SYSTEM IDENTIFICATION: TIME VARYING AND NONLINEAR METHODS

A Dissertation

by

MANORANJAN MAJJI

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2009

Major Subject: Aerospace Engineering

SYSTEM IDENTIFICATION: TIME VARYING AND NONLINEAR METHODS

A Dissertation

by

MANORANJAN MAJJI

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	John L. Junkins
Committee Members,	Aniruddha Datta
	Jer-Nan Juang
	Daniele Mortari
	Othon K. Rediniotis
	Srinivas R. Vadali
Head of Department,	Dimitris C. Lagoudas

May 2009

Major Subject: Aerospace Engineering

ABSTRACT

System Identification: Time Varying and Nonlinear Methods. (May 2009)

Manoranjan Majji, B.E. , Birla Institute of Technology and Science;

M.S., Texas A&M University

Chair of Advisory Committee: Dr. John L. Junkins

Novel methods of system identification are developed in this dissertation. First set of methods are designed to realize time varying linear dynamical system models from input-output experimental data. The preliminary results obtained in a recent paper by the author are extended to establish a new algorithm called the Time Varying Eigensystem Realization Algorithm (TVERA). The central aim of this algorithm is to obtain a linear, time varying, discrete time model sequence of the dynamic system directly from the input-output data. Important results relating to concepts concerning coordinate systems for linear time varying systems are developed (discrete time theory) and an intuitive understanding of equivalent realizations is provided. A procedure to develop first few time step models is detailed, providing a unified solution to the time varying identification problem.

The practical problem of identifying the time varying generalized Markov parameters required for TVERA is presented as the next result. In the process, we generalize the classical time invariant input output AutoRegressive model with an eXogenous input (ARX) models to the time varying case and realize an asymptotically

stable observer as a byproduct of the calculations. It is further found that the choice of the generalized time varying ARX model (GTV-ARX) can be set to realize a time varying dead beat observer.

Methods to use the developed algorithm(s) in this research are then considered for application to the identification of system models that are bilinear in nature. The fact that bilinear plant models become linear for constant inputs is used in the development of an algorithm that generalizes the classical developments of Juang.

An intercept problem is considered as a candidate for application of the time varying identification scheme, where departure motion dynamics model sequence is calculated about a nominal trajectory with suboptimal performance owing to the presence of unstructured perturbations. Control application is subsequently demonstrated.

The dynamics of a particle in a rotating tube is considered next for identification using the time varying eigensystem realization algorithm. Continuous time bilinear system identification method is demonstrated using the particle example and the identification of an automobile brake model.

DEDICATION

To all my teachers and mentors, in particular, Professor John Junkins, for his exceptional mentorship, exemplary scholarship and extraordinary leadership! To the memory of late Dr. Kyong B. Lim, a great colleague we lost to cancer.

ACKNOWLEDGEMENTS

The author wishes to express his gratitude to his advisor, Dr. John L. Junkins, whose encouragement, tutelage, motivation, guidance and support made this dissertation possible. Having taken constant inspiration from his advisor, the author is highly indebted to Dr. Junkins. It is indeed a great honor to have shared some subspace of research activity with him and the author cannot find words to express his inordinate appreciation.

The author feels greatly blessed for having come into the sphere of intellectual influence of several bright scholars in various areas of his interest. Interaction and collaboration with them enriched his graduate studies at A&M, providing a well rounded experience.

Dr. Jer-Nan Juang and the author have shared some memorable research collaborations in the process of conducting the research underlying this dissertation. The author would like to express his gratitude to Dr. Juang for all the motivation, guidance, interaction, insight, training and collaboration he continues to provide the author, with exceptional humility. Dr. Juang is also gratefully acknowledged for his warm hospitality while hosting the author, visiting him in Taiwan. The author anticipates continuing collaborative research in system identification (and mountain climbing) with Dr. Juang, and looks forward to more break-throughs (but no broken bones).

Dr. James D. Turner has been a major influence on the author, providing him great advice, support, guidance and insight on several joint research topics. His scholarship, drive and consistent motivation have often been inspirational to the author. The author is grateful to Dr. Turner for his generous spirit during our collaborations.

Dr. Daniele Mortari is warmly acknowledged for the love, kindness and affection showered by him on the author.

Mrs. Elouise Junkins has dearly cheered the author along with Dr. John Junkins on various phases of his academic development. Mrs. Junkins also marked the happy occasion of the author's successful defense with a fabulous party at the Junkins residence. The author extends his warm appreciation to Elouise. Andrea, Anna and Daniele have given a feeling of home – away from home. Grazie Mille.

Dr. Srinivas R. Vadali influenced the author through his excellent teaching, deep devotion to dynamics and control and ideal scholarship. The author greatly appreciates the inspiration from Dr. Vadali, as he nurtures research rigor and analytical acumen through his gifted mentorship to all aerospace students who study dynamics and control with him.

Dr. Othon Rediniotis helped the author hone his experimental abilities and imparted key elements of aerodynamics theory and experimental research methods. The author extends his gratitude to Dr. Rediniotis for all his help.

The author further extends his thanks to Dr. Aniruddha Datta for teaching him various elements of systems and control theory in three different classes.

The author is delighted to warmly acknowledge the following colleagues: Dr. Maruthi R. Akella, Dr. Kyle T. Alfriend, Dr. Hyochoong Bang, Dr. Raktim Bhattacharya, Dr. Robert H. Bishop, Dr. Christian Bruccoleri, Dr. Suman Chakravorty, Dr. Yang Cheng, Dr. Daniel Choukroun, Dr. John L. Crassidis, Dr. James R. Fisher, Dr. Adam Fosbury, Dr. John E. Hurtado, Dr. Anup Katake, Dr. Richard W. Longman, Dr. F. Landis Markley, Dr. Yaakov Oshman, Dr. Lawrence "Robbie" Robertson, Dr. Hanspeter Schaub, Dr. Prasenjit Sengupta, Dr. Rajnish Sharma, Dr. Puneet Singla, Dr. Kamesh Subbarao, Dr. Monish Tandale, Dr. John Valasek, Dr. Renato Zanetti. These individuals are gratefully acknowledged for their motivation, mentorship and guidance through several technical discussions, insightful suggestions and especially for what the author has learnt from their exemplary scholarship on various topics of interest, throughout my graduate studies.

The author was also influenced and motivated by the writings of Dr. Thomas Kailath, Dr. Leonard Meirovitch, Dr. Arthur E. Bryson, Jr. and (late) Dr. Gene H. Golub. Many thanks are due to these distinguished senior colleagues for their significant efforts in fostering the interests of young researchers in general and the author in particular.

The following graduate student colleagues: Ms. Xiaoli Bai, Ms. Lesley Wietz, Mr. James Doebbler, Mr. Jeremy Davis, Mr. Troy Henderson, Mr. Mrinal Kumar, Mr. Bong Su Koh and Mr. Drew Woodbury have been great friends and collaborators during the author's graduate studies at Texas A&M University. Thanks much for your continuing great camaraderie.

The author is highly indebted to an important member of the dynamics and control family at A&M, Ms. Lisa Willingham. Her exceptional efficiency and hard work have always been a source of support and inspiration to the author. The author wishes to thank Ms. Willingham for her kind help in making administrative things happen.

Acknowledgements are also due to Ms. Karen Knabe for her kindness and affection towards graduate students in The Aerospace Engineering Department.

The author wants to extend his gratitude to the following organizations for their financial support during his graduate studies. Texas Institute for Intelligent Bio-Nano Materials and Structures for Aerospace Vehicles (TIIMS, a NASA University Research, Engineering and Technology Institute), Texas A&M University, the Consortium for Autonomous Space Systems (CASS), funded by the Air Force Research Laboratory and Coordinated by Dr. Lawrence “Robbie” Robertson, endowments from the George J. Eppright Chair in Engineering and the Royce E. Wisenbaker Chair in Engineering, the John V. Breakwell student travel award (administered by the American Astronautical Society) and the student grant from the F. Landis Markley Astronautics Symposium organizing committee (monitored by Dr. John L. Crassidis).

The author thanks his parents Mrs. Sujatha Majji and Mr. Siva Rama Krishna Prasad Majji for showing all the patience, faith and love in the world for him. His sister Ms. Sai Majji, brother-in-law, Dr. Anuj Sharma and cousin Mr. Koteswara Rao Majji have been very supportive by being really close friends. The author is grateful for their support and love.

NOMENCLATURE

ERA	Eigensystem Realization Algorithm
OKID	Observer/Kalman Filter Identification
ARX	Autoregressive Model with Exogenous Input
ARMA(X)	Autoregressive Moving Average Model (with Exogenous Input)
GTVARX	Generalized Time Varying Autoregressive Model with Exogenous Input
TVERA	Time Varying Eigensystem Realization Algorithm
TOKID	Time Varying Observer/Kalman Filter Identification
$\mathbf{a}, \mathbf{x}$	Bold face small case letters denote vectors of a specified state dimension
A, X, Φ	Normal face letters are used to denote matrices of compatible dimensions, including scalars. Distinction between matrices and scalars is made in accordance with the context
$(\ )^{\dagger}$	Dagger is particularly reserved for the Moore-Penrose pseudo inverse of a rectangular matrix

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

FIGURE Page

LIST OF TABLES

TABLE Page

CHAPTER I

INTRODUCTION: MODERN SYSTEM IDENTIFICATION

Mathematical modeling has emerged as a pivotal tool in modern engineering theory and practice. The explosive increase in the computational power has been the central reason behind this in-progress modeling revolution. The Apollo program stands as an important testimony of one of the greatest model-based engineering miracles in history and represents a landmark demonstration of advanced technology in general and in dynamical system modeling and applications in particular. However, the Apollo program was actually the beginning rather than the culmination of modern methods of modeling (including system identification). Primarily, mathematical models of dynamical systems are of analytical and computational in nature and represent the physics of the components of the systems being modeled. The utility of the system models developed for analysis is either to carry out measurement-based study of some intrinsic properties of the dynamical system (called estimation theory) or to examine the interaction of the system under investigation with external influence functions (e.g., external generalized force profiles, magnetic field interactions, etc., known as control theory).

The fidelity of the model used for estimation or control purposes is directly proportional to the level of accuracy achieved in the solutions (for observation/control applications), and model fidelity can be adjusted (simple models usually suffice) to meet

---

This dissertation follows the style of the *Journal of Guidance Control and Dynamics*.

the needs of particular applications. Therefore one typically makes an appropriate judgment on the level of accuracy needed depending on the available computational or analytical resources and modeling capabilities.

System Identification is the branch of mathematical system theory, which deals with the process of constructing differential equation and/or difference equation model(s) of a dynamical system, whose input forces and the sensor outputs are available for measurement. Thus, System Identification is an "inverse problem" where the model is derived from measurements. This is clearly related to the large literature on mathematical modeling, wherein the model is derived from first principles and with assumptions on the system's physical geometry, mass properties, constitutive laws, environmental forces and so on. The process of system identification is depicted in Figure 1. The models thus realized directly from input-output experimental data, are frequently found directly useful for control and estimation purposes. Owing to the emergence of model based control and estimation strategies in modern system theory, coupled with the eternal presence of model error (no matter how one obtains the model), system identification has occupied a center stage in the recent developments of dynamics and control.

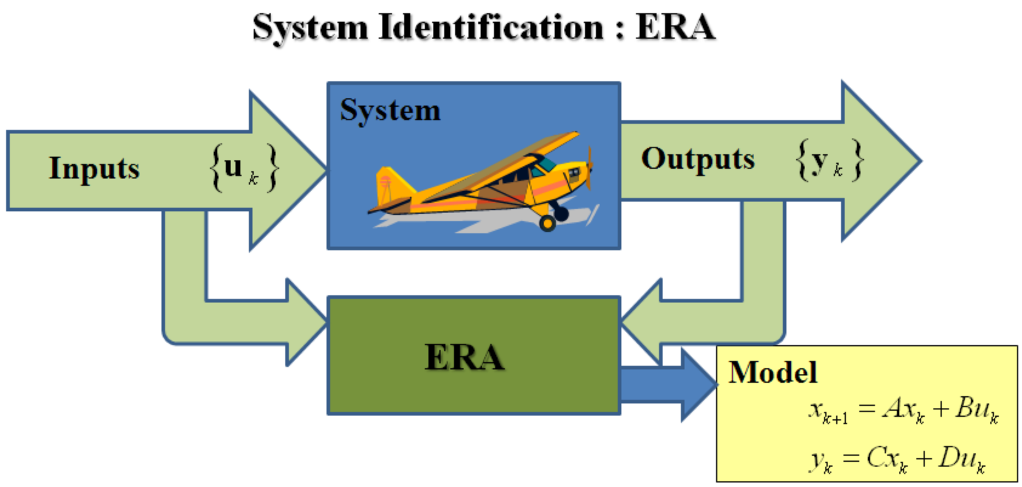

Figure 1. System Identification Process

The methods and ideas central to realize linear time invariant models from input output data are now very well understood and documented widely [1-3] owing to extensive research in this area for the past few decades. An important member of this class of system identification methods is the Eigensystem Realization Algorithm[1]. Key ideas of this popular algorithm are summarized in Figure 2.

ERA: Key Ideas

Input/Output Relationship

$Y_i := CA^{i-1}B -$ Markov Parameters

$$\mathbf{y}_k = CA^k\left(\mathbf{x}_0 = \mathbf{0}\right) + \sum_{i=1}^{k} CA^{i-1}B\mathbf{u}_{k-i} + D\mathbf{u}_k = \sum_{i=1}^{k} Y_i\,\mathbf{u}_{k-i} + D\mathbf{u}_k$$

Hankel Matrix

$$H_{k-1} := \begin{bmatrix} Y_k & Y_{k+1} & \cdots & Y_{k+\beta-1} \\ Y_{k+1} & Y_{k+2} & \cdots & Y_{k+\beta} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{k+\alpha-1} & Y_{k+\alpha} & \cdots & Y_{k+\alpha+\beta-1} \end{bmatrix} = P_\alpha A^{k-1} Q_\beta$$

$$Q_\beta = \begin{bmatrix} B & AB & A^2B & \cdots & A^{\beta-1}B \end{bmatrix}$$

Controllability Grammian

$$P_\alpha = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{\alpha-1} \end{bmatrix}$$

Observability Grammian

$$\boxed{H_0 = R\Sigma S^T = P_\alpha Q_\beta \Rightarrow P_\alpha = R\Sigma^{1/2}, Q_\beta = \Sigma^{1/2}S^T}$$

$$\boxed{H_1 = P_\alpha A Q_\beta} \Rightarrow \boxed{\begin{aligned} \hat{A} &= P_\alpha^\dagger H_1 Q_\beta^\dagger = \Sigma^{-1/2}R^T H_1 S\Sigma^{-1/2} \\ \hat{B} &= \Sigma^{1/2}S^T E_r, \hat{C} = E_m^T R\Sigma^{1/2} \end{aligned}}$$

Figure 2. Key Ideas of ERA

The significant effort of researchers in aerospace engineering in implementing precise control and estimation strategies for flexible spacecraft structures, helped refine the methods of system identification for mechanical system models and make important connections with classical methods of modal analysis[1, 4]. Although several efforts to extend the now classical techniques (time invariant theory) to realize time varying discrete time state space models have been reported in the past[5-7], scope of the available methods for time varying system identification remains limited[8]. This is mostly due to the lack of a consistent theoretical foundations and computational algorithms.

It was found in this research project, that a consistent computational algorithm could not be formulated because of some gaps in implementing the incomplete theoretical ideas formulated by the researchers in the past. The investigations carried out, leading to this dissertation were aimed at ameliorating the practical difficulties and formulating a consistent algorithm.

The outline of the dissertation is as follows. Chapter II presents the details on a novel algorithm and computational procedure for identification of time varying discrete time plant model sequence sets from measured input/output data. This is followed in Chapter III by the extension of the classical OKID algorithm to calculate the generalized Markov parameters for the time varying discrete time plant model sequence sets that are needed in the algorithms developed in the second chapter. Some new results for identification of nonlinear systems with bilinearity in the plant model dynamics are presented in chapter IV. Chapter V applies the methods developed in chapters II and III to problems in guidance and dynamics. Control designs and simulations are carried out using the models realized by using the time varying identification algorithms. Conclusions are presented in the chapter VI along with the new research direction opportunities created as a consequence of the results presented in this dissertation. Three appendices are included to present relevant results supporting the main contents of the chapters outlined above. The first appendix presents an innovations process derivation of the classical Kalman filter equations to aid in the qualitative relationship discussions involved in the time varying extension of the OKID relations. In appendix B, the author details at length the definition and properties of time invariant and time varying deadbeat observers. Numerical example realizes such an observer as a by-product of the OKID procedure, using a generalized time varying ARX (GTV-ARX) model in this appendix. Third appendix summarizes some tools required for the bilinear system identification algorithm of chapter IV.

CHAPTER II

TIME VARYING EIGENSYSTEM REALIZATION ALGORITHM

Introduction

The Eigensystem Realization Algorithm (ERA)[1, 9, 10] has occupied the center stage in the current system identification theory and practice, owing to its ease, efficiency and robustness of implementation in several spheres of engineering. Connections of ERA with modal and principal component analyses made the algorithm an invaluable tool for the analysis of mechanical systems. As a consequence, the associated algorithms have contributed to several successful applications in design, control and model order reduction of mechanical systems. ERA is the member of a class of algorithms derived from system realization theory based on the now classical Ho-Kalman method[3]. Since both left and right singular vector matrices of the singular value decomposition are utilized, ERA is in fact a modest generalization of the subspace methods and as a consequence yields state space realizations that are not only minimal but also balanced[1]. The key utility of ERA has been in the development of discrete time invariant models from input output experimental data. Owing to the one-to-one mapping of linear time invariant dynamical system models between the continuous and discrete time domains, the ERA identified discrete time model is tantamount to the identification of a continuous time model (with the standard assumptions on the sampling theorem). Furthermore, the physical parameters of a mechanical system (natural frequencies, normal modes and damping) can be derived from the identified plant models by using ERA. A variety of system identification methods for such time invariant systems are available, the fundamental unifying features of which are now well understood [2, 11, 12] and can be shown to be related (and/or equivalent) to the corresponding features of ERA.

Several efforts were undertaken in the past to develop a holistic approach for the identification of time varying systems. Specifically, it has been desired for some time to generalize ERA to the case of time varying systems. Earliest efforts in the development of methods for time varying systems involved recursive and fast implementations of the time invariant methods by exploring structural properties of the input/output realizations. The classic paper by Chu et. al, exploring the displacement structure in the Hankel matrices is representative of the efforts of this nature. Subsequently, significant results were obtained by Shokoohi and Silverman [6] and Dewilde and Van der Veen[5], that generalized several concepts in the classical linear time invariant system theory consistently. Verhaegen and coworkers ([7, 13]) subsequently introduced the idea of repeated experiments (termed ensemble I/O data), rendering practical methods to realize the conceptual identification strategies presented earlier. These methods are referred to as ensemble state space model identification problems in the literature. This class of generalized subspace based methods was applied to complex problems such as the modeling the dynamics of human joints, with much success. Liu [8] developed a methodology for developing time varying models from free response data (for systems with an asymptotically stable origin) and made initial contributions to the development of time varying modal parameters and their identification[14].

Although the effects of time varying coordinate systems are shown to exist by these classical developments, it is not clear if the identified plant models (more generally identified model sequence sets) are useful in state propagation. This is because no guarantees are given as to whether the system matrices identified are in fact, all realized in the same coordinate system. This limits the utility of the classical solutions since model sequences identified by different procedures cannot be merged as the sequences would loose compatibility at the time instance at which the algorithm is switched.

In other words, most classical results developed realized models that are topologically equivalent (defined mathematically in subsequent sections) from an input output stand point. However this does not imply that they are in coordinate systems consistent in time, for state propagation purposes. It is straightforward to see that the initial state given in a certain coordinate system cannot be propagated to the next time step unless the state transition and control influence matrices are expressed in the same (or compatible) coordinate system as the initial state of interest. Any misalignment would cause the state propagation to be physically meaningless and the identified plant model(s) are rendered useless.

We cannot emphasize more on the importance of the coordinate transformations and their role in time varying systems. As a practical example of this important feature underpinning the developments of the current chapter, let us consider the following situation. It is not too difficult to consider a version of the method proposed by Liu[8] to obtain the first few time step models. This could in-principle be merged with the plant model sequence realized by using the classical developments of Shookohi and

Silverman[6] (or equivalently Verhaegen and Yu [7]). The fact that the plant model sequences identified appropriately in such a manner, would be incompatible at the junction (discrete time instant) of merger, making both the procedures incomplete. Looking at the facts more transparently, following Liu[8] alone, we would not have the control influence matrix sequence (and the formulations there-in are restricted to plants with an asymptotically stable origin) and alternatively following Shookohi (and others [6]), we would never be able to identify the first few time step (and last few time step) models since negative time indexing is not possible in general. However, following the developments of this chapter, one could indeed realize the complete model sequence without invoking the negative time step experimental data or assuming asymptotic stability of the origin. Furthermore, unlike the preliminary developments of coordinate transformations by Liu[8], the solutions presented here-in are compatible (give back the generalized Markov parameters indicating the arbitrariness of the transformations) and in general valid for the practical case of the number of outputs being less than the state dimension.

In contrast, the methods developed in this chapter arise from a perspective of generalizing the classical Ho-Kalman approach to the case of time varying systems, while utilizing the notation and preliminary developments of past researchers[6-8] on this problem. It is shown that the generalization thus made enables us to identify time varying plant models that are in arbitrary coordinate systems at each time step. Furthermore, the coordinate systems at successive time steps are compatible with one another. This makes the model sequences realized, useful in state propagation. The computational methods of generalized Markov parameters using the input output map are subsequently discussed for the two cases of the presence and absence of zero input response data in the output sequences. This is followed by a discussion on a computational procedure to determine the time varying coordinate transformations with respect to a fixed time step, t_k (most times initial condition time step) using free response experimental data. Numerical examples demonstrating the theoretical developments conclude the chapter.

Linear Discrete Time Varying System Realization Theory

We review the notation and definitions in linear time varying systems following the developments presented in the classic paper by Shokoohi and Silverman[6]. Linear discrete time varying systems are governed by a set of difference equations governing the evolution of the state in time being given by

$$\mathbf{x}_{k+1} = A_k \mathbf{x}_k + B_k \mathbf{u}_k \tag{2.1}$$

together with a corresponding initial state vector $\mathbf{x}_0$. The state variable $\mathbf{x}_k \in \mathbb{R}^n$ is most often related to the output by the measurement equation,

$$\mathbf{y}_k = C_k \mathbf{x}_k + D_k \mathbf{u}_k \tag{2.2}$$

with the outputs and inputs being $\mathbf{y}_k \in \mathbb{R}^m, \mathbf{u}_k \in \mathbb{R}^r$. Together with, $A_k \in \mathbb{R}^{n\times n}$, $B_k \in \mathbb{R}^{n\times r}$, $C_k \in \mathbb{R}^{m\times n}$ and $D_k \in \mathbb{R}^{m\times r}$ being in compatible dimensions. In the following developments, it is assumed that the true state dimension n is constant throughout the time period of interest. It will be transparent in the course of our developments that this assumption could be relaxed but we retain it to facilitate some coordinate transformation results for the special class of mechanical systems; important in applications. The solution of the difference equation relating the initial state and the control inputs to the state at a general time instant is given by

$$\mathbf{x}_k = \Phi\left(k, k_0\right)\mathbf{x}_0 + \sum_{j=k_0}^{k-1} \Phi\left(k, j+1\right) B_j \mathbf{u}_j \tag{2.3}$$

where the state transition matrix is defined in terms of its components by

$$\Phi\left(k, k_0\right) = \begin{cases} A_{k-1}A_{k-2}\ldots A_{k_0}, & \forall k > k_0 \\ I, & k = k_0 \\ \text{undefined}, & \forall k < k_0 \end{cases} \tag{2.4}$$

Using the definition of the compound state transition matrix, the input output relationship is given by,

$$\mathbf{y}_k = C_k \Phi\left(k, k_0\right)\mathbf{x}_0 + \sum_{j=k_0}^{k-1} C_k \Phi\left(k, j+1\right) B_j \mathbf{u}_j + D_k \mathbf{u}_k \tag{2.5}$$

This enables us to define the input output relationship in terms of the two index coefficients as,

$$\mathbf{y}_k = C_k \Phi\left(k, k_0\right)\mathbf{x}_0 + \sum_{j=k_0}^{k-1} h_{k,j} \mathbf{u}_j + D_k \mathbf{u}_k \tag{2.6}$$

where the generalized Markov parameters are defined to be given by,

$$h_{k,i} = \begin{cases} C_k \Phi\left(k, i+1\right) B_i, & \forall i < k-1 \\ C_k B_{k-1}, & i = k-1 \\ 0, & \forall i > k-1 \end{cases} \tag{2.7}$$

In stark contrast to the time invariant (shift invariant) systems, the generalized Markov parameters determine the impulse response characteristics of the true plant in a much more general fashion. Note that the number of independent degrees of freedom to describe the input-output relationship increases tremendously for the case of time varying systems, as the response of the system ($h_{k,i}$) not only depends upon the time difference from the applied input ($\mathbf{u}_i$) but also on the time instant at which the said input is applied (i, t_i).

Similar to the time invariant case, the time varying discrete time systems, when expressed as the input – output map, are invariant to coordinate (similarity) transformations. In fact, the generalized Markov parameters we defined above are invariant to a more general set of transformations called the Lyapunov transformations[5-7, 15]. Using these Lyapunov transformations, several equivalent state space realizations can be obtained. We briefly introduce the Lyapunov transformations in this section. We will use several notions being introduced here, in the subsequent sections (and chapters); particularly, while constructing projection maps to transform all the time varying coordinate systems in to a reference coordinate system.

Following the notions set up by Shokoohi and Silverman[6], the system representation $\left\{A_k, B_k, C_k, D_k\right\}$ is said to be *topologically equivalent* (Gohberg et. al., [16] call this equivalence, *Kinematic Similarity*) to the representation $\left\{\bar{A}_k, \bar{B}_k, \bar{C}_k, \bar{D}_k\right\}$ if there exists a sequence of invertible, square matrices (not necessarily related to each other, Lyapunov transformations), $\{T_k\}$ such that, $D_k = \bar{D}_k$ and,

$$\bar{A}_k = T_{k+1}^{-1} A_k T_k \tag{2.8}$$

$$\bar{B}_k = T_{k+1}^{-1} B_k \tag{2.9}$$

$$\bar{C}_k = C_k T_k \tag{2.10}$$

It is easy to see that the state transition matrices have relationship similar to (2.8) and that all topologically equivalent representations give the same numerical value for the generalized Markov parameters owing to their definition in (2.7). Controllability and Observability grammians are given by the infinite matrices,

$$O_k := \begin{bmatrix} C_k \\ C_{k+1} A_k \\ \vdots \\ C_{k+p} A_{k+p-1} \ldots A_k \\ \vdots \end{bmatrix} \tag{2.11}$$

and

$$R_k = \begin{bmatrix} B_k & A_k B_{k-1} & \ldots \end{bmatrix} \tag{2.12}$$

Although the grammians are infinite matrices, usually for a system which is both Controllable and Observable (minimal), the principal full rank components of the corresponding grammians have most information related to the plant parameters corresponding to the current time step. This fact enables us to construct the time varying realizations without resorting to population of infinite matrices, a central idea of this chapter and a key algorithmic contribution of this dissertation. The Controllability and

Observability grammians transform with topologically equivalent system descriptions of the time varying systems. The relationships between topologically equivalent representations are given by

$$\bar{O}_k := \begin{bmatrix} \bar{C}_k \\ \bar{C}_{k+1}\bar{A}_k \\ \vdots \\ \bar{C}_{k+p}\bar{A}_{k+p-1}\ldots\bar{A}_k \\ \vdots \end{bmatrix} = \begin{bmatrix} C_k \\ C_{k+1}A_k \\ \vdots \\ C_{k+p}A_{k+p-1}\ldots A_k \\ \vdots \end{bmatrix} T_k = O_k T_k \tag{2.13}$$

and

$$\bar{R}_k = \begin{bmatrix} \bar{B}_k & \bar{A}_k\bar{B}_{k-1} & \ldots \end{bmatrix} = T_{k+1}^{-1}\begin{bmatrix} B_k & A_k B_{k-1} & \ldots \end{bmatrix} = T_{k+1}^{-1}R_k \tag{2.14}$$

Similar relations hold for block shifted controllability and observability grammians, which can be easily derived as,

$$\begin{aligned} \bar{O}_k^{\uparrow} &:= \begin{bmatrix} \bar{C}_{k+1}\bar{A}_k \\ \bar{C}_{k+2}\bar{A}_{k+1}\bar{A}_k \\ \vdots \\ \bar{C}_{k+p}\bar{A}_{k+p-1}\ldots\bar{A}_k \\ \vdots \end{bmatrix} \\ &= \begin{bmatrix} C_{k+1}A_k \\ C_{k+2}A_{k+1}A_k \\ \vdots \\ C_{k+p}A_{k+p-1}\ldots A_k \\ \vdots \end{bmatrix} T_k \\ &= O_k^{\uparrow} T_k \end{aligned} \tag{2.15}$$

and

$$
\begin{aligned}
\overline{R}_k^{\leftarrow} &= \begin{bmatrix} \overline{A}_k \overline{B}_{k-1} & \overline{A}_k \overline{A}_{k-1} \overline{B}_{k-2} & \dots \end{bmatrix} \\
&= T_{k+1}^{-1} \begin{bmatrix} A_k B_{k-1} & A_k A_{k-1} B_{k-2} & \dots \end{bmatrix} \\
&= T_{k+1}^{-1} R_k
\end{aligned} \tag{2.16}
$$

The generalized Hankel matrix for time varying systems is defined for every time step k to be the infinite dimensional matrix,

$$
\begin{aligned}
H_k &= \begin{bmatrix} h_{k,k-1} & h_{k,k-2} & \cdots \\ h_{k+1,k-1} & h_{k+1,k-2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \\
&= \begin{bmatrix} C_k \\ C_{k+1} A_k \\ \vdots \end{bmatrix} \begin{bmatrix} B_{k-1} & A_{k-1} B_{k-2} & \cdots \end{bmatrix} \\
&= O_k R_{k-1} \\
&= \overline{O}_k \overline{R}_{k-1}
\end{aligned} \tag{2.17}
$$

In general, assuming the system is uniformly observable and controllable, rank of the generalized Hankel matrix is representative of the state dimension at the given time instant. In the subsequent developments of the dissertation, it is assumed that the state dimension does not change with the time index. It is not difficult to see that this assumption can be relaxed. However, we retain the assumption owing to our focus on mechanical systems, where the connection between physical degrees of freedom and the number of state variables allows us to hold the dimensionality of the state space fixed throughout the time interval of interest. We now elaborate on the time varying coordinate systems for discrete time state equations and some identities governing their structure and properties.

Time Varying Coordinate Systems and Transformations

As was noted from the previous sections, the state propagation for linear time varying systems takes place between time varying coordinate systems. This is very similar to the concept of body fixed rotating reference frames employed to describe rigid body rotation in attitude dynamics[16, 17]. Using the notation developed thus far; consider the state propagation equations in two topologically equivalent realizations of the discrete time varying system. The states being propagated in the equivalent realizations are related by the time varying transformations $\mathbf{z}_k = T_k \mathbf{x}_k$. When the corresponding state evolution equations are written as,

$$\begin{aligned} \mathbf{x}_{k+1} &= A_k \mathbf{x}_k + B_k \mathbf{u}_k \\ \mathbf{y}_k &= C_k \mathbf{x}_k + D_k \mathbf{u}_k \end{aligned} \tag{2.18}$$

and

$$\begin{aligned} \mathbf{z}_{k+1} &= F_k \mathbf{z}_k + G_k \mathbf{u}_k \\ \mathbf{y}_k &= H_k \mathbf{z}_k + D_k \mathbf{u}_k \end{aligned} \tag{2.19}$$

Relationships between the topologically equivalent realizations presented above are considerably different from time invariant systems. We rewrite the relations between the topologically equivalent realizations as,

$$\begin{aligned} F_k &= T_{k+1}^{-1} A_k T_k \\ G_k &= T_{k+1}^{-1} B_k \\ C_k &= H_k T_k \end{aligned} \tag{2.20}$$

The most important distinction is that the system matrices (transition matrices F_k, A_k) do not have the same eigenvalues. Since the system evolution takes place in two different coordinate systems, T_{k+1}, T_k , it does not map the state at the next time step back in to the same state space. This leads the basis vectors for the initial time step and the final time step to be different. Therefore, the situation is quite similar to body fixed, rotating coordinate systems in rigid body dynamics, with the exception that the frames (basis vectors can be thought of as frames) are unknown, arbitrarily assigned by the singular value decomposition (we will see very shortly) and not necessarily orthogonal.

A clear picture of this situation appears in the state transition matrices of continuous time varying systems. To clarify this point we digress at this stage to consider the linear time varying homogeneous system given by the continuous time linear differential equation,

$$\dot{\xi}(t) = \Lambda(t)\xi(t) \tag{2.21}$$

with initial conditions, $\xi(t_0) = \xi_0$ and $\xi(t) : \mathbb{R}^+ \to \mathbb{R}^n$, $\Lambda(t) : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^n$. Then for every initial state $\xi_i(t_0)$, ($i = 1, 2, ..., n$) spanning the state space at initial time, there exists a solution at final time $t(\neq t_0)$, denoted by $\xi_i(t)$. Collecting these solutions in to a matrix $\Psi(t) := \begin{bmatrix} \xi_1(t) & \xi_2(t) & \cdots & \xi_n(t) \end{bmatrix}$, we arrive at the fundamental matrix[15, 18]. Since it constitutes the linearly independent (arbitrary) solutions of the state differential equation, the fundamental matrix satisfies the matrix differential equation,

$$\dot{\Psi}(t) = \Lambda(t)\Psi(t) \tag{2.22}$$

with initial conditions, $\Psi\left(t_0\right)=\Psi_0:=\left[\xi_1\left(t_0\right) \quad \xi_2\left(t_0\right) \quad \cdots \quad \xi_n\left(t_0\right)\right]$. Note that this is not necessarily the $n\times n$ identity matrix, thereby defining the non-orthogonal, skew coordinate frame at the initial (in general at time step t_k) time step.

It can be shown that this fundamental matrix is related to the state transition matrix $\Phi\left(t,t_0\right)$ as,

$$\Phi\left(t,t_0\right)=\Psi\left(t\right)\Psi^{-1}\left(t_0\right) \tag{2.23}$$

where the classical state transition matrix [4, 19] is governed by the matrix differential equation,

$$\frac{\partial}{\partial t}\Phi\left(t,t_0\right)=\Lambda\left(t\right)\Phi\left(t,t_0\right) \tag{2.24}$$

with initial conditions, $\Phi\left(t,t_0\right)=I_n$ as the unit (identity) matrix. If the identified discrete time state transition matrices A_k were all constrained to be the state transition matrices, the coordinate systems of solutions would indeed be compatible. However, this rarely happens and the realized (identified) $\hat{A}_k$ are in fact the more general fundamental matrices (defined in equation (2.22)) with arbitrary sets of initial conditions, due to the arbitrary decompositions of the singular value decomposition, detailed in the next section.

Realizing that the solution structure at any time t is given by $\mathbf{x}\left(t\right)=\Phi\left(t,t_0\right)\mathbf{x}_0$. We point out a stark contrast to time invariant system ($\Lambda\left(t\right)=A_c$), where the state transition matrix is given by $\Phi\left(t,t_0\right)=\exp\left[A_c\left(t-t_0\right)\right]$; noting further that the solution in the time invariant case remains in the same space owing to the power series expansion definition of the matrix exponential. More specifically, this is the space spanned by $\left(I, A_c, A_c^2, ...\right)\mathbf{x}_0$. Such a definition/parallelism cannot be made for time varying systems and hence the state transition matrix, as given by equation (2.23) maps the state in one coordinate system at initial time step t_0 to a possibly (usually) different coordinate system at any subsequent time t. Therefore, it emerges conclusively that the true and identified system matrices in our current discussions are special instances of the fundamental matrices outlined in equation (2.22) with an arbitrary set of basis functions at the corresponding initial time step.

This simple observation is evidently new and of fundamental importance in establishing a complete system identification algorithm for time varying systems. The ideas presented in the section are graphically illustrated in Figure 3, where a 3 dimensional state space is assumed for clarity in demonstration.

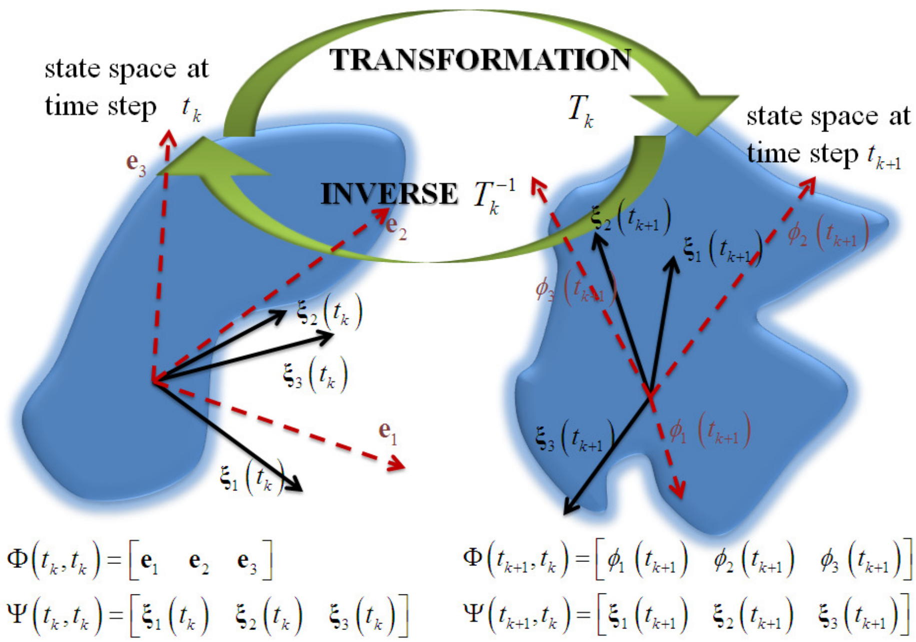

Figure 3. Illustration of Time Varying Coordinate Systems and Transformations

We note in passing that the topologically equivalent realizations, (F_k , A_k related as in (2.20)) in general, are not similar, owing to the fact that $T_{k+1} \neq T_k$. An analyst armed with this piece of information (that the state evolution of discrete time varying systems in general takes place between time varying coordinate systems (different spaces)), is often dangerous. She/ he may conclude that no physics - based information can be derived from such a method, since there appears to be no such information. It turns out that such a speculation is erroneous and one can indeed extract time varying quantities that are representative of the true time varying system behavior from these topologically equivalent (kinematically similar) transformations. These parameters are the eigenvalues of the time varying system matrices (true and identified), all transformed in to a common reference (more generally, projected onto) coordinate system. This is the central result of this chapter. We now detail the procedure to construct such transformations on the topologically equivalent discrete time varying realizations (2.18),(2.19). Applying the general relationship, (2.39) between observability grammians in different coordinate systems to the realizations, (2.18),(2.19), we have that,

$$O_k^{T_k} = O_k T_k \tag{2.25}$$

At any other time step, the same relationship holds, given by $O_{k+p}^{T_{k+p}} = O_{k+p} T_{k+p}$, $\forall p \geq 1$. This enables us to define the quantity,

$$\left(O_k^{T_k}\right)^{\dagger} O_{k+p}^{T_{k+p}} = T_k^{-1} O_k^{\dagger} O_{k+p} T_{k+p} \tag{2.26}$$

$\forall p \geq 1$ where, the identity $\left(O_k^{T_k}\right)^{\dagger} = \left(O_k T_k\right)^{\dagger} = T_k^{-1} O_k^{\dagger}$ was used. Considering the first time step t_k, the relation between the kinematically similar system matrices is given by,

$$F_k = T_{k+1}^{-1} A_k T_k \tag{2.27}$$

Now, we proceed to use the correction ($p=1$) to the left of (2.27) and obtain a corrected system matrix $\overline{F}_k$, as

$$\begin{aligned} \overline{F}_k := \left(O_k^{T_k}\right)^{\dagger} O_{k+1}^{T_{k+1}} \left(F_k\right) &= \left(T_k^{-1} O_k^{\dagger} O_{k+1} T_{k+1}\right) T_{k+1}^{-1} A_k T_k \\ &= T_k^{-1} O_k^{\dagger} O_{k+1} A_k T_k \\ &= T_k^{-1} \overline{A}_k T_k \end{aligned} \tag{2.28}$$

where $\overline{A}_k := O_k^{\dagger} O_{k+1} A_k$ is the correction to the time varying system matrix in the different coordinate system. Note that, at a general time step ($t_{k+p}, p \geq 1$) both the left hand side (T_{k+p+1}) and right hand side (T_{k+p}) coordinates need to be transformed in to the reference coordinate system (T_k in this case,) such that the corrected system matrices become similar. So, at any general time step, we have that,

$$F_{k+p} = T_{k+p+1}^{-1} A_{k+p} T_{k+p} \tag{2.29}$$

In such situations, we should operate on both sides to correct and obtain a transformation to the reference coordinate system (T_k in this case). This is accomplished by employing corrections on both sides given by,

$$\begin{aligned} \overline{F}_{k+p} &= \left(O_k^{T_k}\right)^{\dagger} O_{k+p+1}^{T_{k+p+1}} \left(F_{k+p}\right)\left(\left(O_k^{T_k}\right)^{\dagger} O_{k+p}^{T_{k+p}}\right)^{-1} \\ &= T_k^{-1} O_k^{\dagger} O_{k+p+1} T_{k+p+1} \left(T_{k+p+1}^{-1} A_{k+p} T_{k+p}\right)\left(T_k^{-1} O_k^{\dagger} O_{k+p} T_{k+p}\right)^{-1} \\ &= T_k^{-1} O_k^{\dagger} O_{k+p+1} A_{k+p} \left(O_k^{\dagger} O_{k+p}\right)^{-1} T_k \\ &= T_k^{-1} \overline{A}_{k+p} T_k \end{aligned} \tag{2.30}$$

We point out that the transformations developed above can also be based on the controllability grammian and are easy to derive, following the developments above. Note that in such a situation however, the reference coordinate system to which the system is reset (say some Q_k) is in general independent and different from the ones obtained by using the observability grammians (denoted here by T_k). So the least squares solution to realize a transformation, in general produces a projection on the controllable or observable subspace at the reference time instant. When the analyst has access to sufficient number of sensor outputs however (i.e., $m \geq n$), the solution becomes unique and the projections are made exactly on to the reference coordinates of the state space at the time instant of interest, similar to the solution by Liu[8], for this simpler situation.

For the system identification problem, as was derived in the previous section, the true and identified system parameters are kinematically similar realizations. The identified system matrices and the simulated "true" system should also be corrected for these physical variations to perform a comparison. This is because this true observability (controllability) correction aligns the true system in to the corresponding observability (controllability) subspace at the reference time instance. It was found that the system matrices appropriately corrected share common eigenvalues. Example demonstrations illustrate this fact. The physical nature of these eigenvalues and their role in the evolution of the true system and possible applications are issues that require further investigations.

Time Varying Eigensystem Realization Algorithm

We first present the algorithm to calculate the time varying plant parameter models assuming the availability of the generalized Markov parameters. The important problem of computing the generalized Markov parameters is addressed in the next section. A more practical algorithm for obtaining them is discussed in the next chapter which closely follows the developments of Majji and Junkins[20].

Calculation of Time Varying Discrete Time Models

Consider the generalized Hankel matrix populated using the generalized Markov parameters.

$$
\begin{aligned}
H_k &= \begin{bmatrix} h_{k,k-1} & h_{k,k-2} & \cdots \\ h_{k+1,k-1} & h_{k+1,k-2} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \\
&\approx \begin{bmatrix} h_{k,k-1} & h_{k,k-2} & \cdots & h_{k,k-p} \\ h_{k+1,k-1} & h_{k+1,k-2} & \cdots & h_{k+1,k-p} \\ \vdots & \vdots & \ddots & \vdots \\ h_{k+q,k-1} & h_{k+q,k-2} & \cdots & h_{k+q,k-p} \end{bmatrix} \\
&= H_k^{(p,q)}
\end{aligned}
\tag{2.31}
$$

with the parameters p, q chosen such that the generalized Hankel matrix retains the rank n, the true state dimension. Insight into what numbers must be chosen is often obtained by computing the rank of each element of the Hankel matrix sequence (at every time step). Differing ranks are possible for this generalized time varying Hankel matrix $H_k^{(p,q)}$ at every time step t_k, for the variable state dimension problem. For problems in which the state dimension does not change, rank consistency is indicative of the validity of our assumption of a constant state dimension. Extraneous observable disturbance input states are quite often isolated from the rank sequence plots of the generalized Hankel matrix sequence. Furthermore, rank consistency checking helps the analyst to retain appropriate numbers of row and column blocks in the Hankel matrix at a given time step for computations.

