B-Splines in Joint Parameter, State, and Input Estimation in Linear Time-Varying Systems

Deepak Sridhar

Department of Electrical & Computer Engineering

McGill University

Montreal, Canada

March 2018

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

To my parents, my brother and my teachers.

Abstract

Non-asymptotic observers have received a great deal of attention in recent times due to the advancement in hybrid control system theory. This is because fast and switching control methods are essential for hybrid systems. Conventional observers such as the Luenberger observers and Kalman filters are asymptotic in nature and fail to achieve this. Algebraic state and parameter estimation methods offer the alternative to conventional methods since they are non-asymptotic and have other superior features. Algebraic state and parameter estimation methods have been studied extensively in LTI systems. For LTV systems, there is very limited literature related to state and parameter estimation especially using algebraic methods. This thesis provides a new method of joint state, parameter and input estimation for linear time-varying systems.

The objective of this thesis is to propose and describe a new method to construct non-asymptotic state, parameter and input estimators for LTV systems that employs a kernel functional representation of linear time-varying systems in conjunction with B-spline functional approximation techniques. The double-sided kernel for LTV systems is a generalization of its LTI counterpart developed by Dr.Michalska and her team. Total observability of the estimated system must be assumed. Practical identifiability conditions for parametric estimation are also stated in this thesis.

In the absence of output measurement noise the observer provides almost exact reconstruction of the system state and delivers high fidelity functional estimates of the time varying system parameters. It also shares the usual superior features of algebraic observers such as independence of the initial conditions of the system and good noise attenuation properties. Other advantages of the kernel and B-spline based identification of linear time-varying systems are elucidated. In the presence of output measurement noise the performance of the estimator deteriorates with decreasing signal-to-noise ratios. Results are presented for both noiseless and noisy output measurements. An example is also presented to show that the joint parameter and state estimation is superior using P-Splines than using B-Splines.

Résumé

Les observateurs non-asymptotiques ont reçu beaucoup d'attention ces derniers temps en raison de avancement dans la théorie du système de commande hybride. C'est parce que la commande rapide est essentielle pour les systèmes hybrides. Des observateurs conventionnels tels que l'observateur de Luenberger et les filtres de Kalman sont asymptotiques et ne parviennent pas à atteindre cet objectif. D'autre part, les méthodes algébriques offrent des alternatives d'estimation non-asymptotiques. Cela a été étudié de manière approfondie dans les systèmes LTI. Dans le cas des systèmes LTV, il y a une littérature très limitée pour l'estimation d'état et des paramètres. Cette thèse fournit une nouvelle méthode d'estimation simultanée de l'état et des paramètres pour des systèmes LTV.

L'objectif de cette thèse est de proposer et de décrire la nouvelle méthode dans laquelle un noyau qui sert de représentation fonctionnelle des systèmes LTV est employée en conjonction avec des techniques d'approximation fonctionnelle par les B-Splines pour construire d'estimateurs non-asymptotique. Le noyau double pour les systèmes LTI a été développé par le professeur Michalska et ses anciens étudiants. Observabilité totale du système estimé doit être supposé. Les conditions d'identification pratiques pour l'estimation paramétrique sont également définies.

En l'absence de bruit de mesure, l'observateur fournit reconstruction de l'état du système presque exacte et fournit des estimations de haute fidélité de paramètres du système. Il partage également les caractéristiques supérieures habituelles de observateurs algébriques tels que l'indépendance des conditions initiales du système et la bonne propriétés d'atténuation du bruit. D'autres avantages du noyau pour les systèmes LTV et de l'identification basée sur les B-Splines sont élucidés. En présence de bruit de mesure la performance de l'estimateur se détériore. Les résultats sont présentés pour les mesures de sortie silencieuses et bruyantes. Un exemple est également présenté pour montrer que l'estimation conjointe des paramètres et des états est supérieure en utilisant P-Splines que d'utiliser B-Splines.

Acknowledgments

This work was not possible without the help of several people whom I am obliged to thank here.

Firstly, I would like to thank my supervisor Dr.Hannah Michalska, for providing me the opportunity to work under her supervision and be part of the ongoing innovative research work. She has been of immense help in guiding me through tough times by helping me overcome the roadblocks in my thesis work through her witty suggestions and providing moral and financial support during my research work. I would also like to thank her for being extremely friendly and accommodating to my needs. It is because of this stress-free environment, it was possible to complete the thesis work in a short time.

I would like to thank the department of Electrical and Computer Engineering, McGill University for their financial support in the form of Graduate Excellence Fellowship and Mitacs Inc, for providing me with Globalink Graduate Fellowship which was of immense help to pay my tuition fee. I would also like to thank Dr. Brian Chen for providing me with a part time job as an electrical engineer in his lab which helped me meet my living expenses and made my stay in Montreal an affordable one.

I would like to thank my colleague, Debarshi Patanjali Ghoshal, who was of great help during my research. I am grateful to him for his support in providing an in depth understanding on the advanced concepts and help in troubleshooting problems. I would also like to acknowledge my friends Abhishek Pandey, Nachiket Kale, Praveen Ravichandran, Bashar Navaz, and Shaunak Sinha who are a part of the lab for their continued moral support and company during my research.

Last but not least, I would like to convey my deepest thanks to my mother, my father and my brother for their constant support throughout my life, both financially and emotionally and for their motivation to achieve the best. I am grateful to them for constantly supporting and encouraging my endeavours.

Preface

This is to declare that this thesis work was carried out entirely by myself, but was part of a larger collaborative effort - that of the research team led by my supervisor Dr.Hannah Michalska. The following group of students are members of this team: Debarshi Patanjali Ghoshal, PhD scholar, M.Eng Students Abhishek Pandey, Nachiket Kale, Praveen Ravichandran, Bashar Navaz, and Shaunak Sinha. The group would hold regular research meetings to discuss and pose ideas and exchange work experiences. Although each student in the group was assigned a different research topic, present was a unifying theme of system modeling and estimation by concepts of machine learning and functional data analysis.

Contents

List of Figures

List of Acronyms

LTI	Linear Time Invariant
LTV	Linear Time Varying
B-Spline	Basis Spline
P-Spline	Penalized B-Spline

Chapter 1

Introduction

Control and automation play a crucial role in today's technology by ensuring safety, stability and reliability of engineering systems used in industry. Control systems are ubiquitous. In our homes, we find them in everything from toasters to heating systems to DVD players. Control systems also have widespread applications in science and industry, from steering ships and planes to guiding missiles. Control systems also exist naturally; our bodies contain numerous control systems. Even economic and psychological system representations have been proposed based on control system theory. Control systems are used where power gain, remote control, or conversion of the form of input is required.

The field of Control systems has rich history associated with it. Automatic control systems were first developed over two thousand years ago [2]. The first feedback control device on record is thought to be the ancient water clock of Ktesibios in Alexandria Egypt around the third century B.C. It kept time by regulating the water level in a vessel and, therefore, the water flow from that vessel. From there, the field of control systems has come a long way.

Mathematical techniques made it possible to control, more accurately, and more significantly, complex dynamical systems. These techniques include developments in optimal control in the 1950's and 1960's, followed by progress in stochastic, robust, adaptive and optimal control methods in the 1970's and 1980's. Applications of control methodology have helped make possible space travel and communication satellites, safer and more efficient aircraft, cleaner auto engines, cleaner and more efficient chemical processes, to mention but a few. More recently, futuristic dreams have become a reality such as reusable rockets, self-driving cars, unmanned aerial vehicles, autonomous underwater vehicles, smart homes, and smart grids with the development of novel control methods.

Definition 1.0.1 *[1] A control system consists of subsystems and processes (or plants) assembled for the purpose of obtaining desired output with desired performance, given a specified input.*

Figure 1.1 Simplified description of a control system [1]

Figure 1.1 shows a control system in its simplest form, where the input represents a desired output. There are two major configurations of control systems: open loop and closed loop. We can consider these configurations to be the internal architecture of the total system shown in Figures 1.2 and 1.3.

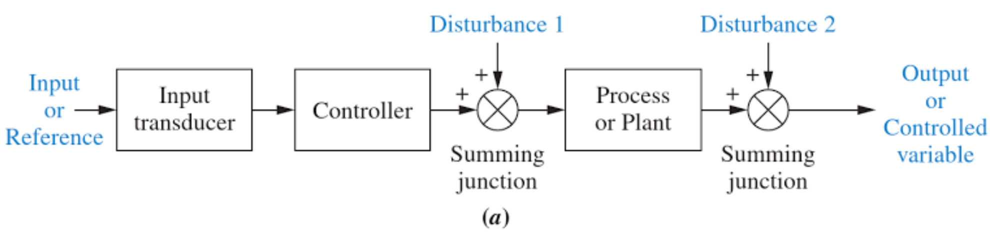

Figure 1.2 Block diagram of an open-loop control system [1]

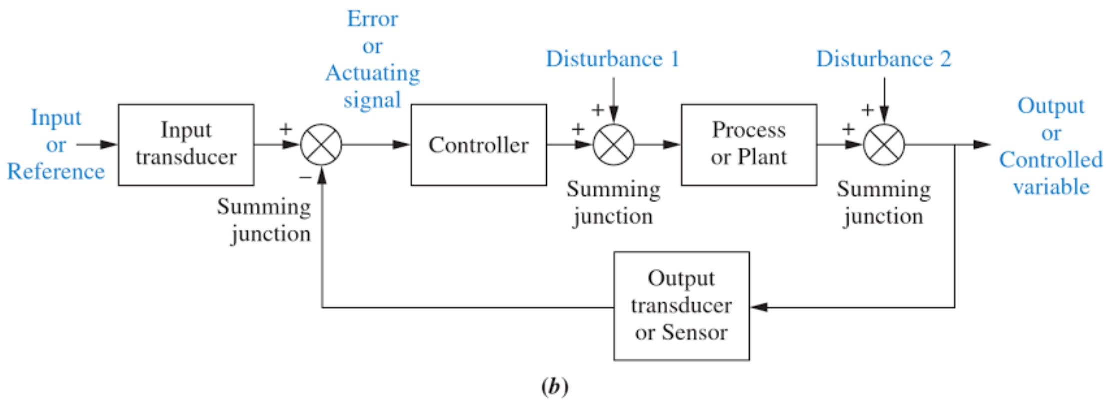

Figure 1.3 Block diagram of a closed-loop control systems [1]

1.0.1 Open Loop Systems

A generic open-loop system is shown in Figure 1.2. It starts with a subsystem called an input transducer, which converts the form of the input to that used by the controller. The controller drives a process or a plant. The input is sometimes called the *reference*, while the output can be called the *controlled variable*. Other signals, such as disturbances, are shown added to the controller and process outputs via summing junctions, which yield the algebraic sum of their input signals using associated signs. The distinguishing characteristic of an open-loop system is that it cannot compensate for any disturbances that add to the controller's driving signal (Disturbance 1 in Figure 1.2). The output of an open loop system is corrupted not only by signals that add to the controller's commands but also by disturbances at the output (Disturbance 2 in Figure 1.2). The system cannot correct for these disturbances, either. Some examples of open loop systems are toasters, mass-spring-damper systems, and open loop temperature controllers that work in coordination with timer devices.

1.0.2 Closed Loop Systems

The disadvantages of open loop systems, namely sensitivity to disturbances and inability to correct disturbances or model system error, may be overcome in closed loop systems. The generic architecture of a closed loop system is shown in Figure 1.3. The input transducer converts the form of the input to the form used by the controller. An output transducer, or sensor, measures the output response and converts it into the form used by the controller. The first summing junction algebraically subtracts the input signal from the output signal, which arrives via the feedback path i.e. the return path from the output to the summing junction; see Figure 1.3. The result is generally called the *actuating signal*. However, in systems where both the input and output transducers employ a unity gain, the actuating signal's value is equal to the actual difference between the input and the output. Under this condition, the actuating signal is called the *control error*. The closed loop system compensates for disturbances by measuring the output response, feeding that measurement back through a feedback path, and computing that response to the input at the summing junction. If there is any difference between the two responses, the system drives the plant, via the actuating signal, to make a correction. If there is no difference, the system does not drive the plant, since the plant's response is already the desired response. Hence, closed loop systems are less sensitive to noise, disturbances, and changes in the environment. Transient response and steady state error can be controlled more conveniently and with greater flexibility in closed loop systems.

In order to implement a good feedback control technique, three main problems must be addressed. These are (a) state estimation, (b) parameter estimation and (c) robustness with respect to external disturbances or modelling errors.

State Estimation

The states of a system are those variables that provide a complete representation of the internal condition or state of the system at a given instant of time [1]. For example, the states of a motor might include the current through the windings, and the position and speed of the motor shaft. State estimation is applicable to virtually all areas of engineering and science. Any discipline that is concerned with the mathematical modelling of its systems is a likely candidate for state estimation.

State estimation is interesting to engineers for the following reasons:

- States are estimated to implement a state-feedback controller. For example, estimate the winding currents of a motor to control its position.

- States are estimated since it is expensive and impractical to measure all the states of the system using sensors.
- Additionally, the measurements are disturbed by noise thereby resulting in uncertainities and this is overcome through state estimation.
- In general, states are also estimated to monitor the health of the system.

Parameter Estimation

Parameter estimation is the process of using observations from a dynamic system to develop parametric mathematical models that adequately represent the system characteristic behavior [1]. In parametric estimation the structure of the model is assumed known while the parameters need to be estimated using special estimation techniques. Mathematical modelling via parameter estimation is one of the ways that leads to deeper understanding of the system's characteristics. These parameters often determine the stability and control behavior of the system. Estimation of these parameters from input-output data (signals) of the system is thus an important step in the analysis of the dynamic system.

The problems of state and parameter estimation have been extensively studied for the case of LTI systems. However, for LTV systems, these problems are still largely untackled.

The estimation in LTV systems is hence the major objective of this thesis. The initial results derived in this thesis are presented in our first conference paper; [3].

1.1 Estimation in LTV systems: Literature

Linear time-varying systems have received considerable attention in recent years. Description and analysis of physical systems by time-varying models, analysis of existing time-varying systems, more effective use of physical devices exhibiting time-varying characteristics, and adaptive feedback control of time-varying systems are some examples of practical as well as theoretical interest. A systematic design procedure for the synthesis of feedback control inputs, which guarantee asymptotic stabilization to zero of the incremental state variables, is possible only with accurate representation of the LTV system. Hence, state and parameter estimation of LTV systems is significant with this regard.

Non-asymptotic state and parameter estimation is attracting increased interest as the development of more efficient control technologies than current methods poses new challenges. Rapidly switching or hybrid control strategies are best matched with dead-beat observers to deliver optimal closed loop performance. A dead-beat observer has the special property that the estimated state will converge to the true state in exactly (at most) n steps, where n is the order of the state-space model. It is the fastest observer that is theoretically possible. Recursive estimation algorithms, with the powerful family of Kalman filters feature good noise rejection properties, but require careful implementation and tuning, as any errors in the assumptions about the system model, noise characteristic, or *initial conditions of the system* can have highly destabilizing effect resulting in their slow convergence.

Non-asymptotic algebraic observers would be an obvious choice if not the consensus of opinion that accurate higher order differentiation of noisy signals is unrealistic. Still, the algebraic differentiation approach based on operational calculus, as first introduced in [4], [5] has remarkably good noise rejection properties as discussed in [6]. Further improvements in the direction of derivative estimation are offered by [7], [8] and lead to the development of high fidelity non-asymptotic state and parameter estimators for LTI systems that do not exhibit singularities and thus deliver fast estimates of the states. A kernel linear system representation is proposed which exploits the obvious differential invariance in homogeneous linear systems that is a direct consequence of the validity of the Cayley Hamilton Theorem. Generally, a differential invariant of a dynamical system is a function $\mathcal{J}(t, y(t), y^{(1)}(t), ..., y^{n}(t))$ of time, output, and a number of its derivatives which remains constant under the action of the flow of the system. Any differential invariant can be regarded as a "deterministic signature" of the system that holds regardless of any uncertainties, a behavioral law that can be harnessed to help eliminate errors in measurement or numerical integration by simply enforcing its validity at all times.

A similar double-sided kernel representation of linear time-invariant systems employs the notion of a controlled invariance and captures the behaviour of systems with exogeuous inputs, as presented in [9]. The LTV kernel system represenation is singularity free, delivers a recursive formula for computation of output derivatives and immediately yields a non-asymptotic state observer on any finite observation window, much like the one presented in [10], but is developed entirely in the time domain and exhibits improved performance. The estimation window is moved forward as new output measurements arrive, which can be carried out safely in the absence of singularities. The observer does not need initialization as the notion of "initial conditions" is removed from the "behavioral" kernel system representation. The measurement noise is "filtered naturally" as all estimates are obtained as output functions of integral operators.

The ensemble of results pertaining to parameter estimation in LTV systems is relatively modest, see e.g. [11], [12], [13], [14] and references therein. The prevailing techniques include least squares optimization and involve assumptions about specific model structure e.g. convex polytopic structure (Takagi-Sugeno model) [13].

The approach presented here build on the ideas of [7], [8], [15], [9] and extends the application of kernel system representations to joint state, input and functional parameter estimation in linear time-varying systems. The kernel system representation for LTV systems is discussed in Chapter 3.

1.2 Detailed Thesis Objectives and Organization

The primary objective of this Master's thesis is to develop a new method for joint parameter, state, and input estimation for linear time-varying systems using B-spline functional approximation techniques combined with a double sided kernel representation of the system. The motivation to use B-splines is that B-splines work well for functional approximations over time and fit a wide variety of functions. To achieve the non-asymptotic state and parameter estimation, the double sided kernel is derived using the differential invariance principle of linear systems and the parameters are estimated by solving a system of linear equations as discussed in Chapter 4. The states are then estimated by evaluating the double sided kernel using the estimated parameters. The joint state, parameter, and input estimation constitute an original contribution as reported in [3].

Chapter 2 gives an overview of B-splines: its definition and its properties. A recursive formula for computing the B-splines is discussed along with its differentiation formula. The chapter also briefly touches upon P-splines and their properties which can provide improved performance over B-splines in the estimation as discussed in Chapter 5.

Chapter 3 discusses the detailed derivation of the double sided kernel for LTV systems. It exploits the differential invariance principle valid for linear systems to derive the forward and backward kernels for the system output and its time derivatives. An example of a third order LTV system is presented to validate the double sided kernel and our approach to estimation of LTV systems .

Chapter 4 proposes the new estimation method which combines B-splines with the kernel representation of the LTV system. It describes the algebraic equations employed to obtain the parameters of the system and the linear identifiability conditions necessary for it. It also describes the technique employed for state and input estimation. The chapter concludes with sections describing joint state, parameter, and input estimation using the proposed approach.

Chapter 5 presents the simulation results of the proposed novel method of estimation in LTV systems. The results are presented for second and third order LTV systems. In order to prove the robustness of the proposed method, four cases are considered for both second and third order LTV systems. The proposed method is valid only for the output corrupted by little or no noise. However, results for the noisy case are also presented for clarification. Finally, the last section presents the simulation results obtained using P-splines.

Chapter 6 concludes the thesis by providing a synopsis of the results and suggesting possible future work.

Chapter 2

B-splines

2.1 Introduction

A B-spline function is a piecewise defined polynomial function with several beneficial properties such as numerical stability of computations, local effects of coefficient changes and built-in smoothness between neighboring polynomial pieces [16]. A common application of B-spline functions, curves and surfaces is fitting of data points. Fitting can either be interpolation or approximation. An interpolating B-spline function passes through the data points, whereas an approximating B-spline function minimizes the residuals between the function and the data but does not pass through the data points in general. The representation using B-splines is popular in computer-aided design, modeling and engineering as well as computer graphics for the geometry of curves, objects and surfaces [17]. It is also used for planning trajectories of computer controlled industrial machines [18] and robots [19, 20].

We define k-th order B-splines as appropriately scaled k-th divided differences of the truncated power function and its properties in the following section.

Definition 2.1.1 *[21] Let $t := (t_i)$ be a nondecreasing sequence (which may be finite or infinite). The i-th (normalized) B-spline of order k for the knot sequence t is denoted by $B_{i,k,t}$ and is defined by the rule*

$$B_{i,k,t}(x) := (t_{i+k} - t_i)[t_i, \ldots.., t_{i+k}](. - x)_+^{k-1}, \quad all \;\; k \in \mathbb{N} \tag{2.1}$$

The "placeholder" notation employed here may be unfamiliar to the reader. It is used here to indicate that the k-th divided difference of the function $(t-x)_+^{k-1}$ of the two variables t and x. The divided difference is calculated by fixing x and considering $(t-x)_+^{k-1}$ as a function of t alone. The resulting number depends, of course, on the particular value of x we chose prior to the calculation, i.e., the resulting number varies as we vary x, and so we obtain eventually the function $B_{i,k,t}$ of x.

2.2 Properties

A function $\mathrm{B}_{i,p,t}(x)$, $B_{i,p,t}(\cdot) : \mathbb{R} \to \mathbb{R}$ is called the i-th B-spline of order p with a knot sequence $t := \{t_1, \cdots, t_{p+1}\}, t_j \leq t_{j+1}$; see [21] for a comprehensive introduction to B-spline spaces, their properties and applications.

A B-spline of order p is a piecewise polynomial function of degree $p-1$ in the variable x (the pieces connect continuously at the knots). A B-spline is uniquely defined by the $p+1$ knots and is zero outside the interval $(t_1, \cdots, t_{p+1})$. B-splines are called "uniform" if the knots are equidistant. Some derivatives of B-splines may also be continuous, depending on whether the consecutive knots are distinct or not. If all knots in a B-spline are all distinct, then their derivatives are continuous up to order $p-1$. If the knots are coincident at a given value of x, the continuity of derivative order is reduced by 1 for each additional coincident knot.

The B-splines have the property that any spline function of order p on a given set of knots t can be expressed as a linear combination

$$S_{p,t}(x) = \sum_i \alpha_i B_{i,p,t}(x) \tag{2.2}$$

hence lending themselves well to function approximation or interpolation. To approximate a desired function over some inteval $[a, b]$ the splines should cover a superset of $[a, b]$.

Expressions for the polynomial pieces $B(i, p, t)$ can be derived by means of the Cox-de Boor recursion formula:

$$\mathrm{B}_{i,1,t}(x) = \begin{cases} 1 & \text{if } t_i \leq x \leq t_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

Higher order B-splines are constructed using the recursion

$$\mathrm{B}_{i,p,t}(x) = \omega_{i,p}(x)\mathrm{B}_{i,p-1,t}(x) + (1 - \omega_{i+1,p}(x))\mathrm{B}_{i+1,p-1,t}(x)$$

where $\omega_{i,p}(x)$ is given by

$$\omega_{i,p}(x) = \frac{x - t_i}{t_{i+p-1} - t_i}$$

The following differentiation formula is valid for B-splines

$$\frac{d}{dx}(\sum_{i=1}^{n} \alpha_i \mathrm{B}_{i,p,t}(x)) = \sum_{i=1}^{n} \alpha_i \mathrm{B}'_{i,p,t}(x) \tag{2.3}$$

where $\mathrm{B}'_{i,p,t}(x)$ is given by

$$\mathrm{B}'_{i,p,t}(x) = \frac{p-1}{t_{i+p-1} - t_i}\mathrm{B}_{i,p-1,t}(x) - \frac{p-1}{t_{i+p} - t_{i+1}}\mathrm{B}_{i+1,p-1,t}(x) \tag{2.4}$$

Again the differentiation formula is recursive in the sense that higher order derivatives are obtained by recursion from equation (2.4). It is given by,

$$\mathrm{B}^{(i)}_{i,p,t}(x) = \frac{p-1}{t_{i+p-1} - t_i}\mathrm{B}^{(i-1)}_{i,p-1,t}(x) - \frac{p-1}{t_{i+p} - t_{i+1}}\mathrm{B}^{(i-1)}_{i+1,p-1,t}(x) \tag{2.5}$$

While writing function approximations by B-splines it is convenient to omit indicating the order and knot sequence in the index of the B-splines, thus denoting the k-th B-spline by B_k while letting the order and knot sequence be clear from the context.

2.3 P-Splines

P-Splines are a variation of B-Splines. There are two components of a P-Spline: B-Splines and discrete penalties. As we have discussed B-Splines in the previous section, let us proceed to discrete penalties.

Discrete Penalties

With the number of B-splines in the basis of polynomial functions we can tune the smoothness of a curve to the data at hand. A smaller number of splines gives a smoother result.

However, this is not the only possibility. We can also use a large basis and additionally constrain the coefficients of the B-splines, to achieve the desired smoothness. A properly chosen penalty allows this.

