Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Moving horizon least-squares input estimation for linear discrete-time stochastic systems ⋆

Yiming Wan * **Tamas Keviczky** * **Michel Verhaegen** *

* *Delft University of Technology, 2628CD, Delft, The Netherlands (e-mail: y.wan@tudelft.nl; t.keviczky@tudelft.nl; m.verhaegen@tudelft.nl).*

Abstract: This paper presents a novel moving horizon least-squares input estimation method for linear discrete-time stochastic systems. For systems with completely unknown initial state and no unstable zeros, some existing work showed that asymptotic input reconstruction is possible in the absence of noises. However, under the same condition but with stochastic noises, most existing input estimators, which are designed to optimally deal with noises, fail to ensure asymptotic unbiasedness. In order to address this limitation for linear discrete-time stochastic systems, we characterize necessary and sufficient conditions for input observability and detectability, and propose a moving horizon least-squares input estimator. Based on the conditions for input observability and detectability, it is proved that our proposed input estimator gives an asymptotically unbiased estimate and has minimal estimation error variance over all linear asymptotically unbiased input estimators. Its effectiveness is illustrated by simulation examples involving aircraft sensor and actuator faults.

Keywords: unknown input estimation, input observability/detectability, moving horizon estimation, minimum variance unbiased estimation

1. INTRODUCTION

Due to its applications in fault diagnosis and fault-tolerant control, the problem of determining the unknown inputs of a dynamic system from available input and output (I/O) data, known as unknown input estimation (UIE), has received considerable attention during the last decades.

In the context of deterministic systems, the UIE problem is often referred to as input reconstruction. One main difficulty is due to the presence of unknown initial state. For continuous-time linear systems, Basile and Marro (1992) defined the notion of unknown-state, unknown-input completely reconstructable system, which was characterized by necessary and sufficient geometric conditions. A similar notion, i.e., l-delay input and initial-state observability, was discussed in Kirtikar et al. (2011) for discrete-time systems. These results reveal the role of invariant zeros on input reconstruction: (a) if the system has invariant zeros, then there exists an initial state such that, for some nonzero input sequences, the output signal is identically zero, which makes exact input reconstruction impossible; (b) if the invariant zeros are all stable, the unobservable input decays, thus asymptotic input reconstruction is still possible. These properties of input reconstruction can be achieved by a joint state-input reconstructor proposed in Gillijns (2007). All the poles of the proposed reconstructor can be arbitrarily assigned except those which are equal to the invariant zeros of the original system. Therefore, when the original system has no unstable zeros, the proposed reconstructor in Gillijns (2007) is guaranteed to be stable through pole placement, which ensures asymptotic input reconstruction. Instead of the joint consideration of state and input reconstruction as in the above literature, Hou and Patton (1998) proposed the notion of input observability and detectability for continuous-time systems, and showed that state observability and controllability were not necessary for input reconstruction.

The UIE problem has also been investigated for linear discrete-time stochastic systems. In this context, the input estimators are designed to optimize certain criteria dealing with stochastic noises. One class of methods is the joint state-input estimation approach with assumed a priori statistics of the initial state, e.g., Gillijns and Moor (2007a,b); Gillijns (2007). Gillijns and Moor (2007a) proved that its proposed input estimator achieved minimum variance over all linear unbiased input estimators. Another class of methods to UIE estimates the unknown input without estimating the state. By replacing the unknown initial state with the past I/O data, a constrained least-squares (LS) approach was proposed in Dong and Verhaegen (2012) which ensures asymptotically unbiased estimation for systems with stable inversion.

Basile and Marro (1992) and Gillijns (2007) showed the possibility of asymptotic input reconstruction when the considered system has neither unstable zeros nor stochastic noises. Under the same condition but with stochastic noises, it is intuitive and reasonable to infer that asymptotically unbiased input estimation should be possible.

⋆ The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7-RECONFIGURE/2007–2013) under grant agreement No. 314544.

Unfortunately, the optimal input estimators developed in the above references fail to meet this expectation for stochastic systems. Actually, the poles of these optimal input estimators have no direct relationship to invariant zeros or other properties of the original systems, and might even be unstable. As a result, they may not yield an asymptotically unbiased estimate, even for systems with no unstable zeros.

In order to address the above limitation of the existing UIE methods for linear discrete-time stochastic systems, this paper proposes a novel moving horizon LS (MHLS) input estimator. First, necessary and sufficient conditions for input observability and detectability are developed, which can be regarded as an extension of the results in Kirtikar et al. (2011). Then, the MHLS input estimator is proposed. Its estimation performance is analyzed based on the conditions for input observability and detectability. For systems with no transmission zeros, the proposed estimator gives an unbiased input estimate if the horizon length is no less than the observability index of the considered system. For systems with only stable zeros, the proposed estimator achieves asymptotic unbiasedness. In both of the above cases, it is shown that the MHLS estimator has minimal variance over the class of all linear asymptotically unbiased estimators. Finally, the effectiveness of the proposed MHLS estimator is illustrated by simulation examples involving aircraft sensor and actuator faults.

Notations: For a matrix X, its range and null space is denoted by $\mathcal{R}(X)$ and $\mathcal{N}(X)$, respectively. X^{-} represents the left inverse satisfying $X^{-}X = I$, while $X^{(1)}$ represents the generalized inverse satisfying $XX^{(1)}X = X$. The Moore-Penrose inverse is denoted by $X^{\dagger}$.

