A New Method of Non-Asymptotic Estimation for Linear Systems

Kumar Vishwanath Gopalakrishnan

Department of Electrical & Computer Engineering
McGill University
Montréal, Canada

December 2016

A thesis submitted to McGill University in partial fulfillment of the requirements of the Degree of Master of Engineering.

To my parents, sister and all my teachers without whom I wouldn't be here.

Abstract

Analysis of autonomous dynamical systems is important due to the advancement in control systems, especially feedback control systems. There are three fundamental problems encountered during designing a feedback controller. They are state estimation, parameter estimation, and robustness to external perturbations. There has been a wide range of methods proposed to estimate states and their time derivatives and parameter estimation, ranging from classical observers to sliding mode observers and algebraic observers. This thesis provides a critique of the approach of algebraic observers proposed by Fliess et al, in detail and aims to provide a solution to the drawbacks associated with the approach such as singularity at t=0 and accumulation of the truncation error in the Taylor series.

The objective of this thesis is to propose and describe a new method to estimate state and time derivatives of the state, as well as estimate the parameters of an unknown system using the knowledge of the model of the system or otherwise an existing differential invariant. A new method has been developed for non-asymptotic estimation of linear systems which is a simple alternative to the derivation of algebraic estimation equations. The method is based on a construction of an integral operator that that effectively implements numerical differentiation of the system output and offers a geometric representation of a linear system over in a Hilbert space. Such an approach readily suggests powerful noise rejection methods in which differential invariance rendered by the Cayley-Hamilton theorem plays a central role. Results are presented comparing our method to another classical algebraic estimation approach and also a Kalman filter.

Résumé

L'analyse de systèmes dynamiques autonomes est importante en raison de l'avancement dans les systèmes de commande, particulièrement des systèmes commande à retour d'information. Il y a trois problèmes fondamentaux rencontrés pendant la conception d'un contrôleur de rétroactif. Ils sont l'évaluation d'état, l'évaluation de paramètre et la robustesse aux perturbations externes. Il y a eu une vaste gamme de méthodes proposées pour évaluer les états et leurs dérivées de temps et l'évaluation de paramètre, s'étendant des observateurs classiques aux observateurs de mode glissants et des observateurs algébriques. Cette thèse fournit une critique de l'approche des observateurs algébriques proposés par des Fliess et al., en détail et a pour but de fournir une solution aux inconvénients associés à l'approche comme la singularité à t=0 et l'accumulation de l'erreur de troncature dans la série de Taylor

L'objectif de cette thèse est de proposer et décrire une nouvelle méthode d'évaluation l'état et ses dérivées, aussi bien qu'évaluer les paramètres d'un système inconnu utilisant la connaissance du modèle du système ou autrement par un invariant différentiel. Une nouvelle méthode a été développée pour l'évaluation non-asymptotique des systèmes linéaires qui est une alternative simple à la dérivation d'équations d'évaluation algébriques. La méthode est basée sur une construction d'un opérateur intégral qui effectivement met en œuvre la différenciation numérique de la sortie du système et offre une représentation géométrique d'un système linéaire dans un espace Hilbert. Une telle approche suggère aisément des méthodes de rejet de bruit puissantes dans lesquelles l'invariance différentielle rendue par le théorème Cayley-Hamilton joue un rôle central. Les résultats sont présentés comparant notre méthode à une autre approche d'évaluation algébrique classique et aussi le filtre de Kalman.

Acknowledgements

This thesis would not have been possible without the help and guidance of several people. I would like to take this opportunity to express my gratitude towards them.

First of all, I would like to sincerely thank my supervisor, Professor Hannah Michalska for giving me the opportunity to work with her. It has been an eye-opening experience working with her and learning from her every day. She has been an inspiration and motivation to pursue high quality research and never to settle for second best. I would like to express my heartfelt gratitude towards her for guiding me every step of the way and helping in conquering all challenges, professional or personal. I sincerely believe that the lessons learned from her will keep me in good stead in life ahead. This thesis would not have been possible if not for her infinite patience, support, and encouragement. I wish to express gratitude to my co-supervisor, Professor Benoit Boulet for his constant encouragement throughout the course of my studies at McGill University. I would like to especially thank Professor Boon-Teck Ooi, for providing a safe haven and allowing me to work in his lab. Without this, I wouldn't have been able to complete my thesis.

I would like to thank the financial support provided by the Department of Electrical Engineering, McGill University in the form of Graduate Excellence Award and casual employment which made my stay at Montréal very affordable. I would like to further express my heartfelt thanks to the McGill Office of Student with Disabilities, especially Ms. Stephanie Larocque and Mr. Andrew, who have offered casual employment throughout my studies in McGill University which helped me make my ends meet, and concentrate on working on my research.

I am indebted to my colleague Debarshi Ghoshal who has been guiding me and has saved me many times with his brilliance and forethought. I would like to place thanks on record to my colleagues, Nandakumar Menon, Namrata Barman, Howard, and Jack , all of whom I have been lucky enough to work with.

I would express my sincerest gratitude towards my parents and my sister who have been supportive of me every step of the way, not only emotionally, but financially, keeping me solvent throughout my studies. I wouldn't be here without their support.

I would be remiss if I were to not thank my friends in Montréal, especially Arun, Kishan, Niladri, Nishanth, Nithin, Sanjeewa and Vinnay who have been tolerant of me and have been my mentors at various points of my study here.

Lastly, I would like to thank God for me giving me the strength to be successful, keeping me healthy and grounded throughout my stay in Montréal.

Preface

This is to declare that this work is a part of a collaborative work done by the team supervised by Professor Hannah Michalska and comprising Mr. Debarshi Ghoshal, Ph.D. scholar, Mr. Nandakumar Menon, Masters student, Ms. Namrata Barman, a Masters graduate and myself.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

We are living in a world which is constantly evolving. We live in an age where driverless cars, unmanned aerial vehicles, automated homes, and factories are a reality. This reality was in part made possible due to the development of control systems during the turn of the century. A control system is a scheme whose objective is to positively affect the performance of a system according to a set of objectives [3]. A set of regulations that allows us to achieve the desired output is called a controller. A controller can be in one of the two types: *open-loop* or *closed-loop* control.

An *open-loop control system* does not have any feedback and the output is based on the input. A conveyor belt would be an example of an open loop control system. The speed of a conveyor belt is directly proportional to the voltage applied across it. A schematic of an open control system which is self-explanatory is given below.

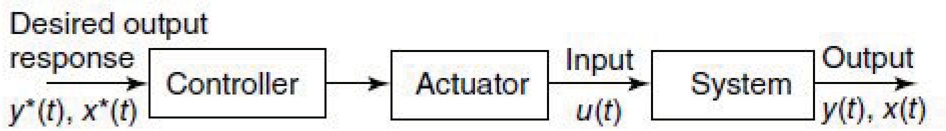

Figure 1.1: Open Loop Control Scheme [3]

A *closed-loop control system* or a *feedback control system* is a control scheme where a sensor measures the error, that is, the difference between the desired output with the current output and then generates a control law so as to change the state or the parameters of the system to bring the system output closer to the desired output. An example in layman terms would be driving a car. Suppose a person is driving a car at a constant speed. When he encounters a traffic signal and it changes to red, the person then takes the feedback of the traffic light and slows down the car to bring it to a halt at the signal. A schematic diagram of a feedback control systems is given below.

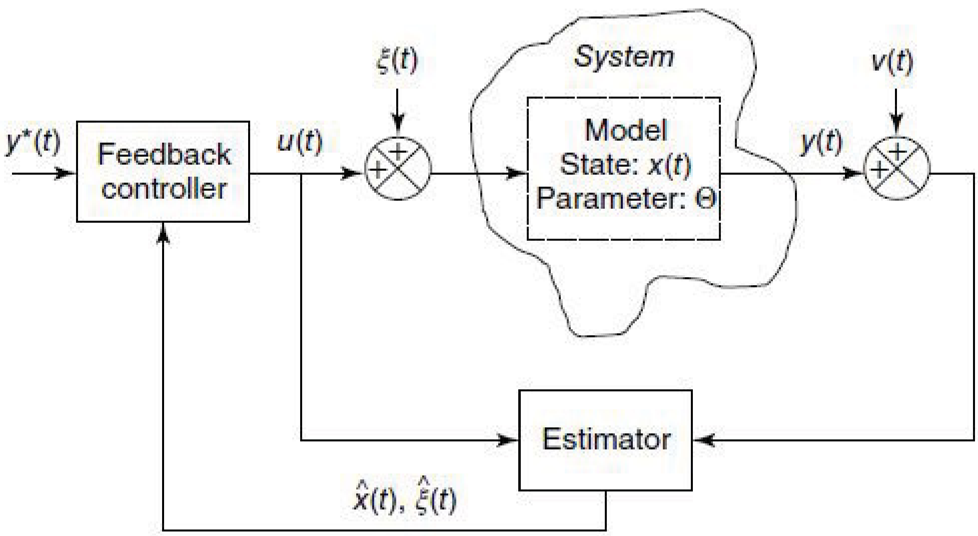

Figure 1.2: Feedback/Closed Control Scheme [3]

Feedback control though complex are required because of the following reasons [3]:

- Firstly, to mitigate the disturbances in the form of undesirable inputs affecting a plant and keeping the output in the operating range.
- Secondly, to improve performance in the presence of uncertainties in model dynamics and in the inputs to the system.
- Thirdly, to make the system stable.

But, to implement a good feedback law, there are three main problems encountered. These are (a) state estimation, (b) parameter estimation and (c)tough towards external perturbations. And these are the issues that will be tackled in the course of the next few sections and the thesis.

1.1 State Estimation - Literature

The first problem in the case of a feedback controller is state estimation. The state of a system can be defined as a set of variable whose present and future values determined by the input signal, completely determine the behaviour of a system [3]. The importance of state estimation in a feedback controller can be understood with the help of fig. 1.2. A very good explanation is given in [3] which helps us understand its significance.

In the case of a continuous-time system as in figure 1.2, the model consists of a set of differential equations which represent the change of the state variables $x(t)$ given a set of parameters θ. The input signal which is known is denoted by $u(t)$ and $v(t)$ is the external disturbance such as sensor noise etc, Hence due to these kinds of perturbations and measurement noise $\chi(t)$, the output $y(t)$ may be tainted. The desired output of the system is $y^{\star}(t)$. As shown in the fig. 1.2, the feedback controller needs the value of the state to produce the control signal to steer the system towards the desired output $y^{\star}(t)$. But, not all state values may be available and due to this, we have to compute the state to ensure that the feedback controller works smoothly, which is done by the estimator. To reduce the influence of external perturbations, the estimator is required to be robust, a quality that should characterize any good estimator.

There are two approaches to the problem of state estimation. One is by using state observers, while the other is by constructing time derivative estimators. The state observers are modeled on the plant and have been studied in great detail for linear systems through the works of Rudolf Kalman [4] and D. Luenberger [5]. The main difference between the methods proposed by [4] and [5] is that a Kalman observer (filter) uses stochastic optimization and needs assumptions about the measurement noise while the Luenberger observer is deterministic and is designed as a stable system with respect to the observer error.

The other way is to estimate the state is algebraically (in finite terms) by numerically differentiating the signal, which is known as estimation using time derivatives. This is applicable only if the system is observable but it has a drawback. It is greatly affected by the noise in the signal. And differentiating a noisy signal results in highly distorted time derivative from which it is tough to determine the state. Hence, it is almost always important that an additional filtering algorithm is implemented along with this method so that the results are acceptable. Rudolf Kalman extensively worked and built the state observer theory on the foundations of the work by Norbert Wiener [6] in stochastic context [4], [7], [8], [9]. While, Luenberger worked on the estimation of inaccessible states using available outputs in a deterministic setting [5], [10], [11].

As already mentioned above, in this method of state estimation using time derivatives, a specific set of methods known as algebraic estimation methods were encountered. These methods are either based on the model of the system or the time derivative estimations of the output. The alluring quality of these methods is that they are not affected by the choice of initial conditions unlike classical observers and have built in mechanism to be immune to noise and other external perturbations. Messrs. S. Diop and M. Fliess proposed an interesting differential algebraic approach to observability in [12]. In their approach, they state that a system is algebraically observable if it can be expressed in terms of the input, the output and their respective time derivatives of finite order. This method has been investigated further by M. Fliess, H. Sira-Ramírez et al, the results of which has been published in their seminal work [13]. This method was also applied to applications and modified to give better results [3], [14], [15].

The following were the reasons that made the algebraic approach of Fliess, H. Sira-Ramírez et al., attractive to use and apply in control systems [3]:

- Firstly, this method was found to be robust by the researchers in the presence of noise and time-polynomial perturbations. This set of methods didn't need a statistical knowledge of noise corrupting the signal and this noise could be filtered by the low pass filtering property of the integrating action incorporated in the state and parameter estimation methods.

- Secondly, this method did not rely on asymptotic convergence and was found to converge to the desired value in a limited amount of time, which really helped in designing much better controllers.

- Thirdly, as stated earlier, the unknown parameters and state estimates were found to be obtained in a very short period of time. They were also accurate and did not require “persistency of excitation condition" which is required in traditional adaptive control schemes. It was also found in the literature that the algebraic requirements needed to be satisfied only in a short interval of time and the identification process can be stopped once the state and the desired parameters are obtained.

- Finally, this method allows us to deal with the three fundamental challenges of controller design, which are; parameter identification, state estimation and robustness to additive perturbations such as ramp, parabolic etc.

The above method has been discussed in detail in section§ 2.2.

1.2 Parameter Identification - Literature

The second problem in the case of a feedback controller is that of parameter identification. A model or a plant in a control system is a mathematical description of a system and when this system is defined in terms of variables and parameters, it is known as a parametric model [3]. Examples of this are parametrized transfer functions of a system. To develop a model-based controller, it is very essential to know the parameters of the system, and if they are unknown, they need to be estimated in order to develop an effective controller. To put it formally, parameter estimation is the empirical determination of the parameter values that govern the behaviour of the system from a contaminated output signal and the output of the hypothesized model, when both are subject to same input signal [3].

The parameter estimation problem has been investigated in depth right from 300 B.C. by Galileo [16] to the seminal work on linear identification by Zadeh [17], to the works on adaptive control by Astrom [18] and [19]. A very detailed survey by Gevers found in [20] which gives us an excellent idea of the historical perspective and evolution of parameter identification.

In this work, we concern ourselves with the work of Messrs. Fliess and Sira-Ramírez [21] and [22], in which they use the algebraic technique developed for state estimation to estimate the parameters. The advantages of their method over the others as summarized in [2] are as follows:

- Firstly, the algebraic estimation is "dead-beat", i.e., parameter estimation is done in a very short amount of time provided the system trajectory is measured continuously.
- Secondly, due to the presence of low-pass filtering in the form of integration, noise rejection property is in-built into the method.
- Lastly, initial conditions are rendered redundant as algebraic estimation operates on output trajectories than data points.

Due to these advantages, the algebraic method is examined in the next chapter in the section § 2.3

1.3 Thesis Objectives and Organisation

The primary objective of this Masters thesis is to develop a new method for state and parameter estimation using the dynamical system knowledge to estimate the state variables of a system and their time derivatives as well as the parameter of the system with the help of a 'Double Sided Kernel K_{DS}. To accomplish this, the algebraic methods discussed in the previous sections , § 1.1 and § 1.2 are used to develop the K_{DS}. The development of the Double Sided Kernel is an original contribution and thus merited publications [1], [2].

Chapter 2 focuses on the development of the algebraic philosophy of estimation as proposed by M. Fliess et al, in [13] and [21], trying to understand the process of as well as ascertaining the advantages as well as disadvantages and the lacunae in the approach. An example is shown to better understand the functioning of the estimator as well as the areas of improvements.

Chapter 3 proposes the new estimation method. The goal is to show the development and the derivation of the expressions for the K_{DS} for the state as well as the time derivatives and the parameters for an example system: a generalized third order linear time invariant system. The derivation can be extrapolated to a linear time-invariant case with an input as well as a linear time varying case by the virtue of the validity of the Cayley-Hamilton theorem. A third order example is discussed to better illustrate the merits of the novel method over the algebraic estimation discussed in Chapter 2.

Chapter 4 presents the performance comparison of the proposed method against the algebraic method as well as a recursive observer such as the Kalman filter. This chapter focuses on different cases based on different values of signal to noise ratio (SNR) in the output signal and the results are presented graphically to clearly elucidate the differences in estimator performance. Also, the parameters of a third order system are estimated using the methods developed and the output signals as well as their derivatives are plotted to highlight the differences in performance.

Chapter 5 concludes the thesis by summarising the results and suggesting recommendations for possible future work.

Chapter 2

The Algebraic Approach to Estimation

2.1 Introduction

As seen in the previous chapter in section § 1.1, many methods have been proposed to tackle the issue of state and parameter estimation. In this chapter, we focus our efforts on understanding the algebraic approach to estimation of Messrs. Fliess et al. to state and parameter estimation and finding out how it mitigates the various problems afflicting the classical observers. It also gives us an idea of the drawbacks of the approach which can be alleviated using a novel approach to this fundamental problem of state and parameter estimation.

But before proceeding to the discussion on the development of the method, let us first describe the example system whose state we are estimating.

2.1.1 System Description

For simplicity we consider a general input-free single output Linear Time Invariant (LTI) model of the form (2.1) with $x(t) \in \mathbb{R}^n$.

$$\begin{aligned} \dot{x}(t) &= Ax(t) \\ y(t) &= Cx(t) \end{aligned} \tag{2.1}$$

The original idea of Fliess et. al. pertaining to the estimation of the above type of LTI systems has its foundation at a special method for signal differentiation. This type of differentiation uses the operational calculus to effectively substitute differentiation by integration. This is explained in detail in the next section.

2.2 Algebraic Signal Differentiation

This section discusses the development of state estimation using the algebraic estimation techniques proposed by Messrs. M. Fliess and H. Sira-Ramírez. *Algebraic estimation is based on the idea of repeated differentiation of a measured system output. The latter cannot be accomplished with satisfactory precision via any finite difference scheme due to the presence of measurement noise. Instead, the operation of differentiation is expressed in terms of the operation of integration*. This is explained step-by-step below; see also [13].

Consider $y_n(t)$ to be a noisy measurement of the output $y(t)$ from the above system, perturbed by an additive noise in the form of $y_n(t) = y(t) + \nu(t)$. Here the additive noise is denoted by $\nu(t)$. Before proceeding with the development, we make two important assumptions. The first assumption is that the signal to be estimated, $y(t)$ is an analytic signal and is differentiable at all points. This means that the signal is smooth and infinitely differentiable which allows us to use its Taylor Series expansion for the estimation of state and time derivatives of the signal $y(t)$. The second is that the Taylor series is convergent at $t = 0$, and hence we can consider the expansion of $y(t)$ as a function of time derivations of $y(t)$ at $t = 0$ [3].

Now, let us consider the Taylor series expansion of $y(t)$ at time $t = 0$ which is given as follows:

$$\begin{aligned} y(t) &= \sum_{i\geq 0}^{\infty} \frac{1}{i!}\, y^{(i)}(0)\, (t-0)^i \\ &= y(0) + \frac{1}{1!}\, y^{(1)}(0)\, (t-0) + \frac{1}{2!}\, y^{(2)}(0)\, (t-0)^2 + \frac{1}{3!}\, y^{(3)}(0)\, (t-0)^3 + \cdots \end{aligned} \tag{2.2}$$

In the above expression, $y^{(1)}(0)$ denotes the first derivative of the signal with respect to time evaluated at 0, i.e., $\frac{\mathrm{d}}{\mathrm{d}t}y(t)\mid_{t=0}$ and $1!$ denotes the factorial of 1. We know that the Taylor series is an infinite series, i.e., it has infinitely many terms. We can truncate the Taylor series to a sufficiently higher order, after which the error due to truncation would be negligible while at the same time would not complicate our calculations. Therefore, we approximate the above equation (2.2), by truncating to an order N. Rewriting (2.2) as a truncation and denoting the approximation of $y(t)$ as $\tilde{y}(t)$,

$$\begin{aligned}\tilde{y}(t) &= \sum_{i\geq 0}^{N} \frac{1}{i!}\, y^{i}(0)\, t^{i} \\ &= y(0) + \frac{1}{1!}\, y^{(1)}(0)\,(t-0) + \frac{1}{2!}\, y^{(2)}(0)\,(t-0)^{2} + \frac{1}{3!}\, y^{(3)}(0)\,(t-0)^{3} + \cdots \\ &\quad + \frac{1}{N!}\, y^{(N)}(0)\,(t-0)^{N}\end{aligned} \tag{2.3}$$

We can infer from (2.3) that the following statement is true

$$\frac{\mathrm{d}^{N+1}}{\mathrm{d}t^{N+1}}\tilde{y}(t) = 0 \tag{2.4}$$

Now, in order to get rid of the problem with the initial conditions which plague the classical observers, (by making them converge slowly) we take (2.4) to the Laplace domain, where we can eliminate the initial values of the output and its derivatives. On doing so we get,

$$s^{N+1}\tilde{Y}(s) - s^{N}y(0) - s^{N-1}y^{(1)}(0) - s^{N-2}y^{(2)}(0) - \cdots - sy^{(N-1)}(0) - y^{(N)}(0) = 0 \tag{2.5}$$

Then (2.5) can be written simply as,

$$s^{N+1}\tilde{Y}(s) - \sum_{j=0}^{N} s^{N-j}y^{(j)}(0) = 0 \tag{2.6}$$

To eliminate the initial conditions from (2.6), we differentiate it $N+1$ times with respect to s. This results in the following expression,

$$\frac{\mathrm{d}^{N+1}}{\mathrm{d}s^{N+1}}\left(s^{N+1}\tilde{Y}(s)\right) = 0 \tag{2.7}$$

In the above equation we observe that highest degree of s is $(N+1)$ and the $(N+1)$-th time derivative of $\tilde{y}(t)$ can be expressed as a linear combination of terms of lower order time derivatives. Multiplying both sides of (2.7) by s^{-j} results in a recursive system of equations using which we can determine the requisite time derivatives of $\tilde{y}(t)$ as the value of j ranges $j = 1, 2, \cdots, N+1$ [15].

$$\begin{aligned}
0 &= s^{-j}\frac{\mathrm{d}^{N+1}}{\mathrm{d}s^{N+1}}\left(s^{N+1}\tilde{Y}(s)\right)\\
&= s^{-j}\frac{\mathrm{d}^{N}}{\mathrm{d}s^{N}}\left((N+1)s^{N}\tilde{Y}(s) + s^{N+1}\frac{\mathrm{d}}{\mathrm{d}s}\tilde{Y}(s)\right)\\
&= s^{-j}\frac{\mathrm{d}^{N-1}}{\mathrm{d}s^{N-1}}\left((N+1)(N)s^{N-1}\tilde{Y}(s) + (N+1)s^{N}\frac{\mathrm{d}}{\mathrm{d}s}\tilde{Y}(s) + (N+1)s^{N}\frac{\mathrm{d}}{\mathrm{d}s}\tilde{Y}(s) + s^{N+1}\frac{\mathrm{d}^2}{\mathrm{d}s^2}\tilde{Y}(s)\right)\\
&= s^{-j}\frac{\mathrm{d}^{N-1}}{\mathrm{d}s^{N-1}}\left((N+1)(N)s^{N-1}\tilde{Y}(s) + 2(N+1)s^{N}\frac{\mathrm{d}}{\mathrm{d}s}\tilde{Y}(s) + s^{N+1}\frac{\mathrm{d}^2}{\mathrm{d}s^2}\tilde{Y}(s)\right)
\end{aligned} \tag{2.8}$$

Differentiating $(N-1)$ more times, we get,

$$\begin{aligned}
0 &= s^{-j}\Bigg((N+1)!\tilde{Y}(s) + N\frac{(N+1)!}{1!}s\frac{\mathrm{d}}{\mathrm{d}s}\tilde{Y}(s) + \cdots + \frac{(N+1)(N+1)!}{N!}s^{N}\frac{\mathrm{d}^N}{\mathrm{d}s^N}\tilde{Y}(s)\\
&\qquad + s^{N+1}\frac{\mathrm{d}^{N+1}}{\mathrm{d}s^{N+1}}\tilde{Y}(s)\Bigg)\\
&= \Bigg((N+1)!s^{-j}\tilde{Y}(s) + N\frac{(N+1)!}{1!}s^{1-j}\frac{\mathrm{d}}{\mathrm{d}s}\tilde{Y}(s) + \cdots + \frac{(N+1)(N+1)!}{N!}s^{N-j}\frac{\mathrm{d}^N}{\mathrm{d}s^N}\tilde{Y}(s)\\
&\qquad + s^{N+1-j}\frac{\mathrm{d}^{N+1}}{\mathrm{d}s^{N+1}}\tilde{Y}(s)\Bigg)
\end{aligned} \tag{2.9}$$

This can be further simplified using Leibniz rule which is recalled as follows:

$$\frac{\mathrm{d}^k}{\mathrm{d}s^k}\big((x(s))\,y(s)\big) = \sum_{i=0}^{k}\frac{k!}{k!(i-k)!}x(s)\,\frac{\mathrm{d}^i}{\mathrm{d}s^i}y(s) \tag{2.10}$$

This with

$$\frac{\mathrm{d}^{k-i}}{\mathrm{d}s^{k-i}}\big(x(s)\big) = \frac{\mathrm{d}^{k-i}}{\mathrm{d}s^{k-i}}s^k \;= \frac{k!}{i!}s^i \tag{2.11}$$

applied in (2.9) results in (2.12) as shown in [23]

$$s^{-j}\left(\sum_{i=0}^{N+1}\frac{((N+1)!)^2}{(N+1-i)!(i!)^2}s^i\frac{\mathrm{d}^i}{\mathrm{d}s^i}\tilde{Y}(s)\right) = 0 \tag{2.12}$$

Now taking s^{-j} inside the bracket we get,

$$\left(\sum_{i=0}^{N+1}\frac{((N+1)!)^2}{(N+1-i)!(i!)^2}s^{i-j}\frac{\mathrm{d}^i}{\mathrm{d}s^i}\tilde{Y}(s)\right) = 0 \tag{2.13}$$

We now transform (2.9) and (2.12) from Laplace domain to time domain. But before doing so, we need to keep the following guidelines in mind.

- $\frac{\mathrm{d}^i}{\mathrm{d}s^i}\tilde{Y}(s)$ in Laplace domain transforms to $(-t)^i\tilde{y}(t)$.
- $s^i\tilde{Y}(s)$ transforms to $\frac{\mathrm{d}^i}{\mathrm{d}t^i}\tilde{y}(t)$.
- $s^{-i}\tilde{Y}(s)$ transforms to $\int_0^t\int_0^{t_1}\dots\int_0^{t_{i-1}}\tilde{y}(t)\mathrm{d}t$. This means that $\tilde{y}(t)$ is integrated i times.

