Observer/Kalman Filter Time Varying System Identification

Manoranjan Majji[1]
Texas A&M University, College Station, Texas, USA

Jer-Nan Juang[2]
National Cheng Kung University, Tainan, Taiwan

and
John L. Junkins[3]
Texas A&M University, College Station, Texas, USA

An algorithm for computation of the generalized Markov parameters of an observer or Kalman filter for discrete time varying systems from input-output experimental data is presented. Relationships between the generalized observer Markov parameters and the system Markov parameters are derived for the time varying case. A systematic procedure to compute the time varying sequence of system Markov parameters and the time varying observer (or Kalman) gain Markov parameter sequence from the generalized time varying observer Markov parameters is presented. This procedure is shown to be a time varying generalization of the recursive relations developed for the time invariant case using an Autoregressive model with an exogenous input (ARX model – in a procedure known as Observer/Kalman filter identification, OKID). These generalized time varying input-output relations with the time varying observer in the loop are referred to in the paper as the Generalized Time Varying Autoregressive model with an exogenous input (GTV-ARX model). The generalized system Markov parameters thus derived are used by the Time Varying Eigensystem Realization Algorithm (TVERA) developed by the authors, to obtain a time varying discrete time state space model. Qualitative relationship of the time varying observer with the Kalman observer in the stochastic environment and an asymptotically

[1] Post Doctoral Research Associate, Aerospace Engineering Department, 616 – D, 3141 TAMU, College Station, Texas, 77843-3141, majji@tamu.edu, Student Member, AIAA.
[2] Professor, National Cheng Kung University, Tainan, Taiwan; Former President, National Applied Research Laboratories, Taipei, Taiwan; Adjunct Professor, Aerospace Engineering Department, 3141 TAMU, College Station, Texas, 77843-3141, jjuang@cox.net, Fellow, AIAA.
[3] Distinguished Professor, Regents Professor, Royce E. Wisenbaker Chair, Aerospace Engineering Department, 722 B, 3141 TAMU, College Station, Texas, 77843-3141, junkins@aeromail.tamu.edu, Fellow, AIAA.

stable realized observer are discussed briefly to develop insights for the analyst. The 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. The time varying observer gains realized in the process are subsequently shown to be in consistent coordinate systems for closed loop state propagation. It is also demonstrated that the observer gain sequence realized in case of the minimum number of experiments corresponds naturally to a time varying deadbeat observer. Numerical examples demonstrate the utility of the concepts developed in the paper.

I. Introduction

SYSTEM identification has emerged as an important topic of research over the past few decades owing to the advancements of model based modern guidance, navigation and control. Eigensystem Realization Algorithm[1] (ERA) is widely acknowledged as a key contribution from the aerospace engineering community to this dynamic research topic. The system identification methods for time invariant systems have seen efforts from various researchers. The methods are now well understood for continuous and discrete time systems including the relationships between the continuous and discrete time system models.

On the other hand, discrete time varying system identification methods are comparatively poorly understood. Several past efforts by researchers have documented the developments in the identification of discrete time varying models. Cho et al.[2] explored the displacement structure in the Hankel matrices to obtain time invariant models from instantaneous input-output data. Shokoohi and Silverman [3] and Dewilde and Van der Veen[4], generalized several concepts of the classical linear time invariant system theory to include the time varying effects. Verhaegen and coworkers [5, 6] subsequently introduced the idea of repeated experiments (termed ensemble i/o data), enabling further research in the development of methods for identification of time varying systems. Liu [7] developed a methodology for developing time varying model sequences 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[8]. An important concept of *kinematic similarity* among linear discrete time varying system models concerns certain time varying transformations involved in the state transition matrices. Gohberg et al.[9]

discuss fundamental developments of this theory using a difference equation operator theoretic approach. In companion papers, Majji et al.[10, 11] extend the ERA, a classical algorithm for system identification of linear time invariant systems to realize discrete time varying models from input-output data following the framework and conventions of the above papers. The time varying eigensystem realization algorithm (TVERA) presented in the companion papers [10, 11] uses the generalized Markov parameters to realize time varying system descriptions by manipulations on Hankel matrix sequences of finite size. The realized discrete time varying models are shown to be in time varying coordinate systems and a method is outlined to transform all the time varying models to a single (observable or controllable subspace) coordinate system at a given time step.

However the algorithm developed there-in requires the determination of the generalized Markov parameters from sets of input-output experimental data. Therefore we need a practical method to calculate them without resorting to a high dimensioned calculation. This calculation becomes further compounded in systems where stability of the origin cannot be ascertained, since the number of potentially significant generalized Markov parameters grows rapidly. In other words, 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) pulse 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. In other words, the number of repeated experiments is proportional to the resolution of the model sequence desired by the analyst. This computational challenge has been among the main reasons for the lack of ready-adoption of the time varying system identification methods.

In this paper, we use an asymptotically stable 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 (TOKID). In addition, the tools systematically presented in this paper give an estimate on the minimum number of experiments one needs to perform for identification and/or recovery of all the Markov parameters of interest until that time instant. Thus, the central result of the current is to make the number of repeated experiments independent of the desired resolution of the model. 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.

Novel models relating input-output data are developed in this paper and are found to be elegant extensions of the ARX models well known in the analysis of time invariant models (cf. Juang et al., [12]). 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. This analogy is shown to go even further and enable us to realize a deadbeat observer gain sequence for time varying systems. The generalization of this deadbeat definition is rather unique and general for the time varying systems as it is shown that not all the closed loop time varying eigenvalues need to be zero for the time varying observer gain sequence to be called dead beat. Further, it is demonstrated that the time varying observer sequence (deadbeat or otherwise) computed from the GTV-ARX model is realized in a compatible coordinate system with the identified plant model sequence. Relations with the time varying Kalman filter are made comparing features of the parameters of the Kalman filter gains with the time varying observer gains realized from the generalized OKID procedure presented in the paper.

II. 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, refer to a companion paper [10] based on [11] and the references therein). The fundamental difference equations governing the evolution of a linear system in discrete time are given by

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

together with the measurement equations

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

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 of the state evolution (the linear time varying discrete time difference equation solution) is given by

$$\mathbf{x}_k = \Phi\left(k, k_0\right)\mathbf{x}_0 + \sum_{i=k_0}^{k-1} \Phi\left(k, i+1\right) B_i \mathbf{u}_i \tag{3}$$

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

$$\Phi(k,k_0)=\begin{cases} A_{k-1}A_{k-2}...A_{k_0}, & \forall k>k_0 \\ I, & k=k_0 \\ \text{undefined}, & \forall k<k_0 \end{cases} \tag{4}$$

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

$$\mathbf{y}_k = C_k\Phi(k,0)\mathbf{x}_0 + \sum_{i=0}^{k-1} C_k\Phi(k,i+1)B_i\mathbf{u}_i + D_k\mathbf{u}_k \tag{5}$$

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

$$\mathbf{y}_k = C_k\Phi(k,0)\mathbf{x}_0 + \sum_{i=0}^{k-1} h_{k,i}\mathbf{u}_i + D_k\mathbf{u}_k \tag{6}$$

where the generalized Markov parameters are defined as

$$h_{k,i}=\begin{cases} C_k\Phi(k,i+1)B_i, & \forall i<k-1 \\ C_kB_{k-1}, & i=k-1 \\ 0, & \forall i>k-1 \end{cases} \tag{7}$$

From now on, we try to 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

$$\mathbf{y}_k = C_kA_{k-1}...A_{k_0+1}A_{k_0}\mathbf{x}_{k_0} + \begin{bmatrix} D_k & C_kB_{k-1} & \cdots & C_kA_{k-1}...A_{k_0+1}B_{k_0} \end{bmatrix}\begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k-1} \\ \vdots \\ \mathbf{u}_{k_0} \end{bmatrix} \tag{8}$$

where k_0 can denote any general time step prior to k (in particular let us assume that it denotes the initial time such that $k_0=0$). As was pointed out in the companion paper, such a relationship between the input and output leads to a problem that increases by $m*r$ parameters for every time step considered. Thus it becomes difficult to compute the increasing number of unknown parameters. In the special case of systems whose open loop is asymptotically stable, this is not a problem. However, frequently, one tries to use identification in problems which do not have a stable origin for control and estimation purposes. In such problems, the analyst may be required to compute time varying model sequences with higher resolution. Hence we need to explore alternative methods in which plant parameter models can be realized from input-output data. A viable alternative to this problem useful to the practicing engineer is developed in the following section.

The central assumption involved in the developments of this paper is that (in order to obtain the generalized system and observer gain 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. The more general case deals with the presence of initial condition response in the output data. In the physical situation of unknown initial conditions, this problem is compounded and the separation of zero input response from the output data becomes an involved procedure. We do not discuss this general situation in the present paper. Most importantly since the connections between time varying ARX model and the state space model, and a discussion on the associated observer are complicated by themselves, we proceed with the presentation of the algorithm under the assumption that each experiment can be performed with zero initial conditions.

III. Input Output Representations: Observer Markov Parameters

The input-output representations for the time varying systems are quite similar to the input output model estimation of a lightly damped flexible spacecraft structure in the time invariant case. In the identification problem involving a lightly damped structure, one has to track a large number of Markov parameters to obtain reasonable accuracy in computation of the modal parameters involved. An effective method for "compressing" experimental input-output data, called observer/Kalman filter Markov parameter identification theory (OKID) was developed by Juang et al. [1, 12, 13]. 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 systems. 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 *a priori*.