Following the identity (Eq. (2.17)) presented in the previous section where we discuss the generalized Hankel matrices and using its singular value decomposition[21, 22], we can write,

$$
H_k^{(p,q)} = O_k^{(q)T_k} R_{k-1}^{(p)T_k} = \left(U_k \Sigma_k^{\frac{1}{2}} \right) \left(\Sigma_k^{\frac{1}{2}} V_k^T \right) \tag{2.32}
$$

such that expressions can be written for the corresponding controllability and observability grammians at a given time step. Notice that this decomposition is non-unique. The fact that the realizations derived from these grammians can be in any of the infinite different coordinate systems (the coordinate systems also change with the order of controllability/observability grammian (p,q) chosen to be computed) is symbolized by using the superscript $(.)^{(p/q)T_k}$ on the grammian calculated by this particular decomposition (at time instant t_k). Using the generalized Markov parameters, now consider block up-shifted Hankel matrix defined as,

$$H_k^{(p,q)\uparrow} := \begin{bmatrix} h_{k+1,k-1} & h_{k+1,k-2} & \cdots & h_{k+1,k-p} \\ h_{k+2,k-1} & h_{k+2,k-2} & \cdots & h_{k+2,k-p} \\ \vdots & \vdots & \ddots & \vdots \\ h_{k+q+1,k-1} & h_{k+q+1,k-2} & \cdots & h_{k+q+1,k-p} \end{bmatrix} \tag{2.33}$$

and the left-shifted Hankel matrix defined as,

$$H_k^{(p,q)\leftarrow} := \begin{bmatrix} h_{k,k-2} & h_{k,k-3} & \cdots & h_{k,k-p-1} \\ h_{k+1,k-2} & h_{k+1,k-3} & \cdots & h_{k+1,k-p-1} \\ \vdots & \vdots & \ddots & \vdots \\ h_{k+q,k-2} & h_{k+q,k-3} & \cdots & h_{k+q,k-p-1} \end{bmatrix} \tag{2.34}$$

Considering the singular value decomposition and the definitions of the up-shifted Hankel matrix, we have that,

$$H_k^{(p,q)\uparrow} = \tilde{O}_k^{(q)\uparrow} \tilde{R}_{k-1}^{(p)} = \left(U_k^{\uparrow} \Sigma_k^{\uparrow\frac{1}{2}} \right) \left(\Sigma_k^{\uparrow\frac{1}{2}} V_k^{\uparrow T} \right) \tag{2.35}$$

Owing to the uniqueness of the singular value decomposition for a given matrix, the controllability grammian in equation (2.35) is different from the controllability grammian calculated from the unshifted Hankel matrix of equation (2.32). These grammians are similar, however, since the plant matrices $\tilde{A}_{k-1}, \tilde{B}_{k-1}$ of the grammian $\tilde{R}_{k-1}$ are similar to the plant matrices, A_{k-1}, B_{k-1} constituting R_{k-1}. Using a technique analogous to Juang [23] we calculate the similarity transformation to set $\tilde{O}_k^{\uparrow}$ in the same coordinate system as $O_k^{T_k \uparrow}$. Recall the transformation definitions from equations (2.13), (2.15), (2.14) and (2.16) that the similar grammians are related as,

$$\tilde{O}_k^{\uparrow} = O_k^{(q)T_k \uparrow} \tilde{Q}_k \tag{2.36}$$

and

$$\tilde{R}_{k-1} = \tilde{Q}_k^{-1} R_{k-1}^{(p)T_k} \tag{2.37}$$

Therefore the transformation matrix is calculated using the equation (2.37) as,

$$\tilde{R}_{k-1} R_{k-1}^{(p)T_k \dagger} = \tilde{Q}_k^{-1} \tag{2.38}$$

where the operator $(.)^{\dagger}$ denotes the pseudo inverse operation[21, 22]. Thus the block shifted observability grammian can be computed at every time step using the equations derived above. Similar calculations can be performed to yield the consistent expression for $R_k^{T_k \leftarrow}$.

Now, let the controllability and observability grammians in the true (usually unknown) coordinate systems at each time step t_k be denoted by the unadorned (no superscripts) symbols, O_k, R_{k-1}. Then, observability (controllability) grammian computed from the populated Hankel matrix is related to the "true" observability (controllability) grammian (of the unknown, true plant model) by,

$$O_k^{(q)T_k} = O_k T_k \qquad \left(R_{k-1}^{(p)T_k} = T_k^{-1} R_{k-1} \right) \tag{2.39}$$

where again, T_k is any invertible square matrix of the state dimension. Note that in problems of varying state dimension this matrix becomes rectangular and hence the coordinate transformations (projections) have to be appropriately defined. Again, we do not wish to include that case in our discussions since in most mechanical system identification problems the dimensionality information can be determined apriori. This gives rise to the estimates for the system matrix as,

$$O_k^{(q)T_k \uparrow} = O_k^{\uparrow} T_k = O_{k+1} A_k T_k \tag{2.40}$$

where the identity $O_k^{\uparrow} = O_{k+1} A_k$ (easily verifiable from the definitions in the previous section) was used. However, to produce a consistent estimate, we do not have the true O_{k+1} from the decomposition of the Hankel matrix at the next time step, namely, $H_{k+1}^{(p,q)} = O_{k+1}^{(q)T_{k+1}} R_k^{(p)T_{k+1}}$. But we know from the previous developments that $O_{k+1} = O_{k+1}^{(q)T_{k+1}} T_{k+1}^{-1}$. Substituting this expression in favor of O_k in equation (2.40), we get,

$$O_k^{(q)T_k \uparrow} = O_{k+1}^{(q)T_{k+1}} T_{k+1}^{-1} A_k T_k \tag{2.41}$$

This allows us to set

$$\hat{A}_k = T_{k+1}^{-1} A_k T_k = O_{k+1}^{(q)T_{k+1} \dagger} O_k^{(q)T_k \uparrow} \tag{2.42}$$

as an estimate for the identified time varying discrete system transition matrix. Notice that $\hat{A}_k$ is related to the unknown true system matrix but NOT the true system matrix. A similar estimate can be derived from the controllability grammian expressions. Considering the left shifted Hankel matrix and appropriately resetting its coordinate system, we have,

$$H_{k+1}^{(p,q)\leftarrow} = O_{k+1}^{(q)T_k} R_k^{(p)T_k \leftarrow} \tag{2.43}$$

Using similar manipulations,

$$\begin{aligned} R_k^{(p)T_{k+1}\leftarrow} &= T_{k+1}^{-1} R_k^{\leftarrow} \\ &= T_{k+1}^{-1} A_k R_{k-1} \\ &= T_{k+1}^{-1} A_k T_k R_k^{(p)T_k} \end{aligned} \tag{2.44}$$

we can obtain a similar estimate for the identified system matrix as,

$$\hat{A}_k = R_k^{(p)T_{k+1}\leftarrow} R_{k-1}^{(p)T_k\,\dagger} \tag{2.45}$$

Since the first r columns of $R_k^{(p)T_{k+1}}$ form an estimate for the identified control influence matrix, $\hat{B}_k$, its relation to the unknown true matrix B_k is given by,

$$\hat{B}_k = T_{k+1}^{-1} B_k = R_k^{(p)T_{k+1}}\left(:,1:r\right) \tag{2.46}$$

similarly, the estimate for the identified C_k is obtained by extracting the first m rows of the calculated observability grammian

$$\hat{C}_k = O_k^{(q)T_k}\left(1:m,:\right) = C_k T_k \tag{2.47}$$

where the notation $M\left(1:a,:\right)$ $\left(\text{or } M(:,1:b)\right)$ denotes the first a rows (or b columns) of M matrix.

Having derived the relationships between the identified and the true system model parameters, we now proceed to the impact of the identified plant parameters in the state propagation problem.

State Propagation Using Identified Time Varying Plant Parameters

Let us consider any general time step t_k and the state vector in the coordinate system of the identified plant parameters be given by $\hat{\mathbf{x}}_k$. Assume that the state vector at this time step is known to be given in the identified plant parameter coordinate system (i.e., $\hat{\mathbf{x}}_k = T_k^{-1}\mathbf{x}_k$ is known, while $\mathbf{x}_k, T_k$ is unknown). This assumption will be relaxed shortly. Using the identified plant parameters at the corresponding time step, we have,

$$\begin{aligned} \hat{\mathbf{x}}_{k+1} &= \hat{A}_k \hat{\mathbf{x}}_k + \hat{B}_k \mathbf{u}_k \\ \mathbf{y}_k &= \hat{C}_k \hat{\mathbf{x}}_k + D_k \mathbf{u}_k \end{aligned} \tag{2.48}$$

Clearly, D_k, being invariant with respect to coordinate transformations, while the true propagation equations (had we known $\mathbf{x}_k, A_k, B_k, C_k$) are written as,

$$\begin{aligned} \mathbf{x}_{k+1} &= A_k \mathbf{x}_k + B_k \mathbf{u}_k \\ \mathbf{y}_k &= C_k \mathbf{x}_k + D_k \mathbf{u}_k \end{aligned} \tag{2.49}$$

Using the derived relationships between the true and identified system matrices (equations (2.42),(2.46) and (2.47)), we can write the equation (2.48) as

$$\begin{aligned} \hat{\mathbf{x}}_{k+1} &= T_{k+1}^{-1}\left(A_k T_k \hat{\mathbf{x}}_k + B_k \mathbf{u}_k \right) \\ \mathbf{y}_k &= C_k T_k \hat{\mathbf{x}}_k + D_k \mathbf{u}_k \end{aligned} \tag{2.50}$$

Similarly propagating to one more step, gives us

$$\begin{aligned}\hat{\mathbf{x}}_{k+2} &= T_{k+2}^{-1}\left(A_{k+1}T_{k+1}\hat{\mathbf{x}}_{k+1} + B_{k+1}\mathbf{u}_{k+1}\right)\\ &= T_{k+2}^{-1}\left(A_{k+1}T_{k+1}T_{k+1}^{-1}\left(A_k T_k\,\hat{\mathbf{x}}_k + B_k\mathbf{u}_k\right) + B_{k+1}\mathbf{u}_{k+1}\right)\\ &= T_{k+2}^{-1}\left(A_{k+1}A_k T_k\,\hat{\mathbf{x}}_k + A_{k+1}B_k\mathbf{u}_k + B_{k+1}\mathbf{u}_{k+1}\right)\end{aligned} \tag{2.51}$$

Thus the state equation in general becomes (after p time steps),

$$\hat{\mathbf{x}}_{k+p} = T_{k+p}^{-1}\left(A_{k+p-1}...A_0 T_0\,\hat{\mathbf{x}}_0 + A_{k+p-1}...A_0 B_0\mathbf{u}_0 + ... + A_{k+p-1}B_{k+p-2}\mathbf{u}_{k+p-2} + B_{k+p-1}\mathbf{u}_{k+p-1}\right)$$

(2.52)

Now considering a state propagation error defined as $\mathbf{e}_k := T_k\,\hat{\mathbf{x}}_k - \mathbf{x}_k$, we have the error dynamics after $p+1$ time steps being given by the evolution equation,

$$\mathbf{e}_{p+1} = A_p A_{p-1}...A_0\left(T_0\,\hat{\mathbf{x}}_0 - \mathbf{x}_0\right) = A_p A_{p-1}...A_0\mathbf{e}_0 \tag{2.53}$$

Using Lyapunov's stability theory[24] for discrete time systems we have that the effect of this initial coordinate system misalignment, $\mathbf{e}_0$ will decay asymptotically to zero for time varying systems with a stable origin (cases of asymptotic and exponential stability of the origin). However, in general, one needs to at least determine the initial conditions in the initial system coordinates (namely $\hat{\mathbf{x}}_0$). If one has the initial conditions in the right coordinate system, the identified plant parameters can be used for state propagation in the deterministic case.

Estimation of Initial Conditions from Identified Plant Model Parameters

Let us now look at a method of calculating initial conditions *after* having identified the plant parameter matrix sequences. The question as to whether it is possible to obtain the plant parameters when the output data is inclusive of the initial condition response deserves some explanation at this point. This "chicken and egg problem" can be addressed in several ways. One possible solution is to write the initial condition response together with the forced response and try to solve a matrix equation relating the initial condition and input data to the sensed (noise free) outputs. This leads to a matrix equation the solution of which, using the free response data matrix is not difficult to obtain. We avoid the associated discussion to stay focused on the central developments of the current algorithm. Alternatively, one can use an observer based (ARX model) calculation as detailed in the next chapter developed along the lines of our recent paper[20]. In certain other special case situations where physical nature of the problem is known to the analyst, one can perform repeated experiments by physically setting the initial conditions to zero (position and velocity). In this section we concern ourselves with the problem of determination of initial conditions (in fact the state at a general time step t_k) after the identified plant model sequence is available.

Writing the input output mapping from a general k th time step, for p more time steps, one obtains a set of equations that can be written in a matrix form as,

$$\begin{aligned}
\mathbf{Y} := \begin{bmatrix} \mathbf{y}_k \\ \mathbf{y}_{k+1} \\ \vdots \\ \mathbf{y}_{k+p} \end{bmatrix} &= \begin{bmatrix} \hat{C}_k \\ \hat{C}_{k+1}\hat{A}_k \\ \vdots \\ \hat{C}_{k+p}\hat{A}_{k+p-1}\ldots\hat{A}_k \end{bmatrix} \hat{\mathbf{x}}_k + \begin{bmatrix} D_k & & & \\ \hat{C}_{k+1}\hat{B}_k & D_{k+1} & & \\ \vdots & \vdots & \ddots & \\ \hat{C}_{k+p}\hat{A}_{k+p-1}\ldots\hat{B}_k & \hat{C}_{k+p}\hat{A}_{k+p-1}\ldots\hat{B}_{k+1} & \cdots & D_{k+p} \end{bmatrix} \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k+1} \\ \vdots \\ \mathbf{u}_{k+p} \end{bmatrix} \\
&= O_k^{(q)T_k} \hat{\mathbf{x}}_k + \Delta\mathbf{U} \\
&= \Pi_k \hat{\mathbf{x}}_k + \Delta\mathbf{U}
\end{aligned} \tag{2.54}$$

which can be solved using the least squares solution,

$$\hat{\mathbf{x}}_k^{LS} = \left(\Pi_k^T \Pi_k\right)^{-1} \Pi_k^T \left(\mathbf{Y} - \Delta \mathbf{U}\right) \tag{2.55}$$

provided p is chosen sufficiently large so as to ensure the full rank of the observability grammian (= dimensionality of the state space). We note, for higher dimensioned systems, the least squares inverse of equation (2.55) should be computed instead using either the QR algorithm or the SVD method[21, 22].

Models for the First /Last Few Time Steps

As pointed out in the introductory section of this chapter, in the problems where time varying model identification is of interest, it is often unclear how to isolate the system models for the first few time steps, since the generalized Hankel matrix sequence at these time steps, has a rank of only less than or equal to the true order of the system $\forall k, rank\left(H_k\right) < n$. The first generalized Hankel matrix in question can be written as,

$$H_1^{(0,q)} = \begin{bmatrix} h_{1,0} \\ h_{2,0} \\ \vdots \\ h_{q,0} \end{bmatrix} = \begin{bmatrix} C_1 B_0 \\ C_2 A_1 B_0 \\ \vdots \\ C_q A_{q-1} \ldots B_0 \end{bmatrix} \tag{2.56}$$

Note that it is difficult to compute the generalized Markov parameters such as $C_0 B_{-1}$ (for the populating the grammian $R_1^{(1)T_k}$) since in practical experiments, inputs cannot be applied at negative time index so as to "feel" its response at the current time (i.e., $\mathbf{u}_{-j}, j \in Z^+$ are not available for measurement/computations from the experiment).

Recall that the methodology detailed in previous sections can be employed only once a full rank Hankel matrix can be populated (that is to say in subsequent time steps only).

We now present a method for computing the first few time step models using an additional set of experimental data, the free response experiments. The output data of the free response experiments (also known as the zero input response) are given by,

$$\begin{aligned} \begin{bmatrix} \mathbf{y}_k^{(f,1)} & \mathbf{y}_k^{(f,2)} & \cdots & \mathbf{y}_k^{(f,N_f)} \\ \mathbf{y}_{k+1}^{(f,1)} & \mathbf{y}_{k+1}^{(f,2)} & \cdots & \mathbf{y}_{k+1}^{(f,N_f)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{y}_{k+q-1}^{(f,1)} & \mathbf{y}_{k+q-1}^{(f,2)} & \cdots & \mathbf{y}_{k+q-1}^{(f,N_f)} \end{bmatrix} &= O_k^{(q-1)T_k^f} X_k^{T_k^f} \\ &= \left(U_{k,f} \Sigma_{k,f}^{\frac{1}{2}} \right)\left(\Sigma_{k,f}^{\frac{1}{2}} V_{k,f}^T \right) \end{aligned} \tag{2.57}$$

$\forall k = 0,1,..., p-1$, forming the corresponding observability grammian in the respective coordinate system as the initial conditions (we denote as T_k^f). Deleting the first block of data, we arrive at the block shifted output matrix that can be written as,

$$\begin{aligned} \begin{bmatrix} \mathbf{y}_{k+1}^{(f,1)} & \mathbf{y}_{k+1}^{(f,2)} & \cdots & \mathbf{y}_{k+1}^{(f,N_f)} \\ \mathbf{y}_{k+2}^{(f,1)} & \mathbf{y}_{k+2}^{(f,2)} & \cdots & \mathbf{y}_{k+2}^{(f,N_f)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{y}_{k+q}^{(f,1)} & \mathbf{y}_{k+q}^{(f,2)} & \cdots & \mathbf{y}_{k+q}^{(f,N_f)} \end{bmatrix} &= O_{k+1}^{(q-1)T_{k+1}^f} X_{k+1}^{T_{k+1}^f} \\ &= \left(U_{k+1,f} \Sigma_{k+1,f}^{\frac{1}{2}} \right)\left(\Sigma_{k+1,f}^{\frac{1}{2}} V_{k+1,f}^T \right) \end{aligned} \tag{2.58}$$

We note that the state variable ensemble at the time step t_{k+1} (denoted by $X_{k+1}^{T_{k+1}^f}$, with the corresponding index number $k+1$) is related to the state ensemble at time step t_k (written as $X_k^{T_k^f}$)by,

$$X_{k+1}^{T_{k+1}^f} = \left(T_{k+1}^{f\ -1} A_k T_k^f\right) X_k^{T_k^f} \tag{2.59}$$

Using this relationship, we can derive estimates for the state transition matrix for time steps $k = 0,1,..., p-1$ given by,

$$\hat{A}_k := \left(T_{k+1}^f\right)^{-1} A_k T_k^f = X_{k+1}^{T_{k+1}^f} \left(X_k^{T_k^f}\right)^{-1} \tag{2.60}$$

The calculation of the corresponding $\hat{C}_k$ is accomplished by setting,

$$\hat{C}_k = O_k^{(q-1)T_k^f} \left(1:m,:\right) \tag{2.61}$$

The partial ($rank < n$) Hankel matrices, similar to the one in equation (2.56) are written for the first few time steps ($k = 0,1,..., p-1$) as,

$$H_{k+1}^{(0,q)} = \begin{bmatrix} h_{k+1,k} \\ h_{k+2,k} \\ \vdots \\ h_{k+q,k} \end{bmatrix} = \begin{bmatrix} C_{k+1}B_k \\ C_{k+2}A_{k+1}B_k \\ \vdots \\ C_{k+q}A_{k+q-1}...B_k \end{bmatrix} \tag{2.62}$$

These are used in the determination of the control influence matrix as shown in the following calculation. From the equation (2.62) above,

$$H_{k+1}^{(0,q)} = \begin{bmatrix} C_{k+1}B_k \\ C_{k+2}A_{k+1}B_k \\ \vdots \\ C_{k+q}A_{k+q-1}...B_k \end{bmatrix} = O_{k+1}^{T_{k+1}^f} B_k \tag{2.63}$$

leading to,

$$\hat{B}_k = \left(T_{k+1}^f\right)^{-1} B_k = \left(O_{k+1}^{T_{k+1}^f}\right)^{\dagger} H_{k+1}^{(0,q)} \tag{2.64}$$

However, the plant parameter estimates determined from the equations (2.60),(2.61) and (2.64) are of little use in practice without the coordinate transformation theory developed in a previous section. As pointed out before, the first few models developed in this manner are in totally different coordinate systems, derived from the free response singular value decomposition. Hence, one cannot use the models, thus developed in state propagation since they have a jump discontinuity at the time step $k = p$ in their coordinate systems. Using the developments of the previous section, we correct the models by transforming (projecting) them consistently in to a reference coordinate system. The transformed models are therefore given by,

$$\hat{A}_k^{T_k^{ref}} := P_{k+1}\left(\left(T_{k+1}^f\right)^{-1} A_k T_k^f\right) P_k^{-1} = P_{k+1}\left(X_{k+1}^{T_{k+1}^f}\left(X_k^{T_k^f}\right)^{-1}\right) P_k^{-1} \tag{2.65}$$

$$\hat{B}_k^{T_k^{ref}} := P_{k+1}\left(\left(T_{k+1}^f\right)^{-1} B_k\right) = P_{k+1}\left(\left(O_{k+1}^{T_{k+1}^f}\right)^{\dagger} H_{k+1}^{(0,q)}\right) \tag{2.66}$$

and

$$\hat{C}_k^{T_k^{ref}} := \left(C_k T_k^f\right) P_k^{-1} == \left(O_k^{T_k^f}\left(1:m,:\right)\right) P_k^{-1} \tag{2.67}$$

where the transformation (projection onto the reference subspace $T_{k_r}^{ref}$, at a reference time step k_r) is defined as,

$$P_k = \left(O_{k_r}^{T_{k_r}^{ref}}\right)^{\dagger} O_k^{T_k^f} \tag{2.68}$$

Using the transformed system models in to a reference coordinate system, and considering the subsequent models in compatible coordinate systems, one therefore obtains a complete sequence of discrete time varying models from time step $k = 0,...,k_f$ as long as desired by the analyst, depending on the availability of multiple experimental data. The first few models for the case of the numerical examples, discussed in a subsequent section were obtained in this manner and the state propagation results were computed employing the transformations developed here-in. The last few time step models have a dual nature in that the system observability grammian cannot be formed fully owing to the rank deficiency of the Hankel matrix. This defect can analogously be corrected using the developments of this section.

Thus, using a framework similar to Liu [8], we arrive at a different set of more general results for the first few time step models. We note in passing that there exist some structural relationships (among the generalized Markov parameters and the Hankel matrices) that lead to suggest that one can avoid repeated free response experiments. However, we could not find any useful manipulations to report at this stage and are forced to use these extra conditions to recover the first few time step models.

Estimation of Markov Parameters from Input Output Data Using Least Squares

As we have seen so far, the generalized Markov Parameters play an important role in the determination of the time varying plant parameters in the Time Varying Eigensystem Realization algorithm. We now address the question: How these Markov Parameters are computed from input output data? For simplicity, we consider only the ideal case where the output data from multiple experiments is devoid of initial condition response. In this case, we assume that all the experiments are performed from zero initial conditions (ideal situation). In the presence of unknown initial conditions in the output data, the determination of Markov parameters is more complicated since one requires more information to separate out the components of the output data caused due to the unknown initial conditions.

The output of the system at the time step t_k (sufficiently later than the initial time t_0) is related to the control inputs up to that time instant as (using equation (2.6) with $\mathbf{x}_0 = \mathbf{0}$),

$$\begin{aligned} \mathbf{y}_k &= \sum_{j=k_0}^{k-1} h_{k,j}\mathbf{u}_j + D_k\mathbf{u}_k \\ &= h_{k,k_0}\mathbf{u}_{k_0} + \ldots + h_{k,k-1}\mathbf{u}_{k-1} + D_k\mathbf{u}_k \\ &= D_k\mathbf{u}_k + C_kB_{k-1}\mathbf{u}_{k-1} + C_kA_{k-1}B_{k-2}\mathbf{u}_{k-2} + \ldots + C_kA_{k-1}\ldots A_{k_0+1}B_{k_0}\mathbf{u}_{k_0} \end{aligned} \tag{2.69}$$

Stacking the generalized Markov parameters in the block matrix notation, we have that,

$$\mathbf{y}_k = \begin{bmatrix} D_k & C_kB_{k-1} & \cdots & C_kA_{k-1}\ldots A_{k_0+1}B_{k_0} \end{bmatrix} \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k-1} \\ \vdots \\ \mathbf{u}_{k_0} \end{bmatrix} \tag{2.70}$$

For input output data from multiple experiments (experiment number denoted by the superscript $(.)^{(j)}$) we consequently have the matrix equation,

$$\begin{bmatrix} \mathbf{y}_k^{(1)} & \mathbf{y}_k^{(2)} & \cdots & \mathbf{y}_k^{(N)} \end{bmatrix}$$
$$= \begin{bmatrix} D_k & C_k B_{k-1} & \cdots & C_k A_{k-1} \ldots A_{k_0+1} B_{k_0} \end{bmatrix} \begin{bmatrix} \mathbf{u}_k^{(1)} & \mathbf{u}_k^{(2)} & \ldots & \mathbf{u}_k^{(N)} \\ \mathbf{u}_{k-1}^{(1)} & \mathbf{u}_{k-1}^{(2)} & \ldots & \mathbf{u}_{k-1}^{(N)} \\ \vdots & \vdots & & \vdots \\ \mathbf{u}_{k_0}^{(1)} & \mathbf{u}_{k_0}^{(2)} & \ldots & \mathbf{u}_{k_0}^{(N)} \end{bmatrix} \quad (2.71)$$

where the number of experiments, N is chosen such that for each output time step of interest, a least square solution for all the Markov parameters (until the initial time step k_0) is possible.

The design of such increasing number of experiments is necessary to obtain a unique solution for the generalized Markov parameters from the input output map. This increase in computations is one of the few reasons behind the lack of popularity among time varying identification methods. In the next chapter, based on the theoretical developments of the recent papers in preparation,[20] we present techniques to remedy this increase and demonstrate the fact that the introduction of an observer in to the identification process enables a dramatic reduction of the number of required experiments, while retaining the level of accuracy in the calculated generalized Markov parameters due to the existence of certain recursive relationships existing in the time varying observer realized. These results generate sufficient optimism for the practical analyst to consider the time varying identification methods as an alternative in analysis and design of models for control and estimation (and/or guidance and navigation).

Numerical Examples

We demonstrate the results of this chapter on two representative examples. The first example has a stable origin but with true system matrix having time varying elements that have an oscillatory nature. The second example is also an oscillator example with stiffness matrix varying with time and no damping. This represents the class of problems which do not have a stable origin. We do not present examples where the solution diverges exponentially to infinity, because the generalized Markov parameters for such problems also go to infinity and hence the input output description may become too highly ill-conditioned to allow stable computations and comparisons.

Example 1: System with a Stable Origin

Consider the time varying system with true matrices, being given by,

$$\begin{aligned} A_k &= \begin{bmatrix} 0.3-0.9\tau_k & 0.1 & 0.7\tau_k' \\ 0.6\tau_k & 0.3-0.8\tau_k' & 0.01 \\ 0.5 & 0.15 & 0.6-0.9\tau_k \end{bmatrix} \\ B_k &= \begin{bmatrix} 1 & 0 \\ 1 & -1 \\ 0 & 1 \end{bmatrix},\ C_k = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix},\ D_k = 0.1\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned} \tag{2.72}$$

where the time varying elements are defined as $\tau_k = \sin\left(10t_k\right),\quad \tau_k' := \cos\left(10t_k\right)$. The first validation is performed by inspection of the rank of the Hankel matrix sequences. As the Figure 4 clearly shows, the rank of the system remains 3 for all time, indicating the order of the system, as discussed previously in this chapter. Using the least squares solution as shown in the previous section, the system Markov parameters are determined from repeated experiments. The identification procedure is carried out and two test control inputs are applied to the true and identified system with zero initial conditions given by $u_1(t_k) = 0.5\sin(12t_k)$ and $u_2(t_k) = \cos(7t_k)$.

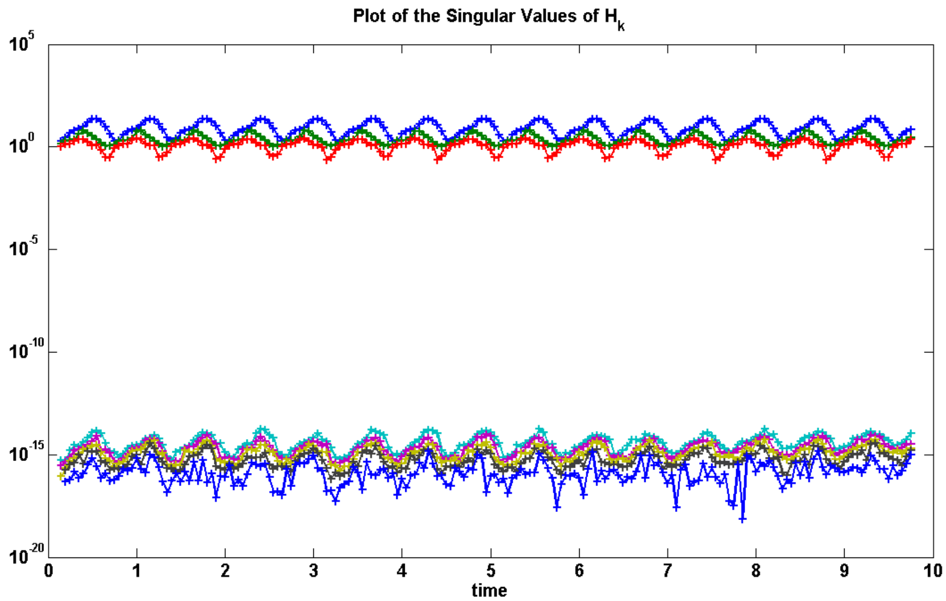

Figure 4. (Ex. 1) Hankel Matrix Sequence Singular Values

The response for these test control inputs obtained from the identified plant model sequence and the true model sequence is compared in Figure 5. Figure 6 plots the output error incurred between the outputs of the true and identified systems to the test control input sequences.

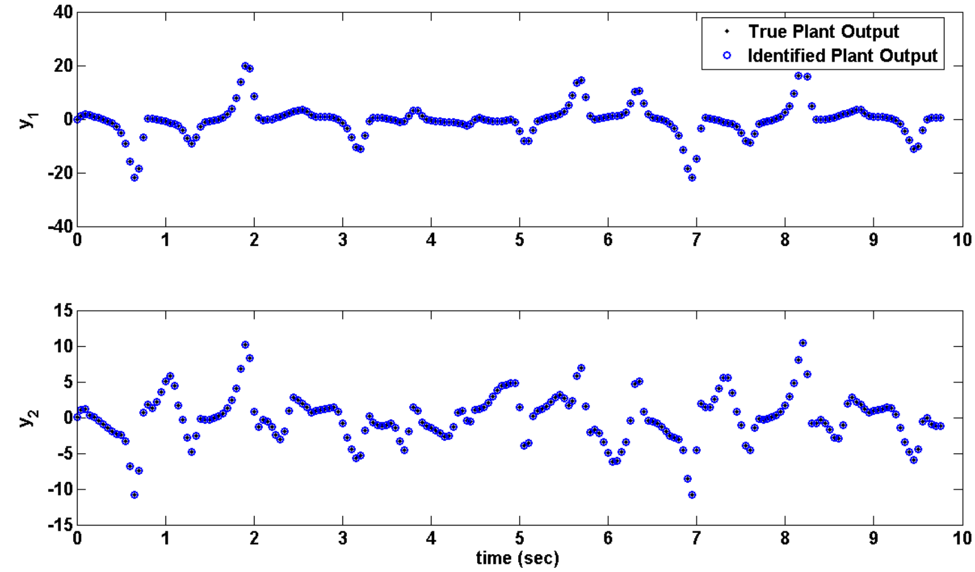

Figure 5. (Ex. 1) Output Comparison: Response to Test Functions

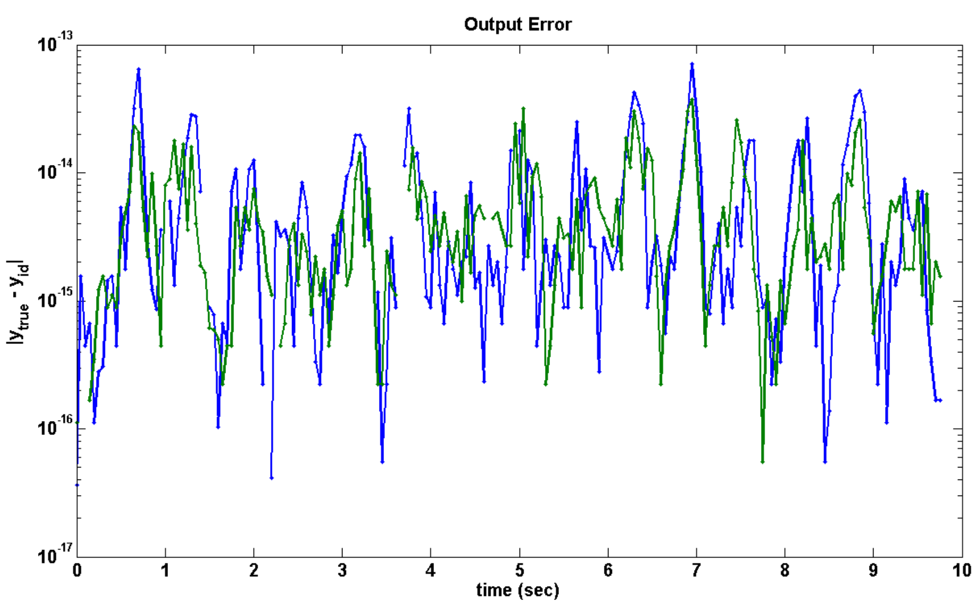

Figure 6. (Ex. 1) Output Error Comparison: Response to Test Functions

Example 2: Oscillatory System (Zero Damping)

Alternatively, we consider the system with an oscillatory nature. In this case the plant system matrix was calculated as

$$
\begin{aligned}
A_k &= \exp\left[A_c * \Delta t\right] \\
B_k &= \begin{bmatrix} 1 & 0 \\ 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad C_k = \begin{bmatrix} 1 & 0 & 1 & 0.2 \\ 1 & -1 & 0 & -0.5 \end{bmatrix}, \\
D_k &= 0.1\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\end{aligned}
\tag{2.73}
$$

where the matrix is given by

$$
A_c = \begin{bmatrix} 0_{2\times2} & I_{2\times2} \\ -K_t & 0_{2\times2} \end{bmatrix} \tag{2.74}
$$

with $K_t = \begin{bmatrix} 4+3\tau_k & 1 \\ 1 & 7+3\tau'_k \end{bmatrix}$ and τ_k, τ'_k are as defined in the Example 1.

The free response of this system from true initial conditions $\mathbf{x}_0^T = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$ is plotted in Figure 7. This clearly shows qualitatively that there is no damping inherent in the system and that time varying stiffness term is present. The singular values of the Hankel matrix sequence, plotted in the Figure 8 reveal that the true order of the system is 4.

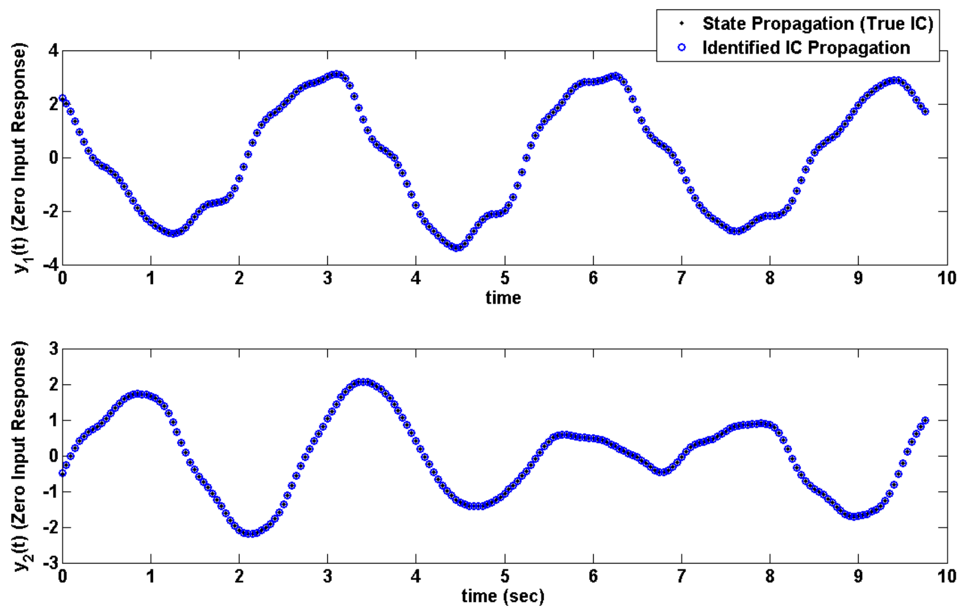

Figure 7. (Ex.2) Oscillatory Output for Non-zero Initial Conditions (Unstable Origin)

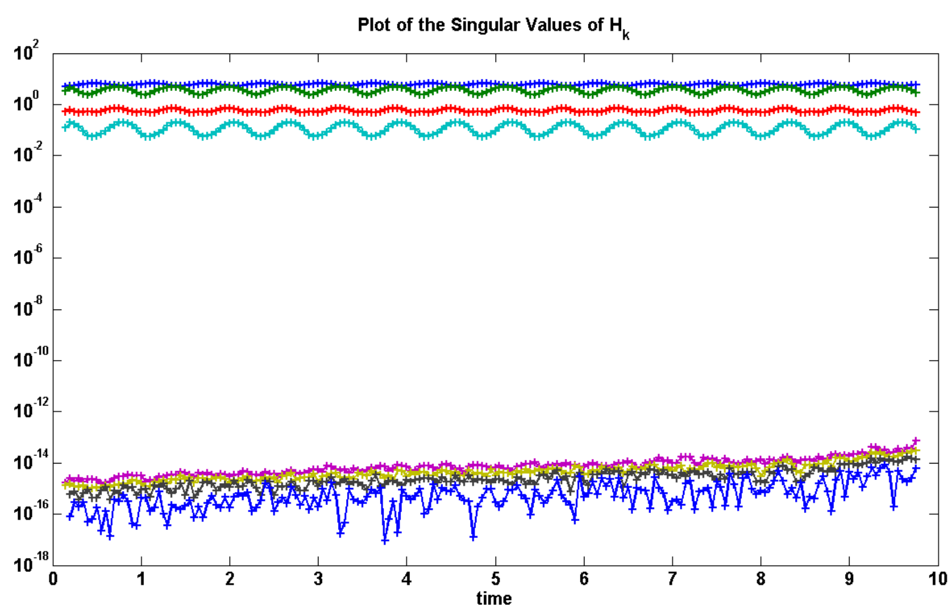

Figure 8. Hankel Matrix Sequence of Singular Values (Ex.2)

The generalized Markov parameters are determined by solving the least squares problem obtained by considering the input output relationship as described in the previous section. The norm of the error incurred in these calculations is plotted in Figure 9. The deterioration of the accuracy towards the end of the simulation is due to the increase in the size of the least squares problem towards the end of the simulation and the deterioration of the absolute error tolerance of the solution of the linear system (for the same level of relative error maintained in the numerical solution). After identification using the Markov parameters, the same test functions as the previous example were employed to the true and identified system model sequence and the responses obtained are compared in Figure 10. The error between the true and the identified response to test functions is plotted in Figure 11.

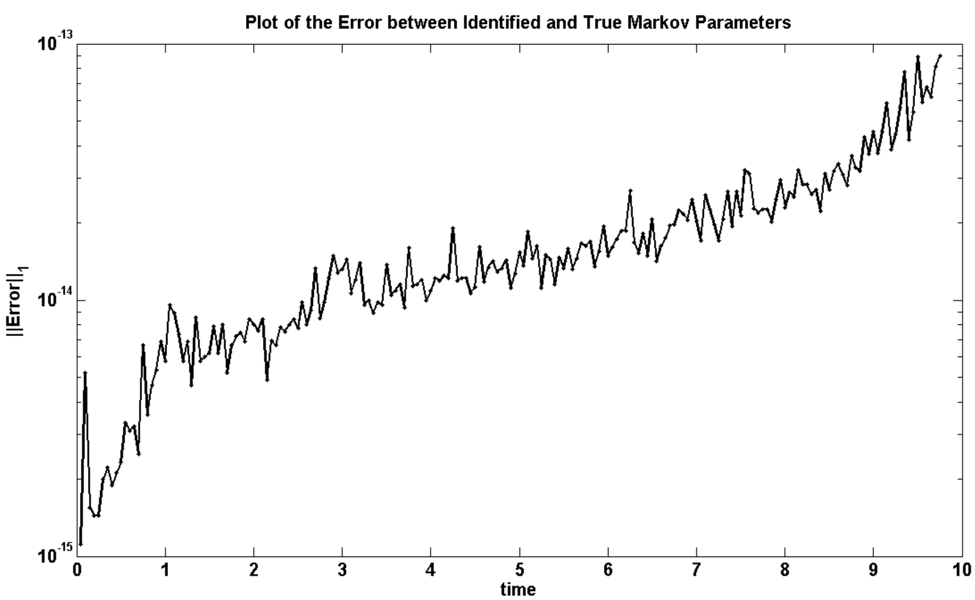

Figure 9. Error in Calculation of Markov Parameters From Zero State Response

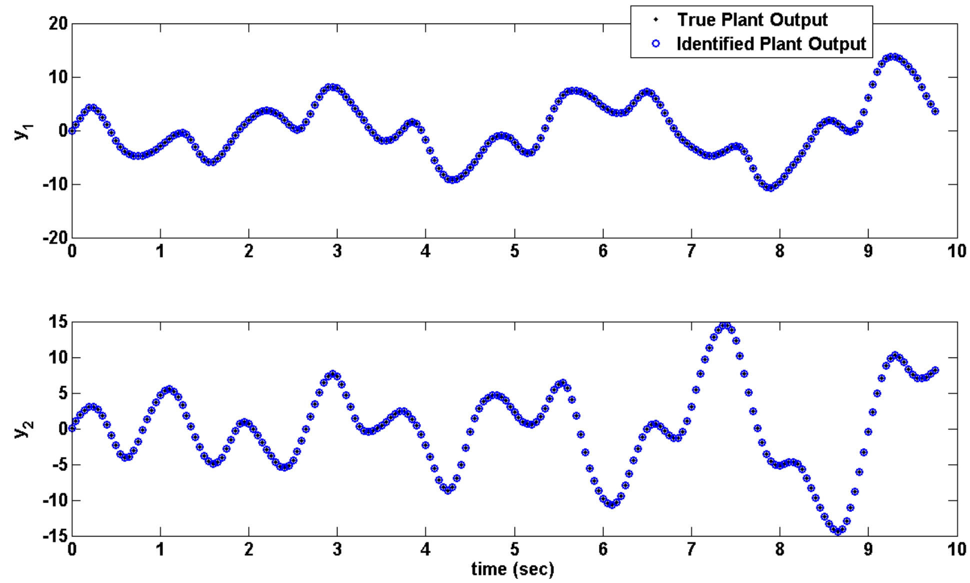

Figure 10. (Ex. 2) Output Comparison (True vs. Identified - Forced Response)

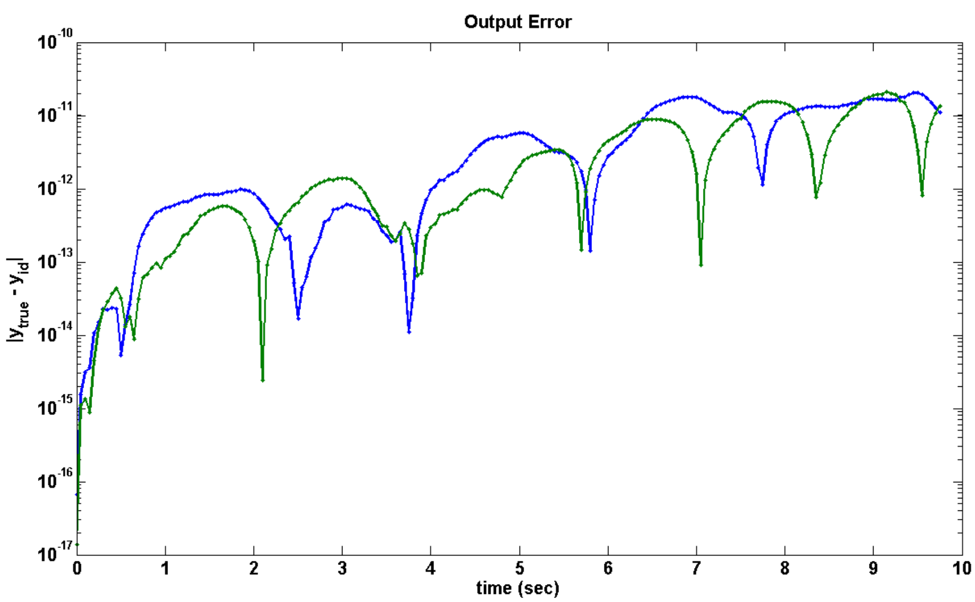

Figure 11. (Ex. 2) Output Comparison (True vs. Identified - Forced Response)

The initial condition determination strategy presented previously in this chapter, was employed on the output data to calculate the initial conditions of the state in the identified initial coordinate system. Choosing $p = 9$ for best accuracy in the normal equations, we obtained an estimate of the initial conditions (in the unknown T_0 coordinate system to be)

$$\hat{\mathbf{x}}_0 = \left[\text{-}0.8833 \quad -1.5612 \quad -2.0021 \quad -0.2452\right]^T \tag{2.75}$$

Using the estimated initial conditions (in their coordinate systems), the state was propagated and the free response was compared as shown in the Figure 7. The output error between the true initial condition response and the determined initial condition response is plotted in Figure 12.