O'Sullivan (1986) first proposed to use a large number of knots and a penalty on the second derivative of the curve to prevent overfitting. The penalty is a way to measure the roughness of a curve and is given by:

$$R = \int_l^u [f^{(2)}(x)]^2 dx$$

where l and u indicate the bounds of the domain of x. If $f(x) = \sum_j a_j B_j(x)$, we can derive a banded matrix P such that $R = a^T P a$. The elements of P are computed as the integrals of products of second derivatives of neighbouring B-splines [22].

The computation of P is not trivial, and it becomes quite tedious when the third or fourth order derivative is used to measure roughness. P-splines circumvent the issue by dropping derivatives and integrals completely. Instead they use a discrete penalty matrix from the start. It is also simple to compute, as it is based on difference formulas. Let $\Delta a_j = a_j - a_{j-1}$, $\Delta^2 a_j = \Delta(\Delta a_j) = a_j - 2a_{j-1} + a_{j-2}$ and in general $\Delta^d a_j = \Delta(\Delta^{d-1} a_j)$. Let D_d be a matrix such that $D_d a = \Delta^d a$. If we replace the penalty by $\lambda ||D_d a||^2 = \lambda a^T D_d^T D_d a = \lambda a^T P a$, we get a similar construction as O'Sullivan's, but with a minimal amount of work. In modern languages like R and Matlab, D_d can be obtained mechanically as the dth order difference of the identity matrix.

2.3.1 Properties

P-splines have a number of useful properties, partially inherited from B-splines [23].

- P-splines show no boundary effects, as many types of kernel smoothers do. By this we mean the spreading of a fitted curve or density outside of the (physical) domain of the data, generally accompanied by bending toward zero.
- P-splines can fit polynomial data exactly. Let data (x_i, y_i) be given. If the y_i are a polynomial in x of degree k, then B-splines of degree k or higher will exactly fit the

data [21]. The same is true for P-splines, if the order of the penalty is $k+1$ or higher, whatever the value of λ.

- The limit of a P-splines fit with strong smoothing is a polynomial. For large values of λ and a penalty of order k, the fitted series will approach a polynomial of degree $k-1$, if the degree of the B-splines is equal to, or higher than, k.

The discrete penalties are somewhat less interpretable in terms of function shape than the traditional derivative based spline penalties, but tend towards penalties proportional to traditional spline penalties in the limit of large basis size. However, part of the point of P-splines is not to use a large basis size. In addition, the spline basis functions arise from solving functional optimization problems involving derivative based penalties, so moving to discrete penalties for smoothing may not always be desirable. Hence, P-splines with derivative based penalties has been proposed by [24]. We use the approach mentioned in [24] to compute the penalties for the joint parameter and state estimation in LTV systems.

As the kernel system representation of [9] is linear in the system parameters, the approach presented here can be seamlessly combined with standard B-spline functional approximation. Under mild identifiability assumptions the sampled input-output data is employed in the kernel system representation to deliver linear algebraic equations for the values of the B-spline coefficients thereby reconstructing the unknown functional parameters in finite time with almost uniform accuracy. State reconstruction follows by way of calculating the time derivatives of the system output. The joint B-spline estimation approach is non-asymptotic. This is discussed in Chapter 4.

Chapter 3

A Double Sided Kernel for an LTV System

3.1 Introduction

As seen in Chapter 1, parameter and state estimation literature for LTV systems is limited compared to its LTI counterpart. We develop a novel method of estimation by deriving a kernel representation for linear time-varying systems using the differential invariance principle . The kernel representation of the LTV system is combined with B-Spline functional approximations to estimate the parameters, input, and the states of a system. The derivation details are presented in the next section.

3.2 Kernel Representation of an LTV System with Exogenous Input

3.2.1 Model Assumptions

The LTV system considered is assumed to be stated in the state space form:

$$\dot{x} = A(t)x + B(t)u \ ; \quad y = C(t)x \tag{3.1}$$

where the state, input and output are: $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ and $y \in \mathbb{R}^d$; with $A(t) \in \mathbb{R}^{n\times n}$, $B(t) \in \mathbb{R}^{n\times m}$, and $C(t) \in \mathbb{R}^{d\times n}$ are known, $n-1$ times differentiable matrix functions of time.
The following definitions and criteria for complete and total observability of a general system (3.1) can be found in [25] and [26]:

Definition 1

- System (3.1) is completely observable on a time interval $[t_0, t_f]$ if any initial state can be determined from the knowledge of the system output and input, $y(t)$ and $u(t)$, on the interval $t \in [t_0, t_f]$.
- System (3.1) is totally observable on a time interval $[t_0, t_f]$ if it is completely observable on any subinterval of $[t_0, t_f]$.

Theorem 1 *[26] System* (3.1) *restricted to the time interval* $[t_0, t_f]$ *is:*

- *completely observable if rank* $\mathcal{O}(t) = n$ *for* $t \in [t_0, t_f]$*;*
- *totally observable if and only if rank* $\mathcal{O}(t) = n$ *on any subinterval of* $[t_0, t_f]$*;*

where the observability matrix is defined by

$$\begin{aligned} &\mathcal{O}(t) = \{S_0(t), S_1(t), \dots, S_{n-1}(t)\} \\ &S_0(t) = C(t)^T; \\ &S_{k+1}(t) = A(t)^T S_k(t) + \dot{S}_k(t); k = 0, \dots, n-2 \end{aligned} \tag{3.2}$$

3.2.2 A Controlled Differential Invariant - the Input-Output Equation

Henceforth, only single-input, single-output and strictly proper systems will be considered, i.e., $m = 1$, $m < n$, with $d = 1$ so $u, y \in \mathbb{R}$.
With assumption of total observability, the state variable of system (3.1) can be expressed in terms of the system input, output, and their derivatives; see, e.g., [10] for a simple proof of this fact:

Theorem 2 *[10] Under the total observability assumption the input-output relation in system* (3.1) *is of the form:*

$$\sum_{i=0}^{n} a_i(t)y^{(i)} + \sum_{i=0}^{m} b_i(t)u^{(i)} = 0 \qquad \text{for } t \in [t_0, t_f] \tag{3.3}$$

with $a_n = 1$, $m < n$. *The state* x *can be expressed as a linear function of the input and output and their derivatives:*

$$x(t) = \begin{pmatrix} \Gamma_0 \\ \Gamma_1 \\ \Gamma_2 \\ \vdots \\ \Gamma_{n-1} \end{pmatrix}^{-1} \left[\begin{pmatrix} y \\ y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(n-1)} \end{pmatrix} - M \begin{pmatrix} u \\ u^{(1)} \\ u^{(2)} \\ \vdots \\ u^{(n-2)} \end{pmatrix} \right] \tag{3.4}$$

where

$$\Gamma_0(t) = C(t)$$

$$\Gamma_k(t) = \left(\left(A(t)^T + \frac{d}{dt} \right)^k C(t)^T \right)^T, \quad 0 < k < n$$

$$M = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ \Delta_{11} & 0 & 0 & \dots & 0 \\ \Delta_{21} & \Delta_{22} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \Delta_{(n-1)1} & \Delta_{(n-1)2} & \Delta_{(n-1)3} & \dots & \Delta_{(n-1)(n-1)} \end{pmatrix}$$

$$\Delta_{k0}(t) = \Gamma_k(t)B(t)$$

$$\Delta_{kj}(t) = \begin{cases} C(t)B(t); & \text{if } j = k \\ \Delta_{(k-1)(j-1)}(t) + \frac{d}{dt}\Delta_{(k-1)j}(t); & \text{if } 1 \leq j < k \end{cases} \tag{3.5}$$

The invertibility of the matrix in (3.4) *is guaranteed by the total observability assumption on the system; see the definition of the observability matrix in Theorem 1.*

Equation (3.3) constitutes a "controlled differential invariant" of system (3.1) which delivers a non-singular integral kernel representation of the system as well as integral transforms for computation of the output derivatives.

Theorem 3 *There exist Hilbert-Schmidt kernels [27]* K_y, K_u, K_y^i, K_u^i, $i = 1, \cdots n-1$, *and functions* f_y^i, f_u^i, $i = 0, \cdots n-2$, *defined respectively on* $[a, b] \times [a, b]$ *and* $[a, b]$ *such that the output* y *of system (3.1) satisfies the following integral equation on* $[a, b]$

$$y(t) = \int_a^b K_y(t,\tau)y(\tau)\ d\tau + \int_a^b K_u(t,\tau)u(\tau)\ d\tau \tag{3.6}$$

while the derivatives of the output $y^{(1)}, \cdots y^{(n-1)}$ *for* $i = 1, \cdots n-1$ *satisfy the recursive relationships :*

$$y^{(i)}(t) = \sum_{k=0}^{i-1} f_y^k(t)y^{(k)}(t) + \sum_{k=0}^{i-1} f_u^k(t)u^{(k)}(t) + \int_a^b K_y^i(t,\tau)y(\tau)\ d\tau + \int_a^b K_u^i(t,\tau)u(\tau)\ d\tau \tag{3.7}$$

□

While the analytical formulae for the kernels for a general n-dimensional system maybe somewhat cumbersome their derivation is surprisingly straightforward. It is to mention that the derivation of such system representation can equivalently be conducted in the Laplace domain or else by direct embedding of (3.3) into a Sobolev space $H_n^2[a, b]$ of functions whose n-th derivatives are absolutely integrable. The proof of the above theorem is done through induction and is presented elsewhere. For clarity and simplicity we will only demonstrate the derivation directly in the time domain with reference to a 3-dimensional example. To this end, let $n = 3$, and consider two equations obtained from (3.3) by pre-multiplication by the respective factors $(\xi - a)$ and $(b - \zeta)$:

$$\begin{aligned}(\xi - a)^3 y^{(3)} + a_2(\xi)(\xi - a)^3 y^{(2)} + a_1(\xi)(\xi - a)^3 y^{(1)} + a_0(\xi)(\xi - a)^3 y \\ + b_2(\xi)(\xi - a)^3 u^{(2)} + b_1(\xi)(\xi - a)^3 u^{(1)} + b_0(\xi)(\xi - a)^3 u = 0\end{aligned} \tag{3.8}$$

$$\begin{aligned}(b - \zeta)^3 y^{(3)} + a_2(\zeta)(b - \zeta)^3 y^{(2)} + a_1(\zeta)(b - \zeta)^3 y^{(1)} + a_0(\zeta)(b - \zeta)^3 y \\ + b_2(\zeta)(b - \zeta)^3 u^{(2)} + b_1(\zeta)(b - \zeta)^3 u^{(1)} + b_0(\zeta)(b - \zeta)^3 u = 0\end{aligned} \tag{3.9}$$

Each of the above is then integrated three times, on the respective intervals $[a, a + \tau]$ and on $[b - \sigma, b]$ while assuming that τ and σ are related by $a + \tau = b - \sigma$. Integration by parts is used whenever it allows to lower the degree of the derivatives appearing under the integrals and the result is then simplified algebraically before proceeding to the next integration. To illustrate this process, integrating the first term of (3.8) yields

$$\begin{aligned}\int_a^{a+\tau} (\xi - a)^3 y^{(3)}(\xi)\, \mathrm{d}\xi &= (\xi - a)^3 y^{(2)}(\xi) \,|_a^{a+\tau} - \int_a^{a+\tau} 3(\xi - a)^2 y^{(2)}(\xi)\, \mathrm{d}\xi \\ &= \tau^3 y^{(2)}(a + \tau) - \left[3(\xi - a)^2 y^{(1)}(\xi) \,|_a^{a+\tau} - \int_a^{a+\tau} 6(\xi - a) y^{(1)}(\xi), \mathrm{d}\xi \right] \\ &= \tau^3 y^{(2)}(a + \tau) - 3\tau^2 y^{(1)}(a + \tau) + 6(\xi - a) y(\xi) \,|_a^{a+\tau} - \int_a^{a+\tau} 6 y(\xi)\, \mathrm{d}\xi \\ &= \tau^3 y^{(2)}(a + \tau) - 3\tau^2 y^{(1)}(a + \tau) + 6\tau y(a + \tau) - \int_a^{a+\tau} 6 y(\xi)\, \mathrm{d}\xi\end{aligned} \tag{3.10}$$

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 (3.10) again,

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

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

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

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

$$= \tau^3y^{(1)}(a+\tau) - 6(\xi^{'}-a)^2y(\xi^{'})\,|_a^{a+\tau} + \int_a^{a+\tau} 12(\xi^{'}-a)y(\xi^{'})\,\mathrm{d}\xi^{'}$$

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

$$= \tau^3y^{(1)}(a+\tau) - 6\tau^2y(a+\tau) + \int_a^{a+\tau} 18(\xi'-a)y(\xi')\,\mathrm{d}\xi' - \int_a^{a+\tau}\int_a^{\xi'} 6y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi' \tag{3.11}$$

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

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

$$= \int_a^{a+\tau}(\xi^{''}-a)^3y^{(1)}(\xi^{''})\mathrm{d}\xi^{''} - \int_a^{a+\tau} 6(\xi^{''}-a)^2y(\xi^{''})\mathrm{d}\xi^{''} + \int_a^{a+\tau}\int_a^{\xi''} 18(\xi'-a)y(\xi')\,\mathrm{d}\xi'\mathrm{d}\xi''$$

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

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

$$- \int_a^{a+\tau} 6(\xi'' - a)^2 y(\xi'') \mathrm{d}\xi'' + \int_a^{a+\tau}\int_a^{\xi''} 18(\xi' - a) y(\xi')\, \mathrm{d}\xi' \mathrm{d}\xi'' - \int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} 6y(\xi)\, \mathrm{d}\xi \mathrm{d}\xi' \mathrm{d}\xi''$$

$$= \tau^3 y(a+\tau) - \int_a^{a+\tau} 9(\xi'' - a)^2 y(\xi'')\, \mathrm{d}\xi''$$

$$+ \int_a^{a+\tau}\int_a^{\xi''} 18(\xi' - a) y(\xi')\, \mathrm{d}\xi' \mathrm{d}\xi'' - \int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} 6y(\xi)\, \mathrm{d}\xi \mathrm{d}\xi' \mathrm{d}\xi'' \tag{3.12}$$

Integrating the second term in (3.8) first time,

$$\int_a^{a+\tau} a_2(\xi)(\xi - a)^3 y^{(2)}(\xi)\, \mathrm{d}\xi$$

$$= a_2(\xi)(\xi - a)^3 y^{(1)}(\xi) \mid_a^{a+\tau} - \int_a^{a+\tau} \left[3a_2(\xi)(\xi - a)^2 + (\xi - a)^3 a_2^{(1)}(\xi)\right] y^{(1)}(\xi)\, \mathrm{d}\xi$$

$$= a_2(\xi)(\xi - a)^3 y^{(1)}(\xi) \mid_a^{a+\tau} - \left[3a_2(\xi)(\xi - a)^2 + (\xi - a)^3 a_2^{(1)}(\xi)\right] y(\xi) \mid_a^{a+\tau}$$

$$+ \int_a^{a+\tau} \left[6(\xi - a)a_2(\xi) + 3(\xi - a)^2 a_2^{(1)}(\xi)\right] y(\xi) \mathrm{d}\xi$$

$$+ \int_a^{a+\tau} \left[3(\xi - a)^2 a_2^{(1)}(\xi) + (\xi - a)^3 a_2^{(2)}(\xi)\right] y(\xi) \mathrm{d}\xi$$

$$= a_2(a+\tau)(\tau)^3 y^{(1)}(a+\tau)$$

$$- \left[3a_2(a+\tau)(\tau)^2 + (\tau)^3 a_2^{(1)}(a+\tau)\right] y(a+\tau)$$

$$+ \int_a^{a+\tau} \left[6(\xi - a)a_2(\xi) + 3(\xi - a)^2 a_2^{(1)}(\xi)\right] y(\xi) \mathrm{d}\xi$$

$$+\int_a^{a+\tau}\left[3(\xi-a)^2a_2^{(1)}(\xi)+(\xi-a)^3a_2^{(2)}(\xi)\right]y(\xi)\mathrm{d}\xi \tag{3.13}$$

As shown earlier,the upper limit again becomes a 'dummy variable' and now we set $\xi^{'} = a+\tau$. Integrating the second time yields,

$$\int_a^{a+\tau}\int_a^{\xi'} a_2(\xi)(\xi-a)^3y^{(2)}(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'$$

$$=\int_a^{a+\tau} a_2(\xi')(\xi'-a)^3y^{(1)}(\xi')\,\mathrm{d}\xi' - \int_a^{a+\tau}\left[3a_2(\xi')(\xi'-a)^2+(\xi'-a)^3a_2^{(1)}(\xi')\right]y(\xi')\,\mathrm{d}\xi'$$

$$+\int_a^{a+\tau}\int_a^{\xi'}\left[6(\xi-a)a_2(\xi)+3(\xi-a)^2a_2^{(1)}(\xi)\right]y(\xi)\mathrm{d}\xi\mathrm{d}\xi'$$

$$+\int_a^{a+\tau}\int_a^{\xi'}\left[3(\xi-a)^2a_2^{(1)}(\xi)+(\xi-a)^3a_2^{(2)}(\xi)\right]y(\xi)\mathrm{d}\xi\mathrm{d}\xi'$$

$$=a_2(\xi')(\xi'-a)^3y(\xi')\mid_a^{a+\tau} - \int_a^{a+\tau}\left[3a_2(\xi')(\xi'-a)^2+(\xi'-a)^3a_2^{(1)}(\xi')\right]y(\xi')\,\mathrm{d}\xi'$$

$$-\int_a^{a+\tau}\left[3a_2(\xi')(\xi'-a)^2+(\xi'-a)^3a_2^{(1)}(\xi')\right]y(\xi')\,\mathrm{d}\xi'$$

$$+\int_a^{a+\tau}\int_a^{\xi'}\left[6(\xi-a)a_2(\xi)+3(\xi-a)^2a_2^{(1)}(\xi)\right]y(\xi)\mathrm{d}\xi\mathrm{d}\xi'$$

$$+\int_a^{a+\tau}\int_a^{\xi'}\left[3(\xi-a)^2a_2^{(1)}(\xi)+(\xi-a)^3a_2^{(2)}(\xi)\right]y(\xi)\mathrm{d}\xi\mathrm{d}\xi'$$

$$=a_2(a+\tau)(\tau)^3y(a+\tau) - \int_a^{a+\tau}2\left[3a_2(\xi')(\xi'-a)^2+(\xi'-a)^3a_2^{(1)}(\xi')\right]y(\xi')\,\mathrm{d}\xi'$$

$$+\int_a^{a+\tau}\int_a^{\xi'}\left[6(\xi-a)a_2(\xi)+3(\xi-a)^2a_2^{(1)}(\xi)\right]y(\xi)\mathrm{d}\xi\mathrm{d}\xi'$$

$$+\int_a^{a+\tau}\int_a^{\xi'}\Big[3(\xi-a)^2a_2^{(1)}(\xi)+(\xi-a)^3a_2^{(2)}(\xi)\Big]y(\xi)\mathrm{d}\xi\mathrm{d}\xi' \tag{3.14}$$

Integrating the second term of (3.8) third time yields,

$$\begin{aligned}
&\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} a_2(\xi)(\xi-a)^3y^{(2)}(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''\\
&=\int_a^{a+\tau} a_2(\xi'')(\xi''-a)^3y(\xi'')\,\mathrm{d}\xi''-\int_a^{a+\tau}\int_a^{\xi''}2\big[3a_2(\xi')(\xi'-a)^2\big]y(\xi')\,\mathrm{d}\xi'\mathrm{d}\xi''\\
&\quad-\int_a^{a+\tau}\int_a^{\xi''}2\big[(\xi'-a)^3a_2^{(1)}(\xi')\big]y(\xi')\,\mathrm{d}\xi'\mathrm{d}\xi''+\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'}\big[6(\xi-a)a_2(\xi)\big]y(\xi)\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''\\
&\quad+\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'}\big[3(\xi-a)^2a_2^{(1)}(\xi)\big]y(\xi)\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''+\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'}\big[3(\xi-a)^2a_2^{(1)}(\xi)\big]y(\xi)\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''\\
&\quad+\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'}\big[(\xi-a)^3a_2^{(2)}(\xi)\big]y(\xi)\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''
\end{aligned} \tag{3.15}$$

Integrating the third term of (3.8) first time yields,

$$\begin{aligned}
&\int_a^{a+\tau} a_1(\xi)(\xi-a)^3y^{(1)}(\xi)\,\mathrm{d}\xi\\
&=a_1(\xi)(\xi-a)^3y(\xi)\,\big|_a^{a+\tau}-\int_a^{a+\tau}3a_1(\xi)(\xi-a)^2y(\xi)\,\mathrm{d}\xi-\int_a^{a+\tau}a_1^{(1)}(\xi)(\xi-a)^3y(\xi)\,\mathrm{d}\xi\\
&=a_1(a+\tau)(\tau)^3y(a+\tau)-\int_a^{a+\tau}3a_1(\xi)(\xi-a)^2y(\xi)\,\mathrm{d}\xi-\int_a^{a+\tau}a_1^{(1)}(\xi)(\xi-a)^3y(\xi)\,\mathrm{d}\xi
\end{aligned} \tag{3.16}$$

Integrating the second time,

$$\int_a^{a+\tau}\int_a^{\xi'} a_1(\xi')(\xi'-a)^3 y^{(1)}(\xi')\,\mathrm{d}\xi\mathrm{d}\xi'$$
$$=\int_a^{\xi'} a_1(\xi')(\xi'-a)^3 y(\xi')\mathrm{d}\xi' - \int_a^{a+\tau}\int_a^{\xi''} 3a_1(\xi')(\xi'-a)^2 y(\xi')\,\mathrm{d}\xi'\mathrm{d}\xi''$$
$$-\int_a^{a+\tau}\int_a^{\xi''} a_1^{(1)}(\xi')(\xi'-a)^3 y(\xi')\,\mathrm{d}\xi'\mathrm{d}\xi'' \tag{3.17}$$

Integrating the third time yields,

$$\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} a_1(\xi)(\xi-a)^3 y^{(1)}(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''$$
$$=\int_a^{a+\tau}\int_a^{\xi''} a_1(\xi')(\xi'-a)^3 y(\xi')\,\mathrm{d}\xi'\mathrm{d}\xi'' - \int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} 3a_1(\xi)(\xi-a)^2 y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''$$
$$-\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} a_1^{(1)}(\xi)(\xi-a)^3 y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi'' \tag{3.18}$$

Finally, the fourth term is

$$\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} a_0(\xi)(\xi-a)^3 y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi'' \tag{3.19}$$

Repeating the same procedure for the next three terms involving the input we obtain the following. Integrating the fifth term yields,

$$\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} b_2(\xi)(\xi-a)^3 u^{(2)}(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''$$

$$= \int_a^{a+\tau} b_2(\xi'')(\xi''-a)^3 u(\xi'')\,\mathrm{d}\xi'' - \int_a^{a+\tau}\int_a^{\xi''} 2\big[3b_2(\xi')(\xi'-a)^2\big]u(\xi')\,\mathrm{d}\xi'\mathrm{d}\xi''$$
$$- \int_a^{a+\tau}\int_a^{\xi''} 2\big[(\xi'-a)^3 b_2^{(1)}(\xi')\big]u(\xi')\,\mathrm{d}\xi'\mathrm{d}\xi'' + \int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} \big[6(\xi-a)b_2(\xi)\big]u(\xi)\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''$$
$$+ \int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} \big[6(\xi-a)^2 b_2^{(1)}(\xi)\big]u(\xi)\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi'' + \int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} \big[(\xi-a)^3 b_2^{(2)}(\xi)\big]u(\xi)\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi'' \tag{3.20}$$

Integrating the sixth term yields,

$$\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} b_1(\xi)(\xi-a)^3 u^{(1)}(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''$$
$$= \int_a^{a+\tau}\int_a^{\xi''} b_1(\xi')(\xi'-a)^3 u(\xi')\,\mathrm{d}\xi'\mathrm{d}\xi'' - \int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} 3b_1(\xi)(\xi-a)^2 u(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''$$
$$- \int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} b_1^{(1)}(\xi)(\xi-a)^3 u(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi'' \tag{3.21}$$

Integrating the last term gives,

$$\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'} b_0(\xi)(\xi-a)^3 u(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi'' \tag{3.22}$$

