Simultaneous Input and State Estimation with a Delay

Sze Zheng Yong [a] Minghui Zhu [b] Emilio Frazzoli [a]

***Abstract*— In this paper, we present recursive algorithms for linear discrete-time stochastic systems that simultaneously estimate the states and unknown inputs in an unbiased minimum-variance sense with a delay. By allowing potential delays in state estimation, the stricter assumptions in a previous work [1] can be relaxed. Moreover, we show that a system property known as strong detectability plays a key role in the existence and stability of the asymptotic estimator with a delay we propose.**

I. Introduction

For linear discrete-time stochastic systems with known inputs, the Kalman filter optimally extracts information about a variable of interest from noisy measurements. However, these inputs that may represent unknown external drivers, instrument faults or attack signals are often not accessible. This problem of simultaneous state and input estimation is found across many disciplines and applications, from the real-time estimation of mean areal precipitation during a storm [2] to input estimation in physiological and transportation systems [1], [3] to fault detection and diagnosis [4].

Literature review. While state estimation for linear stochastic systems with unknown inputs have been widely studied under various assumptions [2], [5]–[7], the problem of *concurrently* obtaining minimum-variance unbiased estimates for both the states and the unknown inputs has received less attention. Initial research was focused on particular classes of linear systems with unknown inputs [8]–[12], and more recently, less restrictive estimators of both state and unknown input have been proposed in [1], [13], [14].

However, these estimators are restricted to estimating the states and unknown inputs at the same time step (i.e., without delay) and thus only apply to a limited class of systems. On the other hand, current results for linear *deterministic* systems [15], [16] suggest that state and input estimation is possible for a broader class of systems if delays are allowed. Such a filter with a delay for stochastic systems has been recently proposed in [17], but only for systems without direct feedthrough and with an emphasis on unbiasedness but not the optimality of the input estimates.

Contributions. We consider simultaneous input and state estimation with a delay (i.e., the estimation of inputs and states up to time step k from the measurements up to time step $k+L$ for some integer $L \geq 0$) with less restrictive assumptions on the system than currently assumed in the literature, and hence for a broader class of systems. We propose recursive algorithms that are optimal in the unbiased minimum-variance sense for these systems, along with necessary and sufficient conditions for the existence of stable estimators. Finally, we relate the stability and existence of our estimators to strong detectability of the system.

[a] S.Z Yong and E. Frazzoli are with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA, USA (e-mail: szyong@mit.edu, frazzoli@mit.edu).

[b] M. Zhu is with the Department of Electrical Engineering, Pennsylvania State University, University Park, PA, USA (e-mail: muz16@psu.edu).

II. Problem Statement

Consider the linear time-invariant discrete-time system

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

where $x_k \in \mathbb{R}^n$ is the state vector at time k, $u_k \in \mathbb{R}^m$ is a known input vector, $d_k \in \mathbb{R}^p$ is an unknown input vector, and $y_k \in \mathbb{R}^l$ is the measurement vector. The process noise $w_k \in \mathbb{R}^n$ and the measurement noise $v_k \in \mathbb{R}^l$ are assumed to be mutually uncorrelated, zero-mean, white random signals with known and bounded covariance matrices, $Q = \mathbb{E}[w_k w_k^\top] \succeq 0$ and $R = \mathbb{E}[v_k v_k^\top] \succ 0$, respectively for all k. The matrices A, B, C, D, G and H are also known. Note that no assumption is made about H either being the zero matrix (no direct feedthrough), or having full column rank when there is direct feedthrough. Without loss of generality, we assume that $n \geq l \geq 1$, $l \geq p \geq 0$ and $m \geq 0$, that the current time variable r is strictly nonnegative, and that x_0 is independent of v_k and w_k for all k and $\text{rk}[G^\top \; H^\top] = p$, where $\text{rk}(M)$ denotes the rank of matrix M.

Given that filtering and smoothing algorithms are typically initiated by an initial state estimate (biased or otherwise), it makes sense to think of an estimator with a delay as an estimator that can uniquely provide state and input estimates at all times after a possible initial delay (cf. the notion of invertibility in Definition 2). In addition, we wish for an estimator whose the state and input estimates are asymptotically unbiased with more observations. More formally, we define the desired asymptotic estimator with a delay as follows:

Definition 1 (Asymptotic/Stable Estimation with a Delay). *For any initial state $x_0 \in \mathbb{R}^n$ and sequence of unknown input $\{d_j\}_{j\in\mathbb{N}}$ in $\mathbb{R}^p$, an asymptotic estimator with a delay L*

(i) uniquely estimates the state $\hat{x}_k$ and the unknown inputs $\{d_i\}_{i=0}^{k-1}$ for all k from observations of outputs up to time step $k+L$, i.e., $\{y_i\}_{i=0}^{k+L}$, and

(ii) provides asymptotically unbiased estimates, i.e., $\mathbb{E}[\hat{x}_k - x_k] \to 0$ and $\mathbb{E}[\hat{d}_{k-1} - d_{k-1}] \to 0$ as $k \to \infty$.

The estimator design problem can be stated as follows: *Given a linear discrete-time stochastic system with unknown inputs* (1)*, design an asymptotic/stable estimator with a possible delay L (cf. Definition 1) that optimally estimates system states and unknown inputs in the unbiased minimum-variance sense.*

III. Preliminary Material

A. Linear System Properties

In this section, we provide definitions of several linear system properties, which we shall relate to the estimator existence and stability in Section IV. Without loss of generality, we assume that $w_k = 0$ and $v_k = 0$ [1], and that $B_k = D_k = 0$ (since u_k is known).

Definition 2 (Invertibility[2])**.** *The system* (1) *is said to be* invertible *if, given the initial state x_0, there exists a nonnegative integer L such that the unknown inputs $\{d_i\}_{i=0}^{k-1}$ (and thus the state x_k) can be uniquely recovered from the outputs up to time step $k+L$, denoted $\{y_i\}_{i=0}^{k+L}$.*

Definition 3 (Strong observability)**.** *The system* (1) *is said to be* strongly observable *if there exists a nonnegative integer L such that x_0 can be uniquely recovered from the outputs up to time step L, denoted $\{y_i\}_{i=0}^{L}$, for any initial state x_0 and any sequence of unknown inputs $\{d_i\}_{i=0}^{L}$. Equivalently, the system* (1) *is* strongly observable *if*

$$y_k = 0 \;\forall\, k \geq 0 \quad \textit{implies} \quad x_k = 0 \;\forall\, k \geq 0$$

for any sequence of unknown inputs $\{d_i\}_{i\in\mathbb{N}}$.

Definition 4 (Strong detectability)**.** *The system* (1) *is said to be* strongly detectable *if*

$$y_k = 0 \;\forall\, k \geq 0 \quad \textit{implies} \quad x_k \to 0 \textit{ as } k \to \infty$$

for any initial state x_0 and input sequence $\{d_i\}_{i\in\mathbb{N}}$.

Next, we consider the Rosenbrock system matrix, also known as the matrix pencil, $\mathcal{R}_S(z)$ of system (1):

$$\mathcal{R}_S(z) := \begin{bmatrix} zI - A & -G \\ C & H \end{bmatrix}. \tag{2}$$

Definition 5 (Invariant Zeros)**.** *The invariant zeros z of the system matrix $\mathcal{R}_S(z)$ in* (2) *are defined as the finite values of z for which the matrix $\mathcal{R}_S(z)$ drops rank, i.e.,*

$$\mathrm{rk}(\mathcal{R}_S(z)) < \mathrm{nrank}(\mathcal{R}_S),$$

where $\mathrm{nrank}(\mathcal{R}_S)$ denotes the normal rank (maximum rank over $z \in \mathbb{C}$) of $\mathcal{R}_S(z)$.

A useful characterization of invertibility, strong observability and strong detectability based on the invariant zeros (see, e.g., [1], [20]–[22] for proofs) are as follows:

Theorem 1 (Invertibility [20], [21])**.** *The system* (1) *is invertible if and only if $\mathrm{rk}(\mathcal{R}_S(z)) = n + p$ for at least one $z \in \mathbb{C}$.*

Theorem 2 (Strong observability [22])**.** *The system* (1) *is strongly observable if and only if $\mathrm{rk}(\mathcal{R}_S(z)) = n+p, \forall z \in \mathbb{C}$.*

Theorem 3 (Strong detectability [1])**.** *The system* (1) *is strongly detectable if and only if $\mathrm{rk}(\mathcal{R}_S(z)) = n + p, \forall z \in \mathbb{C}, |z| \geq 1$.*

From the above theorems, we observe that strong observability implies both invertibility and strong detectability, while strong detectability implies invertibility, i.e., strong observability $\subset$ strong detectability $\subset$ invertibility. Moreover, comparing the above conditions to the PBH test, we observe that strong observability/detectability implies that (A, C) is observable/detectable.

