A Unified Filter for Simultaneous Input and State Estimation of Linear Discrete-time Stochastic Systems ⋆

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

[a] *Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

[b] *Department of Electrical Engineering, Pennsylvania State University, 201 Old Main, University Park, PA 16802, USA*

Abstract

In this paper, we present a unified optimal and exponentially stable filter for linear discrete-time stochastic systems that simultaneously estimates the states and unknown inputs in an unbiased minimum-variance sense, without making any assumptions on the direct feedthrough matrix. We also provide the connection between the stability of the estimator and a system property known as strong detectability, and discuss the global optimality of the proposed filter. Finally, an illustrative example is given to demonstrate the performance of the unified unbiased minimum-variance filter.

Keywords: Optimal filtering; State estimation; Input estimation; Filter stability; Recursive filter

1 Introduction

The term filter or estimator is commonly used to refer to systems that extract information about a quantity of interest from measured data corrupted by noise. Kalman filtering provides the tool needed for obtaining that reliable estimate when the system is linear and when the disturbance inputs are well modeled by a zero-mean, Gaussian white noise. However, in many instances, the exogenous input (e.g., the inputs of other autonomous vehicles) cannot be modeled as a Gaussian stochastic process rendering the estimates unreliable. Nonetheless, we want to be able to estimate the states and inputs of other vehicles based on noisy measurements for purposes of collision avoidance, route planning, etc. Similar problems can be found across a wide range of disciplines, from the real-time estimation of mean areal precipitation during a storm [19] to fault detection and diagnosis [22] to input estimation in physiological systems [6].

Literature review. Much of the research focus has been on state estimation of systems with unknown inputs without actually estimating the inputs. An optimal filter that estimates a minimum-variance unbiased (MVU) state estimate for a system with unknown inputs is first developed for linear systems without direct feedthrough in [19]. This design was extended to a more general parameterized solution by [4], and eventually to state estimation of systems with direct feedthrough in [13, 5, 3]. Similarly, while H_∞ filters (e.g., [28, 27, 20]) can deal with non-Gaussian disturbance inputs in minimizing the worst-case state estimation error, the unknown input is not estimated. However, the problem of estimating the unknown input itself is often as important as the state information, and should also be considered.

Palanthandalam-Madapusi and Bernstein [21] proposed an approach to reconstruct the unknown inputs, in a process that is decoupled from state estimation with an emphasis on unbiasedness, but neglecting the optimality of the estimate. On the other hand, Hsieh [14] and Gillijns and De Moor [10] developed simultaneous input and state filters that are optimal in the minimum-variance unbiased sense, for systems without direct feedthrough. Extensions to systems with a full rank direct feedthrough matrix were proposed by Gillijns and De Moor [11], Fang et al. [9] and Yong et al. [30]. In an attempt to deal with systems with a rank deficient direct feedthrough matrix, Hsieh [15] allowed the input estimate to be biased. Thus, the problem of finding a simultaneous state and input filter for systems with rank deficient direct feedthrough matrix, that is both unbiased and has minimum variance remains open. Moreover, a unified MVU filter that works for all cases remains elusive.

Another set of relevant literature pertains to the stability of the state and input filters, since optimality does not imply stability and vice versa. However, to the best of our knowledge, the literature on this subject is limited to linear time-invariant systems [3, 8, 9]. Yet another related literature is on state and input observability and detectability conditions, also known as strong or perfect observability and detectability, as this will be shown to be related to the stability of the filter dynamics for both linear time-varying and time-invariant systems with un-

⋆ This work was supported by NSF grant CNS-1239182. M. Zhu is partially supported by ARO W911NF-13-1-0421 (MURI) and NSF grant CNS-1505664. The material in this paper was partially presented at the 2013 IEEE Conference on Decision and Control, Florence, Italy. Corresponding author: S.Z. Yong. Tel. +1 339 200 9664. Fax +1 617 324 6819.

Email addresses: szyong@mit.edu (S.Z. Yong), muz16@psu.edu (M. Zhu), frazzoli@mit.edu (E. Frazzoli).

known inputs. Some conditions for state and input observability were derived in [21, 12].

Contributions. We introduce a unified filter for simultaneously estimating both states and unknown inputs such that the estimates are unbiased and have minimum variance with no restrictions on the direct feedthrough matrix of the linear discrete-time stochastic system, which is a generalization of the estimators in the literature, specifically of [10, 11, 30], and the Kalman filter. Furthermore, we derive sufficient conditions for the filter stability of linear time-varying systems with unknown inputs, an important problem that has been previously unexplored; while for linear time-invariant systems, necessary and sufficient conditions for convergence of the filter gains to a steady-state solution are provided. The key insight we gained is that the exponential stability of the filter is directly related to the strong detectability of the time-varying system, without which unbiased state and input estimates cannot be obtained even in the absence of stochastic noise. We shall also show that the proposed filter is globally optimal (i.e., optimal over the class of all linear state and input estimators as in [18]).

In connection to the existing literature, this paper presents a combination of several ideas from [3, 10, 11] and our recent work [30] into a unified filter in a manner that provably preserves and extends the nice properties of these filters. However, there are a number of distinctions between our filter and the above referenced filters. In particular, we show that the state-only filter in [3] implicitly estimates the unknown inputs in a suboptimal manner and so does the approach for input estimation in [11]. In contrast, our filter uses the approaches of our previous work in [30] and of generalized least squares estimation, which lead to the desired optimality of the input estimates. In addition, we gave sufficient conditions for filter stability of linear time-varying systems, which cannot be carried over from the existing literature (including [3, 10, 11]) for linear time-invariant systems. Preliminary versions of the results appeared in a conference paper [30] and on arXiv [29] (in which more details on input and state observability/detectability are provided and a suboptimal filter variant is described).

Notation. We first summarize some notations used throughout the paper. $\mathbb{R}^n$ denotes the n-dimensional Euclidean space, $\mathbb{C}$ the field of complex numbers and $\mathbb{N}$ nonnegative integers. For a random vector, $v \in \mathbb{R}^n$, the expectation is denoted by $\mathbb{E}[v]$. Given a matrix $M \in \mathbb{R}^{p\times q}$, its transpose, inverse, Moore-Penrose pseudoinverse, range, trace and rank are given by $M^\top$, M^{-1}, $M^\dagger$, $\mathrm{Ra}(M)$, $\mathrm{tr}(M)$ and $\mathrm{rk}(M)$. For a symmetric matrix S, $S \succ 0$ ($S \succeq 0$) indicates S is positive (semi-)definite.

2 Problem Statement

Consider the linear time-varying discrete-time system

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k u_k + G_k d_k + w_k \\ y_k &= C_k x_k + D_k u_k + H_k d_k + v_k \end{aligned} \tag{1}$$

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 bounded covariance matrices, $Q_k = \mathbb{E}[w_k w_k^\top] \succeq 0$ and $R_k = \mathbb{E}[v_k v_k^\top] \succ 0$, respectively. The matrices A_k, B_k, C_k, D_k, G_k and H_k are known and bounded. Note that no assumption is made on H_k to be either the zero matrix (no direct feedthrough), or to have full column rank when there is direct feedthrough. Without loss of generality, we assume that $\max_k(\mathrm{rk}[G_k^\top \; H_k^\top]) = p$, $n \geq l \geq 1$, $l \geq p \geq 0$, $m \geq 0$, the current time variable r is strictly nonnegative and x_0 is independent of v_k and w_k for all k.

The estimator design problem, addressed in this paper, can be stated as follows:
Given a linear discrete-time stochastic system with unknown inputs (1)*, design a globally optimal and stable filter that simultaneously estimates system states and unknown inputs in an unbiased minimum-variance manner.*

3 Preliminary Material

3.1 System Transformation

We first decouple the output equation into two components, one with a full rank direct feedthrough matrix and the other without direct feedthrough. In this form, the filter can be designed leveraging existing approaches for both cases (e.g., [10, 30]).

Let $p_{H_k} := \mathrm{rk}(H_k)$. Using singular value decomposition, we rewrite the direct feedthrough matrix H_k as $H_k = \begin{bmatrix} U_{1,k} \; U_{2,k} \end{bmatrix} \begin{bmatrix} \Sigma_k & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_{1,k}^\top \\ V_{2,k}^\top \end{bmatrix}$, where $\Sigma_k \in \mathbb{R}^{p_{H_k}\times p_{H_k}}$ is a diagonal matrix of full rank, $U_{1,k} \in \mathbb{R}^{l\times p_{H_k}}$, $U_{2,k} \in \mathbb{R}^{l\times(l-p_{H_k})}$, $V_{1,k} \in \mathbb{R}^{p\times p_{H_k}}$ and $V_{2,k} \in \mathbb{R}^{p\times(p-p_{H_k})}$, while $U_k := \begin{bmatrix} U_{1,k} \; U_{2,k} \end{bmatrix}$ and $V_k := \begin{bmatrix} V_{1,k} \; V_{2,k} \end{bmatrix}$ are unitary matrices. When there is no direct feedthrough, Σ_k, $U_{1,k}$ and $V_{1,k}$ are empty matrices [1], and $U_{2,k}$ and $V_{2,k}$ are arbitrary unitary matrices.

