Optimal State Estimators for Linear Systems with Unknown Inputs

Shreyas Sundaram and Christoforos N. Hadjicostis

***Abstract*— We present a method for constructing linear minimum-variance unbiased state estimators for discrete-time linear stochastic systems with unknown inputs. Our design provides a characterization of estimators with delay, which eases the established necessary conditions for existence of unbiased estimators with zero-delay. A consequence of using delayed estimators is that the noise affecting the system becomes correlated with the estimation error. We handle this correlation by increasing the dimension of the estimator appropriately.**

I. INTRODUCTION

In practice, it is often the case that a dynamic system can be modeled as having unknown inputs. When the unknown inputs are assumed to have some defined structure (such as being bounded in norm), robust filtering techniques have been developed to estimate the state of the system [1]. Researchers have also proposed methods to optimally estimate the system state when the unknown inputs are completely unconstrained [2], [3], [4], [5]. These latter works revealed that the system must satisfy certain strict conditions in order to reject the unknown inputs. Furthermore, it was shown in [5] that this decoupling might force the system noise to become correlated with the estimation error (i.e., the noise behaves as if it is colored). In [4], Saberi et al. showed through a geometric approach that allowing delays in the estimator can relax the necessary conditions for optimal estimation. Delayed estimators were also studied in [6], but the design procedure outlined in that paper did not fully utilize the freedom in the estimator gain matrix, and ignored the correlation between the noise and error [7].

In this paper, we study delayed estimators for linear systems with (unconstrained) unknown inputs, and develop a method to construct optimal linear estimators. More specifically, our goal is to estimate the entire system state through linear recursive estimators that minimize the mean square estimation error. In addition, we require that the estimator be unbiased (i.e., the expected value of the estimation error must be zero). Our approach is an extension of earlier work on observers for deterministic linear systems with unknown inputs [8], [9]. The fact that we incorporate delays in our design procedure allows us to construct optimal estimators for a much larger class of systems than that considered by [2], [3], [5]. Our approach is purely algebraic and is more direct than the method in [4], which first transforms the estimator into a dual system and then uses techniques from H_2-optimal control in order to obtain the estimator parameters. Finally, our design procedure avoids the problem of colored noise described in [5] by increasing the dimension of the estimator appropriately, and makes full use of the freedom in the gain matrix, which produces better results than the method in [6].

This material is based upon work supported in part by the National Science Foundation under NSF Career Award 0092696 and NSF EPNES Award 0224729, and in part by the Air Force Office of Scientific Research under URI Award No F49620-01-1-0365URI. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of NSF or AFOSR.

The authors are with the Coordinated Science Laboratory and the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, 61801, USA. E-mail {ssundarm, chadjic}@uiuc.edu

II. PRELIMINARIES

Consider a discrete-time stochastic linear system $\mathcal{S}$ of the form

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k \\ y_k &= Cx_k + Du_k + v_k \ , \end{aligned} \tag{1}$$

with state vector $x \in \mathbb{R}^n$, unknown input $u \in \mathbb{R}^m$, output $y \in \mathbb{R}^p$, and system matrices (A, B, C, D) of appropriate dimensions. The noise processes w_k and v_k are assumed to be uncorrelated, white, and zero mean, with covariance matrices Q_k and R_k respectively. Note that we omit known inputs in the above equations for clarity of development. We also assume without loss of generality that the matrix $\left[\begin{smallmatrix} B \\ D \end{smallmatrix}\right]$ is full column rank. This assumption can always be enforced by an appropriate transformation and renaming of the unknown input signals. Note that no assumptions are made on the distribution of the unknown inputs (and that is the reason for distinguishing between u_k and the noise v_k and w_k).

The response of system (1) over $\alpha + 1$ time units is given by

$$\begin{aligned} \underbrace{\begin{bmatrix} y_k \\ y_{k+1} \\ \vdots \\ y_{k+\alpha} \end{bmatrix}}_{Y_{k:k+\alpha}} &= \underbrace{\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{\alpha} \end{bmatrix}}_{\Theta_\alpha} x_k + \underbrace{\begin{bmatrix} v_k \\ v_{k+1} \\ \vdots \\ v_{k+\alpha} \end{bmatrix}}_{V_{k:k+\alpha}} \\ &+ \underbrace{\begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{\alpha-1}B & CA^{\alpha-2}B & \cdots & D \end{bmatrix}}_{M_\alpha} \underbrace{\begin{bmatrix} u_k \\ u_{k+1} \\ \vdots \\ u_{k+\alpha} \end{bmatrix}}_{U_{k:k+\alpha}} \\ &+ \underbrace{\begin{bmatrix} 0 & 0 & \cdots & 0 \\ C & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{\alpha-1} & CA^{\alpha-2} & \cdots & C \end{bmatrix}}_{M_{w,\alpha}} \underbrace{\begin{bmatrix} w_k \\ w_{k+1} \\ \vdots \\ w_{k+\alpha-1} \end{bmatrix}}_{W_{k:k+\alpha-1}} . \end{aligned} \tag{2}$$

The matrices Θ_α and M_α in the above equation can be expressed in a variety of ways. We will be using the following identities in our derivations:

$$\Theta_\alpha = \begin{bmatrix} C \\ \Theta_{\alpha-1}A \end{bmatrix} = \begin{bmatrix} \Theta_{\alpha-1} \\ CA^\alpha \end{bmatrix} \quad , \tag{3}$$

$$M_\alpha = \begin{bmatrix} D & 0 \\ \Theta_{\alpha-1}B & M_{\alpha-1} \end{bmatrix} = \begin{bmatrix} M_{\alpha-1} & 0 \\ C\zeta_{\alpha-1} & D \end{bmatrix} \quad , \tag{4}$$

where $\zeta_{\alpha-1} \equiv \begin{bmatrix} A^{\alpha-1}B & A^{\alpha-2}B & \cdots & B \end{bmatrix}$.

