Simultaneous Input and State Estimation for Linear Discrete-time Stochastic Systems with Direct Feedthrough

Sze Zheng Yong[1]   Minghui Zhu[2]   Emilio Frazzoli[1]

Abstract— In this paper, we present an optimal filter for linear discrete-time stochastic systems with direct feedthrough that simultaneously estimates the states and unknown inputs in an unbiased minimum-variance sense. We argue that the information about the unknown input can be obtained from the current time step as well as the previous one, making it possible to estimate the unknown input in different ways. We then propose one variation of the filter that uses the updated state estimate to compute the best linear unbiased estimate (BLUE) of the unknown input. The comparison of the new filter and the filters in existing literature is discussed in detail and tested in simulation examples.

I. INTRODUCTION

Kalman filtering provides the tool needed for obtaining a reliable estimate from measured data corrupted by noise when the system is linear and when an accurate model of the process dynamics and observations is available. However, in many instances, systematic measurement errors or model uncertainties are inevitable, for example, in the setting of semi-autonomous multi-vehicle systems, the input of the other vehicle is inaccessible/unmeasurable [1]. Nonetheless, we want to be able to estimate the states of the other vehicle based on noisy measurements for purposes of collision avoidance, route planning, etc. Moreover, estimates of the unknown inputs may be used to predict the intention of the other vehicle or to improve control performance.

This same problem can be found across a wide range of disciplines, from the real-time estimation of mean areal precipitation during a storm [2] to fault detection and diagnosis [3] to input estimation in physiological systems [4]. Thus, this filtering problem in the presence of errors and uncertainties, which oftentimes are modeled as unknown disturbance inputs, has steadily made it to the forefront in the recent decades. Research in this field began with state estimation of systems with unknown biases [5] and unknown disturbance of known dynamics [6], but has since moved towards state estimation with arbitrary unknown inputs.

Literature review. An optimal filter that estimates a minimum-variance unbiased (MVU) estimate is first developed for linear systems without direct feedthrough in [2]. This design was extended to a more general parameterized

solution by [7], and eventually to state estimation of systems with direct feedthrough in [8]–[10]. While most of the research focus had only been on state estimation in the presence of unknown inputs, the problem of estimating the unknown input itself is gaining more attention, as researchers realize that the information about the unknown input is often as important as state information. Thus, in addition to MVU state estimation, an approach is proposed in [11] to reconstruct the unknown inputs, in a process that is decoupled from state estimation. This development was followed by the design of simultaneous input and state estimation filters for systems without direct feedthrough in [12], [13], with the objective of *concurrently* obtaining minimum-variance unbiased estimates for both the states and the unknown disturbance inputs to the system. Soon after, extensions to systems with direct feedthrough were proposed by [14]–[16].

Contributions. This paper proposes a new minimum-variance unbiased (MVU) filter for both input and states for systems with direct feedthrough, a problem first considered in [12]–[14]. This is in contrast to the majority of filters, which seek an MVU estimate only of the systems states, e.g. [2], [11]. Interestingly, for the case without direct feedthrough, the MVU state filters of [2], [11], [12], was shown by [13] to implicitly estimate the input in an unbiased minimum variance sense. However, whether this is also true in the case with direct feedthrough is unclear.

More interestingly for systems with direct feedthrough, the unknown inputs affect the system in two ways, both in the state dynamics and in the measurement. Thus, this presents two approaches of estimating the unknown inputs, as shown in [11] – one with and another without a one step delay. The fact that the unknown inputs of systems with and without direct feedthrough are estimated differently with, or without one step delay was also briefly noted in [14], while in [9], there are two proposed filters – an estimator filter that optimally estimates the current state based on current measurements, and a predictor filter which optimally predicts the current state based on previous measurements. Therefore, the unknown input can potentially be estimated in three different ways: solely based on previous measurement and previous state estimate, based on the current measurement and the propagated/predicted current state, or based on the current measurement with the updated/estimated current state.

The MVU input and state estimator in Fang et al. (2011) [16] falls into the first category, as the state and input are estimated with one step delay, in a purely predictive manner. The version of the filter with one step delay in

[1] 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).

[2] M. Zhu is with the Department of Electrical Engineering, Pennsylvania State University, 201 Old Main, University Park, PA 16802, USA (e-mail: muz16@psu.edu). This work was done when M. Zhu was with the Laboratory for Information and Decision Systems at Massachusetts Institute of Technology.

Palanthandalam and Bernstein (2007) [11] is also in this category, although only the states are estimated in an MVU manner. On the other hand, the MVU input and state estimator in Gillijns and De Moor (2007b) [14] belongs to the second category, since the unknown input is estimated from the current measurement and the propagated state estimate, in sort of a half-step delay fashion. In the third category, we have the version of the filter without one step delay in Palanthandalam and Bernstein (2007) [11]. However, as in the case with one step delay, the inputs are reconstructed such that it is unbiased but it is not BLUE. Thus, to bridge this gap, we propose an MVU input and state estimator for linear discrete-time systems with direct feedthrough in Section IV, which estimates the unknown input based on the current measurement and its updated state estimate, as opposed to the predicted state estimate in [14]. Since the updated state estimate is expected to have a lower variance than the propagated state estimate (otherwise the update would be counterproductive), we expect the new filter to be no worse than the one based on the propagated state estimate.

For a more detailed discussion of each of these filters, the readers are referred to Section V, in which we compare and contrast the proposed input and state estimator with the filters in existing literature, as well as discuss the limitations of each of them. We shall then highlight the differences in performance of the different filters using several illustrative examples in Section VI.

