega}_1 + J_{22}\overline{\omega}_2 + J_{23}\overline{\omega}_3) - \overline{\omega}_2(J_{31}\overline{\omega}_1 + J_{32}\overline{\omega}_2 + J_{33}\overline{\omega}_3)] \\
& -e[\overline{\omega}_3(J_{11}\overline{\omega}_1 + J_{12}\overline{\omega}_2 + J_{13}\overline{\omega}_3) + \overline{\omega}_1(J_{31}\overline{\omega}_1 + J_{32}\overline{\omega}_2 + J_{33}\overline{\omega}_3)] \\
& +f[\overline{\omega}_2(J_{11}\overline{\omega}_1 + J_{12}\overline{\omega}_2 + J_{13}\overline{\omega}_3) - \overline{\omega}_1(J_{21}\overline{\omega}_1 + J_{22}\overline{\omega}_2 + J_{23}\overline{\omega}_3)] \\
& +d\tau_1 + e\tau_2 + f\tau_3
\end{aligned} \tag{B-8}
$$

$$
\begin{aligned}
f_6 \stackrel{\text{def}}{=} \dot{\overline{\omega}}_3 \quad = \quad & g[\overline{\omega}_3(J_{21}\overline{\omega}_1 + J_{22}\overline{\omega}_2 + J_{23}\overline{\omega}_3) - \overline{\omega}_2(J_{31}\overline{\omega}_1 + J_{32}\overline{\omega}_2 + J_{33}\overline{\omega}_3)] \\
& -h[\overline{\omega}_3(J_{11}\overline{\omega}_1 + J_{12}\overline{\omega}_2 + J_{13}\overline{\omega}_3) + \overline{\omega}_1(J_{31}\overline{\omega}_1 + J_{32}\overline{\omega}_2 + J_{33}\overline{\omega}_3)] \\
& +i[\overline{\omega}_2(J_{11}\overline{\omega}_1 + J_{12}\overline{\omega}_2 + J_{13}\overline{\omega}_3) - \overline{\omega}_1(J_{21}\overline{\omega}_1 + J_{22}\overline{\omega}_2 + J_{23}\overline{\omega}_3)] \\
& +g\tau_1 + h\tau_2 + i\tau_3
\end{aligned} \tag{B-9}
$$

$$
A = \begin{bmatrix}
\frac{\partial f_1}{\partial \sigma_1} & \frac{\partial f_1}{\partial \sigma_2} & \frac{\partial f_1}{\partial \sigma_3} & \frac{\partial f_1}{\partial \overline{\omega}_1} & \frac{\partial f_1}{\partial \overline{\omega}_2} & \frac{\partial f_1}{\partial \overline{\omega}_3} \\
\frac{\partial f_2}{\partial \sigma_1} & \frac{\partial f_2}{\partial \sigma_2} & \frac{\partial f_2}{\partial \sigma_3} & \frac{\partial f_2}{\partial \overline{\omega}_1} & \frac{\partial f_2}{\partial \overline{\omega}_2} & \frac{\partial f_2}{\partial \overline{\omega}_3} \\
\frac{\partial f_3}{\partial \sigma_1} & \frac{\partial f_3}{\partial \sigma_2} & \frac{\partial f_3}{\partial \sigma_3} & \frac{\partial f_3}{\partial \overline{\omega}_1} & \frac{\partial f_3}{\partial \overline{\omega}_2} & \frac{\partial f_3}{\partial \overline{\omega}_3} \\
\frac{\partial f_4}{\partial \sigma_1} & \frac{\partial f_4}{\partial \sigma_2} & \frac{\partial f_4}{\partial \sigma_3} & \frac{\partial f_4}{\partial \overline{\omega}_1} & \frac{\partial f_4}{\partial \overline{\omega}_2} & \frac{\partial f_4}{\partial \overline{\omega}_3} \\
\frac{\partial f_5}{\partial \sigma_1} & \frac{\partial f_5}{\partial \sigma_2} & \frac{\partial f_5}{\partial \sigma_3} & \frac{\partial f_5}{\partial \overline{\omega}_1} & \frac{\partial f_5}{\partial \overline{\omega}_2} & \frac{\partial f_5}{\partial \overline{\omega}_3} \\
\frac{\partial f_6}{\partial \sigma_1} & \frac{\partial f_6}{\partial \sigma_2} & \frac{\partial f_6}{\partial \sigma_3} & \frac{\partial f_6}{\partial \overline{\omega}_1} & \frac{\partial f_6}{\partial \overline{\omega}_2} & \frac{\partial f_6}{\partial \overline{\omega}_3}
\end{bmatrix} \tag{B-10}
$$

$$B = \begin{bmatrix} \frac{\partial f_1}{\partial \tau_1} & \frac{\partial f_1}{\partial \tau_2} & \frac{\partial f_1}{\partial \tau_3} \\ \frac{\partial f_2}{\partial \tau_1} & \frac{\partial f_2}{\partial \tau_2} & \frac{\partial f_2}{\partial \tau_3} \\ \frac{\partial f_3}{\partial \tau_1} & \frac{\partial f_3}{\partial \tau_2} & \frac{\partial f_3}{\partial \tau_3} \\ \frac{\partial f_4}{\partial \tau_1} & \frac{\partial f_4}{\partial \tau_2} & \frac{\partial f_4}{\partial \tau_3} \\ \frac{\partial f_5}{\partial \tau_1} & \frac{\partial f_5}{\partial \tau_2} & \frac{\partial f_5}{\partial \tau_3} \\ \frac{\partial f_6}{\partial \tau_1} & \frac{\partial f_6}{\partial \tau_2} & \frac{\partial f_6}{\partial \tau_3} \end{bmatrix} \tag{B-11}$$

$$G = \begin{bmatrix} \frac{\partial f_1}{\partial \nu_1} & \frac{\partial f_{0Q}}{\partial \nu_2} & \frac{\partial f_{0Q}}{\partial \nu_3} \\ \frac{\partial f_2}{\partial \nu_1} & \frac{\partial f_{1Q}}{\partial \nu_2} & \frac{\partial f_{1Q}}{\partial \nu_3} \\ \frac{\partial f_3}{\partial \nu_1} & \frac{\partial f_{2Q}}{\partial \nu_2} & \frac{\partial f_{2Q}}{\partial \nu_3} \\ \frac{\partial f_4}{\partial \nu_1} & \frac{\partial f_{3Q}}{\partial \nu_2} & \frac{\partial f_{3Q}}{\partial \nu_3} \\ \frac{\partial f_5}{\partial \nu_1} & \frac{\partial f_{4Q}}{\partial \nu_2} & \frac{\partial f_{4Q}}{\partial \nu_3} \\ \frac{\partial f_6}{\partial \nu_1} & \frac{\partial f_{5Q}}{\partial \nu_2} & \frac{\partial f_{5Q}}{\partial \nu_3} \end{bmatrix} \tag{B-12}$$