B. System Transformation

1) Initial System Transformation: To deal with the potentially rank deficient H, we first carry out a transformation of the system, as is done in our previous work [1], to decouple the output equation into two components, one with a full rank direct feedthrough matrix and the other without direct feedthrough. Let $p_H := \mathrm{rk}(H)$. Using singular value decomposition, we rewrite the matrix H as

$$H = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^\top \\ V_2^\top \end{bmatrix} \tag{3}$$

where $\Sigma \in \mathbb{R}^{p_H \times p_H}$ is a diagonal matrix of full rank, $U_1 \in \mathbb{R}^{l \times p_H}$, $U_2 \in \mathbb{R}^{l\times(l-p_H)}$, $V_1 \in \mathbb{R}^{p\times p_H}$, $V_2 \in \mathbb{R}^{p\times(p-p_H)}$, and $U := \begin{bmatrix} U_1 & U_2 \end{bmatrix}$ and $V := \begin{bmatrix} V_1 & V_2 \end{bmatrix}$ are unitary matrices. Note that when $H = 0$, Σ, U_1 and V_1 are empty matrices[3] while U_2 and V_2 are arbitrary unitary matrices. Then, we define two orthogonal components of the unknown input given by

$$d_{1,k} = V_1^\top d_k, \quad d_{2,k} = V_2^\top d_k. \tag{4}$$

Since V is unitary, $d_k = V_1 d_{1,k} + V_2 d_{2,k}$ and the system (1) can be rewritten as

$$x_{k+1} = Ax_k + Bu_k + G_1 d_{1,k} + G_2 d_{2,k} + w_k \tag{5}$$
$$y_k = Cx_k + Du_k + H_1 d_{1,k} + v_k, \tag{6}$$

where $G_1 := GV_1$, $G_2 := GV_2$ and $H_1 := HV_1 = U_1\Sigma$. Next, we decouple the output y_k using a nonsingular transformation

$$T = \begin{bmatrix} T_1 \\ T_2 \end{bmatrix} = \begin{bmatrix} I_{p_H} & -U_1^\top R U_2 (U_2^\top R U_2)^{-1} \\ 0 & I_{(l-p_H)} \end{bmatrix} \begin{bmatrix} U_1^\top \\ U_2^\top \end{bmatrix} \tag{7}$$

to obtain $z_{1,k} \in \mathbb{R}^{p_H}$ and $z_{2,k} \in \mathbb{R}^{l-p_H}$ given by

$$\begin{aligned} z_{1,k} &= T_1 y_k = C_1 x_k + D_1 u_k + \Sigma d_{1,k} + v_{1,k} \\ z_{2,k} &= T_2 y_k = C_2 x_k + D_2 u_k + v_{2,k} \end{aligned} \tag{8}$$

where $C_1 := T_1 C$, $C_2 := T_2 C = U_2^\top C$, $D_1 := T_1 D$, $D_2 := T_2 D = U_2^\top D$, $v_{1,k} := T_1 v_k$ and $v_{2,k} := T_2 v_k = U_2^\top v_k$. This transform is also chosen such that the measurement noise terms for the decoupled outputs are uncorrelated. The covariances of $v_{1,k}$ and $v_{2,k}$ are:

$$\begin{aligned} R_1 &:= \mathbb{E}[v_{1,k} v_{1,k}^\top] = T_1 R T_1^\top \succ 0, \\ R_2 &:= \mathbb{E}[v_{2,k} v_{2,k}^\top] = T_2 R T_2^\top = U_2^\top R U_2 \succ 0, \\ R_{12,(k,i)} &:= \mathbb{E}[v_{1,k} v_{2,i}^\top] = T_1 \mathbb{E}[v_k v_i^\top] T_2^\top = 0, \;\forall k, i. \end{aligned}$$

Since the initial state, process and measurement noise are assumed to be uncorrelated, it can be verified that $v_{1,k}$ and $v_{2,k}$ are also uncorrelated with the initial state x_0 and process

[1] The analysis can be extended to the case with non-zero w_k and v_k by applying the Gauss-Markov theorem [18, Theorem 3.1.1], and y_k and x_k can be replaced by $\mathbb{E}[y_k]$ and $\mathbb{E}[x_k]$. This simplification also provides a connection to the system properties of deterministic systems.

[2] Note the slightly different definition compared to [19] in which $\{d_i\}_{i=0}^{k}$ is to be uniquely determined from $\{y_i\}_{i=0}^{k+L}$.

[3] We adopt the convention that the inverse of an empty matrix is also an empty matrix and assume that operations with empty matrices are possible.

noise w_k.

2) Further Transformations: As we have seen in [1], that Σ has full rank enables us to estimate $d_{1,k}$ without delay. On the other hand, by substituting (5) into (8) to obtain

$$\begin{aligned} z_{2,k} = & C_2 A x_{k-1} + C_2 B u_{k-1} + C_2 G_1 d_{1,k-1} \\ & + C_2 G_2 d_{2,k-1} + D_2 u_k + v_{2,k}, \end{aligned}$$

we observe that $d_{2,k-1}$ can be estimated with one-step delay if $\mathcal{I}^{(0)} := C_2 G_2$ has full column rank, i.e., $p_{\mathcal{I}^{(0)}} := \text{rk}(C_2 G_2) = p - p_H$, which is the origin for the rank condition that is provided in [1] for the existence of an MVU filter. We emphasize that this rank condition is not only sufficient but also necessary for obtaining a state estimate of x_k without delay due to the influence of $d_{2,k-1}$ on x_k in (5).

We now address the question of whether a potential delay in estimating the state x_k from observations $\{y_i\}_{i=0}^{k+L}$ for some integer $L \geq 1$ would relax the requirement that $C_2 G_2$ be full rank. That is, we consider the scenario when $p_{\mathcal{I}^{(0)}} := \text{rk}(C_2 G_2) < p - p_H$. In this case, we again use singular value decomposition to rewrite the matrix $C_2 G_2$ as

$$C_2 G_2 = \begin{bmatrix} U_3 \; U_4 \end{bmatrix} \begin{bmatrix} \Sigma_3 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_3^\top \\ V_4^\top \end{bmatrix}$$

where $\Sigma_3 \in \mathbb{R}^{p_{\mathcal{I}^{(0)}} \times p_{\mathcal{I}^{(0)}}}$ is a diagonal matrix of full rank, $U_3 \in \mathbb{R}^{(l-p_H) \times p_{\mathcal{I}^{(0)}}}$, $U_4 \in \mathbb{R}^{(l-p_H) \times (l-p_H-p_{\mathcal{I}^{(0)}})}$, $V_3 \in \mathbb{R}^{(p-p_H) \times p_{\mathcal{I}^{(0)}}}$, $V_4 \in \mathbb{R}^{(p-p_H) \times (p-p_H-p_{\mathcal{I}^{(0)}})}$, and $U^{(0)} := \begin{bmatrix} U_3 \; U_4 \end{bmatrix}$ and $V^{(0)} := \begin{bmatrix} V_3 \; V_4 \end{bmatrix}$ are unitary matrices. As before, if $p_{\mathcal{I}^{(0)}} = 0$, then Σ_3, U_3 and V_3 are empty matrices and U_4 and V_4 are arbitrary unitary matrices. We then further decompose $d_{2,k-1}$ into two orthogonal components:

$$d_{3,k-1} = V_3^\top d_{2,k-1}, \quad d_{4,k-1} = V_4^\top d_{2,k-1}. \tag{9}$$

Since $V^{(0)}$ is unitary, $d_{2,k-1}$ can be reconstructed from $d_{3,k-1}$ and $d_{4,k-1}$ using $d_{2,k-1} = V_3 d_{3,k-1} + V_4 d_{4,k-1}$. The system dynamics in (5) can also be rewritten as

$$\begin{aligned} x_{k+1} &= A x_k + B u_k + G_1 d_{1,k} + G_2 V_3 d_{3,k} + G_2 V_4 d_{4,k} + w_k \\ &= A x_k + B u_k + G_1 d_{1,k} + G_3 d_{3,k} + G_4 d_{4,k} + w_k \end{aligned} \tag{10}$$

where $G_3 := G_2 V_3$ and $G_4 := G_2 V_4$. Next, we again decouple the output $z_{2,k}$ using a nonsingular transformation

$$T^{(0)} = \begin{bmatrix} T_3 \\ T_4 \end{bmatrix} = \begin{bmatrix} I_{p_{\mathcal{I}^{(0)}}} & -U_3^\top R_2 U_4 (U_4^\top R_2 U_4)^{-1} \\ 0 & I_{(l-p_H-p_{\mathcal{I}^{(0)}})} \end{bmatrix} \begin{bmatrix} U_3^\top \\ U_4^\top \end{bmatrix} \tag{11}$$

to obtain $z_{3,k} \in \mathbb{R}^{p_{\mathcal{I}^{(0)}}}$ and $z_{4,k} \in \mathbb{R}^{l-p_H-p_{\mathcal{I}^{(0)}}}$ given by

$$\begin{aligned} z_{3,k} &= T_3 z_{2,k} = C_3 x_k + D_3 u_k + v_{3,k} \\ z_{4,k} &= T_4 z_{2,k} = C_4 x_k + D_4 u_k + v_{4,k} \end{aligned} \tag{12}$$

where $C_3 := T_3 C_2$, $C_4 := T_4 C_2 = U_4^\top C_2$, $D_3 := T_3 D_2$, $D_4 := T_4 D_2 = U_4^\top D_2$, $v_{3,k} := T_3 v_{2,k}$ and $v_{4,k} := T_4 v_{2,k} = U_4^\top v_{2,k}$. With this transform, $v_{3,k}$ and $v_{4,k}$ are uncorrelated, with covariances given by:

$$\begin{aligned} R_3 &:= \mathbb{E}[v_{3,k} v_{3,k}^\top] = T_3 R_2 T_3^\top \succ 0, \\ R_4 &:= \mathbb{E}[v_{4,k} v_{4,k}^\top] = T_4 R_2 T_4^\top = U_4^\top R_2 U_4 \succ 0, \\ R_{34,(k,i)} &:= \mathbb{E}[v_{3,k} v_{4,i}^\top] = T_3 \mathbb{E}[v_{2,k} v_{2,i}^\top] T_4^\top = 0, \forall k, i. \end{aligned}$$

Moreover, $v_{3,k}$ and $v_{4,k}$ are uncorrelated with the initial state x_0 and process noise w_k. Next, from (10) and (12), and simplifying, we have

$$\begin{aligned} z_{3,k} &= C_3 A x_{k-1} + C_3 B u_{k-1} + C_3 G_1 d_{1,k-1} + \Sigma_3 d_{3,k-1} \\ &\quad + C_3 w_{k-1} + D_3 u_k + v_{3,k} \\ z_{4,k} &= C_4 A x_{k-1} + C_4 B u_{k-1} + C_4 G_1 d_{1,k-1} + C_4 w_{k-1} \\ &\quad + D_4 u_k + v_{4,k} \end{aligned} \tag{13}$$