II. Preliminary Material

We first summarize the notation used throughout the paper. $\mathbb{R}^n$ denotes the n-dimensional Euclidean space, and $\mathbb{N}$ positive integers. For a vector of random variables, $v \in \mathbb{R}^n$, the expectation is denoted by $\mathbb{E}[v]$. Given a matrix $M \in \mathbb{R}^{p \times q}$, its transpose, inverse and Moore-Penrose pseudoinverse are given by $M^\top$, M^{-1} and $M^\dagger$. For a symmetric matrix S, $S \succ 0$ and $S \succeq 0$ indicates that S is positive definite and positive semidefinite, respectively. We also define some basic notions from estimation theory.

Definition 1 (Minimum-variance unbiased estimate (MVUE)). *An estimate of an unknown signal vector θ, denoted $\hat{\theta}$ is unbiased if $\mathbb{E}[\hat{\theta}] = \theta$, i.e. the estimate error $\tilde{\theta} = \theta - \hat{\theta}$ has a zero bias, $\mathrm{bias}(\hat{\theta}) = \mathbb{E}[\tilde{\theta}] = 0$. Furthermore, the estimate is an minimum-variance unbiased estimate, if the variance of the unbiased estimate $\hat{\theta}$, denoted as $\mathrm{var}(\hat{\theta}) = \mathbb{E}[(\hat{\theta} - \mathbb{E}[\hat{\theta}])^\top (\hat{\theta} - \mathbb{E}[\hat{\theta}])]$ is not higher than any other unbiased estimates for all possible values of the signal vector θ.*

Definition 2 (Best linear unbiased estimate (BLUE)). *An estimate $\hat{\theta}$ of the linear form $\hat{\theta} = Ax$, where x is the measured data, is the best linear unbiased estimate (BLUE) of an unknown signal vector θ if the estimate has no higher variance than any other linear unbiased estimates for all possible values of θ.*

Remark 1. *In the special case with a linear data model given by $x = H\theta + w$ where x is the measured data, H is a known matrix with full rank, θ is to be estimated and w is a Gaussian white noise, a BLUE will be a MVUE [17].*

III. Problem Statement

Consider the time-varying discrete-time linear system with direct feedthrough

$$x_{k+1} = A_k x_k + B_k u_k + G_k d_k + w_k \tag{1}$$

$$y_k = C_k x_k + D_k u_k + H_k d_k + v_k \tag{2}$$

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 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, G_k, C_k, D_k and H_k are known and it is assumed that H_k has full column rank, i.e. $\mathrm{rank}(H_k) = p$. x_0 is also assumed to be independent of v_k and w_k for all k and the unbiased estimate $\hat{x}_0$ of the initial state x_0 is available with covariance matrices $\mathcal{P}_0^x$, $\mathcal{P}_0^{xd}$ and $\mathcal{P}_0^d$. In addition, we assume that the system has *perfect/strong observability*, i.e., the initial condition x_0 and the unknown input sequence $\{d_i\}_{i=0}^{r-1}$ can be uniquely determined from the measured output sequence $\{y_i\}_{i=0}^r$ of a sufficient number of observations, i.e., $r \geq r_0$ for some $r_0 \in \mathbb{N}$ (see, e.g., [18]–[20]).

The objective of this paper is to design an optimal recursive filter algorithm which simultaneously estimates the system state x_k and the unknown input d_k based on an initial unbiased estimate $\hat{x}_0$ and the sequence of measurements up to time k, $\{y_0, y_1, \ldots, y_k\}$. No prior knowledge of the dynamics of d_k is assumed and the unknown input can be a signal of any type.

IV. Minimum-variance Unbiased Filter for Input and State Estimation

We consider a recursive three-step filter of the form

$$\hat{x}_{k|k-1} = A_{k-1}\hat{x}_{k-1|k-1} + B_{k-1}u_{k-1} + G_{k-1}\hat{d}_{k-1} \tag{3}$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + L_k(y_k - C_k \hat{x}_{k|k-1} - D_k u_k) \tag{4}$$

$$\hat{d}_k = M_k(y_k - C_k \hat{x}_{k|k} - D_k u_k) \tag{5}$$

where the matrices $M_k \in \mathbb{R}^{p \times l}$ and $L_k \in \mathbb{R}^{n \times l}$ are yet to be determined. The three steps are *time update*, *measurement update* and *input estimation*. Note that this is similar to the three steps in [14] but in a different order. This seemingly small change actually leads to a rather significantly different estimator, as it turns the input estimate into one which is more of an estimated value, rather than a predicted value. In the next subsections, we will discuss each of the three steps of the filter algorithm given in Algorithm 1.

A. Time Update

Given measurements up to time $k - 1$, let $\hat{x}_{k-1|k-1}$ and $\hat{d}_{k-1}$ denote the optimal unbiased estimates of x_{k-1} and d_{k-1}. Then, the current state is predicted using a copy of the plant given by (3). With this, the error in the propagated state estimate and its covariance matrix are given by

$$\tilde{x}_{k|k-1} := x_k - \hat{x}_{k|k-1}$$
$$= A_{k-1}\tilde{x}_{k-1|k-1} + G_{k-1}\tilde{d}_{k-1} + w_{k-1} \tag{6}$$