Each element of the Jacobian matrix A, B and G are computed as follows:

$$\begin{aligned}
\frac{\partial f_1}{\partial \sigma_1} &= 0.5(\omega_1\sigma_1 + \omega_2\sigma_2 + \omega_3\sigma_3) - 0.5(\nu_1\sigma_1 + \nu_2\sigma_2 + \nu_3\sigma_3) \\
\frac{\partial f_1}{\partial \sigma_2} &= 0.5(-\omega_1\sigma_2 + \omega_2\sigma_1 + \omega_3) + 0.5(\nu_1\sigma_2 - \nu_2\sigma_1 - \nu_3) \\
\frac{\partial f_1}{\partial \sigma_3} &= 0.5(-\omega_1\sigma_3 - \omega_2 + \omega_3\sigma 1) + 0.5(\nu_1\sigma_3 + \nu_2 - \nu_3\sigma_1) \\
\frac{\partial f_1}{\partial \omega_1} &= 0.25(1 + \sigma_1^2 - \sigma_2^2 - \sigma_3^2) = -\frac{\partial f_1}{\partial \nu_1} \\
\frac{\partial f_1}{\partial \omega_2} &= 0.5(\sigma_1\sigma_2 - \sigma_3) = -\frac{\partial f_1}{\partial \nu_2} \\
\frac{\partial f_1}{\partial \omega_3} &= 0.5(\sigma_1\sigma_3 + \sigma_2) = -\frac{\partial f_1}{\partial \nu_2}
\end{aligned} \tag{B-13}$$

$$\begin{aligned}
\frac{\partial f_2}{\partial \sigma_1} &= 0.5(\omega_1\sigma_2 - \omega_2\sigma_1 - \omega_3) + 0.5(-\nu_1\sigma_2 + \nu_2\sigma_1 + \nu_3) \\
\frac{\partial f_2}{\partial \sigma_2} &= 0.5(\omega_1\sigma_1 + \omega_2\sigma_2 + \omega_3\sigma_3) - 0.5(\nu_1\sigma_1 + \nu_2\sigma_2 + \nu_3\sigma_3) \\
\frac{\partial f_2}{\partial \sigma_3} &= 0.5(\omega_1 - \omega_2\sigma_3 + \omega_3\sigma_2) + 0.5(-\nu_1 + \nu_2\sigma_3 - \nu_3\sigma_2) \\
\frac{\partial f_2}{\partial \omega_1} &= 0.5(\sigma_1\sigma_2 + \sigma_3) = -\frac{\partial f_2}{\partial \nu_1} \\
\frac{\partial f_2}{\partial \omega_2} &= 0.25(1 - \sigma_1^2 + \sigma_2^2 - \sigma_3^2) = -\frac{\partial f_2}{\partial \nu_2} \\
\frac{\partial f_2}{\partial \omega_3} &= 0.5(\sigma_2\sigma_3 - \sigma_1) = -\frac{\partial f_2}{\partial \nu_3}
\end{aligned} \tag{B-14}$$

$$\begin{aligned}
\frac{\partial f_3}{\partial \sigma_1} &= 0.5(\omega_1\sigma_3 + \omega_2 - \omega_3\sigma_1) + 0.5(-\nu_1\sigma_3 - \nu_2 + \nu_3\sigma_1) \\
\frac{\partial f_3}{\partial \sigma_2} &= 0.5(-\omega_1 + \omega_2\sigma_3 - \omega_3\sigma_2) + 0.5(\nu_1 - \nu_2\sigma_3 + \nu_3\sigma_2) \\
\frac{\partial f_3}{\partial \sigma_3} &= 0.5(\omega_1\sigma_1 + \omega_2\sigma_2 + \omega_3\sigma_3) - 0.5(\nu_1\sigma_1 + \nu_2\sigma_2 + \nu_3\sigma_3) \\
\frac{\partial f_3}{\partial \omega_1} &= 0.5(\sigma_3\sigma_1 - \sigma_2) = -\frac{\partial f_3}{\partial \nu_1} \\
\frac{\partial f_3}{\partial \omega_2} &= 0.5(\sigma_3\sigma_2 + \sigma_1) = -\frac{\partial f_3}{\partial \nu_2} \\
\frac{\partial f_3}{\partial \omega_3} &= 0.25(1 - \sigma_1^2 - \sigma_2^2 + \sigma_3^2) = -\frac{\partial f_3}{\partial \nu_3}
\end{aligned} \tag{B-15}$$

The partial derivatives of function f_i, where $\{i = 4, 5, 6\}$ with respect to σ_j, ω_j and ν_j, $\forall\, j = \{1, 2, 3\}$ are calculated as i.e.

$$\begin{aligned}
\frac{\partial f_i}{\partial \sigma_j} &= 0 \\
\frac{\partial f_i}{\partial \omega_j} &= \frac{\partial f_{iQ}}{\partial \omega_j} = -\frac{\partial f_i}{\partial \nu_j}
\end{aligned} \tag{B-16}$$

Similarly

$$\frac{\partial f_j}{\partial \tau_j} = 0 \tag{B-17}$$

and

$$\frac{\partial f_4}{\partial \tau_1} = a, \qquad \frac{\partial f_4}{\partial \tau_2} = b, \qquad \frac{\partial f_4}{\partial \tau_3} = c, \qquad \frac{\partial f_5}{\partial \tau_1} = d,$$
$$\frac{\partial f_5}{\partial \tau_2} = e, \qquad \frac{\partial f_5}{\partial \tau_3} = f, \qquad \frac{\partial f_5}{\partial \tau_1} = g, \qquad \frac{\partial f_5}{\partial \tau_2} = h, \tag{B-18}$$
$$\frac{\partial f_5}{\partial \tau_3} = i$$

Substituting the above partial derivations in Equations (B-10), (B-11) and (B-12), the required Jacobian matrices A, B and G can be obtained.