Eigenvalues of the System Matrix

It was found, supporting the discussions of the theoretical developments of this chapter, that the true and identified system matrices are not similar. This is demonstrated by plotting the coefficients of the characteristic polynomial for true system and identified system in the time varying coordinate systems in Figure 13. Owing to the possible existence of complex eigenvalues for a real matrix, we should plot the real and complex components (or the magnitude and phase) of the eigenvalues computed at each time step.

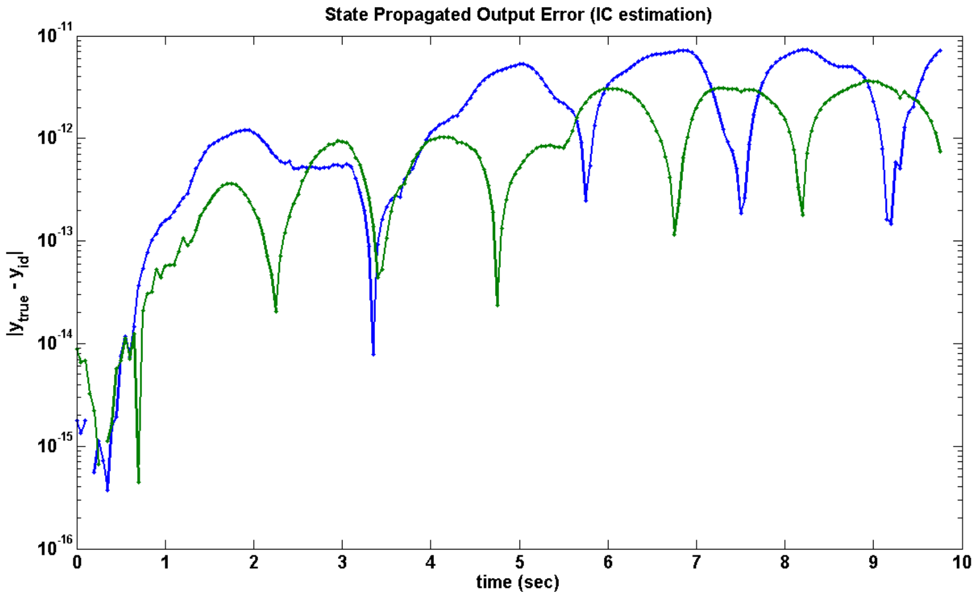

Figure 12. (Ex. 2) Output Error Comparison (True vs. Identified - Initial Condition Determination in Identified Coordinate System)

Using the fundamental theorem of algebra (which guarantees the existence of n solutions to polynomial equations of a fixed order n), this is completely equivalent to plotting the coefficients of the corresponding time varying characteristic polynomials. Figure 13 clearly demonstrates that the eigenvalues of the identified and the true system matrices are indeed different and therefore it appears like no conclusions about the true physics of the problem can be made by such a comparison. This, at first blush agrees with conventional wisdom that no physical significance can be attributed to the (instantaneous) eigenvalues of a time varying system, especially from an input output stand point. Linear systems texts have given examples that show that one can have a system matrix of zero in continuous time domain (identity matrix in discrete time domain).

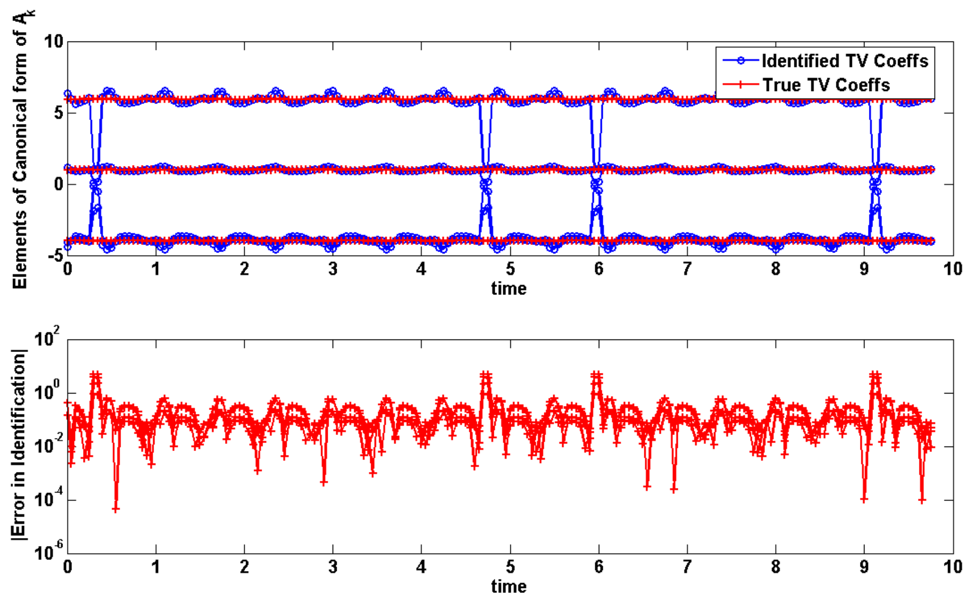

Figure 13. Coefficients of the Characteristic Equation : True and Identified (in Time Varying Coordinate Systems)

Applying the transformations, defined in this chapter, the true and identified eigenvalues (magnitude) as seen in the coordinate system $T_0 = T_0^f$ (in the observable sub space at time t_0, $O_0^{T_0^f}$) are plotted as Figure 14. The corresponding time varying coefficients of the characteristic polynomial are shown to agree in Figure 15.

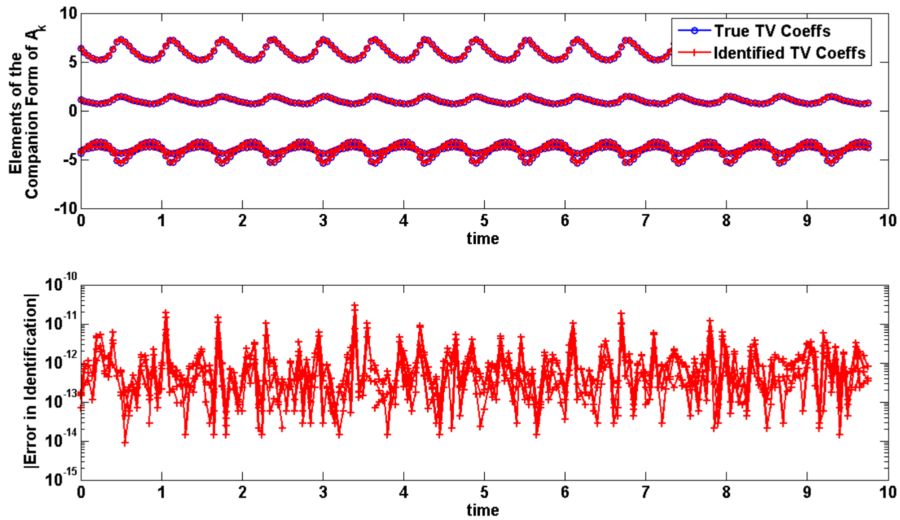

Figure 14. Coefficients of the Characteristic Equation : True and Identified (in Reference Coodinate System)

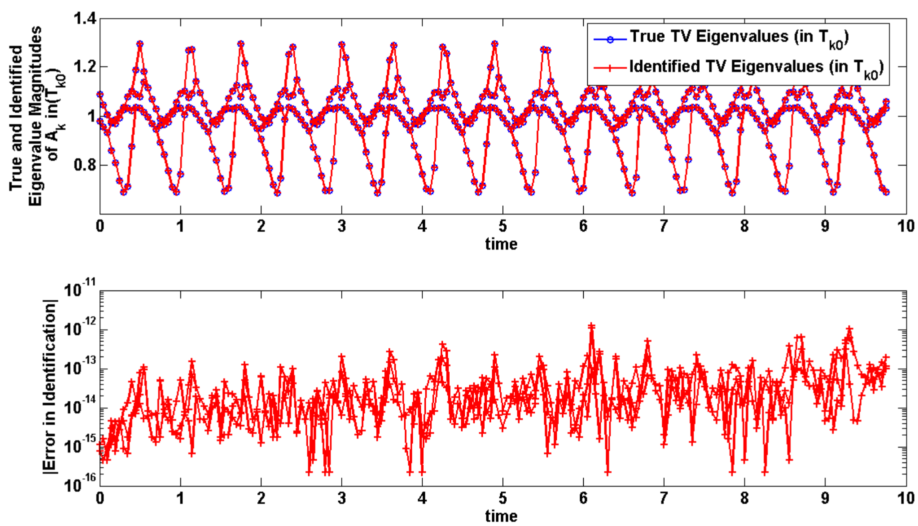

Figure 15. Eigenvalue Magnitudes of the Time Varying System Matrix: True and Identified (in Reference Coodinate System)

Conclusion

This chapter presents extensions of the celebrated Eigensystem Realization Algorithm for the identification of linear time invariant systems to realize linear models that are time varying in the discrete time domain. The time varying extensions are derived using established notions of generalized Markov parameters and the generalized Hankel matrix sequences, thereby extending the classical Ho-Kalman algorithm to include the realization of the time varying discrete time model sequences from input output data. It is shown that the models thus realized are in general obtained in different (arbitrary) coordinate systems, inherent to the general theory of time varying linear systems of differential (and difference) equations. It is found subsequently, that the kinematically similar (topologically equivalent) realizations are indeed similar when observed from a single reference (albeit unknown) coordinate system. This novel result is introduced (and used) as a tool to compare different realizations obtained by several algorithms. A method to transform (more precisely project) the system models thus realized into a (generally unknown) common reference coordinate system is presented by construction of time varying projection operators. It is shown that the transformation matrices constructed project the realized system models into a space spanning the corresponding controllable or observable subspace at the reference time step. A method to isolate the time varying models for the first few (and last few) time steps using free response data from the unknown initial conditions is presented, thereby completing the sequence of models realized by the algorithm to every time step the experimental data is available. A least squares solution is presented to for the determination of generalized

Markov parameters using experimental data from repeated experiments. Numerical examples are presented to demonstrate the algorithmic methodologies and support the theoretical results of the chapter.

CHAPTER III

OBSERVER/KALMAN FILTER TIME VARYING SYSTEM IDENTIFICATION

Introduction

From the developments of the previous chapter, it is clear that the generalized Markov parameters play a central role in the identification of the linear discrete time varying plant model sequences using the time varying eigensystem realization algorithm (TVERA). Also, towards the end of the chapter, it is pointed out that a computationally viable alternative strategy needs to be incorporated to overcome the practical limitation of the increasing number of repeated experiments required for the determination of the large number of generalized Markov parameters used by the TVERA computations. This is because, in practice, one cannot obtain a solution to the equations relating the very high dimensioned input output map to obtain the generalized Markov parameters.

The presence of nonzero initial conditions prevents us from solving the generalized Markov parameters from the linear system of equations in a piece meal fashion. That is to say that solving several sets of equations similar in form to equation (2.54) at fixed time instances, say, t_{k_c} is not practical or even possible since $\hat{\mathbf{x}}_{k_c}$ is nonzero, in general. Furthermore, in systems where stability of the origin cannot be ascertained, the number of potentially significant generalized Markov parameters grows rapidly. This is because in case of the problems with an unstable origin, the output at every time step in the time varying case depends on the linear combinations of the (normalized) unit response functions of all the inputs applied until that instant (causal inputs). Therefore the number of unknowns increase by $m*r$ for each time step in the model sequence and consequently, the analyst is required to perform more experiments if a refined discrete time model is sought. This computational challenge has been among the main reasons for the lack of ready-adoption of the time varying subspace based identification methods.

In this chapter, we introduce an asymptotically stable, time varying observer to remedy this problem of unbounded growth in the number of experiments. The algorithm developed as a consequence is called the time varying observer/Kalman filter system identification (TVOKID). In addition, the tools systematically developed in this chapter give an estimate on the least number of experiments one needs to perform for identification and/or recovery of all the Markov parameters of interest until that time instant. Furthermore, since the frequency response functions for time varying systems are not well known, the method outlined seems to be the one of the first practical ways to obtain the generalized Markov parameters bringing most of the generalized Markov parameter based discrete time varying identification methods to the table of the practicing engineer. Theoretical accomplishments of the chapter are equally important. Novel models relating input output data are developed and are found to be elegant extensions of the ARX models well known in the analysis of time invariant models. This generalization of the classical ARX model to the time varying case admits analogous recursive relations with the system Markov parameters as was developed in the time invariant case. The analogy goes even further and enables us to define a dead beat condition for time varying systems. The generalization of this deadbeat definition is rather unique and general for the time varying systems as it was shown that not all the time varying eigenvalues need to be zero for the closed loop to be called dead beat. Further, it is demonstrated that the time varying observer sequence (dead beat or otherwise) realized from the generalized ARX (GTV-ARX) model is realized in a compatible coordinate system with the identified plant model sequence. Relations with the time varying Kalman observer are made comparing features of the parameters of the Kalman observer with the time varying observer realized from the generalized OKID procedure presented in the chapter.

Basic Formulation

We start by revisiting the relations between the input output sets of vectors via the system Markov parameters as developed in the theory concerning the time varying eigensystem realization algorithm (TVERA) developed in the previous chapter. The fundamental difference equations governing the evolution of a linear system in discrete time are given by (repeated here for convenience of presentation),

$$\mathbf{x}_{k+1} = A_k \mathbf{x}_k + B_k \mathbf{u}_k \tag{3.1}$$

Together with the measurement equations,

$$\mathbf{y}_k = C_k \mathbf{x}_k + D_k \mathbf{u}_k \tag{3.2}$$

with the state, output and input dimensions $\mathbf{x}_k \in \mathbb{R}^n, \mathbf{y}_k \in \mathbb{R}^m, \mathbf{u}_k \in \mathbb{R}^r$ and the system matrices to be of compatible dimensions $\forall k \in \mathbb{Z}$, an index set. The solution for the state evolution (the linear time varying discrete time difference equation solution) is given by,

$$\mathbf{x}_k = \Phi\left(k,k_0\right)\mathbf{x}_0 + \sum_{j=k_0}^{k-1}\Phi\left(k,j+1\right)B_j\mathbf{u}_j \tag{3.3}$$

$\forall k \geq k_0 + 1$, where the state transition matrix, $\Phi(.,.)$ is defined as,

$$\Phi\left(k,k_0\right) = \begin{cases} A_{k-1}A_{k-2}\ldots A_{k_0}, & \forall k > k_0 \\ I, & k = k_0 \\ \text{undefined}, & \forall k < k_0 \end{cases} \tag{3.4}$$

Using the definition of the compound state transition matrix, the input output relationship is given by,

$$\mathbf{y}_k = C_k\Phi\left(k,0\right)\mathbf{x}_0 + \sum_{j=0}^{k-1}C_k\Phi\left(k,j+1\right)B_j\mathbf{u}_j + D_k\mathbf{u}_k \tag{3.5}$$

This enables us to define the input output relationship in terms of the two index coefficients as,

$$\mathbf{y}_k = C_k\Phi\left(k,0\right)\mathbf{x}_0 + \sum_{j=0}^{k-1}h_{k,j}\mathbf{u}_j + D_k\mathbf{u}_k \tag{3.6}$$

where the generalized Markov parameters are defined to be given by,

$$h_{k,i} = \begin{cases} C_k\Phi\left(k,i+1\right)B_i, & \forall i < k-1 \\ C_kB_{k-1}, & i = k-1 \\ 0, & \forall i > k-1 \end{cases} \tag{3.7}$$

From now on, we use the expanded form of the state transition matrix $\Phi(.,.)$ to improve the clarity of the presentation. Thus the output at any general time step t_k is related to the initial conditions and all the inputs as,

$$\mathbf{y}_k = C_k A_{k-1} \ldots A_{k_0+1} A_{k_0} \mathbf{x}_{k_0} + \begin{bmatrix} D_k & C_k B_{k-1} & \cdots & C_k A_{k-1} \ldots A_{k_0+1} B_{k_0} \end{bmatrix} \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k-1} \\ \vdots \\ \mathbf{u}_{k_0} \end{bmatrix} \tag{3.8}$$

where k_0 can denote any general time step prior to k (in particular let us assume it to denote the initial time such that $k_0 = 0$). As was pointed out in the introduction and the previous chapter, such a relationship between the input and output leads to a problem that increases by $m * r$ parameters for every time step considered and the number of unique experiments required in order to obtain a unique solution for the generalized Markov parameters also increases correspondingly. Thus it becomes difficult to compute the increasing number of unknown parameters for reasons of numerical stability (larger system of unknowns and equations) and practicality of carrying out ever larger number of experiments. In the special case of systems whose open loop is asymptotically stable, this is not a problem. However, frequently, one seeks to use identification in problems which do not have a stable origin for control and estimation purposes. Hence in such problems we need to explore alternative methods in which plant parameter models can be realized from input output data. A viable alternative is developed in the following section.

The central assumption involved in the developments of this chapter is that(in order to obtain the system and Observer Markov parameters for all time steps involved), one should start the experiments from zero initial conditions or from the same initial conditions each time the experiment is performed. Although the more general case of the presence of initial condition response included in the output data is soluble, the calculations involved in determining the first few Markov parameters become involved and formulations are considerably more tedious, warranting a separate discussion. Most importantly since the connections between time varying ARX model and the state space model parameters and a discussion on the associated observer are complex enough, we proceed with the presentation of the algorithm under the assumption that each experiment is performed with zero initial conditions.

Input Output Representations: Observer Markov Parameters

The situation we face for the time varying systems is quite analogous to the problem of estimation of the modes for a lightly damped flexible spacecraft structure in the time invariant case (ref. Chapter VI in the book by Juang[1]). In the identification problem involving a lightly damped structure, one has to track a large number of Markov parameters to obtain reasonable accuracy for the modal parameters in question. An effective method for "compressing" experimental input output data, called Observer/Kalman Filter Markov Parameter Identification (OKID) Theory was developed by Juang et. al.,[1, 10, 25] for such problems. In this section, we generalize these classical observer based schemes for determination of generalized Markov parameters. The concept of frequency response functions that enables the determination of system Markov parameters for time invariant system identification does not have a clear analogous theory in case of the time varying models. Therefore, the method described here-in constitutes one of the first efforts to efficiently compute the generalized Markov parameters from experimental data. Importantly, for the first time, we are also able to isolate a minimum number of repeated experiments to help the practicing engineer to plan the experiments required for identification, apriori.

Following the observations of the previous researchers, consider the use of an "output – feedback" style gain (time varying) sequence in the difference equation model Eq. (3.1) governing the linear plant, given by,

$$\begin{aligned}
\mathbf{x}_{k+1} &= A_k \mathbf{x}_k + B_k \mathbf{u}_k + G_k \mathbf{y}_k - G_k \mathbf{y}_k \\
&= \left(A_k + G_k C_k \right) \mathbf{x}_k + \left(B_k + G_k D_k \right) \mathbf{u}_k - G_k \mathbf{y}_k \\
&= \bar{A}_k \mathbf{x}_k + \left[\left(B_k + G_k D_k \right) \ \vdots \ -G_k \right] \begin{bmatrix} \mathbf{u}_k \\ \mathbf{y}_k \end{bmatrix} \\
&= \bar{A}_k \mathbf{x}_k + \bar{B}_k \boldsymbol{\upsilon}_k
\end{aligned} \tag{3.9}$$

with no change in the measurement equations at the time step t_k,

$$\mathbf{y}_k = C_k \mathbf{x}_k + D_k \mathbf{u}_k \tag{3.10}$$

The outputs at the consecutive time steps, starting from the initial time step t_0 $\left(\text{denoted by } k_0 = 0\right)$ are therefore written as,

$$\begin{aligned}
\mathbf{y}_{k_0} &= C_{k_0} \mathbf{x}_{k_0} + D_{k_0} \mathbf{u}_{k_0} \\
\mathbf{y}_{k_0+1} &= C_{k_0+1} \bar{A}_{k_0} \mathbf{x}_{k_0} + D_{k_0+1} \mathbf{u}_{k_0+1} + C_{k_0+1} \bar{B}_{k_0} \boldsymbol{\upsilon}_{k_0} \\
&= C_{k_0+1} \bar{A}_{k_0} \mathbf{x}_{k_0} + D_{k_0} \mathbf{u}_{k_0} + \bar{h}_{k_0+1,k_0} \boldsymbol{\upsilon}_{k_0} \\
\mathbf{y}_{k_0+2} &= C_{k_0+2} \bar{A}_{k_0+1} \bar{A}_{k_0} \mathbf{x}_{k_0} + D_{k_0+2} \mathbf{u}_{k_0+2} + C_{k_0+2} \bar{B}_{k_0+1} \boldsymbol{\upsilon}_{k_0+1} + C_{k_0+2} \bar{A}_{k_0+1} \bar{B}_{k_0} \boldsymbol{\upsilon}_{k_0} \\
&= C_{k_0+2} \bar{A}_{k_0+1} \bar{A}_{k_0} \mathbf{x}_{k_0} + D_{k_0+2} \mathbf{u}_{k_0+2} + \bar{h}_{k_0+2,k_0+1} \boldsymbol{\upsilon}_{k_0+1} + \bar{h}_{k_0+2,k_0} \boldsymbol{\upsilon}_{k_0} \\
&\cdots
\end{aligned} \tag{3.11}$$

with the definition of generalized observer Markov parameters,

$$\bar{h}_{k,i} = \begin{cases} C_k \bar{A}_{k-1} \bar{A}_{k-2} \ldots \bar{A}_{i+1} \bar{B}_i \ , & \forall k > i+1 \\ C_k \bar{B}_{k-1}, & k = i+1 \\ 0, & \forall k < i+1 \end{cases} \tag{3.12}$$

arriving at the general relationship,

$$\mathbf{y}_k = C_k \bar{A}_{k-1} \ldots \bar{A}_{k_0} \mathbf{x}_{k_0} + D_k \mathbf{u}_k + \sum_{j=1}^{k-k_0-1} \bar{h}_{k,k-j} \boldsymbol{\upsilon}_{k-j} \tag{3.13}$$

We point out that the generalized observer Markov parameters have two block components similar to the linear time invariant case shown in the partitions to be,

$$\begin{aligned} \bar{h}_{k,k-j} &= C_k \bar{A}_{k-1} \ldots \bar{A}_{k-j+1} \bar{B}_{k-j} \\ &= \left[C_k \bar{A}_{k-1} \ldots \bar{A}_{k-j+1} \left(B_{k-j} + G_{k-j} C_{k-j} \right) \quad -C_k \bar{A}_{k-1} \ldots \bar{A}_{k-j+1} G_{k-j} \right] \\ &= \left[\bar{h}^{(1)}_{k,k-j} \quad -\bar{h}^{(2)}_{k,k-j} \right] \end{aligned} \tag{3.14}$$

where, as will be derived in the subsequent developments of this chapter, the partitions $\bar{h}^{(1)}_{k,k-j}, \bar{h}^{(2)}_{k,k-j}$ are used in the calculations of system Markov parameters and the time varying observer gain sequence. The closed loop thus constructed, is now forced to have an asymptotically stable origin. The goal of an observer constructed in such a fashion is to enforce certain desirable (stabilizing) characteristics in to the closed loop (e.g., dead beat – like stabilization, etc.).

The first step involved in achieving this goal of closed loop asymptotic stability is to choose the number of time steps p_k (variable each time in general) sufficiently large so that the output of the plant (at t_{k+p_k}) strictly depends on only the $p_k + 1$ previous augmented control inputs $\left\{\boldsymbol{\upsilon}_{k+j-1}\right\}_{j=1}^{p_k}, \mathbf{u}_{k+p_k}$ and independent of the state at every time step t_k. Therefore by writing,

$$\begin{aligned}\mathbf{y}_{k+p_k} &= C_{k+p_k}\overline{A}_{k+p_k-1}\ldots\overline{A}_k\mathbf{x}_k + D_{k+p_k}\mathbf{u}_{k+p_k} + \sum_{j=1}^{p_k}\overline{h}_{k+p_k,k+j-1}\boldsymbol{\upsilon}_{k+j-1} \\ &\approx D_{k+p_k}\mathbf{u}_{k+p_k} + \sum_{j=1}^{p_k}\overline{h}_{k+p_k,k+j-1}\boldsymbol{\upsilon}_{k+j-1}\end{aligned} \tag{3.15}$$

we have set $C_{k+p_k}\overline{A}_{k+p_k-1}\ldots\overline{A}_k\mathbf{x}_k = \mathbf{0}$ (with exact equality assignable i.e., $C_{k+p}\overline{A}_{k+p-1}\ldots\overline{A}_k\mathbf{x}_k = 0$, in the absence of measurement noise $\forall k = 0,1,\ldots,k_f$). This leads to the construction of a generalized time varying autoregressive model with exogenous input (a familiar acronym, GTV-ARX is coined to represent this model) at every time step. Note that the order $\left(p_k\right)$ of the GTV-ARX model can also change with time (the term "generalized" is used to describe this variability in the order of the realized model). This variation and complexity provides a larger number of observer gains at the disposal of the analyst under the time varying OKID framework. In using this input output relationship (Eq.(3.15)) instead of the exact relationship given in Eq.(3.8), we introduce damping into the closed loop. For simplicity and ease in implementation (and understanding), we consider the generally variable order to remain fixed and minimum (dead beat) at each time step. That is to say, $p_k = p = p_{\min}$ where $p_{\min}$ is the smallest positive integer such that $p_{\min} \geq mn$. This restriction, (albeit unnecessary) forces a time varying dead beat observer and includes elements of linear time invariance (shift invariance) in the (closed loop) behavior of observer Markov parameters, providing ease in calculations by requiring only the minimum number of repeated experiments to be performed. It turns out that the dead beat conditions are different in the case of time varying systems, due to the transition matrix product conditions (Eq. (3.15) and Eq. (3.16)) that are set to zero. This situation is in contrast with (and is a modest generalization of the situation in) the time invariant systems where higher powers of the system matrix give sufficient conditions to place all the closed loop system poles at the origin (dead-beat). The nature and properties of the time varying dead beat condition are briefly summarized in the Appendix C, along with an example problem, owing to the fact that considerations of the time varying dead beat condition appear sparse, if not completely heretofore unknown in modern literature.

If the repeated experiments (as derived and presented in[26, 27]) are performed so as to compute a least squares solution to the input output behavior conjectured in Eq. (3.15), we have identified the system (together with the observer-in-the-loop) that achieves zero x_k state response after $p+1$ time steps. In other words, y_{k+p} does not depend on the state x_k due to the least squares solution of the linear system of equations in (3.15). Stating the same in a vector – matrix form, for any time step t_k (denoted by k and $\forall k > p$) we have that,

$$\mathbf{y}_k = \begin{bmatrix} D_k & \overline{h}_{k,k-1} & \overline{h}_{k,k-2} & \cdots & \overline{h}_{k,k-p} \end{bmatrix} \begin{bmatrix} \mathbf{u}_k \\ \boldsymbol{\upsilon}_{k-1} \\ \boldsymbol{\upsilon}_{k-2} \\ \vdots \\ \boldsymbol{\upsilon}_{k-p} \end{bmatrix} \tag{3.16}$$

This represents a set of m equations in $m \times (r + p * (r + m))$ unknowns. Notice that, in contrast to the developments using the generalized system Markov parameters, (to relate the input output data sets; refer to discussions around Eq. (3.8), the previous chapter, the companion papers [20, 26, 27] and the references there-in for more information) the number of unknowns remains constant in this case. This makes the computation of observer Markov parameters possible in practice since the number of repeated experiments required to compute these parameters is now ideally constant (derived below) and does not change with the discrete time step t_k. In fact, it is observed that the minimum number if experiments necessary to determine the observer Markov parameters uniquely is $N_{\exp}^{\min} = (r + p * (r + m))$, and from the developments of the subsequent sections (including chapter II and relevant papers), we can say that $N_{\exp}^{\min}$ is also the minimum number of repeated experiments one should perform in order to realize the time varying system models desired from the TVERA.

$$\begin{aligned}
\mathbf{Y}_k &= \begin{bmatrix} \mathbf{y}_k^{(1)} & \mathbf{y}_k^{(2)} & \cdots & \mathbf{y}_k^{(N)} \end{bmatrix} \\
&= \begin{bmatrix} D_k & \bar{h}_{k,k-1} & \bar{h}_{k,k-2} & \cdots & \bar{h}_{k,k-p} \end{bmatrix} \begin{bmatrix} \mathbf{u}_k^{(1)} & \mathbf{u}_k^{(2)} & & \mathbf{u}_k^{(N)} \\ \boldsymbol{\upsilon}_{k-1}^{(1)} & \boldsymbol{\upsilon}_{k-1}^{(2)} & & \boldsymbol{\upsilon}_{k-1}^{(N)} \\ \boldsymbol{\upsilon}_{k-2}^{(1)} & \boldsymbol{\upsilon}_{k-2}^{(2)} & \cdots & \boldsymbol{\upsilon}_{k-2}^{(N)} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{\upsilon}_{k-p}^{(1)} & \boldsymbol{\upsilon}_{k-p}^{(2)} & & \boldsymbol{\upsilon}_{k-p}^{(N)} \end{bmatrix} \\
&= \mathbf{M}_k \mathbf{V}_k
\end{aligned} \tag{3.17}$$

$\forall k > p$.

Therefore the least squares solution for the generalized observer Markov parameters is given for each time step as,

$$\hat{\mathbf{M}}_k = \mathbf{Y}_k \mathbf{V}_k^{\dagger} \tag{3.18}$$

where $(.)^{\dagger}$ denotes the least squares pseudo inverse of a matrix [21, 22]. The calculation of the system Markov parameters and observer gain Markov parameters is detailed in the next section.

Computation of Generalized System Markov Parameters and Observer Gain Sequence

We first outline a process for the determination of system Markov parameter sequence from the observer Markov parameter sequence calculated in the previous section. A recursive relationship is then given to obtain the system Markov parameters with the index difference of greater than p time steps. Similar procedures are set up for observer gain Markov parameter sequences.

Computation of System Markov Parameters from Observer Markov Parameters

Considering the definition of the generalized observer Markov parameters, we write,

$$\begin{aligned} \bar{h}_{k,k-1} &= C_k \bar{B}_{k-1} \\ &= C_k \left[\left(B_{k-1} + G_{k-1} D_{k-1} \right) \quad -G_{k-1} \right] \\ &= \left[h_{k,k-1}^{(1)} \quad -h_{k,k-1}^{(2)} \right] \end{aligned} \tag{3.19}$$

where the superscript $(.)^{(1,2)}$ notation is used to distinguish between the Markov parameter sequences useful to compute the system parameters and the observer gains respectively. Consider the following manipulation written as,

$$\begin{aligned}\bar{h}^{(1)}_{k,k-1} - \bar{h}^{(2)}_{k,k-1} D_{k-1} &= C_k B_{k-1} \\ &= h_{k,k-1}\end{aligned} \tag{3.20}$$

Considering a similar expression for Markov parameters with two time steps between them, we have that,

$$\begin{aligned}\bar{h}^{(1)}_{k,k-2} - \bar{h}^{(2)}_{k,k-2} D_{k-2} &= C_k \bar{A}_{k-1} \bar{B}_{k-2} - C_k \bar{A}_{k-1} G_{k-2} D_{k-2} \\ &= C_k \bar{A}_{k-1} \left(B_{k-2} + G_{k-2} \right) - C_k \bar{A}_{k-1} G_{k-2} D_{k-2} \\ &= C_k \bar{A}_{k-1} B_{k-2} \\ &= C_k \left(A_{k-1} + G_{k-1} C_{k-1} \right) B_{k-2} \\ &= C_k A_{k-1} B_{k-2} + \bar{h}^{(1)}_{k,k-1} C_{k-1} B_{k-2} \\ &= h_{k,k-2} + \bar{h}^{(1)}_{k,k-1} h_{k-1,k-2}\end{aligned} \tag{3.21}$$

This manipulation leads to an elegant expression for the system Markov parameter $h_{k,k-2}$ to be calculated from observer Markov parameters at the time step t_k and the system Markov parameters at previous time steps. This important recursive relationship was found to hold in general and enables the calculation of the system Markov parameters (unadorned $h_{i,j}$) from the observer Markov parameters $\bar{h}_{i,j}$.

To prove this holds in general, consider the induction step with observer Markov parameters (with p time step separation) given by,

$$\begin{aligned}
\overline{h}^{(1)}_{k,k-p} - \overline{h}^{(2)}_{k,k-p} D_{k-p} &= C_k \overline{A}_{k-1} \overline{A}_{k-2} \ldots \overline{A}_{k-p+1} \left(B_{k-p} + G_{k-p} D_{k-p} \right) - C_k \overline{A}_{k-1} \overline{A}_{k-2} \ldots \overline{A}_{k-p+1} G_{k-p} D_{k-p} \\
&= C_k \overline{A}_{k-1} \overline{A}_{k-2} \ldots \overline{A}_{k-p+1} B_{k-p} \\
&= C_k \overline{A}_{k-1} \overline{A}_{k-2} \ldots \overline{A}_{k-p+2} \left(A_{k-p+1} + G_{k-p+1} C_{k-p+1} \right) B_{k-p} \\
&= C_k \overline{A}_{k-1} \overline{A}_{k-2} \ldots \overline{A}_{k-p+2} A_{k-p+1} B_{k-p} + C_k \overline{A}_{k-1} \overline{A}_{k-2} \ldots \overline{A}_{k-p+2} G_{k-p+1} C_{k-p+1} B_{k-p} \\
&= C_k \overline{A}_{k-1} \overline{A}_{k-2} \ldots \overline{A}_{k-p+2} A_{k-p+1} B_{k-p} + \overline{h}^{(2)}_{k,k-p+1} h_{k-p+1,k-p}
\end{aligned} \tag{3.22}$$

Upon careful examination, we find that the term $C_k \overline{A}_{k-1} \overline{A}_{k-2} \ldots \overline{A}_{k-p+2} A_{k-p+1} B_{k-p}$ can be written as,

$$\begin{aligned}
C_k \overline{A}_{k-1} \overline{A}_{k-2} \ldots \overline{A}_{k-p+2} A_{k-p+1} B_{k-p} &= C_k \overline{A}_{k-1} \ldots \overline{A}_{k-p+3} \left(A_{k-p+2} + G_{k-p+2} C_{k-p+2} \right) A_{k-p+1} B_{k-p} \\
&= C_k \overline{A}_{k-1} \ldots A_{k-p+2} A_{k-p+1} B_{k-p} + C_k \overline{A}_{k-1} \ldots G_{k-p+2} C_{k-p+2} A_{k-p+1} B_{k-p} \\
&= C_k \overline{A}_{k-1} \ldots A_{k-p+2} A_{k-p+1} B_{k-p} + \overline{h}^{(2)}_{k,k-p+2} h_{k-p+2,k-p} \\
&= \ldots \\
&= C_k A_{k-1} \ldots A_{k-p+1} B_{k-p} + \overline{h}^{(2)}_{k,k-1} h_{k-1,k-p} + \overline{h}^{(2)}_{k,k-2} h_{k-2,k-p} + \ldots + \overline{h}^{(2)}_{k,k-p+2} h_{k-p+2,k-p} \\
&= h_{k,k-p} + \overline{h}^{(2)}_{k,k-1} h_{k-1,k-p} + \overline{h}^{(2)}_{k,k-2} h_{k-2,k-p} + \ldots + \overline{h}^{(2)}_{k,k-p+2} h_{k-p+2,k-p}
\end{aligned} \tag{3.23}$$

This manipulation enables us to write,

$$\begin{aligned}
\overline{h}^{(1)}_{k,k-p} - \overline{h}^{(2)}_{k,k-p} D_{k-p} &= h_{k,k-p} + \overline{h}^{(2)}_{k,k-1} h_{k-1,k-p} + \overline{h}^{(2)}_{k,k-2} h_{k-2,k-p} + \ldots + \overline{h}^{(2)}_{k,k-p+1} h_{k-p+1,k-p} \\
&= h_{k,k-p} + \sum_{j=1}^{p-1} \overline{h}^{(2)}_{k,k-1} h_{k-1,k-p}
\end{aligned} \tag{3.24}$$

Writing the derived relationships between the system and observer Markov parameters, we have the following set of equations,

$$
\begin{aligned}
&h_{k,k-1} = \bar{h}^{(1)}_{k,k-1} - \bar{h}^{(2)}_{k,k-1} D_{k-1} \\
&h_{k,k-2} + \bar{h}^{(2)}_{k,k-1} h_{k-1,k-2} = \bar{h}^{(1)}_{k,k-2} - \bar{h}^{(2)}_{k,k-2} D_{k-2} \\
&\ldots \\
&h_{k,k-p} + \bar{h}^{(2)}_{k,k-1} h_{k-1,k-p} + \ldots + \bar{h}^{(2)}_{k,k-p+1} h_{k-p+1,k-p} = \bar{h}^{(1)}_{k,k-p} - \bar{h}^{(2)}_{k,k-p} D_{k-p}
\end{aligned}
\tag{3.25}
$$

Defining $r_{i,j} := \bar{h}^{(1)}_{i,j} - \bar{h}^{(2)}_{i,j} D_j$, we obtain the system of linear equations relating the system and observer Markov parameters as,

$$
\begin{bmatrix} I_m & \bar{h}^{(2)}_{k,k-1} & \bar{h}^{(2)}_{k,k-2} & \cdots & \bar{h}^{(2)}_{k,k-p+1} \\ 0 & I_m & \bar{h}^{(2)}_{k-1,k-2} & \cdots & \bar{h}^{(2)}_{k-1,k-p+2} \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I_m \end{bmatrix}
\begin{bmatrix} h_{k,k-1} & h_{k,k-2} & \cdots & h_{k,k-p} \\ 0 & h_{k-1,k-2} & \cdots & h_{k-1,k-p} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_{k-p+1,k-p} \end{bmatrix}
= \begin{bmatrix} r_{k,k-1} & r_{k,k-2} & \cdots & r_{k,k-p} \\ 0 & r_{k-1,k-2} & \cdots & r_{k-1,k-p} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{k-p+1,k-p} \end{bmatrix}
\tag{3.26}
$$

We note the striking similarity of this equation to the relation between observer Markov parameters and the system Markov parameters in the classical OKID algorithm for time invariant systems (compare coefficient matrix of Eq. (3.26) with equation (6.8) of Juang[1]).

Considering the expressions for $\bar{h}_{k,k-p} := C_k \bar{A}_{k-1} \ldots \bar{A}_{k-p+1} \bar{B}_{k-p}$ and choosing p sufficiently large, we see, (owing to the asymptotic stability of the closed loop - including the observer), that $\bar{h}_{k,k-p} \approx 0$. This fact enables us to establish recursive relationships for the calculation of the system Markov parameters $h_{k,k-i}$, $\forall i > p$.

Generalizing the equation (3.24), (to introduce the variability of the order of the GTV-ARX model) we see that,

$$h_{k,k-i} = \overline{h}^{(1)}_{k,k-i} - \overline{h}^{(2)}_{k,k-i} D_{k-i} - \sum_{j=1}^{i-1} \overline{h}^{(2)}_{k,k-j} h_{k-j,k-i} \tag{3.27}$$

$\forall i > p_k$. Then based on the approximation made in equation (3.16) for the calculation of the generalized observer Markov parameters, all the terms with time step separation greater than p vanish identically, and we obtain the relationship,

$$h_{k,k-i} = -\sum_{j=1}^{p} \overline{h}^{(2)}_{k,k-j} h_{k-j,k-i} \tag{3.28}$$

We remind ourselves in passing that this recursive relation in the general case of variability in the GTV-ARX model order depending on the time step, the corresponding recursions to evaluate time varying system Markov parameters should reflect such a variability, that is to say, for each $i > p_k$

$$h_{k,k-i} = -\sum_{j=1}^{p_k} \overline{h}^{(2)}_{k,k-j} h_{k-j,k-i} \tag{3.29}$$

For maintaining the simplicity of the presentations here-in, we will not make further references to the variable order option in the subsequent developments of the chapter/dissertation. Insight in to the flexibility is provided by appealing to the relations of the identified observer with a linear time varying Kalman filter in the next section of this chapter.

Computation of Observer Gain Markov Parameters from the Observer Markov Parameters

Consider the generalized observer gain Markov parameters defined as,

$$h^{o}_{k,i} = \begin{cases} C_k A_{k-1} A_{k-2} \ldots A_{i+1} G_i \ , & \forall k > i+1 \\ C_k G_{k-1}, & k = i+1 \\ 0, & \forall k < i+1 \end{cases} \tag{3.30}$$

We now derive the relationship between these parameters and the time varying ARX model coefficients, $\overline{h}^{(2)}_{k,j}$. These parameters will be used in the calculation of the observer gain sequence from the input output data in the next subsection, an elegant generalization of the time invariant relations obtained in [1, 10] similar to equation (3.26).

From their corresponding definitions, we note that

$$\overline{h}^{(2)}_{k,k-1} = C_k G_{k-1} = h^{o}_{k,k-1} \tag{3.31}$$

Similarly,

$$\begin{aligned} \overline{h}^{(2)}_{k,k-2} &= C_k \overline{A}_{k-1} G_{k-2} \\ &= C_k \left(A_{k-1} + G_{k-1} C_{k-1} \right) G_{k-2} = h^{o}_{k,k-2} + \overline{h}^{(2)}_{k,k-1} h^{o}_{k-1,k-2} \end{aligned} \tag{3.32}$$

We find that in general an induction step similar to equation (3.22) holds and is given by,

$$\begin{aligned}
\overline{h}^{(2)}_{k,k-p} &= C_k \overline{A}_{k-1}\overline{A}_{k-2}...\overline{A}_{k-p+1}G_{k-p} \\
&= C_k \overline{A}_{k-1}\overline{A}_{k-2}...\overline{A}_{k-p+2}\left(A_{k-p+1}+G_{k-p+1}C_{k-p+1}\right)G_{k-p} \\
&= C_k \overline{A}_{k-1}\overline{A}_{k-2}...\overline{A}_{k-p+2}A_{k-p+1}G_{k-p}+C_k \overline{A}_{k-1}\overline{A}_{k-2}...\overline{A}_{k-p+2}G_{k-p+1}C_{k-p+1}G_{k-p} \\
&= C_k \overline{A}_{k-1}\overline{A}_{k-2}...\overline{A}_{k-p+2}A_{k-p+1}G_{k-p}+\overline{h}^{(2)}_{k,k-p+1}h^{o}_{k-p+1,k-p} \\
&= h^{o}_{k,k-p}+\overline{h}^{(2)}_{k,k-1}h^{o}_{k-1,k-p}+...+\overline{h}^{(2)}_{k,k-p+1}h^{o}_{k-p+1,k-p}
\end{aligned} \tag{3.33}$$

where we used the identity derived in the equation (3.23) (replace B_{k-p} in favor of G_{k-p}).

This enables us to write the general relationship,

$$\overline{h}^{(2)}_{k,k-j} = h^{o}_{k,k-j} + \sum_{i=1}^{j-1} \overline{h}^{(2)}_{k,k-i} h^{o}_{k-i,k-j} \tag{3.34}$$

$\forall j \in \mathbb{Z}^{+}$ analogous to relation (3.27) in case of the system Markov parameters. Also, similar to (3.28) we have the appropriate recursive relationship for the observer gain Markov parameters separated by more than p time steps for each k given as,

$$h^{o}_{k,k-j} = -\sum_{i=1}^{p} \overline{h}^{(2)}_{k,k-i} h^{o}_{k-i,k-j} \tag{3.35}$$

$\forall j > p_k$. Therefore to calculate the observer gain Markov parameters we have a similar upper block – triangular system of linear equations which can be written as,

$$\begin{bmatrix} I_m & \bar{h}^{(2)}_{k,k-1} & \bar{h}^{(2)}_{k,k-2} & \cdots & \bar{h}^{(2)}_{k,k-p+1} \\ 0 & I_m & \bar{h}^{(2)}_{k-1,k-2} & \cdots & \bar{h}^{(2)}_{k-1,k-p+2} \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I_m \end{bmatrix} \begin{bmatrix} h_{k,k-1} & h_{k,k-2} & \cdots & h_{k,k-p} \\ 0 & h_{k-1,k-2} & \cdots & h_{k-1,k-p} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_{k-p+1,k-p} \end{bmatrix}$$
$$= \begin{bmatrix} \bar{h}^{(2)}_{k,k-1} & \bar{h}^{(2)}_{k,k-2} & \cdots & \bar{h}^{(2)}_{k,k-p} \\ 0 & \bar{h}^{(2)}_{k-1,k-2} & \cdots & \bar{h}^{(2)}_{k-1,k-p} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \bar{h}^{(2)}_{k-p+1,k-p} \end{bmatrix} \tag{3.36}$$

to be solved at each time step k. Having outlined a method to compute the observer gain Markov parameters, let us now proceed to look at the procedure to extract the observer gain sequence from them.