Following the observations of the previous researchers, consider the use of a time varying "output – feedback" style gain sequence in the difference equation model Eq. (1) governing the linear plant, given by

$$
\begin{aligned}
\mathbf{x}_{k+1} &= A_k \mathbf{x}_k + B_k \mathbf{u}_k + G_k \mathbf{y}_k - G_k \mathbf{y}_k \\
&= \left(A_k + G_k C_k\right)\mathbf{x}_k + \left(B_k + G_k D_k\right)\mathbf{u}_k - G_k \mathbf{y}_k \\
&= \bar{A}_k \mathbf{x}_k + \left[\left(B_k + G_k D_k\right) \ \vdots \ -G_k\right]\begin{bmatrix}\mathbf{u}_k \\ \mathbf{y}_k\end{bmatrix} \\
&= \bar{A}_k \mathbf{x}_k + \bar{B}_k \mathbf{v}_k
\end{aligned} \tag{9}
$$

with the definitions

$$
\begin{aligned}
\bar{A}_k &= A_k + G_k C_k \\
\bar{B}_k &= \begin{bmatrix} B_k + G_k D_k & -G_k \end{bmatrix} \\
\mathbf{v}_k &= \begin{bmatrix}\mathbf{u}_k \\ \mathbf{y}_k\end{bmatrix}
\end{aligned} \tag{10}
$$

and no change in the measurement equations at the time step t_k

$$
\mathbf{y}_k = C_k \mathbf{x}_k + D_k \mathbf{u}_k \tag{11}
$$

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

$$
\begin{aligned}
\mathbf{y}_{k_0} &= C_{k_0}\mathbf{x}_{k_0} + D_{k_0}\mathbf{u}_{k_0} \\
\mathbf{y}_{k_0+1} &= C_{k_0+1}\bar{A}_{k_0}\mathbf{x}_{k_0} + D_{k_0+1}\mathbf{u}_{k_0+1} + C_{k_0+1}\bar{B}_{k_0}\mathbf{v}_{k_0} \\
&= C_{k_0+1}\bar{A}_{k_0}\mathbf{x}_{k_0} + D_{k_0}\mathbf{u}_{k_0} + \bar{h}_{k_0+1,k_0}\mathbf{v}_{k_0} \\
\mathbf{y}_{k_0+2} &= C_{k_0+2}\bar{A}_{k_0+1}\bar{A}_{k_0}\mathbf{x}_{k_0} + D_{k_0+2}\mathbf{u}_{k_0+2} + C_{k_0+2}\bar{B}_{k_0+1}\mathbf{v}_{k_0+1} + C_{k_0+2}\bar{A}_{k_0+1}\bar{B}_{k_0}\mathbf{v}_{k_0} \\
&= C_{k_0+2}\bar{A}_{k_0+1}\bar{A}_{k_0}\mathbf{x}_{k_0} + D_{k_0+2}\mathbf{u}_{k_0+2} + \bar{h}_{k_0+2,k_0+1}\mathbf{v}_{k_0+1} + \bar{h}_{k_0+2,k_0}\mathbf{v}_{k_0} \\
&\ldots
\end{aligned} \tag{12}
$$

with the definition of generalized observer Markov parameters

$$
\bar{h}_{k,i} = \begin{cases} C_k \bar{A}_{k-1}\bar{A}_{k-2}\ldots\bar{A}_{i+1}\bar{B}_i\,, & \forall k > i+1 \\ C_k \bar{B}_{k-1}, & k = i+1 \\ 0, & \forall k < i+1 \end{cases} \tag{13}
$$

we arrive at the general input-output relationship

$$
\mathbf{y}_k = C_k \bar{A}_{k-1}\ldots\bar{A}_{k_0}\mathbf{x}_{k_0} + D_k \mathbf{u}_k + \sum_{j=1}^{k-k_0-1}\bar{h}_{k,k-j}\mathbf{v}_{k-j} \tag{14}
$$

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

$$
\begin{aligned}
\bar{h}_{k,k-j} &= C_k \bar{A}_{k-1} \ldots \bar{A}_{k-j+1} \bar{B}_{k-j} \\
&= \left[C_k \bar{A}_{k-1} \ldots \bar{A}_{k-j+1} \left(B_{k-j} + G_{k-j} C_{k-j} \right) \quad -C_k \bar{A}_{k-1} \ldots \bar{A}_{k-j+1} G_{k-j} \right] \\
&= \left[\bar{h}_{k,k-j}^{(1)} \quad -\bar{h}_{k,k-j}^{(2)} \right]
\end{aligned} \tag{15}
$$

where the partitions $\bar{h}_{k,k-j}^{(1)}, \bar{h}_{k,k-j}^{(2)}$ are used in the calculations of the generalized system Markov parameters and the time varying observer gain sequence in the subsequent developments of the paper. The closed loop thus constructed, is now forced to have an asymptotically stable origin by the observer design process. The goal of an observer constructed in this fashion is to enforce certain desirable (stabilizing) characteristics into the closed loop (e.g., deadbeat-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 $\left\{ \mathbf{v}_{k+j-1} \right\}_{j=1}^{p_k}, \mathbf{u}_{k+p_k}$ and independent of the state at every time step t_k. Therefore by writing

$$
\begin{aligned}
\mathbf{y}_{k+p_k} &= C_{k+p_k} \bar{A}_{k+p_k-1} \ldots \bar{A}_k \mathbf{x}_k + D_{k+p_k} \mathbf{u}_{k+p_k} + \sum_{j=1}^{p_k} \bar{h}_{k+p_k,k+j-1} \mathbf{v}_{k+j-1} \\
&\approx D_{k+p_k} \mathbf{u}_{k+p_k} + \sum_{j=1}^{p_k} \bar{h}_{k+p_k,k+j-1} \mathbf{v}_{k+j-1}
\end{aligned} \tag{16}
$$

we have set $C_{k+p_k} \bar{A}_{k+p_k-1} \ldots \bar{A}_k \mathbf{x}_k \approx \mathbf{0}$ (with exact equality assignable i.e., $C_{k+p} \bar{A}_{k+p-1} \ldots \bar{A}_k \mathbf{x}_k = \mathbf{0}$, in the absence of measurement noise $\forall k = 0,1,\ldots,k_f$). This leads to the construction of a generalized time varying autoregressive with exogenous input (GTV-ARX) model at every time step. Note that the order p_k of the GTV-ARX model can also change with time (we coin the term "generalized" to describe this variability in the order). This variation and complexity provides a large number of observer gains at the disposal of the analyst under the time varying OKID framework. In using this input-output relationship (Eq.(16)) instead of the exact relationship given in Eq.(8), we introduce damping into the closed loop. For simplicity and ease in implementation and understanding, we set the generally variable order to remain fixed and minimum (time varying deadbeat) 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 deadbeat observer at every time, providing ease in calculations by requiring minimum number of repeated experiments. The deadbeat conditions are different in the case of time varying systems, due to the transition matrix product conditions (Eq. (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 observer system matrix give sufficient conditions to place all the closed loop system poles at the origin (deadbeat). The nature and properties of the time varying deadbeat condition are briefly summarized in the Appendix B, along with an example problem. Considerations of the time varying deadbeat condition appear sparse (Minamide et al.,[14] and Hostetter[15] present some fundamental results on the design of time varying deadbeat observers), if not completely heretofore unknown in modern literature. Therefore the connections made here-in especially in the context of system identification are quite unique in nature.

If the repeated experiments (as derived and presented in [10, 11]) are performed so as to compute a least-squares solution to the input-output behavior conjectured in Eq.(16), we have identified the system (together with the observer-in-the-loop) such that the output y_{k+p} does not depend on the state x_k. Stating the same in a vector – matrix form, for any time step t_k (denoted by k and $\forall k > p$) we have that

$$\mathbf{y}_k = \begin{bmatrix} D_k & \bar{h}_{k,k-1} & \bar{h}_{k,k-2} & \cdots & \bar{h}_{k,k-p} \end{bmatrix} \begin{bmatrix} \mathbf{u}_k \\ \mathbf{v}_{k-1} \\ \mathbf{v}_{k-2} \\ \vdots \\ \mathbf{v}_{k-p} \end{bmatrix} \tag{17}$$

This represents a set of m equations in $m \times \left(r + p * \left(r + m\right)\right)$ unknowns. In contrast to the developments using the generalized system Markov parameters, (to relate the input-output data sets; refer Eq. (8) in the companion paper [10, 11] 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 constant (derived below) and does not change with the discrete time step t_k (resolution of the model sequence desired by the analyst). This is an important result of the current paper. In fact, it is observed that a minimum of $N_{\exp}^{\min} = \left(r + p_{\min} * \left(r + m\right)\right)$ experiments are necessary to determine the observer Markov parameters uniquely. From the developments of the subsequent sections, this is the minimum number of repeated experiments one should perform in order to realize the time varying system models desired from the TVERA. Equations (17) with N repeated experiments yields

$$
\begin{aligned}
\mathbf{Y}_k &= \begin{bmatrix} \mathbf{y}_k^{(1)} & \mathbf{y}_k^{(2)} & \cdots & \mathbf{y}_k^{(N)} \end{bmatrix} \\
&= \begin{bmatrix} D_k & \bar{h}_{k,k-1} & \bar{h}_{k,k-2} & \cdots & \bar{h}_{k,k-p} \end{bmatrix} \begin{bmatrix} \mathbf{u}_k^{(1)} & \mathbf{u}_k^{(2)} & & \mathbf{u}_k^{(N)} \\ \mathbf{v}_{k-1}^{(1)} & \mathbf{v}_{k-1}^{(2)} & & \mathbf{v}_{k-1}^{(N)} \\ \mathbf{v}_{k-2}^{(1)} & \mathbf{v}_{k-2}^{(2)} & \cdots & \mathbf{v}_{k-2}^{(N)} \\ \vdots & \vdots & & \vdots \\ \mathbf{v}_{k-p}^{(1)} & \mathbf{v}_{k-p}^{(2)} & & \mathbf{v}_{k-p}^{(N)} \end{bmatrix} \\
&= \mathbf{M}_k \mathbf{V}_k
\end{aligned}
\tag{18}
$$

$\forall k > p$.

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

$$
\hat{\mathbf{M}}_k = \mathbf{Y}_k \mathbf{V}_k^{\dagger} \tag{19}
$$

where $(.)^{\dagger}$ denotes the pseudo inverse of a matrix [16, 17]. The calculation of the system Markov parameters and observer gain Markov parameters is detailed in the next section.