To get the state $y(t)$ from the above equations, we consider the case of $j = N+1$. Substituting this in Laplace domain in (2.12) we get,

$$\begin{aligned}
0 &= \Bigg((N+1)!s^{-(N+1)}\tilde{Y}(s) + \frac{(N+1)(N+1)!}{1!}s^{-N}\frac{\mathrm{d}}{\mathrm{d}s}\tilde{Y}(s) + \dots + \frac{(N+1)(N+1)!}{N!}s^{-1}\frac{\mathrm{d}^N}{\mathrm{d}s^N}\tilde{Y}(s) \\
&\quad + \frac{\mathrm{d}^{N+1}}{\mathrm{d}s^{N+1}}\tilde{Y}(s)\Bigg) \\
0 &= \left(\sum_{i=0}^{N}\frac{((N+1)!)^2}{(N+1-i)!(i!)^2}s^{i-j}\frac{\mathrm{d}^i}{\mathrm{d}s^i}\tilde{Y}(s) + \frac{\mathrm{d}^{N+1}}{\mathrm{d}s^{N+1}}\tilde{Y}(s)\right)
\end{aligned} \tag{2.14}$$

Transforming the above to time domain using the aforementioned guidelines,

$$0 = \Bigg((N+1)! \int^{(N+1)} \tilde{y}(t)\,\mathrm{d}t - \frac{(N+1)(N+1)!}{1!} \int^{(N)} t\tilde{y}(t)\,\mathrm{d}t + \cdots + \frac{(N+1)(N+1)!}{N!} \int (-t)^N \tilde{y}(t)\,\mathrm{d}t \\ + (-t)^{N+1}\tilde{y}(t) \Bigg) \tag{2.15}$$

Expressing the above equation in terms of summation of series we have,

$$0 = \left(\sum_{i=0}^{N} \frac{((N+1)!)^2}{(N+1-i)!(i!)^2} \int^{((N+1)-i)} (-t)^i \tilde{y}(t)\,\mathrm{d}t + (-t)^{N+1}\tilde{y}(t) \right) \tag{2.16}$$

We can now manipulate algebraically the equation (2.16), to get (2.17).

$$\tilde{y}(t) = \frac{-1}{(-t)^{N+1}} \Bigg((N+1)! \int^{(N+1)} \tilde{y}(t)\,\mathrm{d}t - \frac{N(N+1)!}{1!} \int^{(N)} t\tilde{y}(t)\,\mathrm{d}t + \cdots + \\ \frac{(N+1)(N+1)!}{N!} \int (-t)^N \tilde{y}(t)\,\mathrm{d}t \Bigg) \tag{2.17}$$

We see that $\tilde{y}(t)$ is a expressed as a function of iterated integrals of $\tilde{y}(t)$. Considering that our aim is to estimate the output and the state of the system from a noisy signal $y_n(t)$, we can replace $\tilde{y}(t)$ present on the left hand side (L.H.S) of the above equation with the output to be estimated, let us say $y(t)$ in this case while in the right hand side (R.H.S), we can replace $\tilde{y}(t)$ with the noisy observed signal $y_n(t)$. Doing so,

$$y(t) = \frac{-1}{(-t)^{N+1}} \Bigg((N+1)! \int^{(N+1)} y_n(t)\,\mathrm{d}t - \frac{N(N+1)!}{1!} \int^{(N)} t y_n(t)\,\mathrm{d}t + \cdots \\ + \frac{(N+1)(N+1)!}{N!} \int (-t)^N y_n(t)\,\mathrm{d}t \Bigg) \tag{2.18}$$

Expressing it in terms of a summation,

$$y(t) = \frac{-1}{(-t)^{N+1}} \left(\sum_{i=0}^{N} \frac{((N+1)!)^2}{(N+1-i)!(i!)^2} \int^{((N+1)-i)} (-t)^i y_n(t)\,\mathrm{d}t \right) \tag{2.19}$$

we see that we can estimate the output signal from a noisy or a perturbed signal $y_n(t)$.

Now, let us tackle the problem of estimating the time derivatives using the above expressions. To do so, let us now examine the case where $j < N+1$, for example $j = N$. Let us substitute $j = N$ in the (2.14). This results in,

$$\begin{aligned}
0 &= \Bigg((N+1)!s^{-(N)}\tilde{Y}(s) + \frac{N(N+1)!}{1!}s^{-(N-1)}\frac{\mathrm{d}}{\mathrm{d}s}\tilde{Y}(s) + \cdots + \frac{(N+1)(N+1)!}{N!}\frac{\mathrm{d}^N}{\mathrm{d}s^N}\tilde{Y}(s) \\
&\quad + \frac{\mathrm{d}^{N+1}}{\mathrm{d}s^{N+1}}s\tilde{Y}(s)\Bigg) \\
0 &= \left(\sum_{i=0}^{N-1}\frac{((N+1)!)^2}{(N+1-i)!(i!)^2}s^{i-j}\frac{\mathrm{d}^i}{\mathrm{d}s^i}\tilde{Y}(s) + \frac{(N+1)(N+1)!}{N!}\frac{\mathrm{d}^N}{\mathrm{d}s^N}\tilde{Y}(s) + \frac{\mathrm{d}^{N+1}}{\mathrm{d}s^{N+1}}s\tilde{Y}(s)\right)
\end{aligned} \tag{2.20}$$

Transforming the above back to time domain using the guidelines mentioned before we have,

$$\begin{aligned}
0 = \Bigg((N+1)!\int^{(N)}\tilde{y}(t)\,\mathrm{d}t - \frac{(N+1)(N+1)!}{1!}\int^{(N-1)}t\tilde{y}(t)\,\mathrm{d}t + \cdots + \frac{(N+1)(N+1)!}{N!}(-t)^N\tilde{y}(t) \\
+ \frac{\mathrm{d}}{\mathrm{d}t}((-t)^{N+1}\tilde{y}(t))\Bigg)
\end{aligned} \tag{2.21}$$

simplifying,

$$0 = \left(\sum_{i=0}^{N-1}\frac{((N+1)!)^2}{(N+1-i)!(i!)^2}\int^{(N-i)}(-t)^i\tilde{y}(t)\,\mathrm{d}t + \frac{(N+1)(N+1)!}{N!}(-t)^N\tilde{y}(t) + \frac{\mathrm{d}}{\mathrm{d}t}((-t)^{N+1}\tilde{y}(t))\right) \tag{2.22}$$

We see that we have $\frac{\mathrm{d}}{\mathrm{d}t}((-t)^{N+1}\tilde{y}(t))$ term in the above expression. Differentiating this term yields,

$$\frac{\mathrm{d}}{\mathrm{d}t}((-t)^{N+1}\tilde{y}(t)) = (N+1)(-1)^N(-t)^N\tilde{y}(t) + (-t)^{N+1}\tilde{y}^{(1)}(t) \tag{2.23}$$

where $\tilde{y}^{(1)}(t)$ denotes $\frac{\mathrm{d}}{\mathrm{d}t}\tilde{y}(t)$.

We then move $\tilde{y}^{(1)}(t)$ from the R.H.S in the above equation (2.21) to the L.H.S of the same equation, and then replace $\tilde{y}^{(1)}(t)$ with $y^{(1)}(t)$, while on the L.H.S we replace $\tilde{y}(t)$ with $y_n(t)$, the noisy observation. Thus, we finally obtain the expression for estimation of the first derivative.

$$\begin{aligned} y^{(1)}(t) = \frac{-1}{(-t)^{N+1}} \Bigg(&\sum_{i=0}^{N-1} \frac{((N+1)!)^2}{(N+1-i)!(i!)^2} \int^{(N-i)} (-t)^i y_n(t)\, \mathrm{d}t \\ &+ \frac{(N+1)(N+1)!}{N!} (-t)^N y_n(t) \; + (N+1)(-1)^N (-t)^N y_n(t) \Bigg) \end{aligned} \tag{2.24}$$

Similarly, we can compute higher time derivatives of the signal using the same procedure illustrated above. For example, $y^{(2)}(t)$ is expressed as follows:

$$\begin{aligned} y^{(2)}(t) = \frac{-1}{(-t)^{N+1}} \Bigg(&\sum_{i=0}^{N-2} \frac{((N+1)!)^2}{(N+1-i)!(i!)^2} \int^{(-i)} (-t)^i y_n(t)\, \mathrm{d}t + \frac{((N+1)!)^2}{((N-1)!)^2} ((-t)^{N-1} y_n(t)) + \\ &+ \frac{(N+1)(N+1)!}{N!} (N(-t)^{N-1} y(t) + \\ &(-t)^N y^{(1)}(t)) + N(N+1)(-t)^{N+1} y(t) + 2(N+1)(-t)^N y^{(1)}(t) \Bigg) \end{aligned} \tag{2.25}$$

The expression for a general $n-th$ order time derivative of truncation order of N of the Taylor series is shown in [23] to be:

$$\begin{aligned} y^{(n)}(t) = &\sum_{l=0}^{n-1} \frac{((N+1)!)}{(N+1-n-l)!(n-l)!} \frac{(N+1-l-1)!}{(N+n-n-1)!} \frac{(-1)^{l-n+1}}{t^{n-l}} y^{(l)}(t) + \\ &+ \sum_{m=1}^{N+1-n} \frac{((N+1)!)^2}{(m+n)!(N+1-m-n)!} \frac{(-1)^{1-m-n}}{t^{N+1}} \int^{(m)} t^{N+1-m-n} y_n(t)\, \mathrm{d}t \end{aligned} \tag{2.26}$$

Here m is the order of repeated integration in the above equation. Equation (2.26) provides a way to calculate derivatives of a signal by integration of a measured signal. This estimation can be better explained with the help of an example, shown in the next subsection.

2.2.1 Example

System

The following third order LTI system is considered:

$$\dot{x} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -10 & 0 \end{bmatrix} x \; ; y = x_1 \; ; \; x(0) = [1, 1, 0] \tag{2.27}$$

with the corresponding characteristic equation

$$y^{(3)}(t) + 10y^{(1)}(t) = 0 \tag{2.28}$$

State Estimation

The signal and the estimation results are shown in the figures 2.1 - 2.3. The noiseless signal is taken over a horizon of 10 seconds and 1000 points are sampled every second, giving us 10000 samples.

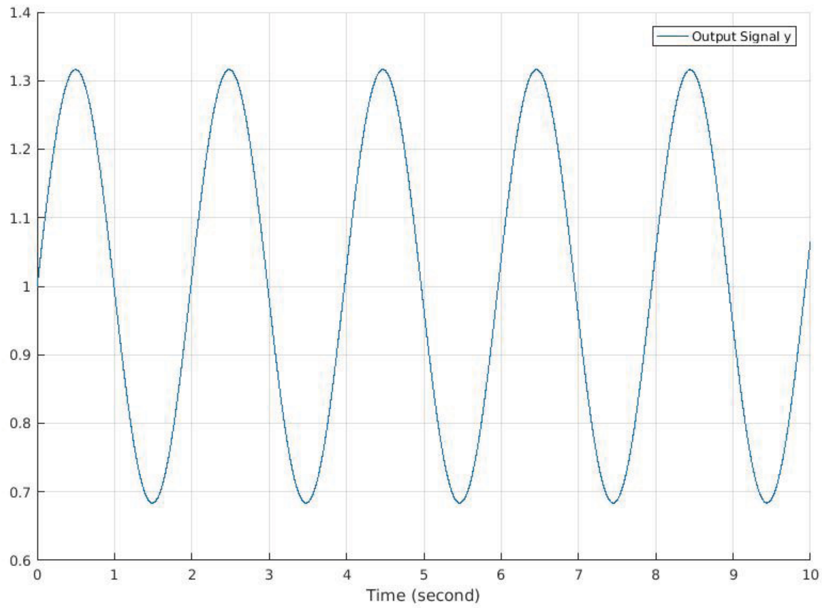

Figure 2.1: True signal $y(t)$

The figures have been enlarged to show clearly the trajectories estimated.

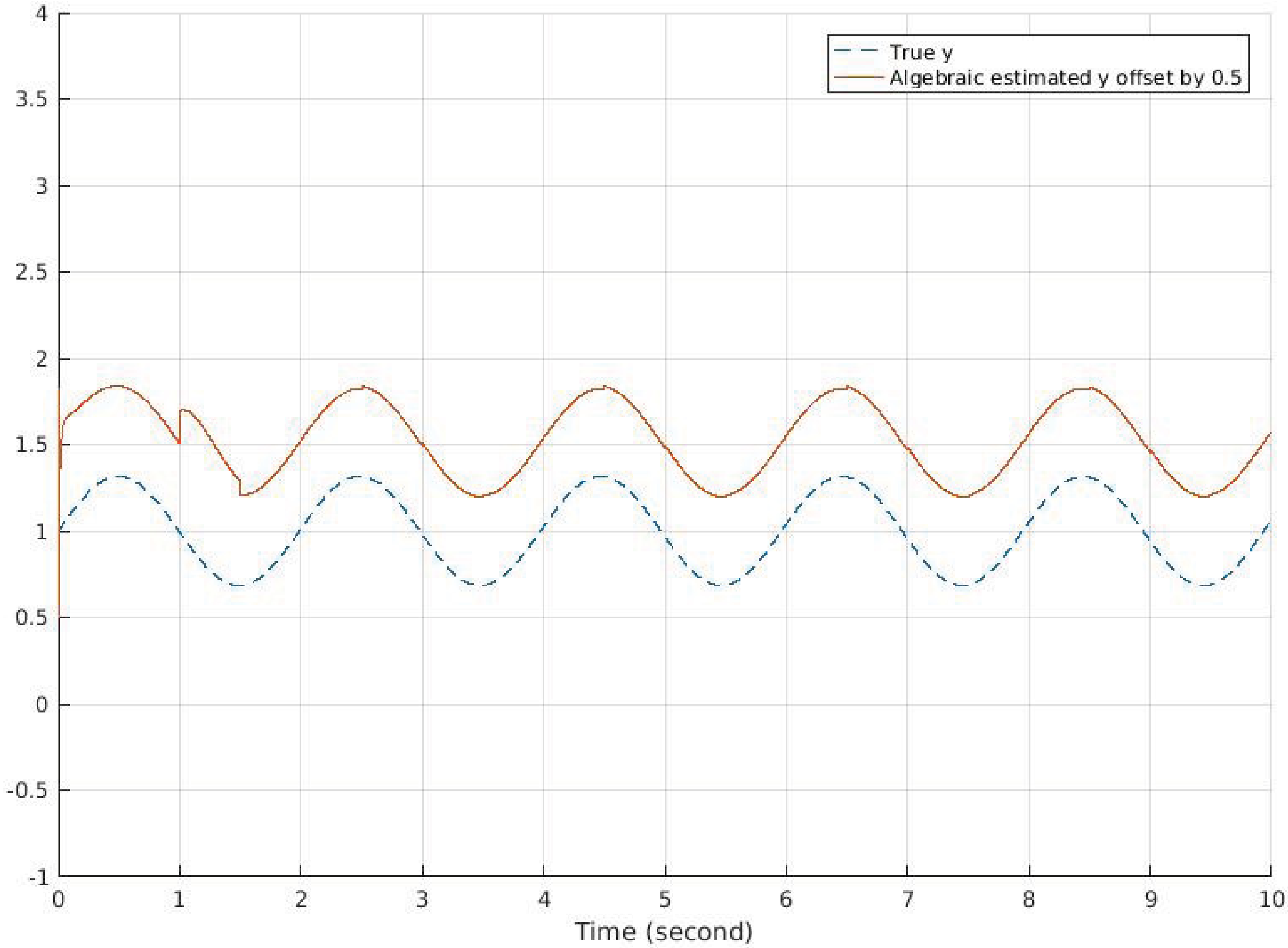

Figure 2.2: True $y(t)$ vs estimated $y(t)$ using the classical algebraic approach offset by +0.5

We see from the above figure that despite there being no initial conditions given to the algebraic estimator, the state and the finite order time derivatives are estimated. But, we can observe that there is a singularity at $t = 0$ and there are small spikes when the estimator is reinitialised at every resetting interval, i.e. after every $1s$ interval. These features are more clearly observable in the figure 2.3.

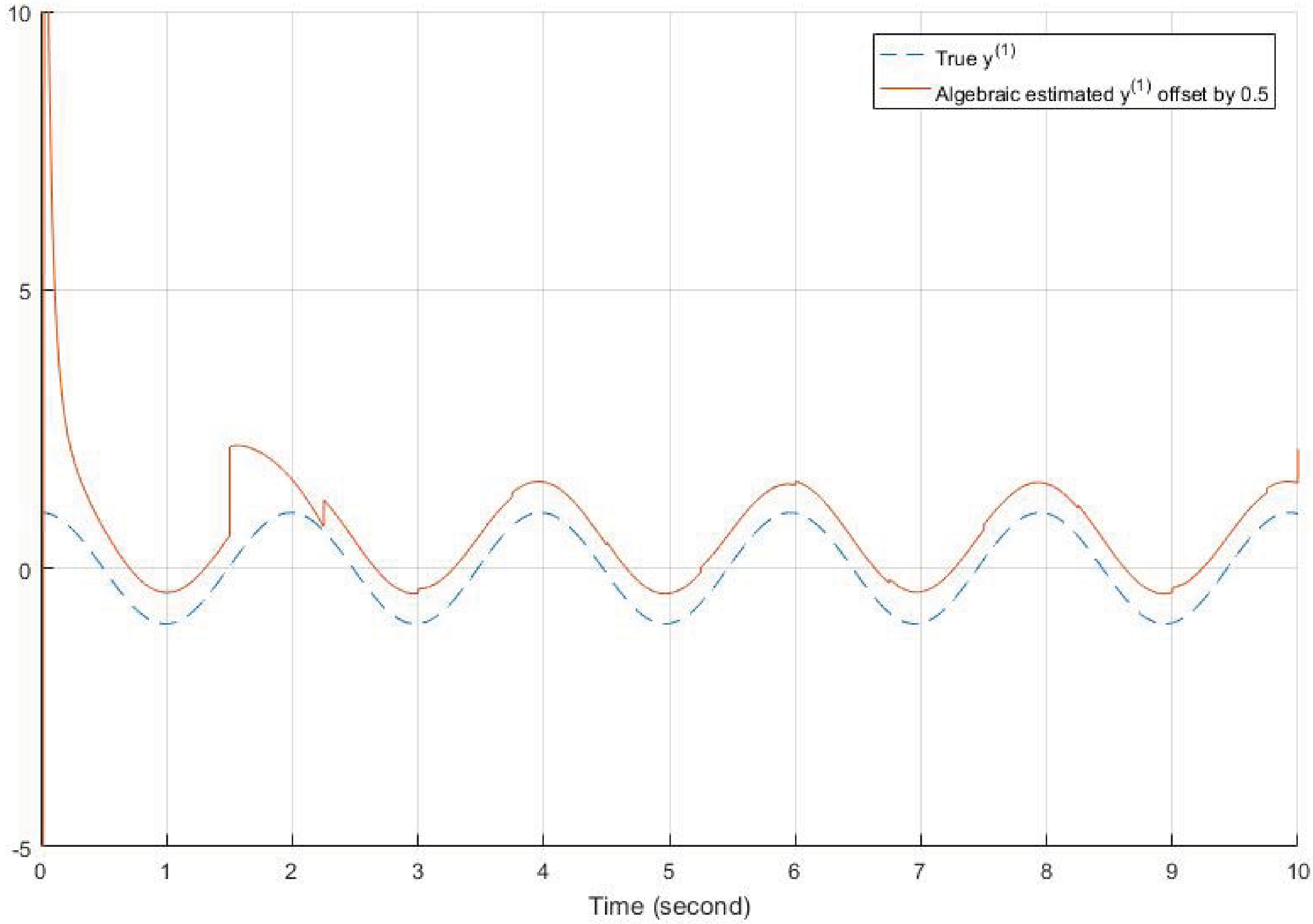

Figure 2.3: True $y^{(1)}(t)$ vs estimated $y^{(1)}(t)$ using the classical algebraic approach offset by +0.5

In figure 2.3 we can observe the singularity much more clearly at $t = 0$ and the spikes due to the reinitiaisation of the estimates are very clearly visible for a noiseless signal case. The performance for a noisy signal case is more flawed and is shown in section § 4.2. The error due to truncation of the Taylor series accumulates and leads to the divergence of the estimation. To counter this divergence, we need to reset the estimation after a fixed interval which is $1s$ here. These are the shortcomings of the algebraic estimator proposed by Fliess et al., and they are discussed in much detail in the section § 2.4.

2.3 Parameter Estimation Using Output Integration

This section discusses algebraic parameter estimation employing *repeated* integration of the system output. The parameter estimation is explained through an example of a general 2nd order system as shown in [21]. But, before proceed to the development of the kernel, we must first understand *linear identifiablilty*. This is an useful and important property for a model to estimate the parameters using a system of equations.

A system is said to be *linearly identifiable* if and only if [21],

$$P \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_r \end{bmatrix} = Q \tag{2.29}$$

where

- P and Q are respectives $r \times r$ and $r \times 1$ matrices.
- the $\det(P) \neq 0$.
- the entries of P and Q belong to the span of the field generated by $y(t)$ and $u(t)$, where $y(t)$ is the output and $u(t)$ is the input to the system.

With this in mind, let us consider the following 2nd order LTI system with input shown below:

$$y^{(2)}(t) + a_1 y^{(1)}(t) + a_0 y(t) = bu(t) + A_0 \tag{2.30}$$

Here a_1, a_0 are the parameters of the A matrix, b is the parameter of the input and A_0 is a constant load perturbation of an unknown magnitude where $A_0 \in \mathrm{R}$.

Transforming the above equation (2.30) into the Laplace domain we get (2.31),

$$s^2 Y(s) - sy(0) - y^{(1)}(0) + a_1(sY(s) - y(0)) + a_0 Y(s) = bU(s) + \frac{A_0}{s} \tag{2.31}$$

multiplying the above equation (2.31) by s we get,

$$s^3Y(s) - s^2y(0) - sy^{(1)}(0) + a_1(s^2Y(s) - sy(0)) + a_0sY(s) = bsU(s) + A_0 \tag{2.32}$$

We now differentiate the above equation (2.32) three times to eliminate the unknown parameter A_0 and the initial conditions $y(0)$, $y^{(1)}(0)$. We get,

$$\begin{aligned} 3b\frac{\mathrm{d}^2U(s)}{\mathrm{d}s^2} + bs\frac{\mathrm{d}^3U(s)}{\mathrm{d}s^3} = 6Y(s) &+ 18s\frac{\mathrm{d}Y(s)}{\mathrm{d}s} + 9s^2\frac{\mathrm{d}^2Y(s)}{\mathrm{d}s^2} + s^3\frac{\mathrm{d}^3Y(s)}{\mathrm{d}s^3} \\ &+ 6a_1\frac{\mathrm{d}Y(s)}{\mathrm{d}s} + 6a_1s\frac{\mathrm{d}^2Y(s)}{\mathrm{d}s^2} + a_1s^2\frac{\mathrm{d}^3Y(s)}{\mathrm{d}s^3} \\ &+ 3a_0\frac{\mathrm{d}^2Y(s)}{\mathrm{d}s^2} + a_0s\frac{\mathrm{d}^3Y(s)}{\mathrm{d}s^3} \end{aligned} \tag{2.33}$$

Multiplying both sides of the resulting expression by s^-3 we have,

$$\begin{aligned} 3s^{-3}b\frac{\mathrm{d}^2U(s)}{\mathrm{d}s^2} + bs^{-2}\frac{\mathrm{d}^3U(s)}{\mathrm{d}s^3} = 6s^{-3}Y(s) &+ 18s^{-2}\frac{\mathrm{d}Y(s)}{\mathrm{d}s} + 9s^{-1}\frac{\mathrm{d}^2Y(s)}{\mathrm{d}s^2} + \frac{\mathrm{d}^3Y(s)}{\mathrm{d}s^3} \\ &+ 6a_1s^{-3}\frac{\mathrm{d}Y(s)}{\mathrm{d}s} + 6a_1s^{-2}\frac{\mathrm{d}^2Y(s)}{\mathrm{d}s^2} + a_1s^{-1}\frac{\mathrm{d}^3Y(s)}{\mathrm{d}s^3} \\ &+ 3a_0s^{-3}\frac{\mathrm{d}^2Y(s)}{\mathrm{d}s^2} + a_0s^{-2}\frac{\mathrm{d}^3Y(s)}{\mathrm{d}s^3} \end{aligned} \tag{2.34}$$

Transforming the above expression back to the time domain using the guidelines in § 2.2 we get the following expression:

$$\begin{aligned} b\big(3\int^{(3)} t^2u(t)\mathrm{d}t - \int^{(2)} t^3u(t)\mathrm{d}t\big) = 6\int^{(3)} y(t)\mathrm{d}t &- 18\int^{(2)} ty(t)\mathrm{d}t + 9\int t^2y(t)\mathrm{d}t - t^3y(t) \\ &+ a_1\big(-6\int^{(3)} ty(t)\mathrm{d}t + 6\int^{(2)} t^2y(t)\mathrm{d}t + \int t^3y(t)\mathrm{d}t\big) \\ &+ a_0\big(3\int^{(3)} t^2y(t)\mathrm{d}t + \int^{(2)} t^3y(t)\mathrm{d}t\big) \end{aligned} \tag{2.35}$$

this can be expressed as relation of the form

$$a_0\pi_1(t) + a_1\pi_2(t) + b\pi_3(t) = q(t) \tag{2.36}$$

with

$$\pi_1(t) = \left[3 \int^{(3)} t^2 y(t)\mathrm{d}t + \int^{(2)} t^3 y(t)\mathrm{d}t\right]$$

$$\pi_2(t) = \left[-6 \int^{(3)} t y(t)\mathrm{d}t + 6 \int^{(2)} t^2 y(t)\mathrm{d}t + \int t^3 y(t)\mathrm{d}t\right]$$

$$\pi_3(t) = \left[3 \int^{(3)} t^2 u(t)\mathrm{d}t - \int^{(2)} t^3 u(t)\mathrm{d}t\right]$$

$$q(t) = \left[6 \int^{(3)} y(t)\mathrm{d}t - 18 \int^{(2)} t y(t)\mathrm{d}t + 9 \int t^2 y(t)\mathrm{d}t - t^3 y(t)\right]$$

Now, in order to identify the parameters a_1, a_0 and b, we need to ensure that the above system is *linearly identifiable*. We need a system of linear equations to make them *linearly identifiable*. And thus, we integrate the (2.36) one time and two times with respect to t. This gives us the following system of linear equation for the estimation of the unknown parameters a_1, a_0 and b.

$$\begin{bmatrix} P_{11}(t) & P_{12}(t) & P_{13}(t) \\ P_{21}(t) & P_{22}(t) & P_{23}(t) \\ P_{31}(t) & P_{32}(t) & P_{33}(t) \end{bmatrix} \begin{bmatrix} a_{0e} \\ a_{1e} \\ b_e \end{bmatrix} = \begin{bmatrix} Q_1(t) \\ Q_2(t) \\ Q_3(t) \end{bmatrix} \tag{2.37}$$

with

$$P_{11}(t) = \pi_1(t),\ P_{12}(t) = \pi_2(t),\ P_{13} = \pi_3(t)$$

$$P_{21}(t) = \int \pi_1(t)\mathrm{d}t,\ P_{22}(t) = \int \pi_2(t)\mathrm{d}t,\ P_{23} = \int \pi_3(t)\mathrm{d}t$$

$$P_{31}(t) = \int^{(2)} \pi_1(t)\mathrm{d}t,\ P_{32}(t) = \int^{(2)} \pi_2(t)\mathrm{d}t,\ P_{33} = \int^{(2)} \pi_3(t)\mathrm{d}t$$

and

$$Q_1(t) = q(t),\ Q_2(t) = \int q(t)\mathrm{d}t,\ Q_3 = \int^{(2)} q(t)\mathrm{d}t$$

Using the above system of equations, we can estimate the unknown parameters while at the same time due to the low pass filtering property, the noise can be alleviated. The above method can be better explained with the help of an example given in the next subsection.