Since $d_{1,k-1}$ can be rewritten from (8) as

$$d_{1,k-1} = \Sigma^{-1}(z_{1,k-1} - C_1 x_{k-1} - D_1 u_{k-1} - v_{1,k-1}),$$

we observe that if Σ_3 has full rank, we can uniquely estimate $d_{3,k-1}$ (with one-step delay). On the other hand, $d_{4,k-1}$ cannot be estimated from $z_{3,k}$ or $z_{4,k}$, but may instead be estimated with two-step delay, i.e., from

$$\begin{aligned} z_{4,k+1} &= C_4 A x_k + C_4 B u_k + C_4 G_1 \Sigma^{-1}(z_{1,k} - C_1 x_k \\ &\quad - D_1 u_k - v_{1,k}) + C_4 w_k + D_4 u_{k+1} + v_{4,k+1} \\ &= C_4 \hat{A} x_k + C_4 B u_k + C_4 G_1 \Sigma^{-1}(z_{1,k} - D_1 u_k \\ &\quad - v_{1,k}) + C_4 w_k + D_4 u_{k+1} + v_{4,k+1} \\ &= C_4 \hat{A} A x_{k-1} + C_4 \hat{A} B u_{k-1} + C_4 \hat{A} G_1 d_{1,k-1} \\ &\quad + C_4 \hat{A} G_3 d_{3,k-1} + C_4 \hat{A} G_4 d_{4,k-1} + C_4 \hat{A} w_{k-1} \\ &\quad + C_4 B u_k + C_4 G_1 \Sigma^{-1}(z_{1,k} - D_1 u_k - v_{1,k}) \\ &\quad + C_4 w_k + D_4 u_{k+1} + v_{4,k+1} \end{aligned} \tag{14}$$

where $\hat{A} := A - G_1 \Sigma^{-1} C_1$. Thus, we see that if $\mathcal{I}^{(1)} := C_4 \hat{A} G_4$ has full column rank, $d_{4,k-1}$ (and thus $d_{2,k-1}$) can be uniquely determined with two-step delay, and the state x_k can be estimated with one-step delay, i.e., $L = 1$. Otherwise, further decomposition procedures as above can be repeated until such a full column rank matrix $\mathcal{I}^{(L)}$ is obtained[4].

Remark 1. *Further decomposition procedures would involve the repetition of all steps in Section III-B.2. That is, for any delay L, we recursively use the singular value decomposition of* $\mathcal{I}^{(L-1)} = \begin{bmatrix} U_{2L+1} \; U_{2L+2} \end{bmatrix} \begin{bmatrix} \Sigma_{2L+1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_{2L+1}^\top \\ V_{2L+2}^\top \end{bmatrix}$ *to further decompose* $d_{2L,k-1}$ *into* $d_{2L+1,k-1} := V_{2L+1}^\top d_{2L,k-1}$ *and* $d_{2L+2,k-1} := V_{2L+2}^\top d_{2L,k-1}$*, as well as* $z_{2L,k}$ *into* $z_{2L+1,k} := T_{2L+1} z_{2L,k}$ *and* $z_{2L+2,k} := T_{2L+2} z_{2L,k}$ *with* $\begin{bmatrix} T_{2L+1}^\top \; T_{2L+2}^\top \end{bmatrix}^\top = T^{(L-1)}$*, where* $T^{(L-1)}$ *is obtained similar to* (11)*. In the process, we recursively define, among others,* $C_{2L+3} := T_{2L+3} C_{2L+2}$*,* $C_{2L+4} := T_{2L+4} C_{2L+2}$*,* $G_{2L+3} := G_{2L+2} V_{2L+3}$ *and* $G_{2L+4} := G_{2L+2} V_{2L+4}$*.*

It can be shown that the L-delay invertibility matrices form a sequence $\{\mathcal{I}^{(L)}\}_{L=0}^{\overline{L}}$ as follows:

$$\begin{aligned} & C_2 G_2, \; C_4 \hat{A} G_4, \; C_6 \hat{A} \hat{A}^{(1)} G_6, \; C_8 \hat{A} \hat{A}^{(1)} \hat{A}^{(2)} G_8, \\ & C_{10} \hat{A} \hat{A}^{(1)} \hat{A}^{(2)} \hat{A}^{(3)} G_{10}, \ldots, C_{2L+2} \hat{A} \hat{A}^{(1)} \ldots \hat{A}^{(L-1)} G_{2L+2}, \\ & \ldots, \; C_{2\overline{L}+2} \hat{A} \hat{A}^{(1)} \ldots \hat{A}^{(\overline{L}-1)} G_{2\overline{L}+2}, \end{aligned} \tag{15}$$

where we have defined $\hat{A}^{(1)} := (I - G_3 \Sigma_3^{-1} C_3) \hat{A}$ and $\hat{A}^{(L)} := (I - G_{2L+1} \Sigma_{2L+1}^{-1} C_{2L+1} \hat{A} \hat{A}^{(1)} \ldots \hat{A}^{(L-2)}) \hat{A}^{(L-1)}$ for all $L = 2, \ldots, \overline{L} - 1$. We denote as $\overline{L}$ the maximum number of delay steps beyond which input estimation (given x_0) is no longer possible. That is, if, after a delay of $\overline{L}$ time steps, d_{k-1} is not uniquely determined with given x_{k-1}, then d_{k-1} and thus x_k cannot be uniquely obtained with

[4]The delay L can be found *a priori* as the index for the first matrix in the sequence in (15) with full rank.

any additional delay, i.e., with $L > \overline{L}$. We next characterize a (conservative) upper bound on this maximum delay.

Lemma 1. *An upper bound on the maximum delay is given by $\overline{L}^u = (n-1)(p-p_H)$.*

Proof. Note that we repeat the process of further decomposing $d_{2L,k-1}$ into $d_{2L+1,k-1}$ and $d_{2(L+1),k-1}$ based on the rank of $\mathcal{I}^{(L-1)}$ only if $\mathcal{I}^{(L-1)}$ is rank deficient (including having zero rank). In the case that $\mathcal{I}^{(L-1)}$ has rank zero, we observe from the above construction of $\mathcal{I}^{(L)}$ that $C_{2L+4} = C_{2L+2}$, $G_{2L+4} = G_{2L+2}$ and $\hat{A}^{(L-1)} = \hat{A}^{(L-2)}$. By the Cayley-Hamilton theorem, if the matrices $\{\mathcal{I}^{(L)}\}$ corresponding to $n-1$ consecutive delays have rank zero, then any further delay cannot increase the rank of the next $\mathcal{I}^{(L)}$ such that d_{k-1} and x_k can be uniquely determined given x_{k-1}. Next, if $\mathcal{I}^{(L)}$ has nonzero but deficient rank, the decomposition leads to a reduction of the dimension of the resulting non-empty $d_{2(L+1),k-1}$ and thus the number of columns of $\mathcal{I}^{(L)}$ by at least 1. Since we started with $p-p_H$ columns of $\mathcal{I}^{(0)}$, this reduction in number of columns can take place at most $p-p_H$ times. Therefore, combining the two worst case scenarios gives us $\overline{L}^u = (n-1)(p-p_H)$. ■

Remark 2. *The ability to uniquely determine the unknown inputs with a delay given a previous state is equivalent to the definition of invertibility in Definition 2. Thus, we can compare the upper bound obtained in Lemma 1 as $\overline{L}^u = n(p-p_H) - p + p_H$ with the upper bound on the inherent delay for invertibility systems given in [19] as $n - p + p_H$. Thus, $\overline{L}^u$ is a more conservative upper bound except when $p = p_H$ or $p = p_H + 1$.*

IV. Algorithms for Simultaneous Input and State Estimation with a Delay

A. Existence Condition for Estimation with a Delay

We first visit the question of when an asymptotic estimator with a delay as defined in Definition 1 exists. The proof of the following claims will be provided in Section V-B.

Lemma 2 (Unique Estimates with a Delay)**.** *Given any any initial estimate $\hat{x}_0$ (biased or otherwise), the state and unknown inputs x_k and d_{k-1} can be uniquely estimated for all k with a delay L if and only if the system* (1) *is invertible.*

Lemma 3 (Asymptotic Unbiasedness)**.** *Given any initial state estimate $\hat{x}_0$ (biased or otherwise), the estimate biases with a delay L exponentially tend to zero if the pairs $(\tilde{A}^{(L)}, \tilde{C}^{(L)})$ and $(\tilde{A}^{(L)}, (\tilde{Q}^{(L)})^{\frac{1}{2}})$ are detectable and stabilizable, respectively, with $\tilde{A}^{(L)}$, $\tilde{C}^{(L)}$ and $\tilde{Q}^{(L)}$ as defined below in Remark 3.*

Theorem 4 (Existence)**.** *An asymptotic estimator with a delay L (based on Definition 1) exists if:*

(i) the system (1) *is invertible, and*

(ii) the pairs $(\tilde{A}^{(L)}, \tilde{C}^{(L)})$ and $(\tilde{A}^{(L)}, (\tilde{Q}^{(L)})^{\frac{1}{2}})$ are detectable and stabilizable, respectively.