Collecting the terms in (3.10) - (3.22) yields

$$\tau^3 y(a+\tau)$$
$$= \int_a^{a+\tau} \Big[9(\xi''-a)^2 - a_2(\xi'')(\xi''-a)^3\Big] y(\xi'')\,\mathrm{d}\xi''$$
$$+ \int_a^{a+\tau}\int_a^{\xi''} \Big[-18(\xi'-a) + 6a_2(\xi')(\xi'-a)^2 + (\xi'-a)^3(2a_2^{(1)}(\xi') - a_1(\xi'))\Big] y(\xi')\,\mathrm{d}\xi'\mathrm{d}\xi''$$

$$+\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'}\Big[6-6a_2(\xi)(\xi-a)-6(\xi-a)^2a_2^{(1)}(\xi)-(\xi-a)^3a_2^{(2)}(\xi)+3a_1(\xi)(\xi-a)^2$$

$$+(\xi-a)^3a_1^{(1)}(\xi)-a_0(\xi)(\xi-a)^3\Big]y(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi''-\int_a^{a+\tau}b_2(\xi'')(\xi''-a)^3u(\xi'')\,\mathrm{d}\xi''$$

$$+\int_a^{a+\tau}\int_a^{\xi''}\Big[6b_2(\xi')(\xi'-a)^2+(\xi'-a)^3(2b_2^{(1)}(\xi')-b_1(\xi'))\Big]u(\xi')\,\mathrm{d}\xi'\mathrm{d}\xi''$$

$$+\int_a^{a+\tau}\int_a^{\xi''}\int_a^{\xi'}\Big[-6b_2(\xi)(\xi-a)-6(\xi-a)^2b_2^{(1)}(\xi)-(\xi-a)^3b_2^{(2)}(\xi)+3b_1(\xi)(\xi-a)^2$$

$$+(\xi-a)^3b_1^{(1)}(\xi)-b_0(\xi)(\xi-a)^3\Big]u(\xi)\,\mathrm{d}\xi\mathrm{d}\xi'\mathrm{d}\xi'' \tag{3.23}$$

Since the integrals in (3.23) are of a special form, this can be reduced to single integrals using Cauchy formula for repeated integration which can be recollected as follows. Let f be a continuous function on the real line, the nth repeated integral of f based at a,

$$f^{(-n)}(x)=\int_a^x\int_a^{\sigma_1}...\int_a^{\sigma_{n-1}}f(\sigma_n)\mathrm{d}\sigma_n...\mathrm{d}\sigma_2\mathrm{d}\sigma_1 \tag{3.24}$$

is given by the single integration

$$f^{(-n)}(x)=\frac{1}{(n-1)!}\int_a^x(x-t)^{n-1}f(t)\mathrm{d}t \tag{3.25}$$

Applying the Cauchy formula for repeated integration stated above on (3.23) while letting $a+\tau=t$, we get,

$$(t-a)^3y(t)=$$

$$\int_a^t\Big[9(\xi-a)^2-a_2(\xi)(\xi-a)^3\Big]y(\xi)\,\mathrm{d}\xi+\int_a^t(t-\xi)\Big[-18(\xi-a)+6a_2(\xi)(\xi-a)^2$$

$$+2(\xi-a)^3a_2^{(1)}(\xi)-a_1(\xi)(\xi-a)^3\Big]y(\xi)\,\mathrm{d}\xi+\frac{1}{2}\int_a^t(t-\xi)^2\Big[6-6a_2^{(1)}(\xi)(\xi-a)$$

$$
\begin{aligned}
&- a_2^{(2)}(\xi)(\xi-a)^3 + 3a_1(\xi)(\xi-a)^2 + a_1^{(1)}(\xi)(\xi-a)^3 - a_0(\xi)(\xi-a)^3\Big] y(\xi)\,\mathrm{d}\xi \\
&+ \int_a^t [-b_2(\xi)(\xi-a)^3] u(\xi)\,\mathrm{d}\xi \\
&+ \int_a^t (t-\xi)\Big[6b_2(\xi)(\xi-a)^2 + 2(\xi-a)^3 b_2^{(1)}(\xi) - b_1(\xi)(\xi-a)^3\Big] u(\xi)\,\mathrm{d}\xi \\
&+ \frac{1}{2}\int_a^t (t-\xi)^2\Big[-6b_2^{(1)}(\xi)(\xi-a) - b_2^{(2)}(\xi)(\xi-a)^3 + 3b_1(\xi)(\xi-a)^2 \\
&+ b_1^{(1)}(\xi)(\xi-a)^3 - b_0(\xi)(\xi-a)^3\Big] u(\xi)\,\mathrm{d}\xi \\
&\triangleq \int_a^t K_{Fy}(t,\tau) y(\tau)\,\mathrm{d}\tau + \int_a^t K_{Fu}(t,\tau) u(\tau)\,\mathrm{d}\tau
\end{aligned}
\tag{3.26}
$$

with $K_{Fy}(t,\tau)$ defined as

$$
\begin{aligned}
K_{Fy}(t,\tau) &\triangleq \Big[9(\tau-a)^2 - (\tau-a)^3 a_2(\tau)\Big] \\
&+ (t-\tau)\Big[-18(\tau-a) + 6(\tau-a)^2 a_2(\tau) + 2(\tau-a)^3 a_2^{(1)}(\tau) - (\tau-a)^3 a_1(\tau)\Big] \\
&+ \frac{(t-\tau)^2}{2}\Big[6 - 6(\tau-a)a_2(\tau) - 6(\tau-a)^2 a_2^{(1)}(\tau) - (\tau-a)^3 a_2^{(2)}(\tau) \\
&\qquad + 3(\tau-a)^2 a_1(\tau) + (\tau-a)^3 a_2^{(1)}(\tau) - (\tau-a)^3 a_0(\tau)\Big]
\end{aligned}
\tag{3.27}
$$

and $K_{Fu}(t,\tau)$ defined as

$$
\begin{aligned}
K_{Fu}(t,\tau) &\triangleq \Big[-(\tau-a)^3 b_2(\tau)\Big] \\
&+ (t-\tau)\Big[6(\tau-a)^2 b_2(\tau) + 2(\tau-a)^3 b_2^{(1)}(\tau) - (\tau-a)^3 b_1(\tau)\Big] \\
&+ \frac{(t-\tau)^2}{2}\Big[-6(\tau-a)b_2(\tau) - 6(\tau-a)^2 b_2^{(1)}(\tau) - (\tau-a)^3 b_2^{(2)}(\tau) \\
&\qquad + 3(\tau-a)^2 b_1(\tau) + (\tau-a)^3 b_2^{(1)}(\tau) - (\tau-a)^3 b_0(\tau)\Big]
\end{aligned}
\tag{3.28}
$$

In order to derive the backward kernel we employ the same procedure developed above on (3.9). See Appendix A for details. From (3.26) and (A.15) we have the following,

$$(t-a)^3 y(t) = \int_a^t K_{Fy}(t,\tau) y(\tau)\,\mathrm{d}\tau + \int_a^t K_{Fu}(t,\tau) u(\tau)\,\mathrm{d}\tau \tag{3.29}$$

$$(b-t)^3 y(t) = \int_t^b K_{By}(t,\tau) y(\tau)\,\mathrm{d}\tau + \int_t^b K_{Bu}(t,\tau) u(\tau)\,\mathrm{d}\tau \tag{3.30}$$

Adding side by side (3.29) and (3.30) while dividing both sides by $[(t-a)^3 + (b-t)^3]$ yields

$$y(t) = \int_a^b K_y(t,\tau) y(\tau)\ d\tau + \int_a^b K_u(t,\tau) u(\tau)\ d\tau \tag{3.31}$$

with

$$K_y(t,\tau) = \frac{1}{C_{[a,b]}(t)} \begin{cases} K_{Fy}(t,\tau) & \text{for } \tau \le t \\ K_{By}(t,\tau) & \text{for } \tau > t \end{cases} \tag{3.32}$$

$$K_u(t,\tau) = \frac{1}{C_{[a,b]}(t)} \begin{cases} K_{Fu}(t,\tau) & \text{for } \tau \le t \\ K_{Bu}(t,\tau) & \text{for } \tau > t \end{cases} \tag{3.33}$$

where $C_{[a,b]}(t) = [(t-a)^3 + (b-t)^3]$.

This delivers the formula for the "double-sided" kernel that effectively combines the operations of forward integration on $[a,t]$ and backward integration on $[b,t]$.

The recursive expressions for the derivatives of the output (3.7) can be derived by proceeding similarly as when deriving the K_y and K_u. To obtain the expression for $y^{(1)}$ the equations (3.8) & (3.9) need to be integrated two times :

$$\begin{aligned}
(t-a)^3 y^{(1)}(t) &= 6(t-a)^2 y(t) - a_2(t)(t-a)^3 y(t) \\
&+ \int_a^t \big[-18(\tau-a) + 6a_2(\tau)(\tau-a)^2 + 2a_2^{(1)}(\tau)(\tau-a)^3 - a_1(\tau)(\tau-a)^3\big] y(\tau)\,\mathrm{d}\tau \\
&+ \int_a^t (t-\tau)\big[6 - 6a_2(\tau)(\tau-a) - 6a_2^{(1)}(\tau)(\tau-a)^2 - a_2^{(2)}(\tau)(\tau-a)^3 + 3a_1(\tau)(\tau-a)^2 \\
&\qquad + a_1^{(1)}(\tau)(\tau-a)^3 - a_0(\tau)(\tau-a)^3\big] y(\tau)\,\mathrm{d}\tau - b_2(t)(t-a)^3 u(t)
\end{aligned}$$

$$
\begin{aligned}
&+ \int_a^t \Big[6b_2(\tau)(\tau-a)^2 + 2b_2^{(1)}(\tau)(\tau-a)^3 - b_1(\tau)(\tau-a)^3\Big] u(\tau)\,\mathrm{d}\tau \\
&+ \int_a^t (t-\tau)\Big[-6b_2(\tau)(\tau-a) - 6b_2^{(1)}(\tau)(\tau-a)^2 - b_2^{(2)}(\tau)(\tau-a)^3 + 3b_1(\tau)(\tau-a)^2 \\
&\qquad + b_1^{(1)}(\tau)(\tau-a)^3 - b_0(\tau)(\tau-a)^3\Big] u(\tau)\,\mathrm{d}\tau
\end{aligned} \tag{3.34}
$$

and

$$
\begin{aligned}
&(b-t)^3 y^{(1)}(t) = -6(b-t)^2 y(t) - a_2(t)(b-t)^3 y(t) \\
&+ \int_t^b \Big[18(b-\tau) - 6a_2(\tau)(b-\tau)^2 + 2a_2^{(1)}(\tau)(b-\tau)^3 + a_1(\tau)(b-\tau)^3\Big] y(\tau)\,\mathrm{d}\tau \\
&+ \int_t^b (t-\tau)\Big[6 + 6a_2(\tau)(b-\tau) - 6a_2^{(1)}(\tau)(b-\tau)^2 + a_2^{(2)}(\tau)(b-\tau)^3 + 3a_1(\tau)(b-\tau)^2 \\
&\qquad - a_1^{(1)}(\tau)(b-\tau)^3 + a_0(\tau)(b-\tau)^3\Big] y(\tau)\,\mathrm{d}\tau - b_2(t)(b-t)^3 u(t) \\
&+ \int_t^b \Big[-6b_2(\tau)(b-\tau)^2 + 2b_2^{(1)}(\tau)(b-\tau)^3 + b_1(\tau)(b-\tau)^3\Big] u(\tau)\,\mathrm{d}\tau \\
&+ \int_t^b (t-\tau)\Big[6b_2(\tau)(b-\tau) - 6b_2^{(1)}(\tau)(b-\tau)^2 + b_2^{(2)}(\tau)(b-\tau)^3 + 3b_1(\tau)(b-\tau)^2 \\
&\qquad - b_1^{(1)}(\tau)(b-\tau)^3 + b_0(\tau)(b-\tau)^3\Big] u(\tau)\,\mathrm{d}\tau
\end{aligned} \tag{3.35}
$$

The final expression for $y^{(1)}(t)$ is obtained by adding the results of (3.34) and (3.35) while dividing by $[(t-a)^3 + (b-t)^3]$.

$$
\begin{aligned}
\Big[(t-a)^3 + (b-t)^3\Big] y^{(1)}(t) &= \Big[6(t-a)^2 - 6(b-t)^2 - [(t-a)^3 + (b-t)^3]a_2(t)\Big] y(t) \\
&\quad - \Big[(t-a)^3 + (b-t)^3\Big] b_2(t) u(t) \\
&\quad + \int_a^t K_{Fy}^1(t,\xi) y(\xi) d\xi + \int_a^t K_{Fu}^1(t,\xi) u(\xi) d\xi \\
&\quad + \int_a^t K_{By}^1(t,\xi) y(\xi) d\xi + \int_a^t K_{Bu}^1(t,\xi) u(\xi) d\xi
\end{aligned} \tag{3.36}
$$

where the kernel functions K^1_{Fy} and K^1_{Fu} are:

$$\begin{aligned}
K^1_{Fy}(t,\tau) &= \big[-18(\tau-a)+6a_2(\tau)(\tau-a)^2+2a_2^{(1)}(\tau)(\tau-a)^3-a_1(\tau)(\tau-a)^3\big]\\
&\quad+(t-\tau)\big[6-6a_2(\tau)(\tau-a)-6a_2^{(1)}(\tau)(\tau-a)^2-a_2^{(2)}(\tau)(\tau-a)^3+3a_1(\tau)(\tau-a)^2\\
&\qquad+a_1^{(1)}(\tau)(\tau-a)^3-a_0(\tau)(\tau-a)^3\big] \qquad (3.37)\\
K^1_{Fu}(t,\tau) &= \big[-6b_2(\tau)(b-\tau)^2+2b_2^{(1)}(\tau)(b-\tau)^3+b_1(\tau)(b-\tau)^3\big]\\
&\quad+(t-\tau)\big[-6b_2(\tau)(\tau-a)-6b_2^{(1)}(\tau)(\tau-a)^2-b_2^{(2)}(\tau)(\tau-a)^3+3b_1(\tau)(\tau-a)^2\\
&\qquad+b_1^{(1)}(\tau)(\tau-a)^3-b_0(\tau)(\tau-a)^3\big] \qquad (3.38)
\end{aligned}$$

and the kernel functions K^1_{By} and K^1_{Bu} are:

$$\begin{aligned}
K^1_{By}(t,\tau) &= \big[18(b-\tau)-6a_2(\tau)(b-\tau)^2+2a_2^{(1)}(\tau)(b-\tau)^3+a_1(\tau)(b-\tau)^3\big]\\
&\quad+(t-\tau)\big[6+6a_2(\tau)(b-\tau)-6a_2^{(1)}(\tau)(b-\tau)^2+a_2^{(2)}(\tau)(b-\tau)^3+3a_1(\tau)(b-\tau)^2\\
&\qquad-a_1^{(1)}(\tau)(b-\tau)^3+a_0(\tau)(b-\tau)^3\big] \qquad (3.39)\\
K^1_{Bu}(t,\tau) &= \big[-6b_2(\tau)(b-\tau)^2+2b_2^{(1)}(\tau)(b-\tau)^3+b_1(\tau)(b-\tau)^3\big]\\
&\quad+(t-\tau)\big[6b_2(\tau)(b-\tau)-6b_2^{(1)}(\tau)(b-\tau)^2+b_2^{(2)}(\tau)(b-\tau)^3+3b_1(\tau)(b-\tau)^2\\
&\qquad-b_1^{(1)}(\tau)(b-\tau)^3+b_0(\tau)(b-\tau)^3\big] \qquad (3.40)
\end{aligned}$$

To obtain a formula for $y^{(2)}(t)$, equations (3.8) & (3.9) need to be integrated only once :

$$\begin{aligned}
(t-a)^3y^{(2)}(t) &= 3(t-a)^2y^{(1)}(t)-a_2(t)(t-a)^3y^{(1)}(t)\\
&\quad+\big[-6(t-a)+3a_2(t)(t-a)^2+a_2^{(1)}(t)(t-a)^3-a_1(t)(t-a)^3\big]y(t)\\
&\quad+\int_a^t\big[6-6a_2(\tau)(\tau-a)-6a_2^{(1)}(\tau)(\tau-a)^2-a_2^{(2)}(\tau)(\tau-a)^3+3a_1(\tau)(\tau-a)^2\\
&\qquad+a_1^{(1)}(\tau)(\tau-a)^3-a_0(\tau)(\tau-a)^3\big]y(\tau)\,\mathrm{d}\tau-b_2(t)(t-a)^3u^{(1)}(t)\\
&\quad+\big[3b_2(t)(t-a)^2+b_2^{(1)}(t)(t-a)^3-b_1(t)(t-a)^3\big]u(t)\\
&\quad+\int_a^t\big[-6b_2(\tau)(\tau-a)-6b_2^{(1)}(\tau)(\tau-a)^2-b_2^{(2)}(\tau)(\tau-a)^3+3b_1(\tau)(\tau-a)^2\\
&\qquad+b_1^{(1)}(\tau)(\tau-a)^3-b_0(\tau)(\tau-a)^3\big]u(\tau)\,\mathrm{d}\tau \qquad (3.41)
\end{aligned}$$

and

$$
\begin{aligned}
(b-t)^3 y^{(2)}(t) &= -3(b-t)^2 y^{(1)}(t) - a_2(t)(b-t)^3 y^{(1)}(t) \\
&+ \big[-6(b-t) - 3a_2(t)(b-t)^2 + a_2^{(1)}(t)(b-t)^3 - a_1(t)(b-t)^3 \big] y(t) \\
&+ \int_t^b \big[6 + 6a_2(\tau)(b-\tau) - 6a_2^{(1)}(\tau)(b-\tau)^2 + a_2^{(2)}(\tau)(b-\tau)^3 + 3a_1(\tau)(b-\tau)^2 \\
&\qquad - a_1^{(1)}(\tau)(b-\tau)^3 + a_0(\tau)(b-\tau)^3 \big] y(\tau)\,\mathrm{d}\tau - b_2(t)(b-t)^3 u^{(1)}(t) \\
&+ \big[-3b_2(t)(b-t)^2 + b_2^{(1)}(t)(b-t)^3 - b_1(t)(b-t)^3 \big] u(t) \\
&+ \int_t^b \big[6b_2(\tau)(b-\tau) - 6b_2^{(1)}(\tau)(b-\tau)^2 + b_2^{(2)}(\tau)(b-\tau)^3 + 3b_1(\tau)(b-\tau)^2 \\
&\qquad - b_1^{(1)}(\tau)(b-\tau)^3 + b_0(\tau)(b-\tau)^3 \big] u(\tau)\,\mathrm{d}\tau
\end{aligned}
\tag{3.42}
$$

The expression for $y^{(2)}(t)$ is obtained by adding (3.41) and (3.42) while dividing by the factor $[(t-a)^3 + (b-t)^3]$.

$$
\begin{aligned}
\Big[(t-a)^3 + (b-t)^3 \Big] y^{(2)}(t) &= \Big[3(t-a)^2 - 3(b-t)^2 - [(t-a)^3 + (b-t)^3] a_2(t) \Big] y^{(1)}(t) \\
&+ \Big[-6(t-a) + 3a_2(t)(t-a)^2 + a_2^{(1)}(t)(t-a)^3 - a_1(t)(t-a)^3 \\
&- 6(b-t) - 3a_2(t)(b-t)^2 + a_2^{(1)}(t)(b-t)^3 - a_1(t)(b-t)^3 \Big] y(t) \\
&- b_2(t) \Big[(t-a)^3 + (b-t)^3 \Big] u^{(1)}(t) \\
&+ \Big[3b_2(t)(t-a)^2 + b_2^{(1)}(t)(t-a)^3 - b_1(t)(t-a)^3 \\
&- 3b_2(t)(b-t)^2 + b_2^{(1)}(t)(b-t)^3 - b_1(t)(b-t)^3 \Big] u(t) \\
&+ \int_a^t K_{Fy}^2(t,\xi) y(\xi) d\xi + \int_a^t K_{Fu}^2(t,\xi) u(\xi) d\xi \\
&+ \int_a^t K_{By}^2(t,\xi) y(\xi) d\xi + \int_a^t K_{Bu}^2(t,\xi) u(\xi) d\xi
\end{aligned}
\tag{3.43}
$$

where the kernel functions K^2_{Fy} and K^2_{Fu} are:

$$K^2_{Fy}(\tau) = \big[6 - 6a_2(\tau)(\tau - a) - 6a_2^{(1)}(\tau)(\tau - a)^2 - a_2^{(2)}(\tau)(\tau - a)^3 + 3a_1(\tau)(\tau - a)^2 + a_1^{(1)}(\tau)(\tau - a)^3 - a_0(\tau)(\tau - a)^3\big] \tag{3.44}$$

$$K^2_{Fu}(\tau) = \big[- 6b_2(\tau)(\tau - a) - 6b_2^{(1)}(\tau)(\tau - a)^2 - b_2^{(2)}(\tau)(\tau - a)^3 + 3b_1(\tau)(\tau - a)^2 + b_1^{(1)}(\tau)(\tau - a)^3 - b_0(\tau)(\tau - a)^3\big] \tag{3.45}$$

and the kernel functions K^1_{By} and K^1_{Bu} are:

$$K^2_{By}(\tau) = \big[6 + 6a_2(\tau)(b - \tau) - 6a_2^{(1)}(\tau)(b - \tau)^2 + a_2^{(2)}(\tau)(b - \tau)^3 + 3a_1(\tau)(b - \tau)^2 - a_1^{(1)}(\tau)(b - \tau)^3 + a_0(\tau)(b - \tau)^3\big] \tag{3.46}$$

$$K^2_{Bu}(\tau) = \big[6b_2(\tau)(b - \tau) - 6b_2^{(1)}(\tau)(b - \tau)^2 + b_2^{(2)}(\tau)(b - \tau)^3 + 3b_1(\tau)(b - \tau)^2 - b_1^{(1)}(\tau)(b - \tau)^3 + b_0(\tau)(b - \tau)^3\big] \tag{3.47}$$

Note that, in general, the kernel functions K^i_{Fy}, K^i_{Fu}, K^i_{By} and K^i_{Bu} would have terms involving t, except when $i = n - 1$.
Refer to [15] for the derivation of the kernels for a second order linear time-varying system which proceeds in a similar fashion.

3.3 Example

To validate the double-sided kernel derived for the third order LTV system we use the following system as an example.
System

$$\begin{aligned}
\dot{x}_1(t) &= x_2(t) \\
\dot{x}_2(t) &= x_3(t) \\
\dot{x}_3(t) &= -(t^2 + 1)x_3(t) - (t^2 + 1)x_2(t) - (t^2 + t)x_1(t) + \frac{1}{t + 1000}u(t)
\end{aligned}$$

with $y = x_1$ as the measured output, x_1, x_2 and x_3 as the states, and with $u = 120\sin(t)$ as the control input.

The controlled system invariance is:

$$y^{(3)}(t) + (t^2+1)y^{(2)}(t) + (t^2+1)y^{(1)}(t) + (t^2+t)y(t) - \frac{1}{t+1000}u(t) = 0 \tag{3.48}$$

Validation by State Estimation

The estimates of the system output and its derivatives are obtained by evaluating the double-sided kernel using the measured system output with known system parameters which are shown in Figures 3.1, 3.2, and 3.3. The measured output signal is defined over a horizon of 5 seconds with 1000 sample points for every second, yielding 5000 sample measurements.

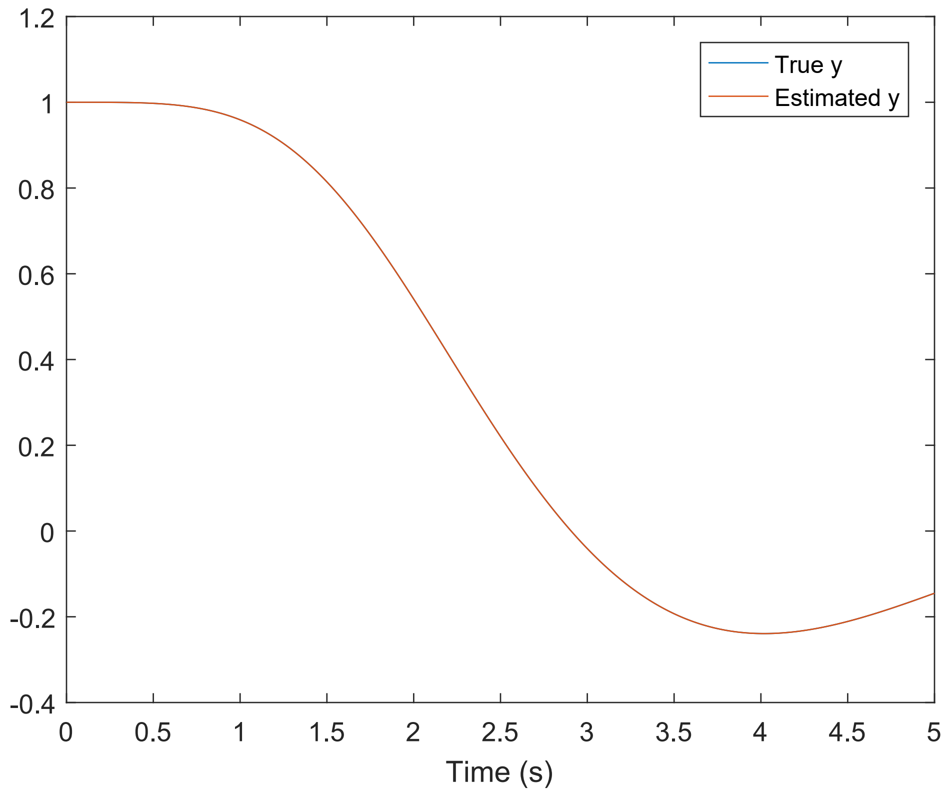

Figure 3.1 True $y(t)$ vs Estimated $y(t)$

The Figure 3.1 shows that the estimated $y(t)$ exactly coincides with the true $y(t)$ thereby validating the derived Double sided Kernel for third order system.