Algorithm 1 Input and State Estimation Algorithm

1: Initialize: $\hat{x}_{0|0} = \mathbb{E}[x_0]$; $\hat{d}_0 = H_0^\dagger(y_0 - C_0\hat{x}_{0|0} - D_0 u_0)$;
$P_{0|0}^x = \mathcal{P}_0^x$; $P_0^{xd} = \mathcal{P}_0^{xd}$; $P_0^d = \mathcal{P}_0^d$;
2: **for** $k = 1$ to N **do**
   ▷ Time update
3:   $\hat{x}_{k|k-1} = A_{k-1}\hat{x}_{k-1|k-1} + B_{k-1}u_{k-1} + G_{k-1}\hat{d}_{k-1}$;
4:   $P_{k|k-1}^x = A_{k-1}P_{k-1|k-1}^x A_{k-1}^\top + G_{k-1}P_{k-1}^{xd\top}A_{k-1}^\top$
          $+A_{k-1}P_{k-1}^{xd}G_{k-1}^\top + G_{k-1}P_{k-1}^d G_{k-1}^\top + Q_{k-1}$;
5:   $\tilde{R}_k = C_k P_{k|k-1}^x C_k^\top + R_k$;
   ▷ Measurement update
6:   $K_k = P_{k|k-1}^x C_k^\top \tilde{R}_k^{-1}$;
7:   $L_k = K_k(I - H_k(H_k^\top \tilde{R}_k^{-1}H_k)^{-1}H_k^\top \tilde{R}_k^{-1})$;
8:   $\hat{x}_{k|k} = \hat{x}_{k|k-1} + L_k(y_k - C_k\hat{x}_{k|k-1} - D_k u_k)$;
9:   $P_{k|k}^x = (I - L_k C_k)P_{k|k-1}^x(I - L_k C_k)^\top + L_k R_k L_k^\top$;
   ▷ Estimation of unknown input
10: $\tilde{R}_k^\star = (I - C_k L_k)\tilde{R}_k(I - C_k L_k)^\top$;
11: $P_k^d = (H_k^\top \tilde{R}_k^{\star-1}H_k)^{-1}$;
12: $M_k = P_k^d H_k^\top \tilde{R}_k^{\star-1}$;
13: $\hat{d}_k = M_k(y_k - C_k\hat{x}_{k|k} - D_k u_k)$;
14: $P_k^{xd} = -P_{k|k}^x C_k^\top M_k^\top + L_k R_k M_k^\top$
15: **end for**

$$P_{k|k-1}^x := \mathbb{E}[\tilde{x}_{k|k-1}\tilde{x}_{k|k-1}^\top]$$
$$= A_{k-1}P_{k-1|k-1}^x A_{k-1}^\top + G_{k-1}P_{k-1}^{dx}A_{k-1}^\top \quad (7)$$
$$+ A_{k-1}P_{k-1}^{xd}G_{k-1}^\top + G_{k-1}P_{k-1}^d G_{k-1}^\top + Q_{k-1}$$

with $\tilde{x}_{k|k} := x_k - \hat{x}_{k|k}$, $\tilde{d}_k := d_k - \hat{d}_k$, $P_{k|k}^x := \mathbb{E}[\tilde{x}_{k|k}\tilde{x}_{k|k}^\top]$, $P_k^d := \mathbb{E}[\tilde{d}_k\tilde{d}_k^\top]$ and $P_k^{dx} = (P_k^{xd})^\top := \mathbb{E}[\tilde{d}_k\tilde{x}_{k|k}^\top]$, which will be derived in the next subsections.

B. Measurement Update

In the measurement update step, the measurement y_k is used to update the propagated estimate $\hat{x}_{k|k-1}$ with (4) to obtain an updated state estimate. Then, from (2) and (4), the updated state estimate error is

$$\tilde{x}_{k|k} = (I - L_k C_k)\tilde{x}_{k|k-1} - L_k H_k d_k - L_k v_k. \quad (8)$$

In order that the state estimate is unbiased ($\mathbb{E}[\tilde{x}_{k|k}] = 0$) for all possible d_k, we observe that the constraint $L_k H_k = 0$ must be imposed for all k. Note also that the latter term in the measurement update step given in (4) does not contain an $H_k\hat{d}_k$ term as would be expected in the residual/innovation computation. Besides the practical reason that this unknown input estimate is not yet available, this term can also be seen as being nullified by the constraint $L_k H_k = 0$. Then, the covariance matrix of the updated state estimate error is

$$P_{k|k}^x = (I - L_k C_k)P_{k|k-1}^x(I - L_k C_k)^\top + L_k R_k L_k^\top. \quad (9)$$

To obtain an unbiased minimum variance estimator, we then proceed to derive the optimal gain matrix L_k, by minimizing the trace of (9), because the trace represents the sum of the estimation error variances of the states, subject to the constraint $L_k H_k = 0$.

Theorem 1. *The minimum-variance unbiased state estimator is obtained with the gain matrix L_k given by*

$$L_k = P_{k|k-1}^x C_k^\top \tilde{R}_k^{-1}(I - H_k(H_k^\top \tilde{R}_k^{-1}H_k)^{-1}H_k^\top \tilde{R}_k^{-1}) \quad (10)$$

where $\tilde{R}_k := C_k P_{k|k-1}^x C_k^\top + R_k$, if and only if $(H_k^\top \tilde{R}_k^{-1}H_k)$ is nonsingular, i.e. $\mathrm{rank}(H_k) = p$.