Remark 3. *We know from [1] that $\tilde{A}^{(0)} = (I - G_2\tilde{M}_2C_2)\hat{A} + G_2\tilde{M}_2C_2$, $\tilde{C}^{(0)} = C_2$ and $\tilde{Q}^{(0)} = (I - G_2\tilde{M}_2C_2)(G_1M_1R_1M_1^\top G_1^\top + Q)(I - G_2\tilde{M}_2C_2)^\top$. For $L \geq 1$, the matrices $\tilde{A}^{(L)}$, $\tilde{C}^{(L)}$ and $\tilde{Q}^{(L)}$ can be obtained directly with further decompositions (cf. Remark 1) and with the procedure outlined in Section V-B. It can be verified that we obtain a sequence $\{\tilde{C}^{(L)}\}_{L=0}^{\overline{L}}$ given by*

$$C_2,\ C_4\hat{A},\ C_6\hat{A}\hat{A}^{(1)},\ C_8\hat{A}\hat{A}^{(1)}\hat{A}^{(2)},\ C_{10}\hat{A}\hat{A}^{(1)}\hat{A}^{(2)}\hat{A}^{(3)},$$
$$\ldots,\ C_{2L+2}\hat{A}\hat{A}^{(1)}\ldots\hat{A}^{(L-1)},\ \ldots,\ C_{2\overline{L}+2}\hat{A}\hat{A}^{(1)}\ldots\hat{A}^{(\overline{L}-1)},$$

and $\tilde{A}^{(L)}$ is of the form of $(I - G_{2L+2}\tilde{M}_{2L+2}\tilde{C}^{(L)})\hat{A}^{(L)} + G_{2L+2}\tilde{M}_{2L+2}\tilde{C}^{(L)}$, whereas the description of $\tilde{Q}^{(L)}$ is much more involved and hence, for brevity, only additionally given for $L = 1$ as $\tilde{Q}^{(1)} = (I - G_4\tilde{M}_4C_4\hat{A})(I - G_3M_3C_3)(G_1M_1R_1M_1^\top G_1^\top + Q)(I - G_3M_3C_3)^\top(I - G_4\tilde{M}_4C_4\hat{A})^\top + (I - G_4\tilde{M}_4C_4\hat{A})G_3M_3R_3M_3^\top G_3^\top(I - G_4\tilde{M}_4C_4\hat{A})^\top + G_4\tilde{M}_4C_4QC_4^\top\tilde{M}_4^\top G_4^\top$, with $M_{2L+1} = \Sigma_{2L+1}^{-1}$ and $\tilde{M}_{2L+2} = \mathcal{I}^{(L)\dagger}$.

B. Estimation Algorithms

As shown in [1], if $\text{rk}(C_2G_2) = p - p_H$, then state estimates can be obtained without delay (i.e., $L = 0$). For the sake of brevity, the reader is referred to [1] for details of the filter algorithm, its derivation and properties. Notably, the stability condition for the filter is detectability and stabilizability of $(\tilde{A}^{(0)}, \tilde{C}^{(0)})$ and $(\tilde{A}^{(0)}, (\tilde{Q}^{(0)})^{\frac{1}{2}})$, respectively (as is given in Remark 3), and strong detectability is a necessary condition for the existence of a stabilizing solution (i.e., convergence of the error covariance to steady-state).

As described in Section III-B.2, if $p_{\mathcal{I}^{(0)}} := \text{rk}(C_2G_2) < p - p_H$ but $p_{\mathcal{I}^{(1)}} := \text{rk}(C_4\hat{A}G_2) = p - p_H - p_{\mathcal{I}^{(0)}}$, then state estimates can be obtained with one step delay. In this case, given measurements up to time step k, we consider the following three-step recursive filter[5]:

Unknown Input Estimation:

$$\hat{d}_{1,k} = M_{1,k}(z_{1,k} - C_1\hat{x}_{k|k+1} - D_1u_k) \quad (16)$$
$$\hat{d}_{3,k-1} = M_{3,k}(z_{3,k} - C_3\hat{x}_{k|k} - D_3u_k) \quad (17)$$
$$\hat{d}_{4,k-1} = M_{4,k}(z_{4,k+1} - D_4u_{k+1} - C_4\hat{A}\hat{x}_{k|k} - C_4Bu_k - C_4\hat{A}G_3\hat{d}_{3,k-1} - C_4G_1\Sigma^{-1}(z_{1,k} - D_1u_k)) \quad (18)$$
$$\hat{d}_{k-1} = V_1\hat{d}_{1,k-1} + V_2V_3\hat{d}_{3,k-1} + V_2V_4\hat{d}_{4,k-1} \quad (19)$$

Time Update:

$$\hat{x}_{k|k} = A\hat{x}_{k-1|k} + Bu_{k-1} + G_1\hat{d}_{1,k-1} \quad (20)$$
$$\hat{x}^\star_{k|k+1} = \hat{x}_{k|k} + G_3\hat{d}_{3,k-1} + G_4\hat{d}_{4,k-1} \quad (21)$$

Measurement Update:

$$\hat{x}_{k|k+1} = \hat{x}^\star_{k|k+1} + \tilde{L}_k(z_{4,k+1} - C_4\hat{A}\hat{x}^\star_{k|k+1} - C_4Bu_k - C_4G_1\Sigma^{-1}(z_{1,k} - D_1u_k) - D_4u_{k+1}) \quad (22)$$

where $\hat{x}_{k-1|k}$, $\hat{d}_{1,k-1}$, $\hat{d}_{3,k-1}$, $\hat{d}_{4,k-1}$ and $\hat{d}_{k-1}$ denote the optimal estimates of x_{k-1}, $d_{1,k-1}$, $d_{3,k-1}$, $d_{4,k-1}$ and d_{k-1}. The matrices $\tilde{L}_k$, $M_{1,k}$, $M_{3,k}$ and $M_{4,k}$ (with appropriate dimensions) are filter gains that are chosen to minimize the state and input error covariances. Note that for the

[5]To initialize the filter, arbitrary initial values of $\hat{x}_{0|1}$, $P^x_{0|1}$ and $\hat{d}_{1,0}$ can be used since we will show that the filter is exponentially stable in Theorem 6. If y_0 and u_0 are available, we can find the minimum variance unbiased initial estimates given in the initialization of Algorithm 1 using the linear minimum-variance-unbiased estimator [18].

measurement update in (22), we only used a component of the measurement given by $z_{4,k+1}$. There is no loss of generality in discarding the rest because it can be verified as in [1] (in which only $z_{2,k}$ is used) that the inclusion of $z_{1,k+1}$ and $z_{3,k+1}$ will result in a biased state estimate.

Algorithm 1 summarizes the filter with delay $L = 1$. Similar to ULISE [1] (with $L = 0$), this filter possesses some nice properties, given by the following theorems. Its derivation and proofs will be provided in Section V.

Theorem 5 (Optimality). *Let the initial state estimate $\hat{x}_{0|1}$ be unbiased. If $\mathrm{rk}(C_4 \hat{A} G_4) = p - p_H - p_{\mathcal{I}^{(0)}}$, then the filter algorithm given in Algorithm 1 provides the unbiased, best linear estimate in the mean square sense of the unknown input and the minimum-variance unbiased estimate of states.*

Theorem 6 (Stability). *Let $\mathrm{rk}(C_4 \hat{A} G_4) = p - p_H - p_{\mathcal{I}^{(0)}}$. Then, that $(\tilde{A}^{(1)}, \tilde{C}^{(1)})$ is detectable is sufficient for the boundedness of the error covariance. Furthermore, if $(\tilde{A}^{(1)}, (\tilde{Q}^{(1)})^{\frac{1}{2}})$ is stabilizable, the filter is exponentially stable (i.e., its expected estimate errors decay exponentially).*

Furthermore, the following proposition shows that the invariant zeros of system (1) are the poles of the input and state filter with delay $L = 1$, specifically of the state error dynamics $\mathbb{E}[\tilde{x}^\star_{k|k+1}]$ (see proof in Section V-C):

Proposition 1. *All invariant zeros of the system* (1) *are eigenvalues of the state matrix $(\tilde{A}^{(1)} - \tilde{A}^{(1)} \tilde{L}_k \tilde{C}^{(1)})$ of the propagated state error dynamics $\mathbb{E}[\tilde{x}^\star_{k|k+1}]$.*

Proposition 1 has the implication that the invariant zeros of the system (1) cannot be stabilized by any choice of filter gain $\tilde{L}_k$. Thus, the invariant zeros of the system (1) must be stable such that the input and state filter algorithm in Algorithm 1 is stable by Theorem 6. In other words, the strong detectability of the system 1 is necessary for the stability of the filter with delay $L = 1$. Moreover, it is observed in simulation (cf. Section VI) that the converse of Proposition 1 does not hold. In that example, there are less invariant zeros than eigenvalues of $(\tilde{A}^{(1)} - \tilde{A}^{(1)} \tilde{L}_k \tilde{C}^{(1)})$.

For $2 \leq L \leq \overline{L}$, we can repeat the decomposition procedure (cf. Remark 1) and correspondingly construct asymptotic filtering algorithms as outlined in this section, provided that the system is invertible. The same optimality and stability properties as Theorems 5 and 6, as well as Proposition 1 can be also verified. However, the description of these cases would require much more notations and hence, for conciseness, is deferred to a later publication.

Remark 4. *Strong detectability of the system* (1) *is necessary for a stable filter because strongly undetectable modes of the system cannot be stabilized by any choice of filter gain (Proposition 1). Since strong detectability implies invertibility (Section III-A), and invertibility is necessary and sufficient for obtaining unique estimates (Lemma 3), we conclude that strong detectability is a key system property for the existence of a stable asymptotic estimator (Theorem 4).*

Algorithm 1 Filtering with a Delay ($L = 1$)

Initialize: $P^x_{0|1} = \mathcal{P}^x_0 = (C_2^\top R_2^{-1} C_2)^{-1}$; $\hat{x}_{0|1} = \mathbb{E}[x_0] = P^x_{0|1} C_2^\top R_2^{-1}(z_{2,0} - D_2 u_0)$; $\hat{A} = A - G_1 \Sigma^{-1} C_1$; $\hat{Q} = G_1 \Sigma^{-1} R_1 \Sigma^{-1} G_1^\top + Q$; $\overline{R}_4 = C_4 G_1 \Sigma^{-1} R_1 \Sigma^{-1} G_1^\top C_4^\top + R_4$; $\hat{d}_{1,0} = \Sigma^{-1}(z_{1,0} - C_1 \hat{x}_{0|1} - D_1 u_0)$; $P^d_{1,0} = \Sigma^{-1}(C_1 P^x_{0|1} C_1^\top + R_1)\Sigma^{-1}$; $P^{xd}_{1,0} = -P^x_{0|1} C_1^\top \Sigma^{-1}$;

for $k = 1$ to $N - 1$ **do**

▷ Estimation of $d_{3,k-1}$, $d_{4,k-1}$ and d_{k-1}

$\hat{x}_{k|k} = A\hat{x}_{k-1|k} + Bu_{k-1} + G_1 \hat{d}_{1,k-1}$;

$\hat{d}_{3,k-1} = \Sigma_3^{-1}(z_{3,k} - C_3 \hat{x}_{k|k} - D_3 u_k)$;