2.3.1 Example

The same system as described in the subsection §§ 2.2.1 is used for parameter estimation albeit for a noiseless case to demonstrate the estimation procedure and its features. In this case, we have chosen every $10th$ sample of the 10000 samples that are present to estimate the parameters. The values of the estimated parameters are as follows.

	a_0	a_1	a_2
True values	0	10	0
Estimated values	0.0153	9.638	-0.0953

Table 2.1: Estimated parameter values using the classical algebraic approach for a noiseless case

Note that the accuracy of estimation is not so good. To better illustrate the effect of the parameter estimation on the estimation of $y(t)$, $y^{(1)}(t)$ and $y^{(2)}(t)$, we use the parameters obtained to plot the graphs of the state and its finite order time derivatives. The graphs are enlarged to show the true as well as trajectory estimated using the paramerters estimated clearly. They are as follows.

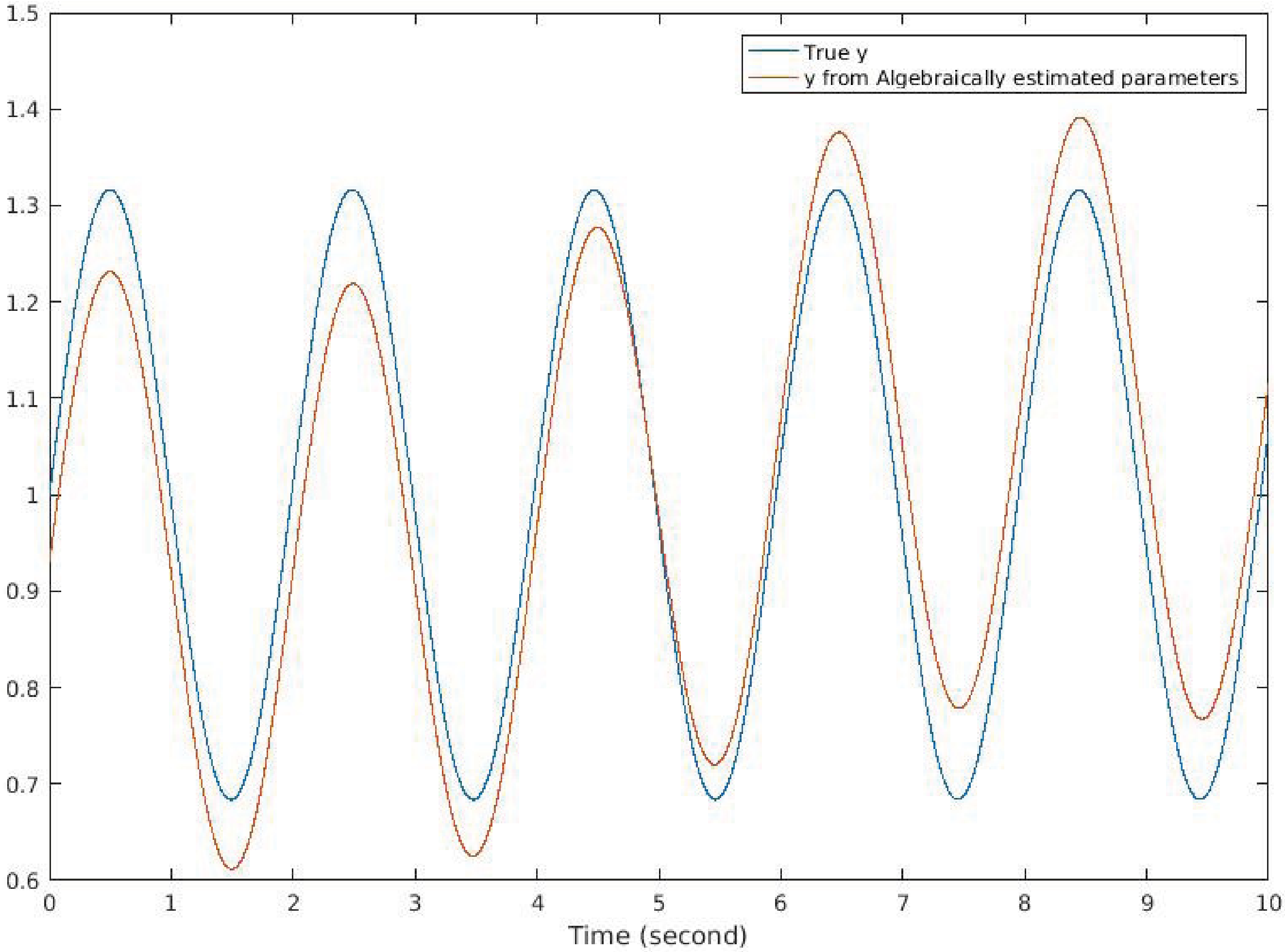

Figure 2.4: True $y(t)$ vs $y(t)$ estimated using parameters estimated using the Algebraic Parameter Estimation procedure

We see that due to the error in estimation, the estimated trajectory of $y(t)$ diverges from the true $y(t)$, coinciding for a very short period of time. Similar results are observed in the subsequent graphs.

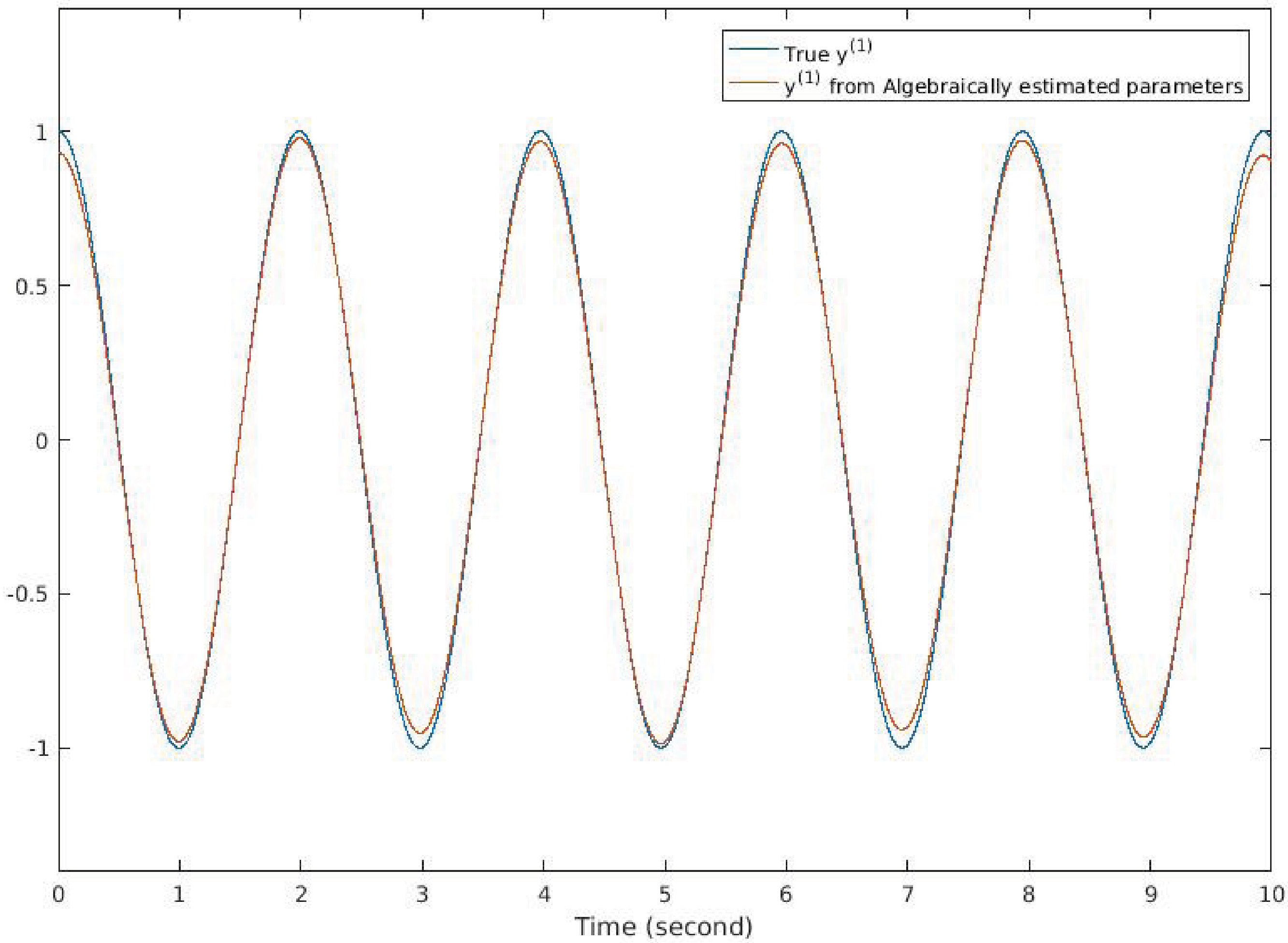

Figure 2.5: True $y^{(1)}(t)$ vs $y^{(1)}(t)$ estimated using parameters estimated using the Algebraic Parameter Estimation procedure

Here in the figure 2.5, the divergence is seen near the peaks of the trajectories, but in the case of noisy signal it is much more pronounced.

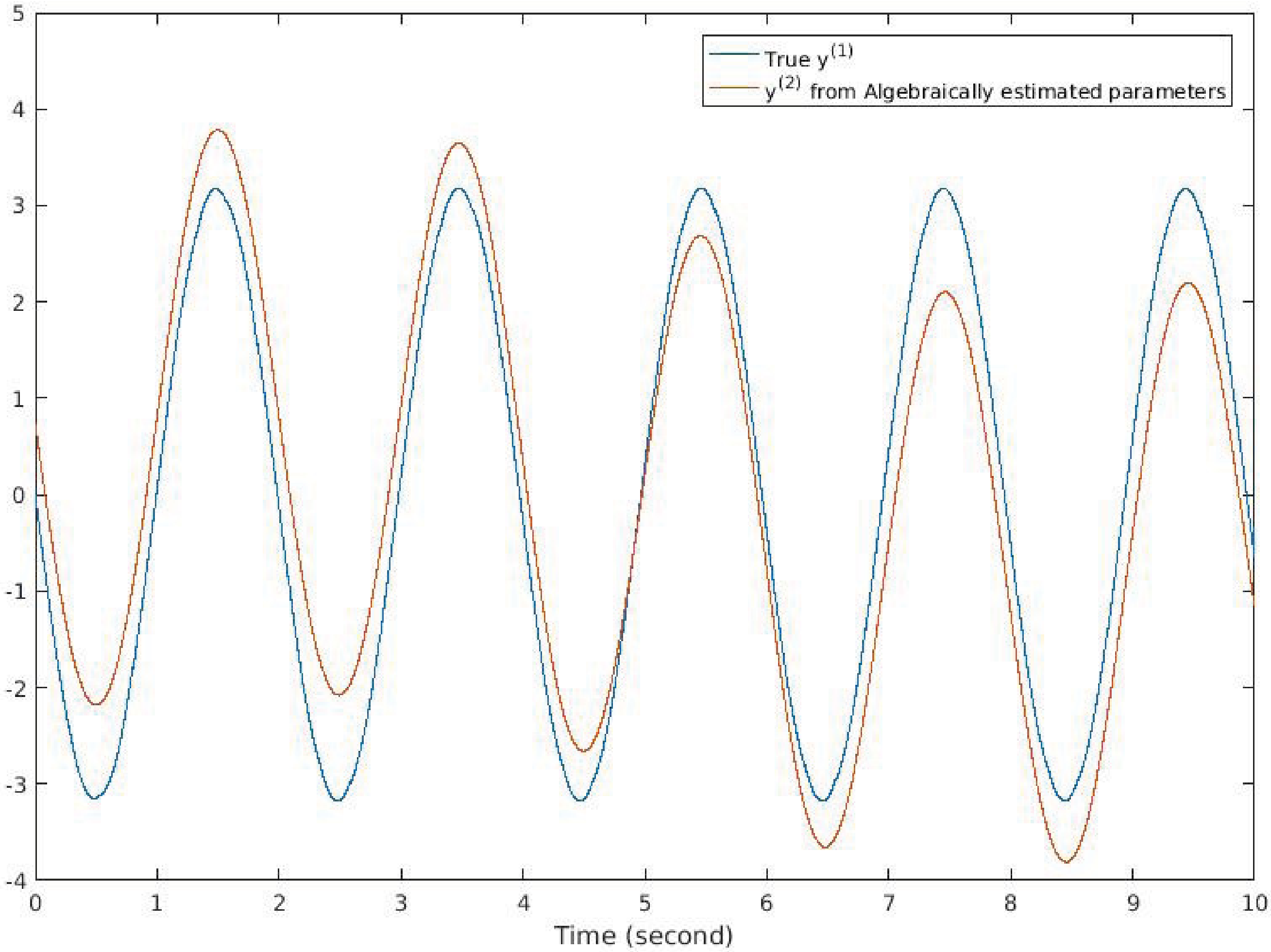

Figure 2.6: True $y^{(2)}(t)$ vs $y^{(2)}(t)$ estimated using parameters estimated using the Algebraic Parameter Estimation procedure

Here in the figure 2.6,the divergence is very conspicuous. As a result of these divergence, we have to design our controllers to compensate for this error. This can make our controller and the system more complicated.

The results for a case with noise are better discussed in Chapter 4.

2.4 Features of Algebraic Differentiation and Estimation

The study of the examples in the previous two sections allows us to notice certain characteristic features of finite-time algebraic observers and estimators described above. They are as follows:

1. The influence of the initial conditions is indeed removed as claimed by Mr. M. Fliess and Mr. Sira-Ramírez. This truly is an improvement over the classical observers which need the right initial conditions. Unknown or incorrect initial conditions invariably entail slow convergence of recursive type of observers. It is exactly this feature that renders the algebraic observers to acquire the classification of "dead-beat" observers

2. It is seen that during estimation of the state, parameter and time derivatives, we integrate $y_n(t)$, the noisy signal. This seems "counter-intuitive" when we consider that we are estimating time derivatives of the state. However by using the Taylor series, we can estimate derivatives by integration.

3. The presence of integrals in the estimation procedure acts like a low pass filter which naturally reduces the influence of noise and external perturbation and hence is good at estimating derivatives from a noisy signal.

4. However, truncation of the Taylor series expansion induces a type of "system modeling error" which propagates leading to divergence of the estimation process on long time horizons. Estimation on long time horizons must thus be accompanied with adequately designed restart procedures after a finite time t.

5. The major drawback of this method is explained as follows. We observe that the R.H.S of the estimation equations is divided by t^{N+1}. So at $t = 0$, the estimate of the state and the parameters cannot be computed because of a singularity. This happens whenever the calculations are reset in the attempt to compensate for the Taylor series truncation error. To overcome this, the researchers proposed a partial

practical “remedy" [3] that is defined as follows:

$$y^{(n)}(t) = \begin{cases} arbitrary \quad \forall t \in [0,\delta) \\ \sum_{l=0}^{n-1} \frac{((N+1)!)}{(N+1-n-l)!(n-l)!} \frac{(N+1-l-1)!}{(N+n-n-1)!} \frac{(-1)^{l-n+1}}{t^{n-l}} y^{(l)}(t) + \\ \qquad + \sum_{m=1}^{N+1-n} \frac{((N+1)!)^2}{(m+n)!(N+1-m-n)!} \frac{(-1)^{1-m-n}}{t^{N+1}} \overset{(m)}{\int} t^{N+1-m-n} y_n(t)\,\mathrm{d}t \\ \quad \forall t \geq \delta \end{cases}$$

Chapter 3

A Double Sided Kernel

3.1 Introduction

In the previous chapter, the algebraic approach of Messrs. M. Fliess and H. Sira-Ramírez was developed and discussed. It was seen that it has some good features but comes with some caveats [2].

- The estimation procedure starts diverging after some time t due to error because of truncation of Taylor series. This compels us to reset the estimation after some period of time t.
- The computation of state and parameter estimates requires division by $time\ t$ and causes singularity at $t = 0$ and whenever we reset the estimation.
- Also, this estimation does not use the underlying dynamics and the knowledge of the system and though some degree of knowledge can be assumed, it is never extended to the estimation of the state and the time derivatives.

To overcome these disadvantages of algebraic estimation, we derive a novel estimation scheme for the state and the derivatives using the the knowledge of the system characteristic equation.

3.2 An Overview of the Double Sided Kernel

With the above motivation, we replace the state equations an with an output reproducing property on an arbitrary time $[t_a, t_b]$ which follows directly from the knowledge of the system characteristic equation. The behavioural model is derived from the differential invariance which is characteristic of the system and eliminating the need of initial conditions and is in the form of a homogeneous Fredholm integral equation of the of the second kind with a Hilbert-Schmidt kernel [1]. The mathematical interpretation as a Reproducing Kernel Hilbert Space (RKHS) of the behavioural model allows us to extract signal and its time derivatives that confirm the system invariance from output measurement subject to noise. The details are presented in the next section.

3.3 Development of the Double Sided Kernel [1]

Before proceeding to the discussion on the development of the method, let us first describe the system whose state we are estimating. We are using this extremely simple model only for simplicity and brevity of exposition. For higher order systems the procedure is identical and its general validity can be proved by mathematical induction.

3.3.1 System Description

Consider the following third order LTI system:

$$\begin{aligned} \dot{x}(t) &= Ax(t) \\ y(t) &= Cx(t) \end{aligned} \tag{3.1}$$

where A, C,$\mathrm{x(t)}$,$\dot{\mathrm{x}}\mathrm{(t)}$ and $\mathrm{y(t)}$ are given as follows

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{3.2}$$

$$C = \begin{bmatrix} c_{11} & c_{12} & c_{13} \end{bmatrix} \tag{3.3}$$

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \tag{3.4}$$

$$\dot{x}(t) = \begin{bmatrix} \dot{x_1}(t) \\ \dot{x_2}(t) \\ \dot{x_3}(t) \end{bmatrix} \tag{3.5}$$

$$y(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix} \tag{3.6}$$

$$\det(A - \lambda I) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} - \begin{vmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{vmatrix}$$

The chracteristic equation for the above system is

$$\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0 = 0 \tag{3.7}$$

Using Cayley Hamilton Theorem, we can rewrite (3.7) as follows.

$$A^3 + a_2A^2 + a_1A + a_0 = 0 \tag{3.8}$$

Now, if we were to multiply (3.8) by $x(t)$ from the right side and multiplying C^T from the left side we get,

$$C^TA^3x(t) + a_2C^TA^2x(t) + a_1C^TAx(t) + a_0C^Tx(t) = 0 \tag{3.9}$$

Differentiating (3.1) with respect to time yields,

$$y(t) = C^T\dot{x}(t) \tag{3.10}$$

Substituting the first part of equation (3.1) into (3.10)

$$y^{(1)}(t) = C^TAx(t) \tag{3.11}$$

Differentiating (3.11) and substituting again two more times we get,

$$y^{(2)}(t) = C^TA^2x(t) \tag{3.12}$$

and

$$y^{(3)}(t) = C^TA^3x(t) \tag{3.13}$$

Now using (3.11) , (3.12), (3.13) and substituting them in (3.9) we finally arrive at,

$$y^{(3)}(t) + a_2y^{(2)}(t) + a_1y^{(1)}(t) + a_0y(t) = 0 \tag{3.14}$$

3.3.2 Derivation of the Double Sided Kernel

We shall now derive the double-sided kernel for a general third order characteristic polynomial described in the previous subsection §§ 3.3.1

$$y^{(3)}(t) + a_2 y^{(2)}(t) + a_1 y^{(1)}(t) + a_0 y(t) = 0 \tag{3.15}$$

on an interval $[a, b]$.

Let us consider two equations obtained from (3.15) by multiplying $(\xi - a)^3$ and $(b - \zeta)^3$ to get the following equations.

$$(\xi - a)^3 y^{(3)}(t) + a_2(\xi - a)^3 y^{(2)}(t) + a_1(\xi - a)^3 y^{(1)}(t) + a_0(\xi - a)^3 y(t) = 0 \tag{3.16}$$

$$(b - \zeta)^3 y^{(3)}(t) + a_2(b - \zeta)^3 y^{(2)}(t) + a_1(b - \zeta)^3 y^{(1)}(t) + a_0(b - \zeta)^3 y(t) = 0 \tag{3.17}$$

Now integrating (3.16) and (3.17) thrice on the interval $[a, a + \tau]$ and $[b - \sigma, b]$. This effectively means that we would be integrating the (3.15) in the forward direction during the interval $[a, a + \tau]$ and in the backward direction on the interval $[b, b - \sigma]$.

Integrating the first term in (3.16) for the first time,

$$\begin{aligned}
\int_a^{a+\tau} (\xi - a)^3 y^{(3)}(\xi)\, \mathrm{d}\xi &= (\xi - a)^3 y^{(2)}(\xi) \mid_a^{a+\tau} - \int_a^{a+\tau} 3(\xi - a)^2 y^{(2)}(\xi)\, \mathrm{d}\xi \\
&= \tau^3 y^{(2)}(a + \tau) - \left[3(\xi - a)^2 y^{(1)}(\xi) \mid_a^{a+\tau} - \int_a^{a+\tau} 6(\xi - a) y^{(1)}(\xi) \mathrm{d}\xi \right] \\
&= \tau^3 y^{(2)}(a + \tau) - 3\tau^2 y^{(1)}(a + \tau) + 6(\xi - a) y^{(1)}(\xi) \mid_a^{a+\tau} - \int_a^{a+\tau} 6y(\xi)\, \mathrm{d}\xi \\
&= \tau^3 y^{(2)}(a + \tau) - 3\tau^2 y^{(1)}(a + \tau) + 6\tau y(a + \tau) - \int_a^{a+\tau} 6y(\xi)\, \mathrm{d}\xi
\end{aligned} \tag{3.18}$$

When we integrate again, the upper limit on the integral becomes a ‘dummy variable’, that is we set $\xi' = a + \tau$ then,

$\tau^3 y^{(2)}(a+\tau)$ is integrated as $(\xi' - a)^3 y^{(2)}(\xi')$

$3\tau^2 y^{(1)}(a+\tau)$ is integrated as $3(\xi' - a)^2 y^{(1)}(\xi')$

$6\tau y(a+\tau)$ is integrated as $6(\xi' - a) y(\xi')$

Integrating (3.18) again,

$$\begin{aligned}
\int_a^{a+\tau}\int_a^{\xi'} (\xi - a)^3 y^{(3)}(\xi)\,\mathrm{d}\xi\mathrm{d}\xi' &= \int_a^{a+\tau} (\xi' - a)^3 y^{(2)}(\xi')\,\mathrm{d}\xi' - \int_a^{a+\tau} 3(\xi' - a)^2 y^{(1')}(\xi)\,\mathrm{d}\xi' \\
&\quad + \int_a^{a+\tau} 6(\xi' - a) y(\xi')\,\mathrm{d}\xi' - \int_a^{a+\tau}\int_a^{\xi'} 6y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi' \\
&= (\xi' - a)^3 y^{(1)}(\xi')\,|_a^{a+\tau} - \int_a^{a+\tau} 3(\xi' - a)^2 y^{(1)}(\xi')\,\mathrm{d}\xi' \\
&\quad - \left[3(\xi' - a)^2 y(\xi')\,|_a^{a+\tau} - \int_a^{a+\tau} 6(\xi' - a) y(\xi')\,\mathrm{d}\xi'\right] \\
&\quad + \int_a^{a+\tau} 6(\xi' - a) y(\xi')\,\mathrm{d}\xi' - \int_a^{a+\tau}\int_a^{\xi'} 6y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi' \\
&= \tau^3 y^{(1)}(a+\tau) - \left[3(\xi' - a)^2 y(\xi')\,|_a^{a+\tau} - \int_a^{a+\tau} 6(\xi' - a) y(\xi')\,\mathrm{d}\xi'\right] \\
&\quad - 3\tau^2 y(a+\tau) + \int_a^{a+\tau} 12(\xi' - a) y(\xi')\,\mathrm{d}\xi' - \int_a^{a+\tau}\int_a^{\xi'} 6y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi' \\
&= \tau^3 y^{(1)}(a+\tau) - 6\tau^2 y(a+\tau) + \int_a^{a+\tau} 18(\xi' - a) y(\xi')\,\mathrm{d}\xi'
\end{aligned} \tag{3.19}$$

As shown earlier, the upper limit again becomes a 'dummy variable' and now we set $\xi'' = a + \tau$. Integrating again for the third time we get,

$$\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} (\xi - a)^3 y^{(3)}(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi'' = \int_a^{a+\tau} (\xi'' - a)^3 y^{(1)}(\xi'')\,\mathrm{d}\xi'' - \int_a^{a+\tau} 6(\xi'' - a)^2 y(\xi'')\,\mathrm{d}\xi''$$

$$
\begin{aligned}
&+\int_a^{a+\tau}\int_a^{\xi''} 18(\xi'-a)y(\xi')\,\mathrm{d}\xi'\mathrm{d}\xi'' - \int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} 6y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi'' \\
&= \tau^3 y(a+\tau) - \int_a^{a+\tau} 3(\xi''-a)^2 y(\xi'')\,\mathrm{d}\xi'' - \int_a^{a+\tau} 6(\xi''-a)y(\xi'')\,\mathrm{d}\xi'' \\
&+\int_a^{a+\tau}\int_a^{\xi''} 18(\xi'-a)y(\xi')\,\mathrm{d}\xi'\mathrm{d}\xi'' - \int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} 6y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi'' \\
&= \tau^3 y(a+\tau) - \int_a^{a+\tau} 9(\xi''-a)^2 y(\xi'')\,\mathrm{d}\xi'' \\
&+\int_a^{a+\tau}\int_a^{\xi''} 18(\xi'-a)y(\xi')\,\mathrm{d}\xi'\mathrm{d}\xi'' - \int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} 6y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''
\end{aligned}
\tag{3.20}
$$

Integrating the second term in (3.16) first time,

$$
\begin{aligned}
\int_a^{a+\tau} a_2(\xi-a)^3 y^{(2)}(\xi)\,\mathrm{d}\xi &= a_2(\xi-a)^3 y^{(1)}(\xi)\,|_a^{a+\tau} - \int_a^{a+\tau} 3(\xi-a)^2 y^{(1)}(\xi)\,\mathrm{d}\xi \\
&= a_2\tau^3 y^{(1)}(a+\tau) - \left[3a_2(\xi-a)^2 y(\xi)\,|_a^{a+\tau} - \int_a^{a+\tau} 6a_2(\xi-a)y(\xi)\,\mathrm{d}\xi\right] \\
&= a_2\tau^3 y^{(1)}(a+\tau) - 3a_2\tau^2 y(a+\tau) + \int_a^{a+\tau} 6a_2(\xi-a)y(\xi)\,\mathrm{d}\xi
\end{aligned}
\tag{3.21}
$$