Then, as suggested in [3], we define two orthogonal components of the unknown input given by

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

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

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k u_k + G_k V_{1,k} d_{1,k} + G_k V_{2,k} d_{2,k} + w_k \\ &= A_k x_k + B_k u_k + G_{1,k} d_{1,k} + G_{2,k} d_{2,k} + w_k \quad (3) \\ y_k &= C_k x_k + D_k u_k + H_k V_{1,k} d_{1,k} + H_k V_{2,k} d_{2,k} + v_k \\ &= C_k x_k + D_k u_k + H_{1,k} d_{1,k} + v_k, \quad (4) \end{aligned}$$

where $G_{1,k} := G_k V_{1,k}$, $G_{2,k} := G_k V_{2,k}$ and $H_{1,k} := H_k V_{1,k} = U_{1,k} \Sigma_k$. Next, we decouple the output y_k using

[1] 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.

a nonsingular transformation $T_k = \begin{bmatrix} T_{1,k}^\top \; T_{2,k}^\top \end{bmatrix}^\top$

$$T_k = \begin{bmatrix} I_{p_{H_k}} & -U_{1,k}^\top R_k U_{2,k}(U_{2,k}^\top R_k U_{2,k})^{-1} \\ 0 & I_{(l-p_{H_k})} \end{bmatrix} \begin{bmatrix} U_{1,k}^\top \\ U_{2,k}^\top \end{bmatrix} \quad (5)$$

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

$$\begin{aligned} z_{1,k} &= T_{1,k} y_k = C_{1,k} x_k + D_{1,k} u_k + \Sigma_k d_{1,k} + v_{1,k} \\ z_{2,k} &= T_{2,k} y_k = C_{2,k} x_k + D_{2,k} u_k + v_{2,k} \end{aligned} \quad (6)$$

where $C_{1,k} := T_{1,k} C_k$, $C_{2,k} := T_{2,k} C_k = U_{2,k}^\top C_k$, $D_{1,k} := T_{1,k} D_k$, $D_{2,k} := T_{2,k} D_k = U_{2,k}^\top D_k$, $v_{1,k} := T_{1,k} v_k$ and $v_{2,k} := T_{2,k} v_k = U_{2,k}^\top v_k$. 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

$$\begin{aligned} R_{1,k} &:= \mathbb{E}[v_{1,k} v_{1,k}^\top] = T_{1,k} R_k T_{1,k}^\top \succ 0, \\ R_{2,k} &:= \mathbb{E}[v_{2,k} v_{2,k}^\top] = T_{2,k} R_k T_{2,k}^\top = U_{2,k}^\top R_k U_{2,k} \succ 0, \\ R_{12,k} &:= \mathbb{E}[v_{1,k} v_{2,k}^\top] = T_{1,k} R_k T_{2,k}^\top = 0, \\ R_{12,(k,i)} &:= \mathbb{E}[v_{1,k} v_{2,i}^\top] = T_{1,k} \mathbb{E}[v_k v_i^\top] T_{2,i}^\top = 0, \ \forall k \neq i. \end{aligned} \quad (7)$$

Moreover, $v_{1,k}$ and $v_{2,k}$ are also uncorrelated with the initial state x_0 and process noise w_k.

3.2 Input and State Detectability

Similar to the stability of the Kalman filter, we will show in Section 5 that the stability of the unified filter is directly related to the notion of input and state detectability, a.k.a. strong detectability. Without loss of generality, we assume that $B_k = D_k = 0$, since u_k is known.

Definition 1 (Strong detectability [2]) *The linear system* (1) *is* strongly detectable *if*

$$\mathbb{E}[y_k] = 0 \; \forall k \geq 0 \quad \textit{implies} \quad \mathbb{E}[x_k] \to 0 \textit{ as } k \to \infty$$

for all initial states and input sequences $\{d_i\}_{i \in \mathbb{N}}$.

Theorem 2 (Strong detectability (time-invariant)) *A linear time-invariant discrete-time system is strongly detectable if and only if either of the following holds:*

(i) $\operatorname{rk} P(z) := \operatorname{rk} \begin{bmatrix} zI - A & -G \\ C & H \end{bmatrix} = n+p, \forall z \in \mathbb{C}, |z| \geq 1,$

(ii) $\operatorname{rk} \begin{bmatrix} zI - \hat{A} & -G_2 \\ C_2 & 0 \end{bmatrix} = n + p - p_H, \forall z \in \mathbb{C}, |z| \geq 1$

where $\hat{A} := A - G_1 \Sigma^{-1} C_1$. *The above conditions are equivalent to the system being minimum-phase (i.e., the invariant zeros of* $P(z)$ *in Condition (i) are stable).*

[2] This definition is a simple extension of strong detectability for deterministic systems [12, 25]. Note that strong detectability we defined is not equivalent to exact detectability [32, 33] ($y_k = 0, \forall k \geq 0 \Rightarrow E\|x_k\|^2 \to 0$).

Proof. To prove that strong detectability is equivalent to Condition (i), we first note without proof that strong observability [3] is equivalent to $\operatorname{rk}(P(z)) = n+p, \forall z \in \mathbb{C}$. Then, Condition (i) is a simple generalization for the case that $P(z)$ is rank deficient for some $z \in Z_0 \subset \mathbb{C}$ but $|z| < 1$. For each such z, there exists $\begin{bmatrix} -x_z^\top \; d_z^\top \end{bmatrix}^\top$ in the null space of $\mathsf{P}(z)$. It can be verified that the input sequence $d_k = z^k d_z$ and the initial state x_z leads to the output is $\mathbb{E}[y_k] = 0$ for all $k \geq 0$ but $\mathbb{E}[x_k] = z^k x_z$, where with a slight abuse of notation, z^k represents the product of any permutations of k numbers from Z_0. Since $|z| < 1$ by assumption, $\mathbb{E}[x_k] \to 0$ as $k \to \infty$. To relate Conditions (i) and (ii), we use the following identity:

$$\begin{aligned} n + p &= \operatorname{rk} \begin{bmatrix} zI - A & -G \\ C & H \end{bmatrix} = \operatorname{rk} \begin{bmatrix} zI - A & -G \\ C & U \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} V^\top \end{bmatrix} \\ &= \operatorname{rk} \begin{bmatrix} I & 0 \\ 0 & T \end{bmatrix} \begin{bmatrix} zI - A & -G \\ C & U \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} V^\top \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & V \end{bmatrix} \\ &= \operatorname{rk} \begin{bmatrix} zI - A & -GV \\ TC & TU \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} \end{bmatrix} = \operatorname{rk} \begin{bmatrix} zI - A & -G_1 & -G_2 \\ C_1 & \Sigma & 0 \\ C_2 & 0 & 0 \end{bmatrix} \\ &= \operatorname{rk} \begin{bmatrix} I & G_1 \Sigma^{-1} & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} zI - A & -G_1 & -G_2 \\ C_1 & \Sigma & 0 \\ C_2 & 0 & 0 \end{bmatrix} \\ &= \operatorname{rk} \begin{bmatrix} zI - \hat{A} & 0 & -G_2 \\ C_1 & \Sigma & 0 \\ C_2 & 0 & 0 \end{bmatrix} = \operatorname{rk} \begin{bmatrix} zI - \hat{A} & -G_2 \\ C_2 & 0 \end{bmatrix} + p_H \end{aligned}$$

for all $z \in \mathbb{C}$, where the final equality holds because Σ is square and has full rank p_H. ∎

4 Minimum-Variance Unbiased Filter for Simultaneous Input and State Estimation

For the design of the *Unified Linear Input & State Estimator* (ULISE), we consider a recursive three-step filter [4] as in [11, 30], composed of an *unknown input estimation* step that uses the current measurement and state estimate to estimate the unknown inputs in the best linear unbiased sense (i.e., the minimum-variance-unbiased among the class of linear estimators), a *time update* step that propagates the state estimate based on the system dynamics, and a *measurement update* step that updates the state estimate using the current measurement.

Given measurements up to time k, the three-step recursive filter [5] can be summarized as follows:

[3] Strong observability is a stronger condition than (and implies) strong detectability. Due to space limitation, the reader is referred to [29, 26] for its definition, properties and proofs.

[4] Note that the restriction to a recursive filter will be relaxed and shown to not lead to suboptimality in Theorem 4.

[5] To initialize the filter, arbitrary initial values of $\hat{x}_{0|0}$, $P^x_{0|0}$

Unknown Input Estimation:

$$\hat{d}_{1,k} = M_{1,k}(z_{1,k} - C_{1,k}\hat{x}_{k|k} - D_{1,k}u_k) \tag{8}$$

$$\hat{d}_{2,k-1} = M_{2,k}(z_{2,k} - C_{2,k}\hat{x}_{k|k-1} - D_{2,k}u_k) \tag{9}$$

$$\hat{d}_{k-1} = V_{1,k-1}\hat{d}_{1,k-1} + V_{2,k-1}\hat{d}_{2,k-1} \tag{10}$$

Time Update:

$$\hat{x}_{k|k-1} = A_{k-1}\hat{x}_{k-1|k-1} + B_{k-1}u_{k-1} + G_{1,k-1}\hat{d}_{1,k-1} \tag{11}$$

$$\hat{x}^{\star}_{k|k} = \hat{x}_{k|k-1} + G_{2,k-1}\hat{d}_{2,k-1} \tag{12}$$

Measurement Update:

$$\begin{aligned}\hat{x}_{k|k} &= \hat{x}^{\star}_{k|k} + L_k(y_k - C_k\hat{x}^{\star}_{k|k} - D_k u_k) \\ &= \hat{x}^{\star}_{k|k} + \tilde{L}_k(z_{2,k} - C_{2,k}\hat{x}^{\star}_{k|k} - D_{2,k}u_k)\end{aligned} \tag{13}$$

where $\hat{x}_{k-1|k-1}$, $\hat{d}_{1,k-1}$, $\hat{d}_{2,k-1}$ and $\hat{d}_{k-1}$ denote the optimal estimates of x_{k-1}, $d_{1,k-1}$, $d_{2,k-1}$ and d_{k-1}; $L_k \in \mathbb{R}^{n\times l}$, $\tilde{L}_k := L_k U_{2,k} \in \mathbb{R}^{n\times(l-p_{H_k})}$, $M_{1,k} \in \mathbb{R}^{p_{H_k}\times p_{H_k}}$ and $M_{2,k} \in \mathbb{R}^{(p-p_{H_k})\times(l-p_{H_k})}$ are filter gain matrices that are chosen to minimize the state and input error covariances. Note that we applied $L_k = L_k U_{2,k}U_{2,k}^{\top}$ in (13), which we will justify in Lemma 8. Algorithm 1 summarizes the three steps of ULISE, in which the estimation of $d_{2,k-1}$ is carried out before the time update, followed by the measurement update and finally, the estimation of $d_{1,k}$. Note that Algorithm 1 is given with significant simplifications and a particular choice of Γ_k that will be further expounded in Section 5. An even more simplified version of ULISE is given in [31, Algorithm 1].