In the rest of the paper, we will be using $E\{\cdot\}$ to denote the expected value of a stochastic parameter. The notation I_r represents the $r \times r$ identity matrix, and $(\cdot)^T$ indicates matrix transpose. We are now in place to proceed with the construction of an estimator for the states in $\mathcal{S}$.

III. Unbiased Estimation

Consider an estimator of the form

$$z_{k+1} = Az_k + K_k \left(Y_{k:k+\alpha} - \Theta_\alpha z_k \right) \quad , \tag{5}$$

where the nonnegative integer α is the estimator delay, and the matrix K_k is chosen to (i) make the estimator unbiased and (ii) minimize the mean square error between z_{k+1} and x_{k+1}. Note that for $\alpha = 0$, the above estimator is in the form of the optimal estimator for systems with no unknown inputs [10]. Using (2), we obtain the estimation error as

$$\begin{aligned} e_{k+1} &\equiv z_{k+1} - x_{k+1} \\ &= \left(A - K_k\Theta_\alpha \right) z_k + K_k Y_{k:k+\alpha} - Ax_k - Bu_k - w_k \\ &= \left(A - K_k\Theta_\alpha \right) e_k + K_k V_{k:k+\alpha} + K_k M_\alpha U_{k:k+\alpha} \\ &\quad - Bu_k + K_k M_{w,\alpha} W_{k:k+\alpha-1} - w_k \quad . \end{aligned} \tag{6}$$

In order for the estimator to be unbiased (i.e., $E\{e_k\} = 0$ for all k, regardless of the values of the unknown inputs), we require that

$$K_k M_\alpha = \begin{bmatrix} B & 0 & \cdots & 0 \end{bmatrix} \quad . \tag{7}$$

The solvability of the above condition is given by the following theorem.

Theorem 1: There exists a matrix K_k satisfying (7) if and only if

$$\text{rank}\left[M_\alpha \right] - \text{rank}\left[M_{\alpha-1} \right] = m \quad . \tag{8}$$

Proof: There exists a K_k satisfying (7) if and only if the row space of the matrix $R \equiv \begin{bmatrix} B & 0 & \cdots & 0 \end{bmatrix}$ is in the space spanned by the rows of M_α. This is equivalent to the condition rank $\left[\begin{smallmatrix} M_\alpha \\ R \end{smallmatrix} \right] = \text{rank}\left[M_\alpha \right]$. Using (4) we get

$$\text{rank} \begin{bmatrix} M_\alpha \\ R \end{bmatrix} = \text{rank} \begin{bmatrix} D & 0 \\ \Theta_{\alpha-1}B & M_{\alpha-1} \\ B & 0 \end{bmatrix} \quad .$$

By our assumption that the matrix $\left[\begin{smallmatrix} B \\ D \end{smallmatrix} \right]$ has full column rank, we get rank $\left[\begin{smallmatrix} M_\alpha \\ R \end{smallmatrix} \right] = m + \text{rank}\left[M_{\alpha-1} \right]$, thereby completing the proof. ■

The result in the above theorem was also obtained in [6] for the specific case of $D = 0$. Note that (8) is the condition for inversion of the inputs with known initial state, as given in [11]. If we set $\alpha = 1$, condition (8) becomes

$$\text{rank} \begin{bmatrix} D & 0 \\ CB & D \end{bmatrix} = m + \text{rank}\left[D \right] \quad ,$$

which is the well known necessary condition for unknown-input observers and unbiased estimators that estimate x_{k+1} based on y_k and y_{k+1} [3], [5]. This is a fairly strict condition, and demonstrates the utility of a delayed estimator. When designing such an estimator, one can start with $\alpha = 0$ and increase α until a value is found that satisfies (8). An upper bound on α is provided by the following theorem.

Theorem 2: Let q be the dimension of the nullspace of D. Then the delay of the unbiased estimator (5) will not exceed $\alpha = n - q + 1$ time-steps. If (8) is not satisfied for $\alpha = n - q + 1$, then unbiased estimation of all the states is not possible with an estimator of the form given in (5).

The proof of the above theorem is immediately obtained by making use of the following result from [12], which considered the problem of system invertibility.

Lemma 1: Let q be the dimension of the nullspace of D. Then there exists an α satisfying (8) if and only if rank $\left[M_{n-q+1} \right] - \text{rank}\left[M_{n-q} \right] = m$.

It is apparent from (6) that for $\alpha > 0$, the error will generally be a function of multiple time-samples of the noise processes w_k and v_k. In other words, the noise becomes colored from the perspective of the error. Darouach et al. studied this situation for $\alpha \in \{0, 1\}$ in [5], and proposed certain strict conditions for the existence of unbiased optimal estimators with dimension n. Consequently, the estimators proposed in that paper can only be applied to a restricted class of systems. In the study of Kalman filters for systems with no unknown inputs, it has been shown that colored noise can be handled by increasing the dimension of the estimator [10]. In the next section, we will apply this technique to construct an optimal estimator for the system in (1). The resulting estimator can be applied to a much larger class of systems than the estimators presented in [5], [2], [3].