Proof. First, we show by induction that the estimates $\hat{x}_{k|k}$ and $\hat{x}_{k|k-1}$ are unbiased, provided that the input estimate is unbiased, which we will ensure in Section IV-C. For the base case, we shall assume that $\mathbb{E}[\tilde{x}_{0|0}] = 0$ and $\mathbb{E}[\tilde{d}_0] = 0$. Thus, by (6) and the fact that the process noise has zero mean, we get $\mathbb{E}[\tilde{x}_{1|0}] = 0$. For the inductive step, we assume that $\mathbb{E}[\tilde{x}_{k|k-1}] = 0$. By (8), $\mathbb{E}[\tilde{x}_{k|k}] = 0$ because we impose the constraint $L_k H_k = 0$ and the measurement noise has zero mean, i.e. $\mathbb{E}[v_k] = 0$. Then, by (6), $\mathbb{E}[\tilde{x}_{k+1|k}] = 0$ since $\mathbb{E}[\tilde{d}_k]$ (see Section IV-C) and $\mathbb{E}[w_k] = 0$. Therefore, by induction, $\mathbb{E}[\tilde{x}_{k|k}] = 0$ and $\mathbb{E}[\tilde{x}_{k|k-1}] = 0$ for all k, which means that $\hat{x}_{k|k}$ and $\hat{x}_{k|k-1}$ are unbiased.

Next, we employ the optimization approach with Lagrange multipliers ($\Lambda_k \in \mathbb{R}^{p \times n}$) in [2], [11], [13], to find the particular gain L_k that minimizes the trace of the covariance matrix $P_{k|k}^x$, while being subjected to the constraint $L_k H_k = 0$ which is a necessary condition for obtaining an unbiased estimate. This constrained optimization problem can be solved using differential calculus with the Lagrangian

$$\mathcal{L}(L_k, \Lambda_k) := \mathrm{trace}(P_{k|k}^x) - 2\,\mathrm{trace}(L_k H_k \Lambda_k^\top).$$

Differentiating the Lagrangian with respect to L_k and Λ_k, and setting it to zero, we obtain

$$\frac{\partial \mathcal{L}}{\partial L_k} = 2(\tilde{R}_k L_k^\top - C_k P_{k|k-1}^x - H_k\Lambda_k^\top) \quad = 0$$
$$\frac{\partial \mathcal{L}}{\partial \Lambda_k} = -2L_k H_k \quad\quad\quad\quad\quad = 0$$

Solving the above linear system of equations, we obtain the optimal gain matrix (10), provided $(H_k^\top \tilde{R}_k^{-1}H_k)$ is nonsingular. ■

C. Input Estimation

Finally, the unknown input can also be estimated, such that it is the best linear unbiased estimate (BLUE). This means that the expected input estimate must be unbiased, i.e. $\mathbb{E}[\hat{d}_k] = d_k$, which will be shown in Theorem 2, and that the mean squared error of the estimate is the lowest possible, as will be shown in Theorem 3.

Theorem 2. *Let $\hat{x}_{k|k}$ be unbiased, then the input estimate given by (5) is unbiased if and only if $M_k H_k = I$, and consequently, $\mathrm{rank}(H_k) = p$.*

Proof. We observe from (2) and (5) that

$$\hat{d}_k = M_k(C_k\tilde{x}_{k|k} + H_k d_k + v_k). \quad (11)$$

By design, the state estimate is unbiased, i.e. $\mathbb{E}[\tilde{x}_{k|k}] = 0$ (see proof of Theorem 1), and the measurement noise is assumed to have zero mean, $\mathbb{E}[v_k] = 0$. Hence, from (11), $\mathbb{E}[\hat{d}_k] = d_k$, i.e. $\hat{d}_k$ is unbiased, if and only if $M_k H_k = I$. It follows that $\mathrm{rank}(H_k) = p$ is a necessary and sufficient condition for the existence of an unbiased input estimate. ■

Theorem 3. *Let $\hat{x}_{k|k}$ be unbiased and $H_k^\top \tilde{R}_k^{\star-1}H_k$ be nonsingular, where $\tilde{R}_k^\star := (I - C_k L_k)\tilde{R}_k(I - C_k L_k)^\top$ is also nonsingular. Then (5) is the best linear input estimate (BLUE) for M_k given by*

$$M_k = (H_k^\top \tilde{R}_k^{\star-1}H_k)^{-1}H_k^\top \tilde{R}_k^{\star-1} \quad (12)$$

while the covariance matrix of the optimal input error estimate and the cross-covariance matrix with $\tilde{x}_{k|k}$ are

$$P_k^d = (H_k^\top \tilde{R}_k^{\star -1} H_k)^{-1} \tag{13}$$

$$P_k^{xd} = (P_k^{dx})^\top = -P_{k|k}^x C_k^\top M_k^\top + L_k R_k M_k^\top. \tag{14}$$

Proof. Let $\tilde{y}_k := y_k - C_k \hat{x}_{k|k} - D_k u_k$. From (2), we have

$$\tilde{y}_k = C_k \tilde{x}_{k|k} + H_k d_k + v_k = H_k d_k + e_k \tag{15}$$

where e_k is defined as $e_k := C_k \tilde{x}_{k|k} + v_k$. The expected value of e_k is $\mathbb{E}[e_k] = 0$, since $\mathbb{E}[\tilde{x}_k] = 0$ (see proof of Theorem 1) and $\mathbb{E}[v_k] = 0$, while its covariance matrix is

$$
\begin{aligned}
\tilde{R}_k^\star &:= \mathbb{E}[e_k e_k^\top] \\
&= C_k P_{k|k}^x C_k^\top + R_k + C_k \mathbb{E}[\tilde{x}_{k|k} v_k^\top] + \mathbb{E}[v_k \tilde{x}_{k|k}^\top] C_k^\top \\
&= (I - C_k L_k) \tilde{R}_k (I - C_k L_k)^\top \tag{16}
\end{aligned}
$$

where we used (8) and the fact that $\mathbb{E}[\tilde{x}_{k|k-1} v_k^\top] = 0$ to get $\mathbb{E}[\tilde{x}_{k|k} v_k^\top] = (\mathbb{E}[v_k \tilde{x}_{k|k}^\top])^\top = -L_k R_k$. Following the estimation approach outlined in [21, pp. 96-98], we can scale (15) by the inverse of any matrix S satisfying $\tilde{R}_k^\star = SS^\top$, which exists for the positive semidefinite matrix $\tilde{R}_k^\star$, to obtain