The proposed unified filter simultaneously estimates the unknown inputs and states for systems with an arbitrary direct feedthrough matrix; thus it relaxes assumptions on the direct feedthrough matrix in [10, 11, 30]. By a suitable system transformation given in (5), the unknown input is decomposed into two components, $d_{1,k}$ and $d_{2,k}$, and similarly, the output equation into two orthogonal projections, $z_{1,k}$ and $z_{2,k}$, one with no direct feedthrough and the other with a full-rank feedthrough matrix. Hence, the $d_{1,k}$ component of the unknown input can be estimated in the best linear unbiased sense by choosing $M_{1,k}$ as in [30] and the $d_{2,k}$ component by choosing $M_{2,k}$ as in [10]. Moreover, the gain matrix L_k is chosen to minimize the state estimate error covariance in an update similar to the Kalman filter. In fact, the proposed filter can be shown to be a generalization of the Kalman filter to systems with unknown inputs and other filters in existing literature, e.g., [19, 10, 11, 3, 30]. The proof is omitted for brevity (cf. [29]).

Moreover, ULISE possesses some nice properties, given by the following lemma and theorems which will be proven in Section 5. To state these claims,

---

and $\hat{d}_{1,0}$ can be used since we will show that the filter is exponentially stable in Theorems 5 and 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 [24].

Algorithm 1 ULISE algorithm

Initialize: $P^x_{0|0} = \mathcal{P}^x_0 = (C_{2,0}^{\top}R_{2,0}^{-1}C_{2,0})^{-1}$; $\hat{x}_{0|0} = \mathbb{E}[x_0] = P^x_{0|0}C_{2,0}^{\top}R_{2,0}^{-1}(z_{2,0} - D_{2,0}u_0)$; $\hat{A}_0 = A_0 - G_{1,0}\Sigma_0^{-1}C_{1,0}$; $\hat{Q}_0 = G_{1,0}\Sigma_0^{-1}R_{1,0}\Sigma_0^{-1}G_{1,0}^{\top} + Q_0$; $\hat{d}_{1,0} = \Sigma_0^{-1}(z_{1,0} - C_{1,0}\hat{x}_{0|0} - D_{1,0}u_0)$; $P^d_{1,0} = \Sigma_0^{-1}(C_{1,0}P^x_{0|0}C_{1,0}^{\top} + R_{1,0})\Sigma_0^{-1}$;

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

▷ Estimation of $d_{2,k-1}$ and d_{k-1}

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

$\tilde{R}_{2,k} = C_{2,k}\tilde{P}_k C_{2,k}^{\top} + R_{2,k}$;

$P^d_{2,k-1} = (G_{2,k-1}^{\top}C_{2,k}^{\top}\tilde{R}_{2,k}^{-1}C_{2,k}G_{2,k-1})^{-1}$;

$M_{2,k} = P^d_{2,k-1}G_{2,k-1}^{\top}C_{2,k}^{\top}\tilde{R}_{2,k}^{-1}$;

$\hat{x}_{k|k-1} = A_{k-1}\hat{x}_{k-1|k-1} + B_{k-1}u_{k-1} + G_{1,k-1}\hat{d}_{1,k-1}$;

$\hat{d}_{2,k-1} = M_{2,k}(z_{2,k} - C_{2,k}\hat{x}_{k|k-1} - D_{2,k}u_k)$;

$\hat{d}_{k-1} = V_{1,k-1}\hat{d}_{1,k-1} + V_{2,k-1}\hat{d}_{2,k-1}$;

$P^d_{12,k-1} = M_{1,k-1}C_{1,k-1}P^x_{k-1|k-1}A_{k-1}^{\top}C_{2,k}^{\top}M_{2,k}^{\top} - P^d_{1,k-1}G_{1,k-1}^{\top}C_{2,k}^{\top}M_{2,k}^{\top}$;

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

▷ Time update

$\hat{x}^{\star}_{k|k} = \hat{x}_{k|k-1} + G_{2,k-1}\hat{d}_{2,k-1}$;

$P^{\star x}_{k|k} = G_{2,k-1}M_{2,k}R_{2,k}M_{2,k}^{\top}G_{2,k}^{\top} + (I - G_{2,k-1}M_{2,k}C_{2,k})\tilde{P}_k(I - G_{2,k-1}M_{2,k}C_{2,k})^{\top}$;

$\tilde{R}^{\star}_k = C_k P^{\star x}_{k|k}C_k^{\top} + R_k - C_k G_{2,k-1}M_{2,k}U_{2,k}^{\top}R_k - R_k U_{2,k}M_{2,k}^{\top}G_{2,k-1}^{\top}C_k$;

▷ Measurement update

$K_k = P^{\star x}_{k|k}C_k^{\top} - G_{2,k-1}M_{2,k}U_{2,k}^{\top}R_k$;

$M^{\star}_{1,k} := \Sigma_k^{-1}(U_{1,k}^{\top}\tilde{R}^{\star\dagger}_k U_{1,k})^{-1}U_{1,k}^{\top}\tilde{R}^{\star\dagger}_k$;

$L_k = K_k(I_l - U_{1,k}\Sigma_k M^{\star}_{1,k})^{\top}\tilde{R}^{\star\dagger}_k$;

$\hat{x}_{k|k} = \hat{x}^{\star}_{k|k} + L_k(y_k - C_k\hat{x}^{\star}_{k|k} - D_k u_k)$;

$P^x_{k|k} = (I - L_k C_k)G_{2,k-1}M_{2,k}U_{2,k}^{\top}R_k L_k^{\top} + L_k R_k U_{2,k}M_{2,k}^{\top}G_{2,k-1}^{\top}(I - L_k C_k)^{\top} + (I - L_k C_k)P^{\star x}_{k|k}(I - L_k C_k)^{\top} + L_k R_k L_k^{\top}$;

▷ Estimation of $d_{1,k}$

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

$M_{1,k} = \Sigma_k^{-1}$;

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

$\hat{d}_{1,k} = M_{1,k}(z_{1,k} - C_{1,k}\hat{x}_{k|k} - D_{1,k}u_k)$;

$\hat{A}_k = A_k - G_{1,k}M_{1,k}C_{1,k}$;

$\hat{Q}_k = G_{1,k}M_{1,k}R_{1,k}M_{1,k}^{\top}G_{1,k}^{\top} + Q_k$;

end for

---

we first define: $\tilde{M}_{2,k} := (C_{2,k}G_{2,k-1})^{\dagger}$, $\hat{Q}_k = Q_k + G_{1,k}\Sigma_k^{-1}R_{1,k}\Sigma_k^{-1\top}G_{1,k}^{\top}$, $\hat{A}_k = A_k - G_{1,k}M_{1,k}C_{1,k}$, $\tilde{A}_k := (I - G_{2,k-1}\tilde{M}_{2,k}C_{2,k})\hat{A}_k + G_{2,k-1}\tilde{M}_{2,k}C_{2,k}$ and $\tilde{Q}_k = (I - G_{2,k-1}\tilde{M}_{2,k}C_{2,k})\hat{Q}_{k-1}(I - G_{2,k-1}\tilde{M}_{2,k}C_{2,k})^{\top}$.

Lemma 3 *Let the initial state estimate $\hat{x}_{0|0}$ be unbiased. If $\mathrm{rk}(C_{2,k}G_{2,k-1}) = p - p_{H_{k-1}}$, then the ULISE 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 4 (Global Optimality) *Let the initial state estimate $\hat{x}_{0|0}$ be unbiased and $\mathrm{rk}(C_{2,k}G_{2,k-1}) = p - p_{H_{k-1}}$. Then, the ULISE algorithm is globally optimal* *(over the class of all linear state and input estimators).*

Theorem 5 (Stability) *Let* $\text{rk}(C_{2,k}G_{2,k-1}) = p - p_{H_{k-1}}$. *Then, that* $(\tilde{A}_k, C_{2,k})$ *is uniformly detectable*[6] *is sufficient for the boundedness of the error covariance of the ULISE algorithm. Furthermore, if* $(\tilde{A}_k, \tilde{Q}_k^{\frac{1}{2}})$ *is uniformly stabilizable*[6], *ULISE is exponentially stable (i.e., its expected estimate errors decay exponentially).*

Theorem 6 (Stability (linear time-invariant)) *Let* $\text{rk}(C_2G_2) = p - p_H$. *Then, that* $(\tilde{A}, C_2)$ *is detectable is sufficient for the boundedness of the error covariance of the ULISE algorithm. Furthermore, if* $(\tilde{A}, \tilde{Q}^{\frac{1}{2}})$ *is stabilizable, ULISE is exponentially stable (i.e., its expected estimate errors decay exponentially). In addition, with* $P^x_{0|0} \succ 0$, *the filter gains of ULISE converge to a unique stationary solution,* $P^x_\infty \succ 0$ *(cf. Lines 3, 13, 19 of Algorithm 1 with* $P^x_{k|k} = P^x_{k-1|k-1} = P^x_\infty$*), if and only if*

(i) The linear time-invariant discrete-time system is strongly detectable, i.e., Theorem 2 holds, and

(ii) $\text{rk}\begin{bmatrix} \hat{A} - e^{j\omega} I & G_2 & \hat{Q}^{\frac{1}{2}} & 0 \\ e^{j\omega}C_2 & 0 & 0 & R_2^{\frac{1}{2}} \end{bmatrix} = n + l - p_H, \forall \omega \in [0, 2\pi].$

Remark 7 *Note the parallels of the convergence and stability conditions above (i.e., strong detectability and a rank condition on the unit circle) to the conditions for the Kalman filter (i.e., detectability and controllability on the unit circle). Conversely, without strong detectability, it is not possible to obtain unbiased estimates of the states and unknown inputs even for the case with no noise.*

5 Filter Description and Analysis

For the analysis of the proposed filter, let $\tilde{x}_{k|k} := x_k - \hat{x}_{k|k}$, $\tilde{x}^\star_{k|k} := x_k - \hat{x}^\star_{k|k}$, $\tilde{d}_k := d_k - \hat{d}_k$, $P^x_{k|k} := \mathbb{E}[\tilde{x}_{k|k}\tilde{x}^\top_{k|k}]$, $P^{\star x}_{k|k} := \mathbb{E}[\tilde{x}^\star_{k|k}\tilde{x}^{\star\top}_{k|k}]$, $P^d_k := \mathbb{E}[\tilde{d}_k\tilde{d}^\top_k]$ and $P^d_{12,k} = (P^d_{21,k})^\top := \mathbb{E}[\tilde{d}_{1,k}\tilde{d}^\top_{2,k}]$, as well as $\tilde{d}_{i,k} := d_{i,k} - \hat{d}_{i,k}$, $P^d_{i,k} := \mathbb{E}[\tilde{d}_{i,k}\tilde{d}^\top_{i,k}]$ and $P^{xd}_{i,k} = (P^{dx}_{i,k})^\top := \mathbb{E}[\tilde{x}_{k|k}\tilde{d}^\top_{i,k}]$, for $i = 1, 2$. We first present a lemma that summarizes the unbiasedness of the state and input estimates for all time steps that is one piece of the claim in Lemma 3.