$\tilde{P}_k = \hat{A} P^x_{k-1|k} \hat{A}^\top + \hat{Q}$;

$\tilde{R}_{3,k} = C_3 \tilde{P}_k C_3^\top + R_3$;

$P^d_{3,k-1} = \Sigma_3^{-1} \tilde{R}_{3,k} \Sigma_3^{-1}$;

$P^{xd}_{3,k-1} = -P^x_{k-1|k} A^\top C_3^\top \Sigma_3^{-1} - P^{xd}_{1,k-1} G_1^\top C_3^\top \Sigma_3^{-1}$;

$P^d_{13,k-1} = \Sigma^{-1} C_1 P^x_{k-1|k} A^\top C_2^\top \Sigma_3^{-1} - P^d_{1,k-1} G_1^\top C_3^\top \Sigma_3^{-1}$;

$\tilde{R}_{4,k} = C_4 \hat{A} \tilde{P}_k \hat{A}^\top C_4^\top + C_4 G_1 \Sigma^{-1} R_1 \Sigma^{-1} G_1^\top C_4^\top + C_4 Q C_4^\top + R_4 + C_4 \hat{A}(A P^{xd}_{3,k-1} + G_1 P^d_{13,k-1} + G_3 P^d_{3,k-1} - Q C_3^\top \Sigma_3^{-1}) G_3^\top \hat{A}^\top C_4^\top + (C_4 \hat{A}(A P^{xd}_{3,k-1} + G_1 P^d_{13,k-1} + G_3 P^d_{3,k-1} - Q C_3^\top \Sigma_3^{-1}) G_3^\top \hat{A}^\top C_4^\top)^\top$;

$P^d_{4,k-1} = (G_4^\top \hat{A}^\top C_4^\top \tilde{R}^{-1}_{4,k} C_4 \hat{A} G_4)^{-1}$;

$M_{4,k} = P^d_{4,k-1} G_4^\top \hat{A}^\top C_4^\top \tilde{R}^{-1}_{4,k}$;

$\hat{d}_{4,k-1} = M_{4,k}(z_{4,k+1} - C_4 \hat{A} \hat{x}_{k|k} - C_4 \hat{A} G_3 \hat{d}_{3,,k-1} - C_4 B u_k - D_4 u_{k+1} - C_4 G_1 \Sigma^{-1}(z_{1,k} - D_1 u_k))$;

$P^{xd}_{4,k-1} = -(P^x_{k-1|k} A^\top + P^{xd}_{1,k-1} G_1^\top + P^{xd}_{3,k-1} G_3^\top) \hat{A}^\top C_4^\top M_{4,k}^\top$;

$P^d_{14,k-1} = -(P^{xd\,\top}_{1,k-1} A^\top + P^d_{1,k-1} G_1^\top + P^d_{13,k-1} G_3^\top) \hat{A}^\top C_4^\top M_{4,k}^\top$;

$P^d_{34,k-1} = -(P^{xd\,\top}_{3,k-1} A^\top + P^{d\,\top}_{13,k-1} G_1^\top + P^d_{3,k-1} G_3^\top) \hat{A}^\top C_4^\top M_{4,k}^\top + M_3 C_3 Q \hat{A} C_4^\top M_{4,k}^\top$;

$P^d_{12,k-1} = P^d_{13,k-1} V_3^\top + P^d_{14,k-1} V_4^\top$;

$P^d_{2,k-1} = \begin{bmatrix} V_3 & V_4 \end{bmatrix} \begin{bmatrix} P^d_{3,k-1} & P^d_{34,k-1} \\ P^{d\,\top}_{34,k-1} & P^d_{4,k-1} \end{bmatrix} \begin{bmatrix} V_3^\top \\ V_4^\top \end{bmatrix}$;

$\hat{d}_{k-1} = V_1 \hat{d}_{1,k-1} + V_2 V_3 \hat{d}_{3,k-1} + V_2 V_4 \hat{d}_{4,k-1}$;

$P^d_{k-1} = V \begin{bmatrix} P^d_{1,k-1} & P^d_{12,k-1} \\ P^{d\,\top}_{12,k-1} & P^d_{2,k-1} \end{bmatrix} V^\top$;

$P^{xd}_{k-1} = P^{xd}_{1,k-1} V_1^\top + P^{xd}_{3,k-1} V_3^\top V_2^\top + P^{xd}_{4,k-1} V_4^\top V_2^\top$;

▷ Time update

$\hat{x}^\star_{k|k+1} = \hat{x}_{k|k} + G_3 \hat{d}_{3,k-1} + G_4 \hat{d}_{4,k-1}$;

$R^{dw}_k = -G_3 \Sigma_3^{-1} C_3 Q - G_4 M_{4,k} C_4 \hat{A}(I - G_3 \Sigma_3^{-1} C_3) Q$;

$P^{\star x}_{k|k+1} = \begin{bmatrix} A & G_1 & G_2 \end{bmatrix} \begin{bmatrix} P^x_{k-1|k} & P^{xd}_{1,k-1} & P^{xd}_{2,k-1} \\ P^{xd\,\top}_{1k-1} & P^d_{1,k-1} & P^d_{12,k-1} \\ P^{xd\,\top}_{2,k-1} & P^{d\,\top}_{12,k-1} & P^d_{2,k-1} \end{bmatrix} \begin{bmatrix} A^\top \\ G_1^\top \\ G_2^\top \end{bmatrix} + Q + R^{dw}_k + R^{dw\,\top}_k$;

▷ Measurement update

$\tilde{R}^\star_{4,k} = C_4 \hat{A} P^{\star x}_{k|k+1} \hat{A}^\top C_4^\top + \overline{R}_4 - C_4 \hat{A} G_4 M_{4,k} \overline{R}_4 - \overline{R}_4 M_{4,k}^\top G_4^\top \hat{A}^\top C_4^\top$;

$\tilde{L}_k = (P^{\star x}_{k|k+1} \hat{A}^\top C_4^\top - G_4 M_{4,k} \overline{R}_4) \tilde{R}^{\star -1}_{4,k}$;

$\hat{x}_{k|k+1} = \hat{x}^\star_{k|k+1} + \tilde{L}_k(z_{4,k+1} - C_4 \hat{A} \hat{x}^\star_{k|k+1} - C_4 B u_k - C_4 G_1 \Sigma^{-1} z_{1,k} + C_4 G_1 \Sigma^{-1} D_1 u_k - D_4 u_{k+1})$;

$P^x_{k|k+1} = (I - \tilde{L}_k C_4 \hat{A}) P^{\star x}_{k|k+1} (I - \tilde{L}_k C_4 \hat{A})^\top + \tilde{L}_k \overline{R}_4 \tilde{L}_k^\top - (I - \tilde{L}_k C_4 \hat{A}) G_4 M_{4,k} \overline{R}_4 \tilde{L}_k^\top - \tilde{L}_k \overline{R}_4 M_{4,k}^\top G_4^\top (I - \tilde{L}_k C_4 \hat{A})^\top$;

▷ Estimation of $d_{1,k}$

$\tilde{R}_{1,k} = C_1 P^x_{k|k+1} C_1^\top + R_1$;

$P^d_{1,k} = \Sigma^{-1} \tilde{R}_{1,k} \Sigma^{-1}$;

$\hat{d}_{1,k} = \Sigma^{-1}(z_{1,k} - C_1 \hat{x}_{k|k+1} - D_1 u_k)$;

$P^{xd}_{1,k} = -P^x_{k|k+1} C_1^\top \Sigma^{-1}$;

end for

Remark 5. *Input and state smoothing with a delay L (i.e., the estimation of $x_{0:N-L}$ and $d_{0:N-L-1}$ from the observations in a fixed time interval given by $y_{0:N}$ and $u_{0:N}$ where $L \leq N - 1$) is also possible with the two-pass approach of*

[13], with the filter in the previous section (cf. Algorithm 1 for $L=1$) as the forward pass. Since the backward pass in [13] is agnostic to whether the filtered estimates are obtained with or without delay, the smoothing algorithm remains the same for all time steps for which filtered estimates can be obtained with a delay L. Moreover, if follows that the smoothed estimates are also unbiased and achieve minimum mean squared error and maximum likelihood [13].

V. Analysis

For the analysis of the results provided in the previous section, let $\tilde{x}_{k|k+1} := x_k - \hat{x}_{k|k+1}$, $\tilde{x}^\star_{k|k+1} := x_k - \hat{x}^\star_{k|k+1}$, $\tilde{d}_k := d_k - \hat{d}_k$, $P^x_{k|k+1} := \mathbb{E}[\tilde{x}_{k|k+1}\tilde{x}^\top_{k|k+1}]$, $P^{\star x}_{k|k+1} := \mathbb{E}[\tilde{x}^\star_{k|k+1}\tilde{x}^{\star\top}_{k|k+1}]$ and $P^d_k := \mathbb{E}[\tilde{d}_k\tilde{d}^\top_k]$, as well as $\tilde{d}_{i,k} := d_{i,k} - \hat{d}_{i,k}$, $P^d_{i,k} := \mathbb{E}[\tilde{d}_{i,k}\tilde{d}^\top_{i,k}]$, $P^{xd}_{i,k} = (P^{xd}_{i,k})^\top := \mathbb{E}[\tilde{x}_{k|k+1}\tilde{d}^\top_{i,k}]$ for $i=1,2,3,4$, and $P^d_{ij,k} = (P^d_{ij,k})^\top := \mathbb{E}[\tilde{d}_{i,k}\tilde{d}^\top_{j,k}]$, for $i,j=1,2,3,4$, $i<j$.

We begin with the derivation of the filter with delay $L=1$, which by design maintains the unbiasedness of the filter and minimizes variance of the estimate errors, thus proving Theorem 5. Then, we derive the stability conditions for the filter in Theorem 6 by means of finding an equivalent system without unknown inputs. Since Lemma 2 is straightforward to verify and Lemma 3 follows from Theorem 6 for $L=1$ and by extension for all L when derived with the same procedure, Theorem 4 holds. Finally, we prove the claim of Proposition 1 that the invariant zeros of the system are poles of the filter regardless of the choice of the filter gain $\tilde{L}_k$.