IV. 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.

A. Computation of System Markov Parameters from Observer Markov Parameters

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

$$
\begin{aligned}
\bar{h}_{k,k-1} &= C_k \bar{B}_{k-1} \\
&= C_k \begin{bmatrix} \left(B_{k-1} + G_{k-1} D_{k-1} \right) & -G_{k-1} \end{bmatrix} \\
&= \begin{bmatrix} h_{k,k-1}^{(1)} & -h_{k,k-1}^{(2)} \end{bmatrix}
\end{aligned}
\tag{20}
$$

where the superscripts (1) and (2) are 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

$$
\begin{aligned}
\bar{h}_{k,k-1}^{(1)} - \bar{h}_{k,k-1}^{(2)} D_{k-1} &= C_k B_{k-1} \\
&= h_{k,k-1}
\end{aligned}
\tag{21}
$$

where the unadorned $h_{i,j}$ are used to denote the generalized system Markov parameters, following the conventions and notations set up in the companion papers[10, 11]. A similar expression for Markov parameters with two time steps between them yields

$$\begin{aligned}\overline{h}^{(1)}_{k,k-2} - \overline{h}^{(2)}_{k,k-2} D_{k-2} &= C_k \overline{A}_{k-1} \overline{B}_{k-2} - C_k \overline{A}_{k-1} G_{k-2} D_{k-2} \\ &= C_k \overline{A}_{k-1} \left(B_{k-2} + G_{k-2} D_{k-2} \right) - C_k \overline{A}_{k-1} G_{k-2} D_{k-2} \\ &= C_k \overline{A}_{k-1} B_{k-2} \\ &= C_k \left(A_{k-1} + G_{k-1} C_{k-1} \right) B_{k-2} \\ &= C_k A_{k-1} B_{k-2} + \overline{h}^{(2)}_{k,k-1} C_{k-1} B_{k-2} \\ &= h_{k,k-2} + \overline{h}^{(2)}_{k,k-1} h_{k-1,k-2}\end{aligned} \tag{22}$$

This elegant manipulation leads to an expression for the generalized 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 recursive relationship was found to hold in general and enables the calculation of the system Markov parameters from the observer Markov parameters $\overline{h}^{(1)}_{i,j}, \overline{h}^{(2)}_{i,j}$.

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

$$\begin{aligned}\overline{h}^{(1)}_{k,k-p} - \overline{h}^{(2)}_{k,k-p} D_{k-p} &= C_k \overline{A}_{k-1} \overline{A}_{k-2} \ldots \overline{A}_{k-p+1} \left(B_{k-p} + G_{k-p} D_{k-p} \right) - C_k \overline{A}_{k-1} \overline{A}_{k-2} \ldots \overline{A}_{k-p+1} G_{k-p} D_{k-p} \\ &= C_k \overline{A}_{k-1} \overline{A}_{k-2} \ldots \overline{A}_{k-p+1} B_{k-p} \\ &= C_k \overline{A}_{k-1} \overline{A}_{k-2} \ldots \overline{A}_{k-p+2} \left(A_{k-p+1} + G_{k-p+1} C_{k-p+1} \right) B_{k-p} \\ &= C_k \overline{A}_{k-1} \overline{A}_{k-2} \ldots \overline{A}_{k-p+2} A_{k-p+1} B_{k-p} + C_k \overline{A}_{k-1} \overline{A}_{k-2} \ldots \overline{A}_{k-p+2} G_{k-p+1} C_{k-p+1} B_{k-p} \\ &= C_k \overline{A}_{k-1} \overline{A}_{k-2} \ldots \overline{A}_{k-p+2} A_{k-p+1} B_{k-p} + \overline{h}^{(2)}_{k,k-p+1} h_{k-p+1,k-p}\end{aligned} \tag{23}$$

Careful examination reveals 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

$$\begin{aligned}C_k \overline{A}_{k-1} \overline{A}_{k-2} \ldots \overline{A}_{k-p+2} A_{k-p+1} B_{k-p} &= C_k \overline{A}_{k-1} \ldots \overline{A}_{k-p+3} \left(A_{k-p+2} + G_{k-p+2} C_{k-p+2} \right) A_{k-p+1} B_{k-p} \\ &= C_k \overline{A}_{k-1} \ldots A_{k-p+2} A_{k-p+1} B_{k-p} + C_k \overline{A}_{k-1} \ldots G_{k-p+2} C_{k-p+2} A_{k-p+1} B_{k-p} \\ &= C_k \overline{A}_{k-1} \ldots A_{k-p+2} A_{k-p+1} B_{k-p} + \overline{h}^{(2)}_{k,k-p+2} h_{k-p+2,k-p} \\ &= \ldots \\ &= C_k A_{k-1} \ldots A_{k-p+1} B_{k-p} + \overline{h}^{(2)}_{k,k-1} h_{k-1,k-p} + \overline{h}^{(2)}_{k,k-2} h_{k-2,k-p} + \ldots + \overline{h}^{(2)}_{k,k-p+2} h_{k-p+2,k-p} \\ &= h_{k,k-p} + \overline{h}^{(2)}_{k,k-1} h_{k-1,k-p} + \overline{h}^{(2)}_{k,k-2} h_{k-2,k-p} + \ldots + \overline{h}^{(2)}_{k,k-p+2} h_{k-p+2,k-p}\end{aligned} \tag{24}$$

This manipulation enables us to write

$$\begin{aligned}\bar{h}^{(1)}_{k,k-p} - \bar{h}^{(2)}_{k,k-p} D_{k-p} &= h_{k,k-p} + \bar{h}^{(2)}_{k,k-1} h_{k-1,k-p} + \bar{h}^{(2)}_{k,k-2} h_{k-2,k-p} + ... + \bar{h}^{(2)}_{k,k-p+1} h_{k-p+1,k-p} \\ &= h_{k,k-p} + \sum_{j=1}^{p-1} \bar{h}^{(2)}_{k,k-j} h_{k-j,k-p}\end{aligned} \tag{25}$$

Writing the derived relationships between the system and observer Markov parameters yields the following set of equations

$$\begin{aligned}&h_{k,k-1} = \bar{h}^{(1)}_{k,k-1} - \bar{h}^{(2)}_{k,k-1} D_{k-1} \\ &h_{k,k-2} + \bar{h}^{(2)}_{k,k-1} h_{k-1,k-2} = \bar{h}^{(1)}_{k,k-2} - \bar{h}^{(2)}_{k,k-2} D_{k-2} \\ &... \\ &h_{k,k-p} + \bar{h}^{(2)}_{k,k-1} h_{k-1,k-p} + ... + \bar{h}^{(2)}_{k,k-p+1} h_{k-p+1,k-p} = \bar{h}^{(1)}_{k,k-p} - \bar{h}^{(2)}_{k,k-p} D_{k-p}\end{aligned} \tag{26}$$

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

$$\begin{bmatrix} I_m & \bar{h}^{(2)}_{k,k-1} & \bar{h}^{(2)}_{k,k-2} & \cdots & \bar{h}^{(2)}_{k,k-p+1} \\ 0 & I_m & \bar{h}^{(2)}_{k-1,k-2} & \cdots & \bar{h}^{(2)}_{k-1,k-p+2} \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I_m \end{bmatrix} \begin{bmatrix} h_{k,k-1} & h_{k,k-2} & \cdots & h_{k,k-p} \\ 0 & h_{k-1,k-2} & \cdots & h_{k-1,k-p} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_{k-p+1,k-p} \end{bmatrix} = \begin{bmatrix} r_{k,k-1} & r_{k,k-2} & \cdots & r_{k,k-p} \\ 0 & r_{k-1,k-2} & \cdots & r_{k-1,k-p} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{k-p+1,k-p} \end{bmatrix} \tag{27}$$

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. (27) with equation (6.8) of Juang[1]).

Considering the expressions for $\bar{h}_{k,k-p} := C_k \bar{A}_{k-1} ... \bar{A}_{k-p+1} \bar{B}_{k-p}$ and choosing p sufficiently large, we have that owing to the asymptotic stability of the closed loop (including the observer) $\bar{h}_{k,k-p} \approx 0$. This fact enables us to establish recursive relationships for the calculation of the system Markov parameters $h_{k,k-\gamma}$, $\forall \gamma > p$. Generalizing Eq. (25) (to introduce the variability of the order of the GTV-ARX model, as proposed else-where in the paper, i.e., setting $p = p_k$) produces

$$h_{k,k-\gamma} = \bar{h}^{(1)}_{k,k-\gamma} - \bar{h}^{(2)}_{k,k-\gamma} D_{k-\gamma} - \sum_{j=1}^{\gamma-1} \bar{h}^{(2)}_{k,k-j} h_{k-j,k-\gamma} \tag{28}$$

$\forall \gamma > p_k$. Then based on the constraint imposed in Eq. (17) for the calculation of the generalized observer Markov parameters, all the terms with time step separation greater than p_k vanish identically, and we obtain the relationship

$$h_{k,k-\gamma} = -\sum_{j=1}^{p_k} \overline{h}_{k,k-j}^{(2)} h_{k-j,k-\gamma} \tag{29}$$

For maintaining the simplicity of the presentations here-in, we will not make any more references to the variable order option in the subsequent developments of the paper. That is to say that the variable order of the GTV-ARX model at each time step is set to realize the time varying deadbeat observer, i.e., $p_k = p_{\min}$ ($= p$ to further clarify the developments). It also implies that only a minimum number of repeated experiments needs to be performed. Insight into the flexibility offered by the variable order of the GTV-ARX model is provided by appealing to the relations of the identified observer with a linear time varying Kalman filter in the next section of this paper.

B. Computation of Observer Gain Markov Parameters from the Observer Markov Parameters

Consider the generalized observer gain Markov parameters defined as

$$h_{k,i}^{o} = \begin{cases} C_k A_{k-1} A_{k-2} \ldots A_{i+1} G_i \,, & \forall k > i+1 \\ C_k G_{k-1}, & k = i+1 \\ 0, & \forall k < i+1 \end{cases} \tag{30}$$

We will now derive the relationship between these parameters and the GTV-ARX model coefficients $\overline{h}_{k,j}^{(2)}$. These parameters will be used in the calculation of the observer gain sequence from the input-output data in the next subsection, a generalization of the time invariant relations obtained in [1, 12] similar to Eq. (27).