$$S^{-1} \tilde{y}_k = S^{-1} H_k d_k + S^{-1} e_k = S^{-1} H_k d_k + e_k^\star$$

where $\mathbb{E}[e_k^\star] = 0$ and $\mathbb{E}[e_k^\star e_k^{\star \top}] = I$. In this form, the Gauss-Markov Theorem [21] assumption of a zero-mean $e_k^\star$ with unit variance is satisfied. Hence, the best linear unbiased estimator (BLUE) for $\hat{d}_k$ is given by

$$
\begin{aligned}
\hat{d}_k &= (H_k^\top (S^{-1})^\top S^{-1} H_k)^{-1} H_k^\top (S^{-1})^\top S^{-1} \tilde{y}_k \\
&= (H_k^\top \tilde{R}_k^{\star -1} H_k)^{-1} H_k^\top \tilde{R}_k^{\star -1} \tilde{y}_k \\
&:= M_k (C_k \tilde{x}_{k|k} + H_k d_k + v_k)
\end{aligned}
$$

where M_k is given by (12). Since $M_k H_k = (H_k^\top \tilde{R}_k^{\star -1} H_k)^{-1} H_k^\top \tilde{R}_k^{\star -1} H_k = I$, from (11), the input estimate error, its covariance matrix and its cross-covariance matrix with $\tilde{x}_{k|k}$ are as follows

$$
\begin{aligned}
\tilde{d}_k &= d_k - \hat{d}_k = -M_k e_k \\
P_k^d &= \mathbb{E}[\tilde{d}_k \tilde{d}_k^\top] = M_k \mathbb{E}[e_k e_k^\top] M_k^\top = (H_k^\top \tilde{R}_k^{\star -1} H_k)^{-1} \\
P_k^{xd} &= \mathbb{E}[\tilde{x}_k \tilde{d}_k^\top] = -\mathbb{E}[\tilde{x}_{k|k} \tilde{x}_{k|k}^\top C_k^\top + \tilde{x}_{k|k} v_k^\top] M_k^\top \\
&= -P_{k|k}^x C_k^\top M_k^\top + L_k R_k M_k^\top \qquad \blacksquare
\end{aligned}
$$

Remark 2. *Moreover, if w_k and v_k are white Gaussian noises, then e_k is white and Gaussian, and (5) is also the minimum variance unbiased (MVU) input estimate.*

V. COMPARISON WITH EXISTING LITERATURE RESULTS

A. With direct feedthrough and with $\mathrm{rank}(H_k) = p$

1) Darouach, Zasadzinski and Boutayeb (2003) [9]: An MVU state estimator is presented for a linear discrete-time system with direct feedthrough was first considered in [9]. The optimal estimator filter derived in that paper assumes that $\mathrm{rank}([C_{k+1}G_k \quad H_{k+1}]) = \mathrm{rank}(G_k) + \mathrm{rank}(H_{k+1})$. However, the unknown input is not reconstructed and hence will not be compared to the new estimator proposed in this paper. Note also that assumption of the optimal estimator filter is more restrictive than the assumption of $\mathrm{rank} H_k = p$ employed in the current MVU input and state estimator.

2) Palanthandalam and Bernstein (2007) [11]: The state and input estimator in [11], decouples the state and input estimation process. It first constructs an MVU state estimator, after which the unknown inputs are reconstructed, without feeding the estimate back to the state estimator.

Remark 3. *This construction implicitly assumes the invertibility of $\Phi_k^\top \tilde{R}_k^{-1} \Phi_k$, where $\tilde{R}_k = C_k P_{k|k-1}^x C_k^\top + R_k$ and $\Phi_k := \begin{bmatrix} -H_k & C_k G_{k-1} \end{bmatrix}$. Since $C_{k+1} P_{k+1|k}^x C_{k+1}^\top \succeq 0$, and by assumption, $R_k \succ 0$, then $\tilde{R}_k \succ 0$, implying that the necessary condition for the invertibility of $\Phi_k^\top \tilde{R}_k^{-1} \Phi_k$ is that $\mathrm{rank}([-H_k \quad C_k G_{k-1}]) = 2p$, i.e $\mathrm{rank}(H_k) = p$ and $\mathrm{rank}(C_k G_{k-1}) = p$. Note that this is more restrictive than the assumption of $\mathrm{rank} H_k = p$, assumed in this paper.*

For the case without direct feedthrough, the reconstructed inputs in [11] has been shown to be BLUE [13]. However, in the case with direct feedthrough, as is the focus of this paper, there is no evidence for the reconstructed inputs to be BLUE. In fact, the authors have proposed two different approaches for input reconstruction – one with and another without a one step delay. To compare the "performance" of both approaches for input reconstruction, we have derived the input estimate error covariance matrices of both approaches

a) Input reconstruction with one step delay:

$$P_{k-1}^d = G_{k-1}^\dagger (L_k^\top \tilde{R}_k^{-1} L_k)(G_{k-1}^\dagger)^\top \tag{17}$$

b) Input reconstruction without one step delay:

$$P_k^d = H_k^\dagger (I - C_k L_k) \tilde{R}_k (I - C_k L_k)^\top (H_k^\dagger)^\top = H_k^\dagger \tilde{R}_k^\star (H_k^\dagger)^\top \tag{18}$$