A. Filter Derivation with $L=1$ (Proof of Theorem 5)

The following lemma shows the unbiasedness of the state and unknown input estimates is preserved for all time steps.

Lemma 4. *Let $\hat{x}_{0|1} = \hat{x}^\star_{0|1}$ be unbiased, then the estimates, $\hat{d}_{k-1}$, $\hat{x}^\star_{k|k+1}$ and $\hat{x}_{k|k+1}$, are unbiased for all k, if and only if $M_{1,k}\Sigma = I$, $M_{3,k}\Sigma_3 = I$ and $M_{4,k}C_4\hat{A}G_4 = I$.*

Proof. From (8), (13), (14), (16), (17) and (18), we have

$$\hat{d}_{1,k} = M_{1,k}(C_1\tilde{x}_{k|k+1} + \Sigma d_{1,k} + v_{1,k}) \tag{23}$$

$$\begin{aligned}\hat{d}_{3,k-1} &= M_{3,k}(C_3(A\tilde{x}_{k-1|k} + G_1\tilde{d}_{1,k-1} + w_{k-1}) \\ &\quad + v_{3,k} + \Sigma_3 d_{3,k-1})\end{aligned} \tag{24}$$

$$\begin{aligned}\hat{d}_{4,k-1} &= M_{4,k}(C_4\hat{A}(A\tilde{x}_{k-1|k} + G_1\tilde{d}_{1,k-1} \\ &\quad + G_3\tilde{d}_{3,k-1} + w_{k-1}) - C_4G_1\Sigma^{-1}v_{1,k} \\ &\quad + C_4w_k + v_{4,k+1} + C_4\hat{A}G_4d_{4,k-1}).\end{aligned} \tag{25}$$

On the other hand, from (20) and (21), the error in the propagated state estimate can be obtained as:

$$\begin{aligned}\tilde{x}^\star_{k|k+1} &= A\tilde{x}_{k-1|k} + G_1\tilde{d}_{1,k-1} + G_3\tilde{d}_{3,k-1} \\ &\quad + G_4\tilde{d}_{4,k-1} + w_{k-1}.\end{aligned} \tag{26}$$

Then, from (14) and (22), the updated state estimate error is

$$\tilde{x}_{k|k+1} = (I - \tilde{L}_kC_4\hat{A})\tilde{x}^\star_{k|k+1} - \tilde{L}_k\overline{v}_{4,k}, \tag{27}$$

where $\overline{v}_{4,k} := v_{4,k} - C_4G_1\Sigma^{-1}v_{1,k}$. It can be easily shown by induction (hence omitted for brevity) that $M_{1,k}\Sigma = I$, $M_{3,k}\Sigma_3 = I$ and $M_{4,k}C_4\hat{A}G_4 = I$, $\forall k$ are necessary and sufficient for unbiasedness of $\hat{d}_{k-1}$, $\hat{x}^\star_{k|k+1}$ and $\hat{x}_{k|k+1}$. ∎

We continue the proof of Theorem 5 in three subsections, one for each step of the three-step recursive filter.

1) Unknown Input Estimation: To obtain an optimal estimate of $\hat{d}_{k-1}$ using (19), we estimate all components of the unknown input as the best linear unbiased estimates (BLUE). This means that the expected input estimate is unbiased, i.e., $\mathbb{E}[\hat{d}_{1,k}] = d_{1,k}$, $\mathbb{E}[\hat{d}_{3,k}] = d_{3,k}$, $\mathbb{E}[\hat{d}_{4,k}] = d_{4,k}$ and $\mathbb{E}[\hat{d}_k] = d_k$, as was shown in Lemma 4, and that the mean squared error of the estimate is the lowest possible.

Theorem 7. *Suppose $\hat{x}_{0|1} = \hat{x}^\star_{0|1}$ are unbiased. Then* (16), (17) *and* (18) *provide the best linear input estimate (BLUE) with $M_{1,k}$, $M_{3,k}$ and $M_{4,k}$ given by*

$$\begin{aligned}&M_{1,k} = \Sigma^{-1}, \quad M_{3,k} = \Sigma_3^{-1}, \\ &M_{4,k} = (G_4^\top\hat{A}^\top C_4^\top\tilde{R}^{-1}_{4,k}C_4\hat{A}G_4)^{-1}G_4^\top\hat{A}^\top C_4^\top\tilde{R}^{-1}_{4,k},\end{aligned} \tag{28}$$

while the input error covariance matrices are

$$\begin{aligned}&P^d_{1,k} = \Sigma^{-1}\tilde{R}_{1,k}\Sigma^{-1}, \quad P^d_{3,k-1} = \Sigma_3^{-1}\tilde{R}_{3,k}\Sigma_3^{-1}, \\ &P^d_{4,k-1} = (G_4^\top\hat{A}^\top C_4^\top\tilde{R}^{-1}_{4,k}C_4\hat{A}G_4)^{-1},\end{aligned} \tag{29}$$

with $\tilde{R}_{1,k}$, $\tilde{R}_{3,k}$ and $\tilde{R}_{4,k}$ defined in Algorithm 1.

Proof. We wish to choose $M_{1,k}$, $M_{3,k}$ and $M_{4,k}$ such that Lemma 4 holds, resulting in input estimate errors given by

$$\begin{aligned}&\tilde{d}_{1,k} = -M_{1,k}e_{1,k}, \quad \tilde{d}_{3,k-1} = -M_{3,k}e_{3,k}, \\ &\tilde{d}_{4,k-1} = -M_{4,k}e_{4,k},\end{aligned} \tag{30}$$

where $e_{1,k} := C_1\tilde{x}_{k|k+1} + v_{1,k}$, $e_{3,k} := C_3(A\tilde{x}_{k-1|k} + G_1\tilde{d}_{1,k-1} + w_{k-1}) + v_{3,k}$ and $e_{4,k} := C_4\hat{A}(A\tilde{x}_{k-1|k} + G_1\tilde{d}_{1,k-1} + G_3\tilde{d}_{3,k-1} + w_{k-1}) - C_4G_1\Sigma^{-1}v_{1,k} + v_{4,k+1}$.

Then, with $\tilde{R}_{i,k} := \mathbb{E}[e_{i,k}e^\top_{i,k}]$ for $i=1,3,4$, we apply the well known generalized least squares (GLS) estimation approach (see, e.g., [18, Theorem 3.1.1]) to obtain the optimal $M_{1,k}$, $M_{3,k}$ and $M_{4,k}$ given by (28), such that the estimates have minimum variance, i.e., are best linear unbiased estimates (BLUE). The corresponding covariance matrices are

$$\begin{aligned}P^d_{1,k} &= \mathbb{E}[\tilde{d}_{1,k}\tilde{d}^\top_{1,k}] = \Sigma^{-1}\tilde{R}_{1,k}\Sigma^{-1}, \\ P^d_{3,k-1} &= \mathbb{E}[\tilde{d}_{3,k-1}\tilde{d}^\top_{3,k-1}] = \Sigma_3^{-1}\tilde{R}_{3,k}\Sigma_3^{-1}, \\ P^d_{4,k-1} &= \mathbb{E}[\tilde{d}_{4,k-1}\tilde{d}^\top_{4,k-1}] = (G_4^\top\hat{A}^\top C_4^\top\tilde{R}^{-1}_{4,k}C_4\hat{A}G_4)^{-1}.\end{aligned}$$

Next, we note the following equality:

$$\mathrm{tr}(\mathbb{E}[\tilde{d}_k\tilde{d}_k^\top]) = \mathrm{tr}(P^d_{1,k}) + \mathrm{tr}(P^d_{3,k}) + \mathrm{tr}(P^d_{4,k}).$$

Since the unbiased estimates of $\hat{d}_{1,k}$ and $\hat{d}_{3,k-1}$ are unique (albeit at different time steps) because Σ and Σ_3 are invertible, we have $\min\mathrm{tr}(\mathbb{E}[\tilde{d}_k\tilde{d}_k^\top]) = \mathrm{tr}(\mathbb{E}[\tilde{d}_{1,k}\tilde{d}^\top_{1,k}]) + \mathrm{tr}(\mathbb{E}[\tilde{d}_{3,k}\tilde{d}^\top_{3,k}]) + \min\mathrm{tr}(\mathbb{E}[\tilde{d}_{4,k}\tilde{d}^\top_{4,k}])$, from which the unbiased estimate $\hat{d}_k$ has minimum variance when $\hat{d}_{1,k}$, $\hat{d}_{3,k}$ and $\hat{d}_{4,k}$ have minimum variances. ∎

2) Time Update: The time update is given by (20) and (21), and the error in the propagated state estimate by (26) and its covariance matrix can be computed as

$$P^{\star x}_{k|k} = \begin{bmatrix}A^\top \\ G_1^\top \\ G_2^\top\end{bmatrix}^\top \begin{bmatrix}P^x_{k-1|k} & P^{xd}_{1,k-1} & P^{xd}_{2,k-1} \\ P^{xd\,\top}_{1,k-1} & P^d_{1,k-1} & P^d_{12,k-1} \\ P^{xd\,\top}_{2,k-1} & P^{d\,\top}_{12,k-1} & P^d_{2,k-1}\end{bmatrix} \begin{bmatrix}A^\top \\ G_1^\top \\ G_2^\top\end{bmatrix} + Q + R^{dw}_k + R^{dw\top}_k,$$

where $R^{dw}_k := -G_4M_{4,k}C_4\hat{A}(I - G_3\Sigma_3^{-1}C_3)Q - G_3\Sigma_3^{-1}C_3Q$.