2. PRELIMINARIES AND PROBLEM FORMULATION

We consider linear discrete-time systems governed by the following state space model:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + Ef(k) + Fw(k) \\ y(k) &= Cx(k) + Du(k) + Gf(k) + v(k). \end{aligned} \tag{1}$$

Here $x(k) \in \mathbb{R}^{n_x}$, $y(k) \in \mathbb{R}^{n_y}$, and $u(k) \in \mathbb{R}^{n_u}$ represent the state, the measurement, and the known control input at time instant k, respectively. The stochastic disturbances are represented by the process noise $w(k) \in \mathbb{R}^{n_w}$ and the measurement noise $v(k) \in \mathbb{R}^{n_v}$, both of which are considered to be white zero-mean Gaussian. A, B, C, D are known real matrices, with bounded norms and appropriate dimensions. $f(k) \in \mathbb{R}^{n_f}$ is the unknown input to be estimated, and we have no a priori knowledge about how it varies with time.

The following assumptions are made in this paper:

Assumption 1. $\operatorname{rank}(G) = n_f$.

Assumption 2. The initial state $x(0)$ is completely unknown.

For the sake of brevity, Assumption 1 is adopted, which is common for sensor faults. However, our proposed methods can be extended to actuator faults as well in a straightforward manner, as explained latter in Remark 3.

Consider the output equation in a sliding window with a length of L sampling instants. Define data vectors in this window as $\mathbf{u}_{k,L}$, $\mathbf{y}_{k,L}$, $\mathbf{f}_{k,L}$, $\mathbf{w}_{k,L}$ and $\mathbf{v}_{k,L}$, respectively for the signals u, y, f, w and v; e.g.,

$$\mathbf{u}_{k,L} = \left[\, u^{\mathrm{T}}(k_0) \, \cdots \, u^{\mathrm{T}}(k) \,\right]^{\mathrm{T}},$$

where $k_0 = k - L + 1$. For the system model (1), let $\mathcal{O}_L$ denote its extended observability matrix with L block elements, and $\mathbf{T}_L^{\star}$ be the lower triangular Toeplitz matrix with L block columns and rows, i.e.,

$$\mathcal{O}_L = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{L-1} \end{bmatrix}, \; \mathbf{T}_L^{u} = \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ CA^{L-2}B & CA^{L-3}B & \cdots & D \end{bmatrix},$$

and $\mathbf{T}_L^{f}$ and $\mathbf{T}_L^{w}$ are defined similarly to $\mathbf{T}_L^{u}$.

With the above definitions, the extended output equation of the system (1) can be written as

$$\mathbf{y}_{k,L} = \mathcal{O}_L x(k_0) + \mathbf{T}_L^{u} \cdot \mathbf{u}_{k,L} + \mathbf{T}_L^{f} \cdot \mathbf{f}_{k,L} + \mathbf{T}_L^{w} \cdot \mathbf{w}_{k,L} + \mathbf{v}_{k,L}. \tag{2}$$

Then a residual vector can be generated as

$$\begin{aligned} \mathbf{r}_{k,L} &= \mathbf{y}_{k,L} - \mathbf{T}_L^{u} \cdot \mathbf{u}_{k,L} \\ &= \mathcal{O}_L x(k_0) + \mathbf{T}_L^{f} \cdot \mathbf{f}_{k,L} + \mathbf{n}_{k,L} \\ &= \mathcal{O}_L x(k_0) + \mathbf{T}_{L,1}^{f} \cdot \mathbf{f}_{k-1,L-1} + \mathbf{T}_{L,2}^{f} \cdot f(k) + \mathbf{n}_{k,L} \\ &= \underbrace{\left[\, \mathcal{O}_L \;\; \mathbf{T}_L^{f} \,\right]}_{\Psi_L} \underbrace{\begin{bmatrix} x(k_0) \\ \mathbf{f}_{k,L} \end{bmatrix}}_{\bar{\mathbf{f}}_{k,L}} + \mathbf{n}_{k,L} \\ &= \underbrace{\left[\, \mathcal{O}_L \;\; \mathbf{T}_{L,1}^{f} \,\right]}_{\Psi_{L,1}} \begin{bmatrix} x(k_0) \\ \mathbf{f}_{k-1,L-1} \end{bmatrix} + \mathbf{T}_{L,2}^{f} \cdot f(k) + \mathbf{n}_{k,L}, \end{aligned} \tag{3}$$

where $\mathbf{n}_{k,L} = \mathbf{T}_L^{w} \cdot \mathbf{w}_{k,L} + \mathbf{v}_{k,L}$, $\mathbf{T}_{L,1}^{f}$ and $\mathbf{T}_{L,2}^{f}$ correspond to the first $(L-1)\, n_f$ columns and the last n_f columns of $\mathbf{T}_L^{f}$, respectively.

The objective is to use the generated residual signals $\mathbf{r}_{k,L}$ to estimate the unknown input $f(k)$, and to analyze the estimation performance.

3. INPUT OBSERVABILITY AND DETECTABILITY

As the basis of MHLS input estimator, this section will analyze some properties of invariant zeros. First, two lemmas about invariant zeros are given.

Lemma 1. With Assumption 1 and Ψ_L, $\bar{\mathbf{f}}_{k,L}$ defined in (3), there exists a nonnegative integer l such that for all $L \geq l$,

$$\mathbf{r}_{k,L}^{d} = \Psi_L \cdot \bar{\mathbf{f}}_{k,L} = 0 \tag{4}$$

if and only if

$$f(k_0 + i) = -G^{-}CK_d^{i}x(k_0), \; i = 0, 1, \cdots, L-1, \tag{5}$$

with

$$K_d \triangleq A - EG^{-}C, \tag{6}$$

and $0 \neq x(k_0) \in \mathcal{S}_1 \cup \mathcal{S}_2$, where

$$K_d\mathcal{S}_1 \subseteq \mathcal{S}_1, \; \mathcal{S}_1 \subseteq \mathcal{N}(C), \tag{7}$$

$$K_d\mathcal{S}_2 \subseteq \mathcal{S}_2, \; \{0\} \neq C\mathcal{S}_2 \subseteq \mathcal{R}(G). \tag{8}$$

The proof of Lemma 1 follows that of Lemma A.1 in Kirtikar et al. (2011) with slight modifications, thus it is omitted for brevity. Lemma 1 shows that the null space

$\mathcal{N}(\Phi_L)$ corresponds to two subspaces $\mathcal{S}_1$ and $\mathcal{S}_2$. The following lemma further reveals that these two subspaces correspond to two types of invariant zeros.