Remark 4. *Note the similarity of the input error covariance matrix of the new MVU input and state estimator in (13) and that of (18). Since the expression for P_k^d in (13) is invertible, then the inverse and pseudoinverse coincide. Thus, if the product of the matrices $H_k^\top \tilde{R}_k^{\star -1} H_k$ has the properties required for $(H_k^\top \tilde{R}_k^{\star -1} H_k)^\dagger = H_k^\dagger (\tilde{R}_k^{\star -1})^\dagger (H_k^\dagger)^\top = H_k^\dagger \tilde{R}_k^\star (H_k^\dagger)^\top$, then the two input estimate error covariance matrices coincide. However, this is in general not the case because the above-mentioned property only holds if H_k has full row rank. Therefore, we can infer that the MVU input and state estimator proposed in this paper is of the kind without one step delay. Moreover, assuming that the value of both state error covariance matrices are the same, the input reconstructed in [11] is also BLUE if H_k has full row rank.*

3) Gillijns and De Moor (2007b) [14]: Similar to the filter proposed in this paper, the MVU input and state estimator in [14] uses a recursive three-step filter but with different choices of the gain matrices, i.e. $M_k = (H_k^\top \tilde{R}_k^{-1} H_k)^{-1} H_k^\top \tilde{R}_k^{-1}$ and $L_k = P_{k|k-1}^x C_k^\top \tilde{R}_k^{-1} (I - H_k M_k)$, where $\tilde{R}_k = C_k P_{k|k-1}^x C_k^\top + R_k$. At first glance, this MVU input and state estimator is almost identical to the one proposed in this paper. In fact, the assumption for the existence of a solution is the same for both estimators, i.e. $\mathrm{rank}(H_k) = p$. The main difference lies in the input estimation, which in this case is based on the propagated state estimate, whereas the input estimator we proposed uses the updated state estimate. However, unlike the approach

in [11], input estimation is coupled with state estimation. Most notably, all else being equal, the input error covariance matrices of [14] and the estimator we proposed, differ only in that the former contains $\tilde{R}_k$ and the latter $(I - L_k C_k)\tilde{R}_k(I - C_k L_k)^\top$. This difference is especially noteworthy because the input estimate error covariance matrix is typically used as a measure of performance of input estimation approaches.

4) Fang, Shi and Yi (2011) [16]: Most recently, an estimator has been proposed in [16] to estimate the unknown input and states of a linear discrete-time system with direct feedthrough. The approach is "purely" predictive, in that the unknown input estimated as a BLUE is used with a one step delay. The state is then propagated based on the known and estimated unknown input as well as the state estimate from the previous step. In the state update step, the propagated state estimate is also updated with the prediction error from the previous step, as opposed to the current prediction error commonly used in Kalman filtering. Similar to the filter in [14] and the new filter we proposed, the only assumption for the existence of a solution is that $\text{rank}(H_k) = p$, which is less restrictive than [11]. However, this filter is only near optimal. Although the state and input estimates are unbiased, their variances are only "near minimum", as restructuring/modification of the filter matrices is necessary to ensure numerical feasibility.

B. With direct feedthrough and with $\text{rank}(H_k) < p$

1) Hsieh (2009) [15]: The estimator in [15] presents an approach to extend the results in [14] to systems with H_k which does not have full column rank. However, this estimator relaxes the unbiasedness condition for input estimates, thus the input estimate is not BLUE or MVUE.

2) Cheng et al. (2009) [10]: The filter proposed in [10] minimizes only the state estimate error variance, while maintaining the unbiasedness of the estimate. However, the strict assumption of [9] on the rank of $\begin{bmatrix} C_{k+1}G_k & H_{k+1} \end{bmatrix}$ is relaxed via singular value decomposition of H_k. However, since the unknown input is not estimated, we are unable to fully compare the performance of this filter in Section V-A.

VI. ILLUSTRATIVE EXAMPLES

This section considers a family of linear discrete-time problems with direct feedthrough based on a simplified version of the discretized DC motor system given in [22], which is also used as a benchmark in [9], [15]:

$$A_k = \alpha \begin{bmatrix} -0.0005 & -0.0084 \\ 0.0517 & 0.8069 \end{bmatrix}; \ B_k = \begin{bmatrix} 0.1815 \\ 1.7902 \end{bmatrix}; \ D_k = \begin{bmatrix} 0 \\ 0 \end{bmatrix};$$

$$C_k = \xi \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \ G_k = \gamma \begin{bmatrix} 0.0129 \\ -1.2504 \end{bmatrix}; \ H_k = \eta \begin{bmatrix} 2 \\ 0 \end{bmatrix};$$

$$Q_k = \chi \begin{bmatrix} 0.0036 & 0.0342 \\ 0.0342 & 0.3250 \end{bmatrix}; \quad R_k = \rho \begin{bmatrix} 0.5 & 0 \\ 0 & 0.16 \end{bmatrix}$$

where $\alpha = \{0, 0.5, 1\}$, $\xi = \{0, 0.5, 1.2\}$, $\gamma = \{0, 0.5, 1.2\}$, $\eta = \{0.6, 0.8, 1.2\}$, $\chi = \{0.1, 1.2, 10\}$ and $\rho = \{0.1, 1.2, 10\}$ are system parameters which we vary for studying the effect of parameter changes to estimator responses. The known and unknown inputs used in the following simulations are
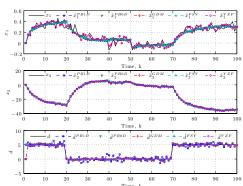

Fig. 1. Actual states x_1 and x_2 and its estimates, as well as unknown input d and its estimates.
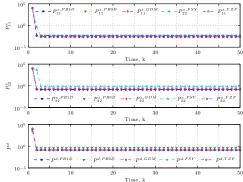

