Observer/Kalman-Filter Time-Varying System Identification

Manoranjan Majji‡
Texas A&M University, College Station, Texas 77843-3141
Jer-Nan Juang‡
National Applied Research Laboratories, Taipei 106 Taiwan, Republic of China and
John L. Junkins‡
Texas A&M University, College Station, Texas 77843-3141

DOI: 10.2514/1.45768

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. The generalized system Markov parameters thus derived are used by the time-varying eigensystem realization algorithm to obtain a time-varying discrete-time state-space model. A qualitative relationship of the time-varying observer with the Kalman filter in the stochastic environment and an asymptotically 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 determined from these developments. The time-varying observer gains realized in the process are subsequently shown to be in consistent coordinate systems for observer state propagation. It is also demonstrated that the observer gain sequence realized in the 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. The eigensystem realization algorithm (ERA) [1] 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 (I/O) 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 Yu [5] 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 [6] 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 [7]. An important concept of kinematic similarity among linear discrete-time-varying system models concerns certain time-varying transformations involved in the state transition matrices. Golberg et al. [8] discussed fundamental developments of this theory using a difference equation operator theoretic approach. In companion papers, Majji et al. [9,10] extended the ERA (a classical algorithm for system identification of linear time-invariant systems) to realize discrete-time-varying models from I/O data, following the framework and conventions of the previously mentioned papers. The time-varying ERA (TVERA) presented in the companion papers [9,10] uses the generalized Markov parameters to realize time-varying system descriptions by manipulations of 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-coordinate system at a given time step.

However, the algorithm developed therein requires the determination of the generalized Markov parameters from sets of I/O experimental data. Therefore, we need a practical method to calculate them without resorting to a high dimensional calculation. This calculation becomes further compounded in systems in which the stability of the origin cannot be ascertained, because the number of potentially significant generalized Markov parameters grows rapidly. In other words, in the case of the problems with unstable origins, the output at every time step in the time-varying case depends on the linear combinations of the pulse response functions and all the inputs applied until that instant. Therefore, the number of unknowns increase by \( m \times r \) 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 needed to perform identification and/or recovery of all the Markov parameters of interest until that time instant. Thus, the central result of the current paper is to make the number of repeated experiments independent of the desired resolution of the model. Furthermore, because the frequency response functions for time-varying systems are not well known, the method outlined seems to be one of the first practical ways to obtain the generalized Markov parameters, bringing the time-varying identification methods to the table of the practicing engineer.

Novel models relating I/O data are developed in this paper and are found to be elegant extensions of the autoregressive exogenous input (ARX) models, well known in the analysis of time-invariant models (cf., Juang et al. [11]). The ARX model is a linear difference equation that relates the input to the output. 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 deadbeat. Furthermore, it is demonstrated that the time-varying observer sequence (deadbeat or otherwise) computed from the 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 with the time-varying observer gains realized from the TOKID procedure presented in the paper.

II. Basic Formulation

We start by revisiting the relations between the I/O sets of vectors via the system Markov parameters, as developed in the theory concerning the TVERA (refer to companion paper [9], based on [10], and the references therein). The fundamental difference equations governing the evolution of a linear system in discrete time are given by

\[ x_{k+1} = A_k x_k + B_k u_k \]

(1)

together with the measurement equations:

\[ y_k = C_k x_k + D_k u_k \]

(2)

with the state, output, and input dimensions \( x_k \in \mathbb{R}^n \), \( y_k \in \mathbb{R}^m \), and \( u_k \in \mathbb{R}^r \), and the system matrices to be of compatible dimensions for all \( k \in \mathbb{Z} \), an index set. The solution of the state evolution (the linear time-varying discrete-time difference equation solution) is given by

\[ x_k = \Phi(k, k_0)x_0 + \sum_{i=k_0}^{k-1} \Phi(k, i + 1) B_i u_i \]

(3)

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

\[ \Phi(k, k_0) = \begin{cases} A_{k-1}, A_{k-2}, \ldots, A_{k_0} & \forall k > k_0 \\ I & k = k_0 \\ \text{undefined} & \forall k < k_0 \end{cases} \]

(4)

Using the definition of the compound state transition matrix, the I/O relationship is given by

\[ y_k = C_k \Phi(k, 0)x_0 + \sum_{i=0}^{k-1} C_k \Phi(k, i + 1) B_i u_i + D_k u_k \]

(5)

This enables us to define the I/O relationship, in terms of the two index coefficients, as

\[ y_k = C_k \Phi(k, 0)x_0 + \sum_{i=0}^{k-1} h_{k,i} u_i + D_k u_k \]

(6)

where the generalized Markov parameters are defined as

\[ h_{k,i} = \begin{cases} C_k \Phi(k, i + 1) B_i & \forall i < k - 1 \\ C_k B_{k-1} & i = k - 1 \\ 0 & \forall i > k - 1 \end{cases} \]

(7)

From now on, we try to use the expanded form of the state transition matrix \( \Phi(\cdot, \cdot) \) to improve the clarity of the presentation. Thus, the output at any general time step \( t_1 \) is related to the initial conditions and all the inputs as

\[ y_k = C_k A_{k-1}, \ldots, A_{k_0 + 1} A_{k-1} x_0 \]

\[ + [D_k C_k B_{k-1} \cdots C_k A_{k_0 + 1} B_{k_0}] \]