Appendix C

Linear Prediction Coefficients through Auto-covariance Method

It is not necessary that the process under consideration would be a stationary one for which the mean and variance of the process are constants. In practice, non-stationary processes where the constant mean and variance conditions might be violated, can frequently encountered into the systems. In those cases, the linear prediction can be carried out as follows:

The correlation coefficients $R[m]$ is given as

$$R[m, 0] = \frac{C_m}{C_0} \tag{C-1}$$

with auto-covariance defined as

$$C_m = \frac{1}{n-m} \sum_{j=1}^{n-m} (y_j - \bar{y})(y_{m+j} - \bar{y}) \tag{C-2}$$

where 'n' is the order of the LPC filter, (or in other words, at maximum n-previous observations can be considered for prediction purposes). Although, conventionally one should write y_{j+m} but since j is a variable and in the later stages j will expand the notation of y_{m+j} is preferred. In the above equation

$$\bar{y} = \frac{1}{n} \sum_{j=1}^{n} y_j \tag{C-3}$$

which means

$$n\bar{y} = \sum_{j=1}^{n} y_j \tag{C-4}$$

The above covariance matrix can be expanded as

$$\begin{aligned} C_m &= \underbrace{\frac{1}{n-m}\sum_{j=1}^{n-m} y_j y_{m+j}}_{B_1} + \underbrace{\frac{1}{n-m}\sum_{j=1}^{n-m}(-y_j\bar{y} - \bar{y}y_{m+j})}_{B_2} + \underbrace{\frac{1}{n-m}\sum_{j=1}^{n-m}\bar{y}^2}_{B_3} \\ &= B_1 + B_2 + B_3 \end{aligned} \tag{C-5}$$

where

$$B_1 = \frac{1}{n-m}\sum_{j=1}^{n-m}\{y_j y_{m+j}\} \tag{C-6}$$

$$B_2 = \frac{1}{n-m}\sum_{j=1}^{n-m}[-y_j\bar{y} - \bar{y}y_{m+j}] \tag{C-7}$$

and

$$B_3 = \frac{1}{n-m}\sum_{j=1}^{n-m}[\bar{y}^2] \tag{C-8}$$

Solving B_2 and B_3 only.

$$\begin{aligned} B_2 &= \frac{-\bar{y}}{n-m}\sum_{j=1}^{n-m}[y_j + y_{m+j}] \\ &= \frac{-\bar{y}}{n-m}[\underbrace{y_1 + y_2 + y_3 + \cdots y_m}_{S_1} + y_{m+1}\cdots y_{n-m-1} + y_{n-m} \\ &\quad + y_{m+1} + y_{m+2}\cdots + \underbrace{y_{n-m+1} + y_{n-m+2}\cdots y_{n-1} + y_n}_{S_2}] \\ &= -\frac{\bar{y}}{n-m}[S_1 + S_3 + S_2] \end{aligned} \tag{C-9}$$

where

$$S_1 = \sum_{j=1}^{m} y_j \tag{C-10}$$

$$S_2 = \sum_{j=n-m+1}^{n} y_j \tag{C-11}$$

$$S_3 = 2\sum_{j=m+1}^{n-m} y_j \tag{C-12}$$

To compute the value of S_3 in terms of $\bar{y}$, S_1 and S_2, look at the expanded version of equation (C-4)

$$\begin{aligned} n\bar{y} &= \underbrace{y_1 + y_2 + \cdots y_{k-1} + y_m}_{S_1} + \underbrace{y_{m+1} + \cdots + y_{n-m}}_{S_3/2} + \underbrace{y_{n-m+1} + \cdots y_n}_{S_2} \\ &= S_1 + S_3/2 + S_2 \end{aligned} \tag{C-13}$$

Or

$$S_3 = 2(n\bar{y} - S_1 - S_2) \tag{C-14}$$

Substituting the value of q in (C-9) yields

$$\begin{aligned} B_2 &= -\frac{\bar{y}}{n-m}[S_1 + 2(n\bar{y} - S_1 - S_2) + S_2] \\ &= -\frac{\bar{y}}{n-m}[2n\bar{y} - (S_1 + S_2)] \\ &= \frac{-2n\bar{y}^2}{n-m} + \frac{\bar{y}}{n-m}(S_1 + S_2) \end{aligned} \tag{C-15}$$

In the similar way

$$B_3 = \frac{1}{n-m}\sum_{j=1}^{n-m} \bar{y}^2 = \bar{y}^2 \tag{C-16}$$

Substituting the value of B_1, B_2 and B_3 in equation (C-5) will result in

$$\begin{aligned} C_m &= \frac{1}{n-m}\sum_{j=1}^{n-m} y_j y_{m+j} + \frac{1}{n-m}[\bar{y}(S_1 + S_2)] - \frac{1}{n-m}(2n\bar{y}^2) + \bar{y}^2 \\ &= \frac{1}{n-m}\sum_{j=1}^{n-m} y_j y_{m+j} + \frac{1}{n-m}\left[\bar{y}(S_1 + S_2) + (n-m-2n)\bar{y}^2\right] \\ &= \frac{1}{n-m}\sum_{j=1}^{n-m} y_j y_{m+j} + \frac{1}{n-m}\left[\bar{y}(S_1 + S_2) + (-m-n)\bar{y}^2\right] \\ &= \frac{1}{n-m}\sum_{j=1}^{n-m} y_j y_{m+j} + \frac{1}{n-m}\left[\bar{y}(S_1 + S_2) - (n+m)\bar{y}^2\right] \end{aligned} \tag{C-17}$$

To go into more details, substituting the observation value

$$y_j = Cx_j + v_j \qquad (E[y_j] = D_j) \tag{C-18}$$

Substituting the value of y_j and $\bar{y}$ as follows:

$$\begin{aligned} C_m &= \underbrace{\frac{1}{n-m}\sum_{t=1}^{n-m}(D_j+v_j)(D_{m+j}+v_{m+j})}_{T_1} + \underbrace{\frac{1}{n(n-m)}\sum_{j=1}^{n}(D_j+v_j)(S_1+S_2)}_{T_2} \\ &\quad \underbrace{-\frac{n+m}{n-m}[\frac{1}{n}\sum_{j=1}^{n}(D_j+v_j)]^2}_{T_3} \end{aligned} \tag{C-19}$$

$$\begin{aligned} T_1 &= \frac{1}{n-m}\sum_{j=1}^{n-m}(D_j+v_j)(D_{m+j}+v_{m+j}) \\ &= \frac{1}{n-m}\sum_{j=1}^{n-m}(D_jD_{m+j}) + \frac{1}{n-m}\sum_{j=1}^{n-m}(v_jv_{m+j}) + \underbrace{\frac{1}{n-m}\sum_{j=1}^{n-m}(D_jv_{m+j}) + \frac{1}{n-m}\sum_{j=1}^{n-m}(D_{m+j}v_j)}_{A_1=0} \\ &= \frac{1}{n-m}\sum_{j=1}^{n-m}(D_jD_{m+j}) + \frac{1}{n-m}\sum_{j=1}^{n-m}(v_jv_{m+j}) \end{aligned} \tag{C-20}$$

and

$$\begin{aligned} T_2 &= \frac{1}{n(n-m)}\sum_{j=1}^{n}(D_j+v_j)(S_1+S_2) \\ &= \frac{1}{n-m}\bar{D}(S_1+S_2) \end{aligned} \tag{C-21}$$