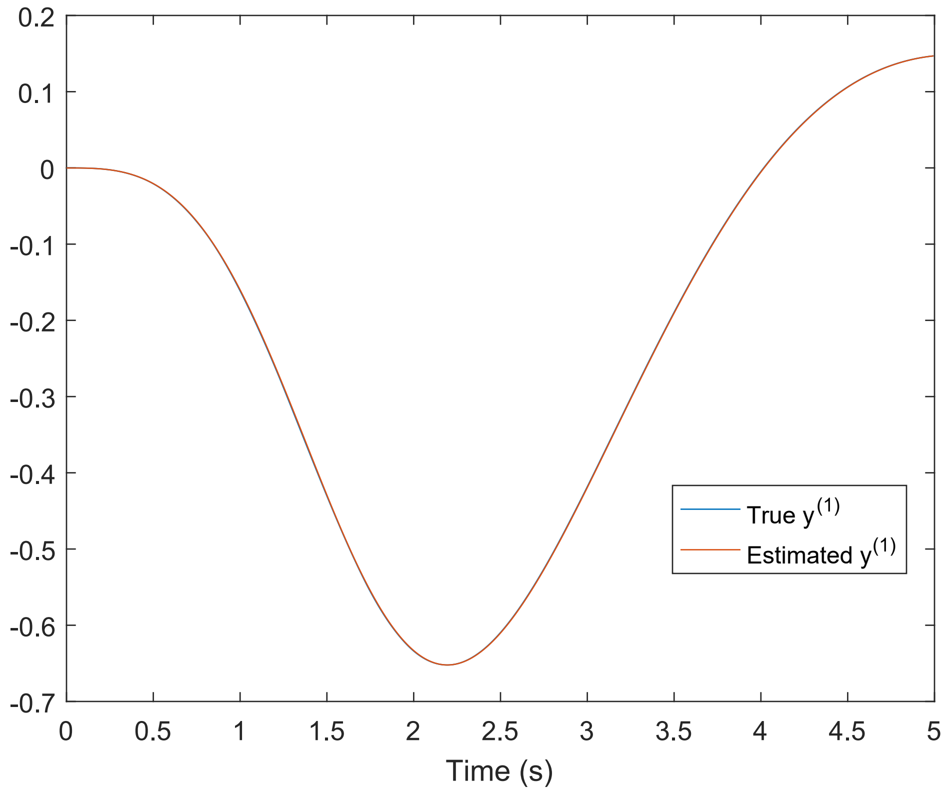

Figure 3.2 True $y^{(1)}(t)$ vs Estimated $y^{(1)}(t)$

The Figure 3.2 shows that the kernel expressions derived for $y^{(1)}(t)$ are accurate.

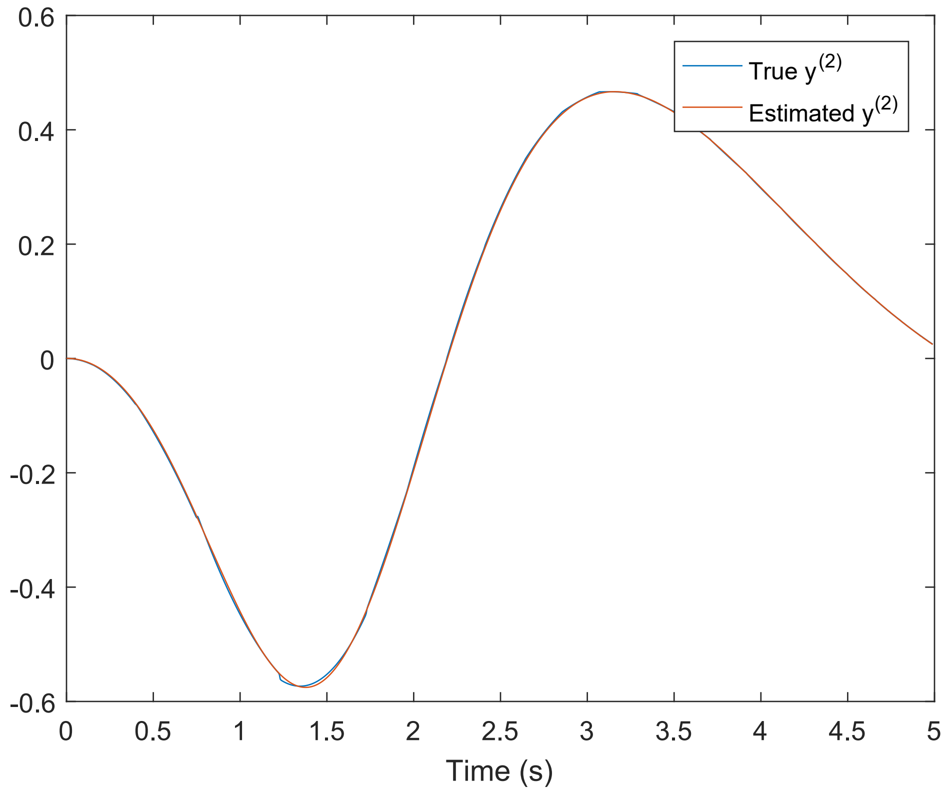

Figure 3.3 True $y^{(2)}(t)$ vs Estimated $y^{(2)}(t)$

The Figure 3.3 shows that the expressions derived for $y^{(2)}(t)$ are also correct. The small mismatch between 1 and 1.5 seconds is due to the error in numerical integration which can be ignored.

Chapter 4

State and Parameter Estimation

4.1 Parameter Estimation for LTV Systems

For the purpose of estimation, the system parameters $a_0, \cdots, a_{n-1}$ and $b_0, \cdots, b_{n-1}$, are first expressed by their respective approximations in the B-spline bases:

$$a_j(t) = \sum_{i \in I_j^a} \alpha_i^j B_i(t), \; b_j(t) = \sum_{i \in I_j^b} \gamma_i^j B_i(t), \quad j = 0, \cdots, n-1 \tag{4.1}$$

Their derivatives are then also linear combinations of B-splines by virtue of the differentiation formula (2.3) - (2.4). This further implies that the kernels in the system representation of Theorem 3, which are themselves linear functions of these parameters and their derivatives, then appear in the form of linear combinations involving the ensemble of the coefficients in the B-spline approximations (4.1). More precisely, there exist functions $g_{y,k}, k \in \mathcal{S}_y$ and $g_{u,j}, j \in \mathcal{S}_u$, such that the kernels of the system representation can be expressed in the following form

$$\begin{aligned} K_y(t,\tau) &= \sum_{k \in \mathcal{S}_y} \beta_k g_{y,k}(t,\tau) B_k(\tau) \\ K_u(t,\tau) &= \sum_{j \in \mathcal{S}_u} \beta_j g_{u,j}(t,\tau) B_j(\tau) \end{aligned} \tag{4.2}$$

where the index sets $\mathcal{S}_y$, $\mathcal{S}_u$ contain the ensemble of indices of B-splines that appear in the B-spline approximations of the system parameters $a_0, \cdots, a_{n-1}$ and $b_0, \cdots, b_{n-1}$, respectively.

It further follows that the evaluation functional (3.6) of Theorem 3 can be written as

$$\begin{aligned} y(t) = & \sum_{k \in \mathcal{S}_y} \beta_k \int_a^b g_{y,k}(t,\tau) B_k(\tau) y(\tau) d\tau \\ & + \sum_{j \in \mathcal{S}_u} \beta_j \int_a^b g_{u,j}(t,\tau) B_j(\tau) u(\tau) d\tau \end{aligned} \tag{4.3}$$

Defining

$$\begin{aligned} h_k(y,t) &:= \int_a^b g_{y,k}(t,\tau) B_k(\tau) y(\tau) d\tau; \quad k \in \mathcal{S}_y \\ h_j(u,t) &:= \int_a^b g_{u,j}(t,\tau) B_j(\tau) u(\tau) d\tau; \quad j \in \mathcal{S}_u \end{aligned} \tag{4.4}$$

and stacking the vectors

$$\begin{aligned} &\beta := [\beta_k; k \in \mathcal{S}_y \mid \beta_j; j \in \mathcal{S}_u]^T; \\ &h(y,u,t) := [h_k(y,t); k \in \mathcal{S}_y \mid h_j(u,t); j \in \mathcal{S}_u] \end{aligned} \tag{4.5}$$

gives $\beta, h(y,u,t) \in \mathbb{R}^s$; $s := \text{card}\{\mathcal{S}_u \cup \mathcal{S}_y\}$ and

$$y(t) = h(y,u,t)\beta \tag{4.6}$$

Let the output y in (4.6) be the measured output y_M . Given distinct time instants $t_1, \cdots, t_m \in (a,b]$, $m \geq 2n$, the last equation is now re-written point-wise as a system of linear algebraic equations

$$q(y_M) = P(y_M,u)\beta \tag{4.7}$$

$$q(y_M) \stackrel{def}{=} \begin{bmatrix} y_M(t_1) \\ \vdots \\ y_M(t_m) \end{bmatrix}; P(y_M,u) \stackrel{def}{=} \begin{bmatrix} h(y_M,u,t_1) \\ \ddots \\ h(y_M,u,t_m) \end{bmatrix} \tag{4.8}$$

where, clearly, $q \in \mathbb{R}^m$ and $P \in \mathbb{R}^m \times \mathbb{R}^s$ while $\beta \in \mathbb{R}^s$.

As no assumptions are made about the noise which may determine the invertibility of the matrix P, the following practical definition of linear identifiability is introduced

4.1.1 Practical Linear Identifiability

Definition 1. The linear time-varying system (3.1) is practically linearly identifiable on $[a, b]$ with respect to a particular realization of the output measurement, $y_M(t), t \in [a, b]$ corresponding to a known input u , if there exist distinct time instants $t_1, \cdots, t_m \in (a, b]$ such that the matrix $P(y_M, u)$ has rank $2n$.

By analogy with the nomenclature used in [4] output trajectories which render the satisfaction of the condition $\text{rank} P(y_M, u) = 2n$ are called *persistent*.

The estimation of the β coefficient vector that deliver the best fitting B-splines approximations of the sytem parameters is best performed solving (4.7) in terms of the pseudo-inverse $P^\dagger$ of P:

$$\beta = P^\dagger(y_M, u)\ q(y_M) \tag{4.9}$$

The coefficients β so obtained are then used in (4.1) to deliver the estimates of the time-varying system parameters.

4.2 State Estimation for LTV Systems

4.2.1 A Non-Asymptotic Observer for LTV Systems

The results in Theorem 2 readily deliver a state estimator for system (3.1) on an arbitrary window $[a, b]$. Given a measured output $y_M : [a, b] \to \mathbb{R}$ in response to a known system input $u_K : [a, b] \to \mathbb{R}$ the estimator equations are:

$$y_E(t) = \int_a^b K_y(t, \tau) y_M(\tau)\ d\tau + \int_a^b K_u(t, \tau) u_K(\tau)\ d\tau$$

$$y_E^{(i)}(t) = \sum_{k=0}^{i-1} f_y^k(t) y_E^{(k)}(t) + \sum_{k=0}^{i-1} f_u^k(t) u_K^{(k)}(t) \tag{4.10}$$
$$+ \int_a^b K_y^i(t,\tau) y_M(\tau)\, d\tau + \int_a^b K_u^i(t,\tau) u_K(\tau)\, d\tau$$

$$x_E(t) = \begin{pmatrix} \Gamma_0 \\ \vdots \\ \Gamma_{n-1} \end{pmatrix}^{-1} \left[\begin{pmatrix} y_E(t) \\ \vdots \\ y_E^{(n-1)}(t) \end{pmatrix} - M \begin{pmatrix} u_K(t) \\ \vdots \\ u_K^{(n-2)}(t) \end{pmatrix} \right]$$

where the input derivatives : $u_K^{(k)} : [a,b] \to \mathbb{R}$ are also considered known, and where $y_E, y_E^{(i)}; i = 1, \ldots n-1$, and x_E are the estimated outputs, output derivatives, and the estimated states, over the estimation window $[a,b]$, respectively.

4.3 Input Estimation

The Input signal is estimated in a similar way as discussed in Section 4.1. The input signal is approximated as a B-spline functional evaluation and combined with the double sided kernel to form a system linear algebraic equations as shown in (4.8) . The coefficients of the B-splines are obtained by solving the system of linear equations. Again, the estimation is possible only if the linear time-varying system is practically linearly identifiable as described in the Section 4.1.1 but with unknown input.

4.4 Joint State, Parameter, and Input Estimation

Theorem 4 *The kernels* K_y *and* K_u *of* (3.6) *are linear functions of the system parameters* $a_0(t), a_1(t), .., a_{n-1}(t)$, $b_0(t), b_1(t), .., b_{n-1}(t)$, *and their time derivatives up to order* $n-1$. *The evaluation functional* (3.6) *is an equivalent LTV system representation in the following sense: any trajectories* $y(\tau)$ *and* $u(\tau)$, *for* $\tau \in [a,b]$ *satisfy the input-output system representation* (3.3) *if and only if they satisfy the integral representation* (3.6). *Clearly, the boundary conditions for* (3.3) *play no role in this equivalence.*

□

Although the kernels in the system representation depend linearly on the system parameters (Theorem 4), the parameters b_i, $i = 0, ..., m$, appear in products with the input function and its derivatives. This follows directly from the original input-output system equation (3.3) in which the right hand side is not jointly linear in $b_0, ..., b_m$ and $u, u^{(1)}, ..., u^{(m)}$ that precludes their simultaneous identifiability based solely on the system output data. To carry out linear identification of all system parameters and also estimate the system input the following two stage procedure is, however, viable:

Stage (1): select a test input such as a monic polynomial of order m, apply it to the system recording its output trajectory; use the recorded test data to estimate system parameters $a_0(t), ..., a_n(t)$, $b_0(t), ..., b_m(t)$, $t \in [a, b]$ employing the algorithm outlined in [3].

Stage (2): acquire the system output data in response to the input signal to be estimated on the time interval $[a, b]$; use this output data and the B-spline approximations of the $b_0(t), ..., b_m(t)$, $t \in [a, b]$, obtained during Stage (1) to re-estimate the parameters $a_0(t), ..., a_n(t)$, $t \in [a, b]$, cross-validating with those computed in Stage (1); simultaneously estimate the unknown system input and the full state vector of the system on $[a, b]$.

It should be clear that Stage (1) will be redundant if $m = 0$ on the right hand side of the system input-output representation in (3.3). In that case the system input can be recovered as the product $b_0(t)u(t)$, $t \in [a, b]$.

To begin with, the unknown system parameters $a_0, \cdots, a_{n-1}$ and the unknown system input u are first expressed by their respective approximations in the B-spline bases for $j = 0, \cdots, n-1$:

$$a_j(t) = \sum_{i \in I_j^a} \alpha_i^j B_i(t), \; u(t) = \sum_{i \in I^u} \gamma_i B_i(t) \tag{4.11}$$

Their derivatives are then also linear combinations of B-splines by virtue of their differentiation formula (2.3) - (2.4). This further implies that the kernels in the system representation of Theorem 3, which are themselves linear functions of these parameters and their derivatives, then appear in the form of linear combinations involving the ensemble of the coefficients in the B-spline approximations (4.11). More precisely, there exist functions $g_{y,k}, k \in \mathcal{S}_y$ and $g_{u,j}, j \in \mathcal{S}_u$, such that the system representation (3.6) of Theorem 3 can be written as

$$y(t) = \sum_{k\in\mathcal{S}_y} \beta_k \int_a^b g_{y,k}(t,\tau)B_k(\tau)y(\tau)d\tau + \sum_{j\in\mathcal{S}_u} \beta_j \int_a^b g_{u,j}(t,\tau)B_j(\tau)d\tau \tag{4.12}$$

where the index sets $\mathcal{S}_y$, $\mathcal{S}_u$ contain the ensemble of indices of B-splines that appear in the B-spline approximations of the system parameters $a_0, \cdots, a_{n-1}$ and u, respectively. In the above discussion there is no mention of the parameters $b_0, ..., b_m$ as these are considered known in Stage (2), i.e. are parts of the functions $g_{u,j}(t,\tau)$. Defining

$$\begin{aligned} h_k(y,t) &:= \int_a^b g_{y,k}(t,\tau)B_k(\tau)y(\tau)d\tau; \quad k\in\mathcal{S}_y \\ h_j(t) &:= \int_a^b g_{u,j}(t,\tau)B_j(\tau)d\tau; \quad j\in\mathcal{S}_u \end{aligned} \tag{4.13}$$

and stacking the vectors

$$\begin{aligned} \beta &:= [\beta_k; k\in\mathcal{S}_y \mid \beta_j; j\in\mathcal{S}_u]^T; \\ h(y,t) &:= [h_k(y,t); k\in\mathcal{S}_y \mid h_j(t); j\in\mathcal{S}_u] \end{aligned} \tag{4.14}$$

gives $\beta, h(y,t) \in \mathbb{R}^s; s := \text{card}\{\mathcal{S}_u \cup \mathcal{S}_y\}$ and

$$y(t) = h(y,t)\beta \tag{4.15}$$

Let the output y in (4.15) be the measured output y_M in response to the unknown input u. Given distinct time instants $t_1, \cdots, t_p \in (a,b]$, $p >> s$, the last equation is now re-written point-wise as a system of linear algebraic equations

$$q(y_M) = P(y_M)\beta \tag{4.16}$$

$$q(y_M) \stackrel{def}{=} \begin{bmatrix} y_M(t_1) \\ \vdots \\ y_M(t_m) \end{bmatrix}; P(y_M) \stackrel{def}{=} \begin{bmatrix} h(y_M,t_1) \\ \ddots \\ h(y_M,t_m) \end{bmatrix} \tag{4.17}$$

where, clearly, $q \in \mathbb{R}^m$ and $P \in \mathbb{R}^m \times \mathbb{R}^s$ while $\beta \in \mathbb{R}^s$.

As no assumptions are made about the noise which may determine the invertibility of the matrix P, the following practical definition of linear identifiability is introduced.

4.4.1 Practical Linear Identifiability

Definition 1. (Identifiability of $a_0, ..., a_n$ and input u)
The linear time-varying system (3.1) is practically linearly identifiable on $[a, b]$ with respect to a particular realization of the output measurement, $y_M(t), t \in [a, b]$ corresponding to an unknown input u , if there exist distinct time instants $t_1, \cdots, t_p \in (a, b]$ such that the matrix $P(y_M)$ has full column rank.

The least squares estimate of the β coefficient vector for the best fitting B-splines approximations of the sytem parameters and the system input signal is then delivered in terms of the pseudo-inverse $P^\dagger$ of P:

$$\beta = P^\dagger(y_M)q(y_M) = [P(y_M)^T P(y_M)]^{-1} P(y_M)^T q(y_M) \tag{4.18}$$

The coefficients β so obtained are then used in (4.11) to deliver the estimates of the time-varying system parameters and input signal.

4.4.2 A Multi-Task Non-Asymptotic Observer for LTV Systems

Non-asymptotic joint state, input, and parameter estimation is carried out by using the raw measurement data y_M to first obtain the estimates for the time-varying system parameters and system input signal employing the approach described above. The estimates can be validated in a number of ways prior to state reconstruction e.g. using prior information about the parameters (Stage (1)) or sentinel ideas; see e.g. [8]. The trusted estimates of the parameters and the estimates of the input signal, denoted by u_E, then deliver estimates of the output and its time derivatives by application of the formulae of Theorem 3, i.e :

$$y_E(t) = \int_a^b K_y(t, \tau) y_M(\tau)\ d\tau + \int_a^b K_u(t, \tau) u_E(\tau)\ d\tau$$

$$y_E^{(i)}(t) = \sum_{k=0}^{i-1} f_y^{i,k}(t) y_E^{(k)}(t) + \sum_{k=0}^{i-1} f_u^{i,k}(t) u_E^{(k)}(t) \tag{4.19}$$

$$+ \int_a^b K_y^i(t,\tau) y_M(\tau)\ d\tau + \int_a^b K_u^i(t,\tau) u_E(\tau)\ d\tau$$

$$x_E(t) = \begin{pmatrix} \Gamma_0 \\ \vdots \\ \Gamma_{n-1} \end{pmatrix}^{-1} \left[\begin{pmatrix} y_E(t) \\ \vdots \\ y_E^{(n-1)}(t) \end{pmatrix} - M \begin{pmatrix} u_E(t) \\ \vdots \\ u_E^{(n-2)}(t) \end{pmatrix} \right]$$

where the estimates of the input derivatives : $u_E^{(k)} : [a,b] \to \mathbb{R}$ are immediately obtained by way of the B-splines differentiation formulae applied to (4.11) (whose coefficients are known from u_E), and where $y_E, y_E^{(k)}; k = 1, \ldots n-1$, and x_E are the estimated outputs, output derivatives, and the estimated states, over the estimation window $[a,b]$, respectively.

Chapter 5

Results

In the previous two chapters, we have shown the development of a Double-Sided Kernel for LTV systems and proposed our new approach of using B-Spline functional approximations for the estimation of parameters, input as well as the state and its time derivatives. We devote this chapter towards presenting the results of the proposed theory with the help of various examples for second order and third order LTV systems with exogenous input.

Due to the inherent properties of B-splines, the approximation precision degrades at the ends of any given estimation interval. In the case when estimation is required over extended periods of time, good approximation precision can be maintained by employing a sliding estimation window. Therefore, the results are presented here on a smaller window where the B-spline estimation is accurate.

We consider four cases of a second order system to show that the proposed method works in disparate situations. We also consider four cases of a third order system to validate our method of estimation. All the examples shown in the following sections assume that the output is not distorted by noise. However, a separate section is dedicated to show the results of parameter and state estimation in the presence of noisy output. Joint state and parameter estimation of LTV systems with noisy measurements is outside the scope of this thesis. n is the number of B-splines used over the entire time horizon and p is the order of the B-spline for the results shown in the next sections. The value of n was chosen by experimenting over a range of values for n ranging from 5 to 100. The optimal value was chosen depending on the computational cost and the estimation accuracy. A value of $p = 3$ was chosen as the order of the B-splines for the results shown in the next section since the parameters are quadratic polynomials. In general, cubic B-splines ($p = 4$) works well for general polynomial functions but is computationally more expensive than using quadratic B-splines. In practice, n is chosen by evaluating over a range of values of n, however, in our experiments $n = 17$ works well for all cases. A non-uniform knot sequence with knots crowded towards the end of the estimation interval was used for the estimation of the parameters, input and states of the system. The motivation for the examples chosen in the next section is to encompass a wide class of LTV systems to show that the proposed method works in disparate situations. The examples chosen are a modification of the simplified model of a DC motor as used in [15].

5.1 Second Order System: Estimation

5.1.1 Case 1: One constant and one time-varying parameter

System Description

$$\begin{aligned}\dot{x}_1(t) &= x_2(t)\\ \dot{x}_2(t) &= -x_2(t) - (t^2+1)x_1(t) + \frac{1}{t+1000}u(t)\end{aligned}$$

with $y = x_1$ as the measured output, x_1 and x_2 as the states, and with $u = 120\sin(t)$ as the control input. The controlled system invariance is:

$$y^{(2)}(t) + y^{(1)}(t) + (t^2+1)y(t) - \frac{1}{t+1000}u(t) = 0 \tag{5.1}$$

The above system consists of one constant parameter and one time-varying parameter. The results were obtained using $n = 16$ and $p = 3$ (quadratic B-spline). The total length of the estimation window is $[0, 10s]$, but the results are shown on an inner window of $[2, 8s]$ where the B-spline estimation is considered to be reliable.