Using the guidelines as shown previously, we introduce a 'dummy variable' again and integrating (3.21) again,

$$
\begin{aligned}
\int_a^{a+\tau}\int_a^{\xi'} a_2(\xi-a)^3 y^{(2)}(\xi)\,\mathrm{d}\xi\mathrm{d}\xi' = &\int_a^{a+\tau} a_2(\xi'-a)^3 y^{(1)}(\xi')\,\mathrm{d}\xi' - \int_a^{a+\tau} 3a_2(\xi'-a)^2 y(\xi)\,\mathrm{d}\xi' \\
&+ \int_a^{a+\tau}\int_a^{\xi'} 6a_2(\xi-a)y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'
\end{aligned}
$$

$$
\begin{aligned}
&= a_2(\xi'-a)^3 y(\xi')\mid_a^{a+\tau} - \int\limits_a^{a+\tau} 3a_2(\xi'-a)^2 y(\xi')\,\mathrm{d}\xi' \\
&\quad - \int\limits_a^{a+\tau} 3a_2(\xi'-a)^2 y(\xi')\mathrm{d}\xi' + \int\limits_a^{a+\tau}\int\limits_a^{\xi'} 6a_2(\xi-a)y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi' \\
&= \tau^3 a_2(a+\tau) - \int\limits_a^{a+\tau} 6a_2(\xi'-a)^2 y(\xi')\,\mathrm{d}\xi' \\
&\quad + \int\limits_a^{a+\tau}\int\limits_a^{\xi'} 6a_2(\xi-a)y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'
\end{aligned}
\tag{3.22}
$$

Integrating the third time we have,

$$
\begin{aligned}
\int\limits_a^{a+\tau}\int\limits_a^{\xi''}\int\limits_a^{\xi'} a_2(\xi-a)^3 y^{(2)}(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi'' &= \int\limits_a^{a+\tau} a_2(\xi''-a)^3 y(\xi'')\,\mathrm{d}\xi'' - \int\limits_a^{a+\tau}\int\limits_a^{\xi''} 6a_2(\xi'-a)^2 y(\xi')\,\mathrm{d}\xi'\mathrm{d}\xi'' \\
&\quad + \int\limits_a^{a+\tau}\int\limits_a^{\xi''}\int\limits_a^{\xi'} 6a_2(\xi-a)y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''
\end{aligned}
\tag{3.23}
$$

Integrating the third term in (3.16) first time,

$$
\begin{aligned}
\int\limits_a^{a+\tau} a_1(\xi-a)^3 y^{(1)}(\xi)\,\mathrm{d}\xi &= a_1(\xi-a)^3 y(\xi)\mid_a^{a+\tau} - \int\limits_a^{a+\tau} 3a_1(\xi-a)^2 y^{(2)}(\xi)\,\mathrm{d}\xi \\
&= a_1\tau^3 y(a+\tau) - \int\limits_a^{a+\tau} 3a_1(\xi-a)^2 y^{(2)}(\xi)\,\mathrm{d}\xi
\end{aligned}
\tag{3.24}
$$

Integrating (3.24) second time yields,

$$
\int\limits_a^{a+\tau}\int\limits_a^{\xi'} a_1(\xi-a)^3 y^{(1)}(\xi)\,\mathrm{d}\xi\mathrm{d}\xi' = \int\limits_a^{a+\tau} a_1(\xi'-a)^3 y(\xi')\,\mathrm{d}\xi' - \int\limits_a^{a+\tau}\int\limits_a^{\xi'} 3a_1(\xi-a)^2 y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi' \tag{3.25}
$$

Integrating (3.25) the last time,

$$
\begin{aligned}
\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} a_1(\xi-a)^3 y^{(1)}(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi'' = &\int_a^{a+\tau}\int_a^{\xi''} a_1(\xi'-a)^3 y(\xi')\,\mathrm{d}\xi' \\
&-\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} 3a_1(\xi-a)^2 y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''
\end{aligned}
\tag{3.26}
$$

Finally, integrating the last term thrice, we get

$$
\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} a_0(\xi-a)^3 y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi'' \tag{3.27}
$$

Collecting the terms in (3.20) to (3.27) we have,

$$
\begin{aligned}
-\tau^3 y(a+\tau) = &\int_a^{a+\tau}\Big[-9(\xi''-a)^2 + a_2(\xi''-a)^3\Big] y(\xi'')\,\mathrm{d}\xi'' \\
&+\int_a^{a+\tau}\int_a^{\xi''}\Big[+18(\xi'-a) - 6a_2(\xi'-a)^2 + a_1(\xi'-a)^3\Big] y(\xi')\,\mathrm{d}\xi'\mathrm{d}\xi'' \\
&+\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'}\Big[-6 + 6a_2(\xi-a) - 3a_1(\xi-a)^2 + a_0(\xi-a)^3\Big] y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''
\end{aligned}
\tag{3.28}
$$

This can be further simplified by recalling Cauchy formula for repeated integration which can be recollected as follows. Let f be continuous function on the real line, the the nth repeated integral of f based at a.

$$
f^{(-n)}(x) = \int_a^x \int_a^{\sigma_1} \cdots \int_a^{\sigma_{n-1}} f(\sigma_n)\mathrm{s}\sigma_n \cdots \mathrm{d}\sigma_2\mathrm{d}\sigma_1 \tag{3.29}
$$

is given by single integration

$$
f^{(-n)}(x) = \frac{1}{(n-1)!}\int_a^x (x-t)^{n-1} f(t)\mathrm{d}t \tag{3.30}
$$

Now we apply the Cauchy formula for repeated integration stated above on (3.28) while letting $a+\tau = t$, getting $\tau = t-a$. Now we get,

$$\begin{aligned}
-(t-a)^3 y(t) = & \int_a^t \Big[-9(\tau-a)^2 + a_2(\tau-a)^3 \Big] y(\tau)\,\mathrm{d}\tau \\
& + \int_a^t (t-\tau)\Big[18(\tau-a) - 6a_2(\tau-a)^2 + a_1(\tau-a)^3 \Big] y(\tau)\,\mathrm{d}\tau \\
& + \frac{1}{2}\int_a^t (t-\tau)^2 \Big[-6 + 6a_2(\tau-a) - 3a_1(\tau-a)^2 + a_0(\tau-a)^3 \Big] y(\tau)\,\mathrm{d}\tau \\
\triangleq & \int_a^t K_F(t,\tau) y(\tau)\,\mathrm{d}\tau
\end{aligned} \tag{3.31}$$

with $K_F(t,\tau)$ defined by,

$$\begin{aligned}
K_F(t,\tau) \triangleq & \Big[-9(\tau-a)^2 + a_2(\tau-a)^3 \Big] + (t-\tau)\Big[18(\tau-a) - 6a_2(\tau-a)^2 + a_1(\tau-a)^3 \Big] \\
& + (t-\tau)^2 \Big[-6 + 6a_2(\tau-a) - 3a_1(\tau-a)^2 + a_0(\tau-a)^3 \Big]
\end{aligned} \tag{3.32}$$

We now turn our attention to the equation (3.17). We start by integrating the first term in (3.17) first time,

$$\begin{aligned}
\int_{b-\sigma}^{b} (b-\zeta)^3 y^{(3)}(\zeta)\,\mathrm{d}\zeta &= (b-\zeta)^3 y^{(2)}(\zeta) \mid_{b-\sigma}^{b} - \int_{b-\sigma}^{b} 3(b-\zeta)^2 y^{(2)}(\zeta)\,\mathrm{d}\zeta \\
&= -\sigma^3 y^{(2)}(b-\sigma) + \Big[3(b-\zeta)^2 y^{(1)}(\zeta) \mid_{b-\sigma}^{b} - \int_{b-\sigma}^{b} 6(b-\zeta) y^{(1)}(\zeta), \mathrm{d}\zeta \Big] \\
&= -\sigma^3 y^{(2)}(b-\sigma) - 3\sigma^2 y^{(1)}(b-\sigma) - 6(b-\zeta) y^{(1)}(\zeta) \mid_{b-\sigma}^{b} + \int_{b-\sigma}^{b} 6y(\zeta)\,\mathrm{d}\zeta \\
&= -\sigma^3 y^{(2)}(b-\sigma) - 3\sigma^2 y^{(1)}(b-\sigma) - 6\sigma y(b-\sigma) + \int_{b-\sigma}^{b} 6y(\zeta)\,\mathrm{d}\zeta
\end{aligned} \tag{3.33}$$

When integrating again the upper limit on the integral becomes a 'dummy variable' as shown previously, i.e. we set $\zeta' = b - \sigma$ then,
$\sigma^3 y^{(2)}(b-\sigma)$ is integrated as $(b-\zeta')^3 y^{(2)}(\zeta')$
$3\sigma^2 y^{(1)}(b-\sigma)$ is integrated as $3(b-\zeta')^2 y^{(1)}(\zeta')$
$6\sigma y(b-\sigma)$ is integrated as $6(b-\zeta')y(\zeta')$.

We also flip the limits of the integration from $(\zeta' \longrightarrow b)$ to $-(b \longrightarrow \zeta')$ and hence a negative sign is introduced. And when we differentiate $(b-\zeta)^n$ with respect to ζ, we get $-n(b-\zeta)^{n-1}$, again a negative sign is introduced. We have to keep these two critical concepts in mind when evaluating the second and third integrals.
Integrating (3.33) again,

$$
\begin{aligned}
\int_{b-\sigma}^{b}\int_{b-\sigma}^{\zeta'} (b-\zeta)^3 y^{(3)}(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta' &= -\int_{b-\sigma}^{b} (b-\zeta')^3 y^{(2)}(\zeta')\,\mathrm{d}\zeta' - \int_{b-\sigma}^{b} 3(b-\zeta')^2 y^{(1)}(\zeta')\,\mathrm{d}\zeta' \\
&\quad - \int_{b-\sigma}^{b} 6(b-\zeta')y(\zeta')\,\mathrm{d}\zeta' - \int_{b-\sigma}^{b}\int_{b}^{\zeta'} 6y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta' \\
&= -(b-\zeta')^3 y^{(1)}(\zeta')\ |_{b-\sigma}^{b} - \int_{b-\sigma}^{b} 3(b-\zeta')^2 y^{(1)}(\zeta')\,\mathrm{d}\zeta' \\
&\quad - \left[3(b-\zeta')^2 y(\zeta')\ |_{b-\sigma}^{b} + \int_{b-\sigma}^{b} 6(b-\zeta')y(\zeta')\,\mathrm{d}\zeta'\right] \\
&\quad - \int_{b-\sigma}^{b} 6(b-\zeta')y(\zeta')\,\mathrm{d}\zeta' + \int_{b}^{b-\sigma}\int_{b}^{\zeta'} 6y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta' \\
&= \sigma^3 y^{(1)}(b-\sigma) - \left[3(b-\zeta')^2 y(\zeta')\ |_{b-\sigma}^{b} + \int_{b-\sigma}^{b} 6(b-\zeta')y(\zeta')\,\mathrm{d}\zeta'\right] \\
&\quad + 3\sigma^2 y(b-\sigma) + \int_{b}^{b-\sigma} 12(b-\zeta')y(\zeta')\,\mathrm{d}\zeta' + \int_{b}^{b-\sigma}\int_{b}^{\zeta'} 6y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta' \\
&= \sigma^3 y^{(1)}(b-\sigma) - \left[-3\sigma^2 y(b-\sigma) + \int_{b-\sigma}^{b} 6(b-\zeta')y(\zeta')\,\mathrm{d}\zeta'\right]
\end{aligned}
$$

$$+\,3\sigma^2 y(b-\sigma) + \int\limits_{b}^{b-\sigma} 12(b-\zeta')y(\zeta')\,\mathrm{d}\zeta' + \int\limits_{b}^{b-\sigma}\int\limits_{b}^{\zeta'} 6y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'$$

$$= \sigma^3 y^{(1)}(b-\sigma) + 6\sigma^2 y(b-\sigma) + \int\limits_{b}^{b-\sigma} 18(b-\zeta')y(\zeta')\,\mathrm{d}\zeta'$$

$$+ \int\limits_{b}^{b-\sigma}\int\limits_{b}^{\zeta'} 6y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta' \tag{3.34}$$

As explained earlier, the upper limit becomes a 'dummy variable' and now we set $\zeta'' = b-\sigma$. Now integrating (3.34) third time,

$$\int\limits_{b-\sigma}^{b}\int\limits_{b-\sigma}^{\zeta''}\int\limits_{b-\sigma}^{\zeta'} (b-\zeta)^3 y^{(3)}(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta'' = \int\limits_{b-\sigma}^{b} (b-\zeta'')^3 y^{(1)}(\zeta'')\,\mathrm{d}\zeta'' + \int\limits_{b-\sigma}^{b} 6(b-\zeta'')^2 y(\zeta'')\,\mathrm{d}\zeta''$$

$$- \int\limits_{b}^{b-\sigma}\int\limits_{b}^{\zeta''} 18(b-\zeta')y(\zeta')\,\mathrm{d}\zeta'\mathrm{d}\zeta'' - \int\limits_{b}^{b-\sigma}\int\limits_{b}^{\zeta''}\int\limits_{b}^{\zeta'} 6y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta''$$

$$= \left[(b-\zeta'')^3 y(\zeta'')\,\Big|_{b-\sigma}^{b} + \int\limits_{b-\sigma}^{b} 3(b-\zeta'')^2 y(\zeta'')\,\mathrm{d}\zeta''\right]$$

$$- \int\limits_{b}^{b-\sigma} 6(b-\zeta'')^2 y(\zeta'')\,\mathrm{d}\zeta'' - \int\limits_{b}^{b-\sigma}\int\limits_{b}^{\zeta''} 18(b-\zeta')y(\zeta')\,\mathrm{d}\zeta'\mathrm{d}\zeta''$$

$$- \int\limits_{b}^{b-\sigma}\int\limits_{b}^{\zeta''}\int\limits_{b}^{\zeta'} 6y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta''$$

$$= -\sigma^3 y(b-\sigma) - \int\limits_{b}^{b-\sigma} 9(b-\zeta'')^2 y(\zeta'')\,\mathrm{d}\zeta''$$

$$- \int\limits_{b}^{b-\sigma}\int\limits_{b}^{\zeta''} 18(b-\zeta')y(\zeta')\,\mathrm{d}\zeta'\mathrm{d}\zeta'' - \int\limits_{b}^{b-\sigma}\int\limits_{b}^{\zeta''}\int\limits_{b}^{\zeta'} 6y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta'' \tag{3.35}$$

Integrating the second term in (3.17) first time,

$$\begin{aligned}\int_{b-\sigma}^{b} a_2(b-\zeta)^3 y^{(2)}(\zeta)\,\mathrm{d}\zeta &= a_2(b-\zeta)^3 y^{(1)}(\zeta)\,|_{b-\sigma}^{b} + \int_{b-\sigma}^{b} 3(b-\zeta)^2 y^{(1)}(\zeta)\,\mathrm{d}\zeta \\ &= -a_2\sigma^3 y^{(1)}(b-\sigma) + \left[3a_2(b-\zeta)^2 y(\zeta)\,|_{b-\sigma}^{b} + \int_{b-\sigma}^{b} 6a_2(b-\zeta)y(\zeta)\,\mathrm{d}\zeta\right] \\ &= -a_2\sigma^3 y^{(1)}(b-\sigma) - 3a_2\sigma^2 y(b-\sigma) - \int_{b}^{b-\sigma} 6a_2(b-\zeta)y(\zeta)\,\mathrm{d}\zeta\end{aligned} \tag{3.36}$$

Using 'dummy variable' again and integrating (3.36) again,

$$\begin{aligned}\int_{b-\sigma}^{b}\int_{b-\sigma}^{\zeta'} a_2(b-\zeta)^3 y^{(2)}(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta' &= -\int_{b-\sigma}^{b} a_2(b-\zeta')^3 y^{(1)}(\zeta')\,\mathrm{d}\zeta' - \int_{b-\sigma}^{b} 3a_2(b-\zeta')^2 y(\zeta')\,\mathrm{d}\zeta' \\ &\quad - \int_{b-\sigma}^{b}\int_{b}^{\zeta'} 6a_2(b-\sigma)y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta' \\ &= -a_2(b-\zeta')^3 y(\zeta')\,|_{b-\sigma}^{b} - \int_{b-\sigma}^{b} 3a_2(b-\zeta')^2 y(\zeta')\,\mathrm{d}\zeta' \\ &\quad + \int_{b}^{b-\sigma} 3a_2(b-\zeta')^2 y(\zeta') + \int_{b-\sigma}^{b}\int_{b}^{\zeta'} 6a_2(b-\zeta)y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta' \\ &= \sigma^3 a_2(b-\sigma) + \int_{b}^{b-\sigma} 6a_2(b-\zeta')^2 y(\zeta')\,\mathrm{d}\zeta' \\ &\quad + \int_{b-\sigma}^{b}\int_{b}^{\zeta'} 6a_2(b-\zeta)y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\end{aligned} \tag{3.37}$$

Integrating third time we have,

$$
\begin{aligned}
\int_{b-\sigma}^{b}\int_{b-\sigma}^{\zeta''}\int_{b-\sigma}^{\zeta'} a_2(b-\zeta-a)^3 y^{(2)}(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta'' &= \int_{b-\sigma}^{b} a_2(b-\zeta'')^3 y(\zeta'')\,\mathrm{d}\zeta'' \\
&\quad + \int_{b-\sigma}^{b}\int_{b}^{\zeta''} 6a_2(b-\zeta')^2 y(\zeta')\,\mathrm{d}\zeta'\mathrm{d}\zeta'' \\
&\quad + \int_{b-\sigma}^{b}\int_{b}^{\zeta''}\int_{b}^{\zeta'} 6a_2(b-\zeta)y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta'' \\
&= -\int_{b}^{b-\sigma} a_2(b-\zeta'')^3 y(\zeta'')\,\mathrm{d}\zeta'' \\
&\quad - \int_{b}^{b-\sigma}\int_{b}^{\zeta''} 6a_2(b-\zeta')^2 y(\zeta')\,\mathrm{d}\zeta'\mathrm{d}\zeta'' \\
&\quad - \int_{b}^{b-\sigma}\int_{b}^{\zeta''}\int_{b}^{\zeta'} 6a_2(b-\zeta)y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta''
\end{aligned}
\tag{3.38}
$$

Integrating the third term in (3.17) first time,

$$
\begin{aligned}
\int_{b-\sigma}^{b} a_1(b-\zeta)^3 y^{(1)}(\zeta)\,\mathrm{d}\zeta &= a_1(b-\zeta)^3 y(\zeta)\,\Big|_{b-\sigma}^{b} + \int_{b-\sigma}^{b} 3a_1(b-\zeta)^2 y^{(2)}(\zeta)\,\mathrm{d}\zeta \\
&= -a_1\sigma^3 y(b-\sigma) - \int_{b}^{b-\sigma} 3a_1(b-\zeta)^2 y^{(2)}(\zeta)\,\mathrm{d}\zeta
\end{aligned}
\tag{3.39}
$$

Integrating (3.39) second time,

$$
\begin{aligned}
\int_{b-\sigma}^{b}\int_{b-\sigma}^{\zeta'} a_1(b-\zeta)^3 y^{(1)}(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta' &= -\int_{b-\sigma}^{b} a_1(b-\zeta)^3 y(\zeta')\,\mathrm{d}\zeta' - \int_{b-\sigma}^{b}\int_{b}^{\zeta'} 3a_1(b-\zeta)^2 y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta' \\
&= \int_{b}^{b-\sigma} a_1(b-\zeta)^3 y(\zeta')\,\mathrm{d}\zeta' + \int_{b}^{b-\sigma}\int_{b}^{\zeta'} 3a_1(b-\zeta)^2 y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'
\end{aligned}
\tag{3.40}
$$

Integrating (3.40) the final time,

$$\int_{b-\sigma}^{b}\int_{b-\sigma}^{\zeta''}\int_{a}^{\zeta'} a_1(b-\zeta)^3 y^{(1)}(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta'' = \int_{b-\sigma}^{b}\int_{b}^{\zeta''} a_1(b-\zeta')^3 y(\zeta')\,\mathrm{d}\zeta' + \int_{b-\sigma}^{b}\int_{b}^{\zeta''}\int_{b}^{\zeta'} 3a_1(b-\zeta)^2 y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta''$$
$$= -\int_{b}^{b-\sigma}\int_{b}^{\zeta''} a_1(b-\zeta')^3 y(\zeta')\,\mathrm{d}\zeta' - \int_{b}^{b-\sigma}\int_{b}^{\zeta''}\int_{b}^{\zeta'} 3a_1(b-\zeta)^2 y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta'' \tag{3.41}$$

Integrating the last term in (3.17) thrice we have,

$$\int_{b-\sigma}^{b}\int_{b-\sigma}^{\zeta''}\int_{b-\sigma}^{\zeta'} a_0(b-\zeta)^3 y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta'' = -\int_{b}^{b-\sigma}\int_{b}^{\zeta''}\int_{b}^{\zeta'} a_0(b-\zeta)^3 y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta'' \tag{3.42}$$

Collecting the thrice integrated terms from (3.34) to (3.42):

$$\sigma^3 y(b-\sigma) = \int_{b}^{b-\sigma}\Big[-9(b-\zeta'')^2 - a_2(b-\zeta'')^3\Big]y(\zeta'')\,\mathrm{d}\zeta''$$
$$+\int_{b}^{b-\sigma}\int_{b}^{\zeta''}\Big[-18(b-\zeta') - 6a_2(b-\zeta')^2 - a_1(b-\zeta')^3\Big]y(\zeta')\,\mathrm{d}\zeta'\mathrm{d}\zeta''$$
$$+\int_{b}^{b-\sigma}\int_{b}^{\zeta''}\int_{b}^{\zeta'}\Big[-6 - 6a_2(b-\zeta) - 3a_1(b-\zeta)^2 + a_0(b-\zeta)^3\Big]y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta'' \tag{3.43}$$

Now, applying the Cauchy formula for repeated integration stated earlier and simultaneously, while letting $b-\sigma = t$ and $\sigma = b-t$. We get:

$$(b-t)^3 y(t) = \int_{b}^{t}\Big[-9(b-\sigma)^2 - a_2(b-\sigma)^3\Big]y(\sigma)\,\mathrm{d}\sigma$$

$$+\int_b^t (t-\sigma)\Big[-18(b-\sigma)-6a_2(b-\sigma)^2-a_1(b-\sigma)^3\Big]y(\sigma)\,\mathrm{d}\sigma$$
$$+\frac{1}{2}\int_b^t (t-\sigma)^2\Big[-6-6a_2(b-\sigma)-3a_1(b-\sigma)^2-a_0(b-\sigma)^3\Big]y(\sigma)\,\mathrm{d}\sigma \tag{3.44}$$

Rewriting (3.44) after flipping the limits on the integrals,

$$\begin{aligned}(b-t)^3y(t) = &-\int_t^b \Big[-9(b-\tau)^2-a_2(b-\tau)^3\Big]y(\tau)\,\mathrm{d}\tau\\ &-\int_t^b (t-\tau)\Big[-18(b-\tau)-6a_2(b-\tau)^2-a_1(b-\tau)^3\Big]y(\tau)\,\mathrm{d}\tau\\ &-\frac{1}{2}\int_t^b (t-\tau)^2\Big[-6-6a_2(b-\tau)-3a_1(b-\tau)^2-a_0(b-\tau)^3\Big]y(\tau)\,\mathrm{d}\tau\\ \triangleq &\int_b^t K_B(t,\tau)y(\tau)\mathrm{d}\tau\end{aligned} \tag{3.45}$$

with $K_B(t,\tau)$ defined by,

$$\begin{aligned}K_B(t,\tau) \triangleq &\Big[9(b-\tau)^2+a_2(b-\tau)^3\Big]\\ &+(t-\tau)\Big[18(b-\tau)+6a_2(b-\tau)^2+a_1(b-\tau)^3\Big]\\ &+\frac{1}{2}(t-\tau)^2\Big[6+6a_2(b-\tau)+3a_1(b-\tau)^2+a_0(b-\tau)^3\Big]\end{aligned} \tag{3.46}$$

Redefining the partial kernels as: *'forward'* & *'backward'*

$$\begin{aligned}K_F(t,\tau) \triangleq \mu(\tau-a)&\Big[9(\tau-a)^2-a_2(\tau-a)^3\Big]\\ &+(t-\tau)\Big[-18(\tau-a)+6a_2(\tau-a)^2-a_1(\tau-a)^3\Big]\\ &+\frac{1}{2}(t-\tau)^2\Big[6-6a_2(\tau-a)+3a_1(\tau-a)^2-a_0(\tau-a)^3\Big]\end{aligned} \tag{3.47}$$

$$
\begin{aligned}
K_B(t,\tau) \triangleq \mu(b-\tau)&\Big[9(b-\tau)^2 + a_2(b-\tau)^3\Big] \\
&+ (t-\tau)\Big[18(b-\tau) + 6a_2(b-\tau)^2 + a_1(b-\tau)^3\Big] \\
&+ \frac{1}{2}\int_a^t (t-\tau)^2\Big[6 + 6a_2(b-\tau) + 3a_1(b-\tau)^2 + a_0(b-\tau)^3\Big]
\end{aligned}
\tag{3.48}
$$

With,

$$
\mu(\tau - a) = \begin{cases} 1 & : \tau \geq a \\ 0 & : \tau < a \end{cases}
$$

and

$$
\mu(b - \tau) = \begin{cases} 1 & : \tau \leq b \\ 0 & : \tau > b \end{cases}
$$

Now (3.47) and (3.48) can be compactly written as :

$$
(t-a)^3 y(t) = \int_a^t K_F(t,\tau) y(\tau)\,\mathrm{d}\tau \tag{3.49}
$$

$$
(b-t)^3 y(t) = \int_t^b K_B(t,\tau) y(\tau)\,\mathrm{d}\tau \tag{3.50}
$$

Now we define:

$$
K_{DS}(t,\tau) \triangleq \begin{cases} K_F(t,\tau) & : \tau \leq t \\ K_B(t,\tau) & : \tau > t \end{cases} \tag{3.51}
$$

Combining (3.50) and (3.51) side by side while dividing both sides by $[(t-a)^3 + (b-t)^3]$ which is always greater than zero. This results in the following expression:

$$
y(t) = \frac{1}{[(t-a)^3 + (b-t)^3]} \int_a^b K_{DS}(t,\tau) y(\tau)\,\mathrm{d}\tau \tag{3.52}
$$

The recursive expressions of the derivatives can be derived by proceeding similarly as we used to derive the K_{DS}. To obtain the expression for $y^{(1)}(t)$ the equations (3.16) and

(3.17) are integrated twice. We now get,

$$
\begin{aligned}
(t-a)^3 y^{(1)}(t) = 6(t-a)^2 y(t) - a_2(t-a)^3 y(t) \\
&+ \int_t^a \Big[-18(\tau-a) + 6a_2(\tau-a)^2 - a_1(\tau-a)^3 \Big] y(\tau)\mathrm{d}\tau \\
&+ \int_t^a (t-\tau)\Big[6 - 6a_2(\tau-a)^2 - 3a_1(\tau-a)^2 - a_0(\tau-a)^3 \Big] y(\tau)\mathrm{d}\tau
\end{aligned}
\tag{3.53}
$$

and

$$
\begin{aligned}
(b-t)^3 y^{(1)}(t) = -6(b-t)^2 y(t) - a_2(b-t)^3 y(t) \\
&+ \int_t^b \Big[18(b-\tau) + 6a_2(b-\tau)^2 - a_1(b-\tau)^3 \Big] y(\tau)\mathrm{d}\tau \\
&+ \int_t^b (t-\tau)\Big[6 - 6a_2(b-\tau)^2 - 3a_1(b-\tau)^2 - a_0(b-\tau)^3 \Big] y(\tau)\mathrm{d}\tau
\end{aligned}
\tag{3.54}
$$

The final expression for $y^{(1)}(t)$ is obtained by adding the results of (3.53) and (3.54) and dividing by $[(t-a)^3 + (b-t)^3]$.