Calculation of the Realized Time Varying Observer Gain Sequence

From the definition of the observer gain Markov parameters, (recall equation (3.30)) we can stack the first few parameters in a tall matrix and observe that,

$$\begin{aligned} P_{k+1} &:= \begin{bmatrix} h^o_{k+1,k} \\ h^o_{k+2,k} \\ \vdots \\ h^o_{k+m,k} \end{bmatrix} \\ &= \begin{bmatrix} C_{k+1}G_k \\ C_{k+2}A_{k+1}G_k \\ \vdots \\ C_{k+m}A_{k+m-1}\ldots A_{k+1}G_k \end{bmatrix} = \begin{bmatrix} C_{k+1} \\ C_{k+2}A_{k+1} \\ \vdots \\ C_{k+m}A_{k+m-1}\ldots A_{k+1} \end{bmatrix} G_k \\ &= O^{(m)}_{k+1}G_k \end{aligned} \tag{3.37}$$

Such that a least squares solution for the gain matrix at each time step is given by,

$$G_k = O_{k+1}^{(m)\dagger} P_k \tag{3.38}$$

However from the discussions about coordinate transformations in the previous chapter, we find that it is indeed impossible to determine the observability grammian in the true coordinate system[26], as suggested by Eq. (3.38) above. The computed (rather decomposed from the generalized Hankel matrix) observability grammian is, in general in a time varying and unknown coordinate system denoted by, $O_{k+1}^{(m)T_{k+1}}$ at the time step t_{k+1}. We will now show that the gain computed from this time varying observability grammian (computed) will be consistent with the time varying coordinates of the plant model computed by the Time Varying Eigensystem Realization Algorithm (TVERA) presented in the previous chapter. Therefore upon using the computed observability grammian (in its own time varying coordinate system) and proceeding with the gain calculation as indicated by the Eq. (3.38) above, we arrive at a consistent computed gain matrix. That is to say that,

$$\begin{aligned} P_{k+1} &= O_{k+1}^{(m)} G_k \\ &= O_{k+1}^{(m)} T_{k+1} T_{k+1}^{-1} G_k \\ &= O_{k+1}^{(m)T_{k+1}} T_{k+1}^{-1} G_k \\ &= O_{k+1}^{(m)T_{k+1}} \hat{G}_k \end{aligned} \tag{3.39}$$

Such that,

$$\hat{G}_k = T_{k+1}^{-1} G_k = \left(O_{k+1}^{(m)T_{k+1}} \right)^\dagger P_{k+1} \tag{3.40}$$

therefore, with no explicit intervention by the analyst, the realized gains are automatically in the right coordinate system for producing the appropriate time varying OKID closed loop. For consistency, it is often convenient, if one obtains the first few time step models as included in the developments of the previous chapter. This automatically gives the observability grammians for the first few time steps to calculate the corresponding observer gain matrix values. To see that the gain sequence computed by the algorithm is indeed in consistent coordinate systems, recall the identified system, control influence and the measurement sensitivity matrices in the time varying coordinate systems, to be derived as (refer to the previous chapter and related paper[26]) :

$$\begin{aligned} \hat{A}_k &= T_{k+1}^{-1} A_k T_k \\ \hat{B}_k &= T_{k+1}^{-1} B_k \\ \hat{C}_k &= C_k T_k \end{aligned} \tag{3.41}$$

The time varying OKID closed loop system matrix, with the realized gain matrix sequence is seen to be consistently given as,

$$\hat{A}_k + \hat{G}_k \hat{C}_k = T_{k+1}^{-1} \left(A_k + G_k C_k \right) T_k \tag{3.42}$$

in a kinematically similar fashion to the true time varying OKID closed loop. The nature of the computed stabilizing (time varying dead – beat) gain sequence are best viewed from a reference coordinate system as opposed to the time varying coordinate systems computed by the algorithm. The projection based transformations can be used for this purpose and are discussed in the previous chapter.

Relationship Between the Identified Observer and a Kalman Filter

We now qualitatively discuss several features of the observer realized from the algorithm presented in the previous section. Constructing the closed loop of the observer dynamics, it can be found to be asymptotically stable as purported at the design stage. Following the developments of the time invariant OKID algorithm[10], we use the well understood time varying Kalman filter theory to make some intuitive observations. These observations help us to further our understanding of important issues concerning how the newly developed GTV-ARX model is a potential generalization/extension of the ARX model well known in the time invariant case. Insight is also obtained as to what happens in the presence of measurement noise. An immediate intuitive leap one can make is that in the practical situation where there is process and measurement noise in the data, the GTV-ARX model becomes a moving average model that can be termed as the GTV-ARMAX (Generalized time varying autoregressive moving average with exogenous input) model (generalized is used to indicate variable order at each time step). A detailed quantitative examination of this situation is beyond the scope of the current investigation and the author limits his discussions to possible speculations on the qualitative relations.

The well known Kalman filter equations for a truth model given in equation (A.1) of the appendix B are given by,

$$\begin{aligned}&\hat{\mathbf{x}}_{k+1}^{-} = A_k \hat{\mathbf{x}}_k^{+} + B_k \mathbf{u}_k \\ &\text{or} \\ &\hat{\mathbf{x}}_{k+1}^{-} = A_k \left[I - K_k C_k \right] \hat{\mathbf{x}}_k^{-} + B_k \mathbf{u}_k + A_k K_k \mathbf{y}_k\end{aligned} \tag{3.43}$$

together with the propagated output equation,

$$\hat{\mathbf{y}}_k^- = C_k \hat{\mathbf{x}}_k^- + D_k \mathbf{u}_k \tag{3.44}$$

where the gain K_k is optimal (expression in equation(A.16)). As documented in the standard estimation theory text books, optimality translates to any one of the equivalent necessary conditions of minimum variance, maximum likelihood, orthogonality or Bayesian schemes well known for linear estimation problems. A brief review of the expressions for the optimal gain sequence is derived in the appendix which also provides an insight into the useful notion of orthogonality of the discrete innovations process, in addition to deriving an expression for the optimal gain matrix sequence (ref. equation (A.16) for an expression for the optimal gain). From an input-output standpoint the innovations approach provides the most insight for analysis and is used in this section. Using the definition of the innovations process $\boldsymbol{\varepsilon}_k := \mathbf{y}_k - \hat{\mathbf{y}}_k^-$, the measurement equation of the estimator in (3.44) can be written in favor of the system outputs as given by,

$$\mathbf{y}_k = C_k \hat{\mathbf{x}}_k^- + D_k \mathbf{u}_k + \boldsymbol{\varepsilon}_k \tag{3.45}$$

Rearranging the state propagation and update equation of (3.43), we arrive at a form given by,

$$\begin{aligned} \hat{\mathbf{x}}_{k+1}^- &= A_k \left[I - K_k C_k \right] \hat{\mathbf{x}}_k^- + B_k \mathbf{u}_k + A_k K_k \mathbf{y}_k \\ &= \tilde{A}_k \hat{\mathbf{x}}_k^- + \tilde{B}_k \mathbf{v}_k \end{aligned} \tag{3.46}$$

with the definitions,

$$\begin{aligned} \tilde{A}_k &= A_k \left[I - K_k C_k \right] \\ \tilde{B}_k &= \left[B_k \quad A_k K_k \right] \\ \mathbf{v}_k &= \begin{bmatrix} \mathbf{u}_k \\ \mathbf{y}_k \end{bmatrix} \end{aligned} \tag{3.47}$$

Notice the structural similarity in the layout of the rearranged equations to the time varying OKID equations in the previous sections. This rearrangement helps in making comparisons and observations as to what are the conditions in which we actually manage to obtain the Kalman filter gain sequence.

Starting from the initial condition, the input-output relation of the Kalman filter equations can be written as,

$$\begin{aligned} \mathbf{y}_0 &= C_0 \hat{\mathbf{x}}_0^- + D_0 \mathbf{u}_0 + \boldsymbol{\varepsilon}_0 \\ \mathbf{y}_1 &= C_1 \tilde{A}_0 \hat{\mathbf{x}}_0^- + D_1 \mathbf{u}_1 + C_1 \tilde{B}_0 \mathbf{v}_0 + \boldsymbol{\varepsilon}_1 \\ \mathbf{y}_2 &= C_2 \tilde{A}_1 \tilde{A}_0 \hat{\mathbf{x}}_0^- + D_2 u_2 + C_2 \tilde{B}_1 \mathbf{v}_1 + C_2 \tilde{A}_1 \tilde{B}_0 \mathbf{v}_0 + \boldsymbol{\varepsilon}_2 \\ &\cdots \\ \mathbf{y}_p &= C_p \tilde{A}_{p-1} \ldots \tilde{A}_0 \hat{\mathbf{x}}_0^- + D_p \mathbf{u}_p + \sum_{j=1}^{p-1} \tilde{h}_{p,p-j} \boldsymbol{\upsilon}_{p-j} + \boldsymbol{\varepsilon}_p \\ &\cdots \end{aligned} \tag{3.48}$$

suggesting the general relationship,

$$\mathbf{y}_{k+p} = C_{k+p} \tilde{A}_{k+p-1} \ldots \tilde{A}_0 \hat{\mathbf{x}}_0^- + D_{k+p} \mathbf{u}_{k+p} + \sum_{j=1}^{k+p-1} \tilde{h}_{k+p,k+p-j} \boldsymbol{\upsilon}_{k+p-j} + \boldsymbol{\varepsilon}_{k+p} \tag{3.49}$$

with the Kalman filter Markov parameters $\tilde{h}_{k,i}$ being defined by,

$$\tilde{h}_{k,i} = \begin{cases} C_k \tilde{A}_{k-1} \tilde{A}_{k-2} \ldots \tilde{A}_{i+1} \tilde{B}_i \,, & \forall k > i+1 \\ C_k \tilde{B}_{k-1}, & k = i+1 \\ 0, & \forall k < i+1 \end{cases} \tag{3.50}$$

Comparing the equations (3.13) and (3.49) we conclude that their input-output representations are identical for a suitable choice of p (i.e., $\forall k > p$), if $G_k = -A_k K_k$ together with the additional condition that $\boldsymbol{\varepsilon}_k = 0, \forall k > p$. Therefore under these conditions our algorithm is expected to produce a gain sequence that is optimal. In the presence of noise in the output data, the additional requirement is to satisfy the orthogonality (innovations property) of the residual sequence, as derived in the appendix.

However, we proceeded to enforce the p (in general p_k) term dependence in equation (3.15) using the additional freedom obtained due to the variability of the time varying observer gains. This enabled us to minimize the number of repeated experiments and the number of computations while also arriving at the fastest observer gain sequence owing to the definitions of time varying dead beat observer notions set up in this dissertation (Appendix C). Notice that the Kalman filter equations are in general not truncated to the first $p\left(p_k\right)$ terms. An immediate question arises as to whether we can ever obtain the "optimal" gain sequence using the truncated representation for gain calculation.

To answer this question qualitatively, we consider the input output behavior of the true Kalman filter in (3.49). Observe that Kalman gains can indeed be constructed so as to obtain matching truncated representations as the GTV-ARX (more precisely GTV – ARMAX) model as in equation (3.15) via the appropriate choice of the tuning parameters P_0 , Q_k . In the GTV-ARMAX parlance using a lower order for p_k (at any given time step) means the incorporation of a forgetting factor which in the Kalman filter framework is tantamount to using larger values for the process noise parameter Q_k (at the same time instant). Therefore, the generalized time varying ARX and ARMA models used for the observer gain sequence and the system Markov parameter sequence in the algorithmic developments of this chapter are intimately tied in to the tuning parameters of the Kalman filter and represent the fundamental balance existing in statistical learning theory between ignorance of the model for the dynamical system and incorporation of new information from measurements. Further research is required to develop a more quantitative relation between the observer identified using the developments of the paper and the time varying Kalman filter gain sequence.

Numerical Example

We now detail the problem of computing the generalized system Markov parameters from the computed observer Markov parameters as outlined in the previous section. Consider the same system presented as one of the example problems in the previous chapter. It has an oscillatory nature and does not have a stable origin. The truth model of the plant system matrix (recall example 2 from chapter II, repeated here for convenience) as

$$
\begin{aligned}
A_k &= \exp\left[A_c * \Delta t\right] \\
B_k &= \begin{bmatrix} 1 & 0 \\ 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad C_k = \begin{bmatrix} 1 & 0 & 1 & 0.2 \\ 1 & -1 & 0 & -0.5 \end{bmatrix}, \\
D_k &= 0.1\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\end{aligned}
\tag{3.51}
$$

where the matrix is given by

$$
A_c = \begin{bmatrix} 0_{2\times 2} & I_{2\times 2} \\ -K_t & 0_{2\times 2} \end{bmatrix}
\tag{3.52}
$$

with $K_t = \begin{bmatrix} 4+3\tau_k & -1 \\ -1 & 7+3\tau'_k \end{bmatrix}$ and τ_k, τ'_k are defined as $\tau_k = \sin\left(10t_k\right), \quad \tau'_k := \cos\left(10t_k\right)$.

The time varying OKID algorithm, as described in the previous sections of this chapter is applied to this example to calculate the system Markov parameters and the observer gain Markov parameters from the simulated repeated experimental data. The system Markov parameters thus computed are used by the TVERA algorithm of the previous chapter to realize system model sequence for all the time steps for which experimental data is available. We demonstrate the computations of the algorithm using the time varying dead-beat observer where the smallest order for the GTV-ARX model is chosen throughout the time history of the identification process. Appendix C details the definition of time varying dead beat observer, for the convenience of the readers along with a representative closed loop sequence result using the example problem presented in this section. Relating to the discussions of the previous section, intuitively, the dead beat observer realized by the computations mean that the process noise is set very high as the forgetting factor of the GTV-ARX model is implied to be largest possible for unique identification of the coefficients. The time history of the open loop and the closed loop eigenvalues as viewed from the coordinate system of the initial condition response decomposition is plotted in the Figure 16.

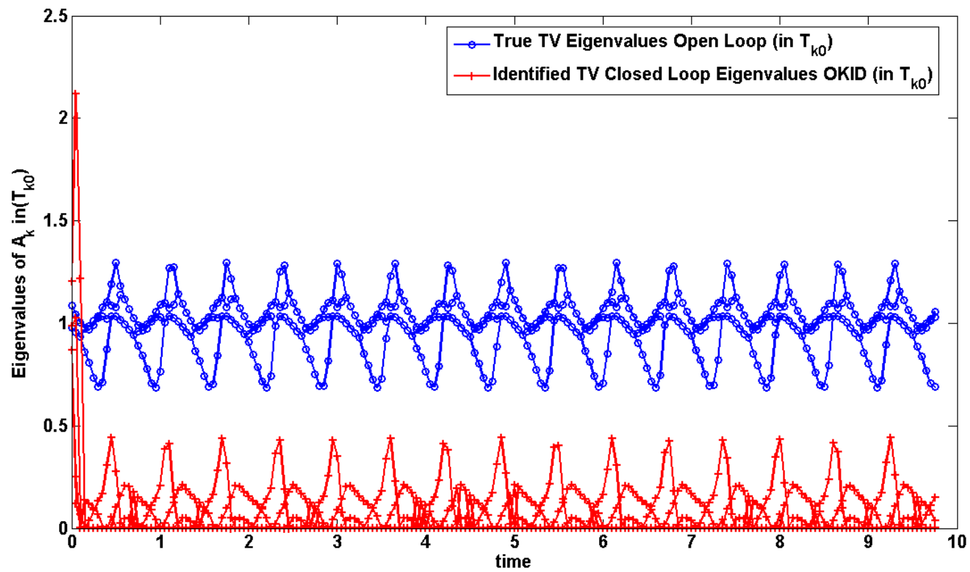

Figure 16. Case 1: Plant Open Loop vs. OKID Closed Loop Pole Locations (Minimum No of Repeated Experiments)

The accuracy of the Markov parameters computed using OKID algorithm presented in the current chapter is compared with the accuracy of the Markov parameters computed using the least squares solution presented in the previous chapter. The agreement is remarkable, as plotted in Figure 17 (especially considering the fact that only $N_{\exp}^{\min} = \left(r + p*\left(r+m\right)\right) = 10$ experiments were performed). Performing larger number of experiments in general leads to similar level of accuracy as shown in Figure 18. However in this case the fastest observer is not realized.

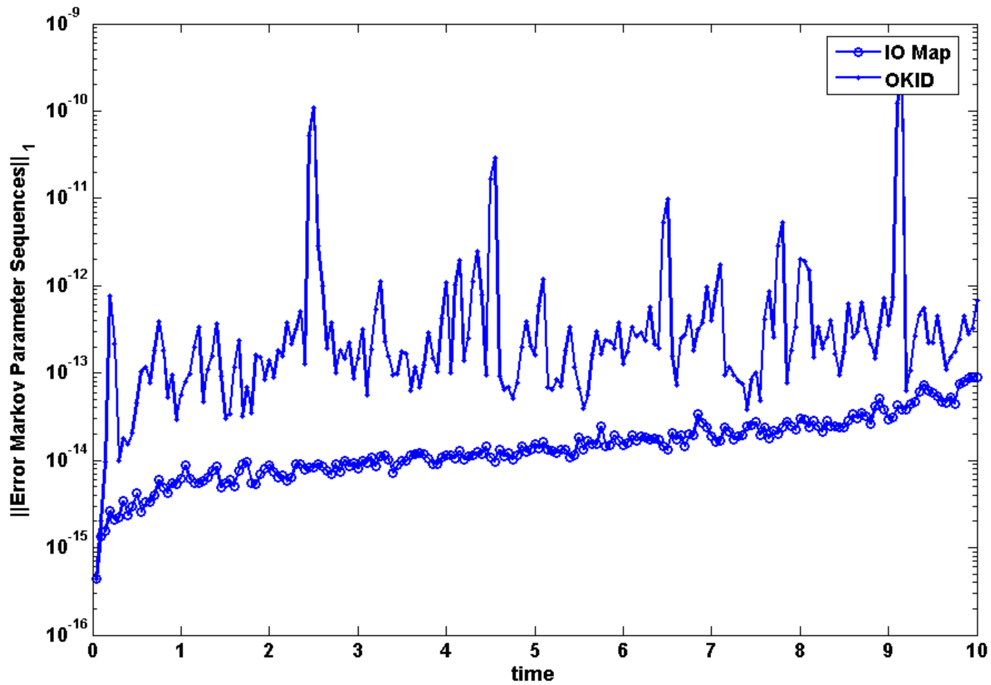

Figure 17. Case 1: Error in System Markov Parameter Calculations (Minimum No of Repeated Experiments = 10)

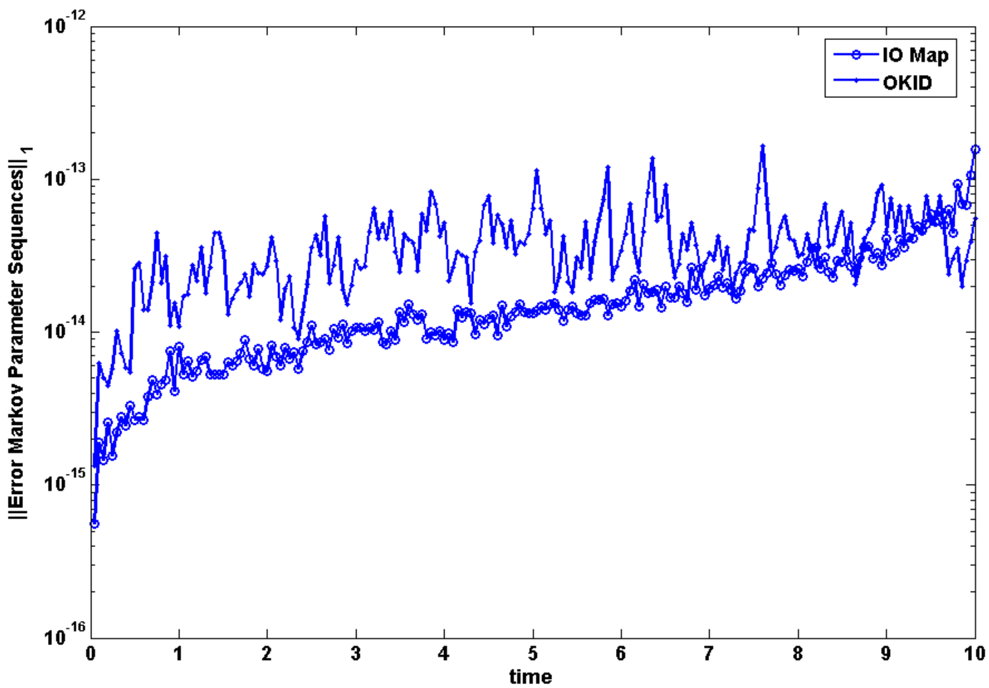

Figure 18. Case 2: Error in Markov Parameters Computations (Non-Minimum Number of Repeated Experiments)

The error incurred in the calculation of the system Markov parameter is directly reflected in the output error between the computed and true system response to test functions. It was found to be of the same order of magnitude (and never greater) in several representative situations incorporating various test case truth models. The corresponding output error plots for Markov parameters with error profiles plotted in Figure 17 and Figure 18 are shown in Figure 19 and Figure 20 respectively.

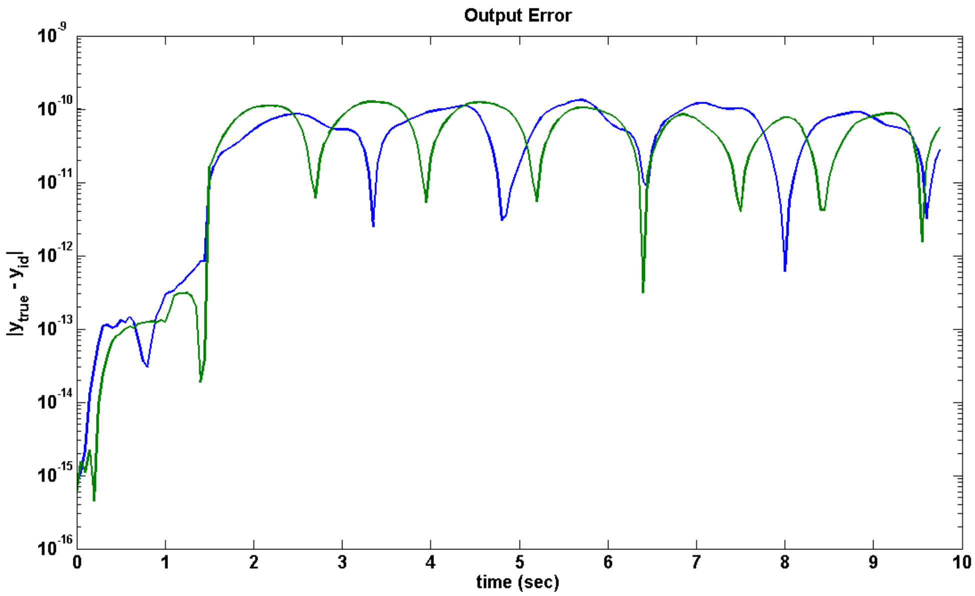

Figure 19. Case 1: Error in Outputs (Minimum No of Repeated Experiments)

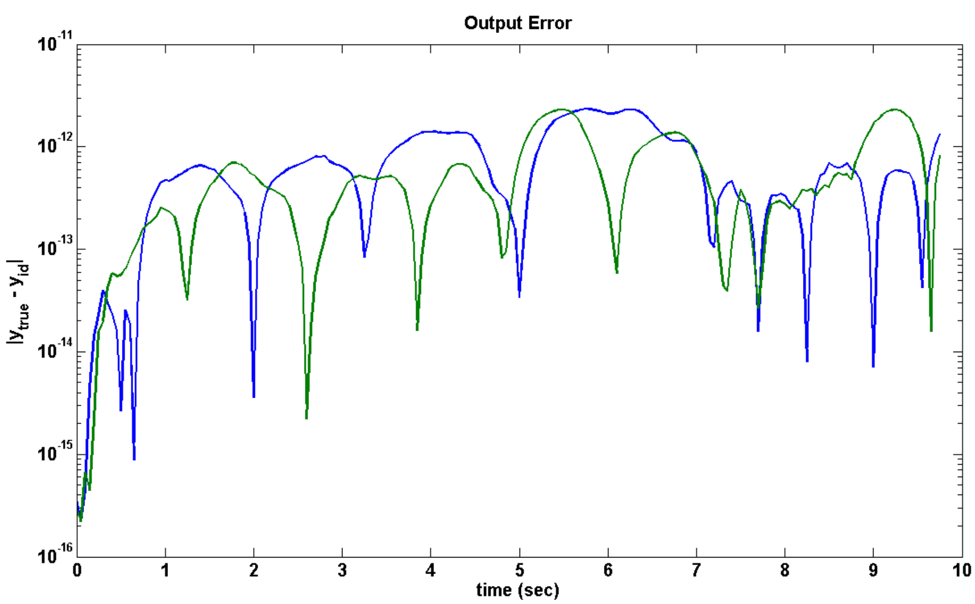

Figure 20. Case 2: Error in Outputs for Test Functions (True vs. Identified Plant Model)

Because the considered system is unstable (oscillatory) in nature, the initial condition response was used to check the nature of decay of the loop system in the presence of the identified observer. The open loop response of the system (with no observer in the loop) and the closed loop response including the realized observer is plotted in Figure 21. Note the oscillatory nature of the open loop outputs (demonstrating the instability of the origin of the system under consideration) while the closed loop system response decays in precisely *two time* steps to zero response. This decay to zero was exponential and too steep to plot for the (time varying) dead beat case. However when the order was chosen to be slightly higher (near dead beat observer is realized in this case and therefore it takes more than two steps for the response to decay to zero), a log scale plot of the magnitudes of output channels shows the steepness of decay of the initial condition response for the open loop and closed loop (with OKID realized observer in the loop) system. This is shown as a demonstration for the achievement of near-exponential time varying feedback stabilization of the origin of the closed loop system. The gain history of the realized time varying observer gains as seen in the initial condition coordinate system is plotted as Figure 22.

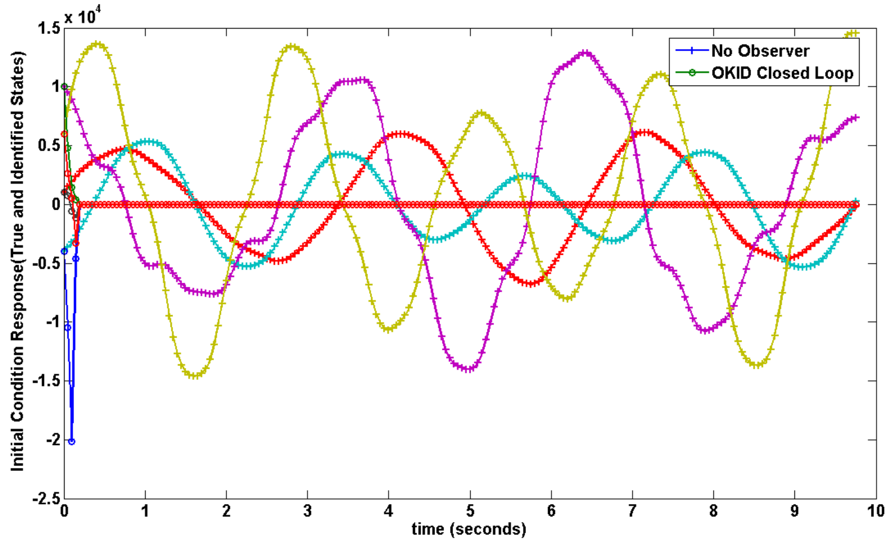

Figure 21. Case 1: Open Loop vs. OKID Closed Loop Response to Initial Conditions

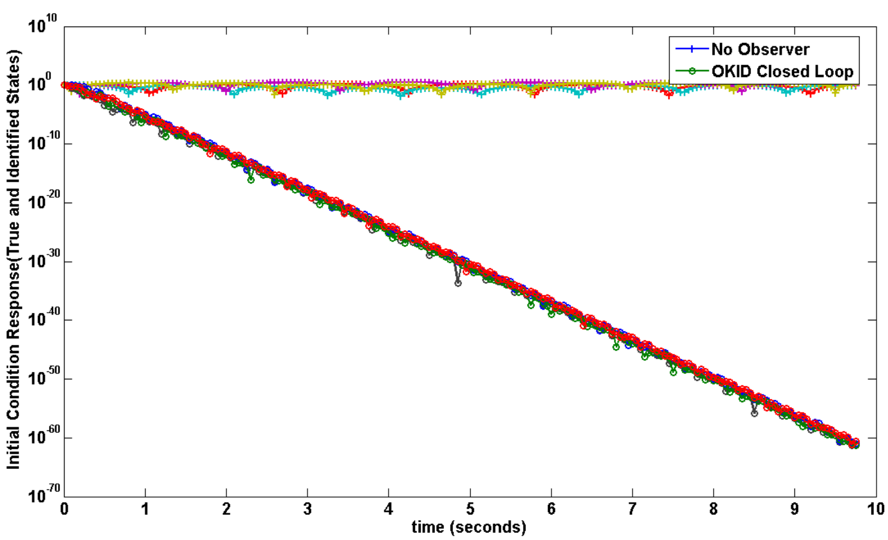

Figure 22. Closed Loop vs. Open Loop Response for a Test Situation Showing Exponential Decay (p=4)

The history of the time varying gains realized in the calculations outlined in this chapter is plotted in the Figure 23. As shown in the theoretical developments of this chapter the gains thus realized were found to be in the time varying coordinate systems. Here we plot the gains after transformation into the reference coordinates.

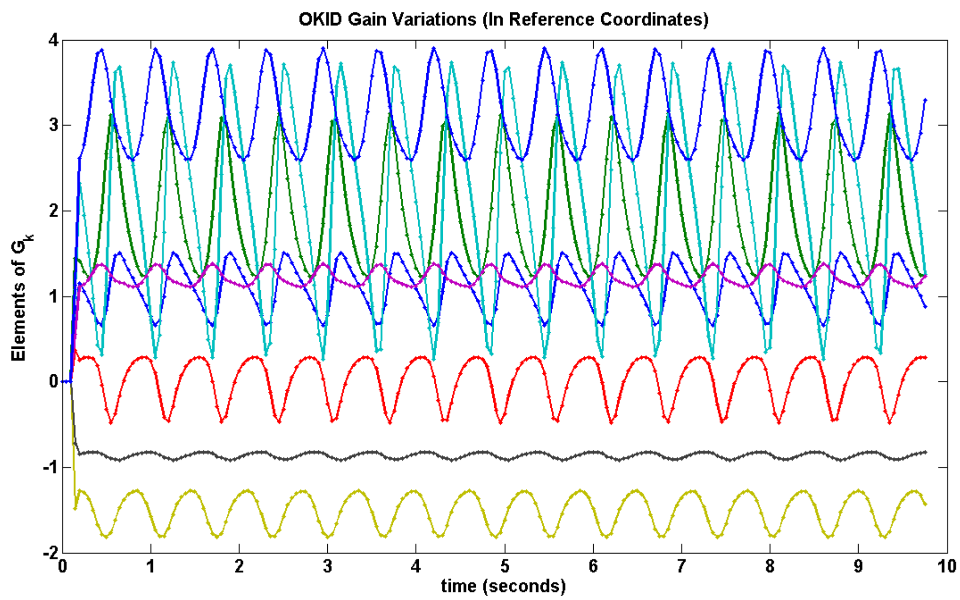

Figure 23. Case 1: Gain History (Minimum No. of Repeated Experiments)

Conclusion

This chapter provides an algorithm for efficient computation of system Markov parameters for use in time varying system identification algorithms. An observer is inserted in the input – output relations and this leads to effective utilization of the data in computation of the system Markov parameters. As a byproduct one gets an observer gain sequence in the same coordinate system as the system models realized by the time varying system identification algorithm. The efficiency of the method in bringing down the number of experiments and computations involved is improved further by truncation of the number of significant terms in the input output description of the closed loop observer, providing a time varying dead beat observer gain sequence. In addition to the flexibility achieved in using a time varying ARX model, it is shown that one could indeed use models with variable order. Relationship with a Kalman filter is detailed from an input-output stand point. It is shown that the flexibility of variable order moving average model realized in the time varying OKID computations is related to the forgetting factor introduced by the process noise tuning parameter of the Kalman filter. The working of the algorithm is demonstrated using a simple example problem.

CHAPTER IV

CONTINUOUS-TIME BILINEAR SYSTEM IDENTIFICATION

Introduction

As we observed in the introduction chapter, advances in sensor and actuator technology of the 21st century has lead to the confluence of many model based state estimation and control strategies for high performance of dynamical systems. Methods for realizing linear time invariant models of dynamical systems have matured and over the past two decades, some advances have been made to realize linear time varying system models and bring the broad field of linear system identification to a mature state of development. However, much work remains to be done in case of the realization of nonlinear models. The simplest nonlinear system model is of a bilinear plant model with a state control input coupling term in addition to a linear term.

Basic Formulation

Following the notations of Juang[23], consider the state vector being denoted by $\mathbf{x}(t) \in \mathbb{R}^n$; together with the control input being denoted by $\mathbf{u}(t) \in \mathbb{R}^r$. The model governing the state evolution for the class of bilinear systems we consider in this dissertation to be given by the ordinary differential equations,

$$\dot{\mathbf{x}} = A_c \mathbf{x} + B_c \mathbf{u} + \sum_{i=1}^{r} N_{ci} \mathbf{x} u_i \tag{4.1}$$

where the system matrix $A_c \in \mathbb{R}^{n\times n}$, control influence matrix $B_c \in \mathbb{R}^{n\times r}$ and the i^{th} state-control coupling matrix $N_{ci} \in \mathbb{R}^{n\times n}$ are assumed to be constant functions of time (time-invariant). We assume the measurement equations are given by the linear measurement model,

$$\mathbf{y}(t) = C\mathbf{x}(t) + D\mathbf{u}(t) \tag{4.2}$$

where $\mathbf{y} \in \mathbb{R}^{m\times 1}$ is the output matrix relating the instantaneous state and control input values to the output, together with the time invariant direct transmission matrix $D \in \mathbb{R}^{m\times r}$ and the measurement sensitivity matrix $C \in \mathbb{R}^{m\times n}$. The algorithm presented here-in relies on the central observation made by Juang [23] that the bilinear system of equations becomes a linear time invariant system upon the application of constant forcing functions. We now present the solution of the bilinear system of equations and show that while the general input output behavior is indeed nonlinear, we can generate an analytical solution for a set of specified inputs.

Bilinear System of Equations: Solution, Input - Output Relations

Considering constant control inputs between any two successive time steps, i.e., $\mathbf{u}(t) = \mathbf{u}^k, \forall t \in [t_k, t_{k+1})$ similar to zero order hold approximation for discretization of a continuous time plant model, we have that the bilinear system becomes a linear time invariant system given by

$$\begin{aligned}\dot{\mathbf{x}} &= A_c\mathbf{x} + B_c\mathbf{u}^k + \sum_{i=1}^{r} N_{ci}\mathbf{x}u_i^k \\ &= \left(A_c + \sum_{i=1}^{r} N_{ci}u_i^k\right)\mathbf{x} + B_c\mathbf{u}^k\end{aligned} \tag{4.3}$$

with initial conditions being given by $\mathbf{x}(t_k) = \mathbf{x}_k$. Solution for the state vector, $\mathbf{x}(t)$ is therefore given by,

$$\begin{aligned}\mathbf{x}(t) &= e^{\left(A_c + \sum_{i=1}^{r} N_{ci}u_i^k\right)(t-t_k)}\mathbf{x}_k + \left[\int_{t_k}^{t_{k+1}} e^{\left[A_c + \sum_{i=1}^{r} N_{ci}u_i^k\right](t_{k+1}-\tau)} d\tau\right] B_c\mathbf{u}^k \\ &= A(\mathbf{u}^k)\mathbf{x}_k + B(\mathbf{u}^k)\mathbf{u}^k\end{aligned} \tag{4.4}$$

$\forall t \in [t_k, t_{k+1})$.

In the more general case, $\mathbf{u}(t)$ is not a constant but a known function of time and differential equations similar to (4.3) result with an additional complexity that the system matrix is time varying. These equations can be written as,

$$\dot{\mathbf{x}} = \left(A_c + \sum_{i=1}^{r} N_{ci}u_i^k(t)\right)\mathbf{x} + B_c\mathbf{u}^k(t) \tag{4.5}$$

The solution in this case is given by[4, 15],

$$\mathbf{x}(t) = \Phi(t, t_k)\mathbf{x}_k + \int_{t_k}^{t} \Phi(t, \tau) B_c\mathbf{u}^k(\tau)\, d\tau \tag{4.6}$$

where the state transition matrix $\Phi(t, t_k)$ is the solution of the matrix differential equation,

$$\begin{aligned}\dot{\Phi}\left(t,t_k\right) &= \left(A_c + \sum_{i=1}^{r} N_{ci} u_i^k\left(t\right)\right)\Phi\left(t,t_k\right) \\ &= F_c\left(t,\mathbf{u}\left(t\right)\right)\Phi\left(t,t_k\right)\end{aligned} \tag{4.7}$$

with initial conditions, $\Phi\left(t_k,t_k\right) = I$. An alternative solution to the state transition matrix differential equation, (for solution of the linear time varying dynamical systems) is given by the Peano-Baker series representation[23]. Writing the solution of the redefined state transition matrix $\Phi\left(t,t_k\right)$ in terms of the Peano-Baker series, we have,

$$\begin{aligned}\Phi\left(t,t_k\right) = I + \int_{t_k}^{t} F_c\left(\tau_1\right)d\tau_1 + \int_{t_k}^{t} F_c\left(\tau_1\right)\int_{t_k}^{\tau_1} F_c\left(\tau_2\right)d\tau_2 d\tau_1 + \ldots \\ + \int_{t_k}^{t} F_c\left(\tau_1\right)\int_{t_k}^{\tau_1} F_c\left(\tau_2\right)\ldots\int_{t_k}^{\tau_{n-1}} F_c\left(\tau_n\right)d\tau_n \ldots d\tau_2 d\tau_1 + \ldots\end{aligned} \tag{4.8}$$

where the implicit dependence of the system matrix $F_c\left(t,\mathbf{u}\left(t\right)\right)$, on the input signal has been suppressed for brevity. Substituting the series expansion solution for the state transition matrix in to the expression governing the outputs of the dynamical system, in equation (4.6) we have,

$$\begin{aligned}\mathbf{y}\left(t\right) = C\Phi\left(t,t_k\right)\mathbf{x}_k + D\mathbf{u}^k\left(t\right) + \int_{t_k}^{t} CB_c\mathbf{u}^k\left(t\right) + \int_{t_k}^{t}\int_{\tau}^{\tau_1} CF_c\left(\tau_1\right)B_c\mathbf{u}^k\left(\tau\right)d\tau_1 d\tau + \ldots \\ + \int_{t_k}^{t} C\left[\int_{\tau}^{t}\int_{\tau}^{\tau_1}\int_{\tau}^{\tau_{n-1}} F_c\left(\tau_1\right)F_c\left(\tau_2\right)\ldots F_c\left(\tau_n\right)d\tau_n\ldots d\tau_2 d\tau_1 + \ldots\right]B_c\mathbf{u}^k\left(\tau\right)\,d\tau + \ldots\end{aligned} \tag{4.9}$$

with the assumed definition that $F_c\left(t\right) = F_c\left(t,\mathbf{u}^k\left(t\right)\right) = \left(A_c + \sum_{i=1}^{r} N_{ci} u_i^k\left(t\right)\right)$.

Therefore clearly for the bilinear problem, the input output behavior is not linear. This series representation of the nonlinear map between the inputs and outputs is known as the Volterra Series [28, 29] representation. It is further noted that, depending on the approximation of interest (eg., zero-order-hold) one can obtain an equivalent discrete time series representation by evaluation of the integrals of the continuous time series of equation (4.9). Having derived the generally nonlinear input-output relationship, we now investigate the zero state response of the nonlinear system subject to piecewise constant inputs (i.e., similar to zero-order hold approximation).

Some Response Characteristics

From the solution of the bilinear state equations developed in the previous subsection, several important relations between the inputs and the response characteristics can be derived. The response characteristics detailed in the following are similar to the relations developed by Juang[23] and are useful in the development of our identification solution. The steps are carried out starting with equation (4.4), since the zero order hold assumption on inputs renders the system to assume a discrete time varying structure amenable for response characterization. Writing the outputs at discrete time instances (i.e., $\mathbf{y}_k := \mathbf{y}\left(t_k\right), \forall k = 0,1,...$) we have the following sequence of expressions for the output,

$$
\begin{aligned}
\mathbf{y}_0 &= C\mathbf{x}_0 + D\mathbf{u}^0 \\
\mathbf{y}_1 &= CA\left(\mathbf{u}^0\right)\mathbf{x}_0 + CB\left(\mathbf{u}^0\right)\mathbf{u}^0 + D\mathbf{u}^1 \\
\mathbf{y}_2 &= CA\left(\mathbf{u}^1\right)A\left(\mathbf{u}^0\right)\mathbf{x}_0 + CA\left(\mathbf{u}^1\right)B\left(\mathbf{u}^0\right)\mathbf{u}^0 + CB\left(\mathbf{u}^1\right)\mathbf{u}^1 + D\mathbf{u}^2 \\
&\ldots \\
\mathbf{y}_k &= C\prod_{j=0,\ldots,k-1}^{(L)} A\left(\mathbf{u}^j\right)\mathbf{x}_0 + C\sum_{i=1}^{k-2}\left(\prod_{j=i,\ldots,k-1}^{(L)} A\left(\mathbf{u}^j\right)\right)B\left(\mathbf{u}_{i-1}\right)\mathbf{u}_{i-1} + CB\left(\mathbf{u}^{k-1}\right)\mathbf{u}^{k-1} + D\mathbf{u}^k
\end{aligned}
\tag{4.10}
$$

with the notation of matrix left product (distinguished by the superscript $\prod^{(L)}$ to account for the non-commutativity of matrix multiplication) defined as $\prod_{j=1,\ldots,m}^{(L)} S_j := S_m S_{m-1} \ldots S_1$. We point out, at this stage, the similarity in the input output representation of the bilinear system with that of a linear discrete time varying system [5, 7](ignoring the implicit dependencies of the system matrices on input sequences). However, the number of unknown parameters (A_c, N_{ci}, B_c, C, D) is limited. This fact is used in the identification algorithm to provide a least squares solution for the unknown parameters using known, judicious nonlinear transformations. This recasting procedure for nonlinear problems in to a linear least squares structure is not unlike the much simpler suggestions outlined in Chapter I (section 1.5, pp. 34 - 36) of Crassidis and Junkins[4].

Furthermore, as pointed out in Juang[23], we note that for all the time instances when inputs are both zero, one obtains the linear part of the bilinear system, that is to say,

$$A\left(\mathbf{u}^k = 0\right) = e^{A_c\left(t_{k+1}-t_k\right)} = A_0$$
$$B\left(\mathbf{u}^k = 0\right) = \left[\int_{t_k}^{t_{k+1}} e^{A_c\left(t_{k+1}-\tau\right)} d\tau\right] B_c = B_0 \tag{4.11}$$

$\forall k$ such that $\mathbf{u}^k = 0$. Using these identities, we can give an expression for the zero input response from an initial state $\mathbf{x}_0$ of the bilinear system as,

$$\begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_k \end{bmatrix} = \begin{bmatrix} C \\ CA_0 \\ CA_0^2 \\ \vdots \\ CA_0^k \end{bmatrix} \mathbf{x}_0 = O_L \mathbf{x}_0 \tag{4.12}$$

where the matrix A_0 denotes the linear component of the system matrix for the bilinear system (defined in equation (4.11)). The observability grammian O_L is subscripted to emphasize this fact that it is (the observability grammian) associated with the linear part of the bilinear model.

Analytical expressions for the zero state response on the other hand are not much different from the general nonlinear response given by the input output relation (4.10), with $\mathbf{x}_0 = 0$. However, when a sequence of applied inputs is followed by free decay ($\mathbf{u}^k = 0, \forall k > p$ for some $p > 0$) the ensuing response has a certain structure that turns out to be quite useful for identification purposes as will be shown in the next section. Let us now examine the structure of the zero state response with different types of forcing followed by free decay.