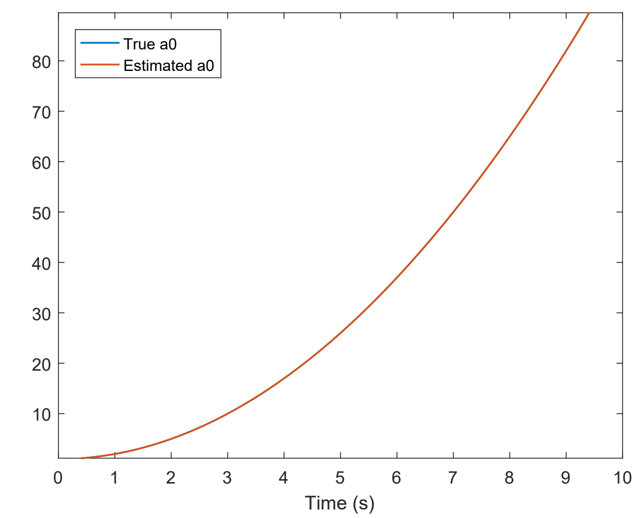

Figure 5.1 True $a_0(t)$ vs Estimated $a_0(t)$

The Figure 5.1 shows that the true value of $a_0(t)$ coincides with the estimated value. The estimated value of the constant a_1 is 1.007. True value of the constant a_1 is 1.

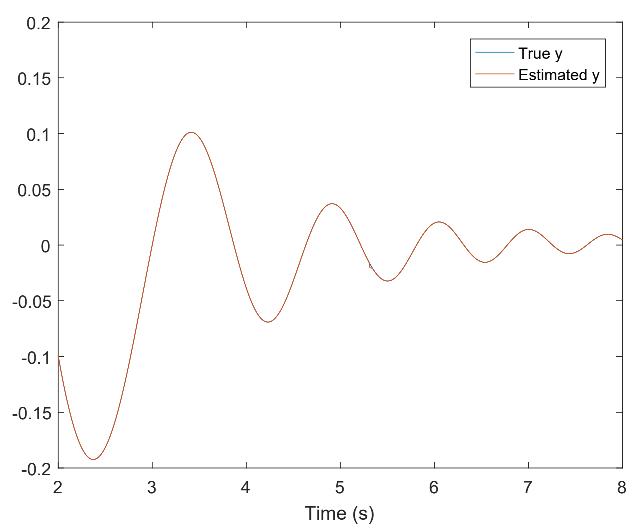

Figure 5.2 True $y(t)$ vs Estimated $y(t)$

Fig. 5.2 shows the estimated $y(t)$ exactly against the true $y(t)$. The small mismatch near $t = 5.5$ s is due to the numerical error in the integration of B-splines in the kernel representation of the system (5.1). Fig. 5.3 shows that the $y^{(1)}(t)$ estimate matches with the true value with high accuracy.

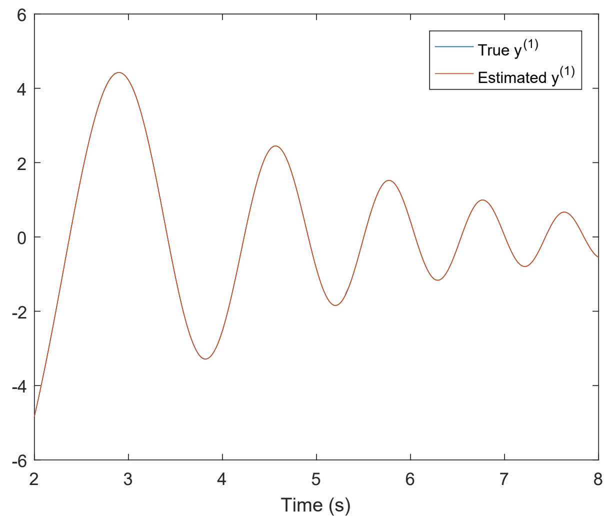

Figure 5.3 True $y^{(1)}(t)$ vs Estimated $y^{(1)}(t)$

5.1.2 Case 2: Two time-varying parameters

System Description

$$\dot{x}_1(t) = x_2(t)$$
$$\dot{x}_2(t) = -(t^2+1)x_2(t) - (t^2+1)x_1(t) + \frac{1}{t+1000}u(t)$$

with $y = x_1$ as the measured output, x_1 and x_2 as the states, and with $u = 120\sin(t)$ as the control input. The controlled system invariance is:

$$y^{(2)}(t) + (t^2+1)y^{(1)}(t) + (t^2+1)y(t) - \frac{1}{t+1000}u(t) = 0 \tag{5.2}$$

The example presented involves two unknown parameters. The parameter estimates obtained as linear combinations of B-splines deliver state estimates of remarkably high accuracy. The estimation method can use an estimation window of arbitrary length which can be dragged forward as necessary to deliver continued time varying estimates.

Parameter Estimation Equations

$$y_E(t) = \int_a^b K_y(t,\xi) y_M(\xi)\ d\xi + \int_a^b K_u(t,\xi) u_K(\xi)\ d\xi \tag{5.3}$$

Let $C_{[a,b]}(t) = [(t-a)^2 + (b-t)^2]$

$$\begin{aligned} y_E(t) &= \frac{1}{C_{[a,b]}(t)} \Big\{ \int_a^t K_{Fy}(t,\xi) y_M(\xi)\ d\xi + \int_t^b K_{By}(t,\xi) y_M(\xi)\ d\xi \\ &+ \int_a^t K_{Fu}(t,\xi) u_K(\xi)\ d\xi + \int_t^b K_{Bu}(t,\xi) u_K(\xi)\ d\xi \Big\} \\ C_{[a,b]}(t) y_E(t) &= \int_a^t K_{Fy}(t,\xi) y_M(\xi)\ d\xi + \int_t^b K_{By}(t,\xi) y_M(\xi)\ d\xi \\ &+ \int_a^t K_{Fu}(t,\xi) u_K(\xi)\ d\xi + \int_t^b K_{Bu}(t,\xi) u_K(\xi)\ d\xi \end{aligned} \tag{5.4}$$

Replacing the kernels with the actual expressions of the system described in (5.2)containing the B-Spline coefficients in equation (5.4) yields,

$$\begin{aligned} &[(t-a)^2 + (b-t)^2] y_E(t) \\ &= \sum_{k=1}^{n} \alpha_k \Big\{ \int_a^t \Big[-(\xi-a)^2 + 2(t-\xi)(\xi-a) \Big] y_E(\xi)\ d\xi \\ &\quad + \int_t^b \Big[(b-\xi)^2 + 2(t-\xi)(b-\xi) B_{k,p,t}(\xi) + (\xi-a)^2 B^{(1)}_{k,p,t}(\xi) \Big] y_E(\xi)\ d\xi \Big\} \\ &\quad + \sum_{k=1}^{n} \beta_k \Big\{ \int_a^t -(t-\xi)(\xi-a)^2 B_{k,p,t}(\xi) y_E(\xi)\ d\xi + \int_t^b (t-\xi)(b-\xi)^2 B_{k,p,t}(\xi) y_E(\xi)\ d\xi \Big\} \\ &\quad + \int_a^t (4(\xi-a) - 2(t-\xi)) y_E(\xi)\ d\xi + \int_t^b (4(b-\xi) + 2(t-\xi)) y_E(\xi)\ d\xi \\ &\quad + \int_a^t ((t-\xi)(\xi-a)^2/(t+1000)) y_E(\xi)\ d\xi - \int_t^b ((t-\xi)(b-\xi)^2/(t+1000)) y_E(\xi)\ d\xi \end{aligned}$$

Bringing the known terms to the left side, we have

$$\Big\{ [(t-a)^2 + (b-t)^2] y_E(t) - \int_a^t (4(\xi-a) - 2(t-\xi)) y_E(\xi)\ d\xi$$

$$
\begin{aligned}
&- \int_t^b (4(b-\xi) + 2(t-\xi))y_E(\xi)\ d\xi - \int_a^t ((t-\xi)(\xi-a)^2/(t+1000))y_E(\xi)\ d\xi \\
&+ \int_t^b ((t-\xi)(b-\xi)^2/(t+1000))y_E(\xi)\ d\xi\Big\} \\
&= \sum_{k=1}^{n} \alpha_k \Big\{ \int_a^t \Big[-(\xi-a)^2 + 2(t-\xi)(\xi-a)\Big] y_E(\xi)\ d\xi \\
&\qquad + \int_t^b \Big[(b-\xi)^2 + 2(t-\xi)(b-\xi)B_{k,p,t}(\xi) + (\xi-a)^2 B^{(1)}_{k,p,t}(\xi)\Big] y_E(\xi)\ d\xi \Big\} \\
&\qquad + \sum_{k=1}^{n} \beta_k \Big\{ \int_a^t -(t-\xi)(\xi-a)^2 B_{k,p,t}(\xi) y_E(\xi)\ d\xi + \int_t^b (t-\xi)(b-\xi)^2 B_{k,p,t}(\xi) y_E(\xi)\ d\xi \Big\}
\end{aligned}
$$

$$
\gamma(t, y_E(t)) = \sum_{k=1}^{n} \alpha_k P_{\alpha k}(t) + \sum_{k=1}^{n} \beta_k P_{\beta k}(t) \tag{5.5}
$$

with $P_{\alpha k}(t)$ as

$$
\begin{aligned}
P_{\alpha k}(t) = &\int_a^t \Big[-(\xi-a)^2 + 2(t-\xi)(\xi-a)\Big] y_E(\xi)\ d\xi \\
&+ \int_t^b \Big[(b-\xi)^2 + 2(t-\xi)(b-\xi)B_{k,p,t}(\xi) + (\xi-a)^2 B^{(1)}_{k,p,t}(\xi)\Big] y_E(\xi)\ d\xi
\end{aligned}
$$

and $P_{\beta k}(t)$ as

$$
P_{\beta k}(t) = \int_a^t -(t-\xi)(\xi-a)^2 B_{k,p,t}(\xi) y_E(\xi)\ d\xi + \int_t^b (t-\xi)(b-\xi)^2 B_{k,p,t}(\xi) y_E(\xi)\ d\xi
$$

Simplfying the equation (5.5) in matrix form gives,

$$
\gamma(t, y_E(t)) = \begin{bmatrix} P_{\alpha k}(t) & P_{\beta k}(t) \end{bmatrix} \begin{bmatrix} \alpha_k \\ \beta_k \end{bmatrix} \qquad k = 1, 2, ..., n \tag{5.6}
$$

Taking distinct time instants $t_1, \cdots, t_m \in (a, b]$, $m \geq 2n$, equation (5.6) is now re-written point-wise as a system of linear algebraic equations.

$$\begin{bmatrix} \gamma(t_1, y_E(t_1)) \\ \vdots \\ \gamma(t_m, y_E(t_m)) \end{bmatrix} ;= \begin{bmatrix} P_{\alpha 1}(t_1) \cdots & P_{\alpha n}(t_1) & P_{\beta 1}(t_1) \cdots & P_{\beta n}(t_1) \\ \vdots & \vdots & \vdots & \vdots \\ P_{\alpha 1}(t_m) \cdots & P_{\alpha n}(t_m) & P_{\beta 1}(t_m) \cdots & P_{\beta n}(t_m) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \\ \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} \tag{5.7}$$

The parameters are then estimated by solving the system of linear equations obtained in (5.7). The state is then estimated using equation (5.3) by substituting the parameters obtained from equation (5.7). The time derivative of the state is obtained similarly using the Kernel expression for $y^{(1)}(t)$.

The results were obtained using $n = 17$ and $p = 3$ (quadratic B-spline). The total length of the estimation window is $[0, 20s]$, but the results are shown on an inner window of $[3, 17s]$ where the B-spline estimation is considered to be reliable.

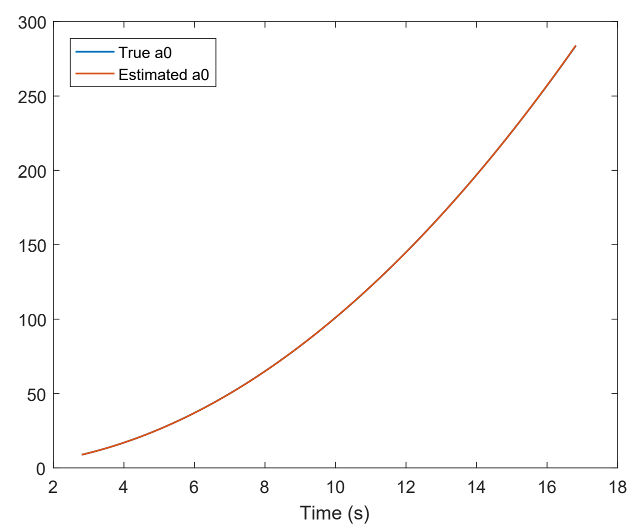

Figure 5.4 True $a_0(t)$ vs Estimated $a_0(t)$

Fig. 5.4 and Fig. 5.5 show that the estimates of the time-varying parameters.

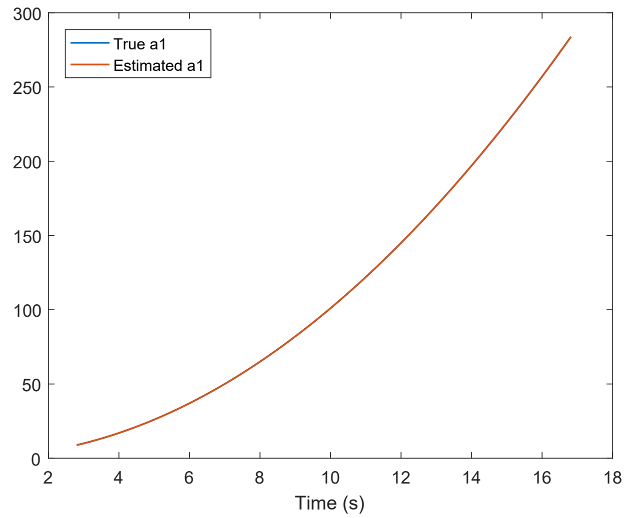

Figure 5.5 True $a_1(t)$ vs Estimated $a_1(t)$

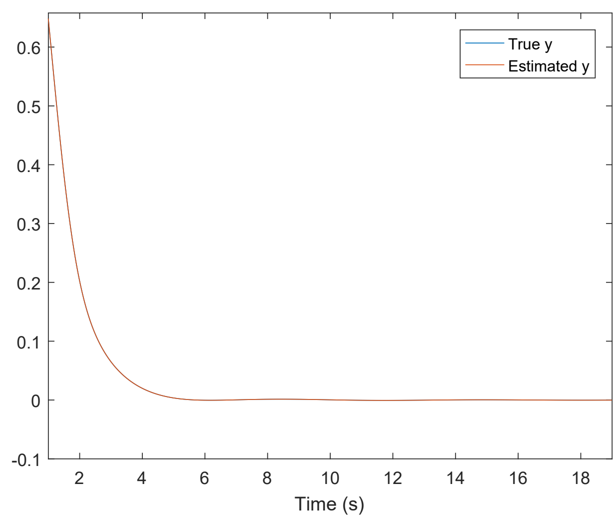

Figure 5.6 True $y(t)$ vs Estimated $y(t)$

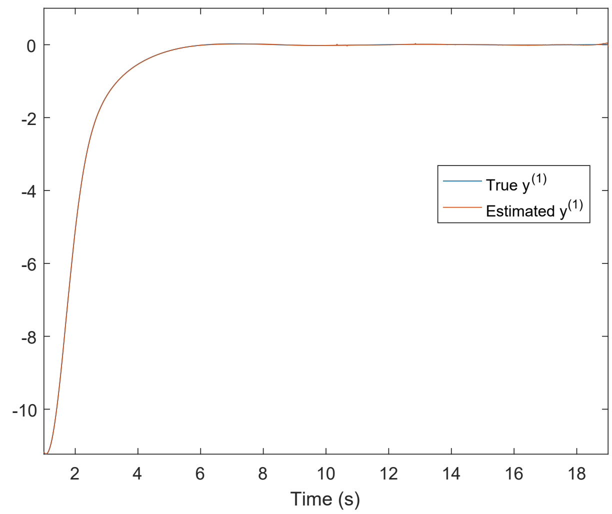

Figure 5.7 True $y^{(1)}(t)$ vs Estimated $y^{(1)}(t)$

Fig. 5.6 shows that the estimated $y(t)$ exactly coincides with the true $y(t)$ thereby validating the estimation method. Fig. 5.7 shows the estimated $y^{(1)}(t)$ against the true $y^{(1)}(t)$.

5.1.3 Case 3: Input

System Description

$$\dot{x}_1(t) = x_2(t)$$
$$\dot{x}_2(t) = -(t^2+1)x_2(t) - (t^2+1)x_1(t) - u(t)$$

with $y = x_1$ as the measured output, x_1 and x_2 as the states, and with $u = 120\sin(t)$ as the control input. The controlled system invariance is:

$$y^{(2)}(t) + (t^2+1)y^{(1)}(t) + (t^2+1)y(t) + u(t) = 0 \tag{5.8}$$

The results were obtained using $n = 17$ and $p = 3$ (quadratic B-spline). The total length of the estimation window is $[0, 5s]$, and the results are shown for the same window length.

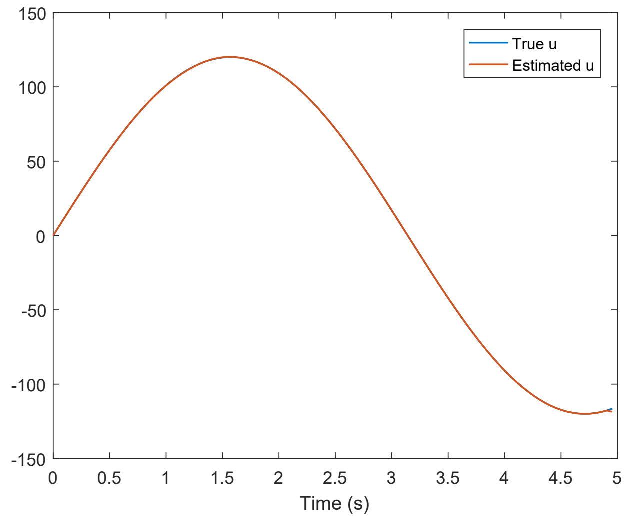

Figure 5.8 True $u(t)$ vs Estimated $u(t)$

Fig. 5.8 shows the estimated input $u(t) = 120\sin(t)$.

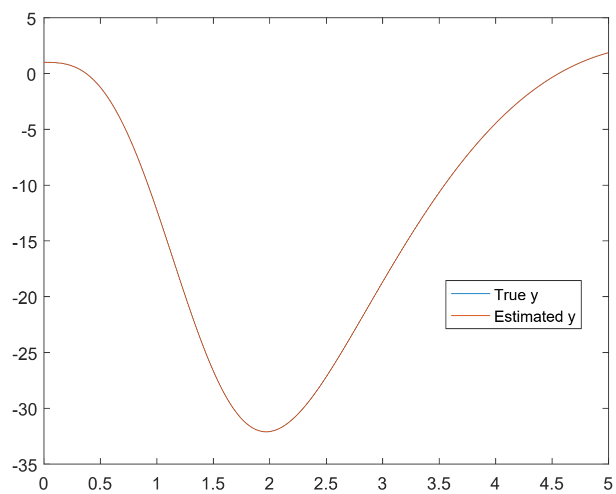

Figure 5.9 True $y(t)$ vs Estimated $y(t)$

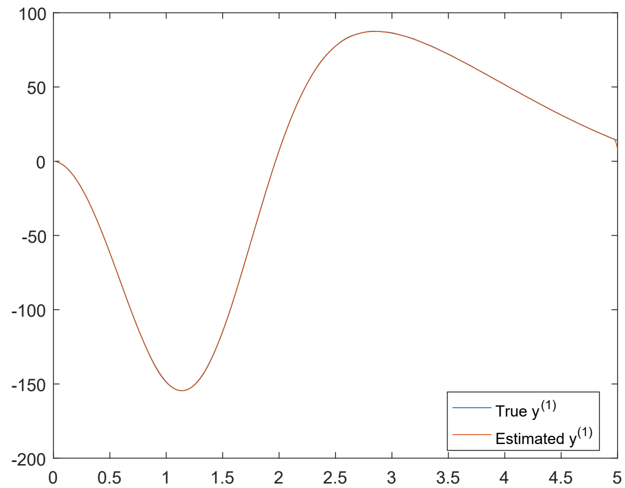

Figure 5.10 True $y^{(1)}(t)$ vs Estimated $y^{(1)}(t)$

Fig. 5.9 shows that the estimated $y(t)$ exactly coincides with the true $y(t)$ and Fig. 5.10 shows the estimated $y^{(1)}(t)$.

5.1.4 Case 4: Two time-varying parameters and Input

System Description

$$\dot{x}_1(t) = x_2(t)$$
$$\dot{x}_2(t) = -(t)x_2(t) - (t^2+1)x_1(t) + u(t)$$

with $y = x_1$ as the measured output, x_1 and x_2 as the states, and with $u = t^2$ as the control input. The controlled system invariance is:

$$y^{(2)}(t) + (t)y^{(1)}(t) + (t^2+1)y(t) + u(t) = 0 \tag{5.9}$$

The results were obtained using $n = 17$ and $p = 3$ (quadratic B-spline). The total length of the estimation window is $[0, 10s]$, but the results are shown on an inner window of $[1, 9s]$ where the B-spline estimation is considered to be reliable.

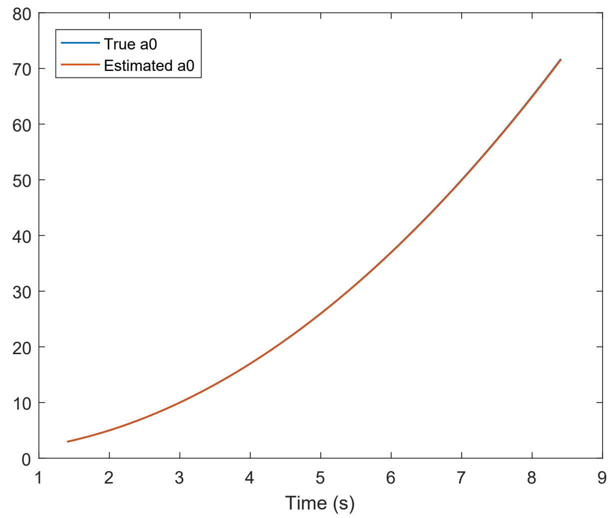

Figure 5.11 True $a_0(t)$ vs Estimated $a_0(t)$

Fig. 5.11 and Fig. 5.12 show that the estimates of the time-varying parameters of the system in (5.9).

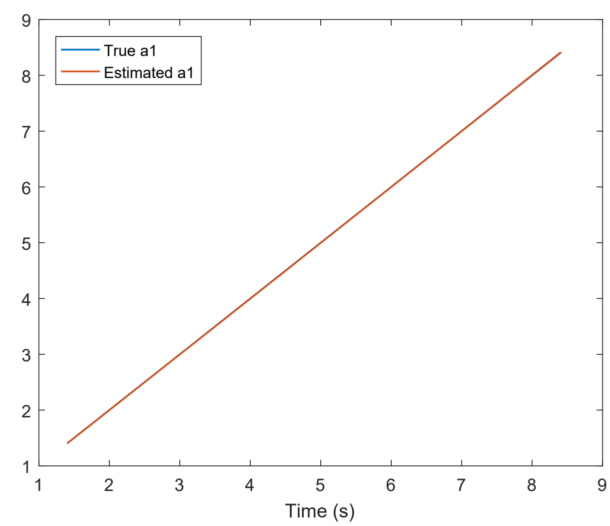

Figure 5.12 True $a_1(t)$ vs Estimated $a_1(t)$

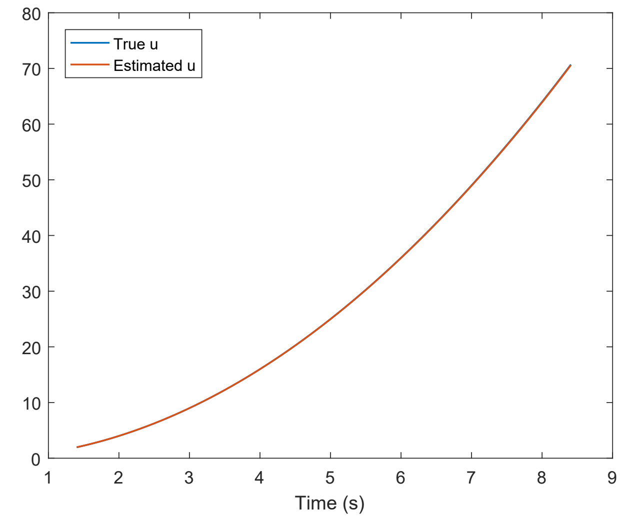

Figure 5.13 True $u(t)$ vs Estimated $u(t)$

Fig. 5.13 shows the estimated input against the true input $u(t) = t^2$.