Lemma 2. Let λ_0 and x_0 denote an eigenvalue and the related eigenvector of K_d defined in (6). If $x_0 \in \mathcal{S}_1 \cup \mathcal{S}_2$, λ_0 is an invariant zero of (A, E, C, G). Furthermore, with Assumption 1, when $x_0 \in \mathcal{S}_1$, λ_0 is an output decoupling zero which implies $x(k_0) \neq 0$ and $\mathbf{f}_{k,L} = 0$ in (4); when $x_0 \in \mathcal{S}_2$, λ_0 is a transmission zero which implies $x(k_0) \neq 0$ and $\mathbf{f}_{k,L} \neq 0$ in (4).

Proof. By defining

$$f_0 = -G^{-}Cx_0, \tag{9}$$

we can obtain $\lambda_0 x_0 = K_d x_0 = Ax_0 + Ef_0$ from (6). Since $x_0 \in \mathcal{S}_1 \cup \mathcal{S}_2$ imply $(I - GG^{-})Cx_0 = 0$ according to (7)-(8), it follows from (9) that $Cx_0 + Gf_0 = GG^{-}Cx_0 + Gf_0 = 0$. Therefore we have

$$\begin{bmatrix} A - \lambda_0 I_{n_x} & E \\ C & G \end{bmatrix} \begin{bmatrix} x_0 \\ f_0 \end{bmatrix} = 0$$

which shows that λ_0 is an invariant zero of (A, E, C, G).

Due to Assumption 1, when x_0 belongs to $\mathcal{S}_1$ defined in (7), there is $f_0 = 0$, and λ_0 is an unobservable mode of (A, C), i.e. an output decoupling zero; when x_0 belongs to $\mathcal{S}_2$ defined in (8), there is $f_0 \neq 0$, and λ_0 is a transmission zero (Zhou et al., 1996). □

Similarly to the concept of input observability and detectability for continuous-time systems in Hou and Patton (1998), we proceed to define input observability and detectability for the discrete-time system (1).

Definition 1. Let l be a nonnegative integer. The input $f(k)$ is observable if there exists $L \geq l$ such that (4) implies $f(k) = 0$, where $f(k)$ corresponds to the last n_f rows in $\mathbf{f}_{k,L}$.

Definition 2. The input $f(k)$ is detectable if (4) implies $f(k) \to 0$ for $L \to \infty$.

According to Lemma 1 and 2, some intuitive explanations about the above two definitions are given as below. By substituting $x(k_0) = x_0$ and (9) into (5), we have

$$f(k_0 + i) = \lambda_0^i f_0, \ i = 0, 1, \cdots, L-1.$$

This implies that there exist a nonzero initial state $x(k_0) = x_0$ and input sequences $\{\lambda_0^i f_0\}_{i=0,1,2,\ldots}$ such that the deterministic residual signal $r_{k,L}^d$ defined in (4) is identically zero. If $x_0 \in \mathcal{S}_1$, the input sequence $\{\lambda_0^i f_0\}_{i=0,1,2,\ldots}$ is identically zero, thus $f(k)$ is observable. If $x_0 \in \mathcal{S}_2$ and the transmission zero λ_0 lies inside the unit disk, it follows that $\lambda_0^k f_0 \to 0$ for $k \to \infty$, which means that $f(k)$ is detectable.

The following results provide necessary and sufficient conditions for input observability and detectability, which will serve as the foundation of the latter sections.

Theorem 1. With Assumption 1, the following statements are equivalent:

(i) The input $f(k)$ of the system (1) is observable.
(ii) $\mathcal{S}_2$ defined in (8) is empty, or equivalently, (A, E, C, G) has no transmission zeros.
(iii) $\mathcal{R}(\Psi_{L,1}) \bigcap \mathcal{R}\left(\mathbf{T}_{L,2}^f\right) = \{0\}$ holds for all $L \geq \nu$, where ν is the observability index of (A, C).

Proof. Due to limited space, here we only prove that (ii) implies (iii). With $L = \nu$, it follows from (5)-(6) that (4) implies for $i = 0, 1, \cdots, \nu - 1$,

$$\begin{aligned} r^d(k_0 + i) &= Cx(k_0 + i) + Gf(k_0 + i) \\ &= \left(I_{n_y} - GG^{-}\right) CK_d^i x(k_0) = 0. \end{aligned} \tag{10}$$

Since any CK_d^j $(j \geq \nu)$ can be rewritten as $CK_d^j = \sum_{i=0}^{\nu-1} \alpha_i CK_d^i$ according to the definition of observability index, we have

$$\left(I_{n_y} - GG^{-}\right) CK_d^j x(k_0) = 0, \text{ for } j \geq \nu. \tag{11}$$

Therefore the satisfaction of (4) for $L = \nu$ can lead to the satisfaction of (4) for all $L \geq \nu$. According to Lemma 1, this implies that $x(k_0)$ belongs to $\mathcal{S}_1$ defined in (7) when $\mathcal{S}_2$ is empty as stated in (ii). From Lemma 2, (4) and $x(k_0) \in \mathcal{S}_1$ indicates that $\mathbf{f}_{k,L} = 0$ for all $L \geq \nu$, which proves (iii). □

Corollary 1. The smallest l which ensures

$$\mathcal{R}(\Psi_{L,1}) \bigcap \mathcal{R}\left(\mathbf{T}_{L,2}^f\right) = \{0\}$$

for all $L \geq l$ is ν.