(8)

where \( k_0 \) can denote any general time step before \( k \) (in particular, let us assume that it denotes the initial time, such that \( k_0 = 0 \)). As was pointed out in the companion paper (cf., [9]), such a relationship between the input and output leads to a problem that increases by \( m \times r \) parameters for every time step considered. Thus, it becomes difficult to compute the increasing number of unknown parameters. In the special case of systems for which the open loop is asymptotically stable, this is not a problem. However, frequently, one tries to use identification in problems that 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 I/O 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 important, because the connections between the time-varying ARX model, the state-space model, and the 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 I/O representations for the time-varying systems are quite similar to the I/O 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 model parameters involved. An effective method for compressing experimental I/O data, called the observer/Kalman-filter Markov parameter identification theory (OKID) was developed by Juang [1], Juang et al. [11], and Phan et al. [12]. 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 the case of the time-varying systems. Therefore, the method described herein 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 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 model [Eq. (1)], producing

$$x_{k+1} = A_k x_k + B_k u_k + G_k y_k - G_k y_k - G_k y_k$$

$$= (A_k + G_k C_k) x_k + (B_k + G_k D_k) u_k - G_k y_k$$

$$= \tilde{A}_k x_k + (B_k + G_k D_k) y_k - G_k y_k$$

$$= \tilde{A}_k x_k + \tilde{B}_k y_k$$

(9)

with the definitions:

$$\tilde{A}_k = A_k + G_k C_k; \quad \tilde{B}_k = [B_k + G_k D_k - G_k]; \quad \nu_k = \begin{bmatrix} u_k \\ y_k \end{bmatrix}$$

(10)

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

$$y_k = C_k x_k + D_k u_k$$

(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

$$y_0 = C_0 x_0 + D_0 u_0$$

$$y_1 = C_1 \tilde{A}_0 x_0 + D_1 u_1 + C_1 \tilde{B}_0 y_0 = C_1 \tilde{A}_0 x_0 + D_0 u_0 + \tilde{h}_{1,0} \nu_0$$

$$y_2 = C_2 \tilde{A}_1 x_0 + D_2 u_2 + C_2 \tilde{B}_1 y_1 = C_2 \tilde{A}_1 x_0 + D_2 u_2 + C_2 \tilde{B}_1 \nu_1$$

$$\vdots$$

(12)

with the definition of the generalized observer Markov parameters:

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

(13)

we arrive at the general I/O relationship:

$$y_k = C_k \tilde{A}_{k-1}, \ldots, \tilde{A}_0 x_0 + D_k u_k + \sum_{j=1}^{k} \tilde{h}_{k,j-1} \nu_{k-j}$$

(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

$$\tilde{h}_{k,j} = \begin{bmatrix} C_k \tilde{A}_{k-1}, \ldots, \tilde{A}_{j+1} \tilde{B}_{k-j} \\ C_k \tilde{B}_{k-j} \\ 0 \end{bmatrix}$$

$$\approx \begin{bmatrix} C_k \tilde{A}_{k-1}, \ldots, \tilde{A}_{j+1} (B_{k-j} + G_{k-j} C_{k-j}) - C_k \tilde{A}_{k-1}, \ldots, \tilde{A}_{j+1} C_{k-j} \\ \tilde{h}_{k,j}^{(1)} - \tilde{h}_{k,j-1}^{(2)} \end{bmatrix}$$

$$\tilde{h}_{k,j}^{(1)}$$

and

$$\tilde{h}_{k,j}^{(2)}$$

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. 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 a 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 only on the $p_k + 1$ previously augmented control inputs:

$$\{v_{k+j-1}\}_{j=1}^{p_k}$$

and

$$u_{k+p_k}$$

and independent of the state at every time step $t_k$. Therefore, by writing

$$y_{k+p} = C_{k+p} \tilde{A}_{k+p-1}, \ldots, \tilde{A}_1 x_k + D_{k+p} u_{k+p}$$

$$+ \sum_{j=1}^{p_k} \tilde{h}_{k+p,j-1} v_{k+j-1} \approx D_{k+p} u_{k+p} + \sum_{j=1}^{p_k} \tilde{h}_{k+p,j-1} v_{k+j-1}$$

(16)

we have set

$$C_{k+p} \tilde{A}_{k+p-1}, \ldots, \tilde{A}_1 x_k \approx 0$$

with exact equality assignable. That is,

$$C_{k+p} \tilde{A}_{k+p-1}, \ldots, \tilde{A}_1 x_k = 0$$

in the absence of measurement noise. This leads to the construction of a generalized time-varying ARX (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 TOKID framework. In using this I/O 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 a 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, in which 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 Appendix B, along with an example problem. Considerations of the time-varying deadbeat condition appear sparse (Minamide et al. [13] and Hostetter [14] present some fundamental results on the design of time-varying deadbeat observers), if not completely heretofore unknown in modern literature.