Lemma 8 *Let* $\hat{x}_{0|0} = \hat{x}^\star_{0|0}$ *be unbiased, then the input and state estimates,* $\hat{d}_{k-1}$, $\hat{x}^\star_{k|k}$ *and* $\hat{x}_{k|k}$, *are unbiased for all* k, *if and only if* $M_{1,k}\Sigma_k = I$, $M_{2,k}C_{2,k}G_{2,k-1} = I$ *and* $L_kU_{1,k} = 0$. *Consequently,* $\text{rk}(C_{2,k}G_{2,k-1}) = p - p_{H_{k-1}}$ *and* $L_k = L_kU_{2,k}U^\top_{2,k}$.

Proof. We observe from (6), (8) and (9) that

$$\hat{d}_{1,k} = M_{1,k}(C_{1,k}\tilde{x}_{k|k} + \Sigma_k d_{1,k} + v_{1,k}) \tag{14}$$

$$\begin{aligned}\hat{d}_{2,k-1} = M_{2,k}(C_{2,k}(A_{k-1}\tilde{x}_{k-1|k-1} + G_{1,k-1}\tilde{d}_{1,k-1} \\ + w_{k-1}) + v_{2,k} + C_{2,k}G_{2,k-1}d_{2,k-1}).\end{aligned} \tag{15}$$

[6] For brevity, the readers are referred to [1, Section 2] for the definitions of uniform detectability and stabilizability. A spectral test for these properties can be found in [23].

From (11) and (12), as well as (4) and (13), the propagated and updated state estimate errors are

$$\begin{aligned}\tilde{x}^\star_{k|k} = A_{k-1}\tilde{x}_{k-1|k-1} + G_{1,k-1}\tilde{d}_{1,k-1} \\ + G_{2,k-1}\tilde{d}_{2,k-1} + w_{k-1}\end{aligned} \tag{16}$$

$$\tilde{x}_{k|k} = (I - L_kC_k)\tilde{x}^\star_{k|k} - L_kU_{1,k}\Sigma_k d_{1,k} - L_kv_k. \tag{17}$$

We show by induction that the estimates $\hat{d}_k$, $\hat{x}_{k|k}$ and $\hat{x}^\star_{k|k}$ are unbiased. For the base case, since $\hat{x}_{0|0}$ and $\hat{x}^\star_{0|0}$ are unbiased and the process and measurement noise are assumed to have zero mean, $\mathbb{E}[w_0] = 0$, $\mathbb{E}[v_0] = 0$, from (14) and (15), we find that $\mathbb{E}[\hat{d}_{1,0}] = d_{1,0}$ and $\mathbb{E}[\hat{d}_{2,0}] = d_{2,0}$, i.e., $\hat{d}_{1,0}$ and $\hat{d}_{2,0}$ are unbiased, if and only if $M_{1,0}\Sigma_0 = I$, and $M_{2,1}C_{2,1}G_{2,0} = I$. Hence, $\hat{d}_0$ is unbiased. In the inductive step, we assume that $\mathbb{E}[\tilde{x}_{k-1|k-1}] = \mathbb{E}[\tilde{x}^\star_{k-1|k-1}] = 0$. Then, the input estimates are unbiased, i.e., $\mathbb{E}[\tilde{d}_{k-1}] = \mathbb{E}[\tilde{d}_{1,k-1}] = \mathbb{E}[\tilde{d}_{2,k-1}] = 0$, if and only if $M_{1,k-1}\Sigma_{k-1} = I$, and $M_{2,k}C_{2,k}G_{2,k-1} = I$. Since the process noise has zero mean, by (16), $\mathbb{E}[\tilde{x}^\star_{k|k}] = 0$. Similarly, from (17) with a zero-mean measurement noise, we impose the constraint $L_kU_{1,k} = 0$ such that we obtain $\mathbb{E}[\tilde{x}_{k|k}] = 0$. Therefore, by induction, $\mathbb{E}[\tilde{x}^\star_{k|k}] = 0$ and $\mathbb{E}[\tilde{x}_{k|k}] = 0$ for all k. Since we require $M_{2,k}C_{2,k}G_{2,k-1} = I$ for all k for the existence of an unbiased input estimate, it follows that $\text{rk}(C_{2,k}G_{2,k-1}) = p - p_{H_k - 1}$ is a necessary and sufficient condition. Furthermore, $L_k = L_kU_kU^\top_k = L_kU_{2,k}U^\top_{2,k}$ since $L_kU_{1,k} = 0$. ∎

Next, we continue the proof of Lemma 3 in three subsections, one for each step of the three-step recursive filter. Then, we present the proof of Theorems 4, 5 and 6.

5.1 Unknown Input Estimation

To obtain an optimal estimate of $\hat{d}_{k-1}$ using (10), we estimate both components of the unknown input as the best linear unbiased estimates (BLUE). This means that the expected input estimate is unbiased, i.e., $\mathbb{E}[\hat{d}_{1,k}] = d_{1,k}$, $\mathbb{E}[\hat{d}_{2,k}] = d_{2,k}$ and $\mathbb{E}[\hat{d}_k] = d_k$, as was shown in Lemma 8, and that the mean squared error of the estimate is the lowest possible, shown next in Theorem 9.

Theorem 9 *Let* $\hat{x}_{0|0} = \hat{x}^\star_{0|0}$ *be unbiased. Then* (8) *and* (9) *provide the best linear input estimate (BLUE) with*

$$M_{1,k} = \Sigma_k^{-1} \tag{18}$$

$$M_{2,k} = (G^\top_{2,k-1}C^\top_{2,k}\tilde{R}^{-1}_{2,k}C_{2,k}G_{2,k-1})^{-1}G^\top_{2,k-1}C^\top_{2,k}\tilde{R}^{-1}_{2,k} \tag{19}$$

while the input error covariance matrices are

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

$$P^d_{2,k-1} = (G^\top_{2,k-1}C^\top_{2,k}\tilde{R}^{-1}_{2,k}C_{2,k}G_{2,k-1})^{-1}$$

where we have defined $\tilde{P}_k := \hat{A}_{k-1}P^x_{k-1|k-1}\hat{A}^\top_{k-1} + \hat{Q}_{k-1}$, $\hat{A}_k := A_k - G_{1,k}M_{1,k}C_{1,k}$, $\hat{Q}_k := Q_k +$

$G_{1,k}M_{1,k}R_{1,k}M_{1,k}^\top G_{1,k}^\top$, $\tilde{R}_{1,k} := C_{1,k}P_{k|k}^x C_{1,k}^\top + R_{1,k}$ *and* $\tilde{R}_{2,k} := C_{2,k}\tilde{P}_k C_{2,k}^\top + R_{2,k}$.

Proof. Let $\tilde{z}_{1,k} := z_{1,k} - C_{1,k}\hat{x}_{k|k} - D_{1,k}u_k$ and $\tilde{z}_{2,k} := z_{2,k} - C_{2,k}\hat{x}_{k|k-1} - D_{2,k}u_k$. Then, we have

$$\tilde{z}_{1,k} = \Sigma_k d_{1,k} + e_{1,k}, \quad (20)$$
$$\tilde{z}_{2,k} = C_{2,k}G_{2,k-1}d_{2,k-1} + e_{2,k}, \quad (21)$$

where we defined $e_{1,k} := C_{1,k}\tilde{x}_{k|k} + v_{1,k}$ and $e_{2,k} := C_{2,k}(A_{k-1}\tilde{x}_{k-1|k-1} + G_{1,k-1}\tilde{d}_{1,k-1} + w_{k-1}) + v_{2,k}$. From the unbiasedness of the state and input estimates (Lemma 8), $\mathbb{E}[e_{1,k}] = 0$ and $\mathbb{E}[e_{2,k}] = 0$. It can be verified that their covariance matrices, $\tilde{R}_{1,k} := \mathbb{E}[e_{1,k}e_{1,k}^\top]$ and $\tilde{R}_{2,k} := \mathbb{E}[e_{2,k}e_{2,k}^\top]$, are as given in the theorem statement. Next, we obtain the estimates for $\hat{d}_{1,k}$ and $\hat{d}_{2,k}$ given by (8), (9), (18) and (19) by applying the well known generalized least squares (GLS) estimate (see, e.g., [24, Theorem 3.1.1]), which are linear minimum-variance unbiased estimates, a.k.a. as best linear unbiased estimates (BLUE). Note that since Σ_k is invertible, there is one unique unbiased estimate of $\hat{d}_{1,k}$. Since $M_{1,k}\Sigma_k = I$ and $M_{2,k}C_{2,k}G_{2,k-1} = I$, the input estimate errors, and their covariance matrices are

$$\begin{aligned}
&\tilde{d}_{1,k} = -M_{1,k}e_{1,k}, \quad \tilde{d}_{2,k-1} = -M_{2,k}e_{2,k} \\
&P_{1,k}^d = \mathbb{E}[\tilde{d}_{1,k}\tilde{d}_{1,k}^\top] = M_{1,k}\mathbb{E}[e_{1,k}e_{1,k}^\top]M_{1,k}^\top = \Sigma_k^{-1}\tilde{R}_{1,k}\Sigma_k^{-1} \\
&P_{2,k-1}^d = \mathbb{E}[\tilde{d}_{2,k-1}\tilde{d}_{2,k-1}^\top] = M_{2,k}\mathbb{E}[e_{2,k}e_{2,k}^\top]M_{2,k}^\top \\
&\qquad = (G_{2,k-1}^\top C_{2,k}^\top \tilde{R}_{2,k}^{-1}C_{2,k}G_{2,k-1})^{-1}.
\end{aligned} \quad (22)$$