Theorem 2. With Assumptions 1, the following statements are equivalent:

(i) The input $f(k)$ of the system (1) is detectable.
(ii) All transmission zeros of (A, E, C, G) are stable.
(iii) $\mathcal{R}(\Psi_{L,1}) \bigcap \mathcal{R}\left(\mathbf{T}_{L,2}^f\right) \to \{0\}$ for $L \to \infty$.

The proof of Theorem 2 follows that of Theorem 1, thus it is omitted.

Remark 1. The statement (ii) of both Theorem 1 and 2 shows that the output decoupling zeros, or equivalently, unobservable states, do not affect input observability and detectability. In contrast, the joint state-input estimation approach in Gillijns and Moor (2007a,b); Gillijns (2007); Kirtikar et al. (2011) explicitly require state observability or detectability.

4. MOVING HORIZON INPUT ESTIMATION

Let $\Sigma_{n,L}$ denote the covariance matrix of $\mathbf{n}_{k,L}$, and define

$$\begin{aligned} &\bar{\mathbf{r}}_{k,L} = \Sigma_{n,L}^{-\frac{1}{2}} \mathbf{r}_{k,L}, \ \bar{\Psi}_L = \Sigma_{n,L}^{-\frac{1}{2}} \Psi_L, \ \bar{\Psi}_{L,1} = \Sigma_{n,L}^{-\frac{1}{2}} \Psi_{L,1}, \\ &\bar{\mathbf{T}}_{L,2}^f = \Sigma_{n,L}^{-\frac{1}{2}} \mathbf{T}_{L,2}^f, \ \bar{\mathbf{n}}_{k,L} = \Sigma_{n,L}^{-\frac{1}{2}} \mathbf{n}_{k,L}. \end{aligned} \tag{12}$$

Then by multiplying both sides of (3) with $\Sigma_{n,L}^{-\frac{1}{2}}$, (3) can be rewritten as

$$\bar{\mathbf{r}}_{k,L} = \underbrace{\left[\bar{\Psi}_{L,1} \ \bar{\mathbf{T}}_{L,2}^f\right]}_{\bar{\Psi}_L} \bar{\mathbf{f}}_{k,L} + \bar{\mathbf{n}}_{k,L}. \tag{13}$$

The following LS problem is formulated to estimate $f(k)$ based on (13):

$$\min_{\hat{\bar{\mathbf{f}}}_{k,L}} \left\| \bar{\mathbf{r}}_{k,L} - \bar{\Psi}_L \hat{\bar{\mathbf{f}}}_{k,L} \right\|_2^2. \tag{14}$$

The solution to the optimization problem (14) may be non-unique, because according to Lemma 1 $\mathcal{N}\left(\bar{\Psi}_L\right) \neq \{0\}$ when (A, E, C, G) has output decoupling zeros or transmission zeros. One solution to (14) is

$$\hat{\bar{\mathbf{f}}}_{k,L} = \left(\bar{\Psi}_L^{\mathrm{T}} \bar{\Psi}_L\right)^{(1)} \bar{\Psi}_L^{\mathrm{T}} \bar{\mathbf{r}}_{k,L}, \tag{15}$$

where $\left(\bar{\Psi}_L^{\mathrm{T}}\bar{\Psi}_L\right)^{(1)}$ is the generalized inverse satisfying

$$\bar{\Psi}_L^{\mathrm{T}}\bar{\Psi}_L\left(\bar{\Psi}_L^{\mathrm{T}}\bar{\Psi}_L\right)^{(1)}\bar{\Psi}_L^{\mathrm{T}}\bar{\Psi}_L=\bar{\Psi}_L^{\mathrm{T}}\bar{\Psi}_L. \quad (16)$$

Since only the estimate $\hat{f}(k)$, i.e., the last n_f entries of $\hat{\bar{\mathbf{f}}}_{k,L}$, is of interest at each time instant, we would like to extract $\hat{f}(k)$ from the solution (15), and analyze its estimation performance in the following subsections.

4.1 Unbiasedness

Given I/O data $\mathbf{u}_{k,L}$ and $\mathbf{y}_{k,L}$, the most general linear input estimator for $f(k)$ can be written as

$$\breve{\mathbf{f}}_{k,L}=\mathbf{M}_k\mathbf{y}_{k,L}+\mathbf{H}_k\mathbf{u}_{k,L}=\mathbf{W}_k\Sigma_{n,L}^{-\frac{1}{2}}\mathbf{y}_{k,L}+\mathbf{H}_k\mathbf{u}_{k,L} \quad (17)$$

with $\mathbf{M}_k\Sigma_{n,L}^{\frac{1}{2}}=\mathbf{W}_k$. A necessary and sufficient condition for (17) to be an unbiased (asymptotically unbiased) estimator for observable (detectable) input $f(k)$, is given in the following theorem.

Theorem 3. The estimator (17) gives an unbiased estimate of observable input $f(k)$ with $L\geq\nu$, or an asymptotically unbiased estimate of detectable input $f(k)$ with $L\to\infty$, if and only if both of the following statements hold:

(i) $\mathbf{W}_k\Sigma_{n,L}^{-\frac{1}{2}}\mathbf{T}_L^u+\mathbf{H}_k=0$;
(ii) $\bar{\Psi}_L\mathbf{W}_k\bar{\Psi}_L=\bar{\Psi}_L$.