If the repeated experiments (as derived and presented in [9,10]) are performed so as to compute a least-squares solution to the I/O 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

$$y_k = \begin{bmatrix} D_k & \tilde{h}_{k,k-1} & \tilde{h}_{k,k-2} & \ldots & \tilde{h}_{k,k-p} \end{bmatrix} \begin{bmatrix} u_k \\ v_{k-1} \\ v_{k-2} \\ \vdots \\ v_{k-p} \end{bmatrix}$$

(17)

This represents a set of $m$ equations in $m \times (r + p = (r + m)$ unknowns. In contrast to the developments using the generalized system Markov parameters (to relate the I/O data sets, refer to Eq. (8) in the companion papers [9,10] and the references therein for more information), the number of unknowns remains constant in this case. This makes the computation of observer Markov parameters possible in practice, because the number of repeated experiments required to compute these parameters is now constant (derived next) and does
not change with the discrete-time step $t_1$. In fact, it is observed that a minimum of $N_{k+1} = \lfloor \alpha + p \ast (\alpha + m) \rfloor$ experiments are necessary to determine the observer Markov parameters uniquely. From the developments of the subsequent sections, this is the maximum number of repeated experiments one should perform in order to realize the time-varying system models desired from the TVERA. Equation (17), with $N$ repeated experiments, yields:

$$Y_k = [y_k(1) \ldots y_k(N)]$$

$$= [D_k \tilde{h}_{k-1} \tilde{h}_{k-2} \ldots \tilde{h}_{k-p}]$$

$$= \begin{bmatrix} u_k(1) & u_k(2) & \ldots & u_k(N) \\ v_k(1) & v_k(2) & \ldots & v_k(N) \\ \vdots & \vdots & \ddots & \vdots \\ v_{k-p} & v_{k-p+1} & \ldots & v_{k-N} \end{bmatrix}$$

$$= M_k V_k$$

where $V_k$ is the least-squares solution for the generalized observer Markov parameters is given for each time step as

$$M_k = Y_k V_k$$

with $(\cdot)^+$ denotes the pseudoinverse of a matrix [15,16]. The calculation of the system Markov parameters and the 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 a 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

$$\tilde{h}_{k,j-1} = C_k \bar{B}_{k-1}$$

$$= C_k [(B_{k-1} + G_{k-1}D_{k-1}) - G_{k-1}]$$

$$= \tilde{h}_{k,j-1} - \tilde{h}_{k,j-1}$$

where the superscripts in Eqs. (1) and (2) are used to distinguish between the Markov parameter sequences useful to compute the system parameters and the observer gains, respectively. Observe the following manipulation, written as

$$\tilde{h}_{k,j-1} = \tilde{h}_{k,j-1} - \tilde{h}_{k,j-1} = C_k B_{k-1} = \tilde{h}_{k,j-1}$$

A similar expression for Markov parameters with two time steps between them yields:

$$\tilde{h}_{k,j-2} - \tilde{h}_{k,j-2} = C_k \bar{A}_{k-1} \bar{B}_{k-2} - C_k \bar{A}_{k-1} G_{k-2} D_{k-2}$$

$$= C_k \bar{A}_{k-1} (B_{k-2} + G_{k-2} D_{k-2}) - C_k \bar{A}_{k-1} G_{k-2} D_{k-2}$$

$$= C_k \bar{A}_{k-1} B_{k-2}$$

$$= C_k (A_{k-1} + G_{k-1}C_{k-1}) B_{k-2}$$

$$= C_k A_{k-1} B_{k-2} + \tilde{h}_{k,j-2}$$

$$= \tilde{h}_{k,j-2} = \tilde{h}_{k,j-2}$$

This elegant manipulation leads to an expression for the system Markov parameter $h_{k,j-2}$ 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 (unlabeled $h_{k,j}$) from the observer Markov parameters $h_{k,j}^*$. To show this holds in general, consider the induction step with the observer Markov parameters (with $p$ time-step separation) given by

$$\tilde{h}_{k,j-2} - \tilde{h}_{k,j-2} = C_k \bar{A}_{k-1} \bar{A}_{k-2} \ldots \bar{A}_{k-p+1} B_{k-p} + G_{k-p} D_{k-p}$$

$$= C_k \bar{A}_{k-1} \bar{A}_{k-2} \ldots \bar{A}_{k-p+1} B_{k-p}$$

$$= C_k \bar{A}_{k-1} \bar{A}_{k-2} \ldots \bar{A}_{k-p+1} (A_{k-p+1} + G_{k-p+1} C_{k-p+1}) B_{k-p}$$

$$= C_k \bar{A}_{k-1} \bar{A}_{k-2} \ldots \bar{A}_{k-p+1} A_{k-p+1} B_{k-p} + C_k \bar{A}_{k-1} \bar{A}_{k-2} \ldots \bar{A}_{k-p+1} B_{k-p}$$

$$= C_k \bar{A}_{k-1} \bar{A}_{k-2} \ldots \bar{A}_{k-p+1} A_{k-p+1} B_{k-p} + C_k \bar{A}_{k-1} \bar{A}_{k-2} \ldots \bar{A}_{k-p+1} B_{k-p}$$

$$= C_k \bar{A}_{k-1} \bar{A}_{k-2} \ldots \bar{A}_{k-p+1} B_{k-p} = \cdots$$

where $j = k - p$.

This manipulation enables us to write

$$\tilde{h}_{k,j-2} = \tilde{h}_{k,j-2}$$

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

$$h_{k,j-1} = h_{k,j-1} - \tilde{h}_{k,j-1}$$

$$h_{k,j-2} = h_{k,j-2} + \tilde{h}_{k,j-2}$$

$$\vdots$$

$$h_{k,j-p} = \tilde{h}_{k,j-p}$$

where $j = k - p$.

Defining $r_{kj} := \tilde{h}_{kj} - \tilde{h}_{kj}$, we obtain the system of linear equations, relating the system and observer Markov parameters as
We note the striking similarity of this equation to the relation between the observer Markov parameters and the system Markov parameters in the classical OKID algorithm for time-invariant systems (compare the coefficient matrix of Eq. (27) with Eq. 6.8 of Juang [1]).