Next, we note the following equality:

$$\begin{aligned}
&\mathrm{tr}(\mathbb{E}[\tilde{d}_k\tilde{d}_k^\top]) = \mathrm{tr}(\mathbb{E}[V_k \begin{bmatrix}\tilde{d}_{1,k} \\ \tilde{d}_{2,k}\end{bmatrix} \begin{bmatrix}\tilde{d}_{1,k} & \tilde{d}_{2,k}\end{bmatrix} V_k^\top]) \\
&= \mathrm{tr}(V_k^\top V_k \mathbb{E}[\begin{bmatrix}\tilde{d}_{1,k} \\ \tilde{d}_{2,k}\end{bmatrix} \begin{bmatrix}\tilde{d}_{1,k} & \tilde{d}_{2,k}\end{bmatrix}]) = \mathrm{tr}(P_{1,k}^d) + \mathrm{tr}(P_{2,k}^d).
\end{aligned} \quad (23)$$

Since the unbiased estimate of $\hat{d}_{1,k}$ is unique (albeit at a different time step), we have $\min \mathrm{tr}(\mathbb{E}[\tilde{d}_k\tilde{d}_k^\top]) = \mathrm{tr}(\mathbb{E}[\tilde{d}_{1,k}\tilde{d}_{1,k}^\top]) + \min \mathrm{tr}(\mathbb{E}[\tilde{d}_{2,k}\tilde{d}_{2,k}^\top])$, from which it can be observed that the unbiased estimate $\hat{d}_k$ has minimum variance when $\hat{d}_{1,k}$ and $\hat{d}_{2,k}$ have minimum variances. ∎

5.2 Time Update

The time update is given by (11) and (12), and the propagated state estimate error covariance matrix is

$$\begin{aligned}
P_{k|k}^{\star x} = &\begin{bmatrix} A_{k-1}^\top \\ G_{1,k-1}^\top \\ G_{2,k-1}^\top \end{bmatrix}^\top \begin{bmatrix} P_{k-1|k-1}^x & P_{1,k-1}^{xd} & P_{2,k-1}^{xd} \\ P_{1,k-1}^{xd\,\top} & P_{1,k-1}^d & P_{12,k-1}^d \\ P_{2,k-1}^{xd\,\top} & P_{12,k-1}^{d\,\top} & P_{2,k-1}^d \end{bmatrix} \begin{bmatrix} A_{k-1}^\top \\ G_{1,k-1}^\top \\ G_{2,k-1}^\top \end{bmatrix} \\
&+ Q_{k-1} - G_{2,k-1}M_{2,k}C_{2,k}Q_{k-1} - Q_{k-1}C_{2,k}^\top M_{2,k}^\top G_{2,k-1}^\top.
\end{aligned}$$

Alternatively, using (22), (18) and (19), the above expression can be reduced to

$$\begin{aligned}
P_{k|k}^{\star x} = &(I - G_{2,k-1}M_{2,k}C_{2,k})\tilde{P}_k(I - G_{2,k-1}M_{2,k}C_{2,k})^\top \\
&+ G_{2,k-1}M_{2,k}R_{2,k}M_{2,k}^\top G_{2,k-1}^\top
\end{aligned} \quad (24)$$

where we applied $L_k = L_kU_{2,k}U_{2,k}^\top$ from Lemma 8 and $T_{1,k}R_kT_{2,k}^\top = 0$ from (7), and where $M_{2,k}$ and $\tilde{P}_k$ are as defined in Theorem 9.

5.3 Measurement Update

In this step, the measurement y_k is used to update the propagated estimate of $\hat{x}_{k|k}^\star$ and $P_{k|k}^{\star x}$. From (4) and (13), the updated state estimate error is given by (17) where the constraint $L_kU_{1,k} = 0$ (Lemma 8) must be imposed for all k such that the state estimate is unbiased ($\mathbb{E}[\tilde{x}_{k|k}] = 0$) for all possible $d_{1,k}$, since Σ_k has full rank. Note that the residual/innovation term in the measurement update step given in (13) appears to not contain an $H_k\hat{d}_k$ term as would be expected. This term is actually present, but has been nullified by the unbiasedness constraint (Lemma 8), since $L_kH_k = L_kU_{1,k}\Sigma_kV_{1,k}^\top = 0$. This is also in line with the practical reason that the unknown input estimate is not yet available. Next, the covariance matrix of the state error is computed as

$$\begin{aligned}
P_{k|k}^x &= (I - L_kC_k)P_{k|k}^{\star x}(I - L_kC_k)^\top + L_kR_kL_k^\top \\
&\quad + (I - L_kC_k)G_{2,k-1}M_{2,k}U_{2,k}^\top R_kL_k^\top \\
&\quad + L_kR_kU_{2,k}M_{2,k}^\top G_{2,k-1}^\top(I - L_kC_k)^\top \\
&:= P_{k|k}^{\star x} + L_k\tilde{R}_k^\star L_k^\top - L_kS_k^\top - S_kL_k^\top
\end{aligned} \quad (25)$$

where $\mathbb{E}[\tilde{x}_{k|k}^\star v_k^\top] = -G_{2,k-1}M_{2,k}U_{2,k}^\top R_k$, and we defined $\tilde{R}_k^\star := C_kP_{k|k}^{\star x}C_k^\top + R_k - C_kG_{2,k-1}M_{2,k}U_{2,k}^\top R_k - R_kU_{2,k}M_{2,k}^\top G_{2,k-1}^\top C_k^\top$ and $S_k := -G_{2,k-1}M_{2,k}U_{2,k}^\top R_k + P_{k|k}^{\star x}C_k^\top$. Using (24), we can rewrite the expression $\tilde{R}_k^\star = N_k\hat{R}_kN_k^\top$ where $\hat{R}_k := C_k\tilde{P}_kC_k^\top + R_k$ and $N_k := I - C_kG_{2,k-1}M_{2,k}U_{2,k}^\top$.

To obtain an unbiased minimum variance estimator, we derive the optimal gain matrix L_k, by minimizing the trace of (25), since the trace represents the sum of the estimation error variances of the states, subject to the constraint $L_kU_{1,k} = 0$. However, the next lemma shows that $\tilde{R}_k^\star = N_k\hat{R}_kN_k^\top$ is singular because N_k is rank deficient, except when $p = p_{H_k}$, i.e., H_k has full rank.

Lemma 10 *Consider $M_{2,k}$ that satisfies* (19)*, then N_k has rank $p_R := l - p + p_{H_{k-1}}$ and $p_{H_{k-1}} \le p_R \le l$.*

Proof. Since $M_{2,k}$ satisfies (19), N_k is an idempotent matrix, i.e., $N_kN_k = N_k$. From [2, Fact 3.12.9 and Proposition 2.6.3] and $\mathrm{rk}(C_{2,k}G_{2,k-1}) = p - p_{H_{k-1}}$, we obtain $p_R := \mathrm{rk}(I_l - C_kG_{2,k-1}M_{2,k}U_{2,k}) = l - \mathrm{rk}(C_kG_{2,k-1}M_{2,k}U_{2,k}) = l - p + p_{H_{k-1}} \le l$. Since we assumed $l \ge p$, we have $p_{H_{k-1}} \le p_R \le l$. ∎

Hence, the optimal gain matrix L_k is in general not unique. Similar to [10], we propose a gain matrix L_k of the form $L_k = \overline{L}_k \Gamma_k$ where $\Gamma_k \in \mathbb{R}^{p_R \times l}$ is an arbitrary matrix which has to be chosen such that $\Gamma_k \tilde{R}_k^\star \Gamma_k^\top$ has full rank. With this, we compute the optimal gain $\overline{L}_k$ and thus L_k in the following theorem.

Theorem 11 *Suppose $\hat{x}_{0|0} = \hat{x}^\star_{0|0}$ are unbiased, and let $\Gamma_k \in \mathbb{R}^{p_R \times l}$ be chosen such that $\Gamma_k \tilde{R}_k^\star \Gamma_k^\top$ has full rank, where $p_R = l - p + p_{H_{k-1}}$. Then, the minimum-variance unbiased state estimator is obtained with*

$$L_k = K_k \check{R}_k (I_l - H_{1,k} M^\star_{1,k}) = K_k (I_l - H_{1,k} M^\star_{1,k})^\top \check{R}_k \tag{26}$$

where $K_k := (P^{\star x}_{k|k} C_k^\top - G_{2,k-1} M_{2,k} U_{2,k}^\top R_k) = (\tilde{P}_k C_k^\top - G_{2,k-1} M_{2,k} U_{2,k}^\top \hat{R}_k) N_k^\top$, $\check{R}_k := \Gamma_k^\top (\Gamma_k \tilde{R}_k^\star \Gamma_k^\top)^{-1} \Gamma_k$, $H_{1,k} = U_{1,k} \Sigma_k$ and $M^\star_{1,k} := \Sigma_k^{-1} (U_{1,k}^\top \check{R}_k U_{1,k})^{-1} U_{1,k}^\top \check{R}_k$, with $M_{2,k}$ and $\tilde{P}_k$ as defined in Theorem 9, and $\hat{R}_k$ and $\tilde{R}_k^\star$ as defined in the text following (25).

