Automatica 103 (2019) 11–19

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Deadbeat unknown-input state estimation and input reconstruction for linear discrete-time systems[☆]

Ahmad Ansari, Dennis S. Bernstein [*]

Department of Aerospace Engineering, The University of Michigan, Ann Arbor, MI 48109, United States

ARTICLE INFO

Article history:
Received 28 June 2017
Received in revised form 11 November 2018
Accepted 31 December 2018
Available online 6 February 2019

Keywords:
Input reconstruction
Unknown-input state estimation
Generalized inverse
Deadbeat observer

ABSTRACT

This paper considers discrete-time input reconstruction and state estimation assuming that the system has no invariant zeros, without assuming that the initial condition is known, and without assuming that at least one Markov parameter has full column rank. Algorithms based on the generalized inverse of a block-Toeplitz matrix are given for unknown-input state estimation and simultaneous input reconstruction and state estimation. In both cases, the unknown input is an arbitrary signal. Both algorithms are deadbeat, which means that exact input reconstruction and state estimation are achieved in a finite number of steps.

1. Introduction

State estimation uses measurements of the output of a system to produce statistically optimal estimates of the states of the system (Crassidis & Junkins, 2011; Lewis, Xie, & Popa, 2007; Simon, 2006). These estimates assume that the exogenous input consists of a known deterministic component, which is replicated in the estimator, and an unknown stochastic disturbance, which is assumed to be white and zero mean. If the deterministic input is unknown, then it cannot be replicated in the observer, and thus the state estimates may be biased. To remedy this problem, state estimators have been developed to provide unbiased state estimates in the presence of unknown, deterministic inputs (Darouach & Zasadzinski, 1997; Glover, 1969; Hou & Patton, 1998a; Kitanidis, 1987; Valcher, 1999).

An alternative approach is to extend state estimation to include input estimation, where the goal is to estimate the deterministic component of the exogenous input (Ansari & Bernstein, 2018; Chakrabarty, Ayoub, Zak, & Sundaram, 2017; Chen & Chen, 2010; Corless & Tu, 1998; Fang, de Callafon, & Cortes, 2013; Fang, Shi, & Yi, 2011; Floquet & Barbot, 2006; Gillijns & De Moor, 2007; Ho & Ma, 2007; Hou & Patton, 1998b; Hsieh, 2000, 2009; Khaloozadeh & Karsaz, 2009; Kirtikar, Palanthandalam-Madapusi, Zattoni, & Bernstein, 2011; Lu, van Kampen, de Visser, & Chu, 2016; Massey & Sain, 1968; Orjuela, Marx, Ragot, & Maquin, 2009; Palanthandalam-Madapusi & Bernstein, 2009; Sain & Massey, 1969; Sanchez & Benaroya, 2014; Willsky, 1974; Xiong & Saif, 2003; Yang, Zhu, & Sun, 2013; Yong, Zhu, & Frazzoli, 2016). In many applications, knowledge of the input signal is of independent interest and, in some cases, may be of greater interest than the estimates of the states (Rajamani, Wang, Nelson, Madson, & Zemouche, 2017). The terminology *input reconstruction* is used in the case of deterministic analysis, just as an observer is the deterministic analogue of an estimator.

In light of state estimation, which assumes a known deterministic input and an unknown zero-mean stochastic input, it may be somewhat surprising that it is indeed possible to estimate not only the states but also, in many cases, the unknown input. The benefit of state and input estimation is the fact that knowledge of the input can often vastly improve the accuracy of the state estimates.

The present paper considers input reconstruction within a discrete-time setting. In particular, novel algorithms are given for unknown-input state estimation and simultaneous input reconstruction and state estimation in terms of the generalized inverse of a block-Toeplitz matrix. In Gillijns and De Moor (2007) and Yong et al. (2016) it is assumed that the first Markov parameter H_1 has full column rank, which implies that the plant has relative degree 1. Likewise, the approach of Kirtikar et al. (2011) is limited to the case where at least one Markov parameter has full column rank. The present paper considers a more general case where no Markov parameter is required to have full column rank.

The algorithms given in the present paper provide deadbeat (that is, finite-step) unknown-input state estimation and simultaneous input reconstruction and state estimation without assuming that the initial condition is known. In this case, the presence of an

☆ This research was supported in part by the Office of Naval Research, United States under grant N00014-14-1-0596. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Brett Ninness under the direction of Editor Torsten Söderström.