To get the expression for $y^{(2)}(t)$, (3.16) and (3.17) are integrated once. We get the following expression:

$$
\begin{aligned}
(t-a)^3 y^{(2)}(t) = 3(t-a)^2 y^{(1)}(t) - a_2(t-a)^3 y^{(1)}(t) \\
&- 6(t-a)y(t) + 3a_2(t-a)^2 y(t) - a_1(t-a)^3 y(t) \\
&+ \int_t^a \Big[6 - 6a_2(\tau-a) + 3a_1(\tau-a)^2 - a_0(\tau-a)^3 \Big] y(\tau)\mathrm{d}\tau
\end{aligned}
\tag{3.55}
$$

and

$$
\begin{aligned}
(b-t)^3 y^{(2)}(t) = -3(b-t)^2 y^{(1)}(t) - a_2(b-t)^3 y^{(1)}(t) \\
&- 6(b-t)y(t) - 3a_2(b-t)^2 y(t) - a_1(b-t)^3 y(t) \\
&+ \int_t^b \Big[6 + 6a_2(b-\tau) + 3a_1(b-\tau)^2 + a_0(b-\tau)^3 \Big] y(\tau)\mathrm{d}\tau
\end{aligned}
\tag{3.56}
$$

the expression for $y^{(2)}(t)$ is obtained by adding (3.55) and (3.56) while dividing by a factor of $[(t-a)^3 + (b-t)^3]$.

This features of this novel method of estimation can be better explained with the help of an example, shown in the next subsection.

3.3.3 Example

We have seen the derivation of the double sided kernel K_{DS} for the estimation of state and its finite order time derivatives. Let us now take a simple third order linear time invariant system whose signal is not corrupted by noise. Signal with noise is discussed in the next chapter and compared with other estimators. We shall now estimate the state and the parameters of the model in an effort to show the virtues of the approach.

System

The following third order LTI system is considered:

$$\dot{x} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -10 & 0 \end{bmatrix} x \,; y = x_1 \,;\; x(0) = [1, 1, 0] \tag{3.57}$$

with the corresponding characteristic equation

$$y^{(3)}(t) + 10y^{(1)}(t) = 0 \tag{3.58}$$

State Estimation

The signal and the estimation results are shown in the figures 3.1 - 3.4. The signal is defined over a horizon of 10 seconds and 1000 points are sampled every second, yielding 10000 sample measurements. The estimations have been offset by +0.5 and enlarged so that they are clearly visible.

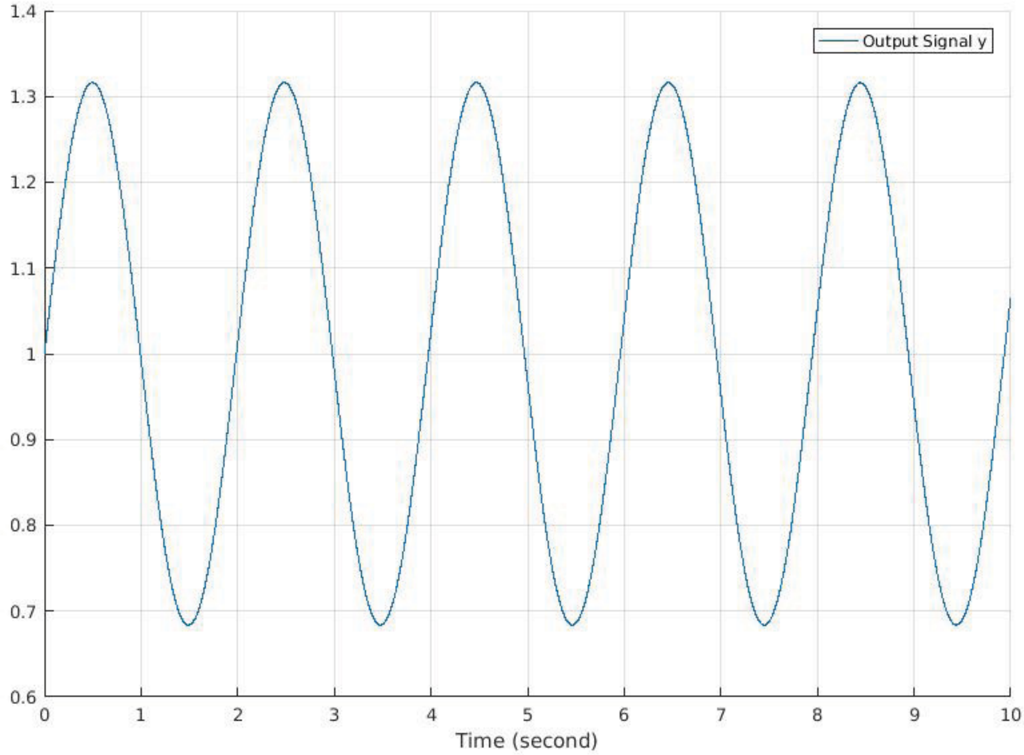

Figure 3.1: True signal $y(t)$

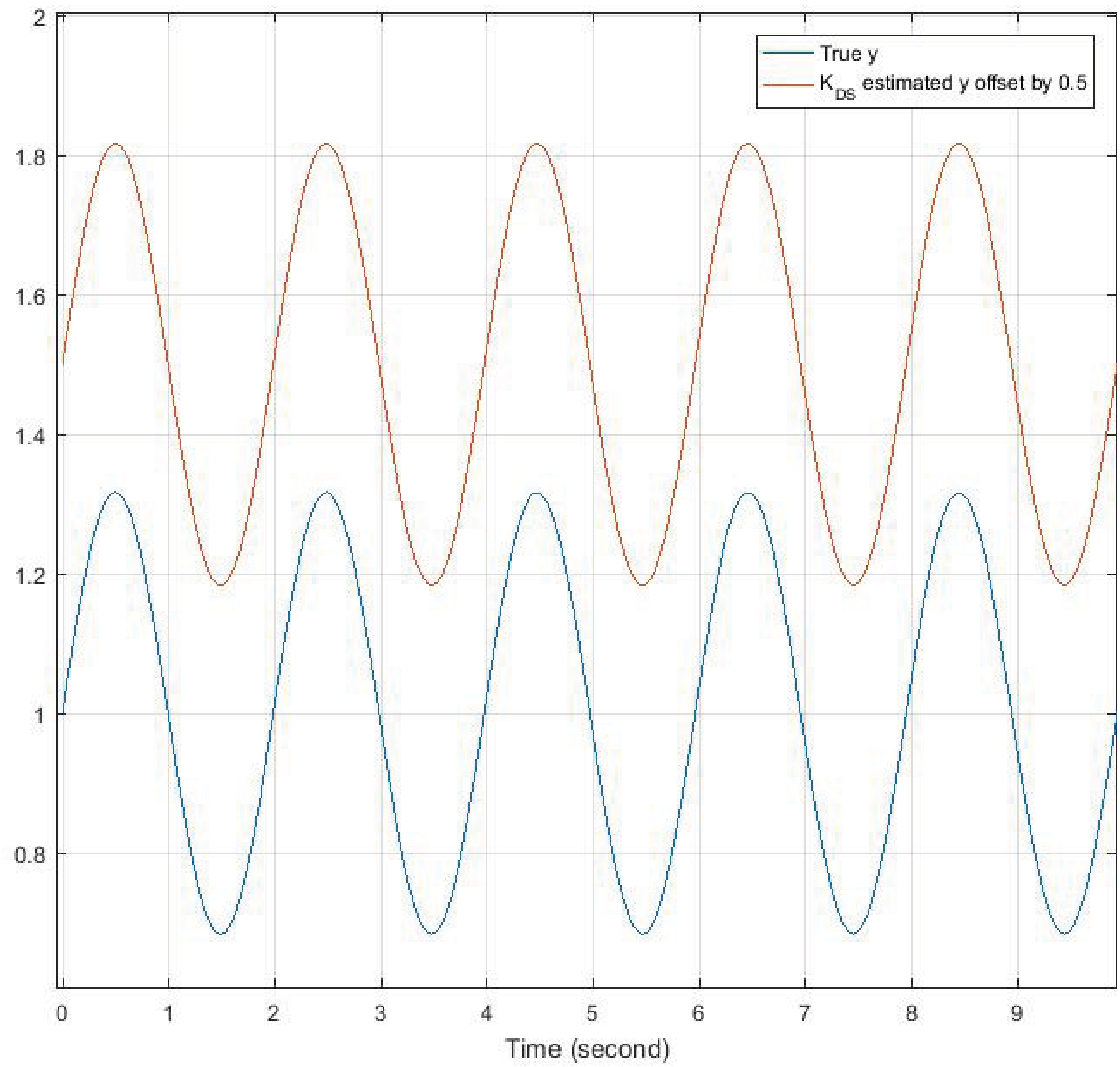

Figure 3.2: True $y(t)$ vs estimated $y(t)$ using the double sided kernel K_{DS} offset by +0.5

The figure 3.2 shows how clearly and well the double sided kernel K_{DS} estimates and there are no spikes when employing the estimator as we use more information about the system to estimate the state instead of taking a truncation of the Taylor series. We do not need any initial conditions to employ this estimator.

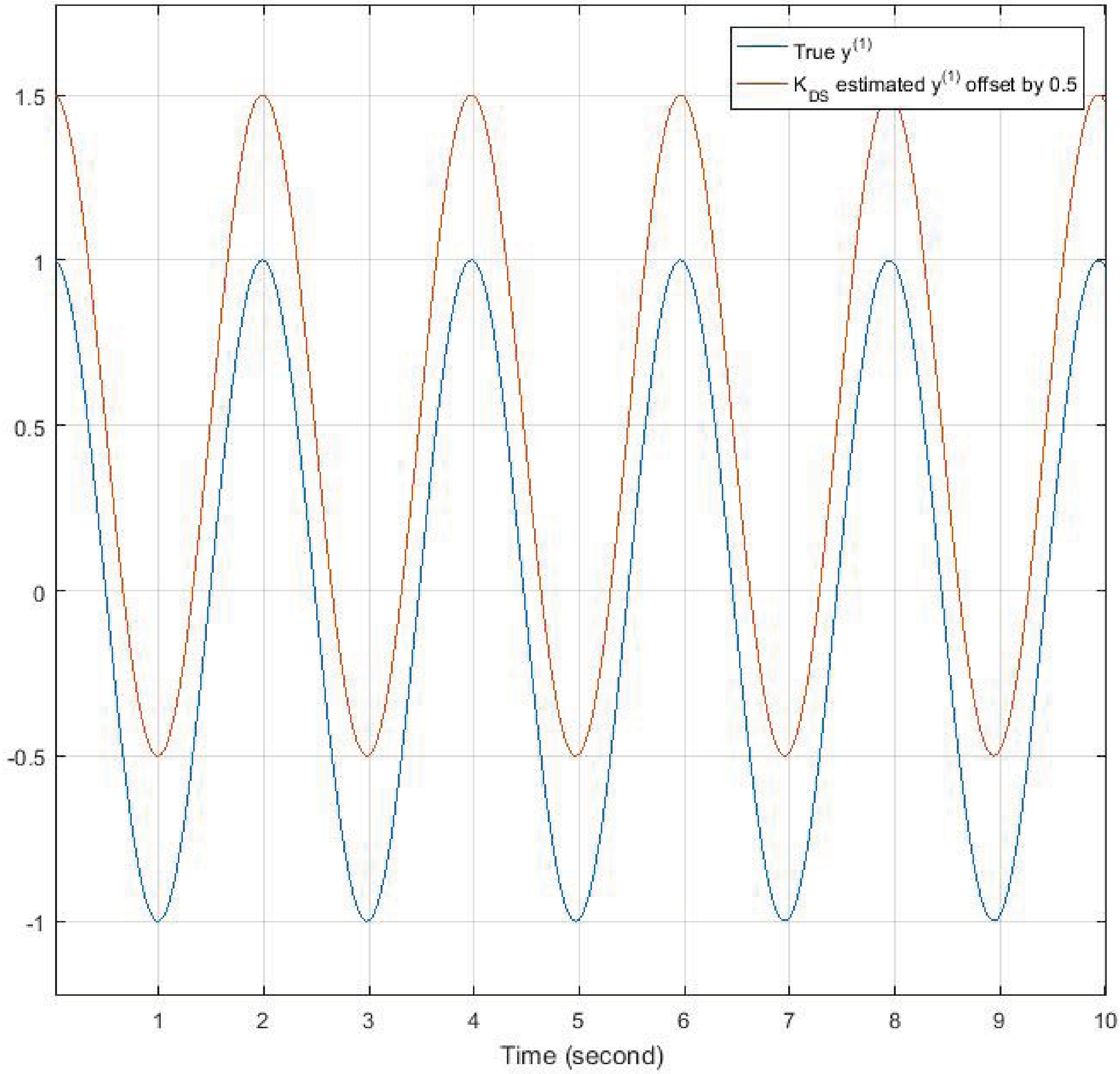

Figure 3.3: True $y^{(1)}(t)$ vs estimated $y^{(1)}(t)$ using the double sided kernel K_{DS} offset by +0.5

As observed in the previous figure, here in figure 3.3 too, the estimate is true to the actual trajectory. On top of this, there are no singularities at the extremities, which means that the K_{DS} is continuous throughout the horizon of estimation, which is better than the classical algebraic approach which is singular at $t = 0$.

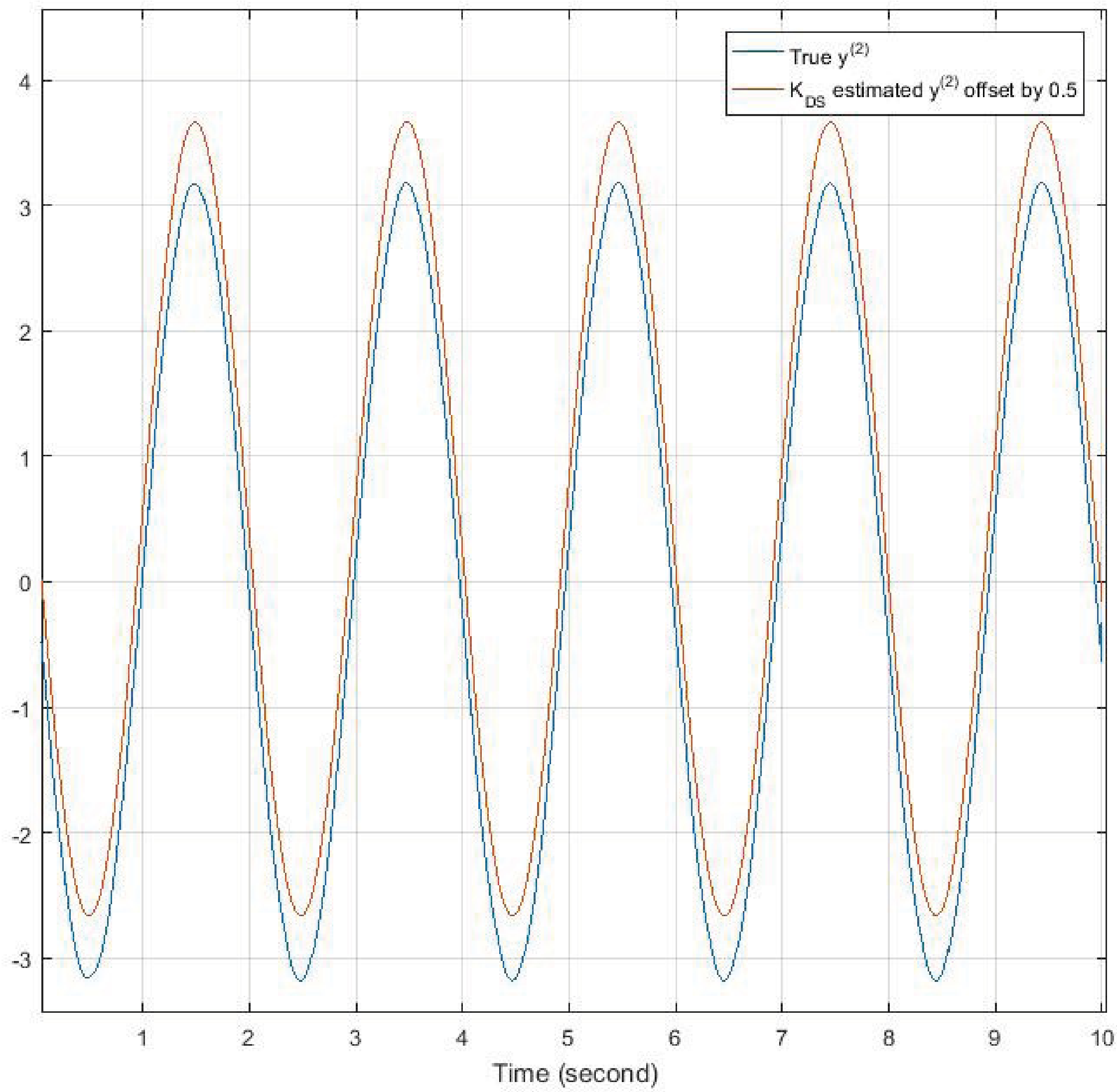

Figure 3.4: True $y^{(2)}(t)$ vs estimated $y^{(2)}(t)$ using the double sided kernel K_{DS} offset by +0.5

Finally in figure 3.4, we observe that there is no divergence in the estimation and hence no need for us to reset calculations at regular intervals.

All these results validate our idea that using the knowledge of the differential invariance of the system (whose output is measured) results in better estimation. A more rigourous examination of the estimator in presence of noise is carried out in Chapter 5.

3.4 Parameter Estimation Using the Double Sided Kernel [2]

This section now discusses the parameter estimation technique using the double sided kernel K_{DS} developed in the subsection §§ 3.3.2 for a generealized third order LTI system for brevity. This can be further adapted for higher order systems. A brief overview of the parameter estimation procedure is as follows.

The development of the Double Sided Kernel K_{DS} shows that it is linear with respect to the system parameters a_i, $i = 0, \cdots, n-1$, i.e. we can write it as follows [2]:

$$y(t) = \int_b^a K_{DS}(t,\tau)y(\tau)\mathrm{d}\tau = \sum_{i=1}^{n} \tilde{a}_i g_i(t,y) + h(t,y) \tag{3.59}$$

where

$$g_i(t,y) \stackrel{def}{=} \int_b^a K_{DS}^{(i)}(t,\tau)y(\tau)\mathrm{d}\tau;\ t \in [a,b] \tag{3.60}$$

where $K_{DS}^{(i)}$; i = 1, n are "component kernels" of K_{DS} and $\tilde{a}_i = a_{i-1}$ for notational convenience. And $h(t,y)$ is the term which is independent of a parameters.

We apply least sqaures fit to the equation (3.59) and given distinct time instants $t_1, \cdots, t_n$ $\in (a,b]$, the (3.59) can be re-written point-wise in the form of a matrix equation [2].

$$\begin{aligned}
&Q(y) = P(y)a\ ;\ \text{mapping trajectories } y : [0,t] \to \mathbb{R} \\
&Q \stackrel{def}{=} \begin{bmatrix} q(t_1) \\ \vdots \\ q(t_n) \end{bmatrix} ; a \stackrel{def}{=} \begin{bmatrix} \tilde{a}_1 \\ \vdots \\ \tilde{a}_n \end{bmatrix} ; P \stackrel{def}{=} \begin{bmatrix} p_1(t_1) \cdots p_n(t_1) \\ \ddots \\ p_1(t_n) \cdots p_n(t_n) \end{bmatrix} \\
&q(t_i) = \int_0^{t_i} y(\sigma)d\sigma;\ p_k(t_i) = \int_0^{t_i} g_k(\sigma,y)\, d\sigma
\end{aligned} \tag{3.61}$$

As we don't have any idea regarding the noise which determine the invertibility of the P matrix, we get the following definition of *linear identifiability*.

Identifiability as defined in [2] for a homogeneous system I=0 as stated in practically linearly identifiable on $[a, b]$ with respect to a particular realization of the output measurement, $y(t)$, $t \in [a, b]$, if and only if there exist distinct time instants $t_1, \cdots, t_n \in (a, b]$ which render the matrix $P(y)$ non-singular. The output trajectories which render det$P \neq 0$ will be called persistent analogous to the terminology used in [21].

The n distinct time instants that are generated can be placed equidistantly on the interval $(a, b]$. As we don't have any knowledge of the noise or the perturbations, the estimate equation (3.61) can be solved in terms of the pseudo-inverse $P^{\dagger}$ of P [2]:

$$a = P^{\dagger}(y)Q(y) \tag{3.62}$$

An alternative method ensuring *linear identifiability* is by integrating (3.60). We get,

$$\int_0^t y(\sigma)\mathrm{d}\sigma = \sum_{i=0}^{n} a_i \int_0^t g_i(\sigma, y)\ \mathrm{d}\sigma;\ t \in [a, b] \tag{3.63}$$

And now given distinct time instants $t_1, \cdots, t_n \in (a, b]$, the (3.63) can be re-written pointwise in the form of a matrix equation [2] as shown in (3.61).

In this thesis, the least sqaures method is used for ensuring *identifiability* for the purpose of estimation. The full derivation of the expressions for parameter estimation for a third order system is given in the next subsection.

3.4.1 Development of Parameter Estimation

Consider a third order linear time invariant system whose characteristic polynomial is as given below:

$$y^{(3)}(t) + a_2 y^{(2)}(t) + a_1 y^{(1)}(t) + a_0 y(t) = 0 \tag{3.64}$$

And whose *forward kernel* is,

$$K_F(t,\tau) = \Big[9(\tau-a)^2 - a_2(\tau-a)^3\Big] + (t-\tau)\Big[-18(\tau-a) + 6a_2(\tau-a)^2 - a_1(\tau-a)^3\Big] + \frac{1}{2}(t-\tau)^2\Big[6 - 6a_2(\tau-a) + 3a_1(\tau-a)^2 - a_0(\tau-a)^3\Big] \tag{3.65}$$

and the *backward kernel* is given as follows,

$$K_B(t,\tau) = \Big[9(b-\tau)^2 + a_2(b-\tau)^3\Big] - (t-\tau)\Big[-18(b-\tau) - 6a_2(b-\tau)^2 - a_1(b-\tau)^3\Big] - \frac{1}{2}(t-\tau)^2\Big[-6 - 6a_2(b-\tau) - 3a_1(b-\tau)^2 - a_0(b-\tau)^3\Big] \tag{3.66}$$

We know from subsection §§ 3.3.2 that,

$$\Big[(t-a)^3 + (b-t)^3\Big]y(t) = \int_a^b K_{DS}(t,\tau)y(\tau)\mathrm{d}\tau \tag{3.67}$$

as well as,

$$\int_a^b K_{DS}(t,\tau)y(\tau)\mathrm{d}\tau = \int_a^t K_F(t,\tau)y(\tau)\mathrm{d}\tau + \int_t^b K_B(t,\tau)y(\tau)\mathrm{d}\tau \tag{3.68}$$

Let,

$$\int_a^b K_{DS}(t,\tau)y(\tau)\mathrm{d}\tau = P \tag{3.69}$$

From (3.67) and (3.69) we get (3.70)

$$\sum\Big[\big[(t-a)^3 + (b-t)^3\big]y(t)\Big] = P \tag{3.70}$$

Bringing P from the R.H.S to the L.H.S in (3.70) and applying the l_2 norm yields,

$$\sum\Big[\Big[\big[(t-a)^3 + (b-t)^3\big]y(t)\Big] - P\Big]^2 = W \tag{3.71}$$

Differentiating (3.71) with respect to a_0,a_1 and a_2 separately we get the equations from (3.72) - (3.74)

$$\sum 2\Bigg[\big[(t-a)^3+(b-t)^3\big]y(t)-P\Bigg]\Bigg[\frac{\mathrm{d}P}{\mathrm{d}a_0}\Bigg]=\frac{\mathrm{d}W}{\mathrm{d}a_0} \tag{3.72}$$

$$\sum 2\Bigg[\big[(t-a)^3+(b-t)^3\big]y(t)-P\Bigg]\Bigg[\frac{\mathrm{d}P}{\mathrm{d}a_1}\Bigg]=\frac{\mathrm{d}W}{\mathrm{d}a_1} \tag{3.73}$$

$$\sum 2\Bigg[\big[(t-a)^3+(b-t)^3\big]y(t)-P\Bigg]\Bigg[\frac{\mathrm{d}P}{\mathrm{d}a_2}\Bigg]=\frac{\mathrm{d}W}{\mathrm{d}a_2} \tag{3.74}$$

from (3.68) and (3.69), P can also be written as,

$$P=\int_a^t K_F(t,\tau)y(\tau)\mathrm{d}\tau+\int_t^b K_B(t,\tau)y(\tau)\mathrm{d}\tau \tag{3.75}$$