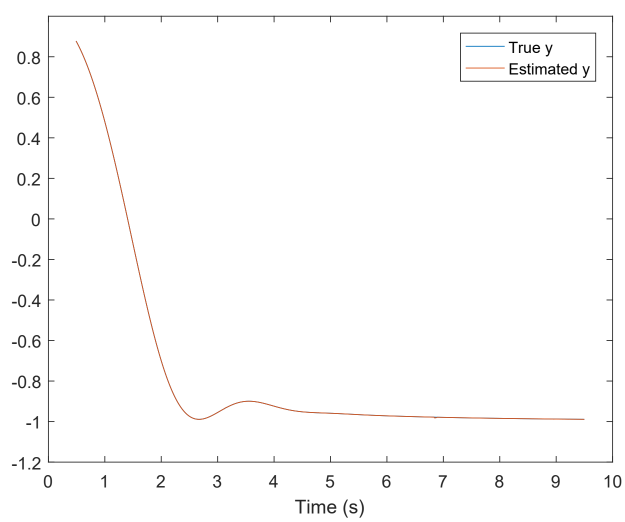

Figure 5.14 True $y(t)$ vs Estimated $y(t)$

Fig. 5.14 shows that the estimated $y(t)$ and Fig. 5.15 shows the estimated $y^{(1)}(t)$. The small mismatch near $t = 7$ s is due to the numerical error in the integration of B-splines in the kernel representation of the system (5.9).

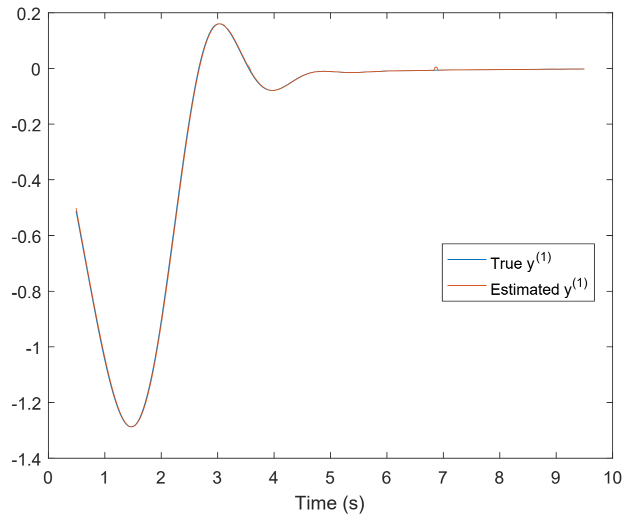

Figure 5.15 True $y^{(1)}(t)$ vs Estimated $y^{(1)}(t)$

5.2 Third Order System: Estimation

The following results were obtained using $n = 17$ and $p = 3$ (quadratic B-spline).

5.2.1 Case 1: One constant and Two time-varying parameters

System Description

$$\dot{x}_1(t) = x_2(t)$$
$$\dot{x}_2(t) = x_3(t)$$
$$\dot{x}_3(t) = -x_3(t) - (t)x_2(t) - (t^2+1)x_1(t) + \frac{1}{t+1000}u(t)$$

with $y = x_1$ as the measured output, x_1, x_2 and x_3 as the states, and with $u = 120\sin(t)$ as the control input. The controlled system invariance is:

$$y^{(3)}(t) + y^{(2)}(t) + (t)y^{(1)}(t) + (t^2+1)y(t) - \frac{1}{t+1000}u(t) = 0 \tag{5.10}$$

The total length of the estimation window is $[0, 10s]$, but the results are shown on an inner window of $[1, 9s]$ where the B-spline estimation is considered to be reliable.

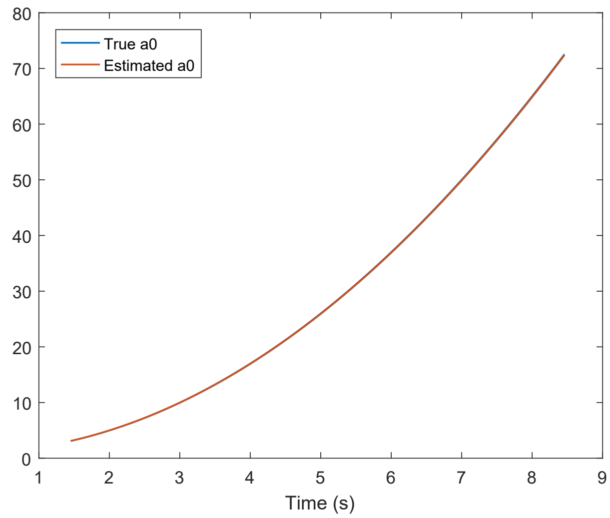

Figure 5.16 True $a_0(t)$ vs Estimated $a_0(t)$

Fig. 5.16 and Fig. 5.17 show the estimates for $a_0(t)$ and $a_1(t)$ of the LTV system in (5.10).

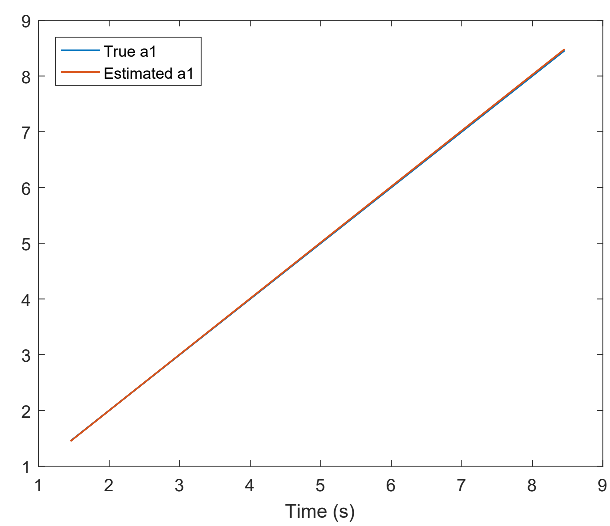

Figure 5.17 True $a_1(t)$ vs Estimated $a_1(t)$

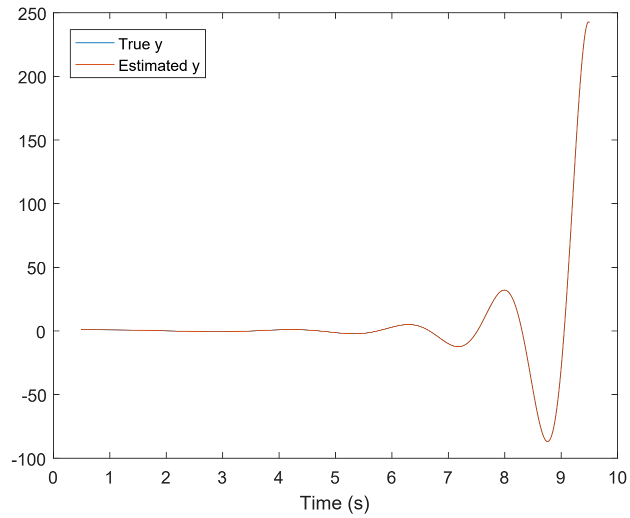

Figure 5.18 True $y(t)$ vs Estimated $y(t)$

Fig. 5.18 shows the estimated $y(t)$ and Fig. 5.19 shows the estimated $y^{(1)}(t)$.

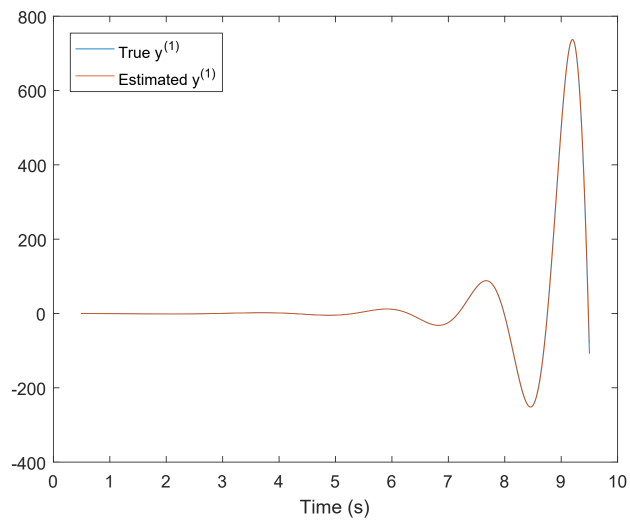

Figure 5.19 True $y^{(1)}(t)$ vs Estimated $y^{(1)}(t)$

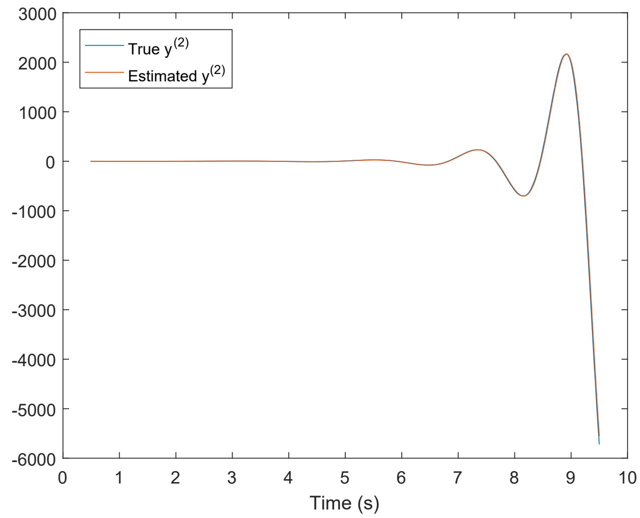

Figure 5.20 True $y^{(2)}(t)$ vs Estimated $y^{(2)}(t)$

Fig. 5.20 shows the estimated $y^{(2)}(t)$ against the true value.

5.2.2 Case 2: Three time-varying parameters

System Description

$$\dot{x}_1(t) = x_2(t)$$
$$\dot{x}_2(t) = x_3(t)$$
$$\dot{x}_3(t) = -(t)x_3(t) - (t^2+1)x_2(t) - (t^2+t)x_1(t) + \frac{1}{t+1000}u(t)$$

with $y = x_1$ as the measured output, x_1, x_2 and x_3 as the states, and with $u = 120\sin(t)$ as the control input. The controlled system invariance is:

$$y^{(3)}(t) + (t)y^{(2)}(t) + (t^2+1)y^{(1)}(t) + (t^2+t)y(t) - \frac{1}{t+1000}u(t) = 0 \tag{5.11}$$

The total length of the estimation window is $[0, 20s]$, but the results are shown on an inner window of $[2, 18s]$ where the B-spline estimation is considered to be reliable.

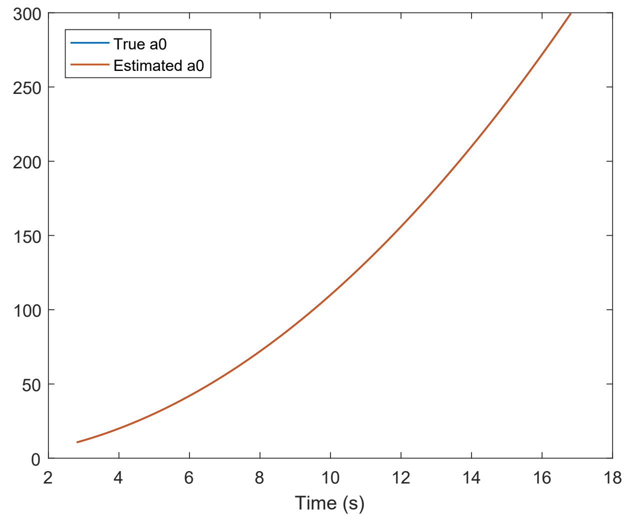

Figure 5.21 True $a_0(t)$ vs Estimated $a_0(t)$

Fig. 5.21 and Fig. 5.22 show the estimates for $a_0(t)$ and $a_1(t)$ against the true value.

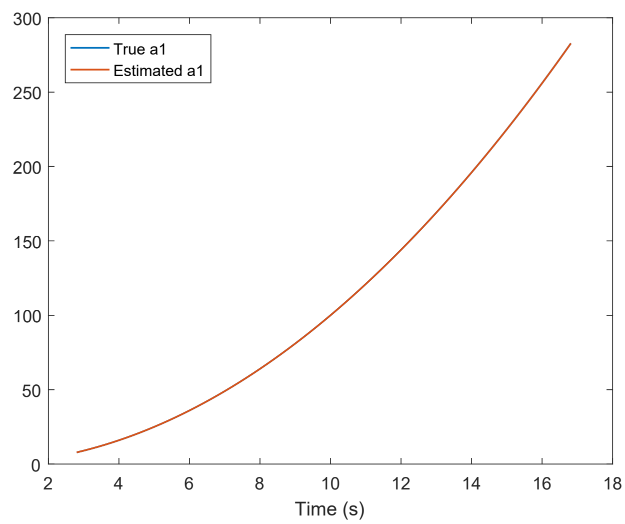

Figure 5.22 True $a_1(t)$ vs Estimated $a_1(t)$

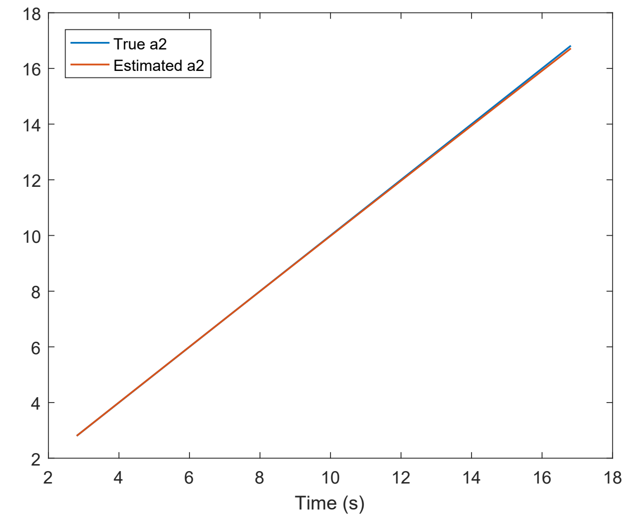

Figure 5.23 True $a_2(t)$ vs Estimated $a_2(t)$

Fig. 5.23 shows the estimated $a_2(t)$ against the true value. The precision of the fit degrades towards the end of the estimation interval due to the knot selection which was obtained by a trial-and-error method. The fit to the parameters can be improved by optimizing for the number and the position of the knots.

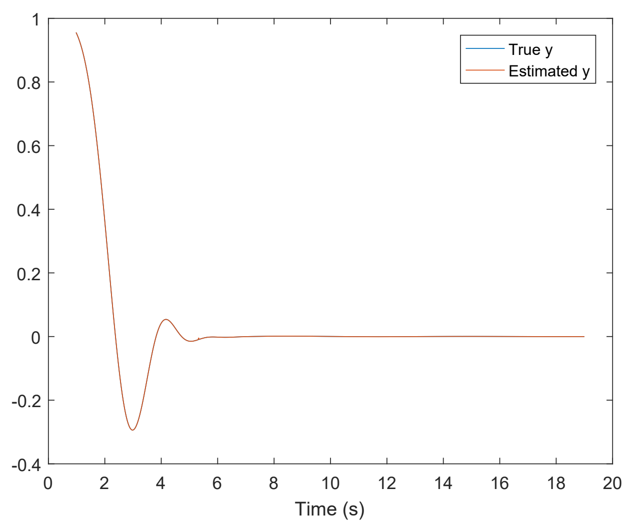

Figure 5.24 True $y(t)$ vs Estimated $y(t)$

Fig. 5.24 shows the estimated $y(t)$ and Fig. 5.25 shows the estimated $y^{(1)}(t)$. The small mismatch near $t = 7$ s is due to the numerical error in the integration of B-splines in the kernel representation of the system (5.11).

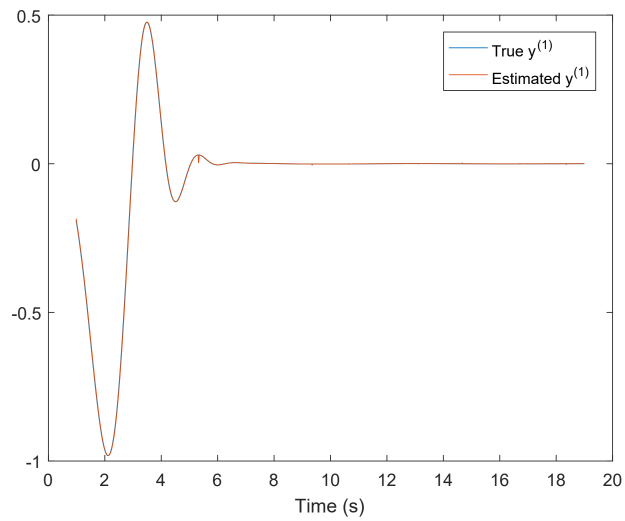

Figure 5.25 True $y^{(1)}(t)$ vs Estimated $y^{(1)}(t)$

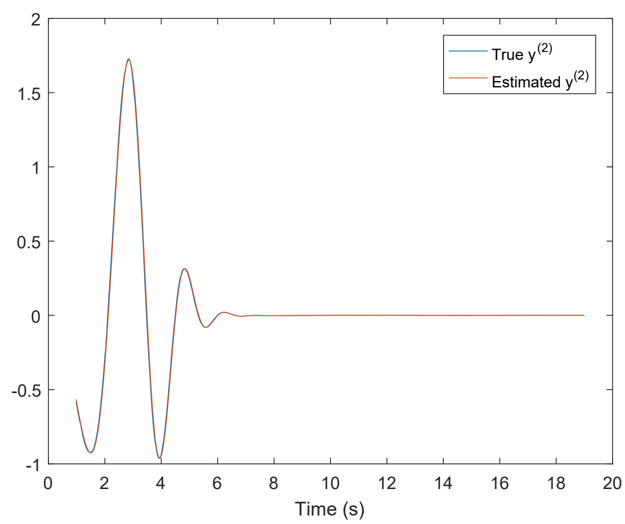

Figure 5.26 True $y^{(2)}(t)$ vs Estimated $y^{(2)}(t)$

Fig. 5.26 shows the estimated $y^{(2)}(t)$ against the true value.

5.2.3 Case 3: Input

System Description

$$
\begin{aligned}
\dot{x}_1(t) &= x_2(t) \\
\dot{x}_2(t) &= x_3(t) \\
\dot{x}_3(t) &= -(t^2+1)x_3(t) - (t)x_2(t) - (t^2+t)x_1(t) - u(t)
\end{aligned}
$$

with $y = x_1$ as the measured output, x_1, x_2 and x_3 as the states, and with $u = -120\sin(t)$ as the unknown input. The controlled system invariance is:

$$
y^{(3)}(t) + y^{(2)}(t) + (t)y^{(1)}(t) + (t^2+t)y(t) + u(t) = 0 \tag{5.12}
$$

The total length of the estimation window is $[0, 5s]$, but the results are shown on an inner window of $[0.5, 4.5s]$ where the B-spline estimation is considered to be reliable.

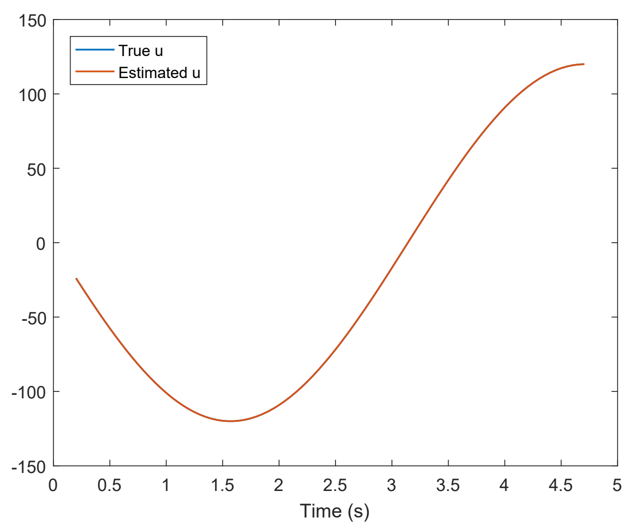

Figure 5.27 True $u(t)$ vs Estimated $u(t)$

Fig. 5.27 shows the estimated input $u(t)$ against the actual value.

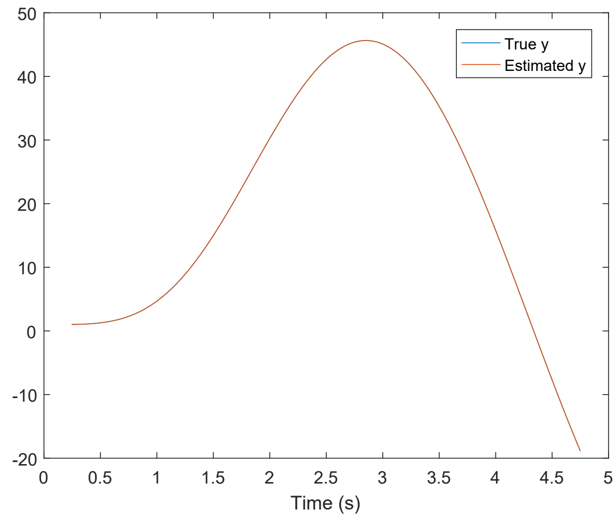

Figure 5.28 True $y(t)$ vs Estimated $y(t)$

Fig. 5.28 shows the estimated $y(t)$ against the true $y(t)$.

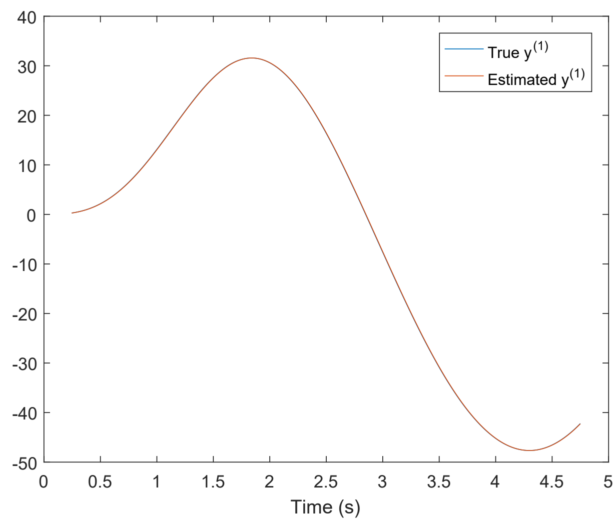

Figure 5.29 True $y^{(1)}(t)$ vs Estimated $y^{(1)}(t)$

Fig. 5.29 shows the estimated $y^{(1)}(t)$ and Fig. 5.30 shows that the estimate for $y^{(2)}(t)$

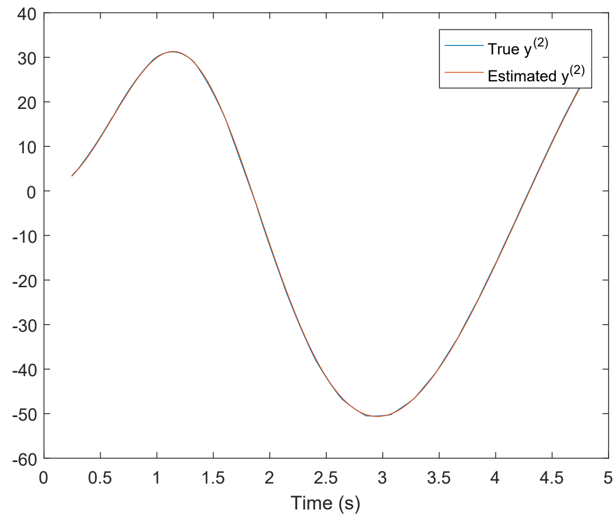

Figure 5.30 True $y^{(2)}(t)$ vs Estimated $y^{(2)}(t)$

5.2.4 Case 4: Two time-varying parameters and Input

System Description

$$\dot{x}_1(t) = x_2(t)$$
$$\dot{x}_2(t) = x_3(t)$$
$$\dot{x}_3(t) = -(t)x_2(t) - (t^2 + t)x_1(t) - u(t)$$

with $y = x_1$ as the measured output, x_1, x_2 and x_3 as the states, and with $u = t$ as the control input. The controlled system invariance is:

$$y^{(3)}(t) + (t)y^{(1)}(t) + (t^2 + t)y(t) + u(t) = 0 \tag{5.13}$$

The above system consists of two time-varying parameters $a_0(t)$ and $a_1(t)$. Both the parameters as well as the input is estimated here. The total length of the estimation window is $[0, 10s]$, but the results are shown on an inner window of $[1, 9s]$ where the B-spline estimation is considered to be reliable.