From their corresponding definitions, note that

$$\overline{h}_{k,k-1}^{(2)} = C_k G_{k-1} = h_{k,k-1}^{o} \tag{31}$$

Similarly

$$\begin{aligned} \overline{h}_{k,k-2}^{(2)} &= C_k \overline{A}_{k-1} G_{k-2} \\ &= C_k \left(A_{k-1} + G_{k-1} C_{k-1} \right) G_{k-2} = h_{k,k-2}^{o} + \overline{h}_{k,k-1}^{(2)} h_{k-1,k-2}^{o} \end{aligned} \tag{32}$$

In general, an induction step similar to Eq. (23) holds and is given by,

$$
\begin{aligned}
\overline{h}^{(2)}_{k,k-p} &= C_k \overline{A}_{k-1}\overline{A}_{k-2}...\overline{A}_{k-p+1}G_{k-p} \\
&= C_k \overline{A}_{k-1}\overline{A}_{k-2}...\overline{A}_{k-p+2}\left(A_{k-p+1} + G_{k-p+1}C_{k-p+1}\right)G_{k-p} \\
&= C_k \overline{A}_{k-1}\overline{A}_{k-2}...\overline{A}_{k-p+2}A_{k-p+1}G_{k-p} + C_k \overline{A}_{k-1}\overline{A}_{k-2}...\overline{A}_{k-p+2}G_{k-p+1}C_{k-p+1}G_{k-p} \\
&= C_k \overline{A}_{k-1}\overline{A}_{k-2}...\overline{A}_{k-p+2}A_{k-p+1}G_{k-p} + \overline{h}^{(2)}_{k,k-p+1}h^{o}_{k-p+1,k-p} \\
&= h^{o}_{k,k-p} + \overline{h}^{(2)}_{k,k-1}h^{o}_{k-1,k-p} + ... + \overline{h}^{(2)}_{k,k-p+1}h^{o}_{k-p+1,k-p}
\end{aligned} \tag{33}
$$

Where the identity derived in Eq. (24) (replace B_{k-p} in favor of G_{k-p}) is used. This enables us to write the general relationship,

$$
\overline{h}^{(2)}_{k,k-\gamma} = h^{o}_{k,k-\gamma} + \sum_{i=1}^{\gamma-1}\overline{h}^{(2)}_{k,k-i}h^{o}_{k-i,k-\gamma} \tag{34}
$$

$\forall \gamma \in \mathbb{Z}^{+}$ analogous to Eq.(28) in case of the system Markov parameters. Also, similar to Eq.(29) we have the appropriate recursive relationship for the observer gain Markov parameters separated by more than p time steps for each k given as

$$
h^{o}_{k,k-\gamma} = -\sum_{i=1}^{p}\overline{h}^{(2)}_{k,k-i}h^{o}_{k-i,k-\gamma} \tag{35}
$$

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

$$
\begin{bmatrix}
I_m & \overline{h}^{(2)}_{k,k-1} & \overline{h}^{(2)}_{k,k-2} \cdots & \overline{h}^{(2)}_{k,k-p+1} \\
0 & I_m & \overline{h}^{(2)}_{k-1,k-2} \cdots & \overline{h}^{(2)}_{k-1,k-p+2} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 \quad \cdots & I_m
\end{bmatrix}
\begin{bmatrix}
h_{k,k-1} & h_{k,k-2} & \cdots & h_{k,k-p} \\
0 & h_{k-1,k-2} & \cdots & h_{k-1,k-p} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & h_{k-p+1,k-p}
\end{bmatrix}
=
\begin{bmatrix}
\overline{h}^{(2)}_{k,k-1} & \overline{h}^{(2)}_{k,k-2} & \cdots & \overline{h}^{(2)}_{k,k-p} \\
0 & \overline{h}^{(2)}_{k-1,k-2} & \cdots & \overline{h}^{(2)}_{k-1,k-p} \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \overline{h}^{(2)}_{k-p+1,k-p}
\end{bmatrix} \tag{36}
$$

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.

C. Calculation of the Realized Time Varying Observer Gain Sequence

From the definition of the observer gain Markov parameters, (recall equation (30)) we can stack the first few parameters in a tall matrix and observe that

$$P_{k+1} := \begin{bmatrix} h^o_{k+1,k} \\ h^o_{k+2,k} \\ \vdots \\ h^o_{k+m,k} \end{bmatrix} = \begin{bmatrix} C_{k+1}G_k \\ C_{k+2}A_{k+1}G_k \\ \vdots \\ C_{k+m}A_{k+m-1}\ldots A_{k+1}G_k \end{bmatrix} = \begin{bmatrix} C_{k+1} \\ C_{k+2}A_{k+1} \\ \vdots \\ C_{k+m}A_{k+m-1}\ldots A_{k+1} \end{bmatrix} G_k = O^{(m)}_{k+1}G_k \tag{37}$$

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

$$G_k = O^{(m)\dagger}_{k+1} P_k \tag{38}$$

However from the developments of the companion paper, we find that it is, in general, impossible to determine the observability grammian in the true coordinate system[10], as suggested by Eq.(38) above. This is because the computed observability grammian is, in general in a time varying and unknown coordinate system denoted by, $O^{(m)}_{k+1}$ at the time step t_{k+1}. We will now show that the gain computed from this time varying observability grammian is consistent with the time varying coordinates of the plant model computed by the Time Varying Eigensystem Realization Algorithm (TVERA) presented in the companion paper. Therefore upon using the computed observability grammian (in its own time varying coordinate system) and proceeding with the gain calculation as indicated by Eq.(38) above, we arrive at a consistent computed gain matrix. That is to say that, given a transformation matrix T_{k+1},

$$\begin{aligned} P_{k+1} &= O^{(m)}_{k+1}G_k \\ &= O^{(m)}_{k+1}T_{k+1}T^{-1}_{k+1}G_k \\ &= \hat{O}^{(m)}_{k+1}T^{-1}_{k+1}G_k \\ &= \hat{O}^{(m)}_{k+1}\hat{G}_k \end{aligned} \tag{39}$$

such that

$$\hat{G}_k = T^{-1}_{k+1}G_k = \left(\hat{O}^{(m)}_{k+1}\right)^{\dagger} P_{k+1} \tag{40}$$

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 companion paper. 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 (cf. companion paper[10]) :

$$\begin{aligned} \hat{A}_k &= T_{k+1}^{-1} A_k T_k \\ \hat{B}_k &= T_{k+1}^{-1} B_k \\ \hat{C}_k &= C_k T_k \end{aligned} \tag{41}$$

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

$$\hat{A}_k + \hat{G}_k \hat{C}_k = T_{k+1}^{-1} \left(A_k + G_k C_k \right) T_k \tag{42}$$

in a kinematically similar fashion to the true time varying OKID closed loop. The nature of the computed stabilizing (deadbeat or near – deadbeat) 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 detail in the companion paper.

V. 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 paper. 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 paper, we use the well understood time varying Kalman filter theory to make some intuitive observations. These observations help us qualitatively address the important issue: "Variable order GTV-ARX model fitting of input-output data – what it all means?". Insight is also obtained as to what happens in the presence of measurement noise. In the practical situation where there is the presence of 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 paper. Hence the authors limit the discussions to qualitative relations.

The Kalman filter equations for a truth model given in Eq.(54) of the appendix are given by

$$\hat{\mathbf{x}}_{k+1}^{-} = A_k \hat{\mathbf{x}}_k^{+} + B_k \mathbf{u}_k \tag{43}$$

or

$$\hat{\mathbf{x}}_{k+1}^{-} = A_k \left[I - K_k C_k \right] \hat{\mathbf{x}}_k^{-} + B_k \mathbf{u}_k + A_k K_k \mathbf{y}_k \tag{44}$$

together with the propagated output equation

$$\hat{\mathbf{y}}_k^{-} = C_k \hat{\mathbf{x}}_k^{-} + D_k \mathbf{u}_k \tag{45}$$

where the gain K_k is optimal (expression in Eq.(69)). As documented in the standard estimation theory textbooks, optimality translates to any one of the equivalent necessary conditions of minimum variance, maximum likelihood, orthogonality or Bayesian schemes. A brief review of the expressions for the optimal gain sequence is derived in the Appendix A 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 (See Eq. (69) in Appendix A for an expression for the optimal gain). From an input-output stand point, 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 shown in Eq.(45) can be written in favor of the system outputs as given by

$$\mathbf{y}_k = C_k \hat{\mathbf{x}}_k^{-} + D_k \mathbf{u}_k + \boldsymbol{\varepsilon}_k \tag{46}$$

Rearranging the state propagation shown in Eq.(43), we arrive at a form given by

$$\begin{aligned} \hat{\mathbf{x}}_{k+1}^{-} &= A_k \left[I - K_k C_k \right] \hat{\mathbf{x}}_k^{-} + B_k \mathbf{u}_k + A_k K_k \mathbf{y}_k \\ &= \tilde{A}_k \hat{\mathbf{x}}_k^{-} + \tilde{B}_k \mathbf{v}_k \end{aligned} \tag{47}$$

with the definitions

$$\begin{aligned} \tilde{A}_k &= A_k \left[I - K_k C_k \right] \\ \tilde{B}_k &= \begin{bmatrix} B_k & A_k K_k \end{bmatrix} \\ \mathbf{v}_k &= \begin{bmatrix} \mathbf{u}_k \\ \mathbf{y}_k \end{bmatrix} \end{aligned} \tag{48}$$