We can express K_F and K_B as $M_1(\tau)$ and $M_2(\tau)$ which are in turn expressed as follows,

$$M_1=M_3+M_4+M_5+M_6 \tag{3.76}$$

$$M_2=M_7+M_8+M_9+M_{10} \tag{3.77}$$

where M_3 and M_7 are the expressions containing a_0; M_4 and M_8 are the expressions containing a_1; M_5 and M_9 are the expressions containing a_2 and lastly, M_6 and M_{10} are the constant terms containing none of the parameters. Now we can rewrite (3.72)- (3.74) as follows,

$$\begin{aligned}a_0\sum\Bigg[-\int_a^t\frac{M_3}{a_0}\mathrm{d}\tau-\int_t^b\frac{M_7}{a_0}\mathrm{d}\tau\Bigg]\Bigg[\frac{-\mathrm{d}P}{\mathrm{d}a_0}\Bigg]&+a_1\sum\Bigg[-\int_a^t\frac{M_4}{a_1}\mathrm{d}\tau-\int_t^b\frac{M_8}{a_1}\mathrm{d}\tau\Bigg]\Bigg[\frac{-\mathrm{d}P}{\mathrm{d}a_0}\Bigg]\\&+a_2\sum\Bigg[-\int_a^t\frac{M_5}{a_2}\mathrm{d}\tau-\int_a^t\frac{M_9}{a_2}\mathrm{d}\tau\Bigg]\Bigg[\frac{-\mathrm{d}P}{\mathrm{d}a_0}\Bigg]\\+\sum\Bigg[\big[(t-a)^3+(b-t)^3\big]y(t)-\int_a^t M_6\mathrm{d}\tau-\int_t^b M_{10}\mathrm{d}\tau\Bigg]\Bigg[\frac{-\mathrm{d}P}{\mathrm{d}a_0}\Bigg]&=0\end{aligned} \tag{3.78}$$

$$a_0 \sum \left[-\int_a^t \frac{M_3}{a_0} \mathrm{d}\tau - \int_t^b \frac{M_7}{a_0} \mathrm{d}\tau \right] \left[\frac{-\mathrm{d}P}{\mathrm{d}a_1} \right] + a_1 \sum \left[-\int_a^t \frac{M_4}{a_1} \mathrm{d}\tau - \int_t^b \frac{M_8}{a_1} \mathrm{d}\tau \right] \left[\frac{-\mathrm{d}P}{\mathrm{d}a_1} \right]$$
$$+ a_2 \sum \left[-\int_a^t \frac{M_5}{a_2} \mathrm{d}\tau - \int_a^t \frac{M_9}{a_2} \mathrm{d}\tau \right] \left[\frac{-\mathrm{d}P}{\mathrm{d}a_1} \right]$$
$$+ \sum \left[\left[(t-a)^3 + (b-t)^3 \right] y(t) - \int_a^t M_6 \mathrm{d}\tau - \int_t^b M_{10} \mathrm{d}\tau \right] \left[\frac{-\mathrm{d}P}{\mathrm{d}a_1} \right] = 0 \tag{3.79}$$

$$a_0 \sum \left[-\int_a^t \frac{M_3}{a_0} \mathrm{d}\tau - \int_t^b \frac{M_7}{a_0} \mathrm{d}\tau \right] \left[\frac{-\mathrm{d}P}{\mathrm{d}a_2} \right] + a_1 \sum \left[-\int_a^t \frac{M_4}{a_1} \mathrm{d}\tau - \int_t^b \frac{M_8}{a_1} \mathrm{d}\tau \right] \left[\frac{-\mathrm{d}P}{\mathrm{d}a_2} \right]$$
$$+ a_2 \sum \left[-\int_a^t \frac{M_5}{a_2} \mathrm{d}\tau - \int_a^t \frac{M_9}{a_2} \mathrm{d}\tau \right] \left[\frac{-\mathrm{d}P}{\mathrm{d}a_2} \right]$$
$$+ \sum \left[\left[(t-a)^3 + (b-t)^3 \right] y(t) - \int_a^t M_6 \mathrm{d}\tau - \int_t^b M_{10} \mathrm{d}\tau \right] \left[\frac{-\mathrm{d}P}{\mathrm{d}a_2} \right] = 0 \tag{3.80}$$

Substitute the following into the equations (3.78)-(3.80),

$$W_1 = \frac{\mathrm{d}P}{\mathrm{d}a_0} \tag{3.81}$$

$$W_2 = \frac{\mathrm{d}P}{\mathrm{d}a_1} \tag{3.82}$$

$$W_3 = \frac{\mathrm{d}P}{\mathrm{d}a_2} \tag{3.83}$$

and

$$V_1 = \left[\int_a^t \frac{M_3}{a_0} \mathrm{d}\tau + \int_t^b \frac{M_7}{a_0} \mathrm{d}\tau \right] \tag{3.84}$$

$$V_2 = \left[\int_a^t \frac{M_4}{a_1} \mathrm{d}\tau + \int_t^b \frac{M_8}{a_1} \mathrm{d}\tau \right] \tag{3.85}$$

$$V_3 = \left[\int_a^t \frac{M_5}{a_1} \mathrm{d}\tau + \int_t^b \frac{M_9}{a_1} \mathrm{d}\tau \right] \tag{3.86}$$

$$V_4 = \left[\int_a^t M_6 \mathrm{d}\tau + \int_t^b M_{10} \mathrm{d}\tau \right] \tag{3.87}$$

We can combine equations (3.81)-(3.87) as follows,

$$V_5 = -V_1 \times -W_1 \,; V_6 = -V_2 \times -W_1 \,; V_7 = -V_3 \times -W_1 \,; V_8 = -V_4 \times -W_1 \tag{3.88}$$

$$V_9 = -V_1 \times -W_2 \,; V_{10} = -V_2 \times -W_2 \,; V_{11} = -V_3 \times -W_2 \,; V_{12} = -V_4 \times -W_2 \tag{3.89}$$

$$V_{13} = -V_1 \times -W_3 \,;\, V_{14} = -V_2 \times -W_3 \,;\, V_{15} = -V_3 \times -W_3 \,;\, V_{16} = -V_4 \times -W_3 \quad (3.90)$$

Now, expressing the above equations as a linear combination of the unknown parameters yields,

$$a_0 V_5 + a_1 V_6 + a_2 V_7 = -V_8 \quad (3.91)$$

$$a_0 V_9 + a_1 V_{10} + a_2 V_{11} = -V_{12} \quad (3.92)$$

$$a_0 V_{13} + a_1 V_{14} + a_2 V_{15} = -V_{16} \quad (3.93)$$

The above set of simultaneous linear equations can be represented as a system of matrices as follows,

$$\begin{bmatrix} V_5 & V_6 & V_7 \\ V_9 & V_{10} & V_{11} \\ V_{13} & V_{14} & V_{15} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} -V_8 \\ -V_{12} \\ -V_{16} \end{bmatrix} \quad (3.94)$$

This can be expressed in the form of (3.62) as follows, thereby computing the unknown parameters

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} V_5 & V_6 & V_7 \\ V_9 & V_{10} & V_{11} \\ V_{13} & V_{14} & V_{15} \end{bmatrix}^{-1} \begin{bmatrix} -V_8 \\ -V_{12} \\ -V_{16} \end{bmatrix} \quad (3.95)$$

The following subsection will give a clearer idea of the parameter estimation and its accuracy.

3.4.2 Example

The same system as described in the subsection §§ 3.3.3 is used for parameter estimation albeit for a noiseless case to demonstrate the procedures and features of our method. In this case, though the horizon was kept as 10 seconds and 10000 measurement samples were assumed to be available, only 1 sample for every 100 samples was chosen for estimating the parameters. The values obtained were as follows.

Note that the accuracy of estimation is very good and is very close to the true values.

	a_0	a_1	a_2
True values	0	10	0
Estimated values (Sampled Case)	0.0000	10.0021	-0.0002

Table 3.1: Estimated parameter values using the double sided kernel K_{DS} for a noiseless case

We can use the above parameters simultaneously to estimate the states which is not possible in the procedure of section §§ 2.3.1. The estimated states are shown in the figures 3.5 - 3.6. In this case, we have not offset the individual plots so as to fully demonstrate the high accuracy of the simultaneously performed state and parameter estimation.

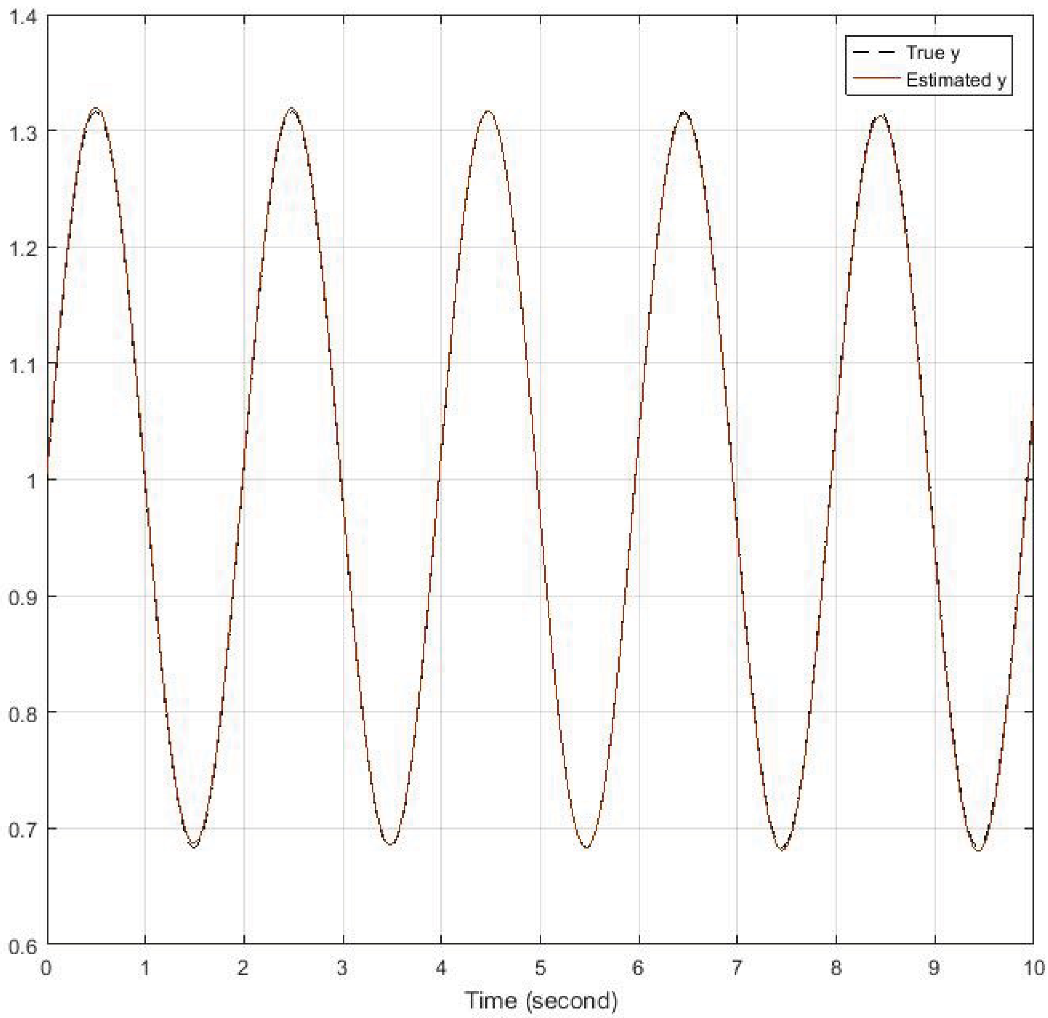

Figure 3.5: True $y(t)$ vs $y(t)$ from parameters estimated by the double sided kernel K_{DS}

The above figure 3.6 shows us the accuracy of the estimator which is very clear. Also, the estimated trajectory and the true trajectory coincide which give us an indication of the precision of the parameter estimation algorithm.

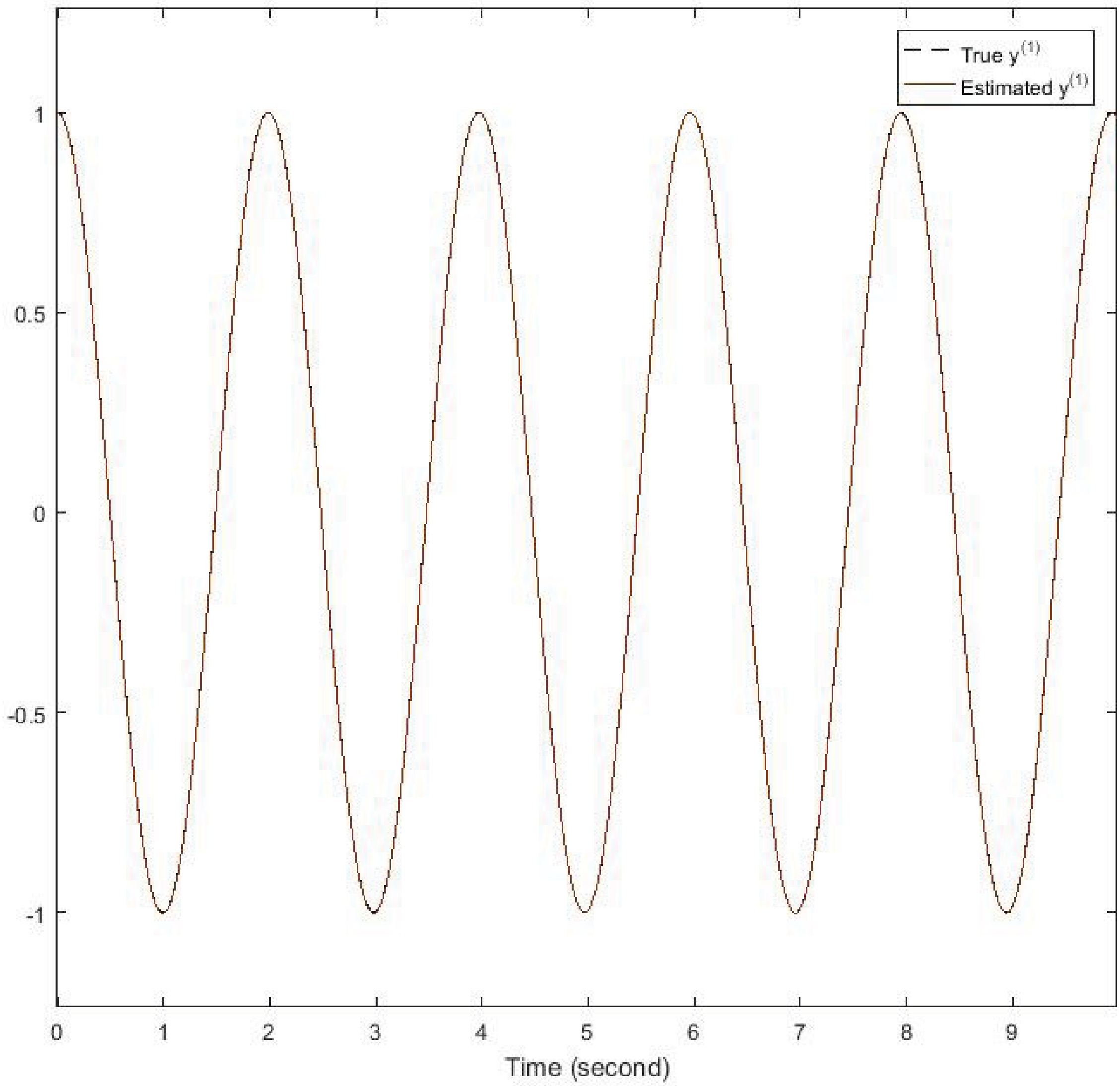

Figure 3.6: True $y^{(1)}(t)$ vs estimated $y^{(1)}(t)$ from parameters estimated by the double sided kernel K_{DS}

Here in the figure 3.7 too, the trajectories overlap giving us a very good idea of the simultaneous parameter and state estimation procedure using the double sided kernel K_{DS}.

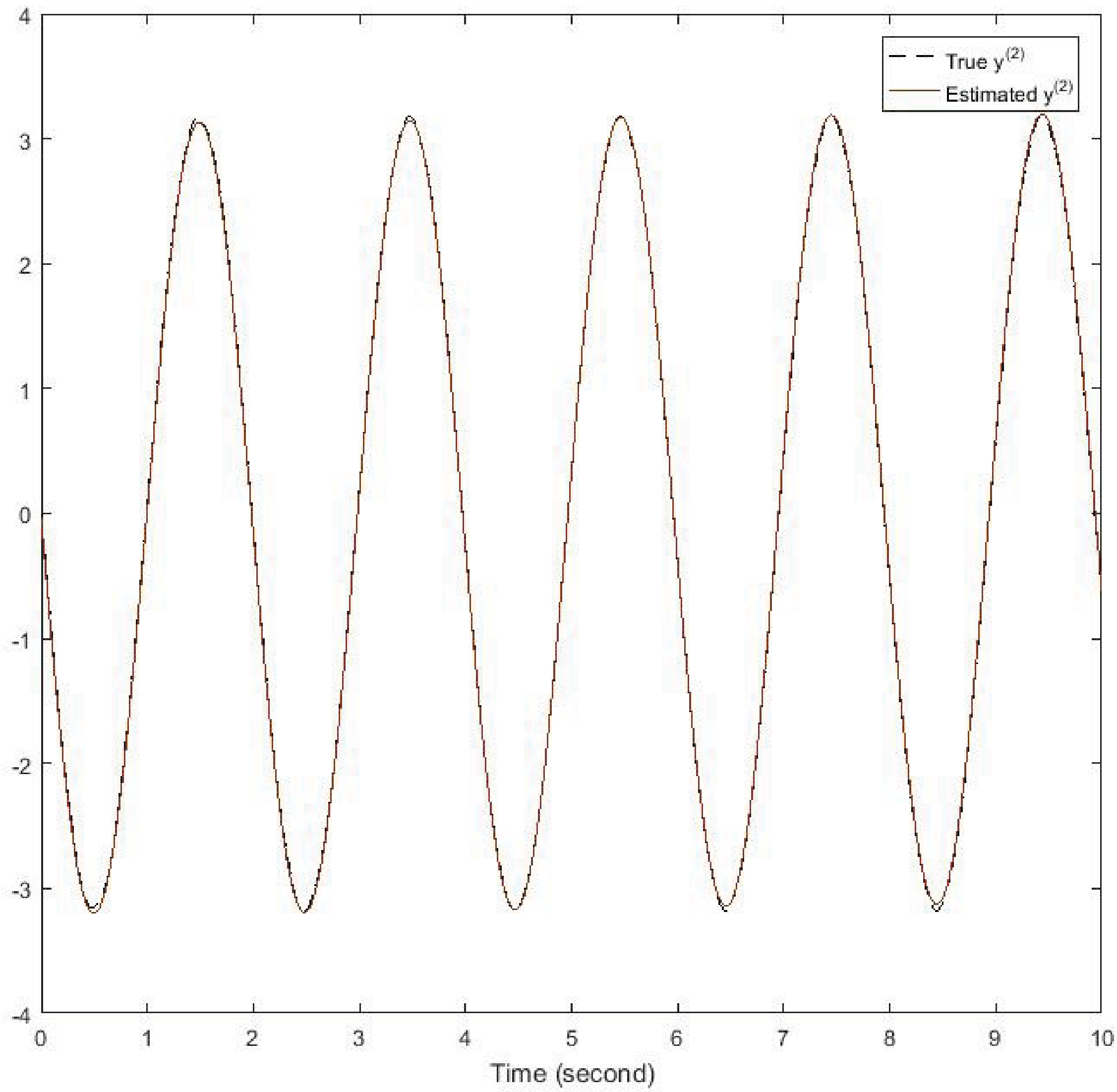

Figure 3.7: True $y^{(2)}(t)$ vs estimated $y^{(2)}(t)$ from parameters estimated by the double sided kernel K_{DS}

The estimator follows the true trajectory without any error in the case of the figure 3.7. This ensures that we do not have to worry about error compensation when we design the controller for our system.

The parameter estimation in the presence of noisy signal is tackled in Chapter 4.

3.5 Features of the Estimation Based on the Double Sided Kernel

Having developed and derived the Double Sided Kernel K_{DS} and the expressions for parameter estimation for a generalised third order linear time invariant system in the sections §§ 3.3.2 and §§ 3.4.1, we can discuss its superior features over the algebraic estimators which were discussed in the § 2.2. These are as follows:

1. The behavioural model, i.e. K_{DS} is derived in a simple manner using the differential invariance characterizing the system (characteristic equation) which eliminates the need for the initial conditions that is otherwise required for state estimation by classical observers such as Kalman filter.

2. We use the knowledge of the dynamical system producing the signal to estimate the derivatives. We use the system knowledge in the form of the Cayley-Hamilton characteristic polynomial, which through manipulation can give us the model in terms of a linear combination of the time derivatives [1].

3. From the construction of the K_{DS} kernel, it can be seen that it does not exhibit any singularities at the extremities of the interval of integration $[a, b]$ because $\min_t[(t-a)^3 + (b-t)^3] = 0.25(b-a)^3 > 0$, that is, the denominator is never zero [1]. Due to this, the double sided integration gives the estimates of the derivatives of the outputs accurate over the entire time interval $[a, b]$. This is a very big advantage over the family of algebraic estimators discussed in § 2.2

4. Since the double sided kernel K_{DS} integrates the measured signal both forwards and backwards while preserving the system invariant in the form of its characteristic equation, the state estimates are not biased with "model error" (unlike in the case of the truncated Taylor series approach). There is thus no need for repeated re-initiation of the estimation process .

5. The parameter estimation can be conducted simultaneously with the state estima-

tion. Assuming system flatness [2], the system states can be expressed as a function of the output and/or its time derivatives. Using the kernels in (3.53) - (3.56) we can compute the derivatives following the parameter identification.

Chapter 4

Results

In the previous two chapters, we have shown the development of the classical algebraic approach as well as *our new approach* employing the proposed double sided kernel K_{DS} for state and its finite order time derivatives as well as parameter estimation. We have also had the opportunity to see the drawbacks of the classical algebraic approach and the virtues of our double sided kernel K_{DS} in application to system state and parameter estimation albeit in the case of a noiseless signal. We devote this chapter towards comparing the two methods with the action of the Kalman filter as an example of a recursive observer.

We take an example of a LTI system whose output is distorted by an addition of Gaussian white noise. We take two cases, the first case is when the signal to noise ratio(SNR) is high, that is, the noise is low and when the SNR is high, that is, the noise is high. In both cases, we give wrong initial conditions to the estimators as well as Kalman filter to show that the estimation is independant of initial conditions in the case of the algebraic approach as well as the double sided kernel K_{DS} and the dependance on the initial conditions by the Kalman filter.

4.1 System Description

We shall first describe the system before proceeding to the comparison. A 3rd order LTI system is considered as follows:

$$\begin{aligned} \dot{x}(t) &= Ax(t) \\ y(t) &= Cx(t) \end{aligned} \tag{4.1}$$

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -10 & 0 \end{bmatrix} \tag{4.2}$$

$$C = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \tag{4.3}$$

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \tag{4.4}$$

$$\dot{x}(t) = \begin{bmatrix} \dot{x_1}(t) \\ \dot{x_2}(t) \\ \dot{x_3}(t) \end{bmatrix} \tag{4.5}$$

$$y(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix} \tag{4.6}$$

The correct initial condition for this system is:

$$x(t) = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \tag{4.7}$$

The model of the above system from the characteristic polynomial is as given below:

$$y^{(3)}(t) + 10y^{(1)}(t) - y(t) = 0 \tag{4.8}$$

Let us consider $y_n(t)$ as the "noisy measurement" of the output signal $y(t)$ which has been corrupted by Gaussian noise. We consider a sampling time of 10 seconds and sample 1000 points in an interval of 1 second, therefore yielding 10000 samples. The standard Kalman filter employing the system was initialised with the incorrect (unknown) intial conditions of $[-0.5, -0.5, -0.5]$, but the correct covariance of noise is left to the Kalman filter to estimate. The resulting "noisy" mesaurement was then used to compute the estimates $y_{est}(t)$, $y_{est}^{(1)}(t)$ and $y_{est}^{(2)}(t)$. Romberg integration was used to evaluate the integral transforms.

4.2 State and Time Derivative Estimation

We know from the section § 2.2 that higher the order of truncation, the better are the estimation of states. This is due to the reduction in the truncation error in the Taylor series. Hence we consider a sixth order of truncation in the Taylor series for the classical algebraic approach. The expressions for the estimation of states and time derivatives in the algebraic approach are as follows:

$$y_{est}(t) = \frac{1}{t^7}\Big[5040\int^{(7)} y_n(t) - 35280\int^{(6)} t y_n(t) + 52920\int^{(5)} t^2 y_n(t) - 29400\int^{(4)} t^3 y_n(t) + 7350\int^{(3)} t^4 y_n(t) - 882\int^{(2)} t^5 y_n(t) + 49\int t^6 y_n(t)\Big] \tag{4.9}$$

$$y^{(1)}_{est}(t) = \frac{1}{t^7}\Big[5040\int^{(6)} y_n(t) - 35280\int^{(5)} t y_n(t) + 52920\int^{(4)} t^2 y_n(t) - 29400\int^{(3)} t^3 y_n(t) + 7350\int^{(2)} t^4 y_n(t) - 882\int t^5 y_n(t) + 49 t^6 y_{est}(t)\Big] \tag{4.10}$$

$$y^{(2)}_{est}(t) = \frac{1}{t^7}\Big[5040\int^{(5)} y_n(t) - 35280\int^{(4)} t y_n(t) + 52920\int^{(3)} t^2 y_n(t) - 29400\int^{(2)} t^3 y_n(t) + 7350\int t^4 y_n(t) - 630 t^5 y_n(t) + 35 t^6 y^{(1)}_{est}(t)\Big] \tag{4.11}$$

The expressions for the estimation of states and time derivatives in the double sided kernel K_{DS} are as follows:

$$\begin{aligned} y_{est}(t) = \frac{1}{[(t-a)^3+(b-t)^3]}\Bigg(\int_a^t \Big[&9(\tau-a)^2 y(\tau)\mathrm{d}\tau + (t-\tau)\big[-18(\tau-a) - 10(\tau-a)^3\big]y(\tau)\mathrm{d}\tau \\ &+ \frac{1}{2}(t-\tau)^2\big[6+30(\tau-a)^2+(\tau-a)^3\big]y(\tau)\mathrm{d}\tau\Big] \\ &+ \int_t^b \Big[9(b-\tau)^2 y(\tau)\mathrm{d}\tau + (t-\tau)\big[18(b-\tau)+10(b-\tau)^3\big]y(\tau)\mathrm{d}\tau \\ &+ \frac{1}{2}(t-\tau)^2\big[6+30(b-\tau)^2-(b-\tau)^3\big]y(\tau)\mathrm{d}\tau\Big]\Bigg) \end{aligned} \tag{4.12}$$

$$
\begin{aligned}
y^{(1)}_{est}(t) = \frac{1}{[(t-a)^3+(b-t)^3]}\Bigg(& 6(t-a)^2 y_{est}(t) + \int\limits_a^t \Bigg[\big[-18(\tau-a) - 10(\tau-a)^3\big] y(\tau)\mathrm{d}\tau \\
& + (t-\tau)\big[6 + 30(\tau-a)^2 + (\tau-a)^3\big] y(\tau)\mathrm{d}\tau\Bigg] \\
& - 6(b-t)^2 y_{est}(t) + \int\limits_t^b \Bigg[\big[18(b-\tau) + 10(b-\tau)^3\big] y(\tau)\mathrm{d}\tau \\
& + \frac{1}{2}(t-\tau)^2\big[6 + 30(b-\tau)^2 - (b-\tau)^3\big] y(\tau)\mathrm{d}\tau\Bigg]\Bigg)
\end{aligned}
\tag{4.13}
$$

$$
\begin{aligned}
y^{(2)}_{est}(t) = \frac{1}{[(t-a)^3+(b-t)^3]}\Bigg(& 3(t-a)^2 y^{(1)}_{est}(t) - 6(t-a) y_{est}(t) - 10(t-a)^3 y_{est}(t) \\
& + \int\limits_a^t \Bigg[\big[6 + 30(\tau-a)^2 + (\tau-a)^3\big] y(\tau)\mathrm{d}\tau\Bigg] \\
& - 3(b-t)^2 y^{(1)}_{est}(t) - 6(b-t) y_{est}(t) - 10(b-t)^3 y_{est}(t) \\
& + \int\limits_t^b \Bigg[\big[6 + 30(b-\tau)^2 - (b-\tau)^3\big] y(\tau)\mathrm{d}\tau\Bigg]\Bigg)
\end{aligned}
\tag{4.14}
$$

In the above equations, the estimated states are denoted as $y_{est}(t)$, $y^{(1)}_{est}(t)$ and $y^{(2)}_{est}(t)$. We estimate the subsequent outputs $y^{(1)}_{est}(t)$ and $y^{(2)}_{est}(t)$ using the previously estimated outputs. Due to the system being differentially flat, the states can be expressed as a function of the output and its derivatives. We now take the first case of high SNR.