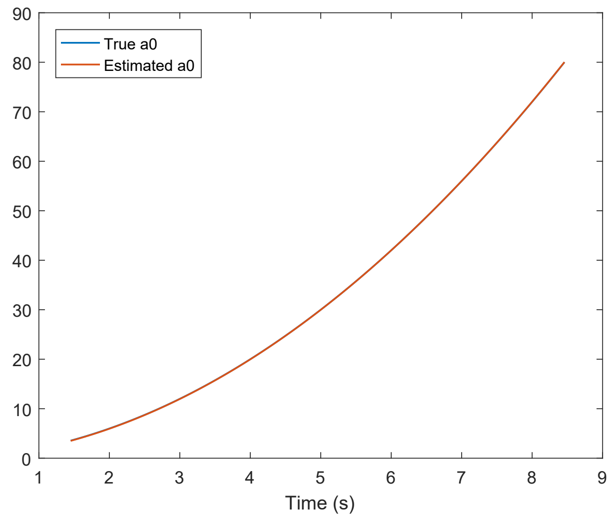

Figure 5.31 True $a_0(t)$ vs Estimated $a_0(t)$

Fig. 5.31 and Fig. 5.32 show that the estimated time-varying parameters of the system in(5.13). The mismatch towards the beginning of the estimation interval in Fig. 5.32 is due to the knot selection which was obtained by a trial-and-error method.

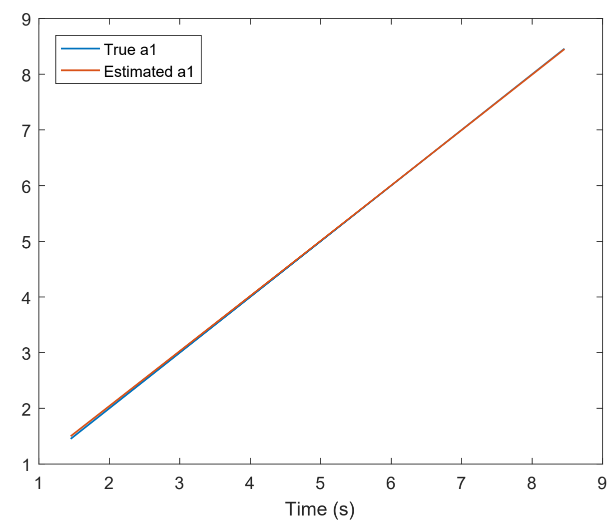

Figure 5.32 True $a_1(t)$ vs Estimated $a_1(t)$

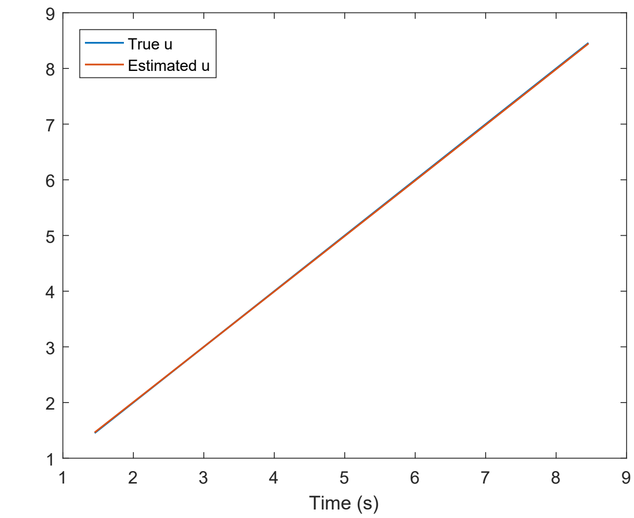

Figure 5.33 True $u(t)$ vs Estimated $u(t)$

Fig. 5.33 shows the estimated $u(t)$ and Fig. 5.34 shows the estimated $y(t)$.

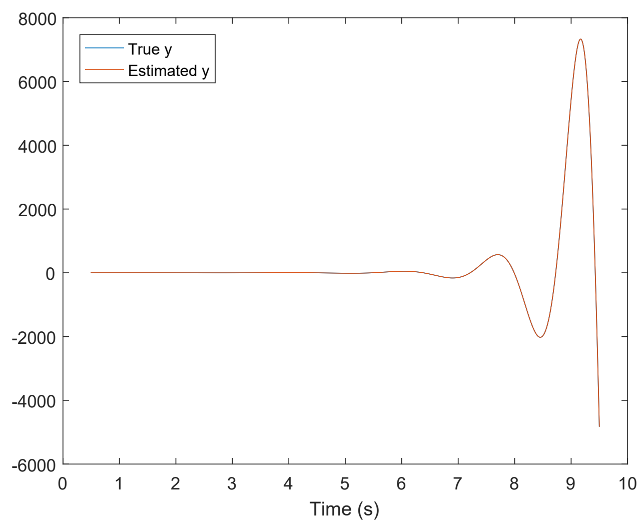

Figure 5.34 True $y(t)$ vs Estimated $y(t)$

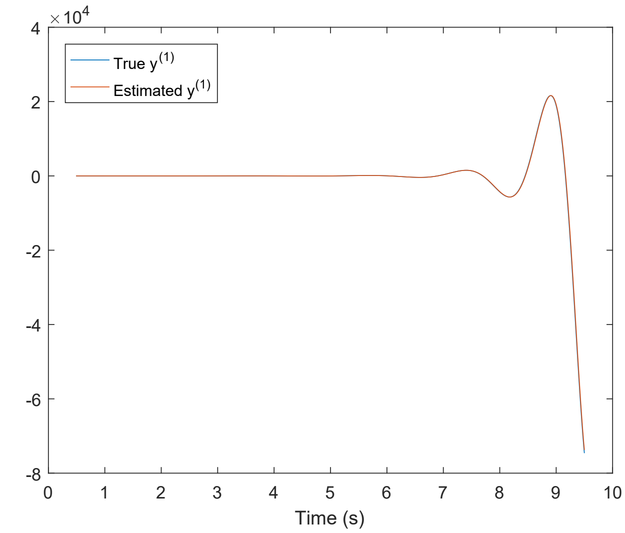

Figure 5.35 True $y^{(1)}(t)$ vs Estimated $y^{(1)}(t)$

Fig. 5.35 and Fig. 5.36 shows the estimated derivatives of the state $y(t)$.

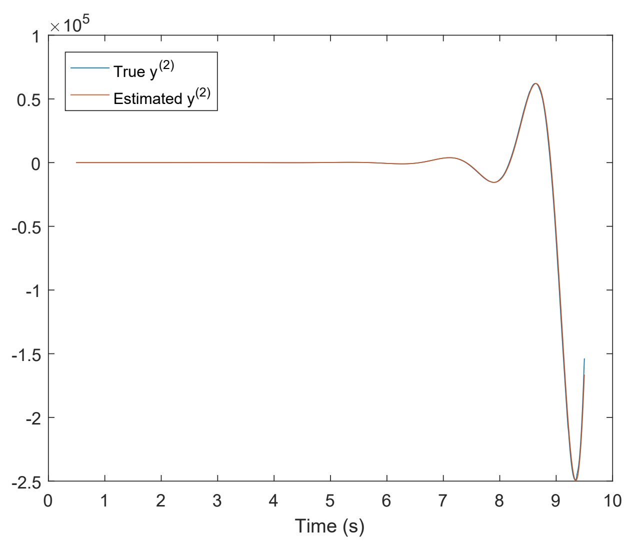

Figure 5.36 True $y^{(2)}(t)$ vs Estimated $y^{(2)}(t)$

5.3 Noisy Case

5.3.1 Second Order System

System Description

$$\dot{x}_1(t) = x_2(t)$$
$$\dot{x}_2(t) = -x_2(t) - (t^2+1)x_1(t) + \frac{1}{t+1000}u(t)$$

with $y = x_1$ as the measured output, x_1 and x_2 as the states, and with $u = 120\sin(t)$ as the control input. The controlled system invariance is:

$$y^{(2)}(t) + y^{(1)}(t) + (t^2+1)y(t) - \frac{1}{t+1000}u(t) = 0 \tag{5.14}$$

The system output is perturbed by 30 SNR (Signal-to-Noise Ratio) noise. The results were obtained using $n = 16$ and $p = 3$ (quadratic B-spline). The total length of the estimation window is $[0, 10s]$.

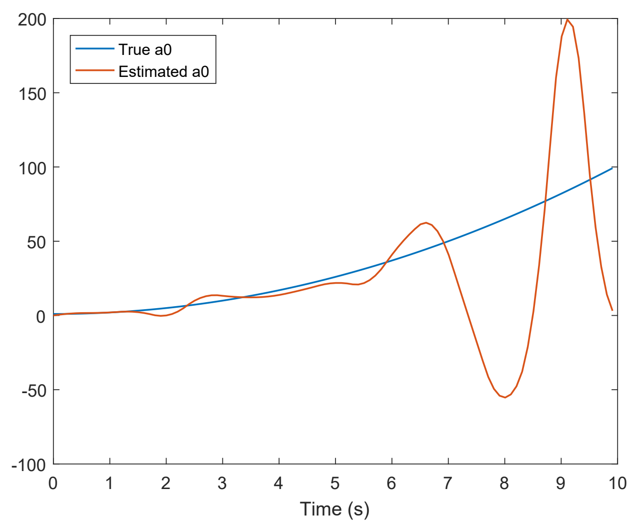

Figure 5.37 True $a_0(t)$ vs Estimated $a_0(t)$

Fig. 5.37 shows the estimated $a_0(t)$. The estimated value of the constant a_1 is 1.1127. True value of the constant a_1 is 1.

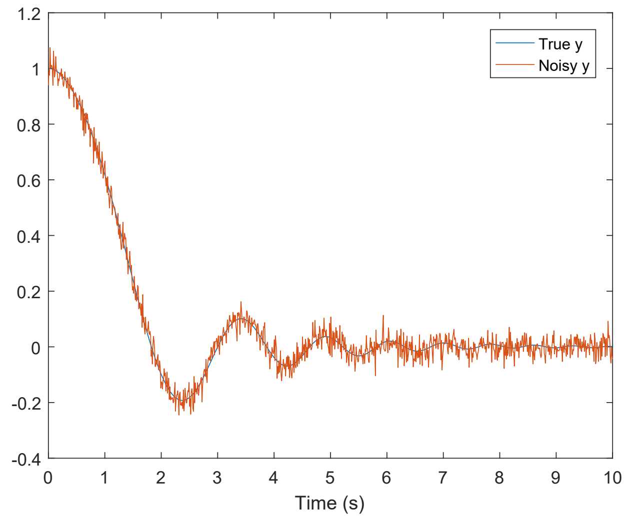

Figure 5.38 True $y(t)$ vs Noisy $y(t)$

Fig. 5.38 shows the output $y(t)$ perturbed with 30 SNR additive white gaussian noise.

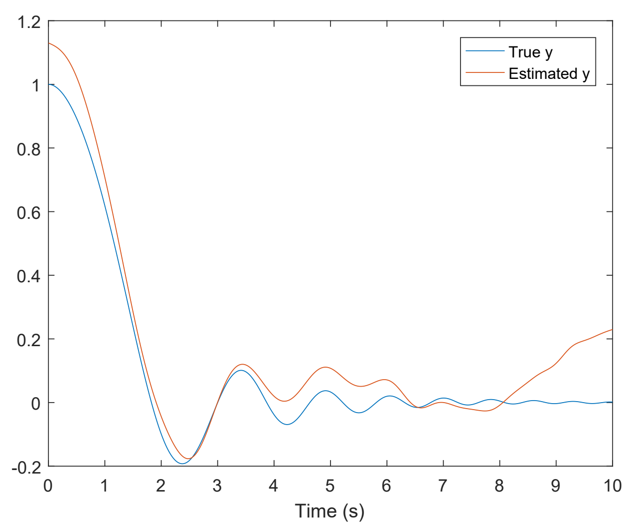

Figure 5.39 True $y(t)$ vs Estimated $y(t)$

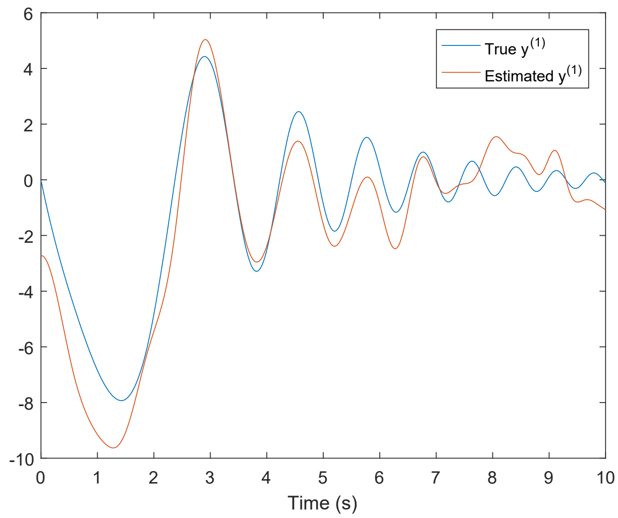

Figure 5.40 True $y^{(1)}(t)$ vs Estimated $y^{(1)}(t)$

Fig. 5.39 shows that the estimated $y(t)$ which diverges from the true value behaving in conjunction with the estimated value. Similarly, the estimated $y^{(1)}(t)$ does not fit the true value as seen in Fig. 5.40.

5.4 Estimation using P-splines

Estimation using P-Splines is similar to the B-Spline estimation but the difference arises from the added penalty to control the smoothness of the fitting. The parameter λ is used to control the smoothness of the fitted curve and is known as the smoothness parameter. P-Splines are superior to B-Splines in terms of the number of splines used for estimation and improved estimation of the derivatives of the state due to the added penalty. The following example illustrates it.

System Description

$$\dot{x}_1(t) = x_2(t)$$
$$\dot{x}_2(t) = -x_2(t) - (t^2+1)x_1(t) + \frac{1}{t+1000}u(t)$$

with $y = x_1$ as the measured output, x_1 and x_2 as the states, and with $u = 120\sin(t)$ as the control input. The controlled system invariance is:

$$y^{(2)}(t) + y^{(1)}(t) + (t^2 + 1)y(t) - \frac{1}{t + 1000}u(t) = 0 \tag{5.15}$$

The above system consists of one constant parameter and one time-varying parameter. The results were obtained using $n = 10$ (fewer than the number of B-splines used) and $p = 3$ (quadratic B-spline) and the smoothness parameter $\lambda = 10^{-5}$. The value of the smoothness parameter was chosen by a trial-and-error method. The total length of the estimation window is $[0, 10s]$, but the results are shown on an inner window of $[1, 9s]$ where the P-spline estimation is considered to be reliable.

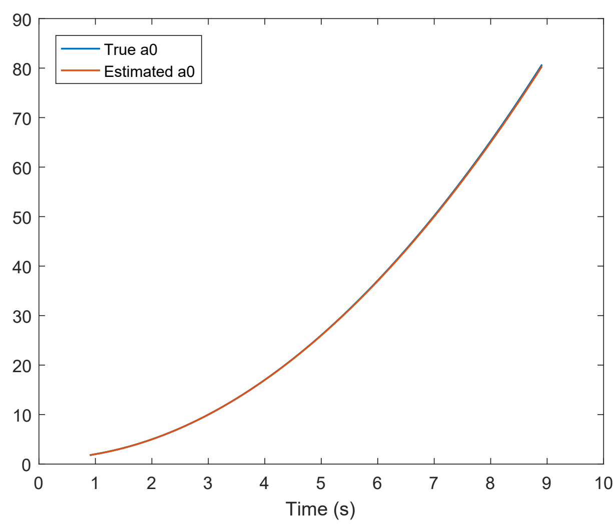

Figure 5.41 True $a_0(t)$ vs Estimated $a_0(t)$

Fig. 5.41 shows that the true value of $a_0(t)$ coincides with the estimated value. The estimated value of the constant a_1 is 1.0039. True value of the constant a_1 is 1.

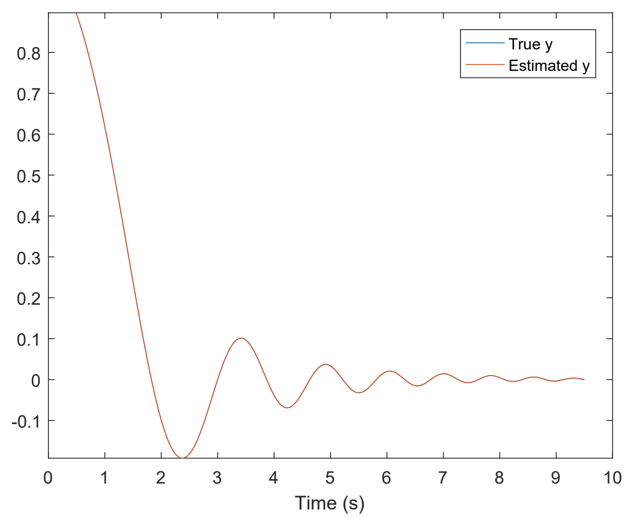

Figure 5.42 True $y(t)$ vs Estimated $y(t)$

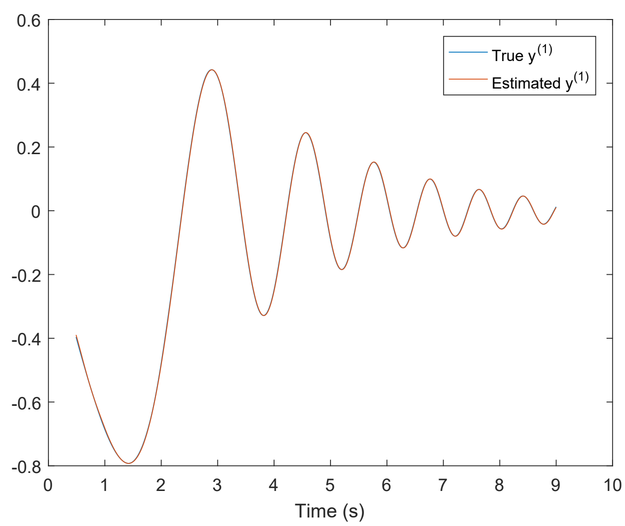

Figure 5.43 True $y^{(1)}(t)$ vs Estimated $y^{(1)}(t)$

Fig. 5.42 shows the estimated $y(t)$ against the true $y(t)$ and Fig. 5.43 shows the $y^{(1)}(t)$ estimate against the true value.

As seen from the above plots, estimation using P-Splines is advantageous relative to using B-Splines since it uses fewer number of splines and has other superior properties as discussed in Section 2.3.1 together with the features of the B-Splines and hence they can be used when a large number of splines are required for estimation. The advantages of using P-Splines are visible more prominently when a large number of splines are required to fit the time-varying parameters (complex) which will be discussed elsewhere.

Chapter 6

Conclusion and Future Work

State and parameter estimation is essential to achieve better control strategies. Most of the research and literature on state and parameter estimation focuses on asymptotic techniques such as Luenberger Observers and Kalman filters. With the development of Hybrid control strategies, fast switching control methods are imperative. Hence, the development of unconventional non-asymptotic estimation techniques is rapidly gaining intense research interest. However, the study of such techniques for LTV systems is limited, primarily due to difficulty involved in the time varying aspect. The proposed method of estimation in LTV systems overcomes this problem and is novel in the sense that the parameters, states, as well as the input is estimated.

Chapter 1 reviewed a brief history of automatic control, discussed open and closed loop systems and delivered a brief literature study on state and parameter estimation for LTV systems.

Chapter 2 discussed the background on B-splines and their properties. It also introduced P-splines and clarified their superior properties over B-Splines.

Chapter 3 dealt with the development of the double sided kernel for LTV systems in detail. The double sided kernel is derived for a third order system and expressions for the system state and its derivatives are also presented.

Chapter 4 proposed the estimation method which combines B-spline functional approximation with the double sided kernel to estimate the parameters, input, and subsequently the states of the system. It also described about the joint parameter, state, and input estimation approach.

Chapter 5 presented the results of the proposed methods taking examples of some simple second and third order systems. The results were primarily presented for the case where the output is free from noise. For the noiseless case, four cases were considered for both second and third order systems in order to show a comprehensive performance of the new estimation technique. A single example is also presented to describe the effect of noisy output measurements on the estimation method. The results show that the proposed approach works remarkably well for the noiseless case and requires denoisification techniques for the noisy case. Finally, a section is devoted to P-splines and an example is presented showing its superior performance over B-splines.

As seen in Chapter 4, the proposed method works remarkably well for the estimation of linear time-varying parameters and input (3 or fewer) in the noiseless case and it is very sensitive to noisy measurements as well as to the B-spline knot placement. This problem can be addressed by optimizing the number and position of the knots prior to the estimation of LTV system parameters. Also, the computational burden for the state, parameter and input estimation can be somewhat reduced using P-Splines instead of B-Splines since P-Splines require only fewer knots.

Future work will focus on simultaneous denoisification and parameter estimation in linear time-varying systems using adequately penalized P-splines with optimal selection of knots. Our future approach will resemble a kernel-based spline ridge regression approach to fully adaptive, non-asymptotic filtering of LTV systems. Finally, the latter ideas will be explored in a Bayesian framework.

Appendix A

Backward Kernel Derivation

Integrating the first term of (3.9) first time,

$$\int_{b-\sigma}^{b} (b-\zeta)^3 y^{(3)}(\zeta)\,\mathrm{d}\zeta = (b-\zeta)^3 y^{(2)}(\zeta)\,\Big|_{b-\sigma}^{b} + \int_{b-\sigma}^{b} 3(b-\zeta)^2 y^{(2)}(\zeta)\,\mathrm{d}\zeta$$

$$= -\sigma^3 y^{(2)}(b-\sigma) + \left[3(b-\zeta)^2 y^{(1)}(\zeta)\,\Big|_{b-\sigma}^{b} + \int_{b-\sigma}^{b} 6(b-\zeta) y^{(1)}(\zeta)\mathrm{d}\zeta\right]$$

$$= -\sigma^3 y^{(2)}(b-\sigma) - 3\sigma^2 y^{(1)}(b-\sigma) + 6(b-\zeta)y(\zeta)\,\Big|_{b-\sigma}^{b} + \int_{b-\sigma}^{b} 6y(\zeta)\,\mathrm{d}\zeta$$

$$= -\sigma^3 y^{(2)}(b-\sigma) - 3\sigma^2 y^{(1)}(b-\sigma) - 6\sigma y(b-\sigma) + \int_{b-\sigma}^{b} 6y(\zeta)\,\mathrm{d}\zeta \quad \text{(A.1)}$$

When we integrate again, the upper limit on the integral becomes a 'dummy variable', that is 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 $6(b-\zeta^{'})y(\zeta^{'})$

Integrating (A.1) again,

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

$$= -(b-\zeta')^3 y^{(1)}(\zeta')\ |_{b-\sigma}^{b} - \int_{b-\sigma}^{b} 3(b-\zeta')^2 y^{(1)}(\zeta')\,\mathrm{d}\zeta' - \int_{b-\sigma}^{b} 3(b-\zeta')^2 y^{(1)}(\zeta')\,\mathrm{d}\zeta'$$

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

$$= -(b-\zeta')^3 y^{(1)}(\zeta')\ |_{b-\sigma}^{b} - \int_{b-\sigma}^{b} 6(b-\zeta')^2 y^{(1)}(\zeta')\,\mathrm{d}\zeta' - \int_{b-\sigma}^{b} 6(b-\zeta')y(\zeta')\,\mathrm{d}\zeta'$$

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

$$= \sigma^3 y^{(1)}(b-\sigma) - 6(b-\zeta')^2 y(\zeta')\ |_{b-\sigma}^{b} - \int_{b-\sigma}^{b} 12(b-\zeta')y(\zeta')\,\mathrm{d}\zeta' - \int_{b-\sigma}^{b} 6(b-\zeta')y(\zeta')\,\mathrm{d}\zeta'$$

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

$$= \sigma^3 y^{(1)}(b-\sigma) + 6\sigma^2 y(b-\sigma) - \int_{b-\sigma}^{b} 18(b-\zeta')y(\zeta')\,\mathrm{d}\zeta' + \int_{b-\sigma}^{b}\int_{\zeta'}^{b} 6y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta' \tag{A.2}$$

As shown earlier,the upper limit again becomes a 'dummy variable' and now we set $\zeta'' = b-\sigma$. Integrating the third time yields,

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

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

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

$$= (b-\zeta'')^3 y(\zeta'') \mid_{b-\sigma}^{b} + \int_{b-\sigma}^{b} 3(b-\zeta'')^2 y(\zeta'')\mathrm{d}\zeta'' + \int_{b-\sigma}^{b} 6(b-\zeta'')^2 y(\zeta'')\mathrm{d}\zeta''$$

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

$$= -\sigma^3 y(b-\sigma) + \int_{b-\sigma}^{b} 9(b-\zeta'')^2 y(\zeta'')\,\mathrm{d}\zeta'' - \int_{b-\sigma}^{b}\int_{\zeta''}^{b} 18(b-\zeta')y(\zeta')\,\mathrm{d}\zeta'\mathrm{d}\zeta''$$

$$+ \int_{b-\sigma}^{b}\int_{\zeta''}^{b}\int_{\zeta'}^{b} 6y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta'' \tag{A.3}$$