* Corresponding author.
E-mail addresses: ansahmad@umich.edu (A. Ansari), dsbaero@umich.edu (D.S. Bernstein).

https://doi.org/10.1016/j.automatica.2019.01.011

Table 1
State Estimation and Input Reconstruction with unknown $x(0)$.

	Asymptotic Estimation	Deadbeat Estimation
Unknown-input state estimation	(Kitanidis, 1987) • unbiased minimum variance filter • assumes that H_1 has full column rank • allows minimum-phase zeros	Theorem 6 • deadbeat FIR filter for state estimation • the inherent delay is μ • requires $\mu+1$ measurements • allows rank-deficient Markov parameters • assumes no invariant zeros
State and input estimation	(Gillijns & De Moor, 2007; Yong et al., 2016) • unbiased minimum variance filter • assumes that H_1 has full column rank • allows minimum-phase zeros (Hsieh, 2017; Yong, Zhu, & Frazzoli, 2015) • unbiased minimum variance filter with a delay • H_1 need not have full column rank • allows minimum-phase zeros (Sundaram & Hadjicostis, 2007) • reduced-order state observers • allows rank-deficient Markov parameters • allows minimum-phase zeros	Theorem 7 • deadbeat FIR filter for input reconstruction • the inherent delay is η • requires $\max(\eta,\mu)+1$ measurements • allows rank-deficient Markov parameters • assumes no invariant zeros

invariant zero makes it impossible to distinguish the zero input with zero initial condition from a nonzero input with a specific initial condition that yields zero response. This case is considered in Hou and Patton (1998b), where an algorithm is given for constructing an input-reconstruction filter based on an observability assumption, which rules out the presence of invariant zeros. For the case where $x(0)$ is unknown and (A, B, C, D) has no invariant zeros, Theorem 6 provides a deadbeat unknown-input observer, that is, an algorithm that exactly estimates the state despite the presence of an unknown, arbitrary input. Although this performance is better than the Kalman filter in the absence of sensor noise, it has to be kept in mind that the estimates are obtained with a delay, which means that the estimator is effectively a smoother. Furthermore, for the case where the system has no invariant zeros, Theorem 7 provides deadbeat input reconstruction and state estimation. In this case, the input-reconstruction delay is η, and the number of required measurements is $\max\{\mu,\eta\}+1$. The algorithms constructed in Theorems 6 and 7 are finite-impulse response (FIR) systems. FIR filters for state estimation are given in Shmaliy, Zhao, and Ahn (2017) and Kim (2010); however, these results assume that the input is known. Furthermore, FIR filters for input estimation are given in Park, Kim, Kwon, and Kwon (2000); however, deadbeat input reconstruction is not considered.

Input reconstruction is related to the problem of system inversion. The inversion techniques of Silverman (1969) and Sain and Massey (1969) are based on constructive algorithms that entail the sequential decomposition of various matrices until a full-rank condition is attained. In contrast to these constructions, Theorems 6 and 7 are given in terms of the generalized inverse of a single matrix.

If the initial condition is unknown and the system has at least one invariant zero, then deadbeat input reconstruction is not possible. In this case, asymptotic input reconstruction must be considered, with careful attention paid to the presence of nonminimum-phase zeros. Table 1 lists various cases that can occur, the relevant literature in each case, and the contribution of the present paper.

The assumption invoked in Theorems 6 and 7 that the system has no invariant zeros is clearly restrictive in the SISO case, since it is unusual for an nth-order SISO system to have relative degree n. Furthermore, since the transmission zeros of a square MIMO transfer function with full normal rank are the roots of the numerator of the determinant, it would be unusual for the system to have no transmission zeros. The situation is different, however, for rectangular systems. For example, a MIMO system with two inputs and four outputs and full normal rank possesses a transmission zero if and only if all six 2×2 embedded transfer functions possess a common transmission zero. Consequently, input reconstruction based on Theorem 7 may be useful for a large class of rectangular systems.

The contents of the paper are as follows. The Section 2 presents the input-reconstruction problem for discrete-time linear systems. Section 3 gives preliminaries on the invertibility of a linear system with an input reconstruction delay. Next, Section 4 defines the minimum delay η for input reconstruction, and gives necessary and sufficient conditions under which η is finite. Finally, Sections 5 and 6 provide Theorems 6 and 7 for μ-delay state estimation and η-delay input reconstruction, respectively.