3) Measurement Update: In the measurement update step, the measurement $z_{4,k+1}$ is used to update the propagated estimate of $\hat{x}^{\star}_{k|k+1}$ and $P^{\star x}_{k|k+1}$. Next, the covariance matrix of the updated state error is computed as

$$\begin{aligned} P^{x}_{k|k+1} &= (I-\tilde{L}_k C_4 \hat{A}) P^{\star x}_{k|k+1} (I-\tilde{L}_k C_4 \hat{A})^\top \\ &\quad + \tilde{L}_k \tilde{R}_4 \tilde{L}_k^\top + (I-\tilde{L}_k C_4 \hat{A}) G_4 M_{4,k} \overline{R}_4 \tilde{L}_k^\top \\ &\quad + \tilde{L}_k \overline{R}_4 M_{4,k}^\top G_4^\top (I-\tilde{L}_k C_4 \hat{A})^\top \\ &:= P^{\star x}_{k|k} + \tilde{L}_k \tilde{R}^{\star}_{4,k} \tilde{L}_k - \tilde{L}_k S_k^\top - S_k \tilde{L}_k^\top \end{aligned} \tag{31}$$

where $\overline{R}_4 := \mathbb{E}[\overline{v}_{4,k}\overline{v}_{4,k}^\top] = C_4 G_1 \Sigma^{-1} R_1 \Sigma^{-1} G_1^\top C_4^\top + R_4$, $\tilde{R}^{\star}_{4,k} := C_4 \hat{A} P^{\star x}_{k|k+1} \hat{A}^\top C_4^\top + \overline{R}_4 - C_4 \hat{A} G_4 M_{4,k} \overline{R}_4 - \overline{R}_4 M_{4,k}^\top G_4^\top \hat{A} C_4^\top$ and $S_k := P^{\star x}_{k|k+1} \hat{A}^\top C_4^\top - G_4 M_{4,k} \overline{R}_4$.

Theorem 8. *Suppose $\hat{x}_{0|1} = \hat{x}^{\star}_{0|1}$ are unbiased. Then, the minimum-variance unbiased state estimator is obtained with the gain matrix $\tilde{L}_k$ given by*

$$\tilde{L}_k = (P^{\star x}_{k|k+1} \hat{A}^\top C_4^\top - G_4 M_{4,k} \overline{R}_4)(\tilde{R}^{\star}_{4,k})^{-1}. \tag{32}$$

Proof. To obtain $\tilde{L}_k$, we minimize $\mathrm{tr}(P^{x}_{k|k+1})$ by using the conventional method of differential calculus. ■

In addition, it can be verified that the (cross-)covariances $P^{xd}_{3,k-1}$, $P^{d}_{13,k-1}$, $P^{xd}_{4,k-1}$, $P^{d}_{14,k-1}$, $P^{d}_{34,k-1}$, $P^{xd}_{2,k-1}$, $P^{d}_{2,k-1}$ and $P^{d}_{12,k-1}$ are as given in Algorithm 1.

B. Stability Condition (Proof of Theorem 6)

To obtain the stability condition for the asymptotic filter with delay $L = 1$, we begin by finding an equivalent system without unknown inputs. From (27), we have $\tilde{x}_{k|k+1} = \tilde{x}^{\star}_{k|k+1} - \tilde{L}_k (C_4 \hat{A} \tilde{x}^{\star}_{k|k+1} + \overline{v}_{4,k})$. Then, substituting (30) into (26) and the above equation, and rearranging, we obtain

$$\begin{aligned} \tilde{x}_{k|k+1} &= \overline{A}^{(1)}_{k-1} \tilde{x}_{k-1|k} + \overline{w}^{(1)}_{k-1} - \tilde{L}_k (C_4 \hat{A} \overline{A}^{(1)}_{k-1} \tilde{x}_{k-1|k} \\ &\quad + C_4 \hat{A} \overline{w}^{(1)}_k + \overline{v}_{4,k}), \end{aligned} \tag{33}$$

where $\overline{A}^{(1)}_{k-1} = (I - G_4 M_{4,k} C_4 \hat{A}) \hat{A}^{(1)}$, $\overline{w}_k = (I - G_4 M_{4,k} C_4 \hat{A})[(I - G_3 \Sigma_3^{-1} C_3)(G_1 \Sigma^{-1} v_{1,k-1} + w_{k-1}) - G_3 \Sigma_3^{-1} v_{3,k}] - G_4 M_{4,k} C_4 w_k - G_4 M_{4,k} \overline{v}_{4,k}$ and $\overline{v}_{4,k} := v_{4,k} - C_4 G_1 \Sigma^{-1} v_{1,k}$. Note that the state estimate error dynamics above is the same for a Kalman filter [23] for a linear system without unknown inputs: $x^e_{k+1} = \overline{A}^{(1)}_k x^e_k + \overline{w}^{(1)}_k$; $y^e_k = C_4 \hat{A} x^e_k + \overline{v}_{4,k}$. Since the objective for both systems is the same, i.e., to obtain an unbiased minimum-variance filter, they are equivalent systems from the perspective of optimal filtering. Furthermore, since the noise terms of this equivalent system are correlated, i.e., $\mathbb{E}[\overline{w}^{(1)}_k \overline{v}_{4,k}^\top] = -G_4 M_{4,k} \overline{R}_4$, we further transform the system into one without correlated noise (see, e.g., [1]): $x^e_{k+1} = \bar{\bar{A}}^{(1)}_k x^e_k + \bar{\bar{u}}^{(1)}_k + \bar{\bar{w}}^{(1)}_k$; $y^e_k = C_4 \hat{A} x^e_k + \overline{v}_{4,k}$, with $\bar{\bar{A}}^{(1)}_k = \overline{A}^{(1)}_k + G_4 M_{4,k} C_4 \hat{A}$, $\bar{\bar{u}}^{(1)}_k = -G_4 M_{4,k} y^e_k$ is a known input and $\bar{\bar{w}}^{(1)}_k = \overline{w}^{(1)}_k + G_4 M_{4,k} \overline{v}_{4,k}$. The noise terms $\bar{\bar{w}}^{(1)}_k$ and $\overline{v}_{4,k}$ are now uncorrelated with covariances $\bar{\bar{Q}}^{(1)}_k := \mathbb{E}[\bar{\bar{w}}^{(1)}_k \bar{\bar{w}}^{(1)\top}_k]$, $\overline{R}_{4,k}$ and $\mathbb{E}[\bar{\bar{w}}^{(1)}_k \overline{v}_{4,k}^\top] = 0$. Finally, if we substitute $M_{4,k}$ by $\tilde{M}_{4,k} := (C_4 \hat{A} G_4)^\dagger$, we obtain the stability condition given in Lemma 3 from standard results of the stability of Kalman filtering (see [1] for the justification of the substitution).

C. Connection between Strong Detectability and Stability (Proof of Proposition 1)

First, we note that the following identity holds

$$\begin{aligned} &\mathrm{rk}\begin{bmatrix} zI-A & -G \\ C & H \end{bmatrix} - p_H = \mathrm{rk}\begin{bmatrix} zI-\hat{A} & -G_2 \\ C_2 & 0 \end{bmatrix} \\ &= \mathrm{rk}\begin{bmatrix} I & 0 \\ -C_2 & I \end{bmatrix}\begin{bmatrix} zI-\hat{A} & -G_2 \\ C_2 & 0 \end{bmatrix} = \mathrm{rk}\begin{bmatrix} zI-\hat{A} & -G_2 \\ C_2\hat{A} & C_2 G_2 \end{bmatrix} \\ &= \mathrm{rk}\begin{bmatrix} I & 0 \\ 0 & T^{(0)} \end{bmatrix}\begin{bmatrix} zI-\hat{A} & -G_2 \\ C_2\hat{A} & U^{(0)}\begin{bmatrix} \Sigma_3 & 0 \\ 0 & 0 \end{bmatrix} V^{(0)\top} \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & V^{(0)} \end{bmatrix} \\ &= \mathrm{rk}\begin{bmatrix} I & G_3\Sigma_3^{-1} & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}\begin{bmatrix} zI-\hat{A} & -G_3 & -G_4 \\ C_3\hat{A} & \Sigma_3 & 0 \\ C_4\hat{A} & 0 & 0 \end{bmatrix} \\ &= \mathrm{rk}\begin{bmatrix} zI-\hat{A}^{(1)} & 0 & -G_4 \\ C_3\hat{A} & \Sigma_3 & 0 \\ C_4\hat{A} & 0 & 0 \end{bmatrix} = \mathrm{rk}\begin{bmatrix} zI-\hat{A}^{(1)} & -G_4 \\ C_4\hat{A} & 0 \end{bmatrix} + p_{\mathcal{I}^{(0)}} \end{aligned}$$

where the first equality is obtained from [1]. Thus, the invariant zeros of system (1) are all $z \in \mathbb{C}$ for which the system matrix $\mathcal{R}^{(1)}_S(z) := \begin{bmatrix} zI-\hat{A}^{(1)} & -G_4 \\ C_4\hat{A} & 0 \end{bmatrix}$ drops rank.

Let z be any invariant zero of $\mathcal{R}^{(1)}_S$. Then, there exists $\begin{bmatrix} \nu^\top & \mu^\top \end{bmatrix}^\top \neq 0$ such that $\mathcal{R}^{(1)}_S(z)\begin{bmatrix} \nu^\top & \mu^\top \end{bmatrix}^\top = 0$, i.e.,

$$(zI - \hat{A}^{(1)})\nu - G_4\mu = 0, \tag{34}$$

$$C_4\hat{A}\nu = 0. \tag{35}$$

Premultiplying (34) with $(I - G_4 M_{4,k} C_4 \hat{A})$ and applying (35) as well as the fact that $M_{4,k} C_4 \hat{A} G_4 = I$, we have

$$\begin{aligned} 0 &= (I - G_4 M_{4,k} C_4 \hat{A})(zI - \hat{A}^{(1)})\nu + (I - G_4 M_{4,k} C_4 \hat{A}) G_4 \mu \\ &= (zI - \tilde{A}^{(1)})\nu = (zI - \tilde{A}^{(1)})\nu + \tilde{A}^{(1)} \tilde{L}_k C_4 \hat{A} \nu \\ &= (zI - (\tilde{A}^{(1)} - \tilde{A}^{(1)} \tilde{L}_k C_4 \hat{A}))\nu. \end{aligned}$$

If $\nu = 0$, then from (34), $G_4\mu = 0$, which implies that $\mu = 0$, which is a contradiction. Hence, $\nu \neq 0$ and the determinant of $zI - (\tilde{A}^{(1)} - \tilde{A}^{(1)} \tilde{L}_k C_4 \hat{A})$ is zero, i.e., any invariant zero of the system matrix $\mathcal{R}^{(1)}_S(z)$ is also an eigenvalue of the error dynamics of $\mathbb{E}[\tilde{x}^{\star}_{k|k+1}] = (\tilde{A}^{(1)} - \tilde{A}^{(1)} \tilde{L}_k C_4 \hat{A}) \mathbb{E}[\tilde{x}^{\star}_{k-1|k}]$.