where

$$\frac{1}{n}\sum_{j=1}^{n}v_j = \bar{v} = 0$$

$$\frac{1}{n}\sum_{j=1}^{n}D_j = \bar{D}$$

Equation (C-21) is simplified by substituting the values of S_1 and S_2;

$$\begin{aligned}
T_2 &= \frac{1}{n-m}\bar{D}[(\sum_{j=1}^{m} y_j) + (\sum_{j=n-m+1}^{n} y_j)] \\
&= \frac{1}{n-m}\bar{D}[(\sum_{j=1}^{m}(D_j + v_j)) + (\sum_{j=n-m+1}^{n}(D_j + v_j))] \\
&= \frac{1}{n-m}[\bar{D}(D_I + \sum_{j=1}^{m} v_j) + \bar{D}(D_F + \sum_{j=n-m+1}^{n} v_j)] \\
&= \frac{1}{n-m}[\bar{D}(D_I + D_F)] + A_2 \qquad \text{(C-22)}
\end{aligned}$$

where

$$A_2 = \frac{\bar{D}}{n-m}[\sum_{j=1}^{m} v_j + \sum_{j=n-m+1}^{n} v_j] \qquad \text{(C-23)}$$

And finally

$$\begin{aligned}
T_3 &= -\frac{n+m}{n-m}[\frac{1}{n}\sum_{j=1}^{n}(D_j + v_j)]^2 \\
&= -\frac{n+m}{n-m}\bar{D}^2 \qquad \text{(C-24)}
\end{aligned}$$

Equation (C-22) is further simplified by substituting the values of D_I and D_F. Recall the substitutions $\bar{D} = \frac{1}{n}\sum_{j=1}^{n} D_j$; $D_I = \sum_{j=1}^{m} D_j$ and $D_F = \sum_{j=n-m+1}^{n} D_j$. Therefore,

$$\begin{aligned}
n\bar{D} &= \sum_{j=1}^{n} D_j \\
&= \underbrace{D_1 + D_2 + \cdots + D_m}_{D_I} + \underbrace{D_{m+1} + \cdots D_{n-m} + D_{n-m+1} + \cdots + D_n}_{\sum_{j=m+1}^{n} D_j = (n-m)\bar{D}_M} \qquad \text{(C-25)}
\end{aligned}$$

where

$$\bar{D}_M = \frac{1}{n-m}\sum_{j=m+1}^{n} D_j \qquad \text{(C-26)}$$

Therefore,

$$\begin{aligned}
n\bar{D} &= D_I + (n-m)\bar{D}_M \\
D_I &= n\bar{D} - (n-m)\bar{D}_M \qquad \text{(C-27)}
\end{aligned}$$

In the same way,

$$n\bar{D} \;=\; \underbrace{D_1 + D_2 + \cdots + D_m + D_{m+1} + \cdots D_{n-m}}_{\sum_{j=1}^{n-m} D_j = (n-m)\bar{D}_m} + \underbrace{D_{n-m+1} + \cdots + D_n}_{D_F} \tag{C-28}$$

where

$$\bar{D}_m = \frac{1}{n-m}\sum_{j=1}^{n-m} D_j \tag{C-29}$$

$$\begin{aligned} n\bar{D} &= (n-m)\bar{D}_m + D_F \\ D_F &= n\bar{D} - (n-m)\bar{D}_m \end{aligned} \tag{C-30}$$

Considering A_2 to be very small, substituting the values of D_I and D_F in Equations (C-22) and (C-24) results in

$$\begin{aligned} T_3 + T_2 &= -\frac{n+m}{n-m}\bar{D}^2 + \frac{1}{n-m}\bar{D}[D_I + D_F] \\ &\frac{1}{n-m}[-(n+m)\bar{D}^2 + \bar{D}(D_F + D_I)] = \\ &= \frac{1}{n-m}[-(n+m)\bar{D}^2 + \bar{D}\{n\bar{D} - (n-m)\bar{D}_M + n\bar{D} - (n-m)\bar{D}_m\}] \\ &= \frac{1}{n-m}[-(n+m)\bar{D}^2 + \bar{D}\{2n\bar{D} - (n-m)(\bar{D}_m + \bar{D}_M)\}] \\ &= \frac{1}{n-m}[-n\bar{D}^2 - m\bar{D}^2 + 2n\bar{D}^2 - \bar{D}(n-m)(\bar{D}_m + \bar{D}_M)] \\ &= \frac{1}{n-m}[n\bar{D}^2 - m\bar{D}^2 - \bar{D}(n-m)(\bar{D}_m + \bar{D}_M)] \\ &= \frac{1}{n-m}[(n-m)\bar{D}^2 - \bar{D}(n-m)(\bar{D}_m + \bar{D}_M)] \\ &= \frac{n-m}{n-m}[\bar{D}^2 - \bar{D}\bar{D}_m - \bar{D}\bar{D}_M] \\ &= \bar{D}^2 - \bar{D}\bar{D}_m - \bar{D}\bar{D}_M \end{aligned} \tag{C-31}$$

substituting T_1, $T_2 + T_3$ in Equation (C-19) yields

$$C_m = \frac{1}{n-m}\sum_{j=1}^{n-m}(D_j D_{m+j}) + \sum_{j=1}^{n-m} v_j v_{m+j} + \bar{D}^2 - \bar{D}\bar{D}_m - \bar{D}\bar{D}_M \tag{C-32}$$

or

$$\begin{aligned}
C_m &= \overbrace{\frac{1}{n-m}\sum_{j=1}^{n-m} D_j D_{m+j} - \bar{D}_m \bar{D}_M}\, + \bar{D}_m \bar{D}_M + \overbrace{\frac{1}{n-m}\sum_{j=1}^{n-m} v_j v_{m+j}} \\
&+ \bar{D}^2 - \bar{D}\bar{D}_m - \bar{D}\bar{D}_M \\
&= Cov(D_j D_{m+j}) + Cov(v_j) + (\bar{D} - \bar{D}_m)(\bar{D} - \bar{D}_M) \qquad \text{(C-33)}
\end{aligned}$$

C_0 is calculated as

$$\begin{aligned}
C_0 &= \frac{1}{n}\sum_{j=1}^{n}(y_j - \bar{y})(y_j - \bar{y}) \\
&= \frac{1}{n}[\sum_{j=1}^{n}(y_j)^2 - 2\sum_{j=1}^{n}(y_j\bar{y}) + \sum_{j=1}^{n}(\bar{y})] \\
&= \frac{1}{n}\sum_{j=1}^{n}(y_j)^2 - \frac{2\bar{y}}{n}\sum_{j=1}^{n}(y_j) + \bar{y}^2 \\
&= \frac{1}{n}\sum_{j=1}^{n}(y_j)^2 - 2\bar{y}^2 + \bar{y}^2 \\
&= \frac{1}{n}\sum_{j=1}^{n}(y_j)^2 - \bar{y}^2 \qquad \text{(C-34)}
\end{aligned}$$