Considering the expressions for

\[ \hat{h}_{k,k,p} := C_k \hat{A}_{k-1} \ldots \hat{A}_{k-p+1} B_{k-p} \]

and choosing a \( p \) sufficiently large, we have that, owing to the asymptotic stability of the closed loop (including the observer), \( \hat{h}_{k,k,p} \approx 0 \). This fact enables us to establish recursive relationships for the calculation of the system Markov parameters \( h_{k,k-\gamma} \) and \( \forall \gamma > p \). Generalizing Eq. (25) produces:

\[ h_{k,k-\gamma} = h_{k,k-\gamma}^{(1)} - \hat{h}_{k,k-\gamma}^{(2)} D_{k-\gamma} - \sum_{j=1}^{\gamma-1} \hat{h}_{k,j}^{(2)} h_{k,j,k-\gamma} \]

for \( \gamma \leq p \). Then, based on the arguments involving Eq. (17) for the calculation of the generalized observer Markov parameters, all the terms with a time-step separation greater than \( p \) vanish identically, and we obtain the relationship:

\[ h_{k,k-\gamma} = - \sum_{\gamma=1}^{p} \hat{h}_{k,k-\gamma}^{(2)} h_{k,k-\gamma} \]

for \( \gamma > p \).

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

Consider the generalized observer gain Markov parameters defined as

\[ h_{k,i}^{\circ} = \begin{cases} C_k A_{k-1}, A_{k-2}, \ldots, A_{i} + G_i & \forall k > i + 1 \\ C_k G_{k-1} & \forall k < i + 1 \end{cases} \]

We will now derive the relationship between these parameters and the time-varying ARX model coefficients \( \hat{h}_{k,i}^{(2)} \). These parameters will be used in the calculation of the observer gain sequence from the I/O data in the next subsection, a generalization of the time-invariant relations obtained in [1,11], similar to Eq. (27).

From their corresponding definitions, note that

\[ \hat{h}_{k,k-1}^{(2)} = C_k \hat{A}_{k-1} = h_{k,k-1}^{(2)} \]

Similarly,

\[ \hat{h}_{k,k-2}^{(2)} = C_k \hat{A}_{k-1} G_{k-2} = C_k (A_{k-1} + G_{k-1} C_{k-1}) G_{k-2} = h_{k,k-2}^{(2)} + \hat{h}_{k,k-1}^{(2)} h_{k-1,k-2} \]

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

\[ \hat{h}_{k,k-p}^{(2)} = C_k \hat{A}_{k-1} \ldots \hat{A}_{k-p+1} G_{k-p} \]

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:

\[ h_{k,k-\gamma}^{\circ} = - \sum_{\gamma=1}^{p} \hat{h}_{k,k-\gamma}^{(2)} h_{k,k-\gamma}^{\circ} \]

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

\[
\begin{bmatrix}
I_m & \hat{h}_{k-1,k}^{(2)} & \hat{h}_{k-2,k}^{(2)} & \ldots & \hat{h}_{k-p+1,k}^{(2)} \\
0 & I_m & \hat{h}_{k-1,k-1}^{(2)} & \ldots & \hat{h}_{k-p+1,k-1}^{(2)} \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & \ldots & I_m \\
\hat{h}_{k,k-1}^{(2)} & \hat{h}_{k,k-2}^{(2)} & \ldots & \hat{h}_{k,k-p+1}^{(2)} \\
0 & \hat{h}_{k-1,k-1}^{(2)} & \ldots & \hat{h}_{k-1,k-p+1}^{(2)} \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & \ldots & \hat{h}_{k-1,k-p+1}^{(2)} \\
0 & 0 & \ldots & \hat{h}_{k,p+1,k-p+1} \\
\end{bmatrix} \times
\begin{bmatrix}
h_{k,k-1}^{(2)} \\
h_{k,k-2}^{(2)} \\
h_{k,k-3}^{(2)} \\
\vdots \\
h_{k,k-p+1}^{(2)} \\
\end{bmatrix} = \begin{bmatrix}
h_{k,k-1,1}^{\circ} \\
h_{k,k-2,1}^{\circ} \\
h_{k,k-3,1}^{\circ} \\
\vdots \\
h_{k,k-p+1,1}^{\circ} \\
\end{bmatrix}
\]

C. Calculation of the Time-Varying Observer Gain Sequence

From the definition of the observer gain Markov parameters [Eq. (30)], we can stack the first few parameters in a tall matrix and observe that
such that a least-squares solution for the gain matrix at each time step is given by

$$G_k = O^{(m)}_{k+1} P_{k+1}$$

(38)

However, from the developments of the companion paper, we find that it is indeed impossible to determine the observability Gramian in the true coordinate system [9], as suggested by Eq. (38). The computed observability Gramian is, in general, in a time-varying and unknown coordinate system denoted by $O^{(m)}$ at the time step $t_{k+1}$. We will now show that the gain computed from this time-varying observability Gramian (computed) will be consistent with the time-varying coordinates of the plant model computed by the TVERA presented in the companion paper. Therefore, upon using the computed observability Gramian (in its own time-varying coordinate system) and proceeding with the gain calculation, as indicated by Eq. (38), we arrive at a consistent computed gain matrix. Given a transformation matrix $T_{k+1}$,

$$P_{k+1} = O^{(m)}_{k+1} G_k = O^{(m)}_{k+1} T_{k+1}^T G_k = \hat{O}^{(m)}_{k+1} T_{k+1}^T G_k = \hat{O}^{(m)}_{k+1} \hat{G}_k$$

(39)

such that

$$\hat{G}_k = T_{k+1}^T G_k = (\hat{O}^{(m)}_{k+1}) P_{k+1}$$

(40)

to determine the gain matrix.