Proof. Sufficiency: It follows from (2), (13) and the statement (i) that (17) can be rewritten as

$$\breve{\mathbf{f}}_{k,L}=\mathbf{M}_k\mathbf{r}_{k,L}=\mathbf{W}_k\bar{\mathbf{r}}_{k,L}. \quad (18)$$

By multiplying both sides of (18) with $\bar{\Psi}_L$ and then taking mathematical expectation, we have

$$\bar{\Psi}_L\mathrm{E}\left(\breve{\mathbf{f}}_{k,L}\right)=\bar{\Psi}_L\mathbf{W}_k\bar{\Psi}_L\bar{\mathbf{f}}_{k,L}. \quad (19)$$

With the statement (ii), the above equation can be reformulated as

$$\bar{\Psi}_L\left(\mathrm{E}\left(\breve{\mathbf{f}}_{k,L}\right)-\bar{\mathbf{f}}_{k,L}\right)=0. \quad (20)$$

According to Theorem 1 and 2, it can be seen from (20) that (i) $\mathrm{E}\left(\hat{f}(k)\right)=f(k)$ with $L\geq\nu$ when the input $f(k)$ is observable, and (ii) $\mathrm{E}\left(\hat{f}(k)\right)\to f(k)$ with $L\to\infty$ when the input $f(k)$ is detectable.

Necessity: By substituting (2) into (17) we have

$$\mathrm{E}\left(\breve{\mathbf{f}}_{k,L}\right)=\mathbf{W}_k\bar{\Psi}_L\bar{\mathbf{f}}_{k,L}+\left(\mathbf{W}_k\Sigma_{n,L}^{-\frac{1}{2}}\mathbf{T}_L^u+\mathbf{H}_k\right)\mathbf{u}_{k,L}. \quad (21)$$

This above equation implies that the statement (i) must hold in order to yield (asymptotically) unbiased estimation. Without loss of generality, let

$$\mathbf{W}_k\bar{\Psi}_L=\left(\bar{\Psi}_L^{\mathrm{T}}\bar{\Psi}_L\right)^{(1)}\bar{\Psi}_L^{\mathrm{T}}\bar{\Psi}_L+S_k, \quad (22)$$

then (21) can be rewritten as

$$\bar{\Psi}_L\left(\mathrm{E}\left(\breve{\mathbf{f}}_{k,L}\right)-\bar{\mathbf{f}}_{k,L}\right)=\bar{\Psi}_LS_k\bar{\mathbf{f}}_{k,L} \quad (23)$$

by multiplying both sides of (21) with $\bar{\Psi}_L$ and utilizing the fact $\bar{\Psi}_L\left(\bar{\Psi}_L^{\mathrm{T}}\bar{\Psi}_L\right)^{(1)}\bar{\Psi}_L^{\mathrm{T}}\bar{\Psi}_L=\bar{\Psi}_L$. Since (asymptotic) unbiasedness implies (20) for arbitrary $\bar{\mathbf{f}}_{k,L}$, it can be concluded from (23) that $\bar{\Psi}_LS_k=0$. Then, multiplying both sides of (22) with $\bar{\Psi}_L$ leads to the statement (i). □

Corollary 2. From Theorem 3, it can be concluded that (15) also gives an unbiased estimate of observable input $f(k)$ with $L\geq\nu$, or an asymptotically unbiased estimate of detectable input $f(k)$ with $L\to\infty$.

4.2 Computation

Based on (15), the following theorem gives the explicit computation form of $\hat{f}(k)$ without calculating other entries of $\hat{\bar{\mathbf{f}}}_{k,L}$, which of course yields an (asymptotically) unbiased estimate as (15) does.

Theorem 4. The estimate $\hat{f}(k)$ in (15) can be computed as

$$\hat{f}(k)=\left(\left(\bar{\mathbf{T}}_{L,2}^f\right)^{\mathrm{T}}\mathbf{P}_{\bar{\Psi},1}^{\perp}\bar{\mathbf{T}}_{L,2}^f\right)^{(1)}\left(\bar{\mathbf{T}}_{L,2}^f\right)^{\mathrm{T}}\mathbf{P}_{\bar{\Psi},1}^{\perp}\bar{\mathbf{r}}_{k,L} \quad (24)$$

with

$$\mathbf{P}_{\bar{\Psi},1}^{\perp}=I-\bar{\Psi}_{L,1}\left(\bar{\Psi}_{L,1}^{\mathrm{T}}\bar{\Psi}_{L,1}\right)^{(1)}\bar{\Psi}_{L,1}^{\mathrm{T}}. \quad (25)$$

Proof. Define $\Pi_{11}=\bar{\Psi}_{L,1}^{\mathrm{T}}\bar{\Psi}_{L,1}$, $\Pi_{12}=\bar{\Psi}_{L,1}^{\mathrm{T}}\bar{\mathbf{T}}_{L,2}^f$ and $\Pi_{22}=\left(\bar{\mathbf{T}}_{L,2}^f\right)^{\mathrm{T}}\bar{\mathbf{T}}_{L,2}^f$. According to (13), Schur complements and inverses of block matrices (Kailath et al., 2000), it is straightforward to verify that

$$\left(\bar{\Psi}_L^{\mathrm{T}}\bar{\Psi}_L\right)^{(1)}=\begin{bmatrix}I & -\Pi_{11}^{(1)}\Pi_{12}\\ 0 & I\end{bmatrix}\begin{bmatrix}\Pi_{11}^{(1)} & 0\\ 0 & \triangle_{11}^{(1)}\end{bmatrix}\begin{bmatrix}I & 0\\ -\Pi_{12}^{\mathrm{T}}\Pi_{11}^{(1)} & I\end{bmatrix} \quad (26)$$

$$\triangle_{11}=\Pi_{22}-\Pi_{12}^{\mathrm{T}}\Pi_{11}^{(1)}\Pi_{12} \quad (27)$$

satisfies (16). From (25) and (27) it follows that

$$\triangle_{11}=\left(\bar{\mathbf{T}}_{L,2}^f\right)^{\mathrm{T}}\mathbf{P}_{\bar{\Psi},1}^{\perp}\bar{\mathbf{T}}_{L,2}^f. \quad (28)$$