Fig. 2. Variance of estimate errors of states x_1 and x_2, and of unknown input $d - P_{11}^x$, P_{22}^x and P^d of the various state and input estimators for the first 50 time steps.

$$u[k] = \begin{cases} -0.3 & k \geq 50 \\ 0.5 & \text{otherwise} \end{cases}$$

$$d[k] = \begin{cases} 5 & 1 \leq k \leq 19, 70 \leq k \leq 100 \\ 0 & \text{otherwise}. \end{cases}$$

To compare the performance of both input and state estimates, we restrict our attention to (i) Palanthandalam-Madapusi and Bernstein filter with one step delay (PB1D) from Section V-A.2.a, (ii) Palanthandalam-Madapusi and Bernstein filter without one step delay (PB0D) from Section V-A.2.b, (iii) Gillijns and De Moor MVU estimator (GDM) from Section V-A.3, (iv) Fang, Shi and Yi filter (FSY) from Section V-A.4 and (v) the MVU estimator presented in this paper (YZF) from Section IV. The simulations were implemented in MATLAB on a 2.2 GHz Intel Core i7 CPU.

Figure 1 shows a comparison of the input and state estimation of the five MVU estimators, when all the system parameters are set to 1. In this case, all estimators considered were reasonably successful at estimating the states as well as the unknown inputs. On the other hand, Figure 2 shows the variances of their states and input estimates, which we use as a metric of estimator performance. Thus, we see that the GDM and YZF estimators are better state estimators than the PB1D, PB0D and FSY filters. For unknown input estimates, the PB0D estimator is the worst, while the FSY, GDM and YZF estimators are the best.

We tested the dependence of the filter performance on system parameters with 18 simulation experiments with different system parameter values and tabulated the results in Table I. Unless otherwise specified, the default values of the parameters are 1. From Table I, we observe that YZF

TABLE I

STEADY-STATE PERFORMANCE OF PB1D, PB0D, GDM AND YZF (MINIMUM VALUE OF EACH TRIAL IN BOLD).

Parameter		PB1D			PB0D			GDM			FSY			YZF		
		P^x_{11}	P^x_{22}	P^d	P^x_{11}	P^x_{22}	P^d	P^x_{11}	P^x_{22}	P^d	P^x_{11}	P^x_{22}	P^d	P^x_{11}	P^x_{22}	P^d
α	0	0.0044	0.1600	0.3102	0.0044	0.1600	0.1261	**0.0021**	**0.1224**	**0.1255**	0.0036	0.5213	**0.1255**	**0.0021**	**0.1224**	**0.1255**
	0.5	0.0044	0.1600	0.3268	0.0044	0.1600	0.1261	**0.0022**	**0.1237**	**0.1255**	0.0036	0.5453	**0.1255**	**0.0022**	**0.1237**	**0.1255**
	1	0.0044	0.1600	0.3767	0.0044	0.1600	0.1261	**0.0024**	**0.1268**	**0.1256**	0.0036	0.6114	**0.1256**	**0.0024**	**0.1268**	**0.1256**
ξ	0	N/A	N//A	N/A	N/A	N/A	N/A	**0.0037**	**1.4969**	**0.1250**	**0.0037**	**1.4969**	**0.1250**	**0.0037**	**1.4969**	**0.1250**
	0.5	0.0044	0.6400	0.8834	0.0044	0.6400	0.1253	**0.0030**	**0.3462**	**0.1252**	0.0036	0.7558	**0.1252**	**0.0030**	**0.3462**	**0.1252**
	1.2	0.0044	0.1111	0.3251	0.0044	0.1111	0.1266	**0.0022**	**0.0935**	**0.1258**	0.0036	0.5893	**0.1258**	**0.0022**	**0.0935**	**0.1258**
γ	0	N/A	N//A	N/A	N/A	N/A	N/A	**0.0016**	**0.1143**	**0.1254**	0.0036	0.4002	**0.1254**	**0.0016**	**0.1143**	**0.1254**
	0.5	0.0044	0.1600	1.5070	0.0044	0.1600	0.1261	**0.0019**	**0.1185**	**0.1255**	0.0036	0.4563	**0.1255**	**0.0019**	**0.1185**	**0.1255**
	1.2	0.0044	0.1600	0.2616	0.0044	0.1600	0.1261	**0.0026**	**0.1302**	0.1257	0.0036	0.7005	0.1257	**0.0026**	**0.1302**	**0.1256**
η	0.6	0.0044	0.1600	0.3767	0.0044	0.1600	0.3502	**0.0030**	**0.1373**	**0.3493**	0.0037	0.9683	0.3494	**0.0030**	**0.1373**	**0.3493**
	0.8	0.0044	0.1600	0.3767	0.0044	0.1600	0.1970	**0.0026**	**0.1311**	**0.1963**	0.0036	0.7252	0.1964	**0.0026**	**0.1311**	**0.1963**
	1.2	0.0044	0.1600	0.3767	0.0044	0.1600	0.0876	**0.0022**	**0.1239**	**0.0872**	0.0036	0.5489	**0.0872**	**0.0022**	**0.1239**	**0.0872**
χ	0.1	0.0005	0.1600	0.1897	0.0005	0.1600	0.1251	**0.0004**	**0.1039**	**0.1251**	**0.0004**	0.2998	**0.1251**	**0.0004**	**0.1039**	**0.1251**
	1.2	0.0052	0.1600	0.4183	0.0052	0.1600	0.1263	**0.0026**	**0.1295**	**0.1257**	0.0044	0.6790	**0.1257**	**0.0026**	**0.1295**	**0.1257**
	10	0.0434	0.1600	2.2476	0.0434	0.1600	0.1359	**0.0052**	**0.1531**	**0.1263**	0.0360	3.5634	**0.1263**	**0.0052**	**0.1531**	**0.1263**
ρ	0.1	0.0043	0.0160	0.2248	0.0043	0.0160	0.0136	**0.0005**	**0.0153**	**0.0126**	0.0036	0.3563	**0.0126**	**0.0005**	**0.0153**	**0.0126**
	1.2	0.0044	0.1920	0.4105	0.0044	0.1920	0.1511	**0.0025**	**0.1490**	**0.1506**	0.0036	0.6656	**0.1506**	**0.0025**	**0.1490**	**0.1506**
	10	0.0045	1.6000	1.8966	0.0045	1.6000	1.2511	**0.0039**	**1.0386**	**1.2510**	**0.0039**	2.9983	**1.2510**	**0.0039**	**1.0386**	**1.2510**