2. Problem statement

Let $A\in\mathbb{R}^{n\times n}$, $B\in\mathbb{R}^{n\times m}$, $C\in\mathbb{R}^{p\times n}$, and $D\in\mathbb{R}^{p\times m}$, assume that (A, B, C, D) is minimal, and consider

$$x(k+1) = Ax(k)+Bu(k), \tag{1}$$

$$y(k) = Cx(k)+Du(k), \tag{2}$$

where, for all $k\ge 0$, $x(k)\in\mathbb{R}^n$, $u(k)\in\mathbb{R}^m$, and $y(k)\in\mathbb{R}^p$. The goal is to use knowledge of $y(k)$ to estimate the unknown input $u(k)$.

For all $l\ge 0$, define the lth Markov parameter

$$H_l \triangleq \begin{cases} D, & l=0,\\ CA^{l-1}B, & l\ge 1.\end{cases} \tag{3}$$

Let r denote a nonnegative integer, and define

$$\mathcal{Y}_r \triangleq \begin{bmatrix} y(0)\\ y(1)\\ \vdots\\ y(r)\end{bmatrix},\quad \mathcal{U}_r \triangleq \begin{bmatrix} u(0)\\ u(1)\\ \vdots\\ u(r)\end{bmatrix},\quad \Gamma_r \triangleq \begin{bmatrix} C\\ CA\\ \vdots\\ CA^r\end{bmatrix}, \tag{4}$$

$$M_r \triangleq \begin{bmatrix} H_0 & 0 & 0 & \cdots & 0\\ H_1 & H_0 & 0 & \cdots & 0\\ H_2 & H_1 & H_0 & \cdots & 0\\ \vdots & \vdots & \ddots & \ddots & \vdots\\ H_r & H_{r-1} & \cdots & H_1 & H_0\end{bmatrix} \in \mathbb{R}^{(r+1)p\times(r+1)m}. \tag{5}$$

It follows from (1), (2) that

$$\mathcal{Y}_r = \Gamma_r x(0) + M_r\mathcal{U}_r = \Psi_r\begin{bmatrix} x(0)\\ \mathcal{U}_r\end{bmatrix}, \tag{6}$$

where

$$\Psi_r \triangleq \begin{bmatrix} \Gamma_r & M_r\end{bmatrix} \in \mathbb{R}^{(r+1)p\times[n+(r+1)m]}. \tag{7}$$

Note that, since (A, C) is observable, it follows that

$$x(0) = \Gamma_n^{+}(\mathcal{Y}_n - M_n\mathcal{U}_n), \tag{8}$$

where Γ_n^{+} is the pseudo inverse of Γ_n. For $r \geq s \geq 0$, it is convenient to partition M_r as

$$M_r = \left[\begin{array}{cccc|cc} H_0 & 0 & \cdots & & \cdots & 0 \\ \vdots & \ddots & \ddots & & \vdots & \vdots \\ H_s & \ddots & \ddots & & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & 0 \\ \underbrace{H_r \ \cdots \ H_s}_{N_{r,s}} & & & & \underbrace{\cdots \ H_0}_{Q_{r,s}} & \end{array}\right] = [\ \underbrace{C_r \ \cdots \ C_s}_{N_{r,s}} \mid \underbrace{\cdots \ C_0}_{Q_{r,s}}\ ], \tag{9}$$

where $N_{r,s} \in \mathbb{R}^{(r+1)p\times(r-s+1)m}$, $Q_{r,s} \in \mathbb{R}^{(r+1)p\times sm}$, and, for all $i \in \{0, \ldots, r\}$, C_i denotes the $(i+1)$th block column of M_r labeled right to left. Furthermore, since, for all $r \geq 0$,

$$M_r = \left[\begin{array}{c|c} H_0 & 0 \\ \hline H_1 & \\ \vdots & M_{r-1} \\ H_r & \end{array}\right] = \left[\begin{array}{ccc|c} & M_{r-1} & & 0 \\ \hline H_r & \cdots & H_1 & H_0 \end{array}\right], \tag{10}$$

it follows that, for all $r \geq 0$,

$$\operatorname{rank} M_{r-1} \leq \operatorname{rank} M_r \tag{11}$$
$$\leq \min\{\operatorname{rank} C_r, \operatorname{rank} R_r\} + \operatorname{rank} M_{r-1} \tag{12}$$
$$\leq m + \operatorname{rank} M_{r-1}, \tag{13}$$

where $R_r \triangleq [H_r \ \cdots \ H_0]$ is the last block row of M_r and M_{-1} is an empty matrix whose rank is 0 and range is $\{0\}$. Finally, note that, if $r > s \geq 0$ and $N_{r,s}$ has full column rank, then, for all $s' \in \{s+1, \ldots, r\}$, $N_{r,s'}$ has full column rank.