Substitute the value of $y_j = D_j + v_j$, with $\bar{v} = 0$ generates

$$\begin{aligned}
C_0 &= \frac{1}{n}\sum_{j=1}^{n}(D_j + v_j)^2 - \bar{D}^2 \\
&= \frac{1}{n}\sum_{j=1}^{n}(D_j^2 + v_j^2 + 2D_j v_j) - \bar{D}^2 \\
&= \frac{1}{n}\sum_{j=1}^{n}(D_j^2) + \frac{1}{n}\sum_{j=1}^{n}(v_j)^2 - \bar{D}^2 \\
&= \underbrace{\frac{1}{n}\sum_{j=1}^{n}(D_j^2) - \bar{D}^2} + \underbrace{\frac{1}{n}\sum_{j=1}^{n}(v_j)^2} \\
&= var(D) + var(v) \qquad \text{(C-35)}
\end{aligned}$$

Therefore, the correlation elements can be computed by substituting Equations (C-33) and (C-35) in Equation (C-1) as follows:

$$R[m,0] = \frac{C_m}{C_0} = \frac{Cov(D_j D_{m+j}) + Cov(v_j) + (\bar{D} - \bar{D}_m)(\bar{D} - \bar{D}_M)}{var(D) + var(v)} \qquad \text{(C-36)}$$

where the $var(D) = \frac{1}{N}\sum_{i=1}^{N}(D_i)^2 - \bar{D}^2$ and $var(v) = \frac{1}{n}\sum_{j=1}^{n}(v_j)^2\ \forall\ i$.

The auto-covariance matrix R_γ and auto-covariance array r_γ are constructed as

$$R_z = \begin{bmatrix} R[1,1] & R[1,2] & \cdots & R[1,p] \\ R[2,1] & R[2,2] & \cdots & R[2,p] \\ \vdots & \vdots & \ddots & \cdots \\ R[p,1] & R[p,2] & \cdots & R[p,p] \end{bmatrix} \tag{C-37}$$

The linear prediction coefficients A_α are derived through

$$A_\alpha = (R_z^{-1})r_\gamma \tag{C-38}$$

The compensated signal is calculated with the help of these computed LPC by the Equation

$$\bar{z}_k = \sum_{j=1}^{n} \alpha_j z_{k-j} \tag{C-39}$$

Bibliography

[1] E. Abbott and D. Powell. Land-Vehicle Navigation Using GPS. *Proceedings of the IEEE*, 87(1):145 – 162, January 1999.

[2] R. Ahmed, D. Gu, and I. Postelwaite. A Case Study on Spacecraft Attitude Control. In *48th IEEE Control and Decision Conference*, pages 7345 – 7350, December 2009.

[3] P. D. Allison. *Missing Data.* Sage Publications, 1st edition, August 2001.

[4] B. O. Anderson and J. B. Moore. *Optimal Filtering.* Prentice Hall, Inc., 1979.

[5] A. Antoniads, H. J. Kim, and S. Sastry. Pursuit-Evasion Strategies for Teams of Multiple Agents with Incomplete Information. In *42nd IEEE Conference on Decision and Control*, USA, December 2003.

[6] B. Azimi-Sadjadi. Stability of Networked control system in the presence of Packet Losses. In *42nd IEEE Conference on Decision and Control*, Spain, December 2003.

[7] R. B. Babu, P. Perez, and P. Bouthemy. Robust Tracking with Motion Estimation and Lock Kernel-based Color Modeling. *Image and Vision Computing*, 25(8):1205 – 1216, 2007.

[8] R. Blind, S. Uhlich, B. Yang, and F. Allgower. Robustification and Optimization of a Kalman Filter with Measurement Loss using Linear Precoding. In *American Control Conference*, pages 2222 – 2227, June 2009.

[9] R. G. Brown and P. Y. C. Hwang. *Introduction to Random Signals and Applied Kalman Filtering.* John Wiley & Sons, 3rd edition, 1997.

[10] A. Censi. On the performance of Kalman filtering with Intermittent Observation: A Geometric Approach with Fractals. In *American Control Conference*, pages 3806 – 3812, USA, June 2009.

[11] C. Chatfield. *The Analysis of Time Series: An Introduction.* Chapman & Hall/CRC, New York, sixth edition, 2004.

[12] G. Chen, Z. Xia, X. Ming, S. Lining, J. Ji, and Z. Du. Camera Calibration based on Extended Kalman Filter using Robot's Arm Motion. In *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pages 1839 – 1844, Suntec Convention and Exhibition Center, Singapore, July 2009.

[13] K.-Y. Choi and A. A. Khan. Optimal State Estimation in the Transesterification stage of a Continuous Polythylene Terephtalate Condensation Polymerization Process. *Chemical Engineering Science*, 43(4):749 – 762, 1988.

[14] W. C. Chu. *Speech Coding Algorithms: Foundation and Evolution of Standardized Coders*. John Wiley and Sons Inc., 2003.

[15] C. K. Chui and G. Chen. *Kalman Filtering with Real-Time Applications*. Springer, fourth edition, 2009.

[16] R. N. Clark. A Simplified Instrument Failure Detection Scheme. *IEEE Transactions on Aerospace and Electronics Systems*, AES-14:558 – 563, 1978.

[17] R. N. Clark, D. C. Fosth, and W. M. Walton. Detecting Instrument Malfunctions in Control Systems. *IEEE Transactions on Aerospace and Electronics Systems*, AES-11:465 – 473, 1975.

[18] J. J. D'Azzo and C. H. Houpis. *Linear Control System Analysis and Design*. McGraw-Hill, 4th edition, 1995.

[19] F. Ding and J. Ding. Parameter Estimation Algorithms for Missing-Data Systems. In *American Control Conference*, pages 5032 – 5036, USA, June 2009.

[20] M. Epstein, L. Shi, A. Tiwari, and R. M. Murray. Probabilistic Performance of State Estimation Across a Lossy Network. *Automatica*, 44:3046 – 3053, December 2008.

[21] J. F., K. Jr., and A. Gelb. *Applied Optimal Estimation*. 65. The M.I.T. Press, TASC, 1974.

[22] H. Fang and Y. Shi. Genetic Adaptive State Estimation with Missing Input/Output Data. *Proc. IMechE System and Control Engineering*, 224(Part I: J.):611 – 617, March 2010.