Notice the structural similarity in the layout of the rearranged equations to the time varying OKID equations in section III. 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

$$
\begin{aligned}
\mathbf{y}_0 &= C_0 \hat{\mathbf{x}}_0^- + D_0 \mathbf{u}_0 + \boldsymbol{\varepsilon}_0 \\
\mathbf{y}_1 &= C_1 \tilde{A}_0 \hat{\mathbf{x}}_0^- + D_1 \mathbf{u}_1 + C_1 \tilde{B}_0 \mathbf{v}_0 + \boldsymbol{\varepsilon}_1 \\
\mathbf{y}_2 &= C_2 \tilde{A}_1 \tilde{A}_0 \hat{\mathbf{x}}_0^- + D_2 \mathbf{u}_2 + C_2 \tilde{B}_1 \mathbf{v}_1 + C_2 \tilde{A}_1 \tilde{B}_0 \mathbf{v}_0 + \boldsymbol{\varepsilon}_2 \\
&\dots \\
\mathbf{y}_p &= C_p \tilde{A}_{p-1} \dots \tilde{A}_0 \hat{\mathbf{x}}_0^- + D_p \mathbf{u}_p + \sum_{j=1}^{p-1} \tilde{h}_{p,p-j} \mathbf{v}_{p-j} + \boldsymbol{\varepsilon}_p \\
&\dots
\end{aligned}
\tag{49}
$$

suggesting the general relationship

$$
\mathbf{y}_{k+p} = C_{k+p} \tilde{A}_{k+p-1} \dots \tilde{A}_0 \hat{\mathbf{x}}_0^- + D_{k+p} \mathbf{u}_{k+p} + \sum_{j=1}^{k+p-1} \tilde{h}_{k+p,k+p-j} \mathbf{v}_{k+p-j} + \boldsymbol{\varepsilon}_{k+p} \tag{50}
$$

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

$$
\tilde{h}_{k,i} = \begin{cases} C_k \tilde{A}_{k-1} \tilde{A}_{k-2} \dots \tilde{A}_{i+1} \tilde{B}_i \, , & \forall k > i+1 \\ C_k \tilde{B}_{k-1}, & k = i+1 \\ 0, & \forall k < i+1 \end{cases} \tag{51}
$$

Comparing the Eqs.(14) and (50) 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 $\boldsymbol{\varepsilon}_k = \mathbf{0}, \forall k > p$. 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 A. Therefore under these conditions, (more specifically the innovations property) our algorithm is expected to produce a gain sequence that is optimal.

However, we proceeded to enforce the p (in general $p_k = p_{\min}$ was set) term dependence in Eq.(16) 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 deadbeat observer notions set up in this paper (following the classical developments of Minamide et al.[14], and Hostetter[15] discussed in Appendix B). Notice that the Kalman filter equations are in general not truncated to the first $p\left(p_k\right)$ terms. Furthermore, the observer realized using the optimality condition (minimum variance for example) are seldom the fastest observers. 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 Eq.(50). 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 (16) 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 paper are intimately tied into 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.

VI. 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 sections.

Consider the same system as presented in an example of the companion paper. It has an oscillatory nature and does not have a stable origin. In case of the time invariant systems, systems of oscillatory nature are characterized by poles on the unit circle and the origin is said to be marginally stable[18, 19]. However, since the system under consideration is not autonomous, the origin is said to be unstable in the sense of Lyapunov[16, 20]. A separate classification has been provided in the theory of nonlinear systems for systems with origin of this type. That is called orbital stability or stability in the sense of Poincare (cf. Meirovitch [21]). We follow the convention of Lyapunov and term the system under consideration unstable. In this case the plant system matrix was calculated as

$$
\begin{aligned}
A_k &= \exp\left[A_c * \Delta t \right] \\
B_k &= \begin{bmatrix} 1 & 0 \\ 1 & -1 \\ 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad C_k = \begin{bmatrix} 1 & 0 & 1 & 0.2 \\ 1 & -1 & 0 & -0.5 \end{bmatrix}, \\
D_k &= 0.1\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\end{aligned} \tag{52}
$$

where the matrix is given by

$$
A_c = \begin{bmatrix} 0_{2\times2} & I_{2\times2} \\ -K_t & 0_{2\times2} \end{bmatrix} \tag{53}
$$

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 main body of the paper, is now 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 companion paper to realize system model sequence for all the time steps for which experimental data is available. We demonstrate the computation of the deadbeat like observer where the smallest order for the GTV-ARX model is chosen throughout the time history of the identification process. Appendix B details the definition of time varying deadbeat observer, for the convenience of the readers along with a representative closed-loop sequence result using this example problem.

Relating to the discussions of the previous section, this implies 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. In this case we were able to realize an asymptotically stable closed loop for the observer equation with OKID. In fact two of the closed loop eigenvalues could be assigned to zero at each time step and there is a certain distribution of closed loop eigenspaces such that the product of any two consecutive closed loop matrices has all the poles at origin. This time varying deadbeat condition realized is demonstrated using the same example in the Appendix B. 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 1.

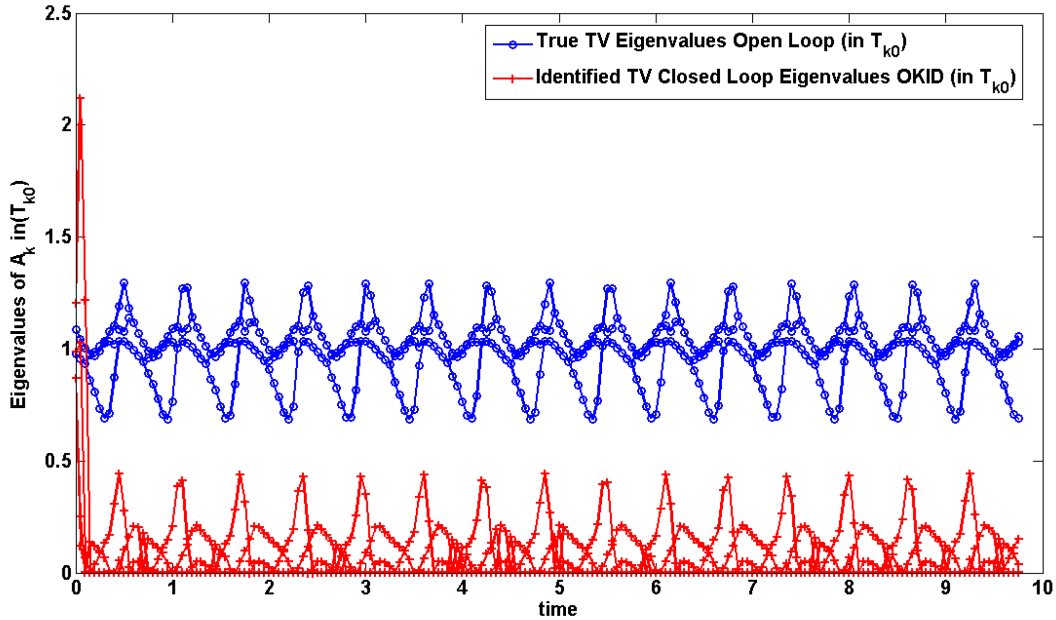

Figure 1. Case 1: Plant Open Loop Vs OKID Closed Loop Pole Locations (Minimum No of Repeated Experiments)

The error incurred in the identification of the system Markov parameters is in two parts. The system Markov parameters for the significant number of steps in the GTV-ARX model are in general computed exactly. However, we would still need extra system Markov parameters to assemble the generalized Hankel matrix sequence for the TV ERA algorithm. These are computed using the recursive relationships. Since the truncation of the input-output relationship even with the observer in the loop is an approximation for the time varying case, we incur some error. The worst case error although it is sufficiently small is incurred in the situation when minimum number of experiments is performed. This is plotted in Figure 2. The comparison in this figure is made with error in system Markov parameters computed from the full input-output relationship of the observer (shown to have the same structure as the Kalman Filter in the no noise case). Performing larger number of experiments in general leads to better accuracy as shown in Figure 3. Note that more experiments give a better condition number for making the pseudo inverse of the matrix $\mathbf{V}_k$ shown in Eq.(18). The accuracy is also improved by retaining larger number of terms (per time step) in the input-output map.

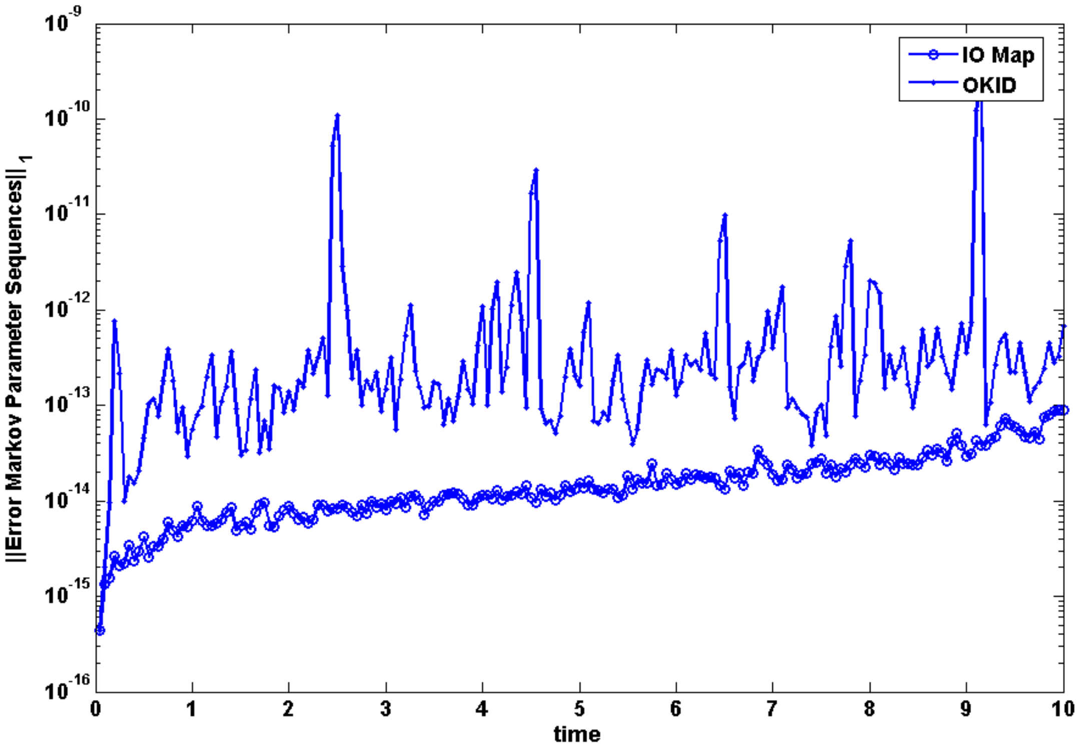

Figure 2. Case 1: Error in System Markov Parameters (Minimum No of Repeated Experiments = 10)