Integrating the second term in (3.9) first time,

$$\int_{b-\sigma}^{b} a_2(\zeta)(b-\zeta)^3 y^{(2)}(\zeta)\,\mathrm{d}\zeta$$

$$= a_2(\zeta)(b-\zeta)^3 y^{(1)}(\zeta) \mid_{b-\sigma}^{b} - \int_{b-\sigma}^{b} \Big[-3a_2(\zeta)(b-\zeta)^2 + (b-\zeta)^3 a_2^{(1)}(\zeta)\Big] y^{(1)}(\zeta)\,\mathrm{d}\zeta$$

$$= a_2(\zeta)(b-\zeta)^3 y^{(1)}(\zeta) \mid_{b-\sigma}^{b} - \Big[-3a_2(\zeta)(b-\zeta)^2 + (b-\zeta)^3 a_2^{(1)}(\zeta)\Big] y(\zeta) \mid_{b-\sigma}^{b}$$

$$+ \int_{b-\sigma}^{b} \Big[6(b-\zeta)a_2(\zeta) - 3(b-\zeta)^2 a_2^{(1)}(\zeta)\Big] y(\zeta)\mathrm{d}\zeta$$

$$+ \int_{b-\sigma}^{b} \Big[-3(b-\zeta)^2 a_2^{(1)}(\zeta) + (b-\zeta)^3 a_2^{(2)}(\zeta)\Big] y(\zeta)\mathrm{d}\zeta$$

$$= -a_2(b-\sigma)(\sigma)^3 y^{(1)}(b-\sigma) + \Big[-3a_2(b-\sigma)(\sigma)^2 + (\sigma)^3 a_2^{(1)}(b-\sigma)\Big] y(b-\sigma)$$

$$+ \int_{b-\sigma}^{b} \Big[6(b-\zeta)a_2(\zeta) - 3(b-\zeta)^2 a_2^{(1)}(\zeta)\Big] y(\zeta)\mathrm{d}\zeta$$

$$+\int_{b-\sigma}^{b}\Big[-3(b-\zeta)^2a_2^{(1)}(\zeta)+(b-\zeta)^3a_2^{(2)}(\zeta)\Big]y(\zeta)\mathrm{d}\zeta \tag{A.4}$$

As shown earlier,the upper limit again becomes a 'dummy variable' and now we set $\zeta^{'} = b-\sigma$. Integrating the second time yields,

$$\int_{b-\sigma}^{b}\int_{\zeta'}^{b} a_2(\zeta)(b-\zeta)^3y^{(2)}(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'$$

$$=\int_{b-\sigma}^{b} -a_2(\zeta')(b-\zeta')^3y^{(1)}(\zeta')\,\mathrm{d}\zeta' + \int_{b-\sigma}^{b}\Big[-3a_2(\zeta')(b-\zeta')^2+(b-\zeta')^3a_2^{(1)}(\zeta')\Big]y(\zeta')\,\mathrm{d}\zeta'$$

$$+\int_{b-\sigma}^{b}\int_{\zeta'}^{b}\Big[6(b-\zeta)a_2(\zeta)-3(b-\zeta)^2a_2^{(1)}(\zeta)\Big]y(\zeta)\mathrm{d}\zeta\mathrm{d}\zeta'$$

$$+\int_{b-\sigma}^{b}\int_{\zeta'}^{b}\Big[-3(b-\zeta)^2a_2^{(1)}(\zeta)+(b-\zeta)^3a_2^{(2)}(\zeta)\Big]y(\zeta)\mathrm{d}\zeta\mathrm{d}\zeta'$$

$$=-a_2(\zeta')(b-\zeta')^3y(\zeta')\mid_{b-\sigma}^{b}+\int_{b-\sigma}^{b}\Big[-3a_2(\zeta')(b-\zeta')^2+(b-\zeta')^3a_2^{(1)}(\zeta')\Big]y(\zeta')\,\mathrm{d}\zeta'$$

$$+\int_{b-\sigma}^{b}\Big[-3a_2(\zeta')(b-\zeta')^2+(b-\zeta')^3a_2^{(1)}(\zeta')\Big]y(\zeta')\,\mathrm{d}\zeta'$$

$$+\int_{b-\sigma}^{b}\int_{\zeta'}^{b}\Big[6(b-\zeta)a_2(\zeta)-3(b-\zeta)^2a_2^{(1)}(\zeta)\Big]y(\zeta)\mathrm{d}\zeta\mathrm{d}\zeta'$$

$$+\int_{b-\sigma}^{b}\int_{\zeta'}^{b}\Big[-3(b-\zeta)^2a_2^{(1)}(\zeta)+(b-\zeta)^3a_2^{(2)}(\zeta)\Big]y(\zeta)\mathrm{d}\zeta\mathrm{d}\zeta'$$

$$=a_2(b-\sigma)(\sigma)^3y(b-\sigma)+\int_{b-\sigma}^{b}2\Big[-3a_2(\zeta')(b-\zeta')^2+(b-\zeta')^3a_2^{(1)}(\zeta')\Big]y(\zeta')\,\mathrm{d}\zeta'$$

$$+\int_{b-\sigma}^{b}\int_{\zeta'}^{b}\Big[6(b-\zeta)a_2(\zeta)-3(b-\zeta)^2a_2^{(1)}(\zeta)\Big]y(\zeta)\mathrm{d}\zeta\mathrm{d}\zeta'$$

$$+\int_{b-\sigma}^{b}\int_{\zeta'}^{b}\Big[-3(b-\zeta)^2a_2^{(1)}(\zeta)+(b-\zeta)^3a_2^{(2)}(\zeta)\Big]y(\zeta)\mathrm{d}\zeta\mathrm{d}\zeta' \tag{A.5}$$

Integrating the second term of (3.9) third time yields,

$$\int_{b-\sigma}^{b}\int_{\zeta''}^{b}\int_{\zeta'}^{b}a_2(\zeta)(b-\zeta)^3y^{(2)}(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta''$$

$$=\int_{b-\sigma}^{b}a_2(\zeta'')(b-\zeta'')^3y(\zeta'')\,\mathrm{d}\zeta''+\int_{b-\sigma}^{b}\int_{\zeta''}^{b}2\big[-3a_2(\zeta')(b-\zeta')^2\big]y(\zeta')\,\mathrm{d}\zeta'\mathrm{d}\zeta''$$

$$+\int_{b-\sigma}^{b}\int_{\zeta''}^{b}2\big[(b-\zeta')^3a_2^{(1)}(\zeta')\big]y(\zeta')\,\mathrm{d}\zeta'\mathrm{d}\zeta''+\int_{b-\sigma}^{b}\int_{\zeta''}^{b}\int_{\zeta'}^{b}\big[6(b-\zeta)a_2(\zeta)\big]y(\zeta)\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta''$$

$$+\int_{b-\sigma}^{b}\int_{\zeta''}^{b}\int_{\zeta'}^{b}2\big[-3(b-\zeta)^2a_2^{(1)}(\zeta)\big]y(\zeta)\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta''$$

$$+\int_{b-\sigma}^{b}\int_{\zeta''}^{b}\int_{\zeta'}^{b}\big[(b-\zeta)^3a_2^{(2)}(\zeta)\big]y(\zeta)\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta'' \tag{A.6}$$

Integrating the third term of (3.9) first time yields,

$$\int_{b-\sigma}^{b}a_1(\zeta)(b-\zeta)^3y^{(1)}(\zeta)\,\mathrm{d}\zeta$$

$$=a_1(\zeta)(b-\zeta)^3y(\zeta)\;\Big|_{b-\sigma}^{b}+\int_{b-\sigma}^{b}3a_1(\zeta)(b-\zeta)^2y(\zeta)\,\mathrm{d}\zeta-\int_{b-\sigma}^{b}a_1^{(1)}(\zeta)(b-\zeta)^3y(\zeta)\,\mathrm{d}\zeta$$

$$=-a_1(b-\sigma)(\sigma)^3y(b-\sigma)+\int_{b-\sigma}^{b}3a_1(\zeta)(b-\zeta)^2y(\zeta)\,\mathrm{d}\zeta-\int_{b-\sigma}^{b}a_1^{(1)}(\zeta)(b-\zeta)^3y(\zeta)\,\mathrm{d}\zeta \tag{A.7}$$

Integrating the second time,

$$\int_{b-\sigma}^{b}\int_{\zeta'}^{b} a_1(\zeta')(b-\zeta')^3 y^{(1)}(\zeta')\,\mathrm{d}\zeta\mathrm{d}\zeta'$$

$$= -\int_{\zeta'}^{b} a_1(\zeta')(b-\zeta')^3 y(\zeta')\mathrm{d}\zeta' + \int_{b-\sigma}^{b}\int_{\zeta''}^{b} 3a_1(\zeta')(b-\zeta')^2 y(\zeta')\,\mathrm{d}\zeta'\mathrm{d}\zeta''$$
$$- \int_{b-\sigma}^{b}\int_{\zeta''}^{b} a_1^{(1)}(\zeta')(b-\zeta')^3 y(\zeta')\,\mathrm{d}\zeta'\mathrm{d}\zeta'' \tag{A.8}$$

Integrating the third time yields,

$$\int_{b-\sigma}^{b}\int_{\zeta''}^{b}\int_{\zeta'}^{b} a_1(\zeta)(b-\zeta)^3 y^{(1)}(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta''$$

$$= -\int_{b-\sigma}^{b}\int_{\zeta''}^{b} a_1(\zeta')(b-\zeta')^3 y(\zeta')\,\mathrm{d}\zeta'\mathrm{d}\zeta'' + \int_{b-\sigma}^{b}\int_{\zeta''}^{b}\int_{\zeta'}^{b} 3a_1(\zeta)(b-\zeta)^2 y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta''$$
$$- \int_{b-\sigma}^{b}\int_{\zeta''}^{b}\int_{\zeta'}^{b} a_1^{(1)}(\zeta)(b-\zeta)^3 y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta'' \tag{A.9}$$

Finally, the fourth term is

$$\int_{b-\sigma}^{b}\int_{\zeta''}^{b}\int_{\zeta'}^{b} a_0(\zeta)(b-\zeta)^3 y(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta'' \tag{A.10}$$

Repeating the same procedure for the next three terms involving the input we obtain the following. Integrating the fifth term yields,

$$\int_{b-\sigma}^{b}\int_{\zeta''}^{b}\int_{\zeta'}^{b} b_2(\zeta)(b-\zeta)^3 u^{(2)}(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta''$$

$$= \int_{b-\sigma}^{b} b_2(\zeta'')(b-\zeta'')^3 u(\zeta'')\,\mathrm{d}\zeta'' + \int_{b-\sigma}^{b}\int_{\zeta''}^{b} 2\big[-3b_2(\zeta')(b-\zeta')^2\big]u(\zeta')\,\mathrm{d}\zeta'\mathrm{d}\zeta''$$
$$+ \int_{b-\sigma}^{b}\int_{\zeta''}^{b} 2\big[(b-\zeta')^3 b_2^{(1)}(\zeta')\big]u(\zeta')\,\mathrm{d}\zeta'\mathrm{d}\zeta'' + \int_{b-\sigma}^{b}\int_{\zeta''}^{b}\int_{\zeta'}^{b}\big[6(b-\zeta)b_2(\zeta)\big]u(\zeta)\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta''$$
$$+ \int_{b-\sigma}^{b}\int_{\zeta''}^{b}\int_{\zeta'}^{b}\big[-6(b-\zeta)^2 b_2^{(1)}(\zeta)\big]u(\zeta)\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta'' + \int_{b-\sigma}^{b}\int_{\zeta''}^{b}\int_{\zeta'}^{b}\big[(b-\zeta)^3 b_2^{(2)}(\zeta)\big]u(\zeta)\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta'' \tag{A.11}$$

Integrating the sixth term yields,

$$\int_{b-\sigma}^{b}\int_{\zeta''}^{b}\int_{\zeta'}^{b} b_1(\zeta)(b-\zeta)^3 u^{(1)}(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta''$$
$$= -\int_{b-\sigma}^{b}\int_{\zeta''}^{b} b_1(\zeta')(b-\zeta')^3 u(\zeta')\,\mathrm{d}\zeta'\mathrm{d}\zeta'' + \int_{b-\sigma}^{b}\int_{\zeta''}^{b}\int_{\zeta'}^{b} 3b_1(\zeta)(b-\zeta)^2 u(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta''$$
$$-\int_{b-\sigma}^{b}\int_{\zeta''}^{b}\int_{\zeta'}^{b} b_1^{(1)}(\zeta)(b-\zeta)^3 u(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta'' \tag{A.12}$$

Integrating the last term gives,

$$\int_{b-\sigma}^{b}\int_{\zeta''}^{b}\int_{\zeta'}^{b} b_0(\zeta)(b-\zeta)^3 u(\zeta)\,\mathrm{d}\zeta\mathrm{d}\zeta'\mathrm{d}\zeta'' \tag{A.13}$$

Collecting the terms in (A.1) - (A.13) yields

$$\sigma^3 y(b-\sigma) = \int_{b-\sigma}^{b}\left[9(b-\zeta'')^2 + a_2(\zeta'')(b-\zeta'')^3\right]y(\zeta'')\,\mathrm{d}\zeta''$$
$$+ \int_{b-\sigma}^{b}\int_{\zeta''}^{b}\left[-18(b-\zeta') - 6a_2(\zeta')(b-\zeta')^2 + (b-\zeta')^3(2a_2^{(1)}(\zeta') - a_1(\zeta'))\right]y(\zeta')\,\mathrm{d}\zeta'\mathrm{d}\zeta''$$

$$+ \int\limits_{b-\sigma}^{b} \int\limits_{\zeta''}^{b} \int\limits_{\zeta'}^{b} \Big[6 + 6a_2(\zeta)(b-\zeta) - 6(b-\zeta)^2 a_2^{(1)}(\zeta)$$

$$+ (b-\zeta)^3 a_2^{(2)}(\zeta) + 3a_1(\zeta)(b-\zeta)^2 - (b-\zeta)^3 a_1^{(1)}(\zeta) + a_0(\zeta)(b-\zeta)^3 \Big] y(\zeta)\,\mathrm{d}\zeta \mathrm{d}\zeta' \mathrm{d}\zeta''$$

$$+ \int\limits_{b-\sigma}^{b} b_2(\zeta'')(b-\zeta'')^3 u(\zeta'')\,\mathrm{d}\zeta''$$

$$+ \int\limits_{b-\sigma}^{b} \int\limits_{\zeta''}^{b} \Big[-6b_2(\zeta')(b-\zeta')^2 + 2b_2^{(1)}(\zeta')(b-\zeta')^3 - b_1(\zeta')(b-\zeta')^3 \Big] u(\zeta')\,\mathrm{d}\zeta' \mathrm{d}\zeta''$$

$$+ \int\limits_{b-\sigma}^{b} \int\limits_{\zeta''}^{b} \int\limits_{\zeta'}^{b} \Big[6b_2(\zeta)(b-\zeta) - 6(b-\zeta)^2 b_2^{(1)}(\zeta) + (b-\zeta)^3 b_2^{(2)}(\zeta) + 3b_1(\zeta)(b-\zeta)^2$$

$$- (b-\zeta)^3 b_1^{(1)}(\zeta) + b_0(\zeta)(b-\zeta)^3 \Big] u(\zeta)\,\mathrm{d}\zeta \mathrm{d}\zeta' \mathrm{d}\zeta'' \tag{A.14}$$

In order to obtain the integrals in the form required to apply the Cauchy formula for repeated integration, we flip the limits of the integration from $(\zeta' \to b)$ to $-(b \to \zeta')$ and hence a negative sign is introduced. Also since the integration variable has a negative sign, a second negative sign is again introduced. Therefore, for odd integrals the sign is unchanged whereas for even integrals the sign changes. Applying the Cauchy formula for repeated integration on (A.14) while letting $b - \sigma = t$, we get,

$$(b-t)^3 y(t)$$

$$= \int\limits_{t}^{b} \Big[9(b-\zeta)^2 + a_2(\zeta)(b-\zeta)^3 \Big] y(\zeta)\,\mathrm{d}\zeta$$

$$+ \int\limits_{t}^{b} (t-\zeta) \Big[18(b-\zeta) + 6a_2(\zeta)(b-\zeta)^2 - 2(b-\zeta)^3 a_2^{(1)}(\zeta) + a_1(\zeta)(b-\zeta)^3 \Big] y(\zeta)\,\mathrm{d}\zeta$$

$$+ \frac{1}{2} \int\limits_{t}^{b} (t-\zeta)^2 \Big[6 + 6a_2(\zeta) - 6a_2^{(1)}(\zeta)(b-\zeta)^2 + a_2^{(2)}(\zeta)(b-\zeta)^3 + 3a_1(\zeta)(b-\zeta)^2$$

$$
\begin{aligned}
&- a_1^{(1)}(\zeta)(b-\zeta)^3 + a_0(\zeta)(b-\zeta)^3\Big] y(\zeta)\,\mathrm{d}\zeta + \int_t^b [b_2(\zeta)(b-\zeta)^3] u(\zeta)\,\mathrm{d}\zeta \\
&+ \int_t^b (t-\zeta)\Big[6b_2(\zeta)(b-\zeta)^2 - 2(b-\zeta)^3 b_2^{(1)}(\zeta) + b_1(\zeta)(b-\zeta)^3\Big] u(\zeta)\,\mathrm{d}\zeta \\
&+ \frac{1}{2}\int_t^b (t-\zeta)^2\Big[6b_2(\zeta)(b-\zeta) - 6b_2^{(1)}(\zeta)(b-\zeta)^2 + b_2^{(2)}(\zeta)(b-\zeta)^3 + 3b_1(\zeta)(b-\zeta)^2 \\
&- b_1^{(1)}(\zeta)(b-\zeta)^3 + b_0(\zeta)(b-\zeta)^3\Big] u(\zeta)\,\mathrm{d}\zeta \\
&\triangleq \int_t^b K_{By}(t,\tau) y(\tau)\,\mathrm{d}\tau + \int_t^b K_{Bu}(t,\tau) u(\tau)\,\mathrm{d}\tau
\end{aligned} \tag{A.15}
$$

with $K_{By}(t,\tau)$ defined as

$$
\begin{aligned}
K_{By}(t,\tau) \triangleq &\Big[9(b-\tau)^2 + (b-\tau)^3 a_2(\tau)\Big] \\
&+ (t-\tau)\Big[18(b-\tau) + 6(b-\tau)^2 a_2(\tau) - 2(b-\tau)^3 a_2^{(1)}(\tau) + (b-\tau)^3 a_1(\tau)\Big] \\
&+ \frac{(t-\tau)^2}{2}\Big[6 + 6(b-\tau)a_2(\tau) - 6(b-\tau)^2 a_2^{(1)}(\tau) + (b-\tau)^3 a_2^{(2)}(\tau) \\
&\qquad + 3(b-\tau)^2 a_1(\tau) - (b-\tau)^3 a_1^{(1)}(\tau) + (b-\tau)^3 a_0(\tau)\Big]
\end{aligned} \tag{A.16}
$$

and $K_{Bu}(t,\tau)$ defined as

$$
\begin{aligned}
K_{Bu}(t,\tau) \triangleq &\Big[(b-\tau)^3 b_2(\tau)\Big] \\
&+ (t-\tau)\Big[6(b-\tau)^2 b_2(\tau) - 2(b-\tau)^3 b_2^{(1)}(\tau) + (b-\tau)^3 b_1(\tau)\Big] \\
&+ \frac{(t-\tau)^2}{2}\Big[6(b-\tau)b_2(\tau) - 6(b-\tau)^2 b_2^{(1)}(\tau) + (b-\tau)^3 b_2^{(2)}(\tau) \\
&\qquad + 3(b-\tau)^2 b_1(\tau) - (b-\tau)^3 b_1^{(1)}(\tau) + (b-\tau)^3 b_0(\tau)\Big]
\end{aligned} \tag{A.17}
$$

Bibliography

[1] N. Nise, *Control Systems Engineering, 7th Edition*. Wiley, 2015.

[2] S. Bennett, "A brief history of automatic control," *IEEE Control Systems*, vol. 16, pp. 17–25, Jun 1996.

[3] D. Sridhar, D. P. Ghoshal, and H. Michalska, "B-splines in joint parameter and state estimation in linear time-varying systems," *submitted to IEEE American Control Conference*, 2018.

[4] 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.

[5] M. Fliess and H. Sira-Ramirez, "Closed-loop parametric identification for continuous-time linear systems via new algebraic techniques," in *Identification of Continuous-time Models from sampled Data*, pp. 363–391, Springer, 2008.

[6] M. Fliess, "Analyse non standard du bruit," *Comptes Rendus Mathematique*, vol. 342, no. 10, pp. 797–802, 2006.

[7] D. P. Ghoshal, K. Gopalakrishnan, and H. Michalska, "Using invariance to extract signal from noise," in *American Control Conference*, 2017.

[8] D. P. Ghoshal, K. Gopalakrishnan, and H. Michalska, "Algebraic parameter estimation using kernel representation of linear systems," in *The 20th World Congress of the International Federation of Automatic Control*, 2017.

[9] D. P. Ghoshal, K. Gopalakrishnan, and H. Michalska, "Kernel-based adaptive multiple model target tracking," in *IEEE Conference on Control Technology and Applications*, 2017.

[10] Y. Tian, T. Floquet, and W. Perruquetti, "Fast state estimation in linear time-varying systems: an algebraic approach," in *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, pp. 2539–2544, IEEE, 2008.

[11] D. Dimogianopoulos and R. Lozano, "Direct least squares estimation for linear time varying systems," in *Control Conference (ECC), 1999 European*, pp. 1884–1889, IEEE, 1999.

[12] L. C. A. Souza and R. M. Palhares, "Parameter estimation on linear time-varying systems," *Journal of the Franklin Institute*, vol. 348, no. 4, pp. 777–789, 2011.

[13] S. Bezzaoucha, B. Marx, D. Maquin, and J. Ragot, "State and parameter estimation for time-varying systems: a takagi-sugeno approach," *IFAC Proceedings Volumes*, vol. 46, no. 2, pp. 761–766, 2013.

[14] S. Mehrkanoon, T. Falck, and J. A. Suykens, "Parameter estimation for time varying dynamical systems using least squares support vector machines," *IFAC Proceedings Volumes*, vol. 45, no. 16, pp. 1300–1305, 2012.

[15] D. P. Ghoshal, K. Gopalakrishnan, and H. Michalska, "Double-sided kernel observer for linear time-varying systems," in *IEEE Conference on Control Technology and Applications*, 2017.

[16] T. Lyche and K. Mørken, "Spline methods draft," 11 2017.

[17] Y. Kineri, M. Wang, H. Lin, and T. Maekawa, "B-spline surface fitting by iterative geometric interpolation/approximation algorithms," *Computer-Aided Design*, vol. 44, no. 7, pp. 697 – 708, 2012.

[18] S. He, D. Ou, C. Yan, and C.-H. Lee, "A chord error conforming tool path b-spline fitting method for nc machining based on energy minimization and lspia," *Journal of Computational Design and Engineering*, vol. 2, no. 4, pp. 218 – 232, 2015. CNCM 2015 Special Issue.

[19] M. Elbanhawi, M. Simic, and R. N. Jazar, "Continuous path smoothing for car-like robots using b-spline curves," *Journal of Intelligent & Robotic Systems*, vol. 80, pp. 23–56, Dec 2015.

[20] B. Cao, G. I. Doods, and G. W. Irwin, "Time-optimal and smooth constrained path planning for robot manipulators," in *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*, pp. 1853–1858 vol.3, May 1994.

[21] C. De Boor, *A practical guide to splines*, vol. 27. Springer-Verlag New York, 1978.

[22] P. H. C. Eilers, B. D. Marx, and M. Durbán, "Twenty years of p-splines," 2015.

[23] P. H. C. Eilers and B. D. Marx, "Flexible smoothing with b-splines and penalties," *STATISTICAL SCIENCE*, vol. 11, pp. 89–121, 1996.

[24] S. N. Wood, "P-splines with derivative based penalties and tensor product smoothing of unevenly distributed data," *Statistics and Computing*, vol. 27, pp. 985–989, Jul 2017.

[25] E. Kreindler and P. Sarachik, "On the concepts of controllability and observability of linear systems," *IEEE Transactions on Automatic Control*, vol. 9, no. 2, pp. 129–136, 1964.

[26] L. M. Silverman and H. Meadows, "Controllability and observability in time-variable linear systems," *SIAM Journal on Control*, vol. 5, no. 1, pp. 64–73, 1967.

[27] Y. Lyubich, "Functional analysis 1: Linear functional analysis," vol. 19, p. 77, 1992.