Then we obtain (24) by substituting (25) and (26)-(28) into (15). □

As defined in (25), $\mathbf{P}_{\bar{\Psi},1}^{\perp}$ is actually the orthogonal projector onto the orthogonal complementary subspace of $\mathcal{R}\left(\bar{\Psi}_{L,1}\right)$. Then there is $\left(\bar{\mathbf{T}}_{L,2}^f\right)^{\mathrm{T}}\mathbf{P}_{\bar{\Psi},1}^{\perp}\bar{\mathbf{T}}_{L,2}^f>0$ if $\mathcal{R}\left(\Psi_{L,1}\right)\bigcap\mathcal{R}\left(\bar{\mathbf{T}}_{L,2}^f\right)=\{0\}$. Based on this fact, the generalized inverse in the estimator (24) should be replaced by the conventional inverse (which is unique) when (A,E,C,G) has no transmission zeros according to Theorem 1.

For (A,E,C,G) that has only stable transmission zeros, the generalized inverse in the estimator (24) is non-unique. But since the estimator (24) is extracted from the unbiased estimate (15), the unbiasedness of the estimator (24) is not affected by the non-unique generalized inverse in (24).

4.3 Optimality

According to (13) and (25), the estimation error covariance matrix of the estimator (24) is

$$\mathrm{cov}\left(\hat{f}(k)-f(k)\right)=\triangle_{11}^{(1)}\triangle_{11}\triangle_{11}^{(1)}=\triangle_{11}^{\dagger} \quad (29)$$

with $\triangle_{11}$ given in (28). Note that the second equation in (29) shows that the estimation error covariance matrix is unique although the generalized inverse $\triangle_{11}^{(1)}$ is non-unique.

Theorem 5. The (asymptotically) unbiased estimator (24) has minimal variance over the class of all linear (asymptotically) unbiased estimators characterized in Theorem 3.

Proof. According to Theorem 3, the class of all linear (asymptotically) unbiased estimators for $f(k)$ can be expressed as

$$\breve{f}(k) = \mathbf{W}_{2,k}\bar{\mathbf{r}}_{k,L}, \tag{30}$$

where $\mathbf{W}_{2,k}$ corresponds to the last n_f rows of $\mathbf{W}_k$ in (18). Then it can be seen that Theorem 5 is proved if we can show that

$$\begin{aligned}&\mathrm{cov}\left(\breve{f}(k) - f(k)\right) - \mathrm{cov}\left(\hat{f}(k) - f(k)\right)\\ =&\mathbf{W}_{2,k}\mathbf{W}_{2,k}^{\mathrm{T}} - \triangle_{11}^{(1)}\triangle_{11}\triangle_{11}^{(1)}\end{aligned} \tag{31}$$

is positive semidefinite.

With the statement (ii) in Theorem 3, we have

$$\mathbf{P}_{\Psi,1}^{\perp}\bar{\Psi}_L\mathbf{W}_k\bar{\Psi}_L = \mathbf{P}_{\Psi,1}^{\perp}\bar{\Psi}_L$$

which can be further reduced to

$$\triangle_{11}\ \mathbf{W}_{2,k}\bar{\Psi}_{L,1} = 0, \tag{32}$$

$$\triangle_{11}\ \mathbf{W}_{2,k}\bar{\mathbf{T}}_{L,2}^{f} = \triangle_{11}\ . \tag{33}$$

through matrix manipulations. Note that $\bar{\Psi}_L$, $\mathbf{P}_{\Psi,1}^{\perp}$ and $\triangle_{11}$ are defined in (13), (25) and (27), respectively. From (32) it follows that

$$\triangle_{11}\ \mathbf{W}_{2,k}\mathbf{P}_{\Psi,1}^{\perp} = \triangle_{11}\ \mathbf{W}_{2,k}, \tag{34}$$

then we have

$$\mathbf{W}_{2,k} = \triangle_{11}^{(1)}\triangle_{11}\ \mathbf{W}_{2,k}\mathbf{P}_{\Psi,1}^{\perp}. \tag{35}$$

Again based on (33) and (34), there is

$$\triangle_{11}\ \mathbf{W}_{2,k}\mathbf{P}_{\Psi,1}^{\perp}\bar{\mathbf{T}}_{L,2}^{f} = \triangle_{11}\ \mathbf{W}_{2,k}\bar{\mathbf{T}}_{L,2}^{f} = \triangle_{11}\ . \tag{36}$$

By substituting (35) and (36) into (31), we have

$$\begin{aligned}&\mathrm{cov}\left(\breve{f}(k) - f(k)\right) - \mathrm{cov}\left(\hat{f}(k) - f(k)\right)\\ =&\ \triangle_{11}^{(1)}\triangle_{11}\ \mathbf{W}_{2,k}\mathbf{P}_{\Psi,1}^{\perp}\mathbf{W}_{2,k}^{\mathrm{T}}\ \triangle_{11}\triangle_{11}^{(1)} - \triangle_{11}^{(1)}\triangle_{11}\triangle_{11}^{(1)}\triangle_{11}\triangle_{11}^{(1)}\\ =&\ \triangle_{11}^{(1)}\triangle_{11}\ \Xi_k\ \triangle_{11}\triangle_{11}^{(1)}\end{aligned} \tag{37}$$

with

$$\begin{aligned}\Xi_k =& \mathbf{W}_{2,k}\mathbf{P}_{\Psi,1}^{\perp}\mathbf{W}_{2,k}^{\mathrm{T}}\\ &- \mathbf{W}_{2,k}\mathbf{P}_{\Psi,1}^{\perp}\bar{\mathbf{T}}_{L,2}^{f}\ \triangle_{11}^{(1)}\left(\bar{\mathbf{T}}_{L,2}^{f}\right)^{\mathrm{T}}\mathbf{P}_{\Psi,1}^{\perp}\mathbf{W}_{2,k}^{\mathrm{T}}.\end{aligned} \tag{38}$$