[23] S. Fekri, M. Athans, and A. Pascoal. Robust Multiple Model Adaptive Control (RMMAC): A Case Study. *International Journal of Adaptive Control And Signal Procession*, 21:1 – 30, November 2007.

[24] H. Feng, Y. Shi, and J. Wu. Parameter Estimation with Missing Input/Output Data. In *American Control Conference*, pages 5061 – 5066, USA, June 2009.

[25] D. L. Fletcher and W. O. Stadlin. Transformer tap position estimation. *IEEE Transactions on Power Apparatus and Systems*, PAS-102(11):3680 – 3686, Nov 1983.

[26] E. Foxlin. Pedestrian Tracking with Shoe-Mounted Inertial Sensors. *IEEE Computer Graphics and Applications*, 25(6):38 – 46, 2005.

[27] P. M. Frank. Advanced Fault Detection and Isolation Schemes Using Nonlinear and Robust Observers. In *10th IFAC World Congress*, Munich, 1987.

[28] M. Fu, C. E. de Souza, and Z.-Q. T. Luo. Finite-Horizon Robust Kalman Filter Design. *IEEE Transactions On Signal Processing*, 49(9):2103 – 2112, September 2001.

[29] T. Gao, Z. Gong, J. Luo, W. Ding, and W. Feng. An Attitude Determination System For A small Unmanned Helicopter Using Low-Cost Sensors. In *Proceedings of the IEEE, International Conference on Robotics and Biomimetics*, pages 1203–1208, China, December 2006.

[30] A. A. Girgis, W. B. Chang, and E. B. Makram. A Aigital Recursive Measurement Scheme for on-line Tracking of Power System Harmonics. *IEEE Transactions on Power Delivery*, 6(3):1153 – 1160, July 1991.

[31] G. C. Goodwin, H. Haimovich, D. E. Quevedo, and J. S. Welsh. A Moving Horizon Approach to Networked Control System Design. *IEEE Transactions on Automatic Control*, 49(9):1427 – 1445, September 2004.

[32] M. S. Grewal and A. P. Andrews. *Kalman Filtering: Theory and Practice using MATLAB*. John Wiley & Sons, Inc., third edition, 2008.

[33] D. Gu, P. H. Petko, and M. M. Konstantinov. *Robust Control Design with MATLAB*. Advanced Textbooks in control and signal processing. Springer, 2005.

[34] G. Guo and B.-F. Wang. Kalman Filtering with Partial Markovain Packet Losses. *International Journal of Automation and Computing*, 04(1):100 – 106, January 2007.

[35] P. D. Hanlon and P. S. Maybeck. Multi-Model Adaptive Estimation Using a Residual Correlation Kalman Filter Bank. *IEEE Transactions on Aerospace and Electronic Systems*, 36(2):393 – 406, 2000.

[36] M. H. Hayes. *Statistical Digital Signal Processing and Modelling*. John Wiley and Sons Inc., 1996.

[37] E. Hendricks, O. Jannerup, and P. H. Sorensen. *Linear System Control, Deterministic and Stochastic Methods*. Springer, 2008.

[38] G. Heredia and A. Ollero. Sensor Fault Detection in Small Autonomous Helicopter using Observer Kalman Filter Identification. Spain, April 2009. Proceeding of the 2009 IEEE International Conference on Mechatronics.

[39] G. Heredia, A. Ollero, M. Bejar, and R. Mahtani. Sensor and Actuator Fault Detection in Small Autonomous Helicopters. *MECHATRONICS*, 18:90 – 99, September 2008.

[40] C. Hu, W. Chen, Y. Chen, and D. Liu. Adaptive Kalman Filtering for Vehicle Navigation. *Journal of Global Positioning Systems*, 2(1):42 – 47, November 2003.

[41] M. Huang and S. Dey. Kalman Filtering with Markovian Packet Losses and Stability. In *45^{th} IEEE Conference on Decision and Control*, pages 5621 – 5626, December 2005.

[42] M. Huang and S. Dey. Stability of Kalman Filtering with Markovian Packet Loss. *Automatica*, 43:598 – 607, October 2007.

[43] Y. Huang, G. V. Reklaitis, and V. Venkatasubramanian. A Heuristic Extended Kalman Filter Based Estimator for Fault Identification in a Fluid Catalytic Cracking Unit. *Industrial Engineering & Chemical Research*, 42:3361 – 3371, 2003.

[44] B. Jiang and F. N. Chowdhury. Fault Estimation and Accomodation for Linear MIMO Discrete-Time Systems. *IEEE Transactions on Control Systems Technology*, 13(3):493 – 499, May 2005.

[45] M. A. Kaafar, L. Mathy, C. Barakat, K. Salamatian, T. Turletti, and W. Dabbous, editors. *Security Internet Coordinate Embedding Systems*, Kyoto, Japan, August 2007. In the proceedings of ACMSIGCOMM.

[46] T. Kailath. An Innovation Approach to Least Square Estimation: Part-1. *IEEE Transactions on Automatic Control*, AC-13(6):646 – 655, December 1968.

[47] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME Journal of Basic Engineering*, pages 35 – 45, March 1960.

[48] P. P. Kanjilal. *Adaptive Prediction and Predictive Control*. 52. Peter Peregrinus Ltd Press, IEE Control Engineering, 1995.

[49] S. Kar, B. Sinopoli, and J. M. F. Moura. Kalman filtering with intermittent observations: Weak convergence to a stationary distribution. *Automatica*, Accepted, 2010.

[50] A. N. Kealy, M. Tsakiri, and M. Stewart. Land Vehicle Navigation in the Urban Canyon - A Kalman Filter Solution using Integrated GPS, GLONASS and Dead Reckoning. *Proceedings of the 12th International Technical Meeting of the Satellite Division of The Institute of Navigation*, pages 509 – 518, September 1999.

[51] A. C. Kermore. *Mechanics of Flights*. Prentice Hall, 11th edition, 2006.

[52] N. Khan, R. Ahmad, and D. Gu. A Fault Tolerant Scheme for Attitude Estimation. In *The 7th IEEE International Bhurban Conference on Applied Sciences & Technology, IBCAST*. Pakistan, January 2010.

[53] N. Khan, S. Fekri, and D. Gu. Improvement on state estimation for discrete-time LTI systems with measurement loss. *Journal of Measurement*, 43(2010):1609 – 1622, December 2010.

[54] N. Khan and D. Gu. Properties of A Robust Kalman Filter. In *The 2nd IFAC Conference on Intelligent Control System and Signal Processing*. Turkey, September 2009.

[55] N. Khan and D. Gu. State Estimation in the Case of Loss of Observations. In *ICROS-SICE International Joint Conference*, pages 1840 – 1845. Japan, Auguest 2009.