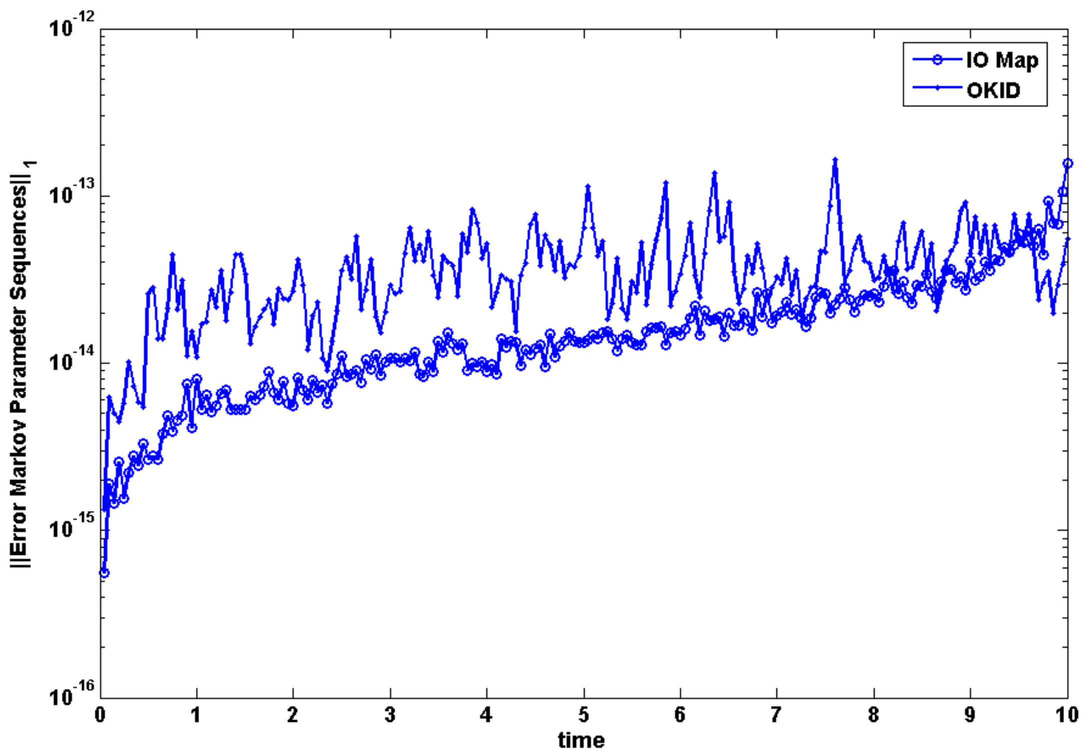

Figure 3. Case 2: Error in Markov Parameters Computations (Non-Minimum Number of Repeated Experiments)

The error incurred in the system Markov parameter computation 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 cases. The corresponding output error plots for Figures 2 and 3 are shown in Figure 4 and Figure 5.

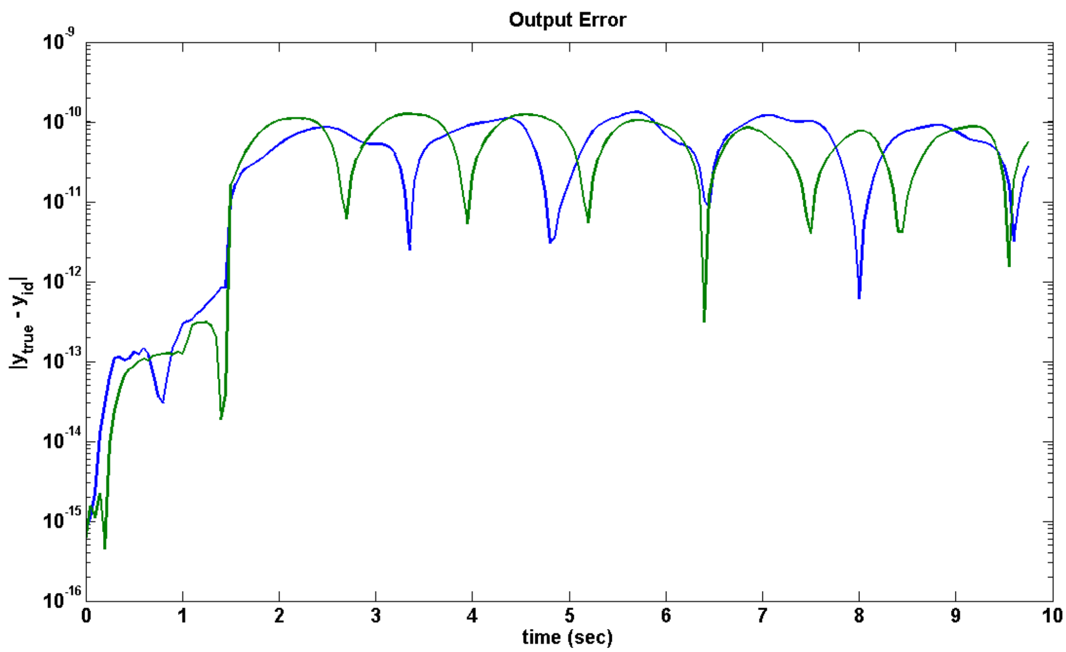

Figure 4. Case 1: Error in Outputs (Minimum No of Repeated Experiments)

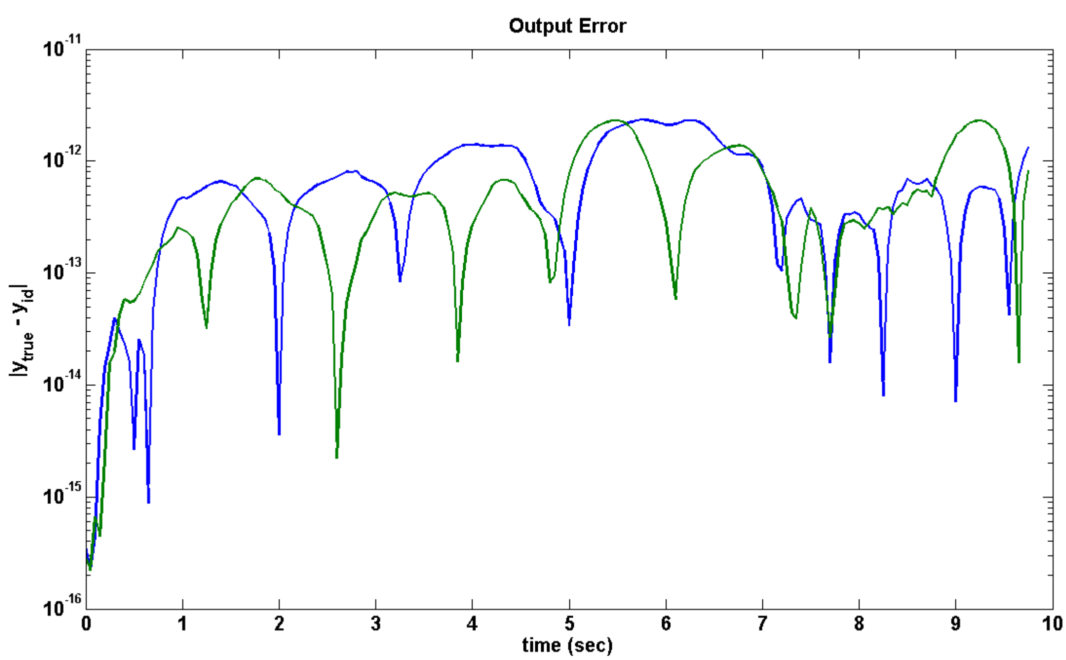

Figure 5. Case 2: Error in Outputs for Test Functions (True Vs. Identified Plant Model)

Because the considered system is unstable (oscillatory) in nature, the initial condition response was used to check the nature of state-error decay of the 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 state-error response including the realized observer are plotted in Figure 6. The plot represents the errors of convergence of a time varying deadbeat observer to the true states of the system. The computed states converge to the true states in precisely *two time* ($p_{\min} = 2$) steps to zero response. An important comment is due at this stage regarding the convergence of the observer in $p_{\min}$ time steps. Since the observer gain Markov parameters for the first few time steps cannot be calculated (because the free decay experiments do not yield any information for the gain calculations - cf. companion paper [10] for details) the corresponding observer gain sequence cannot be determined uniquely. Hence the $p_{\min}$ time steps in implementation implies after the first few time steps – this translates to around $2p_{\min}$ time steps in most implementations. In other words the decay of the deadbeat closed loop starts after the correct determination of unique gains from the time varying OKID procedure. In the example problem, this number $2p_{\min} = 4$ can be clearly seen from the nonzero output error time steps in Figure 6.

This decay to zero was exponential and too steep to plot for the (time varying) deadbeat case. However when the order was chosen to be slightly higher (near deadbeat observer is realized in this case and therefore it takes more than two steps for the response to decay to zero). The gain history of the realized observer as seen in the initial condition coordinate system is plotted as Figure 7.