First consider the case when

$$\text{Type 1}: \quad {}^{[1]}\mathbf{u}^0 = {}^{[1]}\mathbf{v}^0, \quad {}^{[1]}\mathbf{u}^k = 0, \qquad \forall k = 1, 2, \ldots \tag{4.13}$$

In this situation the output sequence (left superscript ${}^{[j]}(.)^0$ notation has been employed, i.e., ${}^{[j]}\mathbf{y}_k$ to indicate the particular type of experiment i.e., type $j,$ involved) can be written as,

$$\begin{aligned}
{}^{[1]}\mathbf{y}_0 &= D\,{}^{[1]}\mathbf{v}^0 \\
{}^{[1]}\mathbf{y}_1^0 &= CB\left({}^{[1]}\mathbf{v}^0\right){}^{[1]}\mathbf{v}^0 \\
{}^{[1]}\mathbf{y}_2^0 &= CA_0B\left({}^{[1]}\mathbf{v}^0\right){}^{[1]}\mathbf{v}^0 \\
&\ldots \\
{}^{[1]}\mathbf{y}_k^0 &= CA_0^{k-1}B\left({}^{[1]}\mathbf{v}^0\right){}^{[1]}\mathbf{v}^0
\end{aligned} \tag{4.14}$$

Let us now consider the response structure of type 2, i.e., the applied inputs take the mathematical form given by,

$$\text{Type 2}: \quad {}^{[2]}\mathbf{u}^0 = {}^{[2]}\mathbf{v}^0, \quad {}^{[2]}\mathbf{u}^1 = {}^{[2]}\mathbf{v}^1 \quad , \ {}^{[2]}\mathbf{u}^k = 0, \quad k = 2, 3, \ldots \tag{4.15}$$

In the case of such inputs, the response of the system has the following structure,

$$\begin{aligned}
{}^{[2]}\mathbf{y}_0 &= D\,{}^{[2]}\mathbf{v}^0 \\
{}^{[2]}\mathbf{y}_1 &= CB\left({}^{[2]}\mathbf{v}^0\right){}^{[2]}\mathbf{v}^0 + D\,{}^{[2]}\mathbf{v}^1 \\
{}^{[2]}\mathbf{y}_2 &= CA\left({}^{[2]}\mathbf{v}^1\right)B\left({}^{[2]}\mathbf{v}^0\right){}^{[2]}\mathbf{v}^0 + CB\left({}^{[2]}\mathbf{v}^1\right){}^{[2]}\mathbf{v}^1 \\
{}^{[2]}\mathbf{y}_3 &= CA_0A\left({}^{[2]}\mathbf{v}^1\right)B\left({}^{[2]}\mathbf{v}^0\right){}^{[2]}\mathbf{v}^0 + CA_0B\left({}^{[2]}\mathbf{v}^1\right){}^{[2]}\mathbf{v}^1 \\
&\ldots \\
{}^{[2]}\mathbf{y}_k &= CA_0^{k-2}A\left({}^{[2]}\mathbf{v}^1\right)B\left({}^{[2]}\mathbf{v}^0\right){}^{[2]}\mathbf{v}^0 + CA_0^{k-2}B\left({}^{[2]}\mathbf{v}^1\right){}^{[2]}\mathbf{v}^1
\end{aligned} \tag{4.16}$$

Similar expressions can be derived in general for the response of the bilinear system to control inputs for fixed number of times steps followed by free decay. However, in the identification problem, it will be shown that the response characteristics for the two types of inputs described above are sufficient for the identification of the time invariant parameters of the bilinear problem.

System Identification Methodology

Having outlined the important details concerning the response characteristics of bilinear system models, we now proceed to describe the system identification method. The methodology presented here-in relies on the central assumption that repeated experiments can be performed on the system and each experiment can be started from zero initial conditions. The requirement of repeated experiments is not new and plays a central role in the identification of time varying systems (sometimes called the ensemble data methods of system identification)[5, 7]. Simply put, owing to the coupling of the parameters to be identified from input – output (I/O) data, the number of parameters happens to be larger than the number of equations they satisfy instantaneously. In the continuous time bilinear system identification this growth in number of parameters is manifested in equation (4.10) making it difficult to obtain unique solutions for the unknown parameters. It is of consequence to note that one of the salient features of the original algorithm by Juang was to perform multiple experiments involving pulse sequences. This chapter relaxes the constraints in the pulse inputs applied to the system those results and hence hinges on the same assumptions. In fact, it was shown by Sontag et. al.,[30] that multiple responses are indeed necessary, as they rigorously answer the question of identifiability of a continuous time bilinear system. Repeated experiments are often possible when the aim of system identification is to achieve reduced order modeling from high fidelity multi-physics simulations. There may be other situations when repeated experiments may not be performed. The applicability of the current method therefore varies on a case by case basis involving the resourcefulness of the analyst. The second requirement of the absence of initial condition response in the output data is being made for clarity of presentation. In the presence of initial condition response, the problem of identification becomes more involved and the procedure presented here-in cannot be applied with simple modifications (in a general situation). We postpone the discussion of an identification solution that considers the initial condition response, as a topic of current research to be reported separately.

Repeated Experiments for Identification

In order to supply the data matrices required in the identification process, we perform multiple experiments consisting of input sequences of the special nature (Types 1 and 2 in equations (4.13) and (4.15)). Two types are involved, as was defined in the previous section. The first type of multiple experiments involve arbitrary inputs (zero order hold type constant inputs) for the first time step (t_0) followed by free decay (input sequences of type 1, i.e., equation(4.13)). Given the response vectors from experiments a sequence of subsets of the response sequences can be collected in compact matrix equations given by,

$$\begin{aligned}{}^{[1]}\mathbf{Y}_0^{N_1} &= \begin{bmatrix} {}^{[1]}\mathbf{y}_0^{(1)} & {}^{[1]}\mathbf{y}_0^{(2)} & \cdots & {}^{[1]}\mathbf{y}_0^{(N_1)} \end{bmatrix} \\ &= D\begin{bmatrix} {}^{[1]}\mathbf{v}^{0,(1)} & {}^{[1]}\mathbf{v}^{0,(2)} & \cdots & {}^{[1]}\mathbf{v}^{0,(N_1)} \end{bmatrix} \\ &= D\,{}^{[1]}\mathbf{V}_0^{N_1}\end{aligned} \tag{4.17}$$

and

$$\begin{aligned}{}^{[1]}\mathbf{Y}_{1,\ldots,k}^{N_1} &= \begin{bmatrix} {}^{[1]}\mathbf{y}_1^{(1)} & {}^{[1]}\mathbf{y}_1^{(2)} & \cdots & {}^{[1]}\mathbf{y}_1^{(N_1)} \\ {}^{[1]}\mathbf{y}_2^{(1)} & {}^{[1]}\mathbf{y}_2^{(2)} & \cdots & {}^{[1]}\mathbf{y}_2^{(N_1)} \\ \vdots & \vdots & \ddots & \vdots \\ {}^{[1]}\mathbf{y}_k^{(1)} & {}^{[1]}\mathbf{y}_k^{(2)} & \cdots & {}^{[1]}\mathbf{y}_k^{(N_1)} \end{bmatrix} \\ &= \begin{bmatrix} C \\ CA_0 \\ \vdots \\ CA_0^{k-1} \end{bmatrix} \begin{bmatrix} B\left({}^{[1]}\mathbf{v}^{0,(1)}\right){}^{[1]}\mathbf{v}^{0,(1)} & B\left({}^{[1]}\mathbf{v}^{0,(2)}\right){}^{[1]}\mathbf{v}^{0,(2)} & \cdots & B\left({}^{[1]}\mathbf{v}^{0,(N_1)}\right){}^{[1]}\mathbf{v}^{0,(N_1)} \end{bmatrix} \\ &= \begin{bmatrix} C \\ CA_0 \\ \vdots \\ CA_0^{k-1} \end{bmatrix} \begin{bmatrix} {}^{[1]}\mathbf{b}^{0,(1)} & {}^{[1]}\mathbf{b}^{0,(2)} & \cdots & {}^{[1]}\mathbf{b}^{0,(N_1)} \end{bmatrix} \\ &= O_L\,{}^{[1]}R_B^{0,N_1}\end{aligned} \tag{4.18}$$

where the linear observability grammian O_L and the bilinear, input dependent controllability grammian ${}^{[1]}R_B^{0,N_1}(0)$ have been defined by the decomposition of the response sequence collection. Number in the left superscript of ${}^{[1]}R_B^{0,N_1}$, $[1]$(enclosed in the square brackets) indicates the type of inputs applied while 0 and N_i in the right superscript indicate the time step involved (in the inputs applied forming the bilinear controllability grammian ${}^{[1]}\mathbf{v}^{0,(j)}$) and the number of repeated experiments.

Similarly, the responses generated from the inputs of second type, represented by the equation (4.16) can be collected in the matrix equations written as,

$$\begin{aligned}
{}^{[2]}\mathbf{Y}_0^{N_2} &= \begin{bmatrix} {}^{[2]}\mathbf{y}_0^{(1)} & {}^{[2]}\mathbf{y}_0^{(2)} & \cdots & {}^{[2]}\mathbf{y}_0^{(N_2)} \end{bmatrix} \\
&= D\begin{bmatrix} {}^{[2]}\mathbf{v}^{0,(1)} & {}^{[2]}\mathbf{v}^{0,(2)} & \cdots & {}^{[2]}\mathbf{v}^{0,(N_2)} \end{bmatrix} \\
&= D\,{}^{[2]}\mathbf{V}_0^{N_2}
\end{aligned} \tag{4.19}$$

$$\begin{aligned}
{}^{[2]}\mathbf{Y}_{2,\ldots,k}^{N_2} &= \begin{bmatrix} {}^{[2]}\mathbf{y}_2^{(1)} & {}^{[2]}\mathbf{y}_2^{(2)} & \cdots & {}^{[2]}\mathbf{y}_2^{(N_2)} \\ {}^{[2]}\mathbf{y}_3^{(1)} & {}^{[2]}\mathbf{y}_3^{(2)} & \cdots & {}^{[2]}\mathbf{y}_3^{(N_2)} \\ \vdots & \vdots & \ddots & \vdots \\ {}^{[2]}\mathbf{y}_k^{(1)} & {}^{[2]}\mathbf{y}_k^{(2)} & \cdots & {}^{[2]}\mathbf{y}_k^{(N_2)} \end{bmatrix} \\
&= O_L\begin{bmatrix} A\left({}^{[2]}\mathbf{v}^{1,(1)}\right){}^{[2]}\mathbf{b}^{0,(1)} & A\left({}^{[2]}\mathbf{v}^{1,(2)}\right){}^{[2]}\mathbf{b}^{0,(2)} & \cdots & A\left({}^{[2]}\mathbf{v}^{1,(N_2)}\right){}^{[2]}\mathbf{b}^{0,(N_2)} \end{bmatrix} \\
&\qquad + O_L\begin{bmatrix} {}^{[2]}\mathbf{b}^{1,(1)} & {}^{[2]}\mathbf{b}^{1,(2)} & \cdots & {}^{[2]}\mathbf{b}^{1,(N_2)} \end{bmatrix}
\end{aligned} \tag{4.20}$$

together with the equation involving the output vectors at the second time step given by

$$\begin{aligned}
{}^{[2]}\mathbf{Y}_1^{N_2} &= \begin{bmatrix} {}^{[2]}\mathbf{y}_1^{(1)} & {}^{[2]}\mathbf{y}_1^{(2)} & \cdots & {}^{[2]}\mathbf{y}_1^{(N_2)} \end{bmatrix} \\
&= C\begin{bmatrix} {}^{[2]}\mathbf{b}^{0,(1)} & {}^{[2]}\mathbf{b}^{0,(2)} & \cdots & {}^{[2]}\mathbf{b}^{0,(N_2)} \end{bmatrix} + D\begin{bmatrix} {}^{[2]}\mathbf{v}^{1,(1)} & {}^{[2]}\mathbf{v}^{1,(2)} & \cdots & {}^{[2]}\mathbf{v}^{1,(N_2)} \end{bmatrix} \\
&= C\,{}^{[2]}R_B^{0,N_2} + D\,{}^{[2]}\mathbf{V}_1^{N_2}
\end{aligned} \tag{4.21}$$

Depending on the choice of inputs involved in the experiments, any subset of equations (4.17) through (4.21) is employed by the identification method. It should be observed that the input set ${}^{[2]}\mathbf{v}^{0,(j)}$ involved in determination of response of the second set of experiments, i.e. ${}^{[2]}\mathbf{Y}_0^{N_2}$ / ${}^{[2]}\mathbf{Y}_1^{N_2}$ / ${}^{[2]}\mathbf{Y}_{2,\ldots,k}^{N_2}$, are in general different from the inputs employed in the first set of experiments performed to compute the response sequences, ${}^{[1]}\mathbf{Y}_0^{N_1}$ / ${}^{[1]}\mathbf{Y}_{1,\ldots,k}^{N_1}$. The use of N_1 / N_2 has been employed to emphasize the fact that the number of experiments of each type needed for identification can in general be different. The nature of the applied input sequence along with the corresponding response is shown in Figure 24, which also serves the purpose of clarifying the notation employed in the developments of this chapter.

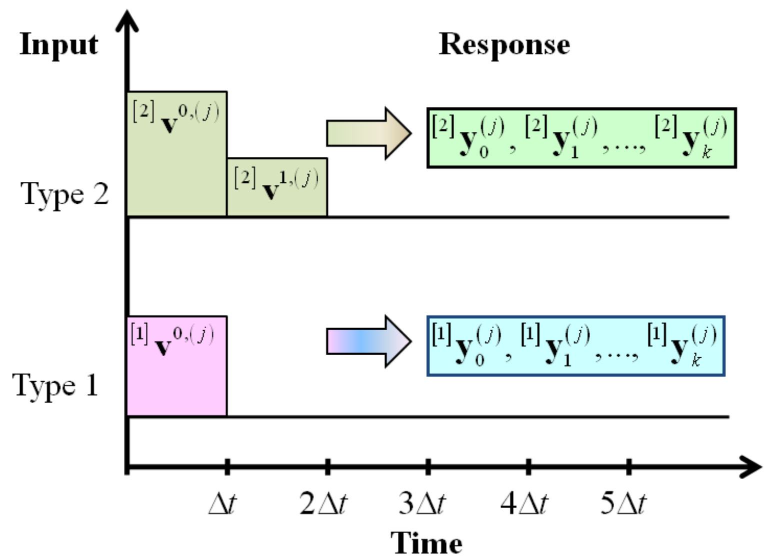

Figure 24. Types of Inputs for Continuous Time Bilinear System Identification

It will be clear in the subsequent developments of this chapter that more experiments from the first type are required in general. The situation where we use the same set of inputs for the second type of experiments (retaining the control forces to same values for two successive time periods), is similar to the case when unit impulse is applied for two successive time steps. We therefore note that the present algorithm allows a modest generalization in the allowable inputs in the sense that at the second time period, arbitrary inputs can be applied while performing the second type of experiments and one can still obtain the continuous time bilinear system parameters.

We now use the above assembled input output sets to compute the bilinear system parameters.

STEP 1: Identification of the Direct Transmission Matrix

The first step involves identification of the direct transmission matrix D. Equations (4.19) or (4.17) could be used for this simple computation. The formal estimate of the direct transmission matrix is given by

$$\begin{aligned}\hat{D} &= {}^{[2]}\mathbf{Y}_0^{N_2}\left({}^{[2]}\mathbf{V}_0^{N_2}\right)^{\dagger} \\ &= {}^{[1]}\mathbf{Y}_0^{N_1}\left({}^{[1]}\mathbf{V}_0^{N_1}\right)^{\dagger}\end{aligned} \tag{4.22}$$

where $(.)^{\dagger}$ denotes the Moore-Penrose pseudo generalized inverse of a matrix[21, 22].

STEP 2: Determination of C and A_0

To calculate the measurement sensitivity matrix C and the linear part of the bilinear system matrix A_0 we start by considering the response sets corresponding to the experiments of the first type (equation (4.18)). Singular value decomposition of the output sets given by,

$$\begin{aligned} {}^{[1]}\mathbf{Y}_{1,\ldots,k}^{N_1} &= U_1 \Sigma_1 V_1^T \\ &= \left(U_1 \Sigma_1^{\frac{1}{2}} \right) \left(\Sigma_1^{\frac{1}{2}} V_1^T \right) \\ &= \hat{O}_L \, {}^{[1]}\hat{R}_B^{0,N_1} \end{aligned} \tag{4.23}$$

Clearly this decomposition is not unique and the corresponding observability and controllability grammians will also reflect this arbitrariness in their corresponding coordinate systems. For example, ${}^{[1]}\mathbf{Y}_{1,\ldots,k}^{N_1} = \left(\hat{O}_L Q \right) \left(Q^{-1} \, {}^{[1]}\hat{R}_B^{0,N_1} \right)$, for any nonsingular matrix Q would be an equivalent and valid set of decompositions. This step is identical to the Hankel matrix formulation and decomposition step in the developments of Juang[23]. Note that the order of the system need not be known and can be given by the number of nonzero singular values of the output collection ${}^{[1]}\mathbf{Y}_{1,\ldots,k}^{N_1}$ in the decomposition of equation (4.23). Due to this non-uniqueness in the decompositions, one can in general realize only similar systems of bilinear models and without more conditions we are not able to obtain the state space realization in the physical coordinates. Therefore the estimated linear controllability grammian is related to the true unknown linear observability grammian by the relation (as shown in Juang[23])

$$\hat{O}_L = O_L T \tag{4.24}$$

Following the subsequent developments of the chapter, we will see that no transformations are required in the present algorithm and we automatically obtain realizations in the same coordinate system.

Using these developments, an estimate for the measurement sensitivity matrix can be set as,

$$\hat{C} = \hat{O}_L\left(1:m,:\right) \tag{4.25}$$

where the notation $X\left(1:m,:\right)$ has been used to indicate the first m rows of a matrix. To compute the corresponding estimate for the linear part of the system matrix, $\hat{A}_0$, we formulate yet another collection of output sequences (from multiple experiments) similar to the equation (4.18) but starting from the third time step, given by,

$$\begin{aligned}
{}^{[1]}\mathbf{Y}_{2,\ldots,k+1}^{N_1} &= \begin{bmatrix} {}^{[1]}\mathbf{y}_2^{(1)} & {}^{[1]}\mathbf{y}_2^{(2)} & \cdots & {}^{[1]}\mathbf{y}_2^{(N_1)} \\ {}^{[1]}\mathbf{y}_3^{(1)} & {}^{[1]}\mathbf{y}_3^{(2)} & \cdots & {}^{[1]}\mathbf{y}_3^{(N_1)} \\ \vdots & \vdots & \ddots & \vdots \\ {}^{[1]}\mathbf{y}_{k+1}^{(1)} & {}^{[1]}\mathbf{y}_{k+1}^{(2)} & \cdots & {}^{[1]}\mathbf{y}_{k+1}^{(N_1)} \end{bmatrix} \\
&= \begin{bmatrix} C \\ CA_0 \\ \vdots \\ CA_0^{k-1} \end{bmatrix} A_0 \left[B\left({}^{[1]}\mathbf{v}^{0,(1)}\right){}^{[1]}\mathbf{v}^{0,(1)} \quad B\left({}^{[1]}\mathbf{v}^{0,(2)}\right){}^{[1]}\mathbf{v}^{0,(2)} \quad \cdots \quad B\left({}^{[1]}\mathbf{v}^{0,(N_1)}\right){}^{[1]}\mathbf{v}^{0,(N_1)} \right] \\
&= \begin{bmatrix} C \\ CA_0 \\ \vdots \\ CA_0^{k-1} \end{bmatrix} A_0 \left[{}^{[1]}\mathbf{b}^{0,(1)} \quad {}^{[1]}\mathbf{b}^{0,(2)} \quad \cdots \quad {}^{[1]}\mathbf{b}^{0,(N_1)} \right] \\
&= O_L A_0 \, {}^{[1]}R_B^{0,N_1}
\end{aligned} \tag{4.26}$$

Using the above relation (equation (4.26)) in conjunction with the estimates calculated by the decomposition presented in equation (4.23), we arrive at the estimate for the linear part of the system matrix given as,

$$\hat{A}_0 = \left(\hat{O}_L\right)^{\dagger}\left({}^{[1]}\mathbf{Y}_{2,\ldots,k+1}^{N_1}\right)\left({}^{[1]}R_B^{0,N_1}\right)^{\dagger} \tag{4.27}$$

where the estimates $\hat{O}_L$, ${}^{[1]}\hat{R}_B^{0,N_1}$ are obtained from the decomposition step developed in the equation (4.23).

An estimate for the continuous time system matrix is then obtained by taking the matrix logarithm of $\hat{A}_0$, similar to the developments of Juang[23], given by,

$$\hat{A}_c = \frac{1}{\Delta t}\log\left(\hat{A}_0\right) \tag{4.28}$$

where Δt corresponds to the step size in which the responses are assumed to be measured while populating the matrices (4.17) through (4.21). We point out that this conversion from discrete time system matrix to continuous time system matrix is ambiguous and non-unique. Sontag et. al.,[30] clarify the ambiguity of the nature of this matrix logarithm in their recent paper. From an engineering and computational stand point, standard subroutines existing in state of the art numerical software, (eg., MATLAB uses the Schur algorithm in[21]) detect the degenerate cases and help the users when the solution may not be unique. A representative situation is the case of system matrices having repeated eigenvalues when the logarithm function may not evaluate accurately owing to the limitation of the subroutines. We do not discuss the computational aspects associated with this matrix logarithm further here, and proceed with subsequent details of the algorithm assuming that the logarithm function in question can be evaluated with sufficient accuracy. Furthermore, if non physical realizations of systems are identified (e.g., complex descriptions for a real system), the analyst is amply warned by a modern numerical implementation indicative of some identifiability/observability issue.

Having determined the estimates for C matrix and the linear component of the system matrix A_c, we now proceed to the computation of estimates for the bilinear system matrices N_{ci}.

STEP 3: Identification of the Bilinear System Matrices

To identify the bilinear system matrices, consider the response sets generated by using the inputs of type 2, as provided by the equation (4.20), repeated here for convenience.

$$\begin{aligned}{}^{[2]}\mathbf{Y}_{2,\ldots,k}^{N_2} = O_L & \left[A\left({}^{[2]}\mathbf{v}^{1,(1)}\right){}^{[2]}\mathbf{b}^{0,(1)} \quad A\left({}^{[2]}\mathbf{v}^{1,(2)}\right){}^{[2]}\mathbf{b}^{0,(2)} \quad \cdots \quad A\left({}^{[2]}\mathbf{v}^{1,(N_2)}\right){}^{[2]}\mathbf{b}^{0,(N_2)} \right] \\ & + O_L\left[{}^{[2]}\mathbf{b}^{1,(1)} \quad {}^{[2]}\mathbf{b}^{1,(2)} \quad \cdots \quad {}^{[2]}\mathbf{b}^{1,(N_2)} \right]\end{aligned} \tag{4.29}$$

Before we proceed further, we need to elaborate more on the second term in the right hand side of the equation (4.29) above, given as,

$$P_2 = O_L\left[{}^{[2]}\mathbf{b}^{1,(1)} \quad {}^{[2]}\mathbf{b}^{1,(2)} \quad \cdots \quad {}^{[2]}\mathbf{b}^{1,(N_2)} \right] \tag{4.30}$$

Recall from our notation beginning equation (4.18) that each column vector, ${}^{[2]}\mathbf{b}^{1,(j)}$ is defined as,

$${}^{[2]}\mathbf{b}^{1,(j)} = B\left({}^{[2]}\mathbf{v}^{1,(j)}\right){}^{[2]}\mathbf{v}^{1,(j)} \tag{4.31}$$

which is no different from ${}^{[2]}\mathbf{b}^{0,(j)} = B\left({}^{[2]}\mathbf{v}^{0,(j)}\right){}^{[2]}\mathbf{v}^{0,(j)}$ for problems with equal first two time step lengths (i.e., $t_1 - t_0 = t_2 - t_1 = \Delta t$). This is because the coefficient matrix $B\left({}^{[1]/[2]}\mathbf{v}^{0/1,(j)}\right)$ is defined by the discretization (equation (4.4)) as,

$$B\left({}^{[1]/[2]}\mathbf{v}^{k,(j)}\right) = \left[\int_{t_k}^{t_{k+1}} e^{\left[A_c + \sum_{i=1}^{r} N_{ci}\, {}^{[1]/[2]}v_i^{k(j)}\right]\left(t_{k+1}-\tau\right)} d\tau\right] B_c \tag{4.32}$$

$\forall k = 0,1$, is time invariant (not an explicit function of time). Therefore the second term of equation (4.29) defined as P_2 in equation (4.30), can now be recognized as the response of inputs from experiments of the first type (same family as that of equations (4.17) and (4.18)). Therefore one may perform an additional set of experiments of the first type or may decide to apply the same set of inputs in the repeated experiments of the second type, in the second time step, namely for the input ensemble ${}^{[2]}\mathbf{V}_1^{N_2}\left(= {}^{[1]}\mathbf{V}_0^{\left(1:N_2 \text{ of total } N_1 \text{ experiments}\right)}\right)$. The analyst may choose to plan the experiments of the (first and second type) depending on his/her convenience with the only constraint that a response of the first type be available for each input (at both time steps) of the second type. We will see shortly that the minimum number of experiments of the second type that is required for unique parameter identification is $r+1$ (r being the number of inputs to the plant – typically not an unreasonably large number of combinations).

Therefore, once the set of inputs for the second time step is fixed, (${}^{[2]}\mathbf{V}_1^{N_2}$ matrix is held constant and therefore P_2 can be determined) we vary the possible inputs for the first time step in the experiments of the second type (namely ${}^{[2]}\mathbf{V}_0^{N_2}$) . This ensures that the following set of equations are obtained for known pulse values, ${}^{[2]}\mathbf{v}^{1,(j)} \in \text{Columns}\left({}^{[2]}\mathbf{V}_1^{N_2}\right)$ (equivalently ${}^{[2]}\mathbf{b}^{1,(j)} \in \text{Columns}\left(P_2\right)$).

$$\begin{aligned}{}^{[2]}\mathbf{Y}_{2,\dots,k}^{N_2} &= O_L A\left({}^{[2]}\mathbf{v}^{1,(j)}\right)\left[{}^{[2]}\mathbf{b}^{0,(1)} \quad {}^{[2]}\mathbf{b}^{0,(2)} \quad \cdots \quad {}^{[2]}\mathbf{b}^{0,(N_2)}\right] + O_L\,{}^{[2]}\mathbf{b}^{1,(j)} \\ &= O_L A\left({}^{[2]}\mathbf{v}^{1,(j)}\right){}^{[2]}R_B^{0,N_2} + O_L\,{}^{[2]}\mathbf{b}^{1,(j)}\end{aligned} \tag{4.33}$$

We are consequently led to a least squares estimate for the system matrix associated with each second time step input value (${}^{[2]}\mathbf{v}^{1,(j)}$) given by

$$\hat{A}\left({}^{[2]}\mathbf{v}^{1,(j)}\right) = \hat{O}_L^{\dagger}\left({}^{[2]}\mathbf{Y}_{2,\dots,k}^{N_2} - \hat{O}_L\,{}^{[2]}\mathbf{b}^{1,(j)}\right)\left({}^{[2]}\hat{R}_B^{0,N_1}\right)^{\dagger} \tag{4.34}$$

$\forall j = 1,\dots,N_2$. We recall that the state transition matrix for a given non zero control input ${}^{[2]}\mathbf{v}^{1,(j)}$ to be (from equation (4.4))

$$\hat{A}\left({}^{[2]}\mathbf{v}^{1,(j)}\right) = e^{\left(\hat{A}_c + \sum_{i=1}^{r}\hat{N}_{ci}\,{}^{[2]}v_i^{1,(j)}\right)\Delta t} \tag{4.35}$$

Therefore, taking the matrix logarithm of the estimate for each j and subtracting the already identified continuous time linear system matrix $\hat{A}_c$, we have the following relationship between the inputs, unknown bilinear system matrices and the identified input dependent transition matrices of the type (4.35),

$$\frac{1}{\Delta t}\log\left[\hat{A}\left({}^{[2]}\mathbf{v}^{1,(j)}\right)\right] - \hat{A}_c = \sum_{i=1}^{r}\hat{N}_{ci}\,{}^{[2]}v_i^{1,(j)} \tag{4.36}$$

In a compact matrix notation, the same set of equations can be written as,

$$\widetilde{LA} = \begin{bmatrix} \hat{N}_{c1} & \hat{N}_{c2} & \cdots & \hat{N}_{cr} \end{bmatrix} \begin{bmatrix} \mathbf{I}\,{}^{[2]}v_1^{1,(1)} & \mathbf{I}\,{}^{[2]}v_1^{1,(2)} \ldots \mathbf{I}\,{}^{[2]}v_1^{1,(N_2)} \\ \mathbf{I}\,{}^{[2]}v_2^{1,(1)} & \mathbf{I}\,{}^{[2]}v_2^{1,(2)} \cdots \mathbf{I}\,{}^{[2]}v_2^{1,(N_2)} \\ \vdots & \vdots \quad \ddots \quad \vdots \\ \mathbf{I}\,{}^{[2]}v_r^{1,(1)} & \mathbf{I}\,{}^{[2]}v_r^{1,(2)} \cdots \mathbf{I}\,{}^{[2]}v_r^{1,(N_2)} \end{bmatrix} \tag{4.37}$$

$$= \begin{bmatrix} \hat{N}_{c1} & \hat{N}_{c2} & \cdots & \hat{N}_{cr} \end{bmatrix} \left({}^{[2]}\mathbf{V}_1^{N_2} \otimes \mathbf{I} \right)$$

where $\mathbf{I} \in \mathbb{R}^{n \times n}$ identity matrix, the matrix product $A \otimes B \in \mathbb{R}^{mn \times nr}$, $\forall A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times r}$ being defined in equation (4.43) and the matrix $\widetilde{LA}$ is defined as,

$$\widetilde{LA} :=$$

$$\left[\frac{1}{\Delta t} \log\left[\hat{A}\left({}^{[2]}\mathbf{v}^{1,(1)} \right) \right] - \hat{A}_c \quad \frac{1}{\Delta t} \log\left[\hat{A}\left({}^{[2]}\mathbf{v}^{1,(2)} \right) \right] - \hat{A}_c \quad \cdots \quad \frac{1}{\Delta t} \log\left[{}^{[2]}\hat{A}\left(\mathbf{v}^{1,(N_2)} \right) \right] - \hat{A}_c \right]$$

(4.38)

The least squares estimate for the estimates of the bilinear system matrices can therefore be given by the relation,

$$\begin{bmatrix} \hat{N}_{c1} & \hat{N}_{c2} & \cdots & \hat{N}_{cr} \end{bmatrix} = \widetilde{LA} \left({}^{[2]}\mathbf{V}_1^{N_2} \otimes \mathbf{I} \right)^{\dagger} \tag{4.39}$$

Observe that the estimates obtained from equation (4.39) are automatically in the same coordinate system as the already identified measurement sensitivity matrix $\hat{C}$ and the linear component of the system matrix $\hat{A}_c$. This is mostly because the formulation allows for the use of the same $\hat{O}_L$ at every step along the way. In addition, one can use power series expansions of the steps involving the matrix exponential and the matrix logarithm to find that the similarity transformations are preserved. This is in turn a consequence of the fact that the functions involved are analytic (refer to [21] for details on analytic functions of a matrix and their transformations).

Having computed the system and measurement sensitivity matrices of the bilinear system, we are ready to perform calculations to obtain the control influence matrix B_c (continuous time domain).

STEP 4: Calculation of B_c

The calculation of B_c presumes that the analyst has computed estimates for the bilinear system matrices A_c, N_{ci} using the steps outlined thus far in this chapter. Consider the definition of ${}^{[1]}\mathbf{b}^{0,(j)}$ as reiterated in the equations (4.31) and (4.32). For $j=1,\ldots,N_1$ the response functions (although nonlinear in $\mathbf{v}^{0,(j)}$) can be written as,

$$
\begin{aligned}
&\begin{bmatrix} {}^{[1]}\mathbf{b}^{0,(1)} & {}^{[1]}\mathbf{b}^{0,(2)} & \cdots & {}^{[1]}\mathbf{b}^{0,(N_1)} \end{bmatrix} \\
&= \begin{bmatrix} B\left({}^{[1]}\mathbf{v}^{0,(1)}\right){}^{[1]}\mathbf{v}^{0,(1)} & B\left({}^{[1]}\mathbf{v}^{0,(2)}\right){}^{[1]}\mathbf{v}^{0,(2)} & \cdots & B\left({}^{[1]}\mathbf{v}^{0,(N_1)}\right){}^{[1]}\mathbf{v}^{0,(N_1)} \end{bmatrix} \\
&= \left\{ \begin{matrix} \left[\displaystyle\int_{t_1}^{t_2} e^{\left[A_c+\sum_{i=1}^{r} N_{ci}\,{}^{[1]}v_i^{0,(1)}\right](t_2-\tau)}\,d\tau\right] B_c\,{}^{[1]}\mathbf{v}^{0,(1)} & \left[\displaystyle\int_{t_1}^{t_2} e^{\left[A_c+\sum_{i=1}^{r} N_{ci}\,{}^{[1]}v_i^{0,(2)}\right](t_2-\tau)}\,d\tau\right] B_c\,{}^{[1]}\mathbf{v}^{0,(2)} \\ & \cdots\left[\displaystyle\int_{t_1}^{t_2} e^{\left[A_c+\sum_{i=1}^{r} N_{ci}\,{}^{[1]}v_i^{0,(N_1)}\right](t_2-\tau)}\,d\tau\right] B_c\,{}^{[1]}\mathbf{v}^{0,(N_1)} \end{matrix} \right\}
\end{aligned}
\tag{4.40}
$$

Note that the unknown parameters, B_c in the above equation (4.40) still appear linearly, in spite of the coefficients being nonlinear functions of the vector test inputs ${}^{[1]}\mathbf{v}^{0,(j)}$. Therefore a linear least squares solution is possible, and developed in the following. We note in passing that, the coefficients involving the matrix exponential convolution operations are most easily evaluated using the "Van-Loan" integrals[31, 32]. A brief discussion on the evaluation of the integrals relevant to the linear system in equation (4.40) is included in the appendix for convenient reference. Therefore, setting $\left[\int_{t_1}^{t_2} e^{\left[A_c + \sum_{i=1}^{r} N_{ci}\, {}^{[1]}v_i^{0,(j)}\right](t_2-\tau)} d\tau\right] := K\left({}^{[1]}\mathbf{v}^{0,(j)}\right)$, equation (4.40) for the unknown parameters B_c becomes,

$$\begin{bmatrix} {}^{[1]}\mathbf{b}^{0,(1)} & {}^{[1]}\mathbf{b}^{0,(2)} & \cdots & {}^{[1]}\mathbf{b}^{0,(N_1)} \end{bmatrix}$$
$$= \begin{bmatrix} K\left({}^{[1]}\mathbf{v}^{0,(1)}\right) B_c\, {}^{[1]}\mathbf{v}^{0,(1)} & K\left({}^{[1]}\mathbf{v}^{0,(2)}\right) B_c\, {}^{[1]}\mathbf{v}^{0,(2)} & \cdots & K\left({}^{[1]}\mathbf{v}^{0,(N_1)}\right) B_c\, {}^{[1]}\mathbf{v}^{0,(N_1)} \end{bmatrix} \tag{4.41}$$

To extract the requisite linear system of equations, we use an identity involving Kronecker products[4], given by,

$$Vec(RSZ) = \left(Z^T \otimes R\right) Vec(S) \tag{4.42}$$

where R, S, Z are matrices of dimensions such that their product can be formed and the $Vec(Z)$ operator is used to stack the columns of Z into a high-dimensioned column vector. The matrix product $A \otimes B \in \mathbb{R}^{mn \times nr}$, $\forall A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times r}$ is defined as,

$$A \otimes B = \begin{bmatrix} a_{1,1}B & \cdots & a_{1,n}B \\ \vdots & \ddots & \vdots \\ a_{m,1}B & \cdots & a_{m,n}B \end{bmatrix} \tag{4.43}$$

with $a_{i,j}$ as the (i,j)th element of the matrix A. Therefore the equivalent linear system for unknown B_c is given by,

$$\begin{bmatrix} {}^{[1]}\mathbf{b}^{0,(1)} \\ {}^{[1]}\mathbf{b}^{0,(2)} \\ \vdots \\ {}^{[1]}\mathbf{b}^{0,(N_1)} \end{bmatrix} = \begin{bmatrix} \left({}^{[1]}\mathbf{v}^{0,(1)T} \otimes K\left({}^{[1]}\mathbf{v}^{0,(1)}\right)\right) \\ \left({}^{[1]}\mathbf{v}^{0,(2)T} \otimes K\left({}^{[1]}\mathbf{v}^{0,(2)}\right)\right) \\ \vdots \\ \left({}^{[1]}\mathbf{v}^{0,(N_1)T} \otimes K\left({}^{[1]}\mathbf{v}^{0,(N_1)}\right)\right) \end{bmatrix} B_c = \mathbf{H}B_c \tag{4.44}$$

Therefore the least squares estimate for the continuous time control influence matrix B_c is given by,

$$\hat{B}_c = \mathbf{H}^{\dagger} \begin{bmatrix} {}^{[1]}\mathbf{b}^{0,(1)} \\ {}^{[1]}\mathbf{b}^{0,(2)} \\ \vdots \\ {}^{[1]}\mathbf{b}^{0,(N_1)} \end{bmatrix} \tag{4.45}$$

We now proceed to demonstrate the steps outlined in the algorithm presented using numerical examples. These numerical examples also provide a basis for optimism with regard to the practical utility of these developments.

Numerical Examples

The examples considered in the demonstration have distinct eigenvalues such that the standard subroutines evaluating the matrix logarithm calculate the matrix function in equation (4.28) without any ambiguity. This is an important step in the identification process and hence dictates the accuracy of the estimates of the continuous time bilinear system parameters.

Example 1

We first demonstrate the working of the algorithms on the bilinear system presented in the paper by Bruni et. al.[33], and also in Juang[23]. The truth model is given by the equation,

$$\begin{aligned}\dot{\mathbf{x}} &= A_c\mathbf{x} + N_{c1}\mathbf{x}u_1 + N_{c2}\mathbf{x}u_2 + B_c\mathbf{u} \\ \mathbf{y} &= C\mathbf{x}\end{aligned} \tag{4.46}$$

where,

$$\begin{aligned}&A_c = \begin{bmatrix} -1 & 0 \\ 1 & -2 \end{bmatrix}; N_{c1} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}; N_{c2} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \\ &B_c = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; C = \begin{bmatrix} 0 & 1 \end{bmatrix}\end{aligned} \tag{4.47}$$

Using the procedure outlined in the paper, the identified plant parameters are given by,

$$\begin{aligned}&\hat{A}_c = \begin{bmatrix} -1.8663 & -3.5311 \\ -0.0328 & -1.1337 \end{bmatrix}; \hat{N}_{c1} = \begin{bmatrix} 1.1116 & -2.9484 \\ 0.0421 & -0.1116 \end{bmatrix}; \\ &\hat{N}_{c2} = \begin{bmatrix} 0.1000 & -0.2652 \\ -0.3393 & 0.9000 \end{bmatrix} \\ &\hat{B}_c = \begin{bmatrix} -0.3323 & -3.6946 \\ 1.1278 & -0.1399 \end{bmatrix}; \hat{C} = \begin{bmatrix} -0.2677 & -0.0789 \end{bmatrix}\end{aligned} \tag{4.48}$$

As is the case for linear systems, the realized system matrices are not unique , because the state space description is not unique. However, the input/output mapping should be unique and the linear part of the identified system matrix should have the same eigenvalues as the true system matrix. The errors in the system matrix eigenvalues (between true and identified) are,

$$\begin{aligned}
\left\|\lambda\left(A_c\right)-\lambda\left(\hat{A}_c\right)\right\| &= \begin{bmatrix} -0.6972 \\ 0.4796 \end{bmatrix}\times 10^{-13} \\
\left\|\lambda\left(N_{c1}\right)-\lambda\left(\hat{N}_{c1}\right)\right\| &= \begin{bmatrix} -0.0809 \\ 0.1232 \end{bmatrix}\times 10^{-13} \\
\left\|\lambda\left(N_{c2}\right)-\lambda\left(\hat{N}_{c2}\right)\right\| &= \begin{bmatrix} -0.0178 \\ -0.1443 \end{bmatrix}\times 10^{-13}
\end{aligned} \tag{4.49}$$

The identified system was subject to some test inputs and the response from the true system to the same test inputs was performed. The test inputs applied to the plants are

$$\begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} = \begin{bmatrix} \sin(7t) \\ \cos(10t) \end{bmatrix} \tag{4.50}$$

Output profiles obtained from the true and identified systems are compared in Figure 25. The error in the response to the test functions is plotted in Figure 26.

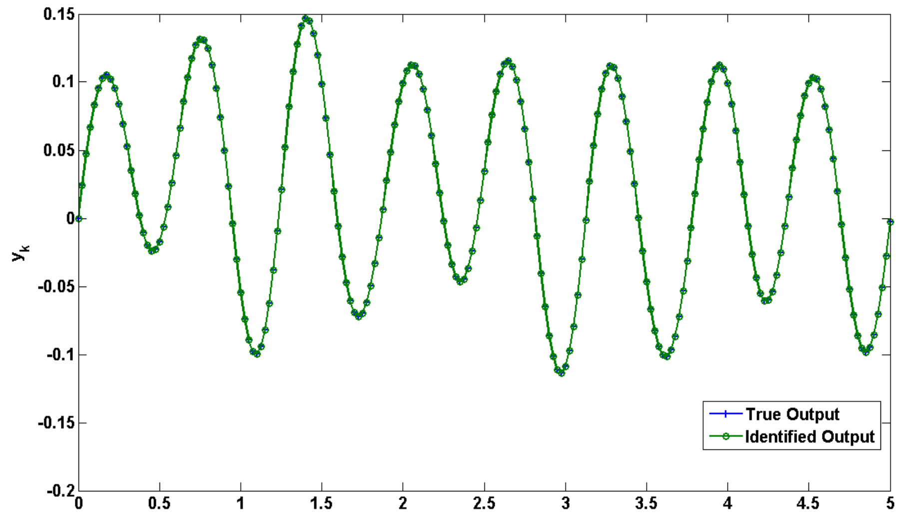

Figure 25. Example 1 (Bruni): Output Comparison (True System vs. Identified System)

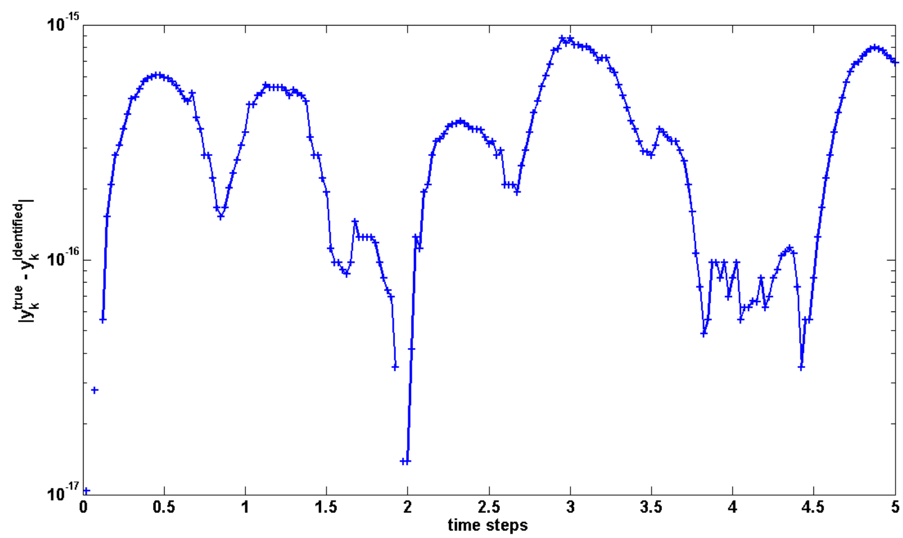

Figure 26. Example 1 (Bruni): Output Error (True System vs. Identified System)

It is evident that the linear system eigenvalues in equation (4.49) and the nonlinear system response in Figure 25 and Figure 26 were very accurately captured in this example.

Example 2

Now we apply the procedure detailed in this chapter to an example where the linear part of the bilinear system matrix is unstable.

$$A_c = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix};$$
$$N_{c1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 2 & 1 \\ 1 & 3 & 4 \end{bmatrix}; N_{c2} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 4 & 2 & 1 \end{bmatrix} \tag{4.51}$$
$$B_c = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}; C = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 2 \end{bmatrix}$$

Using the same procedure we obtain the following estimates for the bilinear plant parameters, as,

$$\hat{A}_c = \begin{bmatrix} 0.0933 & 0.0553 & -0.7048 \\ -0.7767 & -0.3856 & -3.1942 \\ 0.0202 & 0.3606 & 0.5923 \end{bmatrix};$$
$$\hat{N}_{c1} = \begin{bmatrix} 5.8579 & -0.1208 & 5.3429 \\ 1.9337 & 1.5666 & 4.2598 \\ -0.8757 & -0.0147 & -0.4245 \end{bmatrix}; \hat{N}_{c2} = \begin{bmatrix} 2.8457 & -2.1441 & 4.7919 \\ -0.5418 & -0.2191 & 1.6626 \\ -0.0622 & 0.1447 & -1.6267 \end{bmatrix} \tag{4.52}$$
$$\hat{B}_c = \begin{bmatrix} -0.8320 & -1.7599 \\ 1.4318 & -0.8593 \\ -0.0603 & -0.4598 \end{bmatrix}; \hat{C} = \begin{bmatrix} -1.0274 & 0.8101 & 0.2433 \\ -1.9486 & -0.4507 & -0.3987 \end{bmatrix}$$

Errors incurred in the system eigenvalues, between the true and the identified system are given as,

$$\left\|\lambda\left(A_c\right)-\lambda\left(\hat{A}_c\right)\right\|=\begin{bmatrix} \text{-0.2887} \\ \text{0.0167 - 0.5884i} \\ \text{0.0167 + 0.5884i} \end{bmatrix}\times 10^{-14}$$

$$\left\|\lambda\left(N_{c1}\right)-\lambda\left(\hat{N}_{c1}\right)\right\|=\begin{bmatrix} 0.5640 \\ -0.1887 \\ -0.3020 \end{bmatrix}\times 10^{-13} \tag{4.53}$$

$$\left\|\lambda\left(N_{c2}\right)-\lambda\left(\hat{N}_{c2}\right)\right\|=\begin{bmatrix} 0.9770 \\ -0.2554 \\ -0.0444 \end{bmatrix}\times 10^{-13}$$

Response of the identified system and the true system to test functions for the example 2 are compared in the Figure 27. Error between the true and identified system response functions is plotted in Figure 28. As is evident, the eigenvalues of the linear part of the identified system and the full nonlinear response were again captured with high precision in example 2.