3. Preliminaries on *d*-delay invertibility

Let $G \in \mathbb{R}(\mathbf{z})^{p\times m}$ be the $p \times m$ proper rational transfer function corresponding to (1) and (2).

Definition 1. Let d be a nonnegative integer. Then G is d-delay invertible if there exists $\hat{G} \in \mathbb{R}(\mathbf{z})^{m\times p}$ such that $\hat{G}(\mathbf{z})G(\mathbf{z}) = \mathbf{z}^{-d}I_m$. $\hat{G}$ is a d-delay left inverse of G.

Note that, if G is d-delay invertible, then G must have full normal column rank, and thus $m \leq p$, that is, G must be square or tall. Furthermore, if G is d-delay invertible, then, for all $r > d$, G is r-delay invertible.

It follows from (11) and (13) that $\operatorname{rank} M_r \leq m + \operatorname{rank} M_{r-1}$. The following result shows that equality in either the case $r = d$ or $r = n$ is necessary and sufficient for invertibility.

Proposition 2. *The following conditions are equivalent:*

(i) *There exists $d \geq 0$ such that G is d-delay invertible.*
(ii) *G has full column normal rank.*
(iii) $\operatorname{rank} N_{2n,n} = (n+1)m$.
(iv) *There exists $d \geq 0$ such that* $\operatorname{rank} M_d - \operatorname{rank} M_{d-1} = m$.
(v) $\operatorname{rank} M_n - \operatorname{rank} M_{n-1} = m$.

If these conditions hold, then there exists $d \geq 0$ such that $(1/\mathbf{z}^d)[G(\mathbf{z})^{\mathrm{T}} G(\mathbf{z})]^{-1}G(\mathbf{z})^{\mathrm{T}}$ is a d-delay inverse of G.

Proof. The equivalence of (i) and (ii) is immediate. The equivalence of (i) and (iii) is given by Theorem 3 of Sain and Massey (1969). The equivalence of (i) and (iv) is given by Theorem 2 of Sain and Massey (1969) and Theorem 4 of Massey and Sain (1968). The equivalence of (i) and (v) is given by Corollary 1 of Sain and Massey (1969). □

4. Input reconstruction delay

The existence of a d-delay left inverse of G implies that, if $x(0) = 0$, then the output of the cascaded system $\hat{G}G$ is exactly the input sequence $u(0), u(1), \ldots$ delayed by d steps. However, for several reasons, the d-delay inverse $\hat{G}(\mathbf{z}) = (1/\mathbf{z}^d)[G(\mathbf{z})^{\mathrm{T}}G(\mathbf{z})]^{-1}G(\mathbf{z})^{\mathrm{T}}$ given by Proposition 2 may be deficient. In particular, $\hat{G}$ may be unstable; the cascade $\hat{G}G$ may entail nonminimum-phase pole-zero cancellation; and the McMillan degree of $\hat{G}$ may not be the smallest possible value.

It is desirable to achieve the smallest possible delay d such that G is d-delay invertible. We thus define

$$\eta \triangleq \min\{l \geq 0 : \operatorname{rank} M_l = m + \operatorname{rank} M_{l-1}\}. \tag{14}$$

Note that G is d-delay invertible if and only if η is finite. Furthermore, the equivalence of (i) and (iv) of Proposition 2 implies that, if G is d-delay invertible, then η is the smallest delay d such that G is d-delay invertible. Finally, (v) of Proposition 2 implies that $\eta \leq n$. A sharper bound is given in Proposition 3.

We now focus on sufficient and necessary conditions under which η is finite. In the following result, the first three statements are immediate, and the last statement is given by Corollary 1 of Willsky (1974).