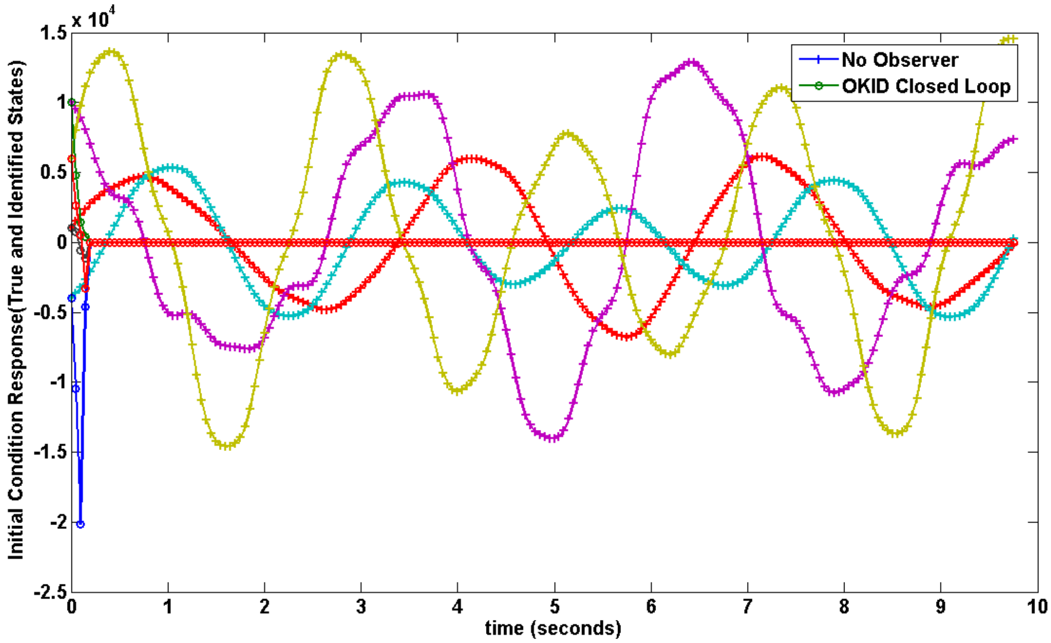

Figure 6. Case 1: Open Loop Vs OKID Closed Loop Response to Initial Conditions

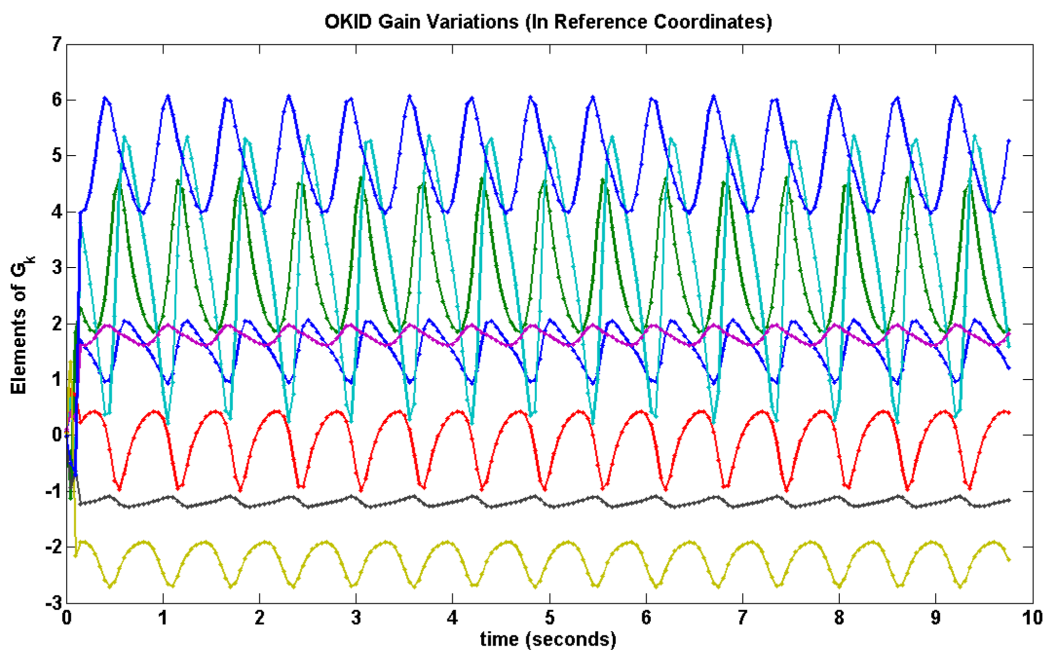

Figure 7. Case 1: Gain History (Minimum No of Repeated Experiments)

VII. Conclusion

The paper provides an algorithm for efficient computation of system Markov parameters for use in time varying system identification. 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 obtains 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. 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.

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[22, 23]) using the innovations approach propounded by Kailath[24] and Mehra[25]. The most commonly used truth model for the linear time varying filtering problem is given by

$$\mathbf{x}_{k+1} = A_k \mathbf{x}_k + B_k \mathbf{u}_k + \Gamma_k \mathbf{w}_k \tag{54}$$

together with the measurement equations given by

$$\mathbf{y}_k = C_k \mathbf{x}_k + D_k \mathbf{u}_k + \boldsymbol{\upsilon}_k \tag{55}$$

The process noise sequence is assumed to be a Gaussian random sequence with zero mean $E\left(\mathbf{w}_i\right) = \mathbf{0}, \forall i$ and a variance sequence $E\left(\mathbf{w}_i \mathbf{w}_j^T\right) = Q_i \delta_{ij}, \forall i, j$ having an uncorrelated profile in time (with itself, as shown by the variance expression) and no correlation with the measurement noise sequence $E\left(\mathbf{w}_i \boldsymbol{\upsilon}_j^T\right) = 0, \forall i, j$. Similarly, the measurement noise sequence is assumed to be a zero mean Gaussian random vector with covariance sequence given by $E\left(\boldsymbol{\upsilon}_i \boldsymbol{\upsilon}_j^T\right) = R_i\,\delta_{ij}$. We 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 [12])

$$\begin{aligned} \hat{\mathbf{x}}_k^+ &= \hat{\mathbf{x}}_k^- + K_k\left[\mathbf{y}_k - \hat{\mathbf{y}}_k\right] \\ &:= \hat{\mathbf{x}}_k^- + K_k\boldsymbol{\varepsilon}_k \end{aligned} \tag{56}$$

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 into 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

$$\begin{aligned} \hat{\mathbf{x}}_{k+1}^- &= A_k\hat{\mathbf{x}}_k^+ + B_k\mathbf{u}_k \\ &= A_k\left[I - K_k C_k\right]\hat{\mathbf{x}}_k^- + B_k\mathbf{u}_k + A_k K_k\mathbf{y}_k \end{aligned} \tag{57}$$

and

$$\hat{\mathbf{y}}_k^- = C_k\hat{\mathbf{x}}_k^- + D_k\mathbf{u}_k \tag{58}$$

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

$$\boldsymbol{\varepsilon}_k = C_k\mathbf{e}_k + \boldsymbol{\upsilon}_k \tag{59}$$

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

$$\begin{aligned} \mathbf{e}_{k+1} &= A_k\left[I - K_k C_k\right]\mathbf{e}_k - A_k K_k\boldsymbol{\upsilon}_k + \Gamma_k\mathbf{w}_k \\ &:= \tilde{A}_k\mathbf{e}_k - A_k K_k\boldsymbol{\upsilon}_k + \Gamma_k\mathbf{w}_k \end{aligned} \tag{60}$$

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

$$P_{k+1} = \tilde{A}_k P_k \tilde{A}_k^T + A_k K_k R_k K_k^T A_k^T + \Gamma_k Q_k \Gamma_k^T \tag{61}$$

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[24] 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

$$E\left[\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_{i-k}^T\right] = 0 \tag{62}$$

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

$$E\left[\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_{i-k}^T\right] = C_i\, E\left[\mathbf{e}_i \mathbf{e}_{i-k}^T\right] C_{i-k}^T + C_i\, E\left[\mathbf{e}_i \boldsymbol{\upsilon}_{i-k}^T\right] = 0 \tag{63}$$

where the two terms $E\left[\boldsymbol{\upsilon}_i \mathbf{e}_{i-k}^T\right] = E\left[\boldsymbol{\upsilon}_i \boldsymbol{\upsilon}_{i-k}^T\right] = 0$ 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 Eq. (63) becomes

$$\begin{aligned} E\left[\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_i^T\right] &= C_i\, E\left[\mathbf{e}_i \mathbf{e}_i^T\right] C_i^T + E\left[\boldsymbol{\upsilon}_i \boldsymbol{\upsilon}_i^T\right] \\ &= C_i\, P_i\, C_i^T + R_i \end{aligned} \tag{64}$$

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

$$\begin{aligned} \mathbf{e}_i &= \tilde{A}_{i-1}\tilde{A}_{i-2}\ldots\tilde{A}_{i-k+1}\tilde{A}_{i-k}\mathbf{e}_{i-k} - \left[\tilde{A}_{i-1}\ldots\tilde{A}_{i-k+1}A_{i-k}K_{i-k}\boldsymbol{\upsilon}_{i-k} + \ldots + \tilde{A}_{i-1}A_{i-2}K_{i-2}\boldsymbol{\upsilon}_{i-2} + A_{i-1}K_{i-1}\boldsymbol{\upsilon}_{i-1}\right] \\ &\quad + \left[\tilde{A}_{i-1}\ldots\tilde{A}_{i-k+1}\Gamma_{i-k}\mathbf{w}_{i-k} + \ldots + \tilde{A}_{i-1}\Gamma_{i-2}\mathbf{w}_{i-2} + \Gamma_{i-1}\mathbf{w}_{i-1}\right] \end{aligned} \tag{65}$$

We obtain expressions for $E\left[\mathbf{e}_i \mathbf{e}_{i-k}^T\right]$ and $E\left[\mathbf{e}_i \boldsymbol{\upsilon}_{i-k}^T\right]$ by operating equation (65) on both sides with $\mathbf{e}_{i-k}^T$ and $\boldsymbol{\upsilon}_{i-k}^T$, and taking the expectation operator

$$\begin{aligned} E\left[\mathbf{e}_i \mathbf{e}_{i-k}^T\right] &= \tilde{A}_{i-1}\tilde{A}_{i-2}\ldots\tilde{A}_{i-k+1}\tilde{A}_{i-k}E\left[\mathbf{e}_{i-k}\mathbf{e}_{i-k}^T\right] \\ &= \tilde{A}_{i-1}\tilde{A}_{i-2}\ldots\tilde{A}_{i-k+1}\tilde{A}_{i-k}P_{i-k} \end{aligned} \tag{66}$$

$$\begin{aligned} E\left[\mathbf{e}_i \boldsymbol{\upsilon}_{i-k}^T\right] &= -\tilde{A}_{i-1}\ldots\tilde{A}_{i-k+1}A_{i-k}K_{i-k}E\left(\boldsymbol{\upsilon}_{i-k}\boldsymbol{\upsilon}_{i-k}^T\right) \\ &= -\tilde{A}_{i-1}\ldots\tilde{A}_{i-k+1}A_{i-k}K_{i-k}R_{i-k} \end{aligned} \tag{67}$$