According to Schur complement lemma (Kailath et al., 2000), Ξ_k in (38) is the Schur complement of the positive semidefinite matrix

$$\begin{bmatrix}\left(\mathbf{T}_{L,2}^{f}\right)^{\mathrm{T}}\mathbf{P}_{\Psi,1}^{\perp}\mathbf{T}_{L,2}^{f} & \left(\bar{\mathbf{T}}_{L,2}^{f}\right)^{\mathrm{T}}\mathbf{P}_{\Psi,1}^{\perp}\mathbf{W}_{2,k}^{\mathrm{T}}\\ \mathbf{W}_{2,k}\mathbf{P}_{\Psi,1}^{\perp}\bar{\mathbf{T}}_{L,2}^{f} & \mathbf{W}_{2,k}\mathbf{P}_{\Psi,1}^{\perp}\mathbf{W}_{2,k}^{\mathrm{T}}\end{bmatrix} \geq 0$$

with regards to its $(2,2)$ block matrix, thus we can conclude $\Xi_k \geq 0$ which finally proves Theorem 5 according to (37). □

4.4 Summary of the proposed MHLS algorithm

Algorithm 1. MHLS input estimator for systems (1) with no unstable transmission zeros

Step 1. Calculate the transmission zeros of (A, E, C, G). If (A, E, C, G) has no transmission zeros, choose the horizon length L to be equal to the observability index ν. If the transmission zeros are all stable, choose sufficiently long horizon to enforce sufficiently small estimation bias.

Step 2. Compute the covariance matrix $\Sigma_{n,L}$ of $\mathbf{n}_{k,L} = \mathbf{T}_L^w \cdot \mathbf{w}_{k,L} + \mathbf{v}_{k,L}$, and generate the residual signal $\bar{\mathbf{r}}_{k,L} = \bar{\mathbf{y}}_{k,L} - \bar{\mathbf{T}}_L^u \cdot \mathbf{u}_{k,L}$ according to (3) and (12).

Step 3. Compute the input estimate $\hat{f}(k)$ according to (24) and (25), where the generalized inverses are replaced with Moore-Penrose inverse.

Remark 2. According to Theorem 3 and 5, the properties of unbiasedness and minimum variance are not affected by the non-unique generalized inverses used to compute (24) and (25). In Step 3 of the above algorithm, we select Moore-Penrose inverse as a special generalized inverse.

Remark 3. Note that Section 3 and 4 focus on sensor faults as stated in Assumption 1. For actuator faults, there would be $G = 0$, and we assume rank $\left(CA^{\tau-1}E\right) = n_f$ where τ is the relative degree of the dynamics from actuator faults to system outputs (Dong and Verhaegen, 2012). In this case, $\mathbf{r}_{k-L+\tau,\tau}$ in (3) has no information about $\mathbf{f}_{k,L}$, hence only $\mathbf{r}_{k,L-\tau}$ is used in input estimation. Based on this, all discussions in Section 3 and 4 can be directly extended by replacing the two matrices C and G with CA^{τ} and $CA^{\tau-1}E$, respectively.

5. SIMULATION STUDIES

Consider the linearized continuous-time VTOL (vertical takeoff and landing) aircraft model used in Dong and Verhaegen (2012). The model has four states, namely horizontal velocity, vertical velocity, pitch rate, and pitch angle. The two inputs are collective pitch control and longitudinal cyclic pitch control, both of which are driven by the second order linear actuator

$$\frac{21.3501s + 162.3867}{s^2 + 17.9994s + 162.3867}.$$

With a sampling rate of 0.5 seconds, the discrete-time model (1) is obtained, with $D = 0$ and $F = I_4$. The process and measurement noise, $w(k)$, $v(k)$, are zero mean white, respectively with a covariance of $Q_w = 0.04 \cdot I_4$ and $Q_v = 0.01 \cdot I_2$.

Two fault scenarios are considered:

(i) Sensor fault: $E = 0$, $G = \left[\begin{smallmatrix}1&0&0&0\\0&1&0&0\end{smallmatrix}\right]^{\mathrm{T}}$, and the fault signals are

$$f(k) = \begin{cases}[0,\ 0]^{\mathrm{T}}, & k \leq 500\\ [0.01\,(k-500),\ 1]^{\mathrm{T}}, & 500 < k \leq 1000.\end{cases} \tag{39}$$

(ii) Actuator fault: $E = B$, $G = 0$, and the fault signals are the same as (39).

In the above fault scenarios, our proposed MHLS input estimator will be compared with the methods of Dong and Verhaegen (2012) and Gillijns and Moor (2007a,b). Note that although the motivation of Dong and Verhaegen (2012) is data-driven design of fault estimation filter, here we use the equations (24)-(25) in Dong and Verhaegen (2012) as a model-based fault estimation filter.