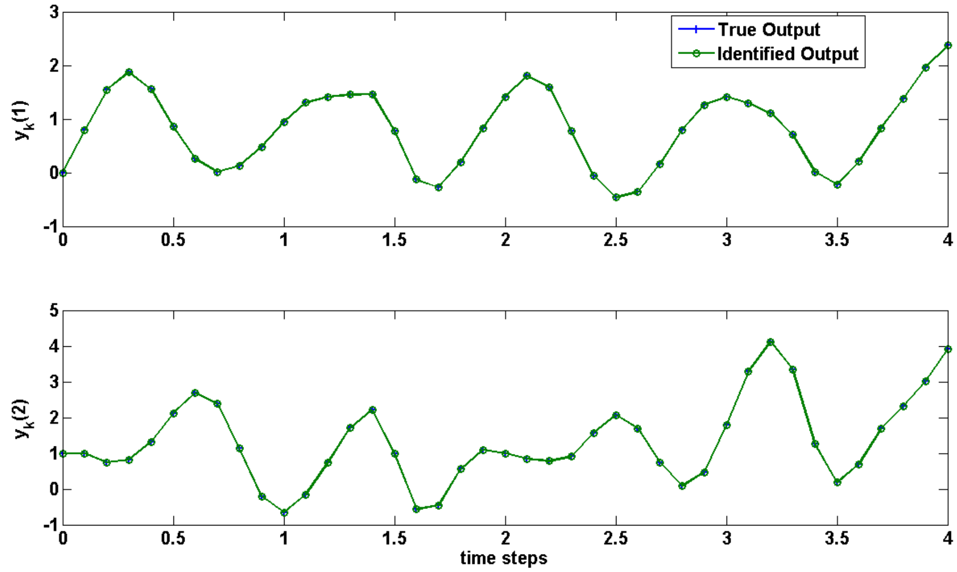

Figure 27. Example 2: Output Comparison (True System vs. Identified System)

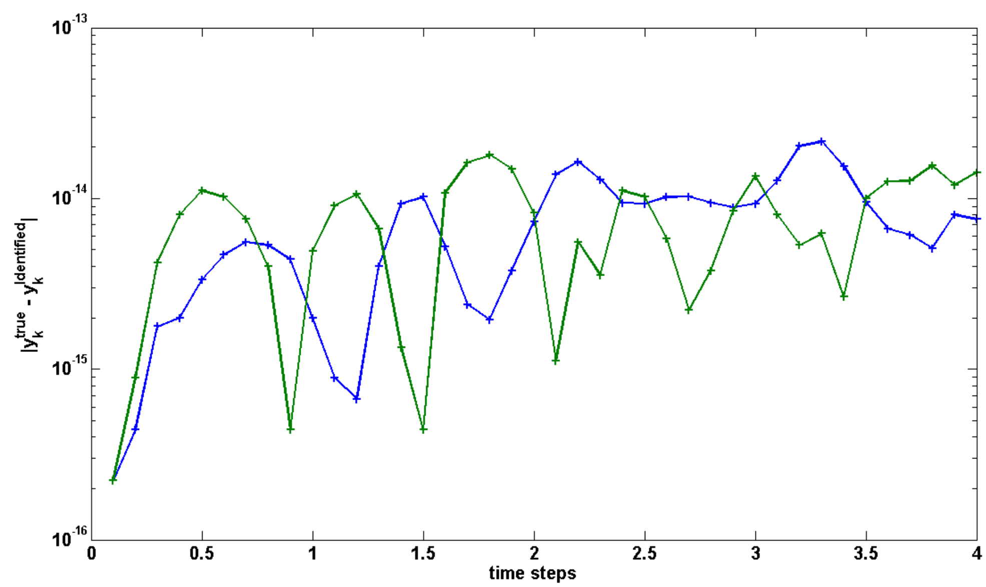

Figure 28. Example 2: Output Error (True System vs. Identified System)

These numerical examples and several others not reported here support the validity of the formulations and provide a basis for optimism with regard to practical importance of these developments.

Conclusions

This chapter introduces fundamental developments that permit the identification of bilinear dynamical systems. These results are believed to be of fundamental significance and represent an important extension of the now classical Ho-Kalman identification methodology (of which the Eigensystem Realization Algorithm is an integral component) that is foundational to linear system identification theory and practice. The structure of the resulting algorithms is attractive for computational and rely only on the collection of a systematic sequence of experimental input/output response measurements.

CHAPTER V

APPLICATION TO PROBLEMS IN GUIDANCE, CONTROL AND DYNAMICS

Introduction

We now apply the techniques developed in previous chapters to practical problems in guidance, control and dynamics. First application considers the guidance of a point mass operating in the presence of significant model error. Perturbation models are identified using the time varying eigensystem realization algorithm (TVERA) presented in the chapter II of this dissertation. Perturbation guidance scheme is presented based on the identified linear departure motion dynamics in the presence of unknown model errors.

Next application involves the dynamics of a point mass in a rotating tube which constitutes a time varying linear system due to the presence of centrifugal force. Time varying model sequence for the system description is obtained along with the corresponding measurement model. The transformation theory developed for linear time varying system identification is clearly explained with this physical example.

Subsequent developments recast the model governing the same system in to a continuous time bilinear system model, with appropriate redefinitions of the variables involved. Results of the continuous time bilinear system identification algorithm developed in the chapter IV of this dissertation are discussed using this example. Dynamics involving an automobile brake mechanism is discussed subsequently as a bilinear control system application problem. Identification results for this problem are also detailed in this chapter.

We now detail the guidance and control example application. This example is rather pedagogical for demonstrative purposes. However, it embodies a typical situation experienced by an analyst in a guidance-type problem set up. It will not be difficult to see that the developments undertaken in this example can be generalized to obtain reduced order perturbation models from a high fidelity multi-physics simulation about a nominal operating point.

Guidance and Control Application Problem

Aerospace engineering control and estimation problems are most often nonlinear in nature. It is central in the actuation and sensing applications, that simplified and accurate models of the dynamics of the aircraft or spacecraft are made available for analysis and control system design. Engineers are also aware of the fact that the accuracy of the models derived or developed is directly reflected in the performance of the control or sensing system being designed. Owing to the increasingly stringent performance requirements on engineers for flight control and sensing systems, the methods of modeling and control cannot be decoupled in modern aerospace engineering research.

One of the subjects of aerospace engineering where this interaction and coupling between modeling and control has profound ramifications is the body of work that goes under the name of guidance, navigation and control. Bryson[34], Battin[19] and others contributed extensively to the problems of guidance, navigation and control and have enabled several complex aircraft, spacecraft orbital and attitude maneuvers documented widely in the historical literature by several researchers[22, 32, 35-38]. The importance of this key handshake in guidance, navigation and control problems has been paraphrased elegantly by Professor Junkins as an important lesson learnt by our community from the Apollo era: "… - theoretical research in dynamics and control methodology and advanced flight implementations not only can comfortably coexist, they belong to the same set" (the von Karman lecture ref. [38]).

An important artifact and tool developed owing to the historical work in guidance and control is the use of perturbation models about a reference trajectory for guidance, navigation, control and analysis. While this local linearization of nonlinear equations of motion of aerospace vehicles about nominal trajectories has become indispensable for controller design and analysis, the central assumption that the "truth" is modeled by the known structure of the nonlinear equations is still restrictive and hence allows the scope for the ever presence of model errors. Inspired from classical developments of aerodynamics, where the theoretical modeling developments often go hand in hand with experimentation, we now propose a methodology to use experimental data to realize the first order, time varying discrete time model of the departure motions of a system from a nominal trajectory.

Problem Statement

A point mass free to move in a planar space with two control inputs is the simple dynamical system chosen for the current demonstration. The mass is being acted upon by two force inputs which can be changed at will by the operator. The application under consideration in the present example requires the calculation of force components required to transfer this point mass from the origin to a position as close as possible to a given target location $\left(x_T, y_T\right)$ in a given amount of time. We do not really care about the components of the velocity at the terminal time. In guidance literature, this type of problem is often called the intercept problem[34]. In appropriate non dimensional units, the equations of motion of the point mass are given by,

$$\begin{aligned} \dot{x}_1 &= x_3 \\ \dot{x}_2 &= x_4 \\ \dot{x}_3 &= u_1 \\ \dot{x}_4 &= u_2 \end{aligned} \tag{5.1}$$

The solution of prescribing the forces (control inputs) to take the point mass sufficiently close to the target location is straightforward. In fact, one can easily obtain this solution in a feedback form formulating the necessary conditions of optimal control theory (or alternatively using Bellman's principle of optimality)[34].

Nominal Solution Generation

Let us assume that the nominal solution to the intercept problem has been designed by the analyst in a feedback form. For clarity of presentation, table 1 presents the continuous time version of the feedback solution to problems with linear system dynamics (with usual assumptions on the notation of the problem and dropping the functional dependencies that are assumed to be understood according to context). It is well known that a feedback solution is preferred over the open loop version owing to the robustness of the feedback solution to plant model uncertainties, plant parametric uncertainties (small in some norm) and its “time-to-go” nature (independence to initial conditions)[34].

Table 1 Summary of Optimal State Feedback Control for the Intercept Application Problem

Dynamical System	$\dot{\mathbf{x}}(t) = A(t)\mathbf{x}(t) + B(t)\mathbf{u}(t)$
Performance Index (tracking problem)	$\min\limits_{\mathbf{u}(t)} J = \frac{1}{2}\boldsymbol{\varepsilon}(t_f)^T Q_f \boldsymbol{\varepsilon}(t_f) + \frac{1}{2}\int_{t_0}^{t_f} \mathbf{u}^T(\tau) R(\tau)\mathbf{u}(\tau) d\tau + + \frac{1}{2}\int_{t_0}^{t_f}\left[\boldsymbol{\varepsilon}(\tau)^T Q(\tau)\boldsymbol{\varepsilon}(\tau)\right] d\tau$ where $\boldsymbol{\varepsilon}(\tau) := \mathbf{x}(\tau) - \mathbf{r}(\tau)$ $\boldsymbol{\varepsilon}(t_f) := \mathbf{x}(t_f) - \mathbf{r}_f$
Optimal Control	$\mathbf{u}^* = -R^{-1}B^T\left(S(t)\mathbf{x}(t) + \mathbf{v}(t)\right)$
Feedback Gain Differential Equation	$\dot{S} = -A^T S - SA + SBR^{-1}B^T S - Q$ with final condition $S(t_f) = Q_f$
Differential Equation for the Feed-Forward term	$\mathbf{v}(t) = -\left[A^T - S(t)BR^{-1}B^T\right]\mathbf{v}(t) + Q(t)\mathbf{r}(t)$ with final condition $\mathbf{v}(t_f) = -Q_f\mathbf{r}_f$

Feedback solution for the intercept problem is therefore obtained from the above table by setting, $\mathbf{r}(t) = \mathbf{0}$, $\mathbf{r}_f^T = \begin{bmatrix} x_T & y_T & 0 & 0 \end{bmatrix}^T$ and choosing a positive semi-definite

Q_f such that the velocity components of the soft constraint in the performance index do not appear. For example,

$$Q_f = \begin{bmatrix} P_f & \mathbf{O}_{2\times2} \\ \mathbf{O}_{2\times2} & \mathbf{O}_{2\times2} \end{bmatrix} \tag{5.2}$$

for some $P_f > 0, P_f \in \mathbb{R}^{2\times2}$ can be chosen.

Considering the example problem in the current discussion, using the usual definition of the state vector to be given as $\mathbf{x}(t) := \begin{bmatrix} x_1(t) & x_2(t) & x_3(t) & x_4(t) \end{bmatrix}^T$, the system matrix and the control influence matrix are written as the constant matrices,

$$\begin{aligned} A &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ B &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned} \tag{5.3}$$

The nominal solution thus obtained for a certain choice of plant parameters,

$$\begin{aligned} & t_f = T = 10\,\text{sec} \\ & x_T = 600, y_T = 600 \\ & Q(t) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; \quad Q_f = \begin{bmatrix} 4 & 1 & 0 & 0 \\ 1 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ & R(t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned} \tag{5.4}$$

is the unperturbed solution plotted in Figure 29.

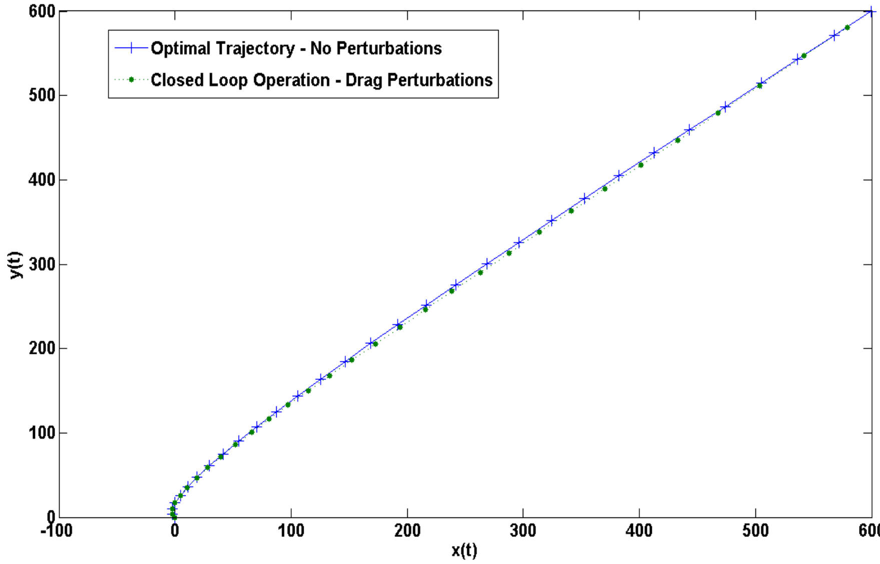

Figure 29. Nominal Trajectory: Reference and Perturbed (x-y Phase Subspace)

Using the state feedback control law (with gains calculated from the differential equations outlined in the table 1),

$$\mathbf{u}^*(t) = -R^{-1}B^T S(t)\mathbf{x}(t) - R^{-1}B^T \mathbf{v}(t) \tag{5.5}$$

The closed loop dynamics of the optimal state trajectory can be written as,

$$\begin{aligned}\dot{\mathbf{x}}^*(t) &= \left[A - BR^{-1}B^T S(t)\right]\mathbf{x}^*(t) - BR^{-1}B^T \mathbf{v}(t) \\ &= A_{CL}(t)\mathbf{x}^*(t) - BR^{-1}B^T \mathbf{v}(t)\end{aligned} \tag{5.6}$$

where $\mathbf{x}^*(t)$ notation is used to denote the optimal trajectory (which by definition means there are no unaccounted perturbations in the closed loop trajectory calculations). With the simulation parameters being considered (in equation (5.4)) the positions achieved by the unperturbed feedback solution are given as $\left(x_T^*, y_T^*\right) = (599.5897,\ 599.7858)$.

We note, in passing, that we could have conveniently redirected the reader to any one of the classical texts on optimal state feedback solution to the tracking problem presented in Table 1. The information presented is to reiterate the fact that the gains obtained from such a finite time problem are in general time varying in nature. Closing the loop with gains calculated in such a fashion makes the closed loop time varying (but still strictly linear plant model) and hence the time varying identification methods developed in this dissertation are of relevance to identify the plant closed loop dynamics while in operation. We will see in the next section that possible perturbations introduce nonlinearities in the form of unmodeled dynamics and the performance is no longer satisfactory and the intercept problem incurs a terminal error which may no longer be within acceptable limits.

Closed Loop Operation: Nominal Solution Operating in the Presence of Unstructured Perturbations (Drag)

As pointed out earlier, the closed loop system for the intercept problem, although designed for operation under the strict conditions that no model uncertainties exist is, in general, forgiving in practice. Let us consider a class of perturbations common in aerospace engineering – the drag perturbations acting on the point mass model. Equations governing the drag model are assumed to be given by

$$\begin{aligned} \dot{x}_1 &= x_3 \\ \dot{x}_2 &= x_4 \\ \dot{x}_3 &= u_1 - \mu V \ x_3 \\ \dot{x}_4 &= u_2 - \mu V \ x_4 \end{aligned} \tag{5.7}$$

where $V=\sqrt{x_3^2+x_4^2}$ denotes the velocity magnitude of the point mass and $\mu=\frac{1}{2}\left(\frac{\rho C_d a}{m}\right)$ is the non-dimensional parameter representing the magnitude of the drag force. The usual notations of ρ being the density of the medium, a denoting the area of the body exposed to the free stream, m being the mass of the body and C_d being the drag coefficient whose value depends on the surface properties of the (bluff) body and the profile shape, have been employed. For small coefficients of drag, in spite of the perturbations, the linear state feedback was able to reach closer to the target.

To exaggerate the effects of the perturbation and subsequently highlight the importance of the identification method, a (perhaps unusually) large value of the non-dimensional drag parameter was chosen ($\mu=2\times10^{-3}$). The performance of the linear state feedback solution was found to be unsatisfactory for this drag parameter value, rendering the final achieved position coordinates to be $\left(x_T^p, y_T^p\right)=\left(578.8761,\ 581.1549\right)$, while in the absence of perturbations, the position coordinates achieved were $\left(x_T^*, y_T^*\right)=\left(599.5897,\ 599.7858\right)$. The deviations from the optimal (conditioned on the unperturbed plant model) path are shown in the Figure 29 (magnified view of the last few time steps is plotted in Figure 30).

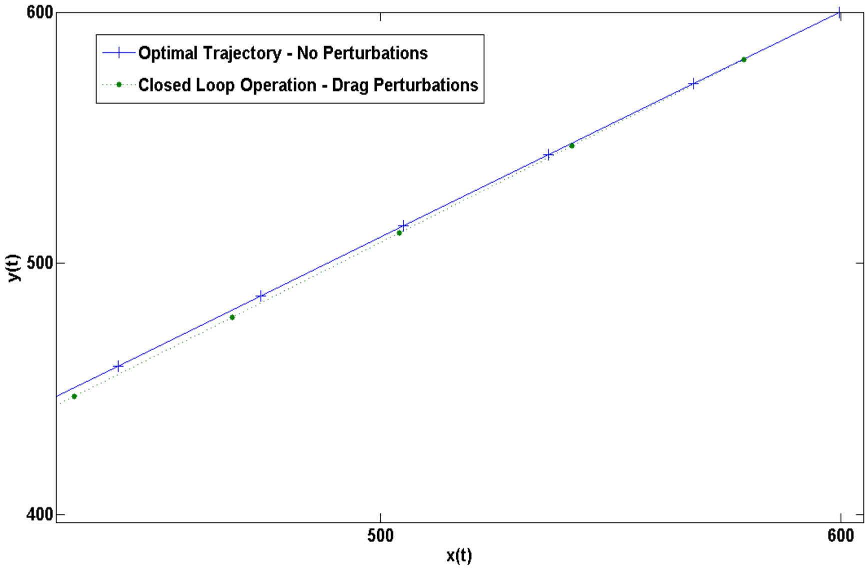

Figure 30. Nominal Trajectory: with and without Perturbations (Zoomed View of x-y Phase Subspace)

The position state profiles of the unperturbed and perturbed solutions are plotted in Figure 31, while Figure 32 details the velocity state unperturbed and perturbed nominal solutions.

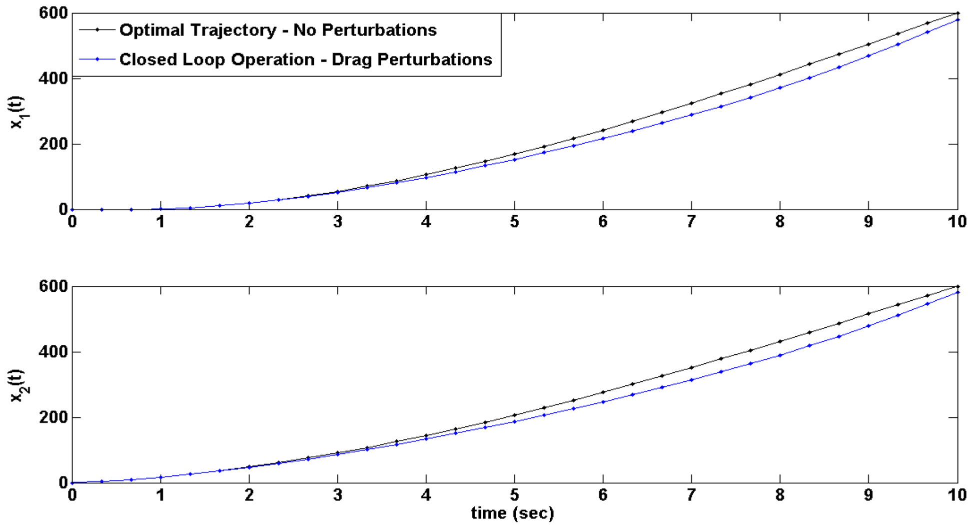

Figure 31. Nominal Trajectory: with and without Drag Perturbations (Position Coordinates)

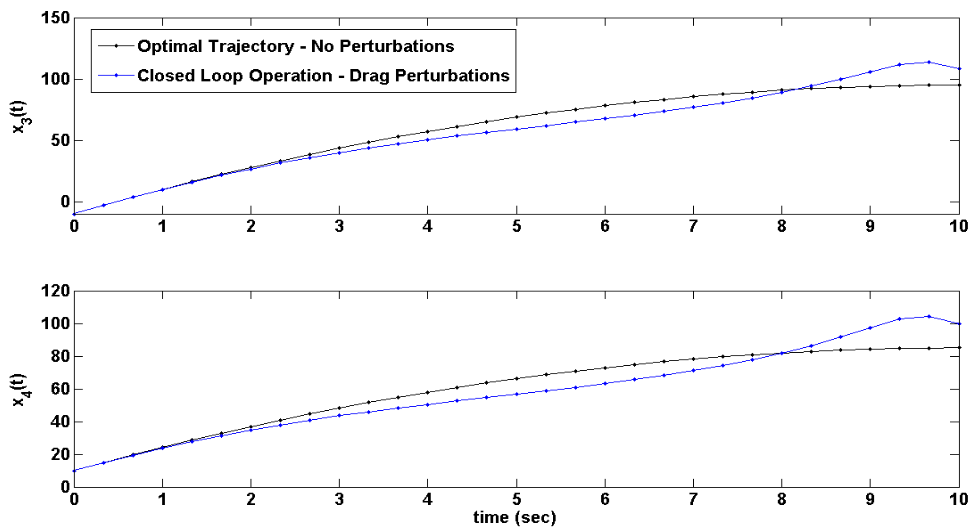

Figure 32. Nominal Trajectory: with and without Drag Perturbations (Velocity Coordinates, x3, x4)

The optimal control profiles applied for these simulations are plotted in Figure 33. Clearly, in the presence of perturbations, the operation of the closed loop system is far from optimal and this fact is reflected in the trajectory profiles.

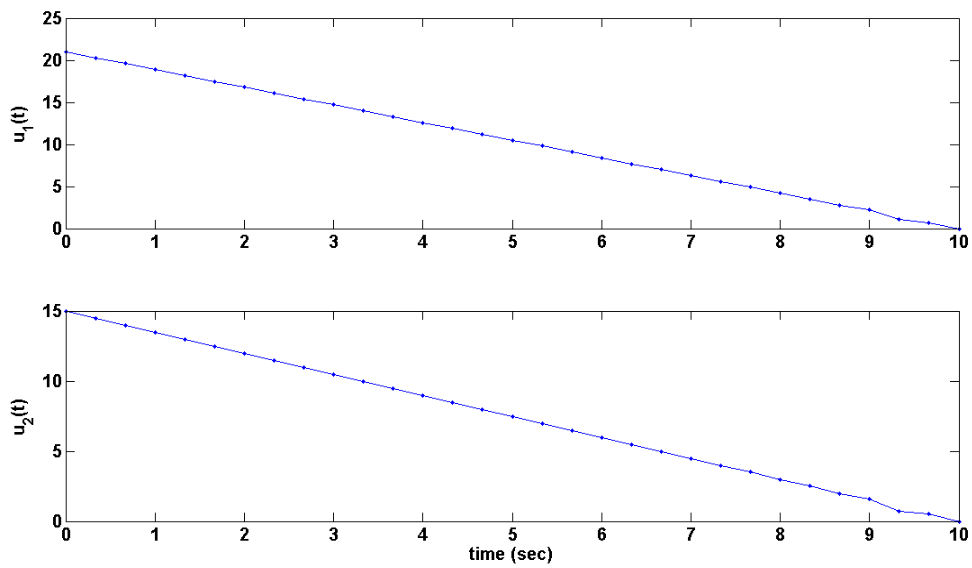

Figure 33. Nominal Solution: Optimal Control Input Profiles (State Feedback Solution)

Application of the Time Varying Eigensystem Realization Algorithm to Identify a Time Varying Linearization Model about the Nominal Solution (perturbed).

Although in the above example, we have explicit knowledge of the structure and model of the perturbations causing the departure motion from the desired trajectory, in practice there is no means of determining a structure let alone a construing a model for the causative perturbations. It can be said that in general, the perturbation models can never be determined uniquely. Hence such perturbations are most often known as unstructured perturbations. Often, these unstructured perturbations are analytic functions of the state variables of the plant model (e.g. aerodynamic forces). Therefore, in the context of the present example, these perturbations can be formally quantified as,

$$\begin{aligned} \dot{\mathbf{x}}^p &= \left(A - BG(t)\right)\mathbf{x}^p + \mathbf{g}\left(\mathbf{x}^p\right) - BR^{-1}B^T\mathbf{v}(t) \\ &= A_{CL}(t)\mathbf{x}^p + \mathbf{g}\left(\mathbf{x}^p\right) \end{aligned} \tag{5.8}$$

with $G(t)$ definition as $G(t) = R^{-1}B^T S(t)$. In the context of the present example, the trajectory $\mathbf{x}^p(t)$ is to be understood as the state variable history obtained by using the state feedback program (Table 1) of the previous section in the realistic system experiencing drag perturbations.

Let us further assume that we can perform experiments about this constructed nominal. That is to say that we could apply control inputs different from the calculated state feedback control law, denoted by $\mathbf{u}(t)$ defined as,

$$\begin{aligned} \mathbf{u}(t) &= \mathbf{u}^* + \delta\mathbf{u}(t) \\ &= -R^{-1}B^T\left(S(t)\mathbf{x}(t) + \mathbf{v}(t)\right) + \delta\mathbf{u}_k \end{aligned} \tag{5.9}$$

$\forall t \in [t_k, t_{k+1})$, similar to a zero order hold approximation[1]. Denoting the trajectory obtained in this fashion by $\mathbf{x}(t)$, the equations of the plant dynamics in this case can be written as,

$$\begin{aligned} \dot{\mathbf{x}} &= \left(A - BG(t)\right)\mathbf{x} + \mathbf{g}(\mathbf{x}) - BR^{-1}B^T\mathbf{v}(t) + B\delta\mathbf{u}_k \\ &= A_{CL}(t)\mathbf{x} + \mathbf{g}(\mathbf{x}) + B\delta\mathbf{u}_k \end{aligned} \tag{5.10}$$

Defining the departure motion state as, $\delta\mathbf{x}(t) := \mathbf{x}(t) - \mathbf{x}^p(t)$ and expanding the disturbance function $\mathbf{g}(\mathbf{x})$ in a Taylor series about the nominal trajectory $\mathbf{x}^p(t)$, we have,

$$\mathbf{g}(\mathbf{x}^p + \delta\mathbf{x}) = \mathbf{g}(\mathbf{x}^p) + \left[\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right]_{\mathbf{x}=\mathbf{x}^p} (\mathbf{x} - \mathbf{x}^p) + \ldots \tag{5.11}$$

Using equations (5.8) and (5.10) together with (5.11), we can write the equations governing the departure motion dynamics, as,

$$\begin{aligned} \delta\dot{\mathbf{x}} &= \left(A_{CL}(t) + \left[\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right]_{\mathbf{x}=\mathbf{x}^p} \right) \delta\mathbf{x} + B\delta\mathbf{u}_k + HOT \\ &\approx \left(A_{CL}(t) + \left[\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\right]_{\mathbf{x}=\mathbf{x}^p} \right) \delta\mathbf{x} + B\delta\mathbf{u}_k \end{aligned} \tag{5.12}$$

with initial conditions $\delta\mathbf{x}(t_0) = \mathbf{x}(t_0) - \mathbf{x}^p(t_0)$ (that can possibly assume a non-zero value). The high order terms are neglected in the equation (5.12) thereby giving us a first order model of the departure motion dynamics. This was done to satisfy the theoretical requirement of the identification algorithm for the true plant dynamics to be linear and time varying.

At this stage, it is quite interesting to emphasize that the method detailed here does not depend on the explicit knowledge of the uncertainty of this nonlinear term"$\mathbf{g}(\mathbf{x})$". Therefore directly from experimental input output data or from the repeated high fidelity multiphysics model simulations (with considered input output data, as we will outline in the subsequent developments of this section), we realize the first order perturbation dynamics model. It is quite fascinating to this author that such a situation is possible. While Bryson (and his co-workers) brought neighboring optimal control solution methods and perturbation guidance schemes to the attention of the practical engineers, this dissertation attempts, for the very first time to develop the models assumed to be derivable from physics realizable from empirical data. Only time can judge the applicability and utility of the developments here-in.

Therefore the equivalent first order perturbation dynamics model to be identified in the discrete time domain takes the usual form,

$$\begin{aligned} \delta\mathbf{x}_{k+1} &= A_k \delta\mathbf{x}_k + B_k \delta\mathbf{u}_k \\ \delta\mathbf{y}_k &= C_k \delta\mathbf{x}_k + D_k \delta\mathbf{u}_k \end{aligned} \tag{5.13}$$

with the assumed definitions that $\delta\mathbf{x}_k := \mathbf{x}(t_k) - \mathbf{x}^p(t_k)$ and the time varying direct transmission terms D_k , have been included for generality.

For simplicity, it will be convenient to assume a number of outputs (sensors) to be equal to the number of states of the system for the subsequent developments of this section. This is certainly not a restrictive assumption, since the estimated state of the dynamic system can be used (assuming that the state estimator has converged effectively) or a technique outlined in the previous chapter can be employed to overcome this restriction. The utility of this assumption is in obtaining the last few time step models. As pointed out in chapter II, we will need to carry out extra set of experiments for the last few time steps as the generalized Hankel matrix populated in such cases will not in general have full row rank. This is because; depending on the particular experimental situation (or a simulation scenario) the generalized Markov parameters

$h_{k_f+1,j}, \forall j = k_f - 1, k_f - 2, \ldots$ may not be available for populating the Hankel matrix $H_{k_f}^{(p,q)}$ (with the assumed definition that k_f is the final time step). This will render the rank of the last step generalized Hankel matrix to be $rank\left(H_{k_f}^{(p,q)}\right) = m < n$. In other words, only the first row block will be only available for use by the TVERA decompositions. This is mathematically an adjoint situation of the first few time step model determination problem discussed earlier. As pointed out in the beginning of this paragraph, the method based on the free response experimental data outlined in the chapter II could be employed to effectively determined the last few time step models, we do not intend to let such details interfere with the main goals of this chapter which is to demonstrate that the methods presented in the dissertation are of importance in practical problems.

Input Output Experimental Data Generation

The key component of obtaining the identified first order perturbation model about the constructed nominal trajectory is the generation of input output experimental data. One of the prime ingredients of this data generation is the determination of the magnitude of the control input deviation sequence $\delta \mathbf{u}_k$. If the input magnitude is too large, the nonlinearity of the problem may reflect in the outputs and the first order deviation model may not be valid anymore. If the input magnitude is too small, then the excitation may not be rich enough to identify all the degrees of freedom of the first order deviation model. For the present example, an input deviation magnitude of

$\left\| \delta \mathbf{u}_k \right\| \approx 2\times10^{-2}$ was chosen by making each element of the input sequence a normal random vector of statistics (mean and covariance), $\delta \mathbf{u}_k \sim N\left(\begin{bmatrix}0\\0\end{bmatrix}, 2\times10^{-2}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\right)$. The Dormand-Prince solver (ode45 subroutine of MATLAB with a relative and absolute tolerance of 10^{-8}) was used to integrate the nonlinear equations of motion to obtain the response of the system with new control input sequence overlaid on to the conventional state feedback controller. For the stiff differential equations, typically this translates to an accuracy of the same order of magnitude of tolerance which usually implies a loss of a digit (or two) of precision in the solutions. A sample of 100 different sets of input output data sets were obtained by choosing the "true" measurement model parameters to be given by,

$$
\begin{aligned}
C_k &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
D_k &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}
\end{aligned}
\tag{5.14}
$$

Determination of the Generalized Markov Parameters (Minimum Number of Experiments)

Time varying OKID algorithm detailed in the presentations of chapter III was applied on the input output data (operating in the time-varying dead beat observer mode) to realize the generalized system Markov parameters required for populating the generalized Hankel matrix sequence.

Generalized Hankel Matrix Singular Value Sequence

An indication as to whether the first order perturbation model is realized (or not) is given by a plot of the singular values of the generalized Hankel matrix sequence. For the present example, the realized generalized singular values of the Hankel matrix can be plotted as shown in Figure 34. It is clear from the Figure 34 that the order of time varying system model realized is the same as the order of the analytic first order perturbation dynamics model (similar to equation (5.12), $n = 4$ in case of the present example). The plot of the generalized Hankel matrix sequence also conveys some other information (from experience with other example situations). If there are nonlinearities detected in the outputs, the Hankel matrix singular value time history reveals changes in order showing the participation of higher order “fake” time varying modes only at certain time instances. It will be clear in such a plot that the linear model identification is not being proper. Appropriate steps of reducing the perturbation input excitation energy can then be taken to obtain much more refined linear perturbation dynamical models. These plots also detect the presence of disturbance inputs to the nominal system. This is manifested by a persistent presence of more degrees of freedom (number of nonzero singular values of the Hankel matrix sequence) than the physical coordinates. Therefore a wealth of additional information can be derived from the methods being developed in this dissertation.

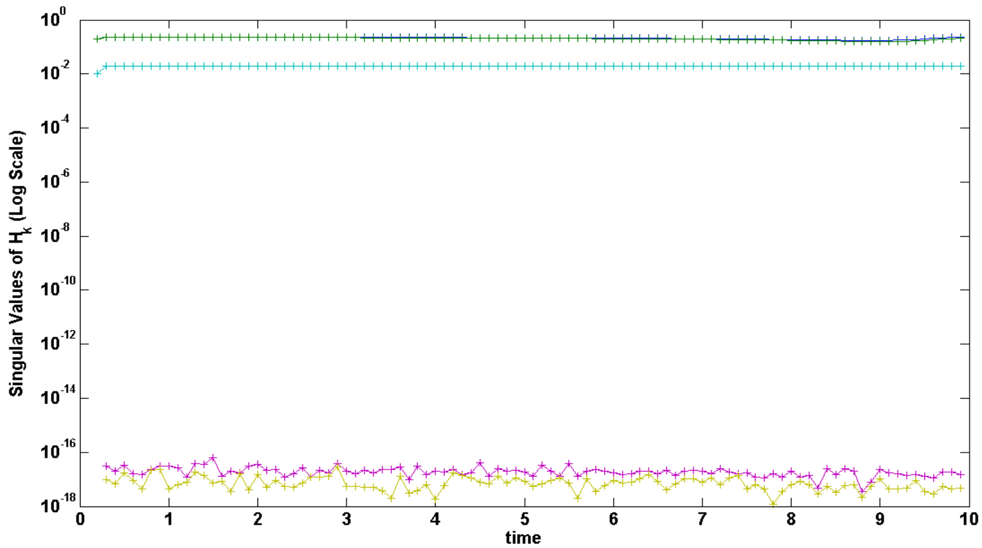

Figure 34. Identification Results: Singular Values of the Hankel Matrix Sequence

Validation of Identified Models

The Time Varying Eigensystem Realization algorithm is then applied to the generalized Hankel matrix sequence to obtain a linear time varying discrete time model sequence for the first order departure motion dynamics. The model sequence thus obtained was applied a sequence of known test function inputs (zero order hold approximation) and the response elicited from the identified model is compared with the response obtained from the numerical integration of the nonlinear equations representing the perturbed truth model. In case of the current example the test forces applied to the identified and the true model are given as,

$$\delta \mathbf{u}\left(t_k\right) = 0.1\begin{bmatrix} \sin\left(7t_k\right) \\ \cos\left(5t_k\right) \end{bmatrix} \tag{5.15}$$

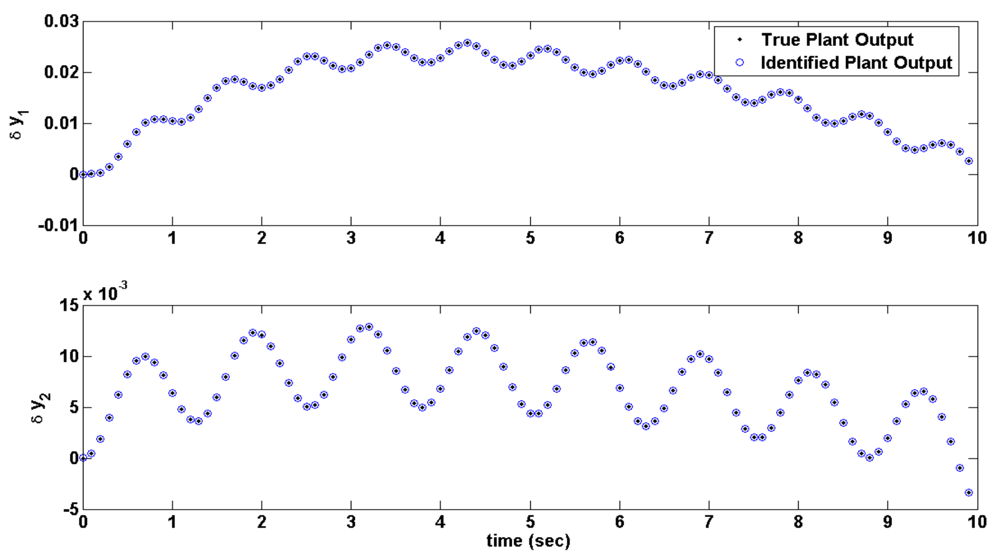

Figure 35. Identification Results: Position Response Comparison for Test Signal Inputs (Identified vs. True Nonlinear System)

Response of the truth model (sampled at the relevant time instances) and the identified model sequence outputs are compared in Figure 35(position coordinates) and Figure 36 (velocity coordinates).

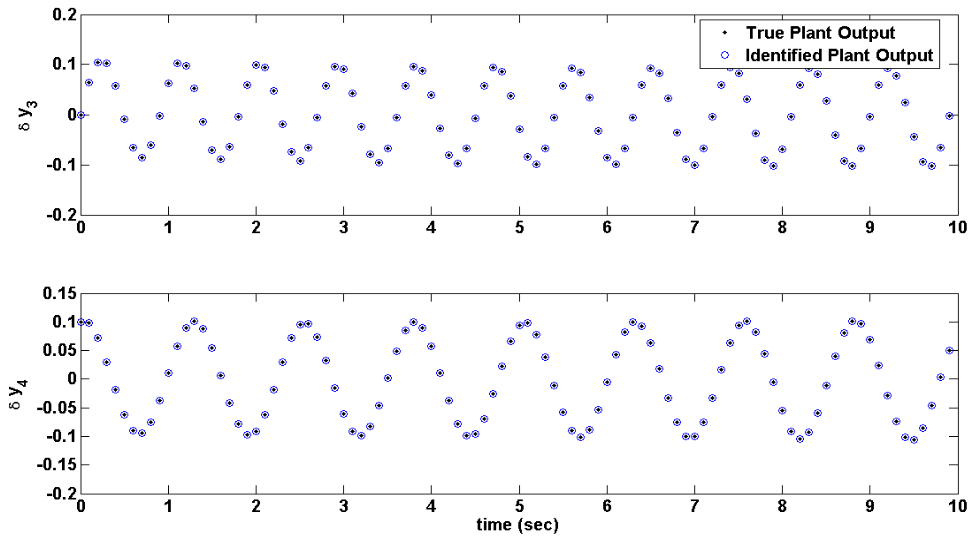

Figure 36. Identification Results: Velocity Response Comparison for Test Signal Inputs (Identified vs. True Nonlinear System)

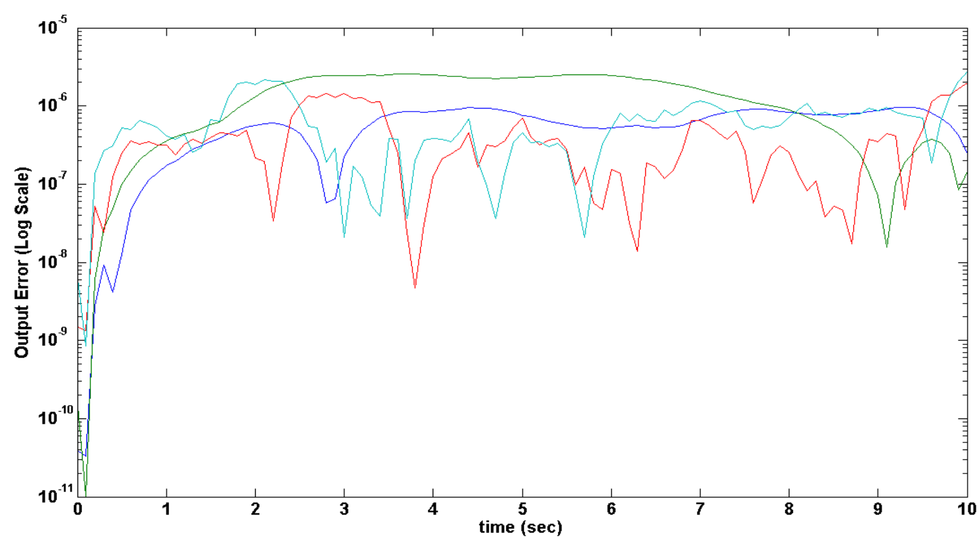

Figure 37. Identification Results: Error in Response Test Signal Inputs (Identified vs. True Nonlinear System)

The error (outputs/response deviations) incurred by the identification process is plotted in Figure 37.

Having been able to identify the first order dynamics of the departure motions about the given nominal solution of the present problem, we proceed to use this model sequence for control purposes. We make a note at this point that the coordinate transformation results presented in the chapter II have been used to transform the models in the same coordinate system for control purposes. Furthermore, since in the particular example, we assume that 4 independent sensor measurements are available throughout the time interval of interest, some interesting observations can be made in this case. Upon transforming the models in to the same coordinate system, the identified $\hat{C}_k$ matrix sequence was found to be time invariant (and therefore constant in time), given by,

$$\hat{C}_k = \begin{bmatrix} -0.0318 & 0.0190 & 0.0246 & 0.0798 \\ -0.0191 & -0.0318 & 0.0798 & -0.0246 \\ -0.3166 & 0.1893 & -0.0025 & -0.0080 \\ -0.1895 & -0.3161 & -0.0080 & 0.0025 \end{bmatrix} \tag{5.16}$$

$\forall k = 1, 2, \ldots 101$. This result of recovering the time invariant measurement sensitivity matrix upon transformation in to the same coordinate system is in principle, the proof of concept demonstration of the time varying coordinate systems and transformations developed in the chapter II. It should also be pointed out that the constant estimate of the measurement sensitivity matrix $\hat{C}_k$, is required in the state feedback controller design to transform the time varying model sequence in the same physical coordinate system as the nominal motion (all the while knowing that the true measurement sensitivity matrix C_k is the identity matrix in this problem).