4.2.1 Case I: Low Noise - 50SNR

The simulated system output was distorted by adding a Gaussian noise of 50 SNR and the signal is as shown in figure 4.1.

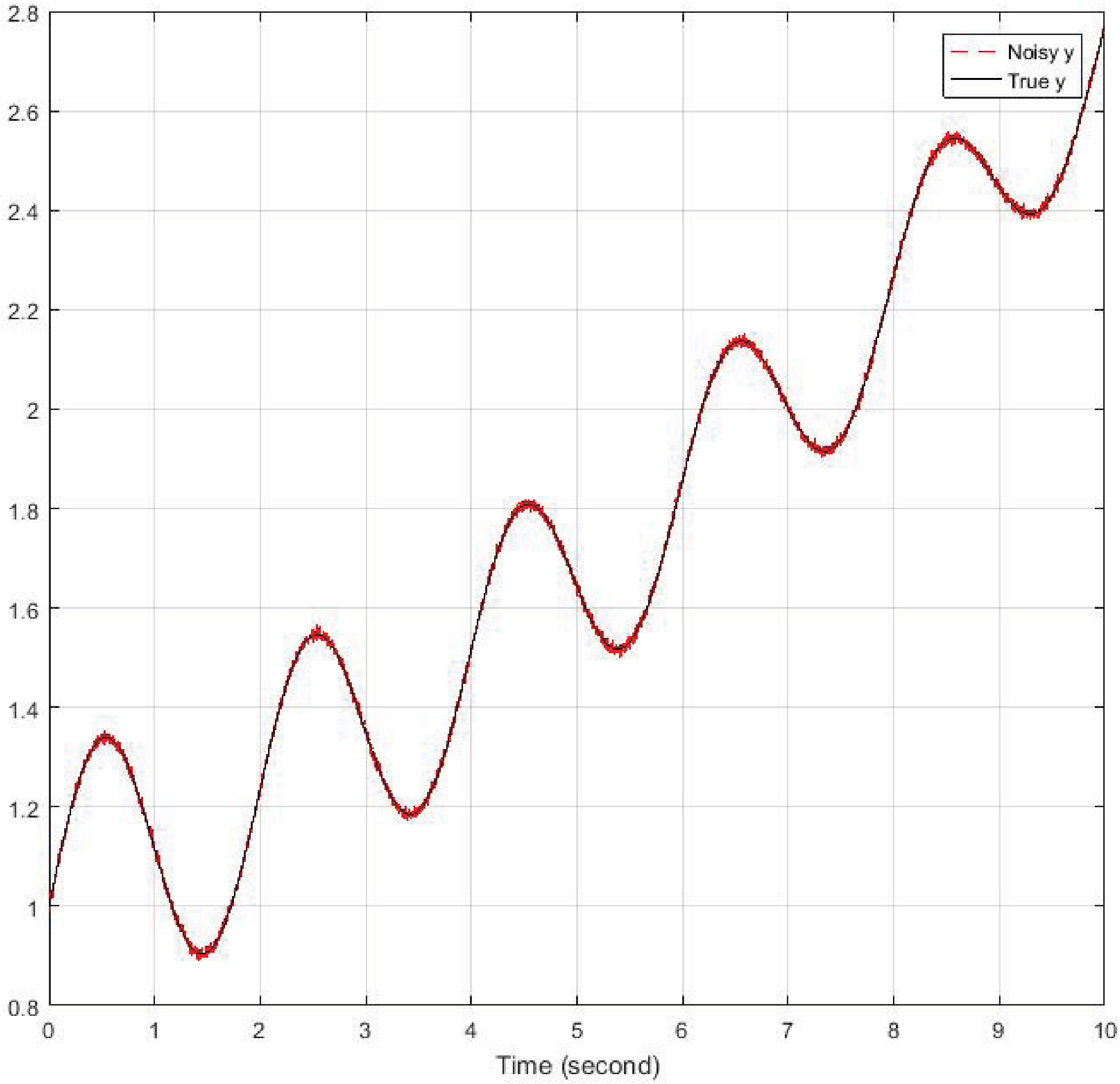

Figure 4.1: Case I: Gaussian 50SNR, noisy $y(t)$ vs the true output $y(t)$

We can see that the noise corruption is relatively small but even such small noise perturbation produces huge inaccuracy of the estimation of state and its time derivatives as the integration error accumulates.

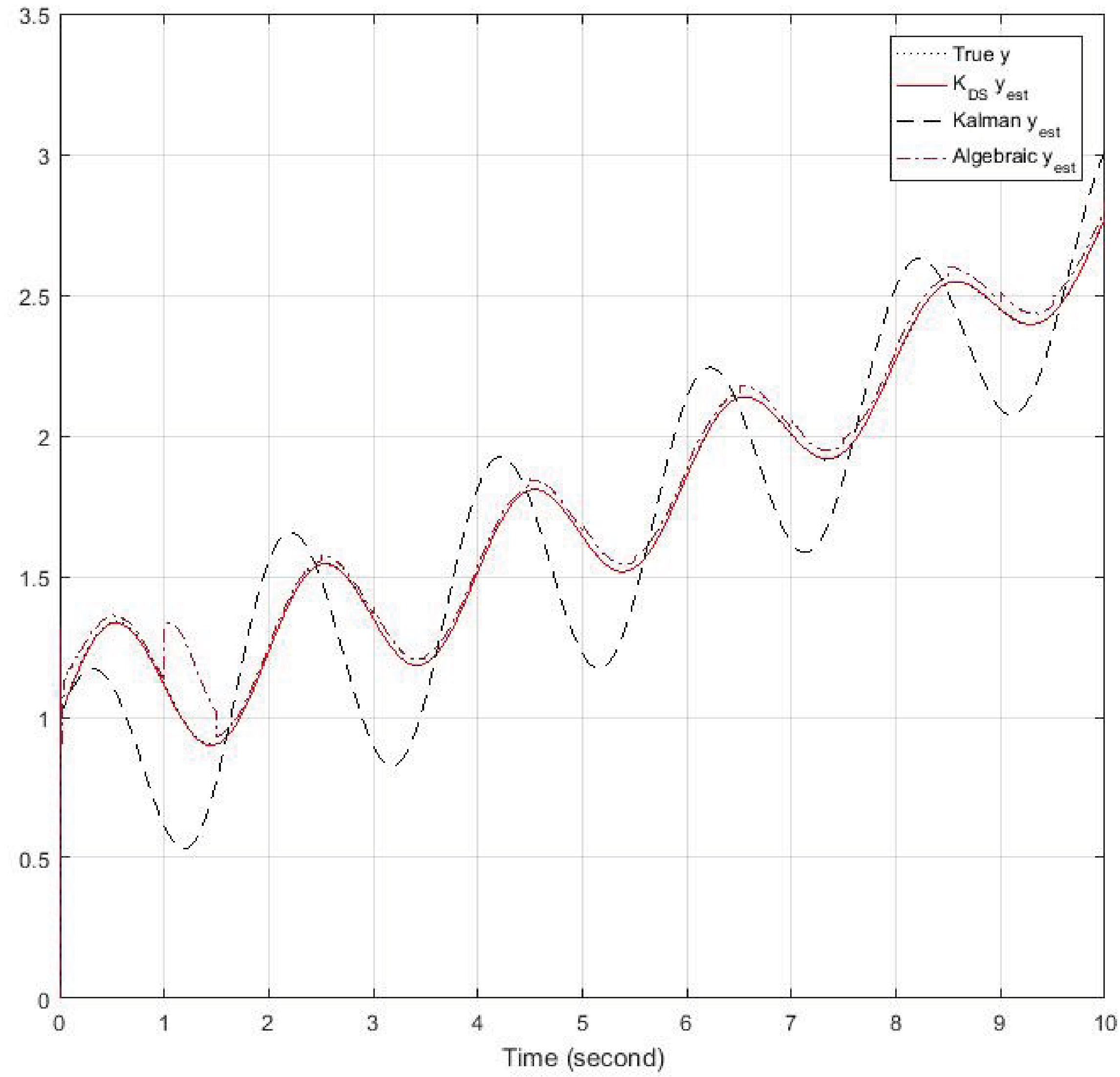

Figure 4.2: Case I: Estimated output $y(t)$ from K_{DS} vs the classical algebraic approach and Kalman filter

In figure 4.2 we observe that the Kalman filter is divergent from the true observation and tends to converge to the true trajectory asymptotically, i.e. converges in infinite time. This is because, the Kalman filter is a recursive observer and depends heavily on the initial conditions. We also see that the estimation by the classical algebraic approach of Fliess et al. is not smooth and has spikes, especially a very prominent spike at $t = 1$ second, the time of resetting the estimation, though it does not require initial condition for estimation. We also notice that the estimates start diverging from the true trajectory before resetting. The K_{DS} based estimator delivers a truly superior estimate and tracks the true trajectory nearly exactly.

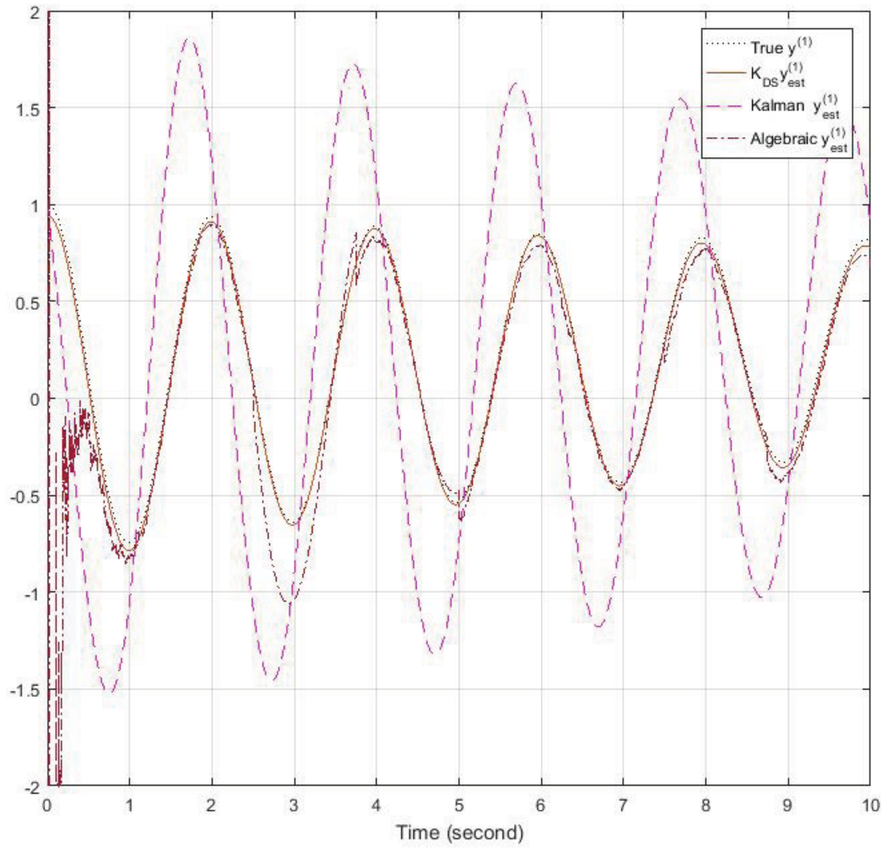

Figure 4.3: Case I: Estimated output $y(t)$ from K_{DS} vs the classical algebraic approach and Kalman filter

In the case of estimation of $y^{(1)}_{est}$, we see that the algebraic estimator has considerable spikes due to singularity at $t = 0$, and diverges at $t = 3s$. This is because of accumulation of error due to truncation of the Taylor series. Note that divergence cannot be avoided completely inspite of reinitialising the estimator (in this case, $0.5s$). Due to this, we can deduce that we need to implement a noise filtering algorithm in addition to this method to improve the estimates. The estimates from the Kalman filter is much more diverged in this case due to the absence of initial conditions and is not converging in our time horizon.

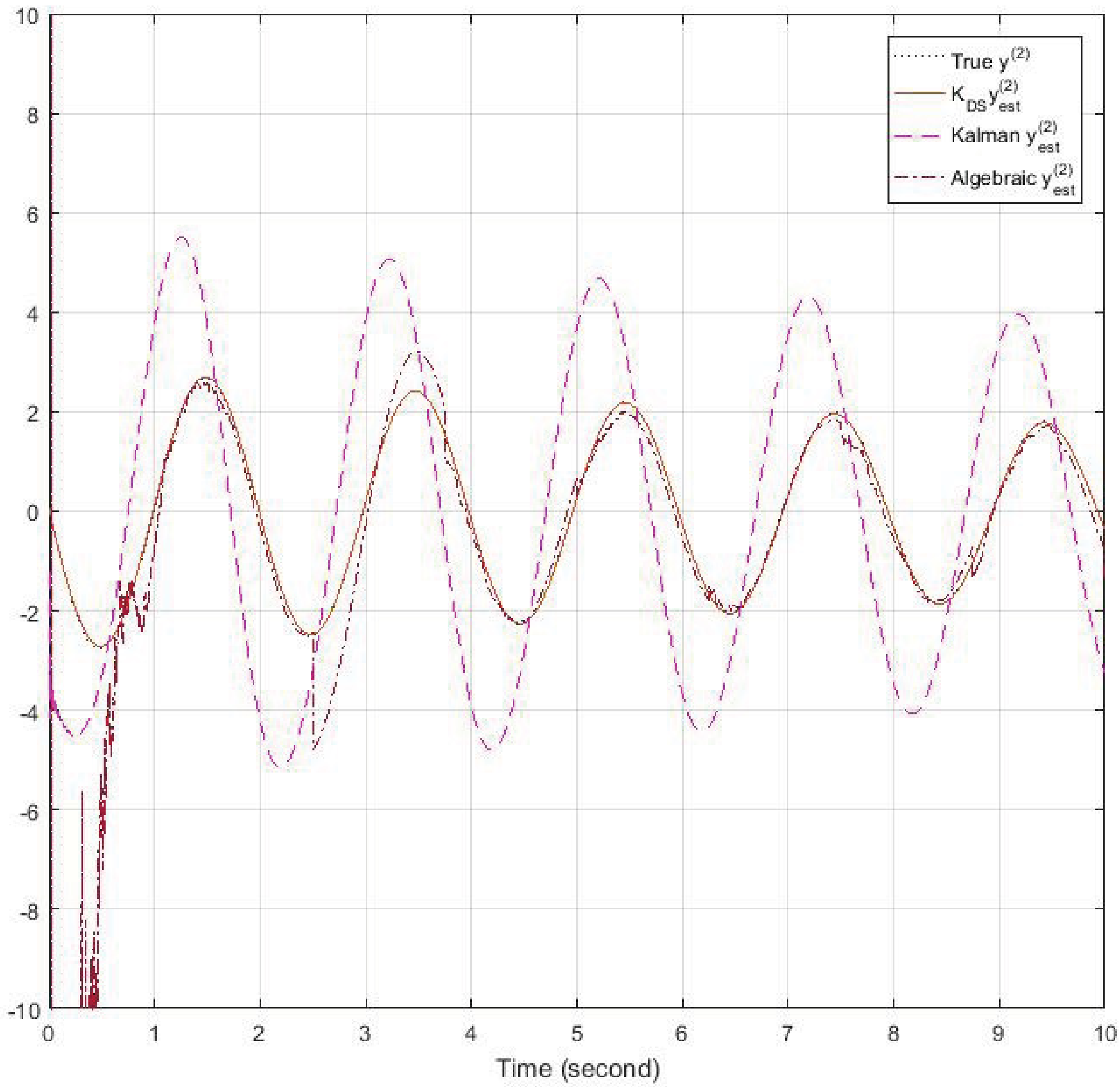

Figure 4.4: Case I: Estimated output derivative $y^{(2)}(t)$ from K_{DS} vs the classical algebraic approach and Kalman filter

When we compare those to the Kalman estimates especially for the case of $y^{(2)}_{est}(t)$, we see that there is an initial spike and the early estimates are not smooth. We can also see that the estimates from the algebraic approach have spikes at $t = 0$, which keeps degrading the estimates as the order of the derivatives increasing. The features observed in the previous two figures, 4.2 and 4.3 in the cases of the algebraic approach and the Kalman filter are much more prominent in the case of estimating $y^{(2)}(t)$, while the performance of the double sided kernel K_{DS} is commendable. This is because, our kernel operates from both the front and the back sides, thereby giving us very good estimates. The estimation is better in case of the double sided kernel K_{DS} due to the fact that we use the knowledge of system invariance.

Let us look at the case where the signal is distorted by a larger Gaussian noise.

4.2.2 Case II: High Noise - 30SNR

The simulated system output was distorted by adding a Gaussian noise of 30 SNR and the noisy signal is as shown below in figure 4.5.

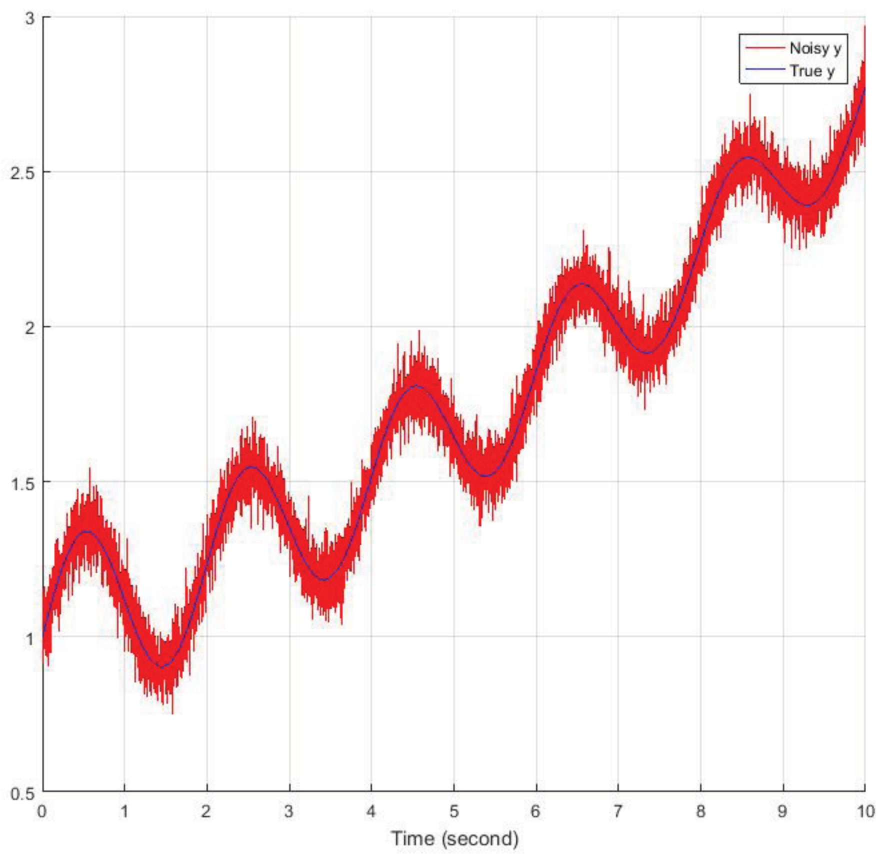

Figure 4.5: Case II: Gaussian 30SNR, noisy $y(t)$ vs the true output $y(t)$

The estimation is greatly affected due to this heavy noise in all cases, as can be seen in the figures below.

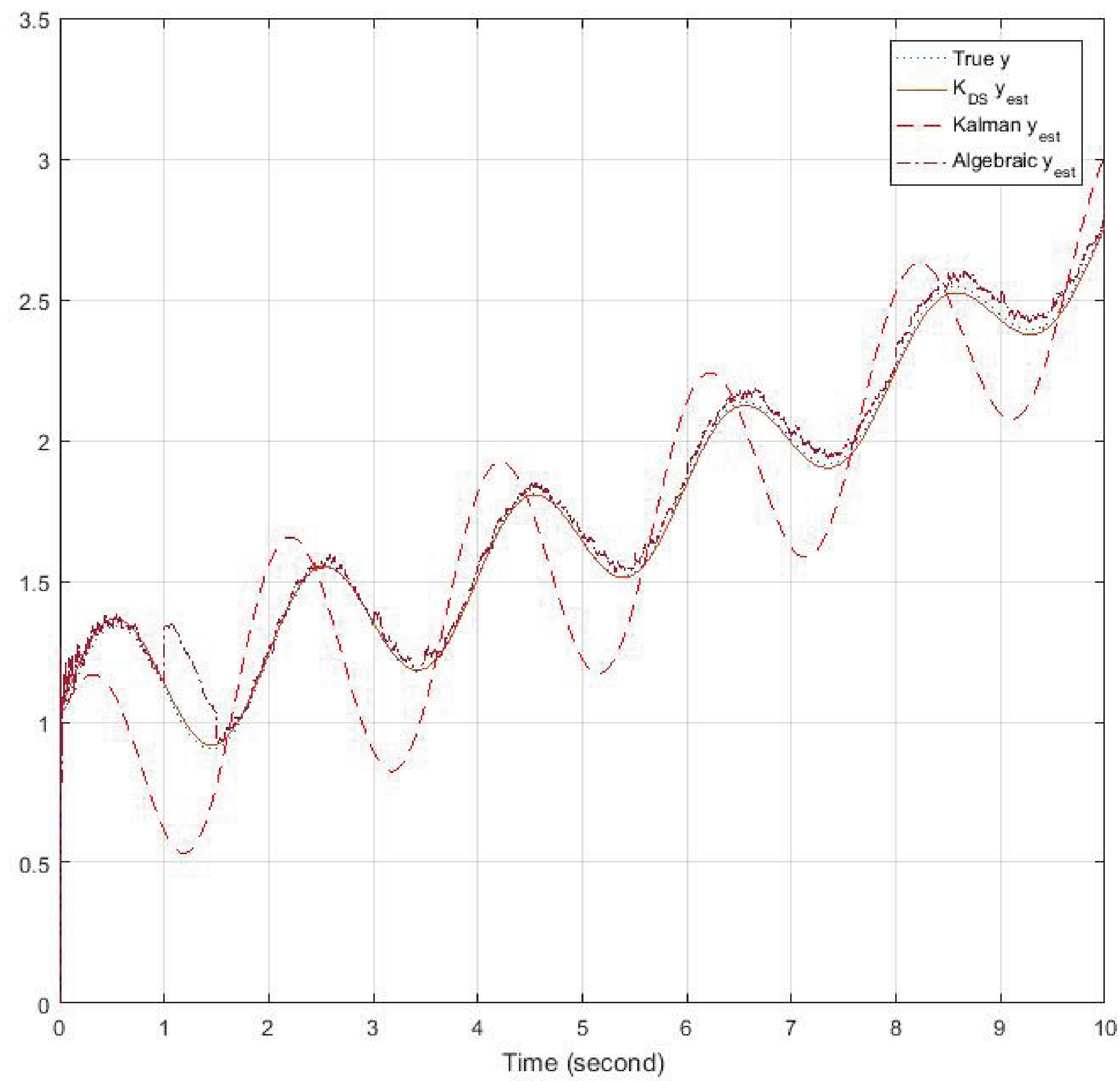

Figure 4.6: Case II: Estimated output $y(t)$ from K_{DS} vs the classical algebraic approach and Kalman filter

In the figure 4.6, we see that the Kalman filter diverges as in the previous case. The classical algebraic approach is also not very robust against strong Gaussian noise. This is very surprising given that the iterated integrals have a low pass filtering property. And despite resetting often (at every $t = 0.5s$ in this case), the $y_{est}(t)$ is very rough and diverges from the true trajectory. On the other hand, the double sided kernel performs better than the other estimators, but still diverges from the true trajectory. This can be attributed to the accumulation of error due to integration of noise.

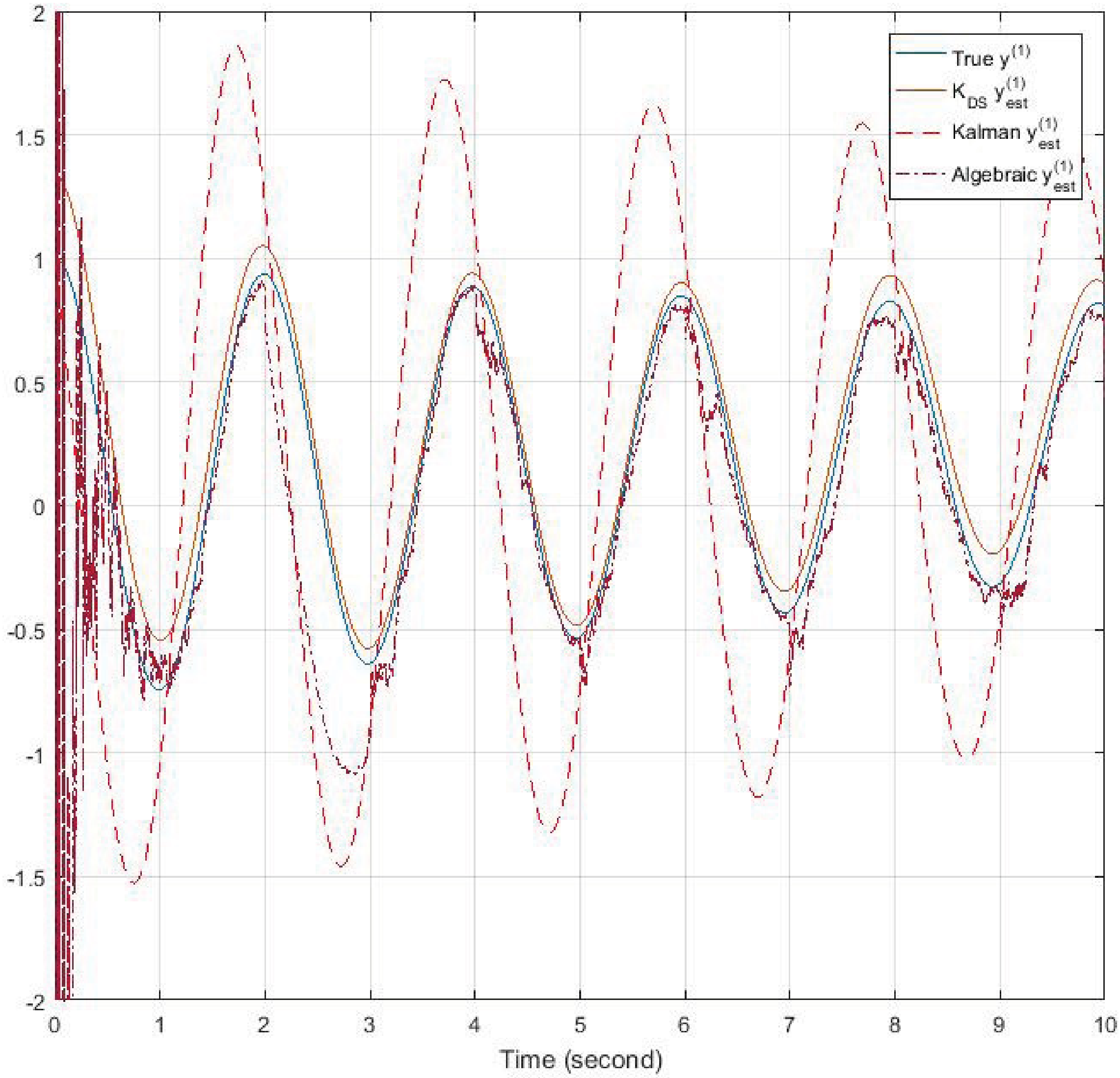

Figure 4.7: Case II: Estimated output $y^{(1)}(t)$ from K_{DS} vs the classical algebraic approach and Kalman filter

This figure 4.7 indicates the greatest flaw of the algebraic approach. The terrible performance in the initial period is due to the singularity at $t = 0$. The estimate also starts diverging early on. We can deduce from the spikes that the controller which employs this method of estimation would require a very good compensator to compensate for the error in estimation. We see that the double sided kernel K_{DS} based estimation diverges from the true trajectory at the peaks. The reason is same as before, i.e., attributed to integral action on the noise.