[56] S.-G. Kim, J. L. Crassidis, Y. Cheng, and A. M. Fosbury. Kalman Filtering for Relative Spacecraft Attitude and Position Estimation. *Journal of Guidance, Control and Dynamics*, 30(1):133 – 143, 2007.

[57] Y. Kim, D. Gu, and I. Postlethwaite. Fault-tolerant Cooperative Target Tracking in Distributed UAV Networks. In *Proceedings of the 17th IFAC World Congress*. Korea, July 2008.

[58] M. Kitamura. Detection of Sensor Failure in Nuclear Plant using Analytical Redundancy. *Transactions Amercian Nuclear Society*, 34:581 – 583, 1980.

[59] T. Kobayashi and D. L. Simon. Application of a Bank of Kalman Filters for Aircraft Engine Sensor Fault Diagnostics. Technical ARL-TR-2955, NASA Centre of Aerospace Information, August 2003.

[60] T. Kobayashi and D. L. Simon. Evaluation of an Enhanced Bank of Kalman Filters for In-Flight Aircraft Engine Sensor Fault Diagnostics. *Journal of Engineering for Gas Turbines and Power*, 127:497 – 504, July 2005.

[61] V. Kodrić, editor. *Kalman Filter*. Intech, India, 2010.

[62] A. Kurian. *Performance analysis of filtering based chaotic synchronization and development of chaotic digital communication schemes*. PhD thesis, National University of Singapore, 2006.

[63] M. B. Larsen, editor. *Synthetic Long Baseline Navigation of Underwater Vechicles*. Proceeding of Oceans, 2000.

[64] E. J. Lefferts, F. L. Markley, and M. D. Shuster. Kalman Filtering for Spacecraft Attitude Estimation. *AIAA Journal of Guidance, Control and Dynamics*, 5(5):417 – 429, September-October 1982.

[65] C. T. Leondes. *Theory and Application of Kalman Filtering*. Technical Editing and Reproduction Ltd, London, 1970.

[66] R. Li and J. H. Olson. Fault Detection and Diagnosis in a Closed-Loop Nonlinear Distillation Process: Application of Extended Kalman Filters. *Ind. Engineering of Chemical Research*, 30(5):898 – 908, 1991.

[67] B. G. Liptak. *Process Control and Optimization*. Instrument Engineering's Handbook. Taylor & Francis, fourth edition, 2006.

[68] X. Liu and A. Goldsmith. Kalman Filtering with Partial Observation Loss. In 43^{rd} *IEEE Conference on Decision and Control*, pages 4180 – 4186, December 2004.

[69] J. Makhoul. Linear Prediction: A Tutorial Review. *Proceedings of the IEEE*, 63(5):561 – 580, April 1975.

[70] M. Micheli. Random Sampling of a Continuous Time Stochastic Dynamical System: Analysis, State Estimation and Applications. Master's thesis, School of Engineering, Barkeley, 2001.

[71] F. M. Mirzaei and S. I. Roumeliotis. A Kalman Filter-Based Algorithm for IMU-Camera Calibration: Observability Analysis and Performance Evaluation. *IEEE Transactions on Robotics*, 24(5):1143 – 1156, October 2008.

[72] K. Mizukami and V. Tews. *State Estimation in a pursuit-evasion-game with incomplete information-exchange*, volume 22 of *Lecture Notes*, pages 241 – 250. Springer, 1980.

[73] N. E. Nahi. Optimal Recursive Estimation with Uncertain Observation. *IEEE transactions on Information Theory*, IT-15(4), July 1969.

[74] J. G. Nielsen, M. Newman, H. Nielsen, and F. Blaabjerg. Control and testing of a Dynamic Voltage Restorer (DVR) at medium voltage level. *IEEE Transactions on Power Electronics*, 19(3):806 – 813, May 2004.

[75] T. Parsons. *Voice and Speech Processing.* McGraw-Hill, 1987.

[76] R. Patton and J. Chen. *Robust Model Based Fault Diagnosis for Dynamic Systems.* Kluwer, 1999.

[77] R. Patton, P. Frank, and R. Clark. *Fault Diagnosis in Dynamic Systems Theory and Applications.* Prentice Hall, 1989.

[78] J. Peyton Z. Peebles. *Probability, Random Variables and Random Signal Principles.* McGraw-Hill, Inc, second edition, 1987.

[79] C. L. Phillips, J. M. Parr, and E. A. Riskin. *Signals, Systems and Transforms.* Prentice Hall, New Jersey, second edition, 1999.

[80] F. N. Pirmoradi, F. Sassani, and C. W. de Silva. Fault Detection and Diagnosis in a Spacecraft Attitude Determination System. *ACTA ASTRONAUTICA*, 65:710 – 729, April 2009.

[81] I. Postlethwaite, D. Gu, Y. Kim, K. Natesan, M. Kothari, N. Khan, and R. Omar. A Robust Fault-Tolerant Tracking Scheme. In *Realising Network Enable Capability*, volume RNEC08, UK, 2008.

[82] A. S. Poznyak. *Advance Mathematical Tools for Automatic Control Engineers*, volume 2: Stochastic Techniques. Elsevier, 2009.

[83] J. G. Proakis and D. G. Manolakis. *Digital Signal Processing: Principles, Algorithms and Applications.* Prentice Hall International Inc., third edition, 1996.

[84] L. Rabiner and B.-H. Juang. *Fundamentals of Speech Recognition.* Prentice Hall International, Inc., 1993.

[85] J. Raol, G. Girija, and J. Singh. *Modelling and Parameter Estimation of Dynamics Systems.* 65. The Institution of Electrical Engineers,, IEE Control Engineering, 2004.

[86] H. Rehbindera and X. Hub. Drift-free Attitude Estimation for Accelerated Rigid Bodies. *Automatica*, 40:653 – 659, 2004.

[87] H. Schaub and J. L. Junkins. *Analytical Mechanics of Space Systems.* AIAA Education Series. American Institute of Aeronautics and Astronautics, Inc., 2003.

[88] L. Schenato. Kalman Filtering for Network Control System with Random Delay and Packet Loss. In *Conference of Mathematical Theory of Networks and Systems (MTNS 06)*, Japan, July 2006.

[89] L. Schenato. Optimal Estimation in Networked Control Systems Subject to Random Delay and Packet Drop. *IEEE Transactions of Automatic Control*, 53(5):1311 – 1316, June 2008.

[90] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry. Foundation of Control and Estimation Over Lossy Network. *Proceeding of The IEEE*, 95(1):163 – 187, January 2007.

[91] C. Shah. Sensorless Control of Stepper Motor using Kalman filter. Master's thesis, Cleveland State University, May 2000.