Therefore, with no explicit intervention by the analyst, the realized gains are automatically in the right coordinate system for producing the appropriate TOKID 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 Gramians 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 and control influence and the measurement sensitivity matrices in the time-varying coordinate systems, to be derived as (refer to companion paper [9])

$$\hat{A}_k = T_{k+1}^{-1} A_k T_k; \quad \hat{B}_k = T_{k+1}^{-1} B_k; \quad \hat{C}_k = C_k T_k$$

(41)

The TOKID closed-loop system matrix, with the realized gain matrix sequence, is seen to be consistently given as

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

(42)

in a kinematically similar fashion to the true TOKID closed loop. The nature of the computed stabilizing (deadbeat or near-deadbeat) gain sequence is 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 [9].

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 address the important issue: variable-order GTV-ARX model fitting of I/O data and what it all means. Insight is also obtained as to what happens in the presence of measurement noise. In the practical situation, for which there is the presence of the process and the measurement noise in the data, the GTV-ARX model becomes a moving average model that can be termed as the GTV autoregressive moving average with exogenous input (GTV-ARMAX) 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: the authors limit the discussions to qualitative relations.

The Kalman-filter equations for a truth model given in Eq. (A1) are given by

$$\dot{x}_{k+1} = A_k \dot{x}_k + B_k u_k$$

(43)

or

$$\dot{\hat{x}}_{k+1} = A_k [I - K_k C_k] \hat{x}_k + B_k u_k + A_k K_k y_k$$

(44)

together with the propagated output equation

$$\hat{y}_k = C_k \hat{x}_k + D_k u_k$$

(45)

where the gain $K_k$ is optimal [see expression in Eq. (A16)]. 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. (A16) for an expression for the optimal gain]. From an I/O standpoint, the innovations approach provides the most insight for analysis and is used in this section. Using the definition of the innovations process $e_k := y_k - \hat{y}_k$, the measurement equation of the estimator [shown in Eq. (45)] can be written in favor of the system outputs, as given by

$$y_k = C_k \hat{x}_k + D_k u_k + e_k$$

(46)

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

$$\hat{x}_{k+1} = A_k [I - K_k C_k] \hat{x}_k + B_k u_k + A_k K_k y_k = \tilde{A}_k \hat{x}_k + \tilde{B}_k y_k$$

(47)

with the definitions

$$\tilde{A}_k = A_k [I - K_k C_k]; \quad \tilde{B}_k = [B_k \quad A_k K_k]; \quad \nu_k = [u_k \quad y_k]$$

(48)

Notice the structural similarity in the layout of the rearranged equations to the TOKID equations in Sec. III. This rearrangement helps in making comparisons and observations as to the conditions in which we actually manage to obtain the Kalman-filter gain sequence (or otherwise).

Starting from the initial condition, the I/O relation of the Kalman-filter equations can be written as
\[ \begin{align*}
y_0 &= C_0 \hat{x}_0 + D_0 u_0 + \epsilon_0; \\
y_1 &= C_1 \hat{A}_0 \hat{x}_0 + D_1 u_1 + C_1 \hat{B}_1 y_0 + \epsilon_1; \\
y_2 &= C_2 \hat{A}_1 \hat{A}_0 \hat{x}_0 + D_2 u_2 + C_2 \hat{B}_1 y_1 + C_2 \hat{B}_2 \hat{B}_1 y_0 + \epsilon_2 \\
& \vdots \\
y_p &= C_p \hat{A}_{p-1} \ldots \hat{A}_0 \hat{x}_0 + D_p u_p + \sum_{j=1}^{p-1} \hat{h}_{p,p-j} y_{p-j} + \epsilon_p \\
& \vdots \\
\end{align*} \]

suggesting the general relationship

\[ \begin{align*}
y_{k+p} &= C_{k+p} \hat{A}_{k+p-1} \ldots \hat{A}_0 \hat{x}_0 + D_{k+p} u_{k+p} \\
& + \sum_{j=1}^{k+p-1} \hat{h}_{k+p,k+p-j} y_{k+p-j} + \epsilon_{k+p} \\
\end{align*} \]

with the Kalman-filter Markov parameters \( \hat{h}_{k,i} \) being defined by

\[ \hat{h}_{k,i}^o = \begin{cases} 
C_i \hat{A}_{i-1} \ldots \hat{A}_0 \hat{x}_0 + D_{k+i} u_k & \forall \ k > i + 1 \\
C_i \hat{B}_{i-1} & k = i + 1 \\
0 & \forall \ k < i + 1
\end{cases} \]

Comparing Eqs. (14) and (50), we conclude that their I/O 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 \( \epsilon_k = 0 \) and \( \forall \ k > p \). 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 Appendix A.