Proof. By Lemma 8, the state estimates are unbiased. We then employ the optimization approach with Lagrange multipliers ($\Lambda_k \in \mathbb{R}^{n \times p_H}$) in [19, 11, 30], to obtain the optimal gain L_k in (26) that minimizes the trace of of the covariance matrix $P^x_{k|k}$, while being subjected to the constraint $L_k U_{1,k} = 0$, which is a necessary condition for obtaining an unbiased estimate. ■

One choice of Γ_k (first proposed in [4] using the singular value decomposition of $\hat{R}_k^{-\frac{1}{2}} C_k G_{2,k-1} = \tilde{U}_k \tilde{\Sigma}_k \tilde{V}_k^\top$) such that $\Gamma_k \tilde{R}_k^\star \Gamma_k^\top$ has full rank, is given by $\Gamma_k = \begin{bmatrix} 0 \; I_{p_R} \end{bmatrix} \tilde{U}_k^\top \hat{R}_k^{-\frac{1}{2}}$, where $\hat{R}_k$ and $\tilde{R}_k^\star$ are defined in the text following (25), and $p_R = l - p - p_{H_{k-1}}$. With this Γ_k, we obtain $\Gamma_k \tilde{R}_k^\star \Gamma_k^\top = I_{p_R}$ which is invertible. Following the procedure in [4, Appendix], (26) reduces to $L_k = K_k (I_l - H_{1,k} M^\star_{1,k})^\top \hat{R}_k^{-1}$, with $M^\star_{1,k} := \Sigma_k^{-1} (U_{1,k}^\top \hat{R}_k^{-1} N_k U_{1,k})^{-1} U_{1,k}^\top \hat{R}_k N_k$, which is independent of $\tilde{U}_k$ and as such, the "expensive" singular value decomposition step can be bypassed. Another choice would be to use the Moore-Penrose pseudoinverse (†) such that $\check{R}_k = (\tilde{R}_k^\star)^\dagger$ in (26). Moreover, we have $L_k = L_k \begin{bmatrix} U_{1,k} \; U_{2,k} \end{bmatrix} \begin{bmatrix} U_{1,k}^\top \\ U_{2,k}^\top \end{bmatrix} = \tilde{L}_k U_{2,k}^\top$ where we defined $\tilde{L}_k := L_k U_{2,k} = K_k (I_l - H_{1,k} M^\star_{1,k})^\top \hat{R}_k^{-1} U_{2,k}$.

In addition, we can compute the (cross-)covariances as

$$\begin{aligned} P^{xd}_{1,k} &= (P^{dx}_{1,k})^\top = -P^x_{k|k} C_{1,k}^\top M_{1,k}^\top \\ P^{xd}_{2,k-1} &= (P^{dx}_{2,k-1})^\top = -P^x_{k-1|k-1} A_{k-1}^\top C_{2,k}^\top M_{2,k}^\top \\ &\quad - P^{xd}_{1,k-1} G_{1,k-1}^\top C_{2,k}^\top M_{2,k}^\top \\ P^d_{12,k-1} &= (P^d_{21,k-1})^\top = -P^{dx}_{1,k-1} A_{k-1}^\top C_{2,k}^\top M_{2,k}^\top \\ &\quad - P^d_{1,k-1} G_{1,k-1}^\top C_{2,k}^\top M_{2,k}^\top \end{aligned}$$

$$P^d_k := \begin{bmatrix} V_{1,k} \; V_{2,k} \end{bmatrix} \begin{bmatrix} P^d_{1,k} & P^d_{12,k} \\ P^d_{21,k} & P^d_{2,k} \end{bmatrix} \begin{bmatrix} V_{1,k}^\top \\ V_{2,k}^\top \end{bmatrix}.$$

5.4 Global optimality of ULISE

In the following, we relax the recursivity assumption of ULISE for both the state and input estimates and consider $\hat{x}_{k|k}$ and $\hat{d}_k$ to be the most general linear combination of the unbiased initial state estimate $\hat{x}_{0|0}$, $\mathcal{Z}_{1,k} = \begin{bmatrix} z_{1,0}^\top \; z_{1,1}^\top \ldots z_{1,k}^\top \end{bmatrix}^\top$ and $\mathcal{Z}_{2,k} = \begin{bmatrix} z_{2,0}^\top \; z_{2,1}^\top \ldots z_{2,k}^\top \end{bmatrix}^\top$. We first prove that the state update of ULISE has the same optimal form as the filter proposed in [3, Remark 3], through which the claim of global optimality of the state estimate over the class of all linear estimators follows from [18]. Then, we prove the global optimality of the input estimate, which completes the proof of Theorem 4.

Proof of Theorem 4. To this end, we rearrange the latter form of (13) of state estimation for ULISE with unknown inputs estimated with (8) and (9), to obtain

$$\begin{aligned} \hat{x}_{k|k} &= \hat{A}_{k-1} \hat{x}_{k-1|k-1} + B_{k-1} u_{k-1} + G_{1,k-1} M_{1,k-1} z_{1,k-1} \\ &\quad + \overline{K}_k (z_{2,k} - C_{2,k} (\hat{A}_{k-1} \hat{x}_{k-1|k-1} + B_{k-1} u_{k-1} \\ &\quad + G_{1,k-1} M_{1,k-1} z_{1,k-1})) \end{aligned} \tag{27}$$

$$\overline{K}_k = G_{2,k-1} M_{2,k} + \tilde{L}_k (I - C_{2,k} G_{2,k-1} M_{2,k}) \tag{28}$$

where $\hat{A}_{k-1} = A_{k-1} - G_{1,k-1} M_{1,k-1} C_{1,k-1}$, as previously defined. Repeating the procedure in Section 5.3, we obtain the optimal gain and the updated state estimate error covariance as

$$\begin{aligned} \tilde{L}_k &= (\tilde{P}_k C_{2,k}^\top - G_{2,k-1} M_{2,k} \tilde{R}_{2,k}) \overline{N}_k^\top (\overline{N}_k \tilde{R}_{2,k} \overline{N}_k^\top)^{-1} \overline{\Gamma}_k \\ P^x_{k|k} &= (I - \overline{K}_k C_{2,k}) \tilde{P}_k (I - \overline{K}_k C_{2,k})^\top + \overline{K}_k R_{2,k} \overline{K}_k^\top \end{aligned}$$

where $\overline{N}_k := \overline{\Gamma}_k (I - C_{2,k} G_{2,k-1} M_{2,k})$, $\tilde{R}_{2,k} := C_{2,k} \tilde{P}_k C_{2,k}^\top + R_{2,k}$ and $\overline{\Gamma}_k$ is an arbitrary matrix such that $\overline{N}_k \tilde{R}_{2,k} \overline{N}_k^\top$ has full rank. Thus, the ULISE's state and state covariance update is almost identical to [3], in which only state estimation is considered. The only difference is in the choice of $M_{2,k}$, where $M_{2,k}$ is replaced by $\tilde{M}_{2,k} := (C_{2,k} G_{2,k-1})^\dagger$ in [3]. More importantly, the state update law is of the optimal form [3, Remark 3] from which the global optimality of the state estimate over the linear class of estimators follows from [18].

To show that the input estimate is also globally optimal, we consider the input estimate $\hat{d}^g_{k-1}$ to be the most general linear combination of the unbiased initial state estimate $\hat{x}_{0|0}$, as well as $\mathcal{Z}_{1,k} = \begin{bmatrix} z_{1,0}^\top \; z_{1,1}^\top \ldots z_{1,k}^\top \end{bmatrix}^\top$ and $\mathcal{Z}_{2,k} = \begin{bmatrix} z_{2,0}^\top \; z_{2,1}^\top \ldots z_{2,k}^\top \end{bmatrix}^\top$. Since $\tilde{z}_{1,i}$ and $\tilde{z}_{2,i}$ as defined for (20) and (21) are linear combinations of $\hat{x}_{0|0}$, $\mathcal{Z}_{1,i}$ and $\mathcal{Z}_{2,i}$, and of $\hat{x}_{0|0}$, $\mathcal{Z}_{1,i-1}$ and $\mathcal{Z}_{2,i}$, respectively,

$\hat{d}^g_{k-1} = \chi_0(k)\hat{x}_{0|0} + \sum_{i=1}^{k}\chi_{1,i}(k)\tilde{z}_{1,i} + \sum_{i=1}^{k}\chi_{2,i}(k)\tilde{z}_{2,i}$.