[92] L. Shi, M. Epstein, A. Tiwari, and R. M. Murray. Estimation with Information Loss: Asymptotic Analysis and Error Bounds. In *44^{th} IEEE Conference on Decision and Control*, pages 1215 – 1221, Spain, December 2005.

[93] P. Shi, E.-K. Boukas, and R. K. Agarwal. Kalman filtering for Continuous-Time Uncertain Systems with Markovian Jumping Parameters. *IEEE Transactions on Automatic Control*, 44(8):1592 – 1597, July 1999.

[94] Y. Shi and H. Fang. Kalman filter-based identification for systems with randomly missing measurements in a network environment. *International Journal of Control*, 83(3):538 – 551, November 2009.

[95] Y. Shi, H. Fang, and M. Yan. Kalman filter-based adaptive control for networked systems with unknown parameters and randomly missing outputs. *International Journal of Robust and Nonlinear Control*, 19:1976 – 1992, November 2009.

[96] S. Simani, C. Fantuzzi, and R. J. Patton. *Model Based Fault Diagnosis in Dynamic Systems Using Identification Techniques*. Advances in Industrial Control. Springer, 2003.

[97] D. Simon. Kalman Filtering. *Embedded Systems Programming*, 14(6):72 – 79, 2001.

[98] D. Simon. *Optimal State Estimation Kalman, H_∞ and Nonlinear Approaches*. John Wiley and Sons, Inc., 2006.

[99] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordon, and S. S. Sastry. Kalman Filtering with Intermittent Observations. *IEEE Transections on Automatic Control*, 49(9):1453 – 1464, September 2004.

[100] J. J. E. Slotine and M. D. Benedetto. Hamiltonian Adaptive Control of Spacecraft. *IEEE Transactions on Automatic Control*, 35(7):848 – 852, July 1990.

[101] J.-J. E. Slotine and W. Li. *Applied Nonlinear Control*. Prentice Hall, 1991.

[102] S. C. Smith and P. Seiler. Estimation with Lossy Measurements: Jump Estimation for Jump Systems. *IEEE Transactions on Automatic Control*, 48(12):2163 – 2171, December 2003.

[103] F. Q. Thomas. *Discrete time Speech Signal Processing; Principles and Practice*. Prentice Hall Signal Processing Series. Prentice Hall, Massachusetts Institute of Technology, Lincoln Lab, 2002.

[104] P. P. Vaidyanathan. *The Theory of Linear Prediction*. Morgan and Claypool Publishers, 2008.

[105] J. Valasek and W. Chen. Observer/Kalman Filter Identification for Online System Identification of Aircraft. *Journal of Guidance, Control and Dynamics*, 26(2):347 – 353, March-April 2003.

[106] S. V. Vaseghi. *Advanced Digital Signal Processing and Noise Reduction*. John Wiley & Sons, Inc., third edition, 2001.

[107] W. S. Veline, editor. *The Control Handbook*, volume I. Jaico Publishing House, CRC Press, 1999.

[108] F. B. Vidal and A. Casanova. Window-Matching Technique with Kalman Filtering for an Improved Object Visual Tracking. In *Proceedings of the IEEE Conference on Automation Science and Engineering*, pages 829 – 834, USA, 2007.

[109] G. Welch and G. Bishop. An Introduction to the Kalman Filter. ACM, Inc., SIGGRAPH, 2001.

[110] J. R. Wertz, editor. *Spacecraft Attitude Determination and Control*. Kluwer Academic Publishers, 1990.

[111] L. L. Whitcomb. Underwater robotics: Out of the research laboratory and into the filed. In *Proceedings of IEEE International Conference on Robotics and Automation*, pages 709 – 716, San Francisco, April 2000.

[112] N. Wiener. *Extrapolation, Interpolation and Smoothing of Stationary Time Series*. The Technology Press and Wiley and Sons, Inc., New York, 1st edition, 1949.

[113] A. S. Willsky. A Survey of Design Methods for Failure Detection in Dynamic Systems. *Automatica*, 12:601 – 611, 1976.

[114] N. Xiao and L. Xie. Peak Covariance Stability of Kalman Filter with Bounded Markovian Packet Losses. In *The 7th World Congress on Intelligent and Automation*, China, June 2008.

[115] E. R. B. Xiaoping Yun, Mariano Lizarraga and R. B. McMGhee. An Improved Quaternion Based Kalman Filter for Real Time Tracking of Rigid Body Orientation. In *Proceedings of the IEEE/RSJ, International Conference on Intelligent Robots and Systems*, pages 1074 – 1079, October 2003.

[116] L. Xie and L. Xie. Peak Covariance Stability of a Random Riccati Equation Arising From Kalman Filtering With Observation Losses. *Journal System Science and Complexity*, 20:262 – 279, January 2007.

[117] W. Xue, Y.-Q. Guo, and X.-D. Zhang. Application of a bank of Kalman filters and a Robust Kalman Filter for Aircraft Engine Sensor/Actuator Fault Diagnosis. *International Journal of Innovative Computing, Information and Control*, 4(12):3161 – 3168, December 2008.

[118] F. Yang, Z. Wang, and Y. S. Hung. Robust Kalman filtering for Discrete Time-varying Uncertain Systems with Multiplicative Noises. *IEEE Transactions on Automatic Control*, 47(7):1179 – 1883, July 2002.

[119] M. Yang, Y.-W. Wang, Y.-H. Huang, and Z.-H. Guan. Stability and Synchronization of Complex Dynamical Network with Random Packet Losses. In *IEEE Multi-Conference on Systems and Control*, pages 1488 – 1492, Japan, September 2010.

[120] S. K. Yang. An experiment of state estimation for predictive maintenance using Kalman filter on a DC motor. *Reliability Engineering & System Safety*, 75:103 – 111, August 2002.

[121] A. Yilmaz, X. li, and M. Shah. Contour-Based Object Tracking with Occlusion Handling in Video Acquired Using Mobile Cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(11):1531 – 1536, 2004.

[122] X. Yun, C. Aparicio, E. R. Bachmann, and R. B. McGhee. Implementation and Experimental Results of a Quaternion Based Kalman Filter for Human Body Motion Tracking. In *Proceeding of IEEE, Internation Conference on Robotics and Automation*, pages 317 – 322. Spain, April 2005.

[123] P. Zarchan. *Fundamentals of Kalman Filtering - A Practical Approach*, volume 208. AIAA, 2005.

[124] Y. Zhang and J. Jiang. Design of Integrated Fault Detection, Diagnosis and Reconfigurable Control Systems. In *38th IEEE Conference on Decision and Control*, pages 3587 – 3592, December 1999.