IV. Optimal Estimator

If Theorem 1 is satisfied for $\alpha = 0$, the noise in (6) will not be colored, and an optimal estimator of dimension n can be constructed. The resulting estimator is called an *optimal predictor* in [5], and can be found by using the method in that paper. To construct an optimal estimator for $\alpha > 0$, we will rewrite system $\mathcal{S}$ to obtain a new system $\bar{\mathcal{S}}$ given by

$$\begin{aligned} \bar{x}_{k+1} &= \bar{A}\bar{x}_k + \bar{B}u_k + \bar{B}_n n_k \\ y_k &= \bar{C}\bar{x}_k + Du_k \quad , \end{aligned} \tag{9}$$

where

$$\bar{x}_k = \left[\begin{array}{c} x_k \\ W_{k:k+\alpha-2} \\ \hline V_{k:k+\alpha-1} \end{array} \right] \quad , \quad n_k = \begin{bmatrix} w_{k+\alpha-1} \\ v_{k+\alpha} \end{bmatrix} \quad ,$$

$$\bar{B} = \left[\begin{array}{ccccc|ccccc} B^T & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{array} \right]^T \quad ,$$

$$\bar{B}_n = \left[\begin{array}{ccccc|ccccc} 0 & 0 & \cdots & 0 & I_n & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & I_p \end{array} \right]^T \quad ,$$

$$\bar{A} = \left[\begin{array}{ccccc|ccccc} A & I_n & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & I_n & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I_n & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ \hline 0 & 0 & 0 & \cdots & 0 & 0 & I_p & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & I_p & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & I_p \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \end{array}\right]$$

$$\bar{C} = \left[\begin{array}{ccccc|ccccc} C & 0 & \cdots & 0 & 0 & I_p & 0 & \cdots & 0 & 0 \end{array}\right] .$$

Note that the state vector in this new system has dimension $\bar{n} = \alpha(n+p)$ for $\alpha > 0$. From (2), the output of this augmented system over $\alpha+1$ time-steps is given by

$$Y_{k:k+\alpha} = \underbrace{\left[\begin{array}{cccc|cccc} C & 0 & \cdots & 0 & I_p & 0 & \cdots & 0 \\ CA & C & \cdots & 0 & 0 & I_p & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & 0 \\ CA^{\alpha-1} & CA^{\alpha-2} & \cdots & C & 0 & 0 & \cdots & I_p \\ CA^{\alpha} & CA^{\alpha-1} & \cdots & CA & 0 & 0 & \cdots & 0 \end{array}\right]}_{\bar{\Theta}_\alpha} \bar{x}_k + M_\alpha U_{k:k+\alpha} + \underbrace{\begin{bmatrix} 0 & 0 \\ C & I_p \end{bmatrix}}_{M_{n,\alpha}} n_k . \quad (10)$$

We notice that a portion of this equation is independent of the noise vector n_k, and so we can obtain some linear functions of the state vector $\bar{x}_k$ directly from the output. We will use the following definition and theorem to characterize these linear functions.

Definition 1 (Rank-d Linear Functional): Let Γ be a $d \times \bar{n}$ matrix with rank d. Then the quantity $\Gamma\bar{x}_k$ will be termed a rank-d linear functional of the state vector $\bar{x}_k$.

Theorem 3: For system (9) with response over $\alpha+1$ time-steps given by (10), let β_α be the dimension of the left nullspace of $M_{\alpha-1}$. Then it is possible to obtain a rank-β_α linear functional of $\bar{x}_k$ directly from the output of the system.

Proof: Let $\bar{\mathcal{P}}$ be a matrix whose rows form a basis for the left nullspace of $M_{\alpha-1}$. Note that the dimension of $\bar{\mathcal{P}}$ is equal to

$$\beta_\alpha = \alpha p - \text{rank}[M_{\alpha-1}] . \quad (11)$$

Define the matrix $\mathcal{P} = \begin{bmatrix} \bar{\mathcal{P}} & 0 \end{bmatrix}$, where the zero matrix has p columns. Using (4), we see that $\mathcal{P}M_\alpha = 0$, and left-multiplying (10) by $\mathcal{P}$, we obtain $\mathcal{P}Y_{k:k+\alpha} = \mathcal{P}\bar{\Theta}_\alpha\bar{x}_k$. From the definition of $\bar{\Theta}_\alpha$ in (10), it is apparent that $\mathcal{P}\bar{\Theta}_\alpha$ is full row rank (with rank β_α), and so the theorem is proved. ■

To estimate the remaining states of $\bar{x}_k$, we choose a matrix $\mathcal{H}$ such that the matrix

$$\mathcal{T} \equiv \begin{bmatrix} \mathcal{P}\bar{\Theta}_\alpha \\ \mathcal{H} \end{bmatrix} \quad (12)$$

is invertible. In particular, $\mathcal{P}$ and $\mathcal{H}$ can be chosen so that $\mathcal{T}$ is orthogonal. Now consider the system $\widehat{\mathcal{S}}$ that is obtained by performing the similarity transformation $\hat{x}_k = \begin{bmatrix} \hat{x}_{1,k} \\ \hat{x}_{2,k} \end{bmatrix} = \mathcal{T}\bar{x}_k$, where $\hat{x}_{1,k}$ represents the first β_α states in $\hat{x}_k$. The evolution of the state vector in this system is given by

$$\begin{bmatrix} \hat{x}_{1,k+1} \\ \hat{x}_{2,k+1} \end{bmatrix} = \underbrace{\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}}_{\mathcal{T}\bar{A}\mathcal{T}^T} \begin{bmatrix} \hat{x}_{1,k} \\ \hat{x}_{2,k} \end{bmatrix} + \begin{bmatrix} \mathcal{P}\bar{\Theta}_\alpha \\ \mathcal{H} \end{bmatrix} \bar{B}u_k + \begin{bmatrix} \mathcal{P}\bar{\Theta}_\alpha \\ \mathcal{H} \end{bmatrix} \bar{B}_n n_k .$$