Perturbation Guidance Using Identified Linear Time Varying Model Sequence

The model sequence thus identified using the Time Varying Eigensystem Realization Algorithm (TVERA) can be used to restore the performance requirements (state trajectory) assumed by the nominal solution in the absence of model errors and perturbations. In order to be able to track the optimal, unperturbed trajectory using perturbation guidance, we first set up a time varying reference trajectory to be tracked. Defining the discrete time reference trajectory, as,

$$\delta \mathbf{r}_k := \mathbf{x}^*\left(t_k\right) - \mathbf{x}^p\left(t_k\right) \tag{5.17}$$

would enable the definition of a tracking error state to be defined as, $\boldsymbol{\varepsilon}_k = \delta \mathbf{x}_k - \delta \mathbf{r}_k$ $= \mathbf{x}\left(t_k\right) - \mathbf{x}^*\left(t_k\right)$. With the definitions of appropriate reference trajectory and a corresponding tracking error, the feedback control law based on the perturbation model to track the reference in finite time involves the solution of an optimal control problem similar to the tracking problem outlined in table 1, but in discrete time domain. The necessary formulae are summarized in the following table, presented in the notation developed in this chapter for convenience of the reader.

Table 2 Summary of Optimal State Feedback Control for the Perturbation Guidance Scheme Using the TVERA Identified Departure Motion Dynamics Model

Dynamical System	$\delta\mathbf{x}_{k+1} = A_k \delta\mathbf{x}_k + B_k \delta\mathbf{u}_k$
Performance Index (Discrete tracking problem)	$\min_{\delta\mathbf{u}_k} J = \frac{1}{2}\boldsymbol{\varepsilon}_N^T Q_N^d \boldsymbol{\varepsilon}_N + \frac{1}{2}\sum_{k=0}^{N-1}\left[\boldsymbol{\varepsilon}_k^T Q_k^d \boldsymbol{\varepsilon}_k + \delta\mathbf{u}_k^T \Delta R_k \delta\mathbf{u}_k\right]$
Optimal Control	$\delta\mathbf{u}_k^* = -\Delta R_k^{-1} B_k^T \left(\Delta S_{k+1}\delta\mathbf{x}_{k+1} + \Delta\mathbf{v}_{k+1}\right)$
Feedback Gain Difference Equation	$\Delta S_k = A_k^T \left(\Delta S_{k+1}^{-1} + B_k \Delta R_k^{-1} B_k^T\right)^{-1} A_k + Q_k^d$ with final condition $\Delta S_N = Q_N^d$
Difference Equation for the Feed-Forward term calculation	$\Delta\mathbf{v}_k = A_k^T \left(\Delta S_{k+1}^{-1} + B_k \Delta R_k^{-1} B_k^T\right)^{-1} \Delta S_{k+1}^{-1}\Delta\mathbf{v}_{k+1} - Q_k^d \delta\mathbf{r}_k$ with final condition $\Delta\mathbf{v}_N = -Q_N^d \delta\mathbf{r}_N$

With the choice of the parameters (for perturbation guidance in finite time), $Q_N^d = 10^3 I_4$, $Q_k^d = 10^2 I_4$ and $\Delta R_k = I_2$, where I_N denotes the $N \times N$ identity matrix, the corrections incorporated from the discrete time guidance scheme following table 2. However, since the simulation runs in continuous time, the pre-calculated gains were used in a gain – scheduling type fashion, with the fixed gain multiplying the time varying departure motion dynamics, while the feed forward term was used in the usual zero order hold assumption. In other words, the implemented controller correction in real time integration was given as,

$$\delta \mathbf{u}_k^{\text{applied}}(t) = -\Delta R_k^{-1} B_k^T \left(\Delta S_{k+1} \delta \mathbf{x}(t) + \Delta \mathbf{v}_{k+1} \right) \tag{5.18}$$

$\forall t \in [t_k, t_{k+1})$, and $k = 1, \ldots, N-1$, which is a reasonable approximation to the optimal solution, provided sufficiently large sampling rate is provided to achieve acceptable performance. In the implementations provided in this chapter, a sampling rate of 10 Hz was chosen. The target position achieved by incorporating the perturbation discrete time guidance scheme as developed in this section is found to be $(x_T, y_T) = (598.6349,\ 598.7890)$ (which is closer to the optimal solution $(x_T^*, y_T^*) = (599.5897,\ 599.7858)$, while the absence of corrections lead to a final position of the point mass in the presence of perturbations to be $(x_T^p, y_T^p) = (578.8761,\ 581.1549)$).

A comparison of the three solutions is plotted in the x-y space in Figure 38. Zoomed view of the solutions focusing on the last few time step states are plotted in Figure 39. The comparison of position trajectories is presented in Figure 40, while the velocity trajectories are compared in Figure 41. The errors incurred from the optimal trajectory in the presence of perturbations and after incorporation of the identified model based compensation scheme are plotted in Figure 42. The extra control effort required in compensation is shown by the control profiles of Figure 43. Clearly a large amount of control expenditure is incurred in suppressing the unstructured uncertainty; however, the required reference is tracked with appreciable accuracy.

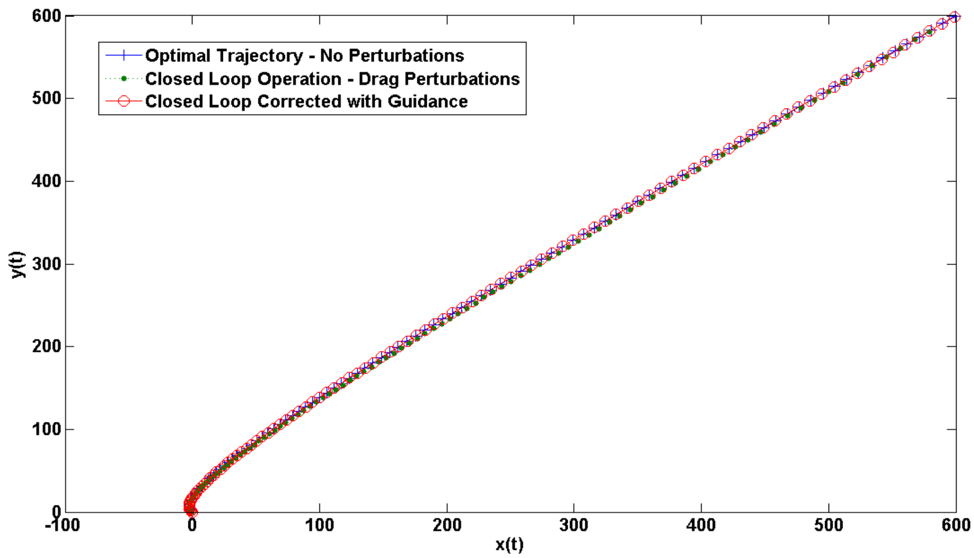

Figure 38. Guidance with Identified Model: Comparison with Other Trajectories (x1-x2 Space View 1)

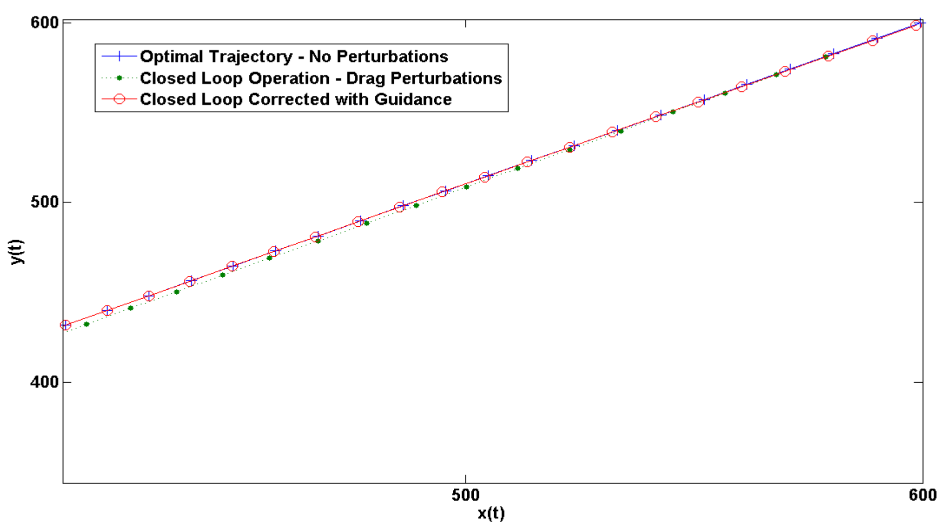

Figure 39. Guidance with Identified Model: Comparison with Other Trajectories (x1-x2 Space Zoomed View)

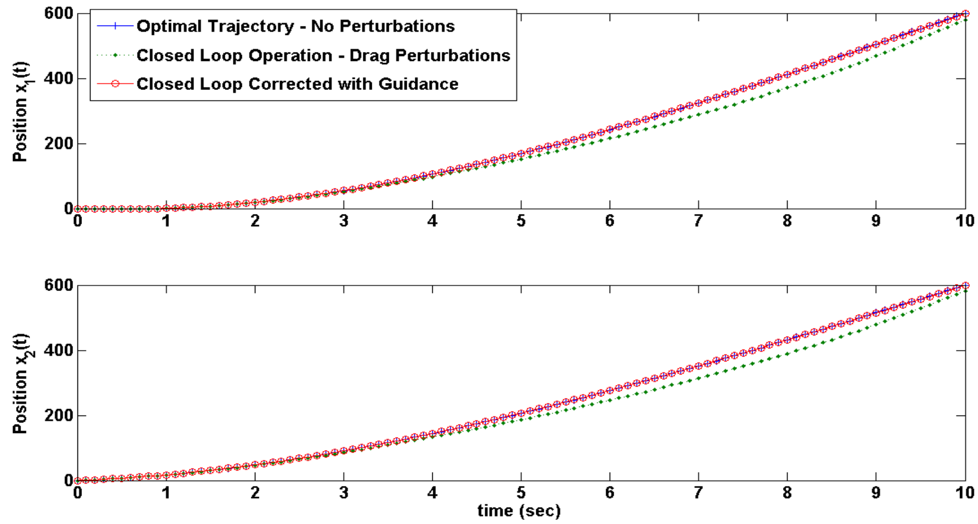

Figure 40. Demonstration of TVERA Identified Perturbation Guidance: Comparison of Position State Variables

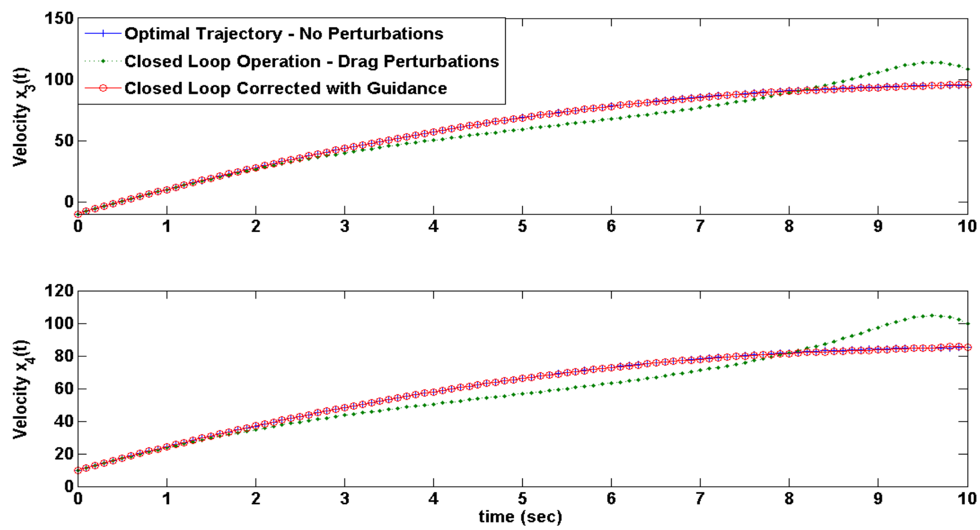

Figure 41. Demonstration of TVERA Identified Perturbation Guidance: Comparison of Velocity State Variables

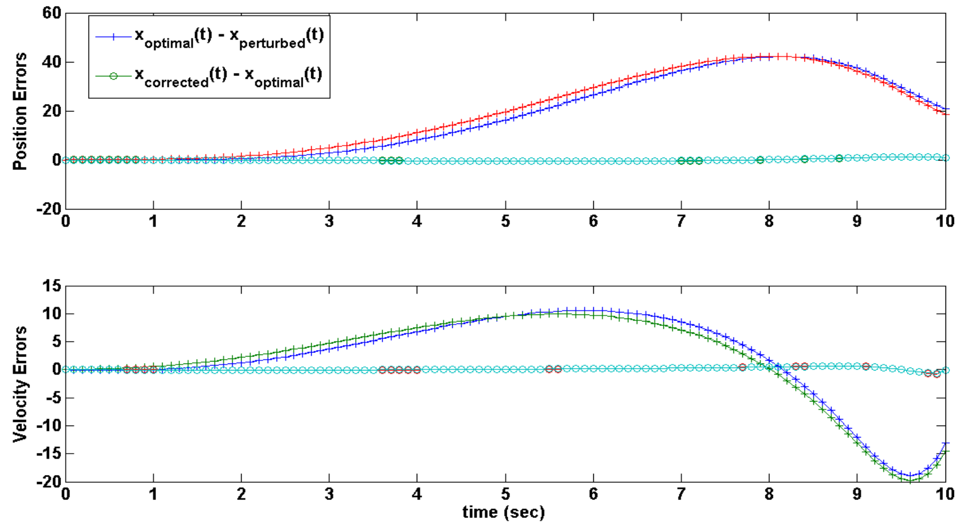

Figure 42. Demonstration of TVERA Identified Perturbation Guidance: State Deviations from the Optimal Trajectory

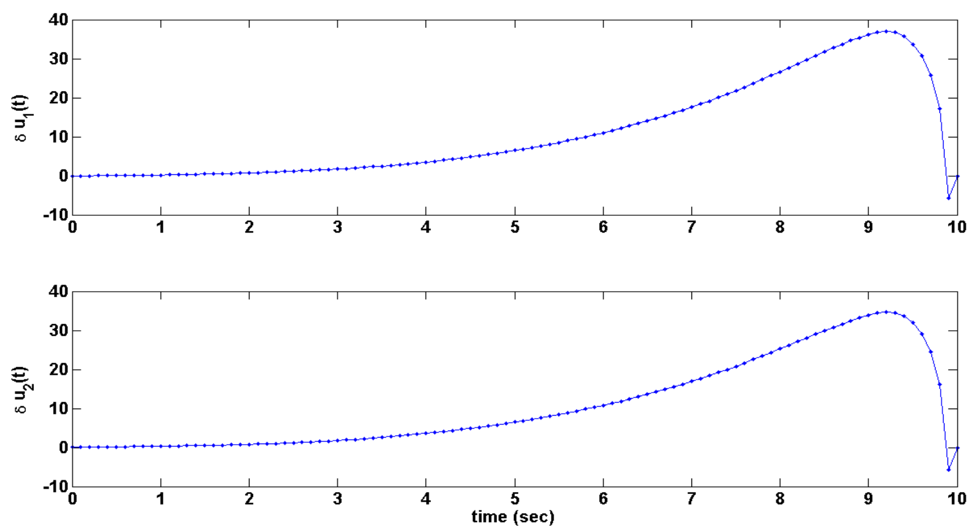

Figure 43. Demonstration of TVERA Identified Perturbation Guidance: Perturbation Guidance Discrete Corrections (Discrete Control)

Application of the Time Varying Identification Technique to a Problem in Dynamics [†]

Consider the dynamics of a point mass in a rotating tube as shown in the schematic of Figure 44.

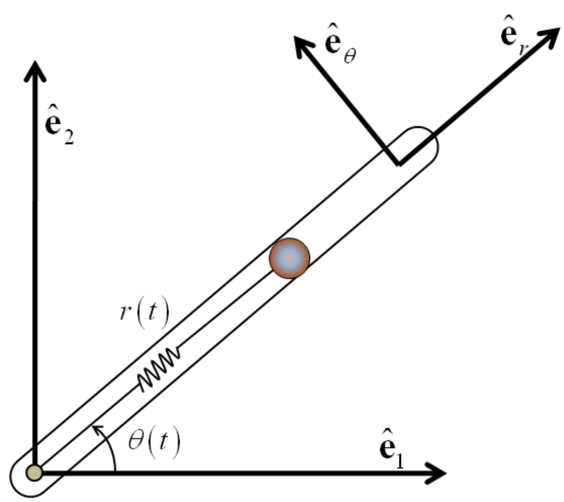

Figure 44. Schematic Depicting the Point Mass in a Rotating Tube

Dynamics of such a point mass is governed by a second order differential equation given by

$$\delta\ddot{r} = \left(\dot{\theta}^2 - \frac{k}{m}\right)\delta r + u(t) + l\dot{\theta}^2 \tag{5.19}$$

where the new variable $\delta r(t) := r(t) - l$, has been introduced, together with the definition of l as the free length of the spring (when no force is applied on it, i.e.,

[†] The author wishes to acknowledge Dr. John E. Hurtado for suggesting this problem.

Hooke's Law applies as $F_s = -k\delta r$). The function $u(t)$ is the radial control force applied on the point mass and the parameters k, m are the spring stiffness and the mass of the point mass of interest. The time variation in this linear system is brought about by the profile of the angular velocity of the rotating tube, $\dot{\theta}(t)$. Choosing the origin of the coordinate system at the position $\mathbf{r}_0 = l\hat{\mathbf{e}}_r$ (with no loss of generality) along the $\hat{\mathbf{e}}_r$ direction, we have the second order differential equations to be given by,

$$\delta\ddot{r} = \left(\dot{\theta}^2(t) - \frac{k}{m}\right)\delta r + u(t) \tag{5.20}$$

where the redefinition of the origin renders the system linear time varying without any extra forcing functions.

In the first order state space form ($x_1(t) := \delta r(t), x_2(t) := \delta\dot{r}(t)$), the equations can be written as,

$$\begin{aligned} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ \dot{\theta}^2(t) - \dfrac{k}{m} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) \\ &= A(t)\mathbf{x}(t) + B(t)u(t) \end{aligned} \tag{5.21}$$

together with the measurement equations,

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + 0.1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} u(t) \tag{5.22}$$

To make comparisons with the identified models, analytical discrete time models were also generated by computing the state transition matrix (equivalent A_k) and the convolution integrals (equivalent B_k , with a zero order hold assumption on the inputs). Since the system matrices are time varying, matrix differential equations given by

$$\begin{aligned}\dot{\Phi}\left(t,t_k\right) &= A\left(t\right)\Phi\left(t,t_k\right)\\ \dot{\Psi}\left(t,t_k\right) &= A\left(t\right)\Psi\left(t,t_k\right)+I\end{aligned} \tag{5.23}$$

$\forall t \in \left[t_k, t_{k+1}\right]$ with initial conditions $\Phi\left(t_k,t_k\right)=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \Psi\left(t_k,t_k\right)=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$ such that,

$$\begin{aligned}A_k &:= \Phi\left(t_{k+1},t_k\right)\\ B_k &:= \Psi\left(t_{k+1},t_k\right)B\end{aligned} \tag{5.24}$$

would represent the equivalent discrete time varying system (truth model). Integration of the matrix differential equations was carried out with a tolerance of 1×10^{-13} (Dormand-Prince solver – subroutine 'ode45' of MATLAB). For the current investigation the time variation profile of the $\dot{\theta}\left(t\right)=3\sin\left(\frac{1}{2}t\right)$ with the mass and stiffness of the system chosen to be $m=1, k=10$. The time interval of interest was held to be 50 seconds, with the discretization sampling frequency of interest set to be $1Hz$.

Time Varying System Identification methods developed in this dissertation were employed using the input-output test data from the models above and model sequences were obtained. Identification process starts with the determination of generalized Markov parameters, similar to the procedure indicated in chapters II and III. Errors incurred in the determination of these Markov parameters are shown in Figure 45.

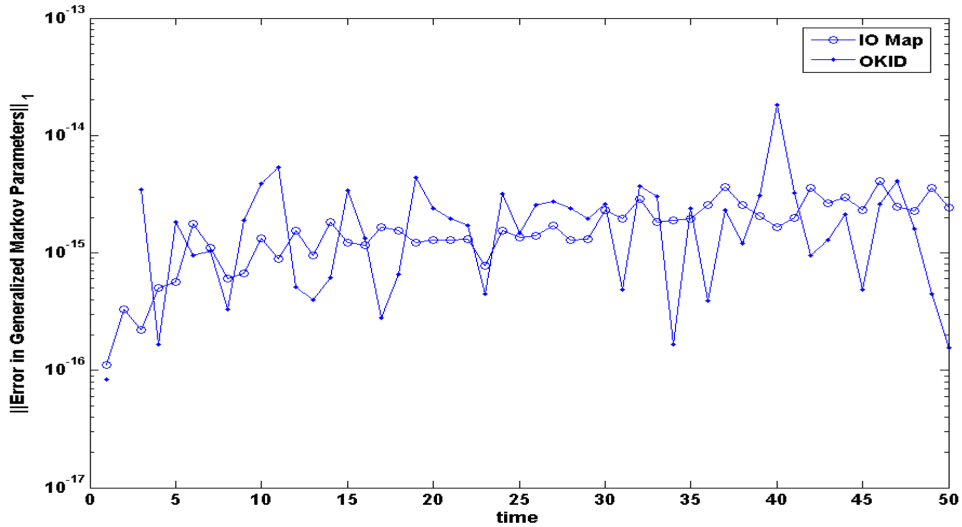

Figure 45. Errors of the Identified Generalized Markov Parameters (Using Least Squares Solution and the Time Varying OKID Procedure)

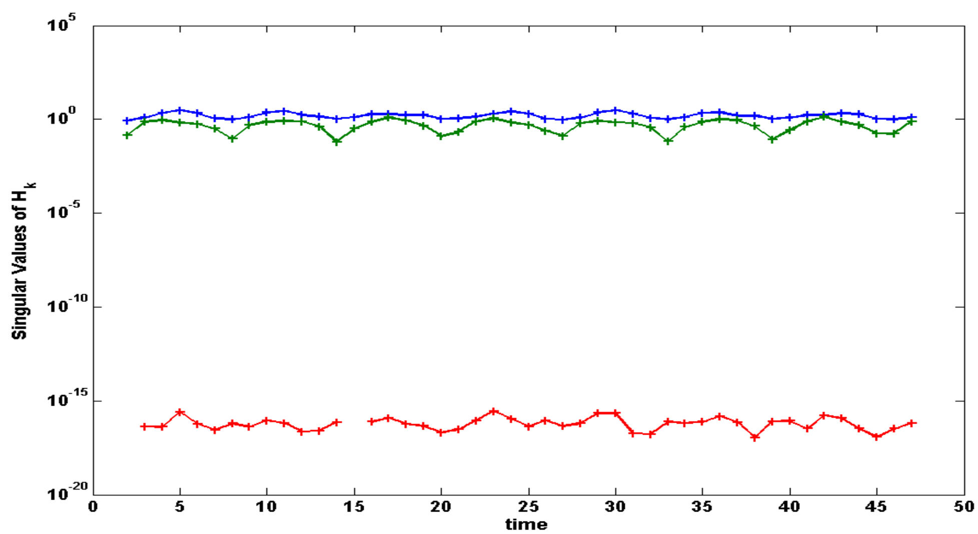

Figure 46. Singular Values of Hankel Matrix (Point Mass in a Rotating Tube)

The singular values of the generalized Hankel matrix sequence are plotted as Figure 46. Applying a test input force $u(t) = \frac{1}{2}\sin(12\ t)$, to the true and identified system matrix sequences, the error incurred in the response is shown in Figure 47 while the response profiles are compared in Figure 48. Response profiles appear jagged to show their sampled nature.

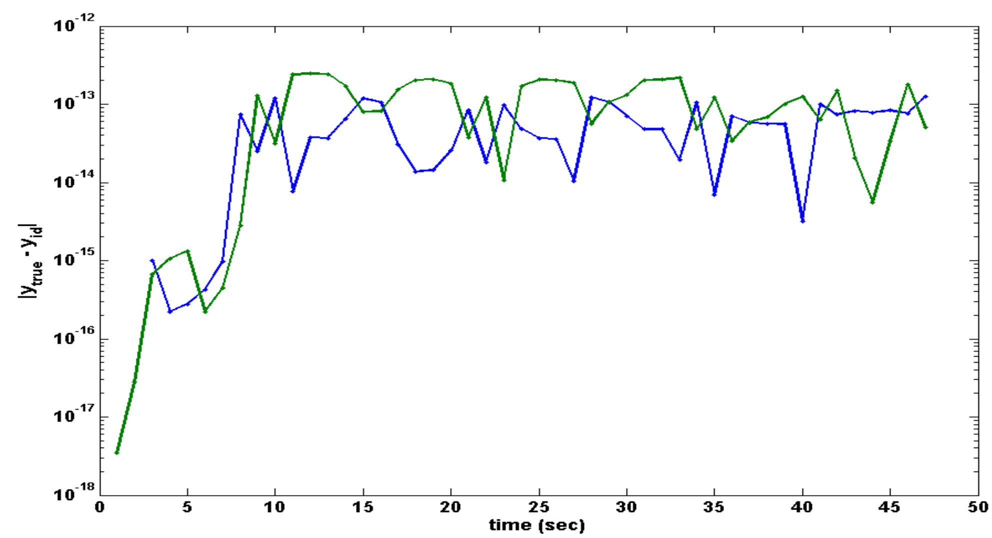

Figure 47. Error in the Identified System Response

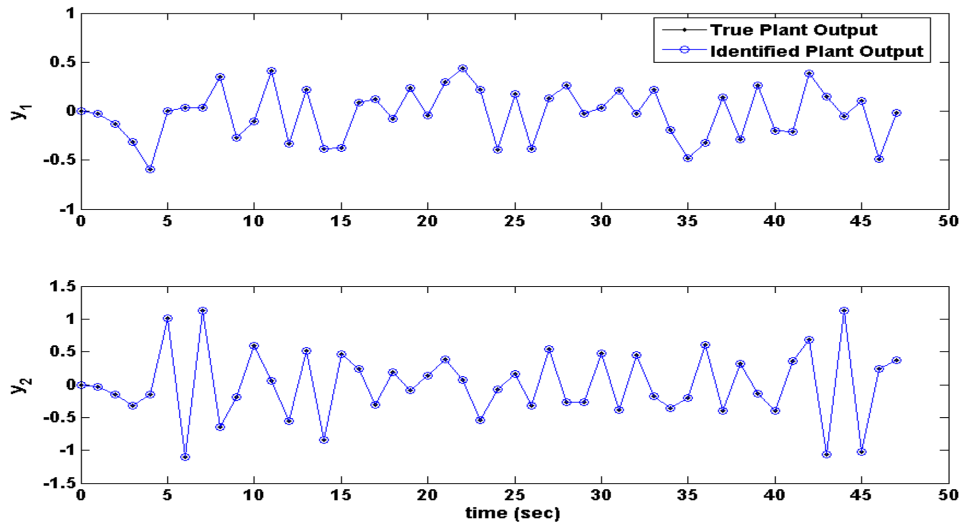

Figure 48. Comparison of System Response (True vs. Identified) to Test Control Input Function

Discussion on the Identified Time Varying Coordinates

This simple physical example helps us explain the time varying coordinate systems and the transformation process. To bring further clarity in to the discussion, we use the same number of sensors as the true dimensionality of the state space ($m = 2$ for this problem). Also, the generalized Hankel matrix is populated with only one redundant time step such that the Observability grammians are non-redundant and hence lead to exact inverse (as opposed to pseudo-inverse) in the transformations. We will first explain the transformations for this simplified situation and then proceed to a short discussion on what happens in the general situation.

For the current problem, the coordinate system calculations are simplified owing to the measurement sensitivity matrix being identity as given by equation (5.22). Recall from chapter II that this implies

$$\hat{C}_k = C_k T_k = T_k \tag{5.25}$$

for this problem. Therefore, if a non identity matrix is realized by the identification procedure, the time varying coordinates are transparently obtained by the problem set up in this simplified setting. Considering four representative time steps, we plot the coordinate systems in Figure 49.

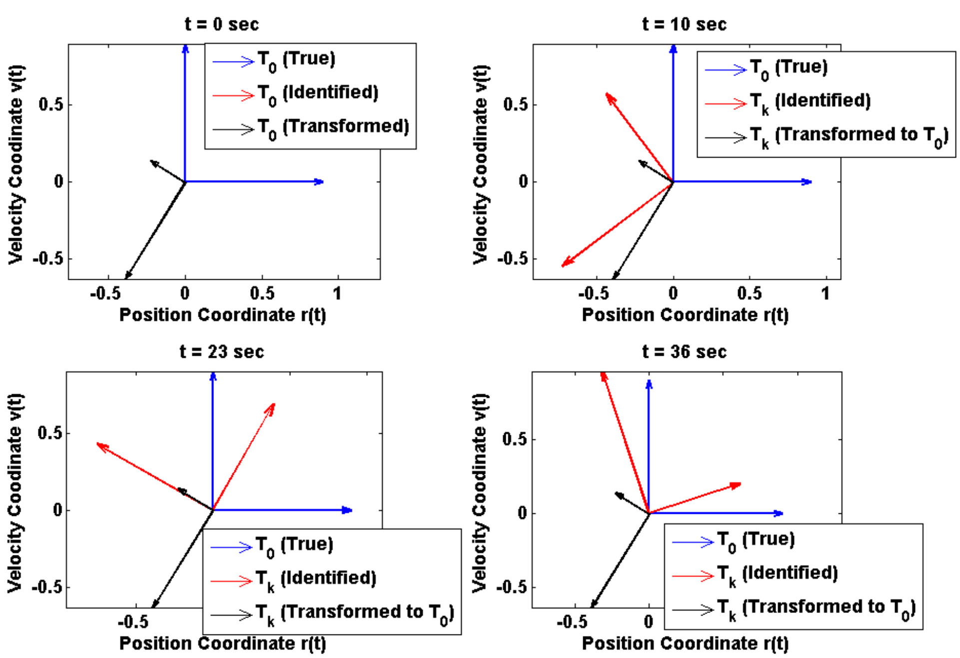

Figure 49. Time Varying Coordinate Systems: Graphical Demonstration of the Transformation Process (Special Case - Number of Sensors Matching the State Dimension)

Note that in Figure 49 above, the blue arrows indicate the reference directions in the state space representing the columns of the true C_k matrix at the corresponding time step. Red arrows plot the columns of the identified $\hat{C}_k^{T_k}$ matrix. They represent the time varying coordinates that are realized by the identification algorithm. The black arrows represent the columns of the identified $\hat{C}_k^{T_0}$ matrix after transformed in to the reference coordinate system. A clear demonstration of the transformation process is obtained by observing that at each time step, the transformed coordinates align with the reference coordinate system (at time t_0 or any other reference time step of interest).

For the more general situation of $m < n$, owing to the arbitrariness of the "free" basis vectors ($n - m$ of them exist at each time step), this elegant projection on to the same subspace is not defined uniquely and hence the basis is completed arbitrarily (at every time step) to produce the necessary inversion (pseudo-inversion to make a precise statement). The arbitrary completion of basis leads to a time varying correction. It also depends upon the number of time steps considered for constructing the Observability grammian through the least squares pseudo inverse constructed in the process of transformation. Considering different time steps would in general lead to a different transformation matrix. We now proceed to show that the same example problem can be represented as a bilinear system and demonstrate the continuous time bilinear system identification algorithm developed in this dissertation.

Application of the Continuous Time Bilinear System Identification Technique to a Problem in Dynamics

Now consider an alternative situation where the model governing the dynamics of the particle in a rotating tube shown in Figure 44 is unforced $u(t)=0$ with the origin defined such that $l \neq 0$ in the system of equations (5.19). For clarity of presentation, these equations are repeated here as,

$$\delta\ddot{r} = \left(-\frac{k}{m}\right)\delta r + \delta r\dot{\theta}^2 + l\dot{\theta}^2 \tag{5.26}$$

Defining the state variables to be $x_1(t) := \delta r(t), x_2(t) := \delta\dot{r}(t)$ and the angular velocity squared as the control input, $u(t) = \dot{\theta}^2(t)$, we have the following system of equations (in the bilinear form)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\dfrac{k}{m} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} u(t) + \begin{bmatrix} 0 \\ l \end{bmatrix} u(t) \tag{5.27}$$

along with the measurement equation,

$$y(t) = x_1(t) + 0.1u(t) \tag{5.28}$$

Note that as opposed to the situation for which the methods of chapter IV were developed, the problem in the current application is limited in scope due to the fact that the input magnitude is constrained to be positive ($u(t) > 0$). Although it was not found to be a limiting factor in the identification problem posed for problems with single degree of freedom, the excitation was found to be insufficient for problems of higher degrees of freedom (more masses attached in the tube). Therefore, in such situations the performance of the identification algorithms was found to be relatively poor.

The identified system matrices (plant parameters) computed using the procedure developed in chapter IV are found to be,

$$\begin{aligned}
\hat{A}_c &= \begin{bmatrix} 2.4332 & -0.014195 \\ 1121.6 & -2.4332 \end{bmatrix} \\
\hat{N}_{c1} &= \begin{bmatrix} 0.023373 & 0.00078133 \\ -0.69916 & -0.023373 \end{bmatrix} \\
\hat{B}_c &= \begin{bmatrix} -1.3086 \\ 39.1459 \end{bmatrix}; \;\; \hat{C} = \begin{bmatrix} -0.0179 & -0.0006 \end{bmatrix}
\end{aligned} \tag{5.29}$$

The errors in system matrix eigenvalues (representative of the identification errors are found to be)

$$\begin{aligned}
\left\| \lambda\left(A_c\right) - \lambda\left(\hat{A}_c\right) \right\| &= \begin{bmatrix} 0.2407 - 0.1394\mathrm{i} \\ 0.2407 + 0.1394\mathrm{i} \end{bmatrix} \times 10^{-10} \\
\left\| \lambda\left(N_{c1}\right) - \lambda\left(\hat{N}_{c1}\right) \right\| &= \begin{bmatrix} -0.0000 + 0.2557\mathrm{i} \\ -0.0000 - 0.2557\mathrm{i} \end{bmatrix} \times 10^{-5}
\end{aligned} \tag{5.30}$$

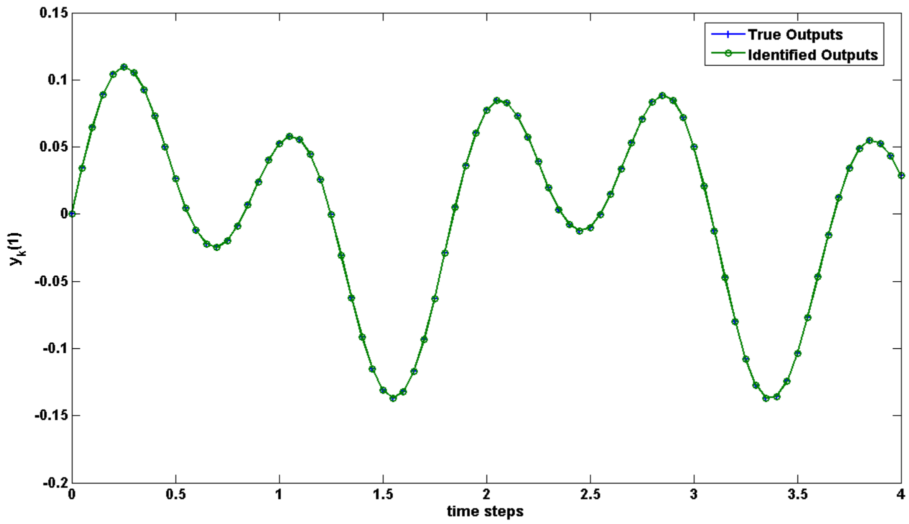

Figure 50. Response Comparison: True and Identified Models (Point Mass in a Rotating Tube - Bilinear)

Note the decrease in the accuracy of identification process. This is owing to the limitation of the admissible forcing functions to positive values. Consequently excitation is not rich enough to extract the unknowns in the problem. Using a test function input profile $u(t)=\sin(7t)$, the response obtained from the identified and true systems are compared in the Figure 50. Errors incurred in the response to this test function are plotted in Figure 51.

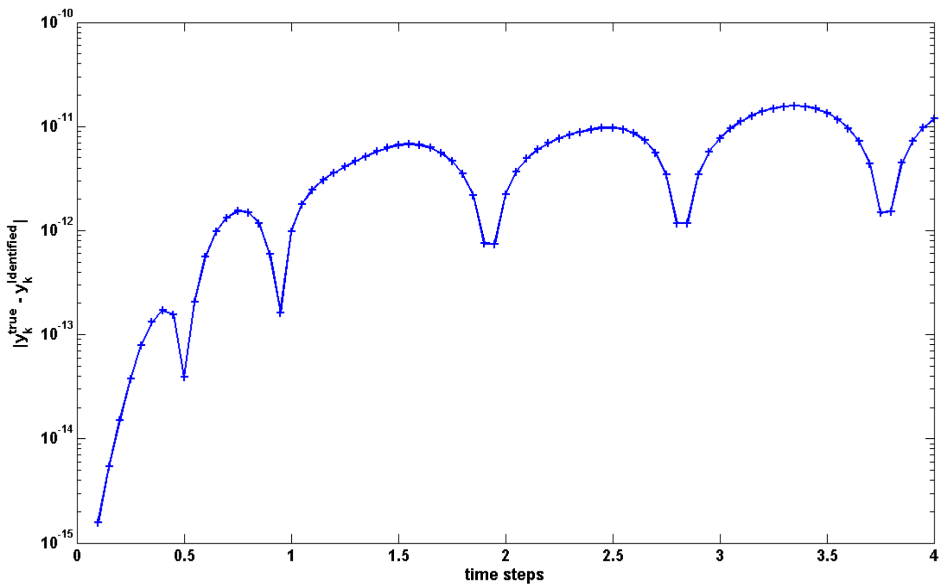

Figure 51. Error in Response between True and Identified Models (Point Mass in a Rotating Tube - Bilinear)

Application of the Continuous Time Bilinear System Identification Technique to an Automobile Brake Problem

Following the developments of Mohler[39], we consider the problem of modeling the deceleration dynamics of an automobile under the action of a braking system. A schematic of the braking mechanism is illustrated in Figure 52

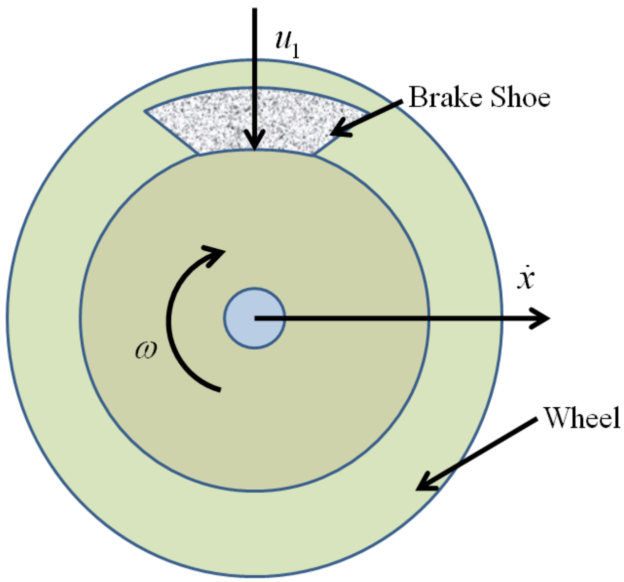

Figure 52. Schematic of an Automobile Braking Mechanism

A mechanical drum brake shown in the schematic operates by the frictional force it produces upon contact with the axle attached to a rotating wheel. Frictional force produced in the process is typically modeled (neglecting Coulomb friction) as,

$$f_b = c_b u_1(t)\dot{x}(t) \tag{5.31}$$

where $u_1(t)$ is the braking force applied to the drum, c_b the coefficient of friction, $\dot{x}(t)$ is the translational velocity of the axle (and the vehicle). Denoting the mass of the automobile by m and designating $u_2(t)$ to represent the engine force acting on the vehicle, a simple model for the dynamics of the braking motion can be obtained as,

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -\frac{c_f}{m} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & -\frac{c_b}{m} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} u_1(t) + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} u_2(t) + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} \tag{5.32}$$

where c_f denotes any further damping present in the dynamics of the problem while the state variables x_1, x_2 represent the position and velocity variables of the vehicle dynamics. It was assumed that both position and velocity of the vehicle were assumed for measurement and the output equation is given by

$$\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + 0.1\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{5.33}$$

Choosing representative values of the damping coefficients to be $c_f = 2m, c_b = 5m$, we get the following identification results. Identified plant model parameters are calculated as,

$$\begin{gathered} \hat{A}_c = \begin{bmatrix} -1.6589 & -0.0092 \\ -61.6885 & -0.3411 \end{bmatrix} \\ \hat{N}_{c1} = \begin{bmatrix} -4.9587 & -0.0274 \\ -7.4734 & -0.0413 \end{bmatrix};\ \hat{N}_{c2} = \begin{bmatrix} -0.2421 & -0.1295 \\ -0.8528 & -0.4205 \end{bmatrix} \times 10^{-12} \\ \hat{B}_c = \begin{bmatrix} 0.3\times10^{-14} & -3.7211 \\ 0.14\times10^{-15} & -5.6082 \end{bmatrix};\ \hat{C} = \begin{bmatrix} -0.0068 & 0.0045 \\ -0.2665 & -0.0015 \end{bmatrix} \end{gathered} \tag{5.34}$$

Identification errors as represented by the errors between true and identified system matrix eigenvalues are given by

$$\left\|\lambda\left(\hat{A}_c\right)-\lambda\left(A_c\right)\right\|=\begin{bmatrix}0.1584\\ \text{-}0.1563\end{bmatrix}\times 10^{-11}$$
$$\left\|\lambda\left(\hat{N}_{c1}\right)-\lambda\left(N_{c1}\right)\right\|=\begin{bmatrix}\text{-}0.4778\\ 0.5950\end{bmatrix}\times 10^{-12} \tag{5.35}$$
$$\left\|\lambda\left(\hat{N}_{c2}\right)-\lambda\left(N_{c2}\right)\right\|=\begin{bmatrix}0.6753\\ \text{-}0.0127\end{bmatrix}\times 10^{-12}$$

Using test input profiles $u_1(t)=\sin(7t), u_2(t)=\cos(5t)$, the response obtained by the identified system parameters and the true system parameters are compared in Figure 53.

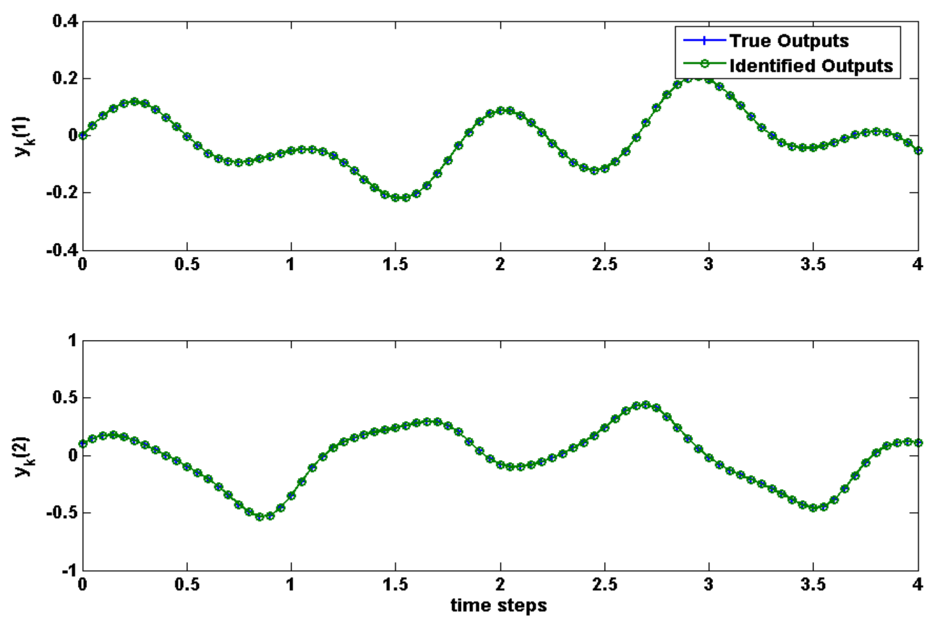

Figure 53. Comparison of Response to Test Function Inputs (Automobile Brake Problem)

Errors incurred in response channels are plotted in Figure 54.

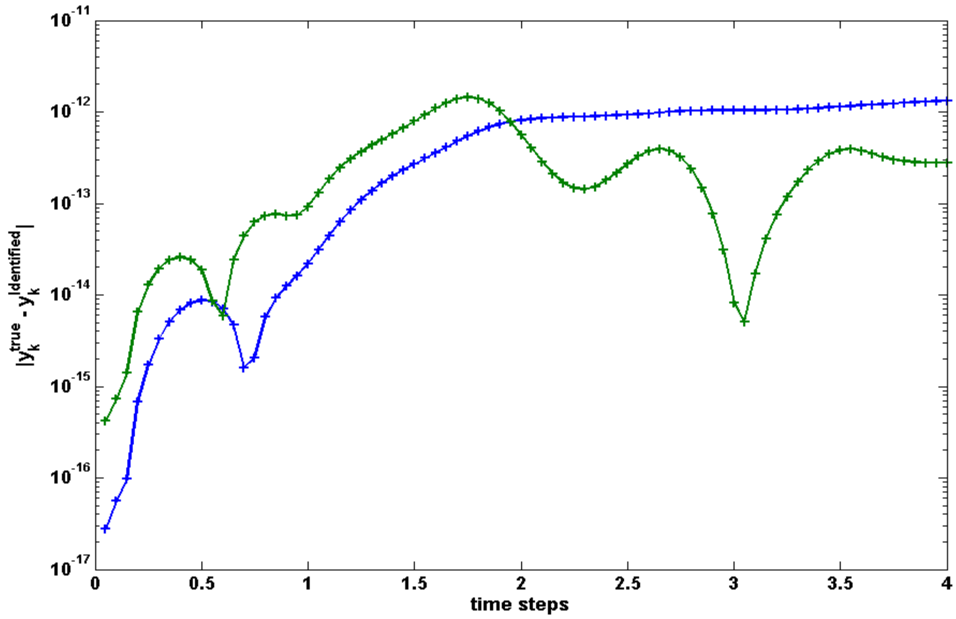

Figure 54. Error in Response to Test Function Inputs (Automobile Brake Problem)

The simple examples of this chapter clearly show the broad applicability of the methods developed in this dissertation. Promising results obtained indicate progress towards the next generation of algorithms for identification of plant models for dynamic systems.