Proposition 3. *The following statements hold:*

(i) *Let $q \geq 0$ be the smallest nonnegative integer such that H_q is nonzero, and assume that H_q has full column rank. Then $\eta = q$.*
(ii) *If $p < m$, then η is infinite.*
(iii) *Assume that, for all $r \geq 0$, either* $\operatorname{rank} R_r < p$ *or* $\operatorname{rank} C_r < m$. *Then η is infinite.*
(iv) *If η is finite, then* $\eta \leq \min\{n, n+1-m+\operatorname{rank} D\}$.

(i) implies that, if $m = 1$, then η is the index of the first nonzero Markov parameter. Therefore, in the SISO case $m = p = 1$, η is the relative degree of G. (ii) shows that η is finite only if G is either square or tall. (iii) implies that, if η is finite, then there exists a nonnegative integer r such that either R_r has full row rank or C_r has full column rank. However, Example 1 shows that the converse of this statement is not true. The second bound in (iv) is given in Willsky (1974).

The following example illustrates the range of possible values of η in the case $p = 3$ and $m = 2$.

Example 1. Let $p = 3$ and $m = 2$, and consider $G(\mathbf{z}) = C(\mathbf{z}I - A)^{-1}B + D$ given by

$$G(\mathbf{z}) = \frac{1}{\mathbf{z}^4}(H_4 + H_3\mathbf{z} + H_2\mathbf{z}^2 + H_1\mathbf{z}^3). \tag{15}$$

Note that $D = H_0 = 0_{3\times 2}$, and thus $\operatorname{rank} M_0 = 0 < m$. If

$$H_1 = \begin{bmatrix} 0 & 1 \\ 1 & 2 \\ 0 & 1 \end{bmatrix}, H_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, H_3 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}, H_4 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 2 \end{bmatrix}, \tag{16}$$

then $\operatorname{rank} M_1 = 2 = m$, and thus $\eta = 1$. Alternatively, if

$$H_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}, H_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, H_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, H_4 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 2 \end{bmatrix}, \tag{17}$$

then $\operatorname{rank} M_1 = 1 < m$, and, for all $l \geq 2$, $\operatorname{rank} M_l - \operatorname{rank} M_{l-1} = 2 = m$, and thus $\eta = 2$. Next, if

$$H_1 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}, H_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, H_3 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}, H_4 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 2 \end{bmatrix}, \tag{18}$$

then, for all $l \le 3$, $\operatorname{rank} M_l - \operatorname{rank} M_{l-1} = 1 < m$, and, for all $l \ge 4$, $\operatorname{rank} M_l - \operatorname{rank} M_{l-1} = 2 = m$, and thus $\eta = 4$. Finally, if

$$H_1 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix},\ H_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},\ H_3 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix},\ H_4 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, \tag{19}$$

then, for all $l \ge 1$, $\operatorname{rank} M_l - \operatorname{rank} M_{l-1} = 1 \ne m$, and thus η is infinite. $\diamond$

The cases (16)–(18) show that η may be finite whether or not at least one Markov parameter has full column rank. Furthermore, the cases (18) and (19) show that, if no Markov parameter has full column rank, then η may be finite or infinite.

The following result, which assumes that η is finite, is used in the proof of Theorem 7. The proof depends on Lemmas B and C. Note that, since, by (*iv*) of Proposition 3, $\eta \le n$, it follows that the first result of Proposition 4 generalizes (*i*) $\Longrightarrow$ (*iii*) of Proposition 2.

Proposition 4. *Assume that η is finite, and let $r \ge \eta$. Then, $N_{r,\eta}$ has full column rank, and*

$$\mathcal{R}(N_{r,\eta}) \cap \mathcal{R}(Q_{r,\eta}) = \{0\}. \tag{20}$$

Proof. First, consider the case $\eta = 0$. Then $M_0 = N_0 = H_0$ and $\operatorname{rank} M_0 = \operatorname{rank} N_0 = \operatorname{rank} H_0 = m$. Since H_0 has full column rank, it follows that $N_{r,\eta}$ has full column rank, and, since $Q_{r,\eta}$ is an empty matrix, (20) holds.