From Theorem 3, we see that $\hat{x}_{1,k}$ is directly available from the output of the system over $\alpha+1$ time-steps, since $\hat{x}_{1,k} = \mathcal{P}\bar{\Theta}_\alpha\bar{x}_k = \mathcal{P}Y_{k:k+\alpha}$. Therefore, we have to construct an estimator for $\hat{x}_{2,k}$, which evolves according to the equation

$$\begin{aligned} \hat{x}_{2,k+1} &= A_{21}\hat{x}_{1,k} + A_{22}\hat{x}_{2,k} + \mathcal{H}\bar{B}u_k + \mathcal{H}\bar{B}_n n_k \\ &= A_{22}\hat{x}_{2,k} + A_{21}\mathcal{P}Y_{k:k+\alpha} + \mathcal{H}\bar{B}u_k + \mathcal{H}\bar{B}_n n_k . \end{aligned} \quad (13)$$

Let $\mathcal{U}$ be a matrix such that $\begin{bmatrix} \bar{\mathcal{P}} \\ \mathcal{U} \end{bmatrix}$ is square and invertible. Define

$$\mathcal{G} \equiv \begin{bmatrix} \mathcal{U} & 0 \\ 0 & I_p \end{bmatrix}, \quad \mathcal{J} \equiv \begin{bmatrix} \mathcal{P} \\ \mathcal{G} \end{bmatrix} . \quad (14)$$

Using (3), (4) and (10), we note that

$$\mathcal{J}M_\alpha = \begin{bmatrix} 0 \\ \mathcal{G}M_\alpha \end{bmatrix}, \quad \mathcal{J}M_{n,\alpha} = \begin{bmatrix} 0 \\ \mathcal{G}M_{n,\alpha} \end{bmatrix} ,$$
$$\mathcal{J}\bar{\Theta}_\alpha\mathcal{T}^T = \begin{bmatrix} I_{\beta_\alpha} & 0 \\ L_1 & L_2 \end{bmatrix} ,$$

where

$$\begin{bmatrix} L_1 & L_2 \end{bmatrix} = \mathcal{G}\bar{\Theta}_\alpha\mathcal{T}^T . \quad (15)$$

Left-multiplying (10) by $\mathcal{J}$, we get

$$\begin{bmatrix} \mathcal{P} \\ \mathcal{G} \end{bmatrix} Y_{k:k+\alpha} = \begin{bmatrix} I_{\beta_\alpha} & 0 \\ L_1 & L_2 \end{bmatrix} \hat{x}_k + \begin{bmatrix} 0 \\ \mathcal{G}M_\alpha \end{bmatrix} U_{k:k+\alpha} + \begin{bmatrix} 0 \\ \mathcal{G}M_{n,\alpha} \end{bmatrix} n_k.$$

We see that the first β_α rows of the above equation do not contain any information about $\hat{x}_{2,k}$. The remaining rows can be written as

$$(\mathcal{G} - L_1\mathcal{P})\, Y_{k:k+\alpha} = L_2\hat{x}_{2,k} + \mathcal{G}M_\alpha U_{k:k+\alpha} + \mathcal{G}M_{n,\alpha}n_k . \quad (16)$$

To estimate $\hat{x}_{2,k}$, we use (13) and (16) to construct an estimator of the form

$$\begin{aligned} z_{k+1} &= A_{22}z_k + A_{21}\mathcal{P}Y_{k:k+\alpha} \\ &\quad + K_k\left((\mathcal{G} - L_1\mathcal{P})\, Y_{k:k+\alpha} - L_2 z_k\right) , \end{aligned} \quad (17)$$

where K_k is chosen to (i) make the estimator unbiased and (ii) minimize the mean square error between z_{k+1} and

$\hat{x}_{2,k+1}$. Using (13) and (16), we find the error between the two quantities to be

$$\begin{aligned} e_{k+1} &\equiv z_{k+1} - \hat{x}_{2,k+1} \\ &= (A_{22} - K_k L_2)\, e_k + \left(K_k \mathcal{G} M_{n,\alpha} - \mathcal{H}\bar{B}_n\right) n_k \\ &\quad + \left(K_k \mathcal{G} M_\alpha - \begin{bmatrix}\mathcal{H}\bar{B} & 0 & \cdots & 0\end{bmatrix}\right) U_{k:k+\alpha} \ . \end{aligned} \tag{18}$$

Note that the noise in (18) is no longer colored from the perspective of the error, allowing us to construct an optimal estimator. For an unbiased estimator (i.e., $E\{e_k\} = 0$), we require that

$$K_k \mathcal{G} M_\alpha = \begin{bmatrix}\mathcal{H}\bar{B} & 0 & \cdots & 0\end{bmatrix} \ . \tag{19}$$

The solvability of the above condition is given by the following theorem.

Theorem 4: There exists a matrix K_k satisfying (19) if and only if $\text{rank}\,[M_\alpha] - \text{rank}\,[M_{\alpha-1}] = m$.