Clearly, if $\chi_{1,k-1}(k) = V_{1,k-1}M_{1,k-1}$ and $\chi_{2,k}(k) = V_{2,k-1}M_{2,k}$ where $M_{1,k-1}$ and $M_{2,k}$ are as in (18) and (19), and if $\chi_0(k)$, $\chi_{1,k}(k)$, $\{\chi_{1,i}(k)\}_{i=0}^{k-2}$ and $\{\chi_{2,i}(k)\}_{i=0}^{k-1}$ are zero, then $\hat{d}^g_{k-1}$ is unbiased. To show the converse, we suppose that $\hat{d}^g_{k-1}$ is unbiased, i.e., $\mathbb{E}[\hat{d}^g_{k-1}] = V_{1,k-1}d_{1,k-1} + V_{2,k-1}d_{2,k-1}$. Since d_k can take on any arbitrary value and $z_{1,k}$ is a function of $d_{1,k}$, $\chi_{1,k}(k) = 0$ such that $\hat{d}^g_{k-1}$ remains unbiased. Moreover, the first measurements containing $d_{1,k-1}$ and $d_{2,k-1}$ are $z_{1,k-1}$ and $z_{2,k}$, then $\mathbb{E}[\chi_{1,k-1}(k)\tilde{z}_{1,k-1}] = V_{1,k-1}d_{1,k-1}$ and $\mathbb{E}[\chi_{2,k}(k)\tilde{z}_{2,k}] = V_{2,k-1}d_{2,k-1}$. Consequently, $\chi_{1,k-1}(k) = V_{1,k-1}M_{1,k-1}$ and $\chi_{2,k}(k) = V_{2,k-1}M_{2,k}$. Moreover, for $\hat{d}^g_{k-1}$ to be unbiased, $\chi_0(k) = 0$, $\{\chi_{1,i}(k)\}_{i=0}^{k-2} = 0$ and $\{\chi_{2,i}(k)C_{2,i}G_{2,i-1}\}_{i=0}^{k-1} = 0$ must hold. Finally, we prove that the mean squared error $\mathbb{E}[\|d_{k-1} - \hat{d}^g_{k-1}\|_2^2]$ is minimized when $\{\chi_{2,i}(k)\}_{i=0}^{k-1} = 0$. From the unbiasedness conditions of $\hat{d}^g_{k-1}$, we have $d_{k-1} - \hat{d}^g_{k-1} = \tilde{d}_{k-1} - \sum_{i=0}^{k-1}\chi_{2,i}(k)\tilde{z}_{2,i}$ where $\tilde{d}_k$ is as defined above Lemma 8. Since it is straightforward to verify (as in [18, Lemmas 1 and 2]) that $\mathbb{E}[\tilde{d}_k(\chi_{2,i}(k)\tilde{z}_{2,i})^\top] = 0$ for all $i \leq k$, it follows that

$$\begin{aligned}\mathbb{E}[\|d_{k-1} - \hat{d}^g_{k-1}\|_2^2] &= \mathrm{tr}\{\mathbb{E}[(\tilde{d}_{k-1} - \textstyle\sum_{i=0}^{k-1}\chi_{2,i}(k)\tilde{z}_{2,i}) \\ &\qquad (\tilde{d}_{k-1} - \textstyle\sum_{i=0}^{k-1}\chi_{2,i}(k)\tilde{z}_{2,i})^\top]\} \\ &= \mathrm{tr}\{\mathbb{E}[\tilde{d}_{k-1}\tilde{d}_{k-1}^\top]\} + \mathbb{E}[\|\textstyle\sum_{i=0}^{k-1}\chi_{2,i}(k)\tilde{z}_{2,i}\|_2^2]\end{aligned}$$

where the first term is minimized by ULISE as is shown in (23) and Theorem 9, while the latter term is minimized when $\sum_{i=0}^{k-1}\chi_{2,i}(k)\tilde{z}_{2,i} = 0$, which occurs when $\{\chi_{2,i}(k)\}_{i=0}^{k-1} = 0$, as desired. Thus, Theorem 4 holds. ■

Remark 12 *ULISE provides a family of optimal state estimators parameterized by Γ_k, while the filter in [3] provides a specific solution with $\overline{N}_k$ as the left null matrix of $C_{2,k}G_{2,k}$, i.e., $\overline{N}_k = \mathrm{Null}((C_{2,k}G_{2,k})^\top)^\top$. More importantly, we have shown that the decorrelation constraint assumed in [3], such that only $z_{2,k}$ can be used in the state update to avoid obtaining a suboptimal estimator, is justified as a direct consequence of the unbiasedness constraint in Lemma 8, i.e., $L_kU_{1,k} = 0$.*

5.5 Stability of ULISE

In this section, we prove the stability of the ULISE filter by first reducing the linear time-varying system with unknown inputs to an equivalent system without unknown inputs. Then, we use existing results on the stability of the Kalman filter [1, Section 5] to obtain the sufficient conditions for the stability of the original system.

Proof of Theorem 5. We begin by reducing the system with unknown inputs to one without unknown inputs. From (13) and (6), we obtain $\tilde{x}_{k|k} = \tilde{x}^\star_{k|k} - \tilde{L}_k(C_{2,k}\tilde{x}^\star_{k|k} + v_{2,k})$. Then, substituting (22) into (16) and the above equation, and rearranging, we obtain

$$\begin{aligned}\tilde{x}_{k|k} = \overline{A}_{k-1}\tilde{x}_{k-1|k-1} + \overline{w}_{k-1} - \tilde{L}_k(C_{2,k}\overline{A}_{k-1}\tilde{x}_{k-1|k-1} \\ + C_{2,k}\overline{w}_{k-1} + v_{2,k}),\end{aligned} \tag{29}$$

where $\overline{A}_{k-1} = (I - G_{2,k-1}M_{2,k}C_{2,k})\hat{A}_{k-1}$ and $\overline{w}_{k-1} = -(I - G_{2,k-1}M_{2,k}C_{2,k})(G_{1,k-1}M_{1,k-1}v_{1,k-1} - w_{k-1}) - G_{2,k-1}M_{2,k}v_{2,k}$. As it turns out, the state estimate error dynamics above is the same for a Kalman filter [17] for a linear system without unknown inputs: $x^e_{k+1} = \overline{A}_k x^e_k + \overline{w}_k$; $y^e_k = C_{2,k}x^e_k + v_{2,k}$. 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. However, the noise terms of this equivalent system are correlated, i.e., $\mathbb{E}[\overline{w}_k v_{2,k}^\top] = -G_{2,k-1}M_{2,k}R_{2,k}$. To transform the system further into one without correlated noise, we employ a common trick of adding a zero term $-G_{2,k-1}M_{2,k}(y^e_k - C_{2,k}x^e_k - v_{2,k})$ to obtain: $x^e_{k+1} = \bar{\bar{A}}_k x^e_k + \bar{\bar{u}}_k + \bar{\bar{w}}_k$; $y^e_k = C_{2,k}x^e_k + v_{2,k}$, where $\bar{\bar{A}}_k = \overline{A}_k + G_{2,k-1}M_{2,k}C_{2,k}$, $\bar{\bar{u}}_k = -G_{2,k-1}M_{2,k}y^e_k$ is a known input and $\bar{\bar{w}}_k = \overline{w}_k + G_{2,k-1}M_{2,k}v_{2,k}$. The new noise terms $\bar{\bar{w}}_k$ and $v_{2,k}$ are uncorrelated with covariances $\bar{\bar{Q}}_k := \mathbb{E}[\bar{\bar{w}}_k\bar{\bar{w}}_k^\top] = (I - G_{2,k-1}M_{2,k}C_{2,k})\hat{Q}_{k-1}(I - G_{2,k-1}M_{2,k}C_{2,k})^\top$, $R_{2,k}$ and $\mathbb{E}[\bar{\bar{w}}_k v_{2,k}^\top] = 0$, where $\hat{Q}_{k-1}$ and $M_{2,k}$ are as defined in Theorem 9.

Ideally, if we can compute $\bar{\bar{A}}$ and $\bar{\bar{Q}}$ prior to applying the ULISE algorithm, then the uniform detectability and stabilizability conditions of [1, Section 5] can be directly applied to obtain the desired stability property. However, this is not the case as these matrices depend on $P^x_{k-1|k-1}$ which is not available a priori. Thus, we substitute $M_{2,k}$ in (9) with $\tilde{M}_{2,k} := (C_{2,k}G_{2,k-1})^\dagger$ to obtain $\tilde{A}_k := (I - G_{2,k-1}\tilde{M}_{2,k}C_{2,k})\hat{A}_{k-1} + G_{2,k-1}\tilde{M}_{2,k}C_{2,k}$ and $\tilde{Q}_k := (I - G_{2,k-1}\tilde{M}_{2,k}C_{2,k})\hat{Q}_{k-1}(I - G_{2,k-1}\tilde{M}_{2,k}C_{2,k})^\top$. This removes the dependence on $P^x_{k-1|k-1}$ from the uniform detectability and stabilizability tests in Theorem 5.

From [1, Lemma 5.1 & Corollary 5.2], if $(\tilde{A}_k, C_{2,k})$ is uniformly detectable, then the corresponding filter error covariance $P^{x,sub}_{k|k}$ is bounded. By the optimality of the ULISE algorithm, it follows that the ULISE error covariance $P^x_{k|k}$ and the filter gain $\tilde{L}_k$ are bounded. Next, by [1, Theorems 4.3 & 5.3], the uniform stability of $(\tilde{A}_k, \tilde{Q}_k^{\frac{1}{2}})$ and the boundedness of $\tilde{L}_k$ imply that the filter (with $\tilde{L}_k$ but with $\tilde{M}_{2,k}$ in the input estimate) is exponentially stable. Finally, using the fact that the ordinary and generalized least squares input estimates have the same expected value (see, e.g., [7, pp. 223-224]), it can be verified from (29) that $\mathbb{E}[\tilde{x}_{k|k}] = (I - \tilde{L}_kC_{2,k})\overline{A}_{k-1}\mathbb{E}[\tilde{x}_{k-1|k-1}] = (I - \tilde{L}_kC_{2,k})\tilde{A}_{k-1}\mathbb{E}[\tilde{x}_{k-1|k-1}]$, from which it follows that the uniform stability of $(\tilde{A}_k, \tilde{Q}_k^{\frac{1}{2}})$ and the boundedness of $\tilde{L}_k$ also imply that ULISE is exponentially stable. ■

Next, we consider the time-invariant case, for which uniform detectability and uniform stabilizability reduce to standard definitions of detectability and stabilizability [23]. Thus, the sufficient conditions of Theorem 6 follow directly. In addition, noting the similarity of ULISE to the state estimator in [3] and that the conditions given in [4] is independent of the choice of $M_{2,k}$ or $\tilde{M}_{2,k}$, it can be shown that the convergence and stability conditions are as given in Theorem 6.