Next, let $r = \eta = 1$ so that

$$\operatorname{rank} M_1 = m + \operatorname{rank} M_0. \tag{21}$$

Since

$$M_1 = \left[\begin{array}{c|c} C_1 & \begin{matrix} 0 \\ M_0 \end{matrix} \end{array}\right] = \left[\begin{array}{c|c} C_1 & C_0 \end{array}\right] = \left[\begin{array}{c|c} N_{1,1} & Q_{1,1} \end{array}\right], \tag{22}$$

it follows from Lemma B that

$$\begin{aligned} \operatorname{rank} M_1 &= \operatorname{rank} C_1 + \operatorname{rank} C_0 - \dim\left(\mathcal{R}(C_1) \cap \mathcal{R}(C_0)\right) \\ &= \operatorname{rank} N_{1,1} + \operatorname{rank} M_0 - \dim\left(\mathcal{R}(N_{1,1}) \cap \mathcal{R}(Q_{1,1})\right). \end{aligned} \tag{23}$$

Combining (21) with (23) yields

$$0 \le \dim\left(\mathcal{R}(N_{1,1}) \cap \mathcal{R}(Q_{1,1})\right) = \operatorname{rank} N_{1,1} - m \le 0,$$

which implies that $N_{1,1}$ has full column rank and $\mathcal{R}(N_{1,1}) \cap \mathcal{R}(Q_{1,1}) = \{0\}$.

Next, let $r \ge 2$ and $\eta \in \{1, \ldots, r\}$ so that

$$\operatorname{rank} M_\eta = m + \operatorname{rank} M_{\eta-1}. \tag{24}$$

Noting

$$M_\eta = \left[\begin{array}{c|c} C_\eta & \begin{matrix} 0 \\ M_{\eta-1} \end{matrix} \end{array}\right] = \left[\begin{array}{c|ccc} C_\eta & C_{\eta-1} & \cdots & C_0 \end{array}\right], \tag{25}$$

it follows from Lemma B that

$$\begin{aligned} \operatorname{rank} M_\eta &= \operatorname{rank}\ C_\eta + \operatorname{rank}\ [C_{\eta-1}\ \cdots\ C_0] \\ &\quad - \dim\left(\mathcal{R}(C_\eta) \cap \mathcal{R}([C_{\eta-1}\ \cdots\ C_0])\right) \\ &= \operatorname{rank}\ C_\eta + \operatorname{rank}\ M_{\eta-1} \\ &\quad - \dim\left(\mathcal{R}(C_\eta) \cap \mathcal{R}([C_{\eta-1}\ \cdots\ C_0])\right). \end{aligned} \tag{26}$$

Combining (24) with (26) yields

$$0 \le \dim\left(\mathcal{R}(C_\eta) \cap \mathcal{R}([C_{\eta-1}\ \cdots\ C_0])\right) = \operatorname{rank}\ C_\eta - m \le 0,$$

which implies that C_η has full column rank and

$$\mathcal{R}(C_\eta) \cap \mathcal{R}([C_{\eta-1}\ \cdots\ C_0]) = \{0\}. \tag{27}$$

It thus follows from Lemma C that $N_{r,\eta} = [C_r\ \cdots\ C_\eta]$ has full column rank and

$$\begin{aligned} \mathcal{R}(N_{r,\eta}) \cap \mathcal{R}(Q_{r,\eta}) &= \mathcal{R}([C_r\ \cdots\ C_\eta]) \cap \mathcal{R}([C_{\eta-1}\ \cdots\ C_0]) \\ &= \{0\}. \quad \square \end{aligned}$$

5. Deadbeat unknown-input state estimation for systems without invariant zeros

Define

$$\mu \triangleq \min\{l \ge 0 : \operatorname{rank} \Psi_l = n + \operatorname{rank} M_l\}. \tag{28}$$

The index μ is the smallest integer such that Γ_μ has full column rank and the disjointness condition (29) is valid.

The following result is used in the proofs of Theorems 6 and 7.

Proposition 5. *Assume that (A, B, C, D) has no invariant zeros. Then μ is finite, and, for all $r \ge \mu$, Γ_r has full column rank and*

$$\mathcal{R}(\Gamma_r) \cap \mathcal{R}(M_r) = \{0\}. \tag{29}$$

Proof. Since (A, B, C, D) has no invariant zeros, Theorem A.1 of Kirtikar et al. (2011) implies that there exists $l \ge n$ such that

$$\mathcal{R}(\Gamma_l) \cap \mathcal{R}(M_l) = \{0\}. \tag{30}$$

Since (A, C) is observable, it follows that

$$\operatorname{rank} \Gamma_l = n. \tag{31}$$

Noting $\Psi_l = [\Gamma_l\ M_l]$ and using (30), (31), and Lemma B, it follows that

$$\begin{aligned} \operatorname{rank} \Psi_l &= \operatorname{rank} \Gamma_l + \operatorname{rank} M_l - \dim(\mathcal{R}(\Gamma_l) \cap \mathcal{R}(M_l)) \\ &= n + \operatorname{rank} M_l. \end{aligned} \tag{32}$$