Proof: Arguing as in the proof of Theorem 1, there exists a K_k satisfying (19) if and only if the row space of the matrix $S \equiv \begin{bmatrix}\mathcal{H}\bar{B} & 0 & \cdots & 0\end{bmatrix}$ is in the space spanned by the rows of $\mathcal{G}M_\alpha$. This is equivalent to the condition

$$\text{rank}\begin{bmatrix}\mathcal{G}M_\alpha \\ S\end{bmatrix} = \text{rank}\,[\mathcal{G}M_\alpha] \ . \tag{20}$$

Since $\mathcal{P}M_\alpha = 0$, we use (4) and (14) to get

$$\begin{aligned} \text{rank}\begin{bmatrix}\mathcal{G}M_\alpha \\ S\end{bmatrix} &= \text{rank}\begin{bmatrix}\mathcal{P}M_\alpha \\ \mathcal{G}M_\alpha \\ S\end{bmatrix} = \text{rank}\begin{bmatrix}M_\alpha \\ S\end{bmatrix} \\ &= \text{rank}\begin{bmatrix} D & 0 \\ \bar{\Theta}_{\alpha-1}\bar{B} & M_{\alpha-1} \\ \mathcal{H}\bar{B} & 0 \end{bmatrix} \ . \end{aligned}$$

Left-multiplying the second row in the above expression by $\bar{\mathcal{P}}$ and inserting the result into the matrix, we obtain

$$\text{rank}\begin{bmatrix}\mathcal{G}M_\alpha \\ S\end{bmatrix} = \text{rank}\begin{bmatrix} D & 0 \\ \bar{\Theta}_{\alpha-1}\bar{B} & M_{\alpha-1} \\ \bar{\mathcal{P}}\bar{\Theta}_{\alpha-1}\bar{B} & 0 \\ \mathcal{H}\bar{B} & 0 \end{bmatrix} \ .$$

Since $\mathcal{P} = \begin{bmatrix}\bar{\mathcal{P}} & 0\end{bmatrix}$, we have $\bar{\mathcal{P}}\bar{\Theta}_{\alpha-1} = \mathcal{P}\bar{\Theta}_\alpha$. This, along with the definition of $\mathcal{T}$ in (12) and our assumption that the matrix $\left[\begin{smallmatrix}B \\ D\end{smallmatrix}\right]$ has full column rank, gives us

$$\begin{aligned} \text{rank}\begin{bmatrix}\mathcal{G}M_\alpha \\ S\end{bmatrix} &= \text{rank}\begin{bmatrix} D & 0 \\ \bar{\Theta}_{\alpha-1}\bar{B} & M_{\alpha-1} \\ \bar{B} & 0 \end{bmatrix} \\ &= m + \text{rank}\,[M_{\alpha-1}] \ . \end{aligned}$$

Finally, we note that rank$[\mathcal{G}M_\alpha]$ = rank$[\mathcal{J}M_\alpha]$ = rank$[M_\alpha]$ in (20), and this concludes the proof. ■

The condition in the above theorem is the same as the one in Theorem 1. This means that the upper bound on the delay provided by Theorem 2 also applies to the reduced order estimator in (17).

To solve (19), we note that if the condition in Theorem 4 is satisfied, then the rank of $\mathcal{G}M_\alpha$ will also be m+rank$[M_{\alpha-1}]$. Let $\mathcal{N}$ be a matrix whose rows form a basis for the left nullspace of the last αm columns of $\mathcal{G}M_\alpha$. In particular, we can assume without loss of generality that $\mathcal{N}$ satisfies

$$\mathcal{N}\mathcal{G}M_\alpha = \begin{bmatrix} 0 & 0 \\ I_m & 0 \end{bmatrix} \ . \tag{21}$$

Note that $\mathcal{G}M_\alpha$ has $(\alpha+1)p - \beta_\alpha$ rows, and the last αm columns of the matrix have the same rank as $M_{\alpha-1}$. Using the expression for β_α given in (11), the number of rows in $\mathcal{N}$ is given by

$$\dim(\mathcal{N}) = (\alpha+1)p - \beta_\alpha - \text{rank}[M_{\alpha-1}] = p \ .$$

From (19), we see that K_k must be of the form

$$K_k = \widehat{K}_k \mathcal{N} \tag{22}$$

for some $\widehat{K}_k = \begin{bmatrix}\widehat{K}_{1,k} & \widehat{K}_{2,k}\end{bmatrix}$, where $\widehat{K}_{1,k}$ has $p - m$ columns and $\widehat{K}_{2,k}$ has m columns. Equation (19) then becomes

$$\begin{bmatrix}\widehat{K}_{1,k} & \widehat{K}_{2,k}\end{bmatrix}\begin{bmatrix} 0 & 0 \\ I_m & 0 \end{bmatrix} = \begin{bmatrix}\mathcal{H}\bar{B} & 0\end{bmatrix} \ , \tag{23}$$

from which it is obvious that $\widehat{K}_{2,k} = \mathcal{H}\bar{B}$ and $\widehat{K}_{1,k}$ is a free matrix.

Returning to equation (18), define

$$\begin{bmatrix}\Phi_1 \\ \Phi_2\end{bmatrix} \equiv \mathcal{N}L_2, \quad \begin{bmatrix}\Psi_1 \\ \Psi_2\end{bmatrix} \equiv \mathcal{N}\mathcal{G}M_{n,\alpha} \ , \tag{24}$$

where Φ_2 and Ψ_2 have m rows each. Substituting (22) into (18), the expression for the estimation error can be written as

$$\begin{aligned} e_{k+1} = &\left(\left(A_{22} - \mathcal{H}\bar{B}\Phi_2\right) - \widehat{K}_{1,k}\Phi_1\right) e_k \\ &+ \left(\mathcal{H}\left(\bar{B}\Psi_2 - \bar{B}_n\right) + \widehat{K}_{1,k}\Psi_1\right) n_k \ . \end{aligned} \tag{25}$$