However, we proceeded to enforce the \( p \) (in general \( p_k \)) term dependence in Eq. (16), using the additional freedom obtained because of 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 the time-varying deadbeat observer notions set up in this paper (following the classical developments of Minamide et al. [13] and Hostetter [14], as discussed in Appendix B). Notice that the Kalman-filter equations are, in general, not truncated to the first \( p \) 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 I/O behavior of the true Kalman filter in Eq. (50). Observe that Kalman gains can indeed be constructed so as to obtain matching truncated representations of the GTV-ARX (more precisely GTV-ARMAX) model, as in Eq. (16), via the appropriate choice of the tuning parameters \( P_0 \) and \( Q_k \), for measurement and process noises. In the GTV-ARMAX parlance, using a lower order for \( P \) (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 GTV-ARX and ARMAX 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 the statistical learning theory between the ignorance of the model for the dynamical system and the incorporation of new information from the 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

Consider the same system, as presented in an example of the companion paper [9]. It has an oscillatory nature and does not have a stable origin. In the 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 [17,18]. However, because the system under consideration is not autonomous, the origin is said to be unstable in the sense of Lyapunov, as described in [15,19]. A separate classification has been provided in the theory of nonlinear systems for systems with an origin of this type. That is called orbital stability or stability, in the sense of Poincaré (cf., Meirovitch [20]). We follow the convention of Lyapunov and term the system under consideration as unstable. In this case, the plant system matrix was calculated as.

![Fig. 1 Open-loop vs TOKID closed-loop pole locations [10 repeated experiments (i.e., minimum number)]. (TV denotes time-varying.)](image-url)
where the matrix is given by

\[
A_k = \exp[A, \Delta t]; \quad B_k = \begin{bmatrix}
1 & 0 \\
1 & -1 \\
-1 & 0
\end{bmatrix}; \quad C_k = \begin{bmatrix}
1 & 0 & 1 & 0.2 \\
1 & -1 & 0 & -0.5
\end{bmatrix}; \quad D_k = 0.1 \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}
\]

with

\[
A_k = \begin{bmatrix}
0_{2 \times 2} & I_{2 \times 2} \\
-K_k & 0_{2 \times 2}
\end{bmatrix}, \quad K_k = \begin{bmatrix}
4 + 3 \tau_k & 1 \\
1 & 7 + 3 \tau_k'
\end{bmatrix}
\]

and \(\tau_k\) and \(\tau_k'\) are defined as \(\tau_k = \sin(10t_k)\) and \(\tau_k' = \cos(10t_k)\). The TOKID algorithm, as described in the main body of the paper, was...
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 the system model sequence for all the time steps for which experimental data is available. We demonstrate the computation of the deadbeat-like observer, for which the smallest order for the GTV-ARX model is chosen throughout the time history of the identification process. Appendix B details the definition of the time-varying deadbeat observer for the convenience of the readers, along with a representative closed-loop sequence result using this example problem.

In this case, we were able to realize an asymptotically stable closed loop for the observer equation with TOKID. In fact, two of the closed-loop eigenvalues could be assigned to zero. The time history of the open-loop and closed-loop eigenvalues, as viewed from the coordinate system of the initial condition response decomposition, is plotted in Fig. 1.

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 TVERA algorithm. These are computed using the recursive relationships. The truncation of the I/O relationship with the observer in the loop is exact in theory. Numerical errors still play a role. The worst-case error, although it is sufficiently small, is incurred in the situation when a minimum number of experiments is performed. The magnitude of this error is plotted in Fig. 2 as the matrix 1 norm of the system Markov parameter error sequence (denoted by \(\|\cdot\|_1\)). The comparison in this figure is made with an error in the system Markov parameters computed from the full I/O relationship without an observer. Performing a larger number of experiments, in general, leads to better accuracy, as shown in Fig. 3. Note that more experiments give a better condition number for making the pseudoinverse of the matrix \(V_k\), shown in Eq. (18). The accuracy is also improved by retaining a larger number of terms (per time step) in the I/O map.

The error incurred in the system Markov parameter computation is directly reflected in the output error between the computed and true

---

**Fig. 4** Output error comparison between true and identified plant models (10 repeated experiments).

**Fig. 5** Output error comparison between true and identified plant models (22 repeated experiments).
system response to the 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 plots for Figs. 2 and 3 are shown in Figs. 4 and 5.

Because the considered system is unstable (oscillatory) in nature, the initial condition response was used to check the nature of the 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 Fig. 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 steps to zero error response. The analyst should count the two time steps in question after the first few time steps to enable the unique identification of the observer gains (a total of four time steps for this example can be seen in the plots). 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, a near-deadbeat observer was realized. 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 Fig. 7.

VII. Conclusions

This paper provides an algorithm for efficient computation of system Markov parameters for use in time-varying system identification. An observer is inserted in the I/O relations, and this leads
to effective utilization of the data in computation of the system Markov parameters. As a by-product, one obtains an observer gain sequence in the same coordinate system as the system models realized by the time-varying system identification algorithm. This observer is further found to be a time-varying deadbeat observer of the system. Relationship of this observer with a Kalman filter is detailed from an I/O standpoint. It is shown that the flexibility of a variable-order moving average model realized in the TORKID 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: Review of the Structure and Properties of the Kalman Filter

I. Linear Estimators of the Kalman Type: 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 [21,22]) using the innovations approach propounded by Khalil [23] and Mehra [24]. The most commonly used truth model for the linear time-varying filtering problem is given by

\[ x_{k+1} = A_k x_k + B_k u_k + \Gamma_k w_k \]  \hspace{1cm} (A1)

together with the measurement equations given by

\[ y_k = C_k x_k + D_k u_k + v_k \]  \hspace{1cm} (A2)