It thus follows from (32) that μ is finite and satisfies $0 \le \mu \le l$. Next, note that

$$\operatorname{rank} \Psi_\mu = n + \operatorname{rank} M_\mu. \tag{33}$$

Furthermore, noting $\Psi_\mu = [\Gamma_\mu\ M_\mu]$ and using Lemma B yields

$$\operatorname{rank} \Psi_\mu = \operatorname{rank} \Gamma_\mu + \operatorname{rank} M_\mu - \dim(\mathcal{R}(\Gamma_\mu) \cap \mathcal{R}(M_\mu)). \tag{34}$$

Combining (33) with (34) yields

$$0 \le \dim\left(\mathcal{R}(\Gamma_\mu) \cap \mathcal{R}(M_\mu)\right) = \operatorname{rank}\ \Gamma_\mu - n \le 0,$$

which implies that Γ_μ has full column rank and

$$\mathcal{R}(\Gamma_\mu) \cap \mathcal{R}(M_\mu) = \{0\}. \tag{35}$$

Since Γ_μ has full column rank, it thus follows from (4) that, for all $r \ge \mu$, Γ_r has full column rank. Finally, note that

$$\mathcal{R}(\Gamma_{\mu+1}) \cap \mathcal{R}(M_{\mu+1}) = \mathcal{R}\left(\begin{bmatrix} \Gamma_\mu \\ CA^{\mu+1} \end{bmatrix}\right) \cap \mathcal{R}\left(\left[\begin{array}{ccc|c} & M_\mu & & 0 \\ \hline H_{\mu+1} & \cdots & H_1 & H_0 \end{array}\right]\right). \tag{36}$$

Since Γ_μ has full column rank and $\mathcal{R}(\Gamma_\mu) \cap \mathcal{R}(M_\mu) = \{0\}$, it follows from (36) and Lemma A that

$$\mathcal{R}(\Gamma_{\mu+1}) \cap \mathcal{R}(M_{\mu+1}) = \{0\}.$$

By similar arguments, it follows that, for all $r \ge \mu$, $\mathcal{R}(\Gamma_r) \cap \mathcal{R}(M_r) = \{0\}$. $\square$

The following example compares μ and η for a system that has no invariant zeros.

Example 2. Since (15) has no invariant zeros for (16)–(18), Proposition 5 implies that μ is finite for each of these cases. For (16), $n = 7$, $\operatorname{rank}\Psi_0 - \operatorname{rank}M_0 = 3 < n$, $\operatorname{rank}\Psi_1 - \operatorname{rank}M_1 = 4 < n$, $\operatorname{rank}\Psi_2 - \operatorname{rank}M_2 = 5 < n$, $\operatorname{rank}\Psi_3 - \operatorname{rank}M_3 = 6 < n$, and, for all $l \geq 4$, $\operatorname{rank}\Psi_l - \operatorname{rank}M_l = 7 = n$. Hence, $\mu = 4 > \eta = 1$.

For (17), $n = 6$, $\operatorname{rank}\Psi_0 - \operatorname{rank}M_0 = 3 < n$, $\operatorname{rank}\Psi_1 - \operatorname{rank}M_1 = 5 < n$, and, for all $l \geq 2$, $\operatorname{rank}\Psi_l - \operatorname{rank}M_l = 6 = n$. Hence, $\mu = \eta = 2$.

For (18), $n = 7$, $\operatorname{rank}\Psi_0 - \operatorname{rank}M_0 = 3 < n$, $\operatorname{rank}\Psi_1 - \operatorname{rank}M_1 = 5 < n$, $\operatorname{rank}\Psi_2 - \operatorname{rank}M_2 = 6 < n$, and, for all $l \geq 3$, $\operatorname{rank}\Psi_l - \operatorname{rank}M_l = 7 = n$. Hence, $\mu = 3 < \eta = 4$.