Denoting

$$\begin{aligned} &\mathcal{A} \equiv A_{22} - \mathcal{H}\bar{B}\Phi_2, \ \ \mathcal{B} \equiv \mathcal{H}\left(\bar{B}\Psi_2 - \bar{B}_n\right) \ , \\ &\Pi_k \equiv E\{n_k n_k^T\} = \begin{bmatrix} Q_{k+\alpha-1} & 0 \\ 0 & R_{k+\alpha} \end{bmatrix} \ , \end{aligned} \tag{26}$$

we use (25) to obtain the error covariance matrix

$$\begin{aligned} \Sigma_{k+1} &\equiv E\{e_{k+1}e_{k+1}^T\} \\ &= \mathcal{A}\Sigma_k\mathcal{A}^T + \mathcal{B}\Pi_k\mathcal{B}^T - \widehat{K}_{1,k}\left(\mathcal{A}\Sigma_k\Phi_1^T - \mathcal{B}\Pi_k\Psi_1^T\right)^T \\ &\quad - \left(\mathcal{A}\Sigma_k\Phi_1^T - \mathcal{B}\Pi_k\Psi_1^T\right)\widehat{K}_{1,k}^T \\ &\quad + \widehat{K}_{1,k}\left(\Phi_1\Sigma_k\Phi_1^T + \Psi_1\Pi_k\Psi_1^T\right)\widehat{K}_{1,k}^T \ . \end{aligned} \tag{27}$$

The matrix $\widehat{K}_{1,k}$ must be chosen to minimize the mean square error (or equivalently, the trace of the error covariance matrix). Recall from the definition of $\widehat{K}_k$ in (22) that $\widehat{K}_{1,k}$ will have $p - m$ columns. This means that there will be no freedom to minimize the mean square error if the number of outputs is equal to the number of unknown inputs. If $p > m$, we take the gradient of (27) with respect to $\widehat{K}_{1,k}$ and set it equal to zero to obtain the optimal gain as

$$\widehat{K}_{1,k} = \left(\mathcal{A}\Sigma_k\Phi_1^T - \mathcal{B}\Pi_k\Psi_1^T\right)\left(\Phi_1\Sigma_k\Phi_1^T + \Psi_1\Pi_k\Psi_1^T\right)^{-1} \ . \tag{28}$$

Substituting this expression into (27), we get the optimal covariance update equation to be

$$\Sigma_{k+1} = \mathcal{A}\Sigma_k\mathcal{A}^T + \mathcal{B}\Pi_k\mathcal{B}^T - \left(\mathcal{A}\Sigma_k\Phi_1^T - \mathcal{B}\Pi_k\Psi_1^T\right)\left(\Phi_1\Sigma_k\Phi_1^T + \Psi_1\Pi_k\Psi_1^T\right)^{-1} \times \left(\mathcal{A}\Sigma_k\Phi_1^T - \mathcal{B}\Pi_k\Psi_1^T\right)^T \quad . \quad (29)$$

Note that equations (28) and (29) require the calculation of a matrix inverse. If this inverse fails to exist, it can be replaced with a pseudo-inverse [10].

We can now obtain an estimate of the original system states as follows. Using (22) and (23), we get the estimator gain in (17) to be

$$K_k = \begin{bmatrix}\widehat{K}_{1,k} & \mathcal{H}\bar{B}\end{bmatrix}\mathcal{N} \ , \quad (30)$$

where $\widehat{K}_{1,k}$ is specified in (28). The estimator is initialized with initial state $z_0 = \mathcal{H}E\{\bar{x}_0\}$, which will ensure that e_0 will have an expected value of zero. The estimate of the original state vector in (1) is given by

$$x_k = \begin{bmatrix}I_n & 0\end{bmatrix}\mathcal{T}^T\hat{x}_k \approx \begin{bmatrix}I_n & 0\end{bmatrix}\mathcal{T}^T\begin{bmatrix}\mathcal{P}Y_{k:k+\alpha}\\ z_k\end{bmatrix} \ , \quad (31)$$

and the estimation error for the state vector $\hat{x}_k$ is given by

$$\hat{e}_{k+1} \equiv \begin{bmatrix}\mathcal{P}Y_{k+1:k+\alpha+1}\\ z_{k+1}\end{bmatrix} - \begin{bmatrix}\hat{x}_{1,k+1}\\ \hat{x}_{2,k+1}\end{bmatrix} = \begin{bmatrix}0\\ e_{k+1}\end{bmatrix} \ ,$$

where e_{k+1} is defined in (18). The error for the augmented system in (9) (with state vector $\bar{x}_k$) can then be obtained as

$$\bar{e}_{k+1} = \mathcal{T}^T\hat{e}_{k+1} = \mathcal{H}^T e_{k+1} \ , \quad (32)$$

and the error covariance matrix of the original state vector is given by

$$\Sigma_{x,k+1} = \begin{bmatrix}I_n & 0\end{bmatrix}E\{\bar{e}_{k+1}\bar{e}_{k+1}^T\}\begin{bmatrix}I_n\\ 0\end{bmatrix} = \begin{bmatrix}I_n & 0\end{bmatrix}\mathcal{H}^T\Sigma_{k+1}\mathcal{H}\begin{bmatrix}I_n\\ 0\end{bmatrix} \ , \quad (33)$$

where Σ_{k+1} is given by (29). The trace of the above covariance matrix will be the mean square estimation error for the state vector in the original system. The initial error covariance matrix for the update equation (29) can also be obtained from (32) as $\Sigma_0 = \mathcal{H}\bar{\Sigma}_0\mathcal{H}^T$, where $\bar{\Sigma}_0 \equiv E\{\bar{e}_0\bar{e}_0^T\}$ is the initial error covariance matrix for the augmented system in (9).