VI. Illustrative Example

In this example, we consider the state estimation and fault identification problem when the system dynamics is plagued by faults, d_k, that influence the system dynamics and the outputs through G and H, as well as zero-mean Gaussian white noises. Specifically, the linear discrete-time problem we consider is based on the system given in [1], [7]:

$$A = \begin{bmatrix} 0.5 & 2 & 0 & 0 & 0 \\ 0 & 0.2 & 1 & 0 & 1 \\ 0 & 0 & 0.3 & 0 & 1 \\ 0 & 0 & 0 & 0.7 & 1 \\ 0 & 0 & 0 & 0 & 0.1 \end{bmatrix};\ G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix};\ H = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix};$$

$B = 0_{5\times 1};\ C = I_5;\ D = 0_{5\times 1};$

$$Q = 10^{-1}\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0.5 & 0 & 0 \\ 0 & 0.5 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix};\ R = 10^{-2}\begin{bmatrix} 1 & 0 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 0 & 0.3 \\ 0 & 0 & 1 & 0 & 0 \\ 0.5 & 0 & 0 & 1 & 0 \\ 0 & 0.3 & 0 & 0 & 1 \end{bmatrix}.$$

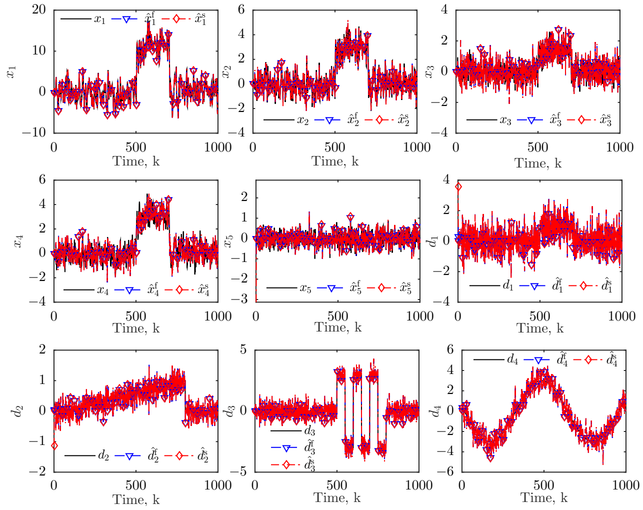

Fig. 1: Actual states x_1, x_2, x_3, x_4, x_5, unknown inputs d_1, d_2, d_3, d_4 and their filtered ('f') and smoothed ('s') estimates.

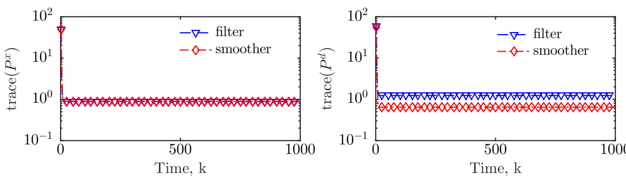

Fig. 2: Trace of estimate error covariance of states, tr(P^x), and unknown inputs, tr(P^d).

The unknown inputs used in this example are

$$d_{k,1} = \begin{cases} 1, & 500 \leq k \leq 700 \\ 0, & \text{otherwise} \end{cases}$$

$$d_{k,2} = \begin{cases} \frac{1}{700}(k-100), & 100 \leq k \leq 800 \\ 0, & \text{otherwise} \end{cases}$$

$$d_{k,3} = \begin{cases} 3, & 500 \leq k \leq 549, 600 \leq k \leq 649, 700 \leq k \leq 749 \\ -3, & 550 \leq k \leq 599, 650 \leq k \leq 699, 750 \leq k \leq 799 \\ 0, & \text{otherwise} \end{cases}$$

$$d_{k,4} = 3\sin(0.01k + 3),\ \forall k.$$

The invariant zeros of the system matrix $\mathcal{R}_S(z)$ are $\{0.7, -0.7\}$. Thus, this system is strongly detectable. Since $\text{rk}(C_2G_2) = 0$ and $\text{rk}(C_4\hat{A}G_4) = 1$, the states and unknown inputs can be estimated with delay $L = 1$.

We observe from Figure 1 and 2 that the proposed algorithm is able to estimate the system states and unknown inputs. For the sake of comparison, we have included smoothed estimates (cf. Remark 5), which shows a lower error covariance, as expected. Moreover, with the steady-state $\tilde{L}_\infty$ obtained in the simulation, we find the eigenvalues of $(\tilde{A}^{(1)} - \tilde{A}^{(1)}\tilde{L}_\infty\tilde{C}^{(1)})$ to be $\{0.7, -0.7, 0, 0, 0.0908\}$. Hence, as is predicted in Proposition 1, all invariant zeros of the system are eigenvalues of the filter.

VII. Conclusion

We presented recursive algorithms that simultaneously estimate the states and unknown inputs in an unbiased minimum-variance sense with a possible delay. The stricter requirement to ensure estimation without delay is relaxed and an asymptotic estimator is developed for this broader class of systems. Notably, strong detectability is identified as a key system property that dictates the existence and stability of an input and state estimator with a delay.

Acknowledgments

This work was supported by the National Science Foundation, grant #1239182. M. Zhu is partially supported by ARO W911NF-13-1-0421 (MURI) and NSF CNS-1505664.

References

[1] S.Z. Yong, M. Zhu, and E. Frazzoli. A unified filter for simultaneous input and state estimation of linear discrete-time stochastic systems. *Automatica*, 2015. Provisionally accepted. Extended version first appeared on *ArXiv* in September 2013. Available from: `http://arxiv.org/abs/1309.6627`.

[2] P.K. Kitanidis. Unbiased minimum-variance linear state estimation. *Automatica*, 23(6):775–778, November 1987.

[3] G. De Nicolao, G. Sparacino, and C. Cobelli. Nonparametric input estimation in physiological systems: Problems, methods, and case studies. *Automatica*, 33(5):851–870, 1997.

[4] R. Patton, R. Clark, and P.M. Frank. *Fault diagnosis in dynamic systems: theory and applications*. Prentice-Hall international series in systems and control engineering. Prentice Hall, 1989.

[5] M. Darouach and M. Zasadzinski. Unbiased minimum variance estimation for systems with unknown exogenous inputs. *Automatica*, 33(4):717–719, 1997.

[6] M. Hou and R.J. Patton. Optimal filtering for systems with unknown inputs. *IEEE Trans. on Automatic Control*, 43(3):445–449, 1998.

[7] Y. Cheng, H. Ye, Y. Wang, and D. Zhou. Unbiased minimum-variance state estimation for linear systems with unknown input. *Automatica*, 45(2):485–491, 2009.

[8] C. Hsieh. Robust two-stage Kalman filters for systems with unknown inputs. *IEEE Trans. on Automatic Control*, 45(12):2374–2378, December 2000.

[9] S. Gillijns and B. De Moor. Unbiased minimum-variance input and state estimation for linear discrete-time systems. *Automatica*, 43(1):111–116, January 2007.

[10] S. Gillijns and B. De Moor. Unbiased minimum-variance input and state estimation for linear discrete-time systems with direct feedthrough. *Automatica*, 43(5):934–937, March 2007.

[11] H. Fang, Y. Shi, and J. Yi. On stable simultaneous input and state estimation for discrete-time linear systems. *International Journal of Adaptive Control and Signal Processing*, 25(8):671–686, 2011.

[12] S.Z. Yong, M. Zhu, and E. Frazzoli. Simultaneous input and state estimation for linear discrete-time stochastic systems with direct feedthrough. In *Conference on Decision and Control*, pages 7034–7039, December 2013.

[13] S.Z. Yong, M. Zhu, and E. Frazzoli. Simultaneous input and state smoothing for linear discrete-time stochastic systems with unknown inputs. In *Conference on Decision and Control*, pages 3388–3394, December 2014.

[14] S.Z. Yong, M. Zhu, and E. Frazzoli. Simultaneous input and state estimation for linear time-invariant continuous-time stochastic systems with unknown inputs. In *American Control Conference*, pages 2511–2518, July 2015.

[15] T. Floquet and J.-P. Barbot. State and unknown input estimation for linear discrete-time systems. *Automatica*, 42(11):1883–1889, 2006.

[16] S. Sundaram and C.N. Hadjicostis. Delayed observers for linear systems with unknown inputs. *Automatic Control, IEEE Transactions on*, 52(2):334–339, February 2007.

[17] R.A. Chavan, K. Fitch, and H.J. Palanthandalam-Madapusi. Recursive input reconstruction with a delay. In *American Control Conference*, pages 628–633, June 2014.

[18] A.H. Sayed. *Fundamentals of Adaptive Filtering*. Wiley, 2003.

[19] A.S. Willsky. On the invertibility of linear systems. *IEEE Trans. on Automatic Control*, 19(3):272–274, June 1974.

[20] P.J. Moylan. Stable inversion of linear systems. *IEEE Transactions on Automatic Control*, 22(1):74–78, February 1977.

[21] S. Sundaram. Fault-tolerant and secure control systems. University of Waterloo, Lecture Notes, 2012.

[22] S. Sundaram and C.N. Hadjicostis. Distributed function calculation via linear iterative strategies in the presence of malicious agents. *IEEE Trans. on Automatic Control*, 56(7):1495–1508, 2011.

[23] R.E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D):35–45, 1960.