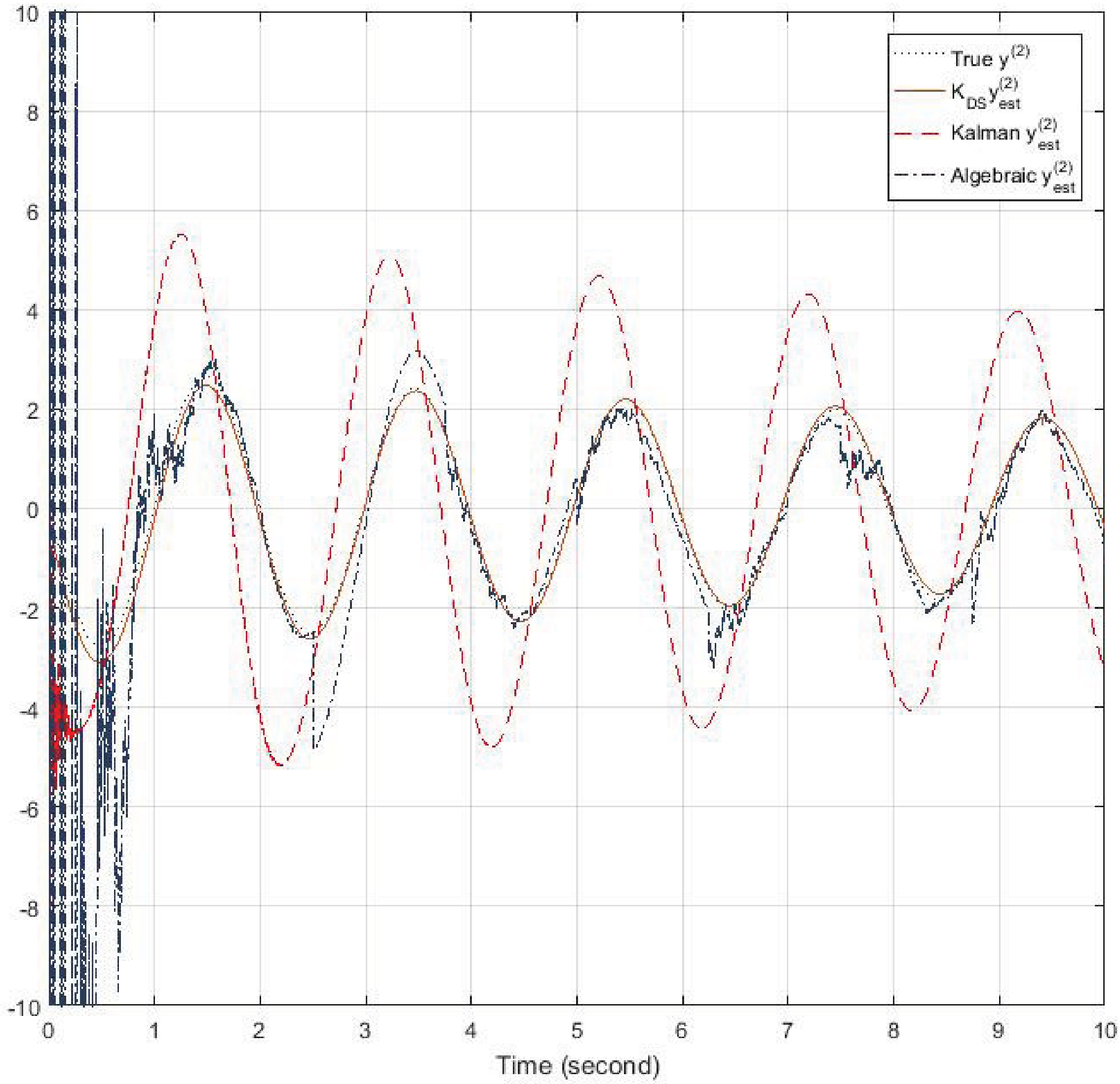

Figure 4.8: Case II: Estimated output derivative $y^{(2)}(t)$ from K_{DS} vs the classical algebraic approach and Kalman filter

Similar observations as earlier can be made in the case of all the estimators in fig 4.8. The flaws are more pronounced in this case. The algebraic approach estimates have spikes and are not smooth and even the Kalman estimate is not smooth. And the divergence is greater in the case of the K_{DS}.

This case is shown on purpose, because it shows that despite being very good and robust, the double sided kernel K_{DS} cannot eliminate noise on its own. This calls for the development of stronger noise rejection properties in future generation of K_{DS} based filters. Such filters are in fact investigated by our group's work [24].

4.3 Parameter Estimation

The same system was considered for parameter estimation. The signal was corrupted with a noise of 30 SNR. The expressions used were as derived and developed in the sections § 2.3 and § 3.4 and the estimation of the parameters as well as the states was carried ou simultaneously in the case of the double sided kernel K_{DS}. The signal shown in the figure 4.5 was used.

The parameters found from the algebraic approach and K_{DS} is as shown as follows:

	a_0	a_1	a_2
True values	-1	10	0
Estimated values (from K_{DS})	-0.9861	9.8171	0.0153
Estimated values (from classical algebraic approach)	-0.8171	9.153	0.121

Table 4.1: Estimated parameter values using the double sided kernel K_{DS} and the classical algebraic appraoch for a signal with noise of 30SNR

We see that the accuracy of the parameter estimation is better in the case of the double sided kernel K_{DS} over the classical algebraic approach. We also observe the effect of the parameters on the state estimation and the estimation of time derivatives. The plots of the estimated states and time derivatives are given below.

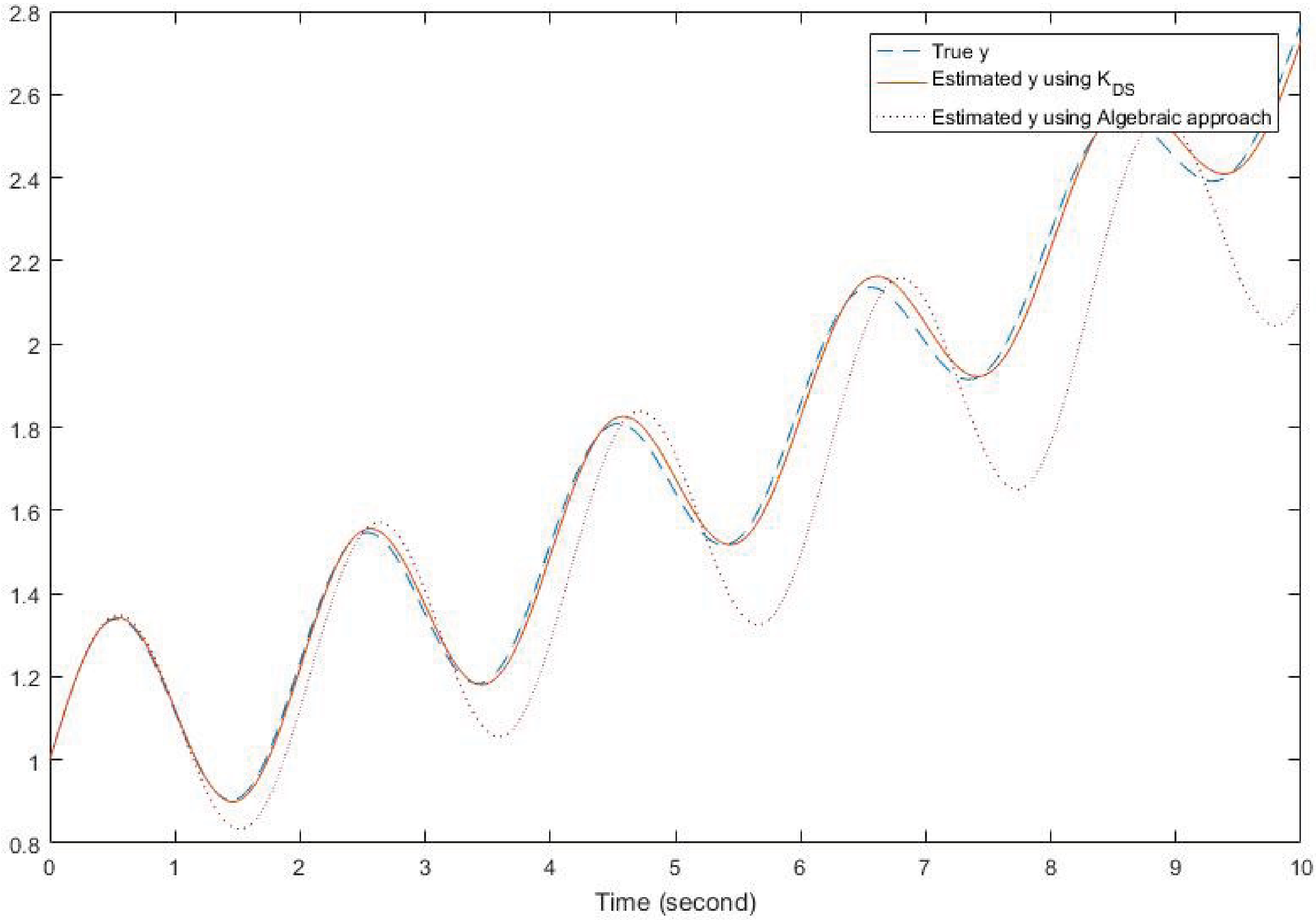

Figure 4.9: Estimated $y(t)$ using parameters from K_{DS} and the classical algebraic approach vs true $y(t)$

The figure 4.9 shows clearly the divergence in both the estimations, but it is more properly seen in the case of the classical algebraic estimation based parameter estimation. While the double sided kernel K_{DS} based parameter estimation is seen to be diverging, it is much less pronounced than that of the classical algebraic method.

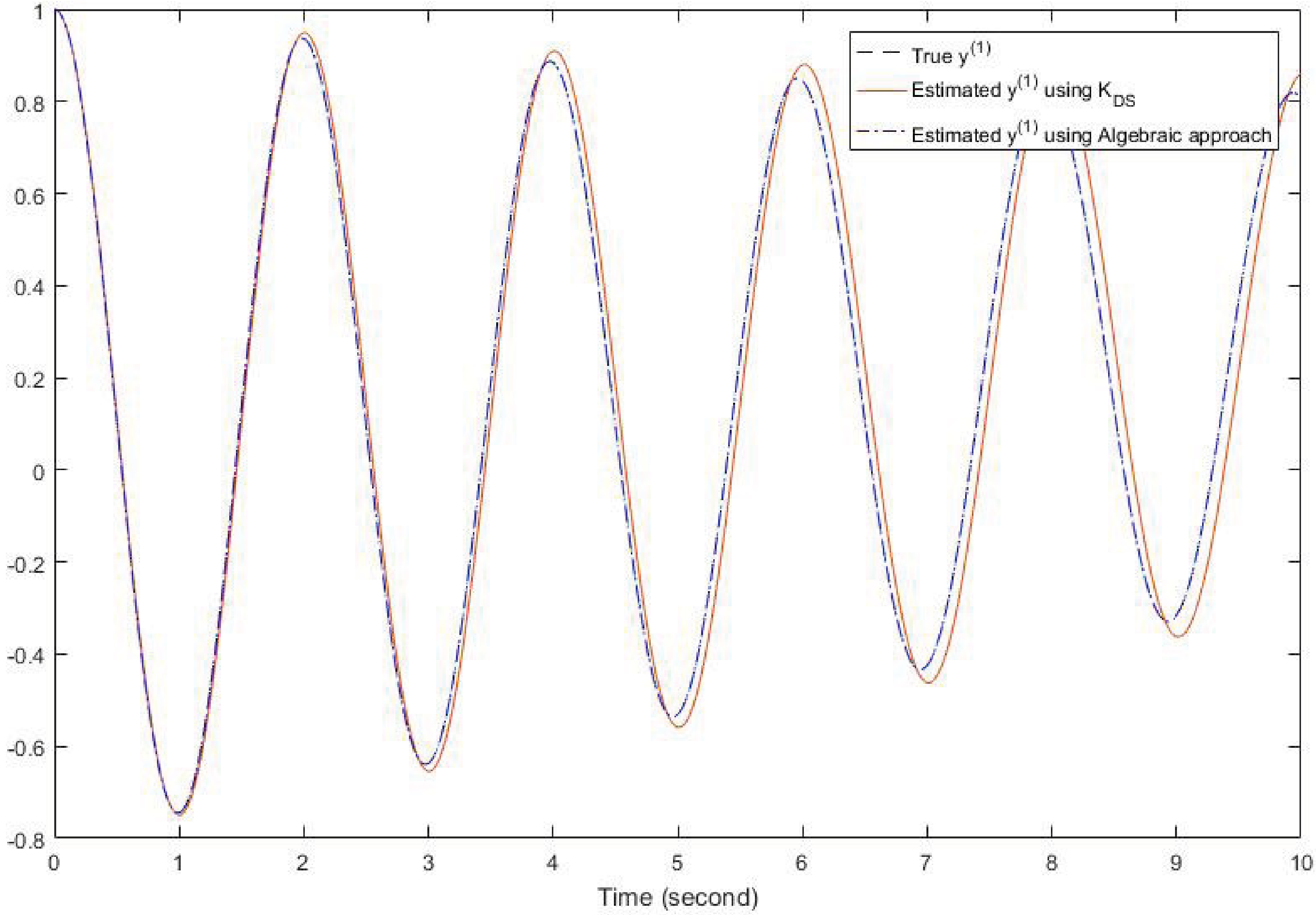

Figure 4.10: Estimated $y^{(1)}(t)$ using parameters from K_{DS} and the classical algebraic approach vs true $y^{(1)}(t)$

In the case of figure 4.10, due to the usage of the simultaneous parameter and state estimation, the estimated $y^{(1)}(t)$ from the K_{DS} is follows the true trajectory, while the estimated $y^{(1)}(t)$ from the classical algebraic approach is seen to be diverging. The reason is the error in estimating the parameters, which cannot be handled by low pass filtering alone.

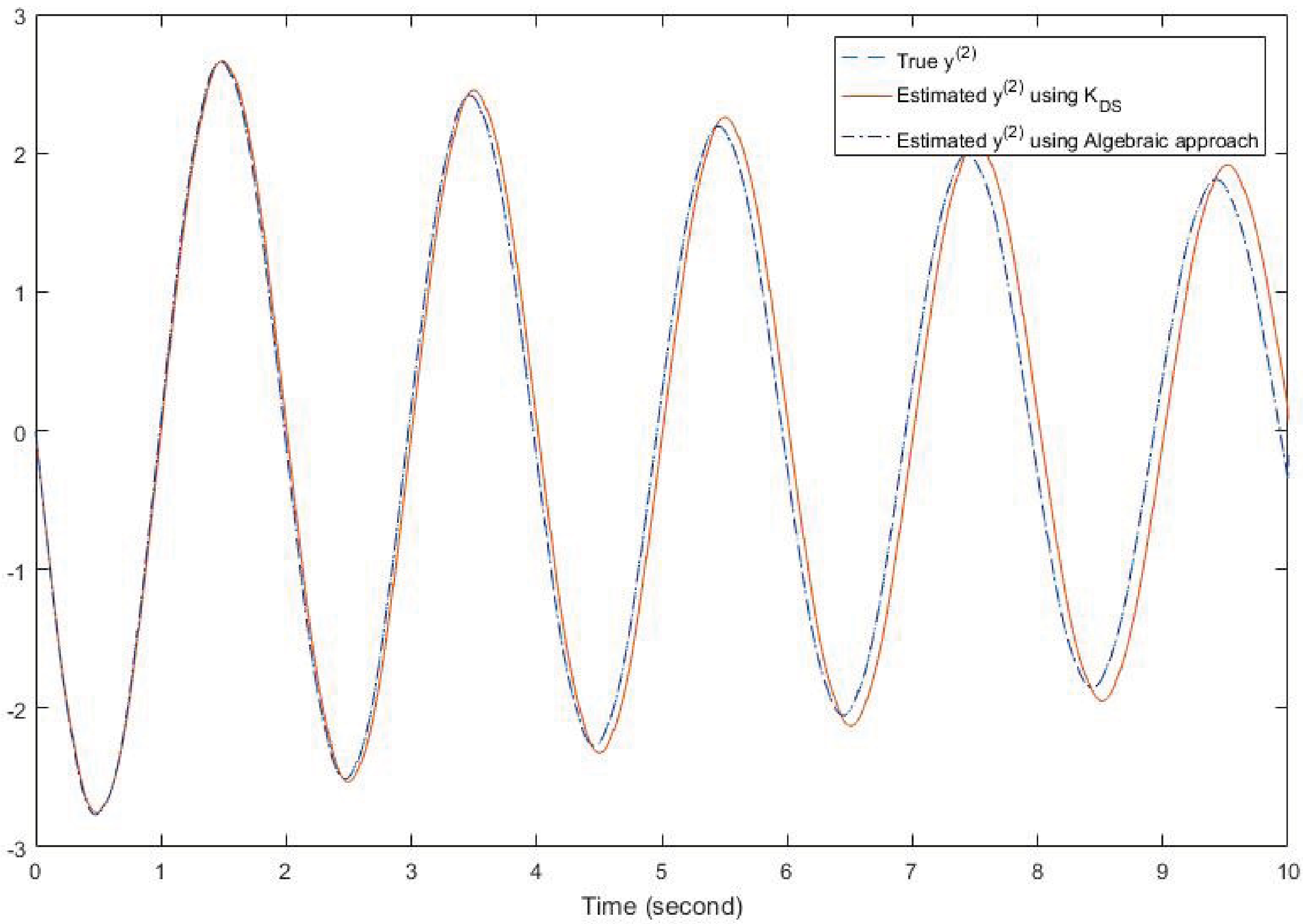

Figure 4.11: Estimated $y^{(2)}(t)$ using parameters from K_{DS} and the classical algebraic approach vs true $y^{(2)}(t)$

Lastly in the figure 4.10 the performance is similar to the previous two estimations. The main takeaway from this case is that despite the presence of integral low pass filters, these filters are not robust enough to handle such heavy, and in some cases quantised, noise. To ensure that the estimation is carried out smoothly, we need to come up with a filtering scheme in alliance with the estimation algorithm to ensure better estimations and simpler, more efficient controllers.

Chapter 5

Conclusion and Future Work

State and parameter estimation are very important in the context of modern technology. Information gathered from the states and its higher order time derivatives can be used to create a better controller. While classical recursive filters are still dominant tools in conventional system estimation, there is an increasing demand for the development of dead-beat methods. This is motivated by the growth of nonlinear, hybrid, and fast switching system control technologies. This thesis presents some new ideas leading to high fidelity dead-beat estimator algorithms. Specifically, several vulnerabilities of standard non-asymptotic estimators are removed by the development of the double sided kernel K_{DS} presented in [1] and [2]. Chapter 1 gave a brief overview of the various approaches to estimate the state and its finite order time derivatives as well as estimation of parameters of a system and the need for these estimation approaches. In Chapter 2, the development of the algebraic method of Fliess et al., was studied in more detail and despite having very good features, it has some serious drawbacks like singularity at $t = 0$ and accumulation of error due to truncation of the Taylor series requiring us to reinitialise the estimator at regular intervals. In order to overcome these drawbacks, in Chapter 3, a novel method for estimation was proposed which ensured no singularity and used the knowledge of the system to better estimate the time derivatives. It also provided us with an option of estimating parameters of the system as well as the states and its finite order time derivatives at the same time. Chapter 4 provided a comprehensive performance comparision between the double sided kernel K_{DS}, the classical algebraic approach of Fliess et al., and a recursive filter such as Kalman filter. As seen, the accuracy is unquestionably superior in the case of the double sided kernel K_{DS} as compared to both the Kalman filter and the classical algebraic estimator. The spikes while employing the classical algebraic estimator are very noticeable and the estimates are not very smooth. These spikes increase as the noise in the signal increases. However, the double sided kernel K_{DS} is not without with its own flaws especially when the noise is higher.

It should be stressed that the present algebraic estimation approaches (both classical and the K_{DS} based) do not depend on the characteristics of the noise affecting the output measurement. The output estimation $y(t)$ is a convolution, similar to the expression in [25] and can be written as follows:

$$y(t) = \int K_{DS}(t, \tau)\, y(\tau) + \nu \mathrm{d}\tau \tag{5.1}$$

Here, the output estimation $y(t)$ is the noisy input $y(\tau) + \nu$ where ν is the noise convolved with $K_{DS}(t, \tau)$ which acts as an impulse response function. The multiple integrations serve as natural filters and hence does not depend on the characteristics of the noise.

Complete de-noisification procedures can be derived [24] but are beyond the scope of the work presented here. The exact characteristcs of attenuated noise is a subject of further detailed study. The double sided kernel K_{DS} lends itself to generalizations for multivariate systems, linear time varying systems as well as nonlinear systems. This will be demonstrated and discussed elsewhere.

Bibliography

[1] D. Ghoshal, K. Gopalakrishnan, and H. Michalska, "Using invariance to extract signal from noise - submitted," in *American Control Conference, 2017. ACC 2017*, 2017 forthcoming.

[2] ——, "Algebraic parameter estimation using kernel representation of linear systems - submitted," in *IFAC World Congress, 2017. 20th World Congress*, 2017 forthcoming.

[3] H. Sira-Ramírez, C. G. Rodríguez, J. C. Romero, and A. L. Juárez, *Algebraic identification and estimation methods in feedback control systems*. John Wiley & Sons, 2014.

[4] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[5] D. Luenberger, "Observers for multivariable systems," *IEEE Transactions on Automatic Control*, vol. 11, no. 2, pp. 190–197, 1966.

[6] N. Wiener, *Extrapolation, interpolation, and smoothing of stationary time series*. MIT press Cambridge, 1949, vol. 2.

[7] R. Kalman, "On the general theory of control systems," *IRE Transactions on Automatic Control*, vol. 4, no. 3, pp. 110–110, 1959.

[8] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Journal of basic engineering*, vol. 83, no. 1, pp. 95–108, 1961.

[9] R. E. Kalman, "Mathematical description of linear dynamical systems," *Journal of the Society for Industrial and Applied Mathematics, Series A: Control*, vol. 1, no. 2, pp. 152–192, 1963.

[10] D. G. Luenberger, "Observing the state of a linear system," *IEEE transactions on military electronics*, vol. 8, no. 2, pp. 74–80, 1964.

[11] G. David, "An introduction to observers," *IEEE Transactions on automatic control*, vol. 16, no. 6, pp. 596–602, 1971.

[12] S. Diop and M. Fliess, "Nonlinear observability, identifiability, and persistent trajectories," in *Decision and Control, 1991., Proceedings of the 30th IEEE Conference on*. IEEE, 1991, pp. 714–719.

[13] M. Fliess, M. Mboup, H. Mounier, and H. Sira-Ramírez, "Questioning some paradigms of signal processing via concrete examples," in *Algebraic Methods in Flatness, Signal Processing and State Estimation*. Editorial Lagares, 2003, pp. pp–1.

[14] M. Fliess and H. Sira-Ramírez, "Reconstructeurs d'état," *Comptes Rendus Mathematique*, vol. 338, no. 1, pp. 91–96, 2004.

[15] R. Morales, E. Segura, J. Somolinos, L. Núñez, and H. Sira-Ramírez, "Online signal filtering based on the algebraic method and its experimental validation," *Mechanical Systems and Signal Processing*, vol. 66, pp. 374–387, 2016.

[16] G. Favier, *Filtrage, modélisation et identification de systèmes linéaires stochastiques à temps discret*. Editions du CNRS, 1982.

[17] L. Zadeh, "On the identification problem," *IRE Transactions on Circuit Theory*, vol. 3, no. 4, pp. 277–281, 1956.

[18] K.-J. Åström and T. Bohlin, "Numerical identification of linear dynamic systems from normal operating records," in *Theory of self-adaptive control systems*. Springer, 1966, pp. 96–111.

[19] K. Äström and B. Wittenmark, "Adaptive control," *NY: Addison-Wesley*, 1989.

[20] M. Gevers, "A personal view of the development of system identification: A 30-year journey through an exciting field," *IEEE Control Systems*, vol. 26, no. 6, pp. 93–105, 2006.

[21] M. Fliess and H. Sira-Ramírez, "An algebraic framework for linear identification," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 9, pp. 151–168, 2003.

[22] ——, "Closed-loop parametric identification for continuous-time linear systems via new algebraic techniques," in *Identification of Continuous-time Models from sampled Data*. Springer, 2008, pp. 363–391.

[23] J. Reger, H. S. Ramírez, and M. Fliess, "On non-asymptotic observation of nonlinear systems," in *Proceedings of the 44th IEEE Conference on Decision and Control*. IEEE, 2005, pp. 4219–4224.

[24] D. Ghoshal, K. Gopalakrishnan, and H. Michalska, "Parameter and state estimation in linear systems under arbitrary noise - submitted," in *IFAC World Congress, 2017. 20th World Congress*, 2017 forthcoming.

[25] M. Mboup, C. Join, and M. Fliess, "Numerical differentiation with annihilators in noisy environment," *Numerical algorithms*, vol. 50, no. 4, pp. 439–467, 2009.