and GDM have the least variance in all categories, whereas FSY is almost equally good in unknown input estimation.

On the other hand, the PB1D and PB0D filters fail when $\xi = 0$ or $\gamma = 0$. In these cases, $\mathrm{rank}(C_k G_{k-1}) < p$ which violates the assumptions of the filters (see Remark 3). Therefore, we conclude that the GDM, FSY and YZF estimators are better estimators than the PB1D and PB0D filters. It also appears that the performance of both the YZF and GDM estimators are comparable. This is consistent with the findings in the literature, since the GDM filter is shown to be globally optimal in [23], while the state update law of YZF filter proposed in this paper can be shown to be a special case of the state estimator in [10] which is also proven to be globally optimal.

VII. CONCLUSION

This paper presented a variation of an optimal filter that simultaneously estimates the states and unknown inputs in an unbiased minimum-variance sense for linear discrete-time stochastic systems with direct feedthrough. We argued that the information about the unknown input can be obtained from the current time step as well as the previous one, making it possible to estimate the unknown in different ways. In contrast to previous filters which predicted the unknown inputs based on the information in the previous step and a hybrid of the current and previous step, the new filter we proposed utilizes the most current updated state estimate to compute the best linear unbiased estimate of the input. Simulation results have shown that the new filter performs just as well as the previous ones, if not better in all test trials.

REFERENCES

[1] R. Verma and D. Del Vecchio, "Semiautonomous multivehicle safety," *IEEE Robotics Automation Magazine*, vol. 18, no. 3, pp. 44–54, 2011.

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

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

[4] G. D. Nicolao, G. Sparacino, and C. Cobelli, "Nonparametric input estimation in physiological systems: Problems, methods, and case studies," *Automatica*, vol. 33, no. 5, pp. 851 – 870, 1997.

[5] B. Friedland, "Treatment of bias in recursive filtering," *IEEE Transactions on Automatic Control*, vol. 14, no. 4, pp. 359–367, Aug. 1969.

[6] E. Davison, "The output control of linear time-invariant multivariable systems with unmeasurable arbitrary disturbances," *IEEE Transactions on Automatic Control*, vol. 17, no. 5, pp. 621–630, 1972.

[7] M. Darouach and M. Zasadzinski, "Unbiased minimum variance estimation for systems with unknown exogenous inputs," *Automatica*, vol. 33, no. 4, pp. 717 – 719, 1997.

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

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

[10] Y. Cheng, H. Ye, Y. Wang, and D. Zhou, "Unbiased minimum-variance state estimation for linear systems with unknown input," *Automatica*, vol. 45, no. 2, pp. 485–491, 2009.

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

[12] C.-S. Hsieh, "Robust two-stage Kalman filters for systems with unknown inputs," *IEEE Transactions on Automatic Control*, vol. 45, no. 12, pp. 2374–2378, Dec. 2000.

[13] S. Gillijns and B. De Moor, "Unbiased minimum-variance input and state estimation for linear discrete-time systems," *Automatica*, vol. 43, no. 1, pp. 111–116, Jan. 2007.

[14] S. Gillijns and B. D. Moor, "Unbiased minimum-variance input and state estimation for linear discrete-time systems with direct feedthrough," *Automatica*, vol. 43, no. 5, pp. 934 – 937, 2007.

[15] C.-S. Hsieh, "Extension of unbiased minimum-variance input and state estimation for systems with unknown inputs," *Automatica*, vol. 45, no. 9, pp. 2149 – 2153, 2009.

[16] 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*, vol. 25, no. 8, pp. 671–686, 2011.

[17] S. Li. Estimation theory. [Online]. Available: http://www.mcn.ece.ufl.edu/public/ShijieLi/PDF/estimation.pdf

[18] H. Payne and L. Silverman, "On the discrete time algebraic riccati equation," *IEEE Transactions on Automatic Control*, vol. 18, no. 3, pp. 226–234, 1973.

[19] L. Silverman, "Discrete Riccati equations: Alternative algorithms, asymptotic properties, and system theory interpretations," *Control and Dynamic Systems*, vol. 12, pp. 313–386, 1976.

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

[21] T. Kailath, A. Sayed, and B. Hassibi, *Linear estimation*, ser. Prentice-Hall information and system sciences series. Prentice Hall, 2000.

[22] S. H. Park, P. S. Kim, O.-K. Kwon, and W. H. Kwon, "Estimation and detection of unknown inputs using optimal FIR filter," *Automatica*, vol. 36, no. 10, pp. 1481 – 1488, 2000.

[23] C. Hsieh, "On the global optimality of unbiased minimum-variance state estimation for systems with unknown inputs," *Automatica*, vol. 46, no. 4, pp. 708 – 715, 2010.