Remark 1: Note that the estimator uses the output of the system up to time-step $k+\alpha$ in order to estimate x_k. In the Kalman filter literature, such estimators are known as *fixed-lag smoothers* [10], where a lag of α is used to obtain a better estimate of the state $x_{k-\alpha}$ (i.e., to reduce the mean square estimation error). In contrast, for the systems considered in this paper (which have unknown inputs), delays are generally necessary in order to ensure that the estimate is unbiased.

V. Design Procedure

We now summarize the steps that can be used in designing a delayed estimator for the system given in (1).

1) Find the smallest α such that $\text{rank}[M_\alpha] - \text{rank}[M_{\alpha-1}] = m$. If the condition is not satisfied for $\alpha = n - q + 1$ (where q is the dimension of the nullspace of D), it is not possible to obtain an unbiased estimate of the entire system state.
2) Construct the augmented system given in (9).
3) Choose $\mathcal{P} = \begin{bmatrix}\bar{\mathcal{P}} & 0\end{bmatrix}$ and $\mathcal{H}$ so that $\mathcal{T} = \left[\begin{smallmatrix}\mathcal{P}\bar{\Theta}_\alpha\\ \mathcal{H}\end{smallmatrix}\right]$ is orthogonal, and $\bar{\mathcal{P}}$ is a basis for the left nullspace of $M_{\alpha-1}$. Also choose $\mathcal{U}$ and form the matrix $\mathcal{G}$ given in (14).
4) Find the rank-p matrix $\mathcal{N}$ satisfying (21).
5) Form the matrices L_1, L_2, $\left[\begin{smallmatrix}\Phi_1\\ \Phi_2\end{smallmatrix}\right]$, and $\left[\begin{smallmatrix}\Psi_1\\ \Psi_2\end{smallmatrix}\right]$ from (15) and (24).
6) At each time-step k, calculate $\widehat{K}_{1,k}$ using equation (28), and update the error covariance matrix using equation (29).
7) Use (30) to obtain the estimator gain K_k.
8) The final estimator is given by equation (17). The estimate of the original state vector is given by (31), with error covariance matrix given by (33).

VI. Examples

A. Example 1

Consider the system given by the matrices

$$A = \begin{bmatrix}0.1 & 1\\ 0 & 0.2\end{bmatrix}, \ B = \begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}, \ Q_k = 0.01I_2 \ ,$$
$$C = \begin{bmatrix}0 & 1\\ 1 & 1\end{bmatrix}, \ D = \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}, \ R_k = 0.04I_2 \ .$$

It is found that condition (8) holds for $\alpha = 2$, so our estimator must have a minimum delay of two time-steps. The state and noise vectors in the augmented system will be

$$\bar{x}_k = \begin{bmatrix}x_k^T & w_k^T & v_k^T & v_{k+1}^T\end{bmatrix}^T, \ n_k = \begin{bmatrix}w_{k+1}\\ v_{k+2}\end{bmatrix} \ .$$

Using Theorem 1, we find $\beta_\alpha = 2$ and choose

$$\bar{\mathcal{P}} = \begin{bmatrix}0.64 & -0.34 & 0.34 & 0\\ -0.36 & -0.34 & 0.34 & 0\end{bmatrix}, \ \mathcal{U} = \begin{bmatrix}0 & 1 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix},$$

and

$$\mathcal{H} = \begin{bmatrix}0 & 0 & 1 & 0 & 0 & 0 & 0 & 0\\ 0.4 & 0.13 & 0 & 0.85 & -0.13 & 0.15 & -0.15 & 0\\ 0.4 & -0.64 & 0 & -0.03 & 0.64 & 0.03 & -0.03 & 0\\ -0.4 & -0.13 & 0 & 0.15 & 0.13 & 0.85 & 0.15 & 0\\ 0.4 & 0.13 & 0 & -0.15 & -0.13 & 0.15 & 0.85 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1\end{bmatrix} \ .$$

The matrix $\mathcal{N}$ in (21) is then given by $\mathcal{N} = \left[\begin{smallmatrix}-0.9 & 1 & -1 & 0\\ 0.5 & 0 & 0 & 0\end{smallmatrix}\right]$. We use (15) to obtain the matrices L_1 and L_2, and from equation (24), we get

$$\begin{bmatrix}\Phi_1\\ \Phi_2\end{bmatrix} = \begin{bmatrix}1 & -0.43 & -0.39 & -0.57 & -1.23 & 1\\ 0 & 0.73 & -0.19 & 0.27 & 0.73 & 0\end{bmatrix} \ ,$$

$$\begin{bmatrix}\Psi_1\\ \Psi_2\end{bmatrix} = \mathcal{N}\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&1&1&0\\1&1&0&1\end{bmatrix} = \begin{bmatrix}0&-1&-1&0\\0&0&0&0\end{bmatrix} .$$

Since Φ_2 and Ψ_2 have $m = 2$ rows, Φ_1 and Ψ_1 are empty matrices. Therefore, in this example, we have no freedom to minimize the trace of the covariance matrix (i.e., all of the freedom in the estimator gain is used to obtain an unbiased estimate). Using (30) (and the fact that $\widehat{K}_{1,k}$ is the empty matrix), we get the gain to be

$$K_k = \begin{bmatrix} 0 & 0 & 0 & 0\\ -0.1108 & 0.4436 & -0.4436 & 0\\ -0.4888 & 0.4221 & -0.4221 & 0\\ 0.1108 & -0.4436 & 0.4436 & 0\\ -0.1108 & 0.4436 & -0.4436 & 0\\ 0 & 0 & 0 & 0 \end{bmatrix} .$$