Conclusion

A simple two dimensional intercept problem is used to demonstrate the capabilities and potential applications of the time varying eigensystem realization (TVERA) algorithm developed in this dissertation. Considering the embedded state feedback gains the closed loop of the plant becomes time varying and incurs errors in operation due to unstructured uncertainties in the form of modeling errors (Drag perturbations are considered).

Subsequent application is a model governing the dynamics of a point mass rotating in a tube forming a naturally time varying system. Time varying system identification methods (TVERA, TOKID) are applied to this problem and reliable model sequences are obtained for a given angular velocity profile. The time varying coordinate systems and the nature of associated transformations are discussed clearly in the context of this physical example.

The same problem with simple redefinition of variables is shown to be bilinear in nature and the continuous time bilinear system identification methods are applied for the identification of the model parameters in this domain. An automobile brake problem is then detailed as a demonstration application of the continuous time bilinear system identification algorithms developed in this dissertation.

CHAPTER VI

CONCLUSIONS AND FUTURE DIRECTIONS

"The outcome of any serious research can only be to make 'n' questions grow where only one grew before."

- Paraphrased from a quote by Thorstein Veblen by Prof. John Junkins

In this spirit, it can be surmised that the efforts undertaken for the investigations comprising of this dissertations are very serious. For most part, the techniques developed in this dissertation raise more questions and bring more interesting problems into focus that call for more investigation. However, that being said, it is felt that the results developed are of sufficient gravity and maturity that near term implementations will result.

We outline the accomplishments made and the challenges and opportunities presented by this dissertation topic-wise below.

Time Varying Eigensystem Realization Algorithm

The first chapter details an identification algorithm called the Time Varying Eigensystem Realization Algorithm (TVERA) is proposed to realize discrete time varying plant models from input output experimental data. It is shown that this singular value decomposition based method is a generalization of the celebrated Eigensystem Realization Algorithm developed by Juang et. al.[1], to realize time invariant models from pulse response sequences. Using the results from discrete time identification theory, the generalized Markov parameter and the generalized Hankel matrix sequences are computed via a least squares problem associated with the input-output map. The computational procedure under investigation outlines a methodology to extract a state space plant model sequence from the generalized Hankel matrix sequence in several different time varying coordinate systems. The concept of free response experiments is recommended to identify the subspace of the unforced system response, providing a consistent methodology to realize system models for the first few time steps. The algorithm developed leads to a tool set (presented in this dissertation) that enables seam-less integration of model sequences which might have been obtained from different algorithms. For the special case of systems with fixed state space dimension, the free response subspace is used to construct a uniform coordinate system for the realized models at different time steps. Numerical simulation results on general systems are presented to investigate the effectiveness of the algorithms developed.

Although the developments of this chapter bring much of the literature in time varying system realization theory to a much more mature state of evolution, the new tools developed also open new opportunities of investigation. The physical nature of time varying eigenvalues begs for some immediate investigation. It must definitely mean something important for the invariance of eigenvalues to result when viewed from a certain family of consistent coordinate systems (a controllable/observable subspace). The change of eigenvalues and the concept of time varying modes has attracted some recent attention[14]. Also, the assumption of zero initial condition needs to be relaxed at some point. While the time varying OKID procedure partially redresses this issue, the first few time steps of the time varying dead beat observer realized needs a specific solution, which we have not presented in this dissertation. The constrained projection scheme to develop the transformation matrices subject to known constraints is also not presented here. A closed form solution for such transformations has been obtained by the author. These discussions were avoided to make the presentation details of the central ideas clear. These results become important only once the core of the algorithms are communicated. The author and his collaborators plan to make these results accessible within a year from the publication of the main algorithms (TVERA/TOKID).

While the requirement of repeated experiments can be overcome by using special types of input sequence sets, measured for a large time interval, this approach may not be possible for systems that are not necessarily periodic or quasi periodic. Extensive work still needs to be done to classify and separate systems based on whether they can be identified from a single set of experimental data (design of experiments). Theoretical identifiability and realizability needs to be distinguished in this context and approaches that lead to practical algorithms need to be developed. The author feels that significant amount of work still needs to be done in this regard. Future research efforts for this topic would focus on such extensions and generalizations (relaxation of assumptions) while bearing sufficient emphasis on practical applications of the theoretical ideas.

Observer Markov Parameter Theory for Time Varying Eigensystem Realization Algorithm

In the second chapter, an algorithm for computation of the Markov parameters of an observer or Kalman filter from input output experimental data is discussed. The relationships between the observer Markov parameters and the system Markov parameters are derived for the time varying case and are found to be generalizations of the developments for the OKID algorithm for the time invariant systems developed by researchers in the past. The time varying sequence of system Markov parameters and the time varying observer (or Kalman filter) gain Markov parameter sequence are projected to be obtainable using time varying generalizations of the recursive relations developed in the time invariant case from the generalized time varying observer Markov parameters. The system Markov parameters thus derived are to be used by the time varying Eigensystem realization algorithm developed in the first chapter, to obtain a time varying discrete time state space model for controller design purposes. Connections with the Kalman observer in the stochastic environment and an asymptotically stable realized observer are qualitatively discussed to develop insights for the analyst. A minimum number of repeated experiments for accurate recovery of the system Markov parameters is derived from these developments, which is vital for the practicing engineer to design multiple experiments before analysis and model computations. Numerical examples demonstrate the utility of the approach presented.

Topics developed in relation to the time varying OKID procedure have in fact opened up new avenues of research for the future. The concept of time varying moving average models (GTV-ARMAX), an obvious stochastic extension to the GTV-ARX models developed in this dissertation are of imminent interest for development of time varying filters. A first look was made at the relationship of the GTV-ARX observer and the Kalman filter and the analogy is near-perfect. This qualitative analogy looks at the fading memory element of the Kalman filter process noise with the number of terms in the time varying moving average model. However, quantitative relationships between the realized observers and the time varying versions of the discrete time Kalman filter need to be studied in greater detail to address the following issues:

1. In case of the time invariant theory, explicit relations between the autoregressive moving average models and the associated observers exist. This allows the analyst to gain useful information on the spectral content of the filter that gets discarded from the original signal in order to minimize the state estimation error covariance (some discussions are presented in section 4.5 of Anderson and Moore[40]). Since frequency domain methods cease to exist for time varying systems, these connections are difficult to make. A significant amount of work needs to be done in order to understand the underpinnings of these connections.
2. On the other hand, in case of the time invariant systems, one often is able to arrive at an estimate of the error statistics of the process noise. While this is central to adaptive filtering[41], similar ideas are used in the improvement of identification results in a method known as “Residual whitening” (based on a version of the OKID algorithm) quite successfully[42]. Similar to the previous

statement, the connections for this inverse problem are not straightforward to make.

3. Considerable research needs to be done in order to investigate the role of an optimal observer in the identification process. Although the time varying deadbeat observer realized by the GTV-ARX model is the fastest possible observer, one cannot obtain the same speed from an optimality perspective where we are interested in minimizing the error covariance. This seemingly dual problem is an interesting issue for further investigations.

Further analysis and research is required in the area of time varying dead beat observers discussed in this dissertation. Specifically, upon close inspection, the central ideas of the computational procedure developed here-in and the classical procedures developed to compute the time varying dead beat observers[43, 44] have some strong similarities. A direct demonstration of their identical nature (if the steps involved are indeed the same) or a relationship between the algorithms (direct correspondence between the steps involved in the calculations) would aid in unifying the theory presented here-in with existing literature.

Continuous Time Bilinear System Identification

Identification of continuous time bilinear system plant models, from input output data associated with multiple experiments is the topic of discussion in chapter III. Making use of recent advances in bilinear system identification, the results of the chapter take advantage of the experimental data from multiple experiments and set up a procedure to obtain bilinear system models. It is shown that the special pulse inputs employed by earlier research can be avoided and accurate identification of the continuous time plant model is possible by performing multiple experiments incorporating a class of piece wise constant control input sequences introduced in this dissertation. Avoiding the practical difficult step of pulse input generation and application, makes the algorithm proposed in this chapter more attractive in practice for the identification of bilinear systems. Furthermore, using the developments designed in this chapter, one obtains the plant models in the same coordinate system automatically. Numerical examples demonstrate a basis for optimism for the methodology developed to solve many members this class of problems.

Bilinear system identification theory has several significant prospective research opportunities for exploration. The first effort is to extend the Markov parameter determination procedure (input dependent) to develop an OKID type algorithm, giving rise to a natural nonlinear autoregressive model with exogenous input (NARX), where the feedback is quadratic in the residual error. Connections to the existing nonlinear estimation theory [45] are important problems for investigation. High order perturbation models in the form of state transition tensors (Volterra Kernels), were recently found to be attractive for applications in nonlinear estimation and trajectory optimization[45]. According to Rugh, the state transition tensors are actually time varying bilinear systems derivable from the equations of motion of a nonlinear system using a technique called the *Carleman linearization* about a reference trajectory[29]. Thus extensions of the bilinear time invariant system identification methods presented here-in to realize time varying bilinear models appears to be a promising avenue for further research, owing to the increasing interest in theory and applications of Volterra kernel based nonlinear identification methods.

Applications to Problems in Guidance, Control and Dynamics

The theory of time varying system identification is applied (in chapter V) to realize the first order departure motion dynamics model sequence about a reference trajectory for an intercept problem operating in the presence of unstructured uncertainty of drag perturbations. The model sequence set thus realized is used to design a controller to stay close to the optimal trajectory in the absence of the unmodeled dynamics, with satisfactory performance and tracking.

The author notes that this is just one simple representative of a plethora of possible applications and dynamical systems. In principle, the current demonstration is directly applicable to obtain first order perturbation models for several complex nonlinear systems operating in a reference trajectory. Reduced order modeling (from high fidelity multiphysics model simulations) is yet another application, where the information from complex simulations can be effectively “compressed” in to a time varying perturbation linear model, based on which control decisions can be made. Since this is the first (and only, to date) example solved, it is difficult to extrapolate the impact of this approach until more example problems with higher dimensionality are addressed. Thus the perturbation guidance problem is recommended for further study.

As shown in the simple dynamics problem, several interesting problems exist where the dynamics is inherently time varying. Such physical problems need to be studied more to understand the true nature of the time varying eigenvalues and their significance and possible utility in the analysis and design of dynamical systems.

The physical applications discussed, demonstrating the effectiveness of continuous time bilinear system identification algorithm show optimism on their broad applicability. However, the method outlined here does not apply for two situations. One of the problems involves the case when the true $B_c = 0$ (this happens in a practical application of the models of a nuclear power plant / reactor [39]) and the other when $A_c = 0$ (although this case technically falls out of the scope of the present study since the system is not observable). Simple modifications to the existing technique are found by the author to circumvent the associated problems. The first problem is circumvented by using initial condition response (free decay experiments) while the second problem is detected and circumvented using the rank test of the output sequence and by-passing some steps in the current algorithm. They will be reported in a separate communication made available to the research community within the next year.

Thus the methods developed in the current dissertation are representative of a wide variety of applications quite useful for analytical and practical engineers. Numerical examples offer optimism to the author that at least many of the algorithms and their near-term technical descendants will stand the test of time.

REFERENCES

[1] Juang, J.-N., *Applied System Identification*, Prentice Hall, Englewood Cliffs, NJ, 1994.

[2] Ljung, L., and Soderstorm, T., *Theory and Practice of Recursive Identification*, MIT Press, Cambridge, MA, 1983.

[3] Ho, B. L., and Kalman, R. E., "Effective Construction of Linear State Variable Models from Input/Output Data," *Regelungstechnik,* vol. 41, 1966, pp. 545 - 548. doi: NA

[4] Crassidis, J. L., and Junkins, J. L., *Optimal Estimation of Dynamic Systems*, CRC Press, Boca Raton, FL, 2004.

[5] Dewilde, P., and Van der Veen, A. J., *Time Varying Systems and Computations*, Kluwer Academic Publishers, Boston, MA, 1998.

[6] Shokoohi, S., and Siverman, L. M., "Identification and Model Reduction of Time Varying Discrete Time Systems," *Automatica,* vol. 23, no. 4, 1987, pp. 509 - 521.
doi:10.1016/0005-1098(87)90080-X

[7] Verhaegen, M., and Yu, X., "A Class of Subspace Model Identification Algorithms to Identify Periodically and Arbitrarily Time Varying Systems," *Automatica,* vol. 31, no. 2, 1995, pp. 201 - 216.
doi:10.1016/0005-1098(94)00091-V

[8] Liu, K., "Identification of Linear Time Varying Systems," *Journal of Sound and Vibration,* vol. 206, no. 4, 1997, pp. 487 - 505.
doi:10.1006/jsvi.1997.1105

[9] Juang, J. N., and Pappa, R. S., "An Eigensystem Realization Algorithm (ERA) for Modal Parameter Identification and Model Reduction," *Journal of Guidance Control and Dynamics,* vol. 8, no. 5, 1985, pp. 620 - 627.
doi:10.2514/3.20031

[10] Juang, J. N., Phan, M., Horta, L., and Longman, R. W., “Identification of Observer/Kalman Filter Parameters: Theory and Experiments,” *Journal of Guidance Control and Dynamics,* vol. 16, no. 2, 1993, pp. 320 - 329. doi:10.2514/3.21006

[11] Schoukens, J., and Pintelton, R., *Identification of Linear Systems: A Practical Guideline to Accurate Modeling*, Pergamon Press, Englewood Cliffs, NJ, 1991.

[12] Soderstorm, T., and Stoica, P. G., *Instrumental Variable Methods for System Identification*, MIT Press, Cambridge, MA, 1983.

[13] Verhaegen, M., “Identification of time-varying state space models from input-output data,” in Workshop on Advanced Algorithms and their Realization, Bonas, France, 1991.

[14] Liu, K., “Extension of Modal Analysis to Linear Time Varying Systems,” *Journal of Sound and Vibration,* vol. 226, no. 1, 1999, pp. 149-167. doi:10.1006/jsvi.1999.2286

[15] Kailath, T., *Linear Systems*, Prentice Hall, Englewood Cliffs, NJ,, 1980.

[16] Gohberg, I., Kaashoek, M. A., and Kos, J., “Classification of Linear Time Varying Difference Equations under Kinematic Similarity,” *Integral Equations and Operator Theory,* vol. 25, 1996, pp. 445 - 480. doi:10.1007/BF01203027

[17] Schaub, H., and Junkins, J. L., *Analytical Mechanics of Space Systems*, American Institute of Aeronatics and Astronautics, Reston, VA, 2003.

[18] Chen, C. T., *Linear System Theory and Design*, Holt, Rinehart and Winston, New York, 1984.

[19] Battin, R. H., *An Introduction to the Mathematics and Methods of Astrodynamics*, American Institute of Aeronautics and Astronautics, Inc., Reston, VA, 1999.

[20] Majji, M., Juang, J. N., and Junkins, J. L, "Observer Markov Parameter Identification Theory for Time Varying Eigensystem Realization Algorithm," *Technical Reports, Center for Mechanics and Control*, Texas A&M University, College Station, 2008.

[21] Golub, G., and Van Loan, C. F., *Matrix Computations*, Johns Hopkins University Press, Baltimore, MD, 1996.

[22] Junkins, J. L., and Kim, Y., *Introduction to Dynamics and Control of Flexible Structures*, American Institute of Aeronautics and Astronautics, Washington, DC, 1991.

[23] Juang, J.-N., "Continuous-Time Bilinear System Identification," *Nonlinear Dynamics,* vol. 39, 2005, pp. 79-94.
doi:10.1007/s11071-005-1915-z

[24] Vidyasagar, M., *Nonlinear Systems Analysis*, 2 ed., Society of Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002.

[25] Phan, M. Q., Horta, L. G., Juang, J. N., and Longman, R. W., "Linear System Identification via an Asymptotically Stable Observer," *Journal of Optimization Theory and Applications,* vol. 79, no. 1, 1993, pp. 59 - 86.
doi:10.1007/BF00941887

[26] Majji, M., Juang, J. N., and Junkins, J. L., "Time Varying Eigensystem Realization Algorithm with Repeated Experiments," *In preparation for submission to the Journal of Guidance Control and Dynamics*, 2009.
doi: NA

[27] Majji, M., and Junkins, J. L., "Time Varying Eigensystem Realization Algorithm with Repeated Experiments," in AAS/AIAA Space Flight Mechanics Meeting, Galveston, TX, 2008, pp. 1069 - 1078 (Paper No. AAS 08-169).

[28] Silva, W., "Identification of Nonlinear Aeroelastic Systems Based on the Volterra Theory: Progress and Opportunities," *Nonlinear Dynamics,* vol. 39, 2005, pp. 25-62.
doi: 10.1007/s11071-005-1907-z

[29] Rugh, W. J., *Nonlinear System Theory: The Volterra Weiner Approach*, The Johns Hopkins University Press, Baltimore, MD, 1981.

[30] Sontag, E. D., Wang, Y., and Megretski, A.,, "Input Classes for Identification of Bilinear Systems," *IEEE Transactions on Automatic Control,* vol. Submitted, 2008.
doi:NA

[31] Van Loan, C. F., "Computing Integrals Involving the Matrix Exponential," *IEEE Transactions on Automatic Control,* vol. AC-23, no. 3, 1978, pp. 395 - 404.
doi:10.1109/TAC.1978.1101743

[32] Junkins, J. L., and Turner, J. D., *Optimal Spacecraft Rotational Maneuvers*, Elsevier, Amsterdam, 1986.

[33] Bruni, C., Dipillo, G., and Koch, G., "On the Mathematical Models of Bilinear Systems," *Ricerche di Automatica,* vol. 2, no. 1, 1971, pp. 11-26.
doi:NA

[34] Bryson, A. E., and Ho, Y. C., *Applied Optimal Control* Hemisphere Publishing, Washington, DC,, 1975.

[35] Bryson, A. E., *Control of Spacecraft and Aircraft*, Princeton University Press, Princeton, NJ, 1994.

[36] Juang, J. N., and Phan, M. Q., *Identification and Control of Mechanical Systems*, Cambridge University Press, Cambridge, U. K., 2001.

[37] Wie, B., *Space Vehicle Dynamics and Control*, American Institute of Aeronautics and Astronautics, Reston, VA, 1998.

[38] Junkins, J. L., "Adventures on the Interface of Dynamics and Control : von Karman Lecture," *Journal of Guidance Control and Dynamics,* vol. 20, no. 6, 1997, pp. 1058-1071.
doi:10.2514/2.4176

[39] Mohler, R. R., *Bilinear Control Processes: With Applications to Engineering, Ecology and Medicine*, Academic Press, Orlando, FL, 1973.

[40] Anderson, B. D. O., and Moore, J. B., *Optimal Filtering*, Prentice Hall, Englewood Cliffs, NJ, 1979.

[41] Mehra, R. K., "On the Identification of Variances and Adaptive Kalman Filtering," *IEEE Transactions on Automatic Control,* vol. AC-15, no. 2, 1970, pp. 175-184.
doi:10.1109/TAC.1970.1099422

[42] Phan, M. Q., Horta, L. G., Juang, J. N., and Longman, R. W., "Improvement of Observer/Kalman Filter Identification (OKID) by Residual Whitening," *Journal of Vibration and Acoustics,* vol. 117, no. 2, 1995, pp. 232-240.
doi:10.1115/1.2873927

[43] Minamide, N., Ohno, M., and Imanaka, H., "Recursive Solutions to Deadbeat Observers for Time Varying Multivariable Systems," *International Journal of Control,* vol. 45, no. 4, 1987, pp. 1311-1321.
doi:10.1080/00207178708933810

[44] Hostetter, G. H., *Digital Control System Design*, Holt, Rinehart and Winston, Inc., New York, 1988.

[45] Majji, M., Turner, J. D., and Junkins, J. L., "High Order Methods for Estimation of Dynamic Systems Part 1: Theory," in AAS - AIAA Spaceflight Mechanics Meeting, Galveston, TX, 2008, pp. (In Print).

[46] Kalman, R. E., and Bucy, R. S., "New Methods and Results in Linear Prediction and Filtering Theory," *Transactions of ASME: Journal of Basic Engineering,* vol. 83, no. D, 1961, pp. 95-108.
doi: NA

[47] Kailath, T., "An Innovations Approach to Least-Squares Estimation Part I: Linear Filtering in Additive White Noise," *IEEE Transactions on Automatic Control,* vol. AC-13, 1968, pp. 646-655.
doi:10.1109/TAC.1968.1099025

APPENDIX A

LINEAR ESTIMATORS OF THE KALMAN TYPE: A REVIEW OF THE STRUCTURE AND PROPERTIES

We review the structure and properties of the state estimators for linear discrete time varying dynamical systems (Kalman Filter Theory[4, 46]) using the innovations approach propounded by Kailath[47] and Mehra[41]. The most commonly used truth model for the linear time varying filtering problem is given by

$$\mathbf{x}_{k+1} = A_k \mathbf{x}_k + B_k \mathbf{u}_k + \Gamma_k \boldsymbol{\omega}_k \tag{A.1}$$

together with the measurement equations given by,

$$\mathbf{y}_k = C_k \mathbf{x}_k + D_k \mathbf{u}_k + \mathbf{v}_k \tag{A.2}$$

The process noise sequence is assumed to be a Gaussian random sequence with zero mean $E\left(\boldsymbol{\omega}_i\right) = \mathbf{0}, \forall i$ and a variance sequence $E\left(\boldsymbol{\omega}_i \boldsymbol{\omega}_j^T\right) = Q_i \,\delta_{ij}, \forall i, j$ having an uncorrelated profile in time (with itself, as shown by the variance expression) and no correlation with the measurement noise sequence $E\left(\boldsymbol{\omega}_i \mathbf{v}_j^T\right) = 0, \forall i, j$. Similarly, the measurement noise sequence is assumed to be a zero mean Gaussian random vector with covariance sequence given by $E\left(\mathbf{v}_i \mathbf{v}_j^T\right) = R_i \,\delta_{ij}$, where the Kronecker delta is denoted as $\delta_{ij} = 0, \forall i \neq j$ and $\delta_{ij} = 1 \;\forall i = j$ along with the usual notation $E(.)$ for the expectation operator of random vectors. A typical estimator of the Kalman type (optimal) assumes the structure (following the notations of [10]),

$$\begin{aligned}\hat{\mathbf{x}}_k^+ &= \hat{\mathbf{x}}_k^- + K_k\left[\mathbf{y}_k - \hat{\mathbf{y}}_k\right]\\ &:= \hat{\mathbf{x}}_k^- + K_k\boldsymbol{\varepsilon}_k\end{aligned} \tag{A.3}$$

where the term $\boldsymbol{\varepsilon}_k := \mathbf{y}_k - \hat{\mathbf{y}}_k$ represents the so called innovations process. In classical estimation theory, this innovations process is defined to represent the new information brought in to the estimator dynamics through the measurements made at each time instant. The state transition equations and the corresponding propagated measurements (most often used to compute the innovations process) of the estimator are given by,

$$\begin{aligned}&\hat{\mathbf{x}}_{k+1}^- = A_k\hat{\mathbf{x}}_k^+ + B_k\mathbf{u}_k\\ &\text{or}\\ &\hat{\mathbf{x}}_{k+1}^- = A_k\left[I - K_kC_k\right]\hat{\mathbf{x}}_k^- + B_k\mathbf{u}_k + A_kK_k\mathbf{y}_k\end{aligned} \tag{A.4}$$

and

$$\hat{\mathbf{y}}_k^- = C_k\hat{\mathbf{x}}_k^- + D_k\mathbf{u}_k \tag{A.5}$$

Defining the state estimation error to be given by, $\mathbf{e}_k := \mathbf{x}_k - \hat{\mathbf{x}}_k^-$ (for analysis purpose), the innovations process is related to the state estimation error as,

$$\boldsymbol{\varepsilon}_k = C_k\mathbf{e}_k + \mathbf{v}_k \tag{A.6}$$

while the propagation of the estimation error dynamics (estimator in the loop, similar to the time varying OKID developments of the paper) is governed by,

$$\begin{aligned}\mathbf{e}_{k+1} &= A_k\left[I - K_kC_k\right]\mathbf{e}_k - A_kK_k\mathbf{v}_k + \Gamma_k\boldsymbol{\omega}_k\\ &:= \tilde{A}_k\mathbf{e}_k - A_kK_k\mathbf{v}_k + \Gamma_k\boldsymbol{\omega}_k\end{aligned} \tag{A.7}$$

Defining the uncertainty associated by the state estimation process, quantified by the covariance to be $P_k := E\left[\mathbf{e}_k\mathbf{e}_k^T\right]$, covariance propagation equations are given by,

$$P_{k+1} = \tilde{A}_k P_k \tilde{A}_k^T + A_k K_k R_k K_k^T A_k^T + \Gamma_k Q_k \Gamma_k^T \tag{A.8}$$

Instead of the usual, minimum variance approach in developing the Kalman recursions for the discrete time varying linear estimator, let us use the orthogonality of the innovations process, a necessary condition for optimality and to obtain the Kalman filter recursions. This property is usually called the innovations property is the conceptual basis for projection methods[47] in a Hilbert space setting. As a consequence of this property we have the following condition.

If the gain in the observer gain is optimal, then the resulting recursions should render the innovations process orthogonal (uncorrelated) with respect to all other terms of the sequence. That is to say that for any time step t_i and a time step t_{i-k} (denoted as $i-k$), $k > 0$ steps behind the ith step, we have that

$$E\left[\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_{i-k}^T\right] = 0 \tag{A.9}$$

Using the definitions for the innovations process and the state estimation error, we use the relationship between them to arrive at the following expression for the necessary condition that,

$$E\left[\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_{i-k}^T\right] = C_i E\left[\mathbf{e}_i \mathbf{e}_{i-k}^T\right] C_{i-k}^T + C_i E\left[\mathbf{e}_i \mathbf{v}_{i-k}^T\right] = 0 \tag{A.10}$$

where the two terms $E\left[\mathbf{v}_i \mathbf{e}_{i-k}^T\right] = E\left[\mathbf{v}_i \mathbf{v}_{i-k}^T\right] = 0$ drop out because of the lack of correlation, in lieu of the standard assumptions of the Kalman filter theory. For the case of $k = 0$, it is easy to see that equation (A.10) becomes,

$$\begin{aligned} E\left[\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^T\right] &= C_i E\left[\mathbf{e}_i \mathbf{e}_i^T\right] C_i^T + E\left[\mathbf{v}_i \mathbf{v}_i^T\right] \\ &= C_i P_i C_i^T + R_i \end{aligned} \tag{A.11}$$

Applying the evolution equation for the estimation error dynamics for k time steps backward in time from t_i , we have that,

$$\begin{aligned}\mathbf{e}_i = \tilde{A}_{i-1}\tilde{A}_{i-2}...\tilde{A}_{i-k+1}\tilde{A}_{i-k}\mathbf{e}_{i-k} - \left[\tilde{A}_{i-1}...\tilde{A}_{i-k+1}A_{i-k}K_{i-k}\mathbf{v}_{i-k} + ... + \tilde{A}_{i-1}A_{i-2}K_{i-2}\mathbf{v}_{i-2} + A_{i-1}K_{i-1}\mathbf{v}_{i-1}\right] \\ + \left[\tilde{A}_{i-1}...\tilde{A}_{i-k+1}\Gamma_{i-k}\boldsymbol{\omega}_{i-k} + ... + \tilde{A}_{i-1}\Gamma_{i-2}\boldsymbol{\omega}_{i-2} + \Gamma_{i-1}\boldsymbol{\omega}_{i-1}\right]\end{aligned} \tag{A.12}$$

We obtain expressions for $E\left[\mathbf{e}_i\mathbf{e}_{i-k}^T\right], E\left[\mathbf{e}_i\mathbf{v}_{i-k}^T\right]$ by operating equation (A.12) on both sides with $\mathbf{e}_{i-k}^T, \mathbf{v}_{i-k}^T$ on both sides and taking the expectation operator.

$$\begin{aligned}E\left[\mathbf{e}_i\mathbf{e}_{i-k}^T\right] &= \tilde{A}_{i-1}\tilde{A}_{i-2}...\tilde{A}_{i-k+1}\tilde{A}_{i-k}E\left[\mathbf{e}_{i-k}\mathbf{e}_{i-k}^T\right] \\ &= \tilde{A}_{i-1}\tilde{A}_{i-2}...\tilde{A}_{i-k+1}\tilde{A}_{i-k}P_{i-k}\end{aligned} \tag{A.13}$$

$$\begin{aligned}E\left[\mathbf{e}_i\mathbf{v}_{i-k}^T\right] &= -\tilde{A}_{i-1}...\tilde{A}_{i-k+1}A_{i-k}K_{i-k}E\left(\mathbf{v}_{i-k}\mathbf{v}_{i-k}^T\right) \\ &= -\tilde{A}_{i-1}...\tilde{A}_{i-k+1}A_{i-k}K_{i-k}R_{i-k}\end{aligned} \tag{A.14}$$

Substituting equations (A.14) and (A.13) in to the expression for the inner product (A.10), we arrive at the expressions for Kalman gain sequence as a function of the statistics of the state estimation error dynamics for all time instances up to $t_{i-1}\left(i-1\right)$ as,

$$\begin{aligned}E\left[\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_{i-k}^T\right] &= C_i\,\tilde{A}_{i-1}\tilde{A}_{i-2}...\tilde{A}_{i-k+1}\tilde{A}_{i-k}P_{i-k}C_{i-k}^T - C_i\,\tilde{A}_{i-1}...\tilde{A}_{i-k+1}A_{i-k}K_{i-k}R_{i-k} \\ &= C_i\,\tilde{A}_{i-1}\tilde{A}_{i-2}...\tilde{A}_{i-k+1}\left[\tilde{A}_{i-k}P_{i-k}C_{i-k}^T - A_{i-k}K_{i-k}R_{i-k}\right] \\ &= C_i\,\tilde{A}_{i-1}\tilde{A}_{i-2}...\tilde{A}_{i-k+1}A_{i-k}\left[P_{i-k}C_{i-k}^T - K_{i-k}\left(R_{i-k} + C_{i-k}P_{i-k}C_{i-k}^T\right)\right] \\ &= 0\end{aligned} \tag{A.15}$$

which is necessary to hold for all Kalman type estimators with the familiar update structure, $\forall k > 0$

$$K_{i-k} = P_{i-k} C_{i-k}^{T} \left(R_{i-k} + C_{i-k} P_{i-k} C_{i-k}^{T} \right)^{-1} \tag{A.16}$$

because of the innovations property involved. Qualitative relationship between the identified observer realized from the time varying OKID calculations (GTV-ARX model) and the classical Kalman filter is explained in the main body of the paper using the innovations property of the optimal filter developed above. A minor detail pointed out at this stage is that the optimality in the sense of Kalman, viewed from the perspective of orthogonality conditions in this appendix does not interfere with the dead-beat conditions discussed elsewhere. Clearly appropriate choice of weights and tuning parameters of the Kalman filter can be chosen to approach the dead beat condition. The quantitative connections however need to be explored more rigorously at the present time, along with the large sample behavior in the stochastic setting.

APPENDIX B

A REVIEW OF THE VAN-LOAN METHOD FOR COMPUTING INTEGRALS INVOLVING A MATRIX EXPONENTIAL

We briefly review the application of "Van-Loan" integral formula (detailed and more general developments can be studied in[31]) to the evaluation of the integral involving matrix exponential for evaluating the input dependent coefficients in the matrix equation (4.40) for the determination of a least squares estimate of B_c. The integral in question is similar to the form given by (using the notations of [31] and employing the necessary change of variables and obvious redefinitions),

$$G_1\left(\Delta t\right) = \int_0^{\Delta t} e^{\left[\hat{A}_c + \sum_{i=1}^{r} \hat{N}_{ci} u_i\right](\Delta t - \tau)} d\tau \tag{B.1}$$

Consider an augmented block matrix $\Omega \in \mathbb{R}^{2n\times 2n}$ given by,

$$\Omega = \begin{bmatrix} \left[\hat{A}_c + \sum_{i=1}^{r} \hat{N}_{ci} u_i\right] & \mathbf{I}_n \\ 0_{n\times n} & 0_{n\times n} \end{bmatrix} \tag{B.2}$$

where $\mathbf{I}_n \in \mathbb{R}^{n\times n}$ identity matrix and $0_{n\times n} \in \mathbb{R}^{n\times n}$ matrix of zeros. Then the associated matrix differential equation,

$$\dot{X}\left(t\right) = \Omega\, X\left(t\right) \tag{B.3}$$

with initial conditions, $X\left(t_0\right) = \mathbf{I}_n$ has a block solution of the form given by,

$$X\left(t\right) = \begin{bmatrix} F_1\left(t\right) & G_1\left(t\right) \\ 0_{n\times n} & F_2\left(t\right) \end{bmatrix} = e^{\Omega\, t} \tag{B.4}$$

By substituting the solution (B.4) in to the matrix differential equation (B.3) and comparing block by block gives us three matrix differential equations that can be written as,

$$\dot{F}_1(t) = \left[\hat{A}_c + \sum_{i=1}^{r} \hat{N}_{ci} u_i\right] F_1(t) \tag{B.5}$$

with initial conditions $F_1(t_0) = \mathbf{I}_n$, and

$$\dot{G}_1(t) = \left[\hat{A}_c + \sum_{i=1}^{r} \hat{N}_{ci} u_i\right] G_1(t) + \mathbf{I}_n \tag{B.6}$$

with initial conditions $G_1(t_0) = 0_n$, where the (rather obvious) solution for the differential equation $\dot{F}_2(t) = 0_n$, $F_2(t) = F_2(t_0) = \mathbf{I}_n$ has been used. A cursory inspection reveals that (B.1) is indeed the solution of the differential equation (B.6) and is accurately evaluated as the upper - right $n \times n(1 \times 2)$ block of the matrix exponential solution $X(t)$ (i.e., $X(t)(1:n, n+1:2n)$ block) in equation (B.4). This procedure enables us to accurately compute the coefficients involved in equations (4.40) in the identification of the continuous time $\hat{B}_c$.

APPENDIX C

THE TIME VARYING DEAD BEAT CONDITION

It was found in chapter III that the generalization of the ARX model in the time varying case gives rise to an observer that could be set to a dead beat condition that has different properties and structure when compared to its linear time invariant counterpart. The topic of extension of the dead beat observer design to time varying systems has not been pursued aggressively in the literature and only scattered results exist in this context. Paper by Minamide et. al.[43], develops a similar definition of the time varying dead beat condition and present an algorithm to systematically assign the observer gain sequence to achieve the generalized condition thus derived. In contrast, through the definition of the time varying ARX model we arrive at this definition quite naturally and we further develop plant models and corresponding dead beat observer models directly from input output data, which is an elegant development of this dissertation.

First we recall the definition of a dead beat observer in case of the linear time invariant system and present a simple example to illustrate the central ideas. Following the conventions of Juang[1] and Kailath[15], if a linear discrete time dynamical system is characterized by the evolution equations given by,

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + B\mathbf{u}_k \tag{C.1}$$

along with the measurement equations (with an additional condition that (C, A) is an observable pair),

$$\mathbf{y}_k = C\mathbf{x}_k + D\mathbf{u}_k \tag{C.2}$$

where the usual assumptions on the dimensionality of the state space are made, $\mathbf{x}_k \in \mathbb{R}^n$, $\mathbf{y}_k \in \mathbb{R}^m$, $\mathbf{u}_k \in \mathbb{R}^r$ and C, A, B are matrices of compatible dimensions. Then the gain matrix G is said to produce a dead beat observer, if and only if the following condition is satisfied (the so-called dead beat condition):

$$(A+GC)^p = [0]_{n\times n} \tag{C.3}$$

where p is the smallest integer such that $m*p \geq n$ and $[0]_{n\times n}$ is an $n\times n$ matrix of zeros.

Example:

Let us consider the following simple linear time invariant example to fix the ideas.

$$\begin{aligned} A &= \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} \\ C &= \begin{bmatrix} 0 & 1 \end{bmatrix} \end{aligned} \tag{C.4}$$

Now the necessary and sufficient conditions for a dead beat observer design give rise to a gain matrix $G = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix}$ such that,

$$\begin{aligned} (A+GC)^2 &= \begin{bmatrix} 1+g_1 & g_1\left(3+g_2\right) \\ 3+g_2 & g_1 + \left(2+g_2\right)^2 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \end{aligned} \tag{C.5}$$

giving rise to the gain matrix $G = \begin{bmatrix} -1 \\ -3 \end{bmatrix}$ (it is easy to see that $p = 2$ for this problem).

The closed loop can be verified to be given by,

$$\left(A+GC\right)=\begin{bmatrix}1 & -1\\ 1 & -1\end{bmatrix} \tag{C.6}$$

which can be verified to be a singular, defective (repeated roots at the origin) and nilpotent matrix. Therefore the deadbeat observer is the fastest observer that could possibly be achieved since in the time invariant case, it designs the observer feedback such that the closed loop poles are placed at the origin. However, it is quite interesting to note that the necessary conditions, albeit redundant nonlinear functions in fact have a solution that exists (one typically does not have to resort to least squares solutions) since some of the conditions are dependent on each other (not necessarily linear dependence). This nonlinear structure of the necessary conditions to realize a dead beat observer makes the problem interesting and several techniques are available to compute solutions in the time invariant case, for both cases when plant models are available (Minamide solution [43]) and when only experimental data is available (OKID solution).

Now considering the time varying system and following the notation of chapter III, the time varying dead beat definition is made. Recall (from equation (3.15)) that in constructing the generalized time varying ARX (GTV-ARX) model of chapter III, we have already used this definition.

A linear time varying discrete time observer is said to be dead beat, if, there exists a gain sequence G_k such that

$$\left(A_{k+p-1}+G_{k+p-1}C_{k+p-1}\right)\left(A_{k+p-2}+G_{k+p-1}C_{k+p-1}\right)\ldots\left(A_k+G_kC_k\right)=\left[0\right]_{n\times n} \tag{C.7}$$

for every k , where p is the smallest integer such that the condition $p*m\geq n$ is satisfied.

Example

To fix the ideas, we demonstrate the observer realized on the same problem used in the chapter III and follow the example by a short discussion on the nature and properties of the time varying dead beat condition in case of the observer design. The parameters involved in the example problem are given by the equation,

as

$$
\begin{aligned}
A_k &= \exp\left[A_c * \Delta t\right] \\
B_k &= \begin{bmatrix} 1 & 0 \\ 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad C_k = \begin{bmatrix} 1 & 0 & 1 & 0.2 \\ 1 & -1 & 0 & -0.5 \end{bmatrix}, \\
D_k &= 0.1\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\end{aligned}
\tag{C.8}
$$

where the matrix is given by

$$
A_c = \begin{bmatrix} 0_{2\times2} & I_{2\times2} \\ -K_t & 0_{2\times2} \end{bmatrix}
\tag{C.9}
$$

with $K_t = \begin{bmatrix} 4+3\tau_k & -1 \\ -1 & 7+3\tau'_k \end{bmatrix}$ and τ_k, τ'_k are defined as $\tau_k = \sin\left(10t_k\right), \quad \tau'_k := \cos\left(10t_k\right)$.

Clearly since $m = 2, n = 4$ for the example, the choice of $p = 2$ is made. Considering the time step $k = 36$, for demonstration purposes, the closed loop (with the observer gain equation in the output feedback style is given by)

$$A_{36}+G_{36}C_{36}=\begin{bmatrix} \text{-2.0405} & 0.3357 & 0.0016 & 0.5965 \\ \text{-1.7735} & \text{-0.7681} & \text{-3.2887} & \text{-0.0289} \\ 1.7902 & \text{-0.0270} & 0.8290 & \text{-0.3852} \\ \text{-6.9208} & 1.4773 & 1.0572 & 2.1980 \end{bmatrix}$$
$$\lambda\left(A_{36}+G_{36}C_{36}\right)=\begin{bmatrix} 0.31545 \\ \text{-0.097074} \\ 1.1878\times10^{-15} \\ 1.2252\times10^{-13} \end{bmatrix} \tag{C.10}$$

while the closed loop for the previous time step is calculated as,

$$A_{35}+G_{35}C_{35}=\begin{bmatrix} \text{-1.7924} & 0.4678 & 0.1630 & 0.5778 \\ \text{-0.7301} & -0.4380 & \text{-2.8865} & \text{-0.1330} \\ 1.1874 & \text{-0.1662} & 0.5671 & \text{-0.2986} \\ \text{-5.7243} & 1.8475 & 2.1805 & 2.0524 \end{bmatrix}$$
$$\lambda\left(A_{35}+G_{35}C_{35}\right)=\begin{bmatrix} 0.43716 \\ \text{-0.048167} \\ \left(\text{-2.1173 +7.4549i}\right)\times10^{-14} \\ \left(\text{-2.1173 -7.4549i}\right)\times10^{-14} \end{bmatrix} \tag{C.11}$$

and the closed loop for the consecutive time step is found to be given by,

$$A_{37}+G_{37}C_{37}=\begin{bmatrix} \text{-2.4701} & 0.1432 & \text{-0.2323} & 0.6315 \\ \text{-2.3403} & \text{-0.8353} & \text{-3.3551} & 0.0362 \\ 2.0767 & 0.0773 & 0.8335 & \text{-0.4165} \\ \text{-8.8651} & 0.6963 & \text{-0.2452} & 2.3719 \end{bmatrix}$$
$$\lambda\left(A_{37}+G_{37}C_{37}\right)=\begin{bmatrix} \text{-0.14861} \\ 0.048661 \\ 4.0371\times10^{-12} \\ \text{-5.5501}\times10^{-15} \end{bmatrix} \tag{C.12}$$

While clearly each of the closed loop member sequence, $\bar{A}_{35,36,37}$ has only two zero eigenvalues (individually non-deadbeat in the time invariant sense, since all closed loop poles are NOT placed at the origin), let us now consider the product matrices,

$$\left(A_{37}+G_{37}C_{37}\right)\left(A_{36}+G_{36}C_{36}\right)=10^{-12}\times\begin{bmatrix} -0.0959 & 0.0070 & -0.0326 & 0.0238 \\ -0.1192 & 0.0035 & -0.0187 & 0.0235 \\ -0.0564 & 0.0003 & -0.0307 & 0.0123 \\ 0.1137 & 0.0075 & 0.0551 & -0.0187 \end{bmatrix} \tag{C.13}$$

and

$$\left(A_{36}+G_{36}C_{36}\right)\left(A_{35}+G_{35}C_{35}\right)=10^{-13}\times\begin{bmatrix} -0.0844 & -0.1443 & 0.0888 & -0.0711 \\ 0.4660 & -0.2783 & 0.4528 & -0.2652 \\ -0.2265 & 0.1987 & -0.2076 & 0.1610 \\ -0.6217 & -0.4086 & 0.1243 & -0.1332 \end{bmatrix} \tag{C.14}$$

The examples clearly indicate that the composite transition matrices taken p (= 2 for this example) at a time can form a null matrix, while still retaining nonzero eigenvalues individually. This is the generalization that occurs in the definition of dead-beat condition in case of the time varying systems. Similar to the case of time invariant systems, we still see that the observer which is dead beat happens to be the fastest observer even in the case of the time varying systems.

We reiterate the fact that in case of the computations and algorithms of this dissertation, the dead-beat observer can be realized naturally along with the plant model sequence being identified. It is not difficult to construct the generalized ARX (GTV-ARX) model and derive the observer gain sequence using the time varying OKID procedure for the case when plant parameters are known. It is of consequence to observe that the procedure due to time varying OKID is developed directly in the reduced dimensional input–output space while the schemes developed to compute the gain sequences in the paper by Minamide et al. [43], which is quite similar to the method outlined by Hostetter[44] are based on projections of the state space on to the outputs.

VITA

Name:	Manoranjan Majji
Address:	616-D, HRBB, Aerospace Engineering Department, 3141 –TAMU, Texas A&M University, College Station, TX 77843-3141.
Email Address:	majji@tamu.edu, majji.neuron@gmail.com

Education:	B.E. (honors), Mechanical Engineering, Birla Institute of Technology and Science, 2003
	M.S., Aerospace Engineering, Texas A&M University, 2006
	Ph D., Aerospace Engineering, Texas A&M University, 2009