The process noise sequence is assumed to be a Gaussian random sequence with zero mean \( E(w) = 0 \) \( \forall \) \( i \) and a variance sequence \( E(w, w') = \Omega_{ij} \forall i, j \), having an uncorrelated profile in time and no correlation with the measurement noise sequence \( E(w, v') = 0 \) \( \forall \) \( i, j \). Similarly, the measurement noise sequence is assumed to be a zero mean Gaussian random vector with a covariance sequence given by \( E(v, v') = R_{ij} \). The Kronecker delta is denoted as \( \delta_{ij} = 1 \) \( \forall \) \( i = j \) and \( = 0 \) \( \forall \) \( i \neq j \), 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 [11]):

\[ \hat{x}_k = \hat{x}_k^0 + K_k [y_k - \tilde{y}_k] = \hat{x}_k^0 + K_k e_k \]  \hspace{1cm} (A3)

where the term \( e_k := y_k - \hat{x}_k \) 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

\[ \hat{x}_{k+1} = A_k \hat{x}_k + B_k u_k = A_k [I - K_k C_k] \hat{x}_k + B_k u_k + A_k K_k y_k \]  \hspace{1cm} (A4)

and

\[ \tilde{y}_k = C_k \hat{x}_k + D_k u_k \]  \hspace{1cm} (A5)

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

\[ e_k = C_k e_k + v_k \]  \hspace{1cm} (A6)

whereas the propagation of the estimation error dynamics (estimator in the loop, similar to the TORKID developments of the paper) is governed by

\[ e_{k+1} = A_k [I - K_k C_k] e_k - A_k K_k v_k + \Gamma_k w_k \]  \hspace{1cm} (A7)

Defining the uncertainty associated by the state estimation process, quantified by the covariance to be \( P_k := E(e_k e_k') \), the covariance propagation equations are given by

\[ P_{k+1} = \tilde{A}_k P_k \tilde{A}_k' + A_k K_k R_k A_k' + \Gamma_k \Omega_k \Gamma_k' \]  \hspace{1cm} (A8)

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, called the innovations property, is the conceptual basis for projection methods [23] in a Hilbert space setting. As a consequence of this property, we have the following condition.

If the gain in the observer equation is optimal, then the resulting recursions should render the innovations process orthogonal (uncorrelated) with respect to all other terms of the sequence. For any time step \( t \) and a time step \( t_{k+1} = (k + 1) \) steps behind the \( k \)th step, we have that

\[ E(e_k e_{k+1}') = 0 \]  \hspace{1cm} (A9)

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

\[ E(e_k e_{k+1}') = C_k E(e_k e_k') C_k' + E(v_k v_{k+1}') = 0 \]  \hspace{1cm} (A10)

where the two terms \( E(e_k e_{k+1}') = E(v_k v_{k+1}') = 0 \) 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. (A10) becomes

\[ E(e_k e_k') = C_k E(e_k e_k') C_k' + E(v_k v_k') = C_k P_k C_k' + R_k \]  \hspace{1cm} (A11)

Applying the evolution equation for the estimation error dynamics for \( k \) time steps backward in time from \( t \), we have that

\[ e_k = \tilde{A}_{k-1} \tilde{A}_{k-2} \ldots \tilde{A}_{k-n} e_k \]  \hspace{1cm} (A12)

where \( \tilde{A}_k := A_k - K_k C_k \) and \( \tilde{A}_k \) is defined by Eqs. (A11) and (A12) on both sides with \( e_k' \) and \( v_k' \) taking the expectation operator

\[ E(e_k e_k') = \tilde{A}_{k-1} \tilde{A}_{k-2} \ldots \tilde{A}_{k-n} \tilde{A}_{k-n} E(e_k e_k') \]  \hspace{1cm} (A13)

\[ E(v_k v_k') = \tilde{A}_{k-1} \tilde{A}_{k-2} \ldots \tilde{A}_{k-n} \tilde{A}_{k-n} E(v_k v_k') \]  \hspace{1cm} (A14)

Substituting Eqs. (A13) and (A14) into the expression for the inner product shown in Eq. (A10), we arrive at the expressions for the Kalman gain sequence as a function of the statistics of the state estimation error dynamics for all time instants up to \( t_{k+1} \) as

\[ E(e_k e_k') = C_k \tilde{A}_{k-1} \tilde{A}_{k-2} \ldots \tilde{A}_{k-n} \tilde{A}_{k-n} P_{k-1} C_{k-1}' \]  \hspace{1cm} (A15)

where the definition \( \tilde{A}_k := A_k - K_k C_k \) has been used to derive the third equality. Equation (A15) is necessary to hold for all Kalman type estimators with the familiar update structure, \( \forall \ k > 0 \):

\[ K_{k+1} = P_k C_{k-1}' C_k (R_{k-1} + C_k P_k C_{k-1})^{-1} \]  \hspace{1cm} (A16)
because of the innovations property involved. The qualitative relationship between the identified observer realized from the TOKID calculations (GTV-ARX model) and the classical Kalman filter is explained in the main body of the paper.

Appendix B: Time-Varying Deadbeat Observers

I. 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 with its linear time-invariant counterpart. The topic of extension of the deadbeat observer design to the time-varying systems has not been pursued aggressively in the literature, and only scattered results exist in this context. Minamide et al. [13] developed a similar definition of the time-varying deadbeat condition and presented 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 the I/O data.