The final estimator is given by (17) and an estimate of the original system states can be obtained via equation (31). To test this estimator, the system is initialized with a random state with zero mean and covariance matrix I_2. A set of sinusoidal signals is used as the unknown inputs to the system. The estimator is initialized with an initial state of zero, and the resulting estimates are shown in Fig. 1. Note that the estimated state should technically be delayed by two time-steps (since our estimator has a delay of $\alpha = 2$), but we have shifted the estimate forward for purposes of comparison. After allowing equation (29) to run for several time-steps, we find the error covariance matrix for the original system state (equation (33)) converges to

$$\Sigma_x = \begin{bmatrix} 0.1156 & -0.0320\\ -0.0320 & 0.0400 \end{bmatrix} ,$$

which has a trace of 0.1556.

B. Example 2

Consider the following example from [6].

$$A = \text{diag}(0.1, 0.2, 0.3, 0.9),\ \ B = \begin{bmatrix}I_2\\ I_2\end{bmatrix},\ \ Q_k = 0.01I_4\ \ ,$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0\\ -1 & 1 & 1 & -1\end{bmatrix},\ \ D = \begin{bmatrix}0&0\\0&0\end{bmatrix},\ \ R_k = 0.01I_2\ \ .$$

Once again, we find that (8) is satisfied for $\alpha = 2$. Following the procedure outlined in this paper, we find that the error covariance matrix converges to a matrix with a trace of 0.2824. In contrast, the trace of the error covariance matrix for the estimator constructed in [6] converges to 0.3778, which is approximately 34% worse than the performance achieved by our estimator. This discrepancy arises from the fact that the procedure in [6] ignores the correlation between the error and the noise.

VII. Summary

We have provided a characterization of linear minimum-variance unbiased estimators for linear systems with unknown inputs, and have provided a design procedure to obtain the estimator parameters. We have shown that it will generally be necessary to use delayed measurements in order to obtain an unbiased estimate, and that these delays cause the system noise to become colored. We increased the dimension of our estimator in order to handle this colored noise. The resulting estimator can be applied to a more general class of systems than those currently studied in the literature.

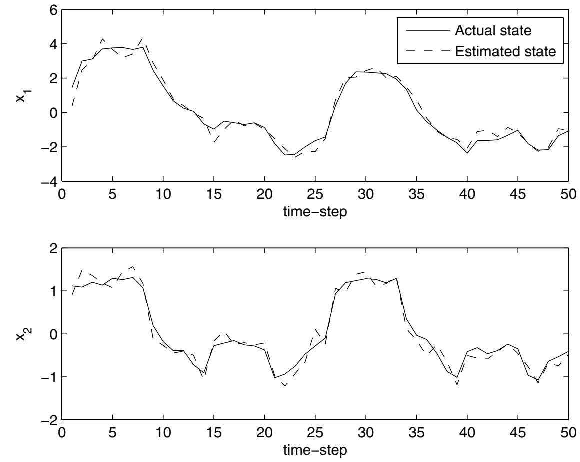

Fig. 1. Simulation of system and estimator.

References

[1] S. O. R. Moheimani, A. V. Savkin, and I. R. Peterson, “Robust filtering, prediction, smoothing, and observability of uncertain systems,” *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications*, vol. 45, no. 4, pp. 446–457, Apr. 1998.

[2] P. K. Kitanidis, “Unbiased minimum-variance linear state estimation,” *Automatica*, vol. 23, no. 6, pp. 775–778, Nov. 1987.

[3] M. Hou and R. J. Patton, “Optimal filtering for systems with unknown inputs,” *IEEE Transactions on Automatic Control*, vol. 43, no. 3, pp. 445–449, Mar. 1998.

[4] A. Saberi, A. A. Stoorvogel, and P. Sannuti, “Exact, almost and optimal input decoupled (delayed) observers,” *International Journal of Control*, vol. 73, no. 7, pp. 552–581, May 2000.

[5] M. Darouach, M. Zasadzinski, and M. Boutayeb, “Extension of minimum variance estimation for systems with unknown inputs,” *Automatica*, vol. 39, no. 5, pp. 867–876, May 2003.

[6] J. Jin and M.-J. Tahk, “Time-delayed state estimator for linear systems with unknown inputs,” *International Journal of Control, Automation, and Systems*, vol. 3, no. 1, pp. 117–121, Mar. 2005.

[7] S. Sundaram and C. N. Hadjicostis, “Comments on ‘Time-delayed state estimator for linear systems with unknown inputs’,” *International Journal of Control, Automation and Systems*, vol. 3, no. 4, pp. 646–647, Dec. 2005.

[8] M. E. Valcher, “State observers for discrete-time linear systems with unknown inputs,” *IEEE Transactions on Automatic Control*, vol. 44, no. 2, pp. 397–401, Feb. 1999.

[9] S. Sundaram and C. N. Hadjicostis, “On delayed observers for linear systems with unknown inputs,” in *Proceedings of the 44th IEEE Conference on Decision and Control and European Control Conference 2005*, Dec. 2005, pp. 7210–7215.

[10] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1979.

[11] M. K. Sain and J. L. Massey, “Invertibility of linear time-invariant dynamical systems,” *IEEE Transactions on Automatic Control*, vol. AC-14, no. 2, pp. 141–149, Apr. 1969.

[12] A. S. Willsky, “On the invertibility of linear systems,” *IEEE Transactions on Automatic Control*, vol. 19, no. 2, pp. 272–274, June 1974.