6 Illustrative Example

In this example, we consider the state estimation and fault identification problem when the system dynamics is plagued by faults, d_k, as well as zero-mean Gaussian white noise. Specifically, the linear discrete-time problems we consider are based on the system given in [3] with three different H matrices to illustrate the effect of parameter changes on filter performance:

$$A = \begin{bmatrix} 0.5 & 2 & 0 & 0 & 0 \\ 0 & 0.2 & 1 & 0 & 1 \\ 0 & 0 & 0.3 & 0 & 1 \\ 0 & 0 & 0 & 0.7 & 1 \\ 0 & 0 & 0 & 0 & 0.1 \end{bmatrix}; \begin{matrix} B = 0_{5\times 1}; \\ C = I_5; \\ D = 0_{5\times 1}; \end{matrix} \quad G = \begin{bmatrix} 1 & 0 & -0.3 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix};$$

$$Q = 10^{-4}\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0.5 & 0 & 0 \\ 0 & 0.5 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}; R = 10^{-2}\begin{bmatrix} 1 & 0 & 0 & 0.5 & 0 \\ 0 & 1 & 0 & 0 & 0.3 \\ 0 & 0 & 1 & 0 & 0 \\ 0.5 & 0 & 0 & 1 & 0 \\ 0 & 0.3 & 0 & 0 & 1 \end{bmatrix};$$

$$H^1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \quad H^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}; \quad H^3 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

The unknown inputs used in this example are as given in Figure 1. With the above H matrices, the invariant zeros of the system (cf. Theorem 2) are respectively $\{0.3, 0.8\}$, $\{0.1, 0.3, 0.5, 0.7, 0.8\}$, and $\{0.1, 0.7, 0.3, -0.8, 0.35\}$. Thus, all three systems are strongly detectable. Moreover, H^2 and H^3 are full rank.

To illustrate the performance of the unified simultaneous input and state estimators, measured by the steady-state trace of the error covariance matrices, we compare the performance of the following filters: (i) Cheng et al. filter [3], augmented by implicit estimates the unknown input, i.e., using (8) (with $\tilde{M}_{2,k}$) and (9) (CYWZ), (ii) ULISE from Section 4, as well as the filters for systems with full-rank H matrix: (iii) Gillijns and De Moor filter (GDM) [11], (iv) Fang et al. filter (FSY) [9] and (v) Yong et al. filter (YZF) [30]. The simulations were implemented in MATLAB on a 2.2 GHz Intel Core i7 CPU.

Figures 1 and 2 show a comparison of the input and state estimates as well as the error covariances of the first two MVU estimators for the system with H^1. In this case, CYWZ and ULISE were equally successful at estimating the states and the unknown inputs. Note also that ULISE is consistently the best filter (cf. Table 1), which agrees with the claim in Section 5.4 of being globally optimal over the class of all linear unbiased state and input estimates for systems with unknown inputs, while CYWZ performs just as well, which shows that in this particular example, the replacement of the generalized least squares estimate of $d_{2,k}$ with the ordinary least squares estimate has little impact on the filter performance. When the direct feedthrough matrix has full rank, as with H^2 and H^3, GDM and YZF performed just as well as CYWZ and ULISE, which is consistent with the claim of global optimality of GDM in [16].

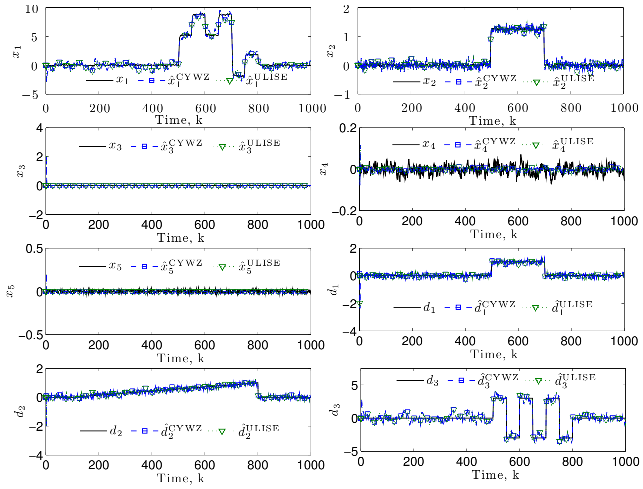

Fig. 1. Actual states x_1, x_2, x_3, x_4, x_5 and their estimates, as well as unknown inputs d_1, d_2 and d_3 and their estimates.

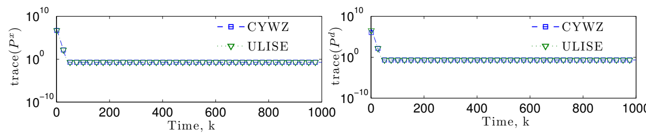

Fig. 2. Trace of estimate error covariance of states, $\mathrm{tr}(P^x)$, and unknown inputs, $\mathrm{tr}(P^d)$.

Table 1
Steady-state Performance of CYWZ, ULISE, GDM, FSY and YZF.

		P^x_{11}	P^x_{22}	P^x_{33}	P^x_{44}	P^x_{55}	P^d_{11}	P^d_{22}	P^d_{33}
H^1	CYWZ	0.1843	0.0091	0.0002	0.0004	0.0001	0.0099	0.0102	0.1923
	ULISE	0.1843	0.0091	0.0002	0.0004	0.0001	0.0099	0.0102	0.1923
	GDM	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	FSY	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	YZF	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
H^2	CYWZ	0.1494	0.0052	0.0002	0.0004	0.0001	0.0097	0.0102	0.1574
	ULISE	0.1494	0.0052	0.0002	0.0004	0.0001	0.0097	0.0102	0.1574
	GDM	0.1494	0.0052	0.0002	0.0004	0.0001	0.0097	0.0102	0.1574
	FSY	0.1724	0.0108	0.0002	0.0004	0.0001	0.0097	0.0102	0.1648
	YZF	0.1494	0.0052	0.0002	0.0004	0.0001	0.0097	0.0102	0.1574
H^3	CYWZ	0.0076	0.0218	0.0002	0.0004	0.0001	0.0309	0.0102	0.0097
	ULISE	0.0076	0.0218	0.0002	0.0004	0.0001	0.0309	0.0102	0.0097
	GDM	0.0076	0.0218	0.0002	0.0004	0.0001	0.0309	0.0102	0.0097
	FSY	0.0315	0.0232	0.0002	0.0004	0.0001	0.0310	0.0102	0.0100
	YZF	0.0076	0.0218	0.0002	0.0004	0.0001	0.0309	0.0102	0.0097

7 Conclusion

This paper presented a unified filter for simultaneously estimating the states and unknown inputs in an unbiased minimum-variance sense for linear discrete-time stochastic systems, without any restriction on the direct feedthrough matrix of the system. We proved that ULISE is globally optimal over the class of all linear unbiased state and input estimators for systems with unknown inputs and provided stability conditions for the filter, which are shown to be closely related to the strong detectability of the system. Simulation results have shown that ULISE was the best estimator in all the test trials.

References

[1] B.D.O. Anderson and J.B. Moore. Detectability and stabilizability of time-varying discrete-time linear systems. *SIAM Journal on Control and Optimization*, 19(1):20–32, 1981.

[2] D.S. Bernstein. *Matrix Mathematics: Theory, Facts, and Formulas (Second Edition)*. Princeton reference. Princeton University Press, 2009.

[3] 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.

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

[5] M. Darouach, M. Zasadzinski, and M. Boutayeb. Extension of minimum variance estimation for systems with unknown inputs. *Automatica*, 39(5):867–876, 2003.

[6] 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.

[7] N.R. Draper and H. Smith. *Applied Regression Analysis (Wiley Series in Probability and Statistics)*. Wiley-Interscience, third edition, April 1998.

[8] H. Fang, Y. Shi, and J. Yi. A new algorithm for simultaneous input and state estimation. In *American Control Conference*, pages 2421–2426, 2008.

[9] 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.

[10] 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.

[11] 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.

[12] M.L.J. Hautus. Strong detectability and observers. *Linear Algebra and its Applications*, 50:353–368, 1983.

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

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

[15] C. Hsieh. Extension of unbiased minimum-variance input and state estimation for systems with unknown inputs. *Automatica*, 45(9):2149–2153, 2009.

[16] C. Hsieh. On the global optimality of unbiased minimum-variance state estimation for systems with unknown inputs. *Automatica*, 46(4):708–715, 2010.

[17] 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.

[18] W.S. Kerwin and J.L. Prince. On the optimality of recursive unbiased state estimation with unknown inputs. *Automatica*, 36(9):1381–1383, 2000.

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

[20] X. Li and H. Gao. Robust finite frequency filtering for uncertain 2-D systems: The FM model case. *Automatica*, 49(8):2446–2452, 2013.

[21] H.J. Palanthandalam-Madapusi and D.S. Bernstein. Unbiased minimum-variance filtering for input reconstruction. In *American Control Conference*, pages 5712–5717, 2007.

[22] 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.

[23] M.A. Peters and P.A. Iglesias. A spectral test for observability and reachability of time-varying systems. *SIAM Journal on Control Optimization*, 37(5):1330–1345, August 1999.

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

[25] L.M. Silverman. *Discrete Riccati equations: Alternative algorithms, asymptotic properties, and system theory interpretations*, volume 12 of *Control and Dynamic Systems*. Academic Press, 1976.

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

[27] Z. Wang, H. Dong, B. Shen, and H. Gao. Finite-horizon H_∞ filtering with missing measurements and quantization effects. *IEEE Transactions on Automatic Control*, 58(7):1707–1718, July 2013.

[28] Z. Wang, Y. Liu, and X. Liu. H_∞ filtering for uncertain stochastic time-delay systems with sector-bounded nonlinearities. *Automatica*, 44(5):1268–1277, 2008.

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

[30] 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.

[31] S.Z. Yong, M. Zhu, and E. Frazzoli. Generalized innovation and inference algorithms for hidden mode switched linear stochastic systems with unknown inputs. *Conference on Decision and Control*, pages 3388–3394, 2014.

[32] W. Zhang, Y. Huang, and L. Xie. Infinite horizon stochastic H_2/H_∞ control for discrete-time systems with state and disturbance dependent noise. *Automatica*, 44:2306–2316, 2008.

[33] W. Zhang, H. Zhang, and B.-S. Chen. Generalized Lyapunov equation approach to state-dependent stochastic stabilization/detectability criterion. *IEEE Transactions on Automatic Control*, 53(7):1630–1642, 2008.