First, we recall the definition of a deadbeat observer in the 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 [18], if a linear discrete-time dynamical system is characterized by evolution equations given by

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + B\mathbf{u}_k \quad (B1)$$

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

$$\mathbf{y}_k = C\mathbf{x}_k + D\mathbf{u}_k \quad (B2)$$

where the usual assumptions on the dimensionality of the state space are made, $\mathbf{x}_k \in \mathbb{R}^n$, $\mathbf{y}_k \in \mathbb{R}^m$, $\mathbf{u}_k \in \mathbb{R}^p$, and $A, C,$ and $B$ 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):

$$(A + GC)^p = 0_{m \times n} \quad (B3)$$

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

II. Example of a Time-Invariant Deadbeat Observer

Let us consider the following simple linear time-invariant example to show the ideas:

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix} \quad C = \begin{bmatrix} 0 & 1 \end{bmatrix} \quad (B4)$$

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

$$(A + GC)^2 = \begin{bmatrix} 1 + g_1 & g_1(3 + g_2) \\ 3 + g_2 & g_1(2 + g_2)^2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (B5)$$

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

$$(A + GC) = \begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix} \quad (B6)$$

which is a defective (repeated roots at the origin) and nilpotent matrix. Therefore, the deadbeat observer is the fastest observer that could possibly be achieved, because 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), because 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 et al. solution [13]) and when only experimental data are 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 condition appears to have been naturally made. Recall [from Eq. (16)] that in constructing the 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

$$(A_{k+p-1} + G_{k+p-1}C_{k+p-1}), (A_{k+p-2} + G_{k+p-2}C_{k+p-2}), \ldots, (A_k + G_kC_k) = 0_{m \times n} \quad (B7)$$

for every $k$, where $p$ is the smallest integer, such that the condition is satisfied.

We now illustrate this definition using an example problem.

III. Example of a Time-Varying Deadbeat Observer

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

$$A_k = \exp[A_k \Delta t]; \quad B_k = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}; \quad C_k = \begin{bmatrix} 1 & 0 & 1 & 0.2 \\ 1 & -1 & 0 & -0.5 \end{bmatrix}; \quad D_k = 0.1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (B8)$$

where the matrix is given by

$$A_k = \begin{bmatrix} 0 & I_{2 \times 2} \\ -K_k & 0_{2 \times 2} \end{bmatrix} \quad (B9)$$

with

$$K_k = \begin{bmatrix} 4 + 3\tau_k & 1 \\ 1 & 7 + 3\tau_k \end{bmatrix}$$

and $\tau_k$ and $\tau_k'$ are defined as $\tau_k = \sin(10t_k)$ and $\tau_k' := \cos(10t_k)$. Clearly, because $m = 2$ and $n = 4$ 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 system matrix and its eigenvalues are [where $\lambda(M)$ denotes the vector of eigenvalues of the matrix $M$] computed as
 whereas the closed-loop system matrix (and its eigenvalues) for the previous time step is calculated as

\[
A_{35} + GC_{35} = \begin{bmatrix}
-1.3884 & 0.3040 & -0.1597 & -0.2623 \\
-0.2505 & 0.8103 & -2.7307 & 0.0933 \\
0.4456 & 0.1297 & -0.7617 & 0.1265 \\
7.2102 & -0.9015 & -1.5928 & 1.4887 \\
\end{bmatrix},
\]

\[
\lambda(A_{35} + GC_{35}) = \begin{bmatrix}
0.14986 \\
-0.0093564 \\
2.4455 \times 10^{-11} \\
-9.2301 \times 10^{-14} \\
\end{bmatrix} \tag{B11}
\]

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

\[
A_{37} + GC_{37} = \begin{bmatrix}
-1.8156 & 0.3184 & -0.3186 & -0.2933 \\
-1.3196 & 0.8332 & -2.9451 & 0.0113 \\
0.5186 & 0.0855 & -0.7044 & 0.1496 \\
9.7693 & -1.309 & -0.1074 & 1.729 \\
\end{bmatrix},
\]

\[
\lambda(A_{37} + GC_{37}) = \begin{bmatrix}
0.086117 \\
-0.043863 \\
1.599 \times 10^{-12} \\
2.5635 \times 10^{-13} \\
\end{bmatrix} \tag{B12}
\]

Although clearly, each of the closed-loop member sequence \(A_{35,36,37}\) has only two zero eigenvalues (individually non-deadbeat according to the time-invariant definition, because all closed-loop poles are not placed at the origin), let us now consider the product matrices:

\[
(A_{37} + GC_{37})(A_{36} + GC_{36}) = 10^{-12} \times \begin{bmatrix}
-0.12568 & -0.007105 & 0.01476 & 0.003664 \\
0.14537 & -0.019406 & 0.03852 & 0.014513 \\
0.07660 & 0.00247 & 0.00483 & -0.003081 \\
-0.24691 & 0.26557 & -0.29265 & -0.19860 \\
\end{bmatrix} \tag{B13}
\]

and

\[
(A_{36} + GC_{36})(A_{35} + GC_{35}) = 10^{-12} \times \begin{bmatrix}
-0.0591 & 0.00617 & -0.001388 & -0.005052 \\
-0.1888 & 0.10671 & -0.098574 & -0.099892 \\
-0.00311 & 0.021205 & -0.020995 & -0.022121 \\
-0.55067 & 0.20917 & -0.15277 & -0.18829 \\
\end{bmatrix} \tag{B14}
\]

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 iterate 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 TOKID (the 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 TOKID is developed directly in the reduced dimensional I/O space, whereas the schemes developed to compute the gain sequences in the paper by Minamide et al. [13], which is quite similar to the method outlined by Hostetter [14], are based on projections of the state space onto the outputs.

### Acknowledgments

The authors wish to acknowledge the support of the Texas Institute of Intelligent Bio Nano Materials and Structures for Aerospace Vehicles 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 NASA.

### References