In the sensor fault scenario, the dynamics from unknown fault signals to the system outputs, i.e., (A, E, C, G),

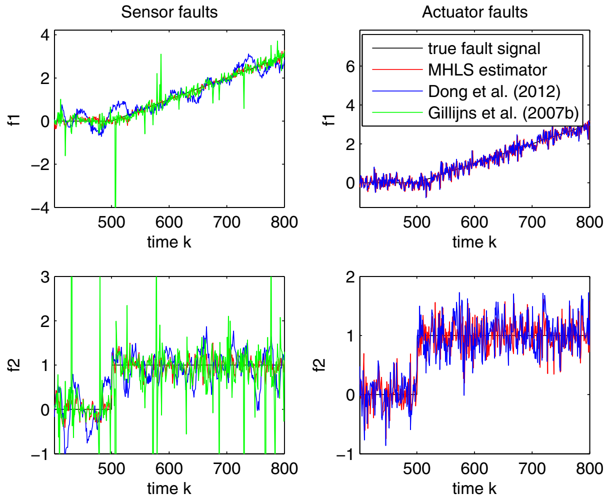

Fig. 1. Results of MHLS input estimator ($L = 15$) and the estimators in Dong and Verhaegen (2012) and Gillijns and Moor (2007b)

has no transmission zeros. This means the sensor fault signals are observable according to Theorem 1. Note that the fault estimation filter in Dong and Verhaegen (2012) would be unstable if it is directly applied to the open-loop plant in this case. To solve this problem, Dong and Verhaegen (2012) suggests to use a controller $\mathcal{C}_u$ such that the plant is stabilized and the sensor faults are in the kernel of $\mathcal{C}_u$. Then the closed-loop system model is utilized to design the fault estimation filter. The other two input estimators are based on the open-loop plant model. The simulation results are shown in Fig. 1 with emphasis on a time window of 400 samples for better illustration. There exist oscillations in the estimates of the fault estimation filter in Dong and Verhaegen (2012). The method of Gillijns and Moor (2007b) suffers from instability in certain time intervals. The estimation performance of our MHLS estimator with different horizon L is shown in Fig. 2, and compared against that of the fault estimation filter in Dong and Verhaegen (2012). It can be seen that the MHLS estimator gives smaller bias, and its error variance decreases with increased horizon length.

In the actuator fault scenario, the relative degree of the dynamics from actuator faults to outputs is $\tau = 2$. According to Remark 3 and Theorem 2, since $\left(A, E, CA^2, CAE\right)$ has only two stable transmission zeros, the actuator fault signal is detectable. The method of Gillijns and Moor (2007a,b) is not used in this scenario because it is limited to systems with relative degree $\tau = 0$ or 1. The simulation results in Fig. 1 and 2 show that the MHLS estimator gives smaller bias and error variance than the fault estimation filter in Dong and Verhaegen (2012).

6. CONCLUSIONS

This paper proposes a novel MHLS input estimation method for linear discrete-time stochastic systems. Based on necessary and sufficient conditions for input observability and detectability, it is shown that our MHLS estimator gives an unbiased estimate for observable unknown input, and an asymptotically unbiased estimate for detectable unknown input. Furthermore, the MHLS estimator has

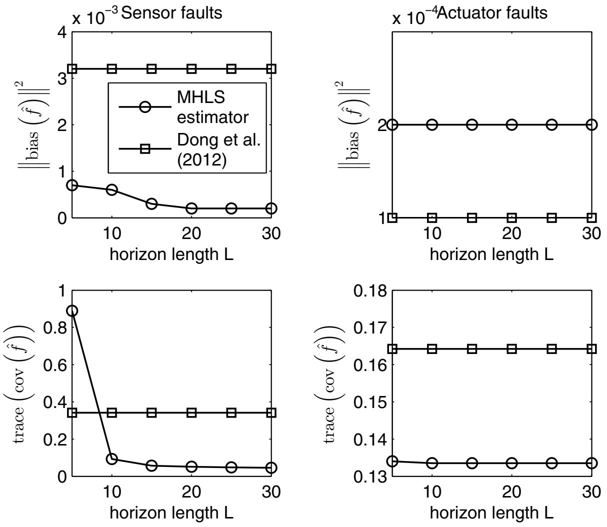

Fig. 2. Bias and estimation error variance of MHLS input estimator and the estimator in Dong and Verhaegen (2012)

minimal variance over the class of all linear asymptotically unbiased estimators. Its effectiveness is illustrated by aircraft sensor and actuator fault examples.

ACKNOWLEDGEMENTS

The authors would like to thank B. Kulcsár and J. Dong for helpful discussions of the ideas in this paper.

REFERENCES

Basile, G. and Marro, G. (1992). *Controlled and Conditioned Invariants in Linear System Theory*. Prentice Hall, Englewood Cliffs.

Dong, J. and Verhaegen, M. (2012). Identification of fault estimation filter from I/O data for systems with stable inversion. *IEEE Transactions on Automatic Control*, 57(6), 1347–1360.

Gillijns, S. (2007). *Kalman Filtering Techniques for System Inversion and Data Assimilation*. Ph.D. thesis, Katholieke University Leuven.

Gillijns, S. and Moor, B.D. (2007a). Unbiased minimum-variance input and state estimation for linear discrete-time systems. *Automatica*, 43(1), 111–116.

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

Hou, M. and Patton, R.J. (1998). Input observability and input reconstruction. *Automatica*, 34(6), 789–794.

Kailath, T., Sayed, A.H., and Hassibi, B. (2000). *Linear Estimation*. Prentice Hall, Upper Saddle River, New Jersey.

Kirtikar, S., Palanthandalam-Madapusi, H., Zattoni, E., and Bernstein, D.S. (2011). l-delay input and initial-state reconstruction for discrete-time linear systems. *Circuits, systems, and signal processing*, 30(1), 233–262.

Zhou, K., Doyle, J., and Glover, K. (1996). *Robust and Optimal Control*. Prentice Hall, Upper Saddle River, New Jersey.