Substituting Eqs. (67) and (66) in to the expression for the inner product shown in Eq. (63), 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} as

$$
\begin{aligned}
E\left[\boldsymbol{\varepsilon}_i \boldsymbol{\varepsilon}_{i-k}^T\right] &= C_i\,\tilde{A}_{i-1}\tilde{A}_{i-2}\ldots\tilde{A}_{i-k+1}\tilde{A}_{i-k}P_{i-k}C_{i-k}^T - C_i\,\tilde{A}_{i-1}\ldots\tilde{A}_{i-k+1}A_{i-k}K_{i-k}R_{i-k} \\
&= C_i\,\tilde{A}_{i-1}\tilde{A}_{i-2}\ldots\tilde{A}_{i-k+1}\left[\tilde{A}_{i-k}P_{i-k}C_{i-k}^T - A_{i-k}K_{i-k}R_{i-k}\right] \\
&= C_i\,\tilde{A}_{i-1}\tilde{A}_{i-2}\ldots\tilde{A}_{i-k+1}A_{i-k}\left[P_{i-k}C_{i-k}^T - K_{i-k}\left(R_{i-k} + C_{i-k}P_{i-k}C_{i-k}^T\right)\right] \\
&= 0
\end{aligned}
\tag{68}
$$

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

$$
K_{i-k} = P_{i-k}C_{i-k}^T\left(R_{i-k} + C_{i-k}P_{i-k}C_{i-k}^T\right)^{-1} \tag{69}
$$

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.

Appendix B

Time Varying Deadbeat Observers

It was shown in the paper that the generalization of the ARX model in the time varying case gives rise to an observer that could be set to a deadbeat condition that has different properties and structure when compared to its linear time invariant counterpart. The topic of extension of the deadbeat 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.[14], develops a similar definition of the time varying deadbeat 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 deadbeat observer models directly from input-output data.

First we recall the definition of a deadbeat 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[19], if a linear discrete time dynamical system is characterized by the evolution equations given by

$$
\mathbf{x}_{k+1} = A\mathbf{x}_k + B\mathbf{u}_k \tag{70}
$$

with the measurement equations (assuming that (C, A) is an observable pair)

$$
\mathbf{y}_k = C\mathbf{x}_k + D\mathbf{u}_k \tag{71}
$$

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 deadbeat observer, if and only if the following condition is satisfied (the so-called deadbeat condition):

$$(A + GC)^p = [0]_{n \times n} \tag{72}$$

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.

D. Example of a Time Invariant Deadbeat Observer:

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

$$\begin{aligned} A &= \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} \\ C &= \begin{bmatrix} 0 & 1 \end{bmatrix} \end{aligned} \tag{73}$$

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

$$\begin{aligned} (A + GC)^2 &= \begin{bmatrix} 1 + g_1 & g_1 \left(3 + g_2\right) \\ 3 + g_2 & g_1 + \left(2 + g_2\right)^2 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \end{aligned} \tag{74}$$

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

$$(A + GC) = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \tag{75}$$

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 deadbeat 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 [14]) and when only experimental data is available (OKID solution).

Now considering the time varying system and following the notation developed in the main body of the paper, this time varying deadbeat definition appears to have been naturally made. Recall (from Eq.(16)) that in constructing the generalized time varying ARX (GTV-ARX) model of this paper, we have already used this definition. Thus we can formally write the definition of a time varying deadbeat observer as follows

Definition:

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

$$\left(A_{k+p-1}+G_{k+p-1}C_{k+p-1}\right)\left(A_{k+p-2}+G_{k+p-2}C_{k+p-2}\right)\ldots\left(A_k+G_kC_k\right)=\left[0\right]_{n\times n} \tag{76}$$

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

We now illustrate this definition using an example problem.

E. Example of a Time Varying Dead Beat Observer

To show the ideas, we demonstrate the observer realized on the same problem used in the Numerical Example section (Section VI) of the paper and follow the example by a short discussion on the nature and properties of the time varying deadbeat condition in case of the observer design. The parameters involved in the example problem are given (we repeat here for convenience) as

$$\begin{aligned}
&A_k=\exp\left[A_c*\Delta t\right]\\
&B_k=\begin{bmatrix}1&0\\1&-1\\0&1\\-1&0\end{bmatrix},\ C_k=\begin{bmatrix}1&0&1&0.2\\1&-1&0&-0.5\end{bmatrix},\\
&D_k=0.1\begin{bmatrix}1&0\\0&1\end{bmatrix}
\end{aligned} \tag{77}$$

where the matrix is given by

$$A_c=\begin{bmatrix}0_{2\times2}&I_{2\times2}\\-K_t&0_{2\times2}\end{bmatrix} \tag{78}$$

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 $\tau_k = \sin\left(10t_k\right)$, $\tau'_k := \cos\left(10t_k\right)$. Clearly since $m=2, n=4$ for the example, the choice of $p=2$ is made to realize the time varying deadbeat condition. 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) system matrix and its eigenvalues are computed as

$$
\begin{aligned}
A_{36}+G_{36}C_{36} &= \begin{bmatrix} -2.0405 & 0.3357 & 0.0016 & 0.5965 \\ -1.7735 & -0.7681 & -3.2887 & -0.0289 \\ 1.7902 & -0.0270 & 0.8290 & -0.3852 \\ -6.9208 & 1.4773 & 1.0572 & 2.1980 \end{bmatrix} \\
\lambda\left(A_{36}+G_{36}C_{36}\right) &= \begin{bmatrix} 0.31545 \\ -0.097074 \\ 1.1878\times 10^{-15} \\ 1.2252\times 10^{-13} \end{bmatrix}
\end{aligned}
\tag{79}
$$

while the closed loop system matrix (and its eigenvalues) for the previous time step is calculated as

$$
\begin{aligned}
A_{35}+G_{35}C_{35} &= \begin{bmatrix} -1.7924 & 0.4678 & 0.1630 & 0.5778 \\ -0.7301 & -0.4380 & -2.8865 & -0.1330 \\ 1.1874 & -0.1662 & 0.5671 & -0.2986 \\ -5.7243 & 1.8475 & 2.1805 & 2.0524 \end{bmatrix} \\
\lambda\left(A_{35}+G_{35}C_{35}\right) &= \begin{bmatrix} 0.43716 \\ -0.048167 \\ \left(-2.1173+7.4549\mathrm{i}\right)\times 10^{-14} \\ \left(-2.1173-7.4549\mathrm{i}\right)\times 10^{-14} \end{bmatrix}
\end{aligned}
\tag{80}
$$

For the consecutive time step these values are found to be given by

$$
\begin{aligned}
A_{37}+G_{37}C_{37} &= \begin{bmatrix} -2.4701 & 0.1432 & -0.2323 & 0.6315 \\ -2.3403 & -0.8353 & -3.3551 & 0.0362 \\ 2.0767 & 0.0773 & 0.8335 & -0.4165 \\ -8.8651 & 0.6963 & -0.2452 & 2.3719 \end{bmatrix} \\
\lambda\left(A_{37}+G_{37}C_{37}\right) &= \begin{bmatrix} -0.14861 \\ 0.048661 \\ 4.0371\times 10^{-12} \\ -5.5501\times 10^{-15} \end{bmatrix}
\end{aligned}
\tag{81}
$$

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

$$\left(A_{37}+G_{37}C_{37}\right)\left(A_{36}+G_{36}C_{36}\right)=10^{-12}\times\begin{bmatrix} -0.0959 & 0.0070 & -0.0326 & 0.0238 \\ -0.1192 & 0.0035 & -0.0187 & 0.0235 \\ -0.0564 & 0.0003 & -0.0307 & 0.0123 \\ 0.1137 & 0.0075 & 0.0551 & -0.0187 \end{bmatrix} \quad (82)$$

and

$$\left(A_{36}+G_{36}C_{36}\right)\left(A_{35}+G_{35}C_{35}\right)=10^{-13}\times\begin{bmatrix} -0.0844 & -0.1443 & 0.0888 & -0.0711 \\ 0.4660 & -0.2783 & 0.4528 & -0.2652 \\ -0.2265 & 0.1987 & -0.2076 & 0.1610 \\ -0.6217 & -0.4086 & 0.1243 & -0.1332 \end{bmatrix} \quad (83)$$

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 deadbeat condition in the case of time varying systems. Similar to the case of time invariant systems, the observer which is deadbeat happens to be the fastest observer for the given (or realized) time varying system model.

We reiterate the fact that in the current developments, the deadbeat observer (gain sequence) is realized naturally along with the plant model sequence being identified. It is not difficult to see that the time varying OKID procedure (the generalized ARX (GTV-ARX) model construction and the deadbeat observer calculation) subsumes the special case when the time varying discrete time plant model is 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. [14], which is quite similar to the method outlined by Hostetter[15], are based on projections of the state space on to the outputs.

Acknowledgments

The authors wish to acknowledge the support of Texas Institute of Intelligent Bio Nano Materials and Structures for Aerospace Vehicles (TIIMS) funded by NASA Cooperative Agreement No. NCC-1-02038. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of National Aeronautics and Space Administration.

References

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

[2] Cho, Y. M., Xu, G., and Kailath, T.,, "Fast Recursive Identification of State Space Models via Exploitation of Displacement Structure," *Automatica,* vol. 30, no. 1, 1994, pp. 45 - 59.
doi:10.1016/0005-1098(94)90228-3

[3] 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

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

[5] 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

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

[7] 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

[8] 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

[9] 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

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

[11] 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).

[12] 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

[13] 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

[14] 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

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

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

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

[18] Chen, C. T., *Linear System Theory and Design*, Holt, Rinehart and Winston, New York, 1984, pp. 130 -131

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

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

[21] Meirovitch, L., *Elements of Vibration Analysis*, McGraw-Hill Book Company, New York, NY, 1986, pp. 352 - 353.

[22] 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.

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

[24] 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

[25] 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