Finally, for (19), (A, B, C, D) has an invariant zero at 0, and thus Proposition 5 is not applicable. $\diamond$

The following result shows that, if (A, B, C, D) has no invariant zeros, then deadbeat state estimation is possible with a delay of μ steps without knowledge of u. The proof depends on Proposition 5 and Lemma D.

Theorem 6. *Assume that (A, B, C, D) has no invariant zeros. Then, for all $k \geq 0$ and $r \geq \mu$,*

$$x(k) = \begin{bmatrix} I_n & 0_{n\times(r+1)m} \end{bmatrix} \Psi_r^+ \begin{bmatrix} y(k) \\ y(k+1) \\ \vdots \\ y(k+r) \end{bmatrix}. \tag{37}$$

Proof. Let $k \geq 0$. Since (A, B, C, D) has no invariant zeros, it follows from Proposition 5 that, for all $r \geq \mu$, Γ_r has full column rank, and

$$\mathcal{R}(\Gamma_r) \cap \mathcal{R}(M_r) = \{0\}. \tag{38}$$

Using (38), it follows from Lemma D that, for all $r \geq \mu$,

$$\Psi_r^+\Psi_r = \begin{bmatrix} \Gamma_r^+\Gamma_r & 0 \\ 0 & M_r^+M_r \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ 0 & M_r^+M_r \end{bmatrix}. \tag{39}$$

Next, multiplying (39) by $[x^{\mathrm{T}}(k)\ u^{\mathrm{T}}(k)\ \cdots\ u^{\mathrm{T}}(k+r)]^{\mathrm{T}}$ implies that, for all $r \geq \mu$,

$$\begin{bmatrix} I_n & 0 \\ 0 & M_r^+M_r \end{bmatrix} \begin{bmatrix} x(k) \\ u(k) \\ \vdots \\ u(k+r) \end{bmatrix} = \Psi_r^+ \begin{bmatrix} y(k) \\ y(k+1) \\ \vdots \\ y(k+r) \end{bmatrix}. \tag{40}$$

Finally, multiplying (40) by $[I_n\ 0_{n\times(r+1)m}]$ implies that, for all $r \geq \mu$, (37) holds. $\square$

Example 3. Consider the mass–spring–damper system with masses m_1, m_2 and unknown input force u applied to m_1, as shown in Fig. 1. The dynamics are given by

$$\dot{x} = A_{\mathrm{c}}x + B_{\mathrm{c}}u, \tag{41}$$

where

$$A_{\mathrm{c}} \triangleq \begin{bmatrix} 0_{2\times2} & I_{2\times2} \\ \Omega_1 & \Omega_2 \end{bmatrix},\ B_{\mathrm{c}} \triangleq \begin{bmatrix} 0_{2\times1} \\ \Omega_3 \end{bmatrix},\ \Omega_1 \triangleq \begin{bmatrix} -\frac{k_1+k_2}{m_1} & \frac{k_2}{m_1} \\ \frac{k_2}{m_2} & -\frac{k_2}{m_2} \end{bmatrix},$$

$$\Omega_2 \triangleq \begin{bmatrix} -\frac{c_1+c_2}{m_1} & \frac{c_2}{m_1} \\ \frac{c_2}{m_2} & -\frac{c_2}{m_2} \end{bmatrix},\ \Omega_3 \triangleq \begin{bmatrix} \frac{1}{m_1} \\ 0 \end{bmatrix},$$

x_1 and x_2 are the displacements and x_3 and x_4 are the velocities of m_1 and m_2, respectively. Letting $m_1 = m_2 = 1$ kg, $k_1 = k_2 = 10$ N/m, and $c_1 = c_2 = 5$ kg/sec, we discretize (41) as

$$A = e^{A_{\mathrm{c}}T_{\mathrm{s}}},\quad B = A_{\mathrm{c}}^{-1}(A - I)B_{\mathrm{c}}, \tag{42}$$

where $T_{\mathrm{s}} = 1$ sec is the sampling time. Letting

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix},\quad D = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

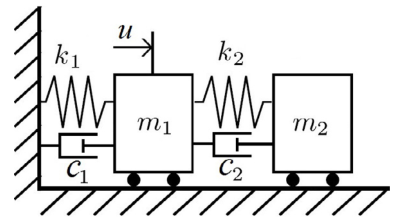

Fig. 1. Mass–spring–damper system, where the disturbance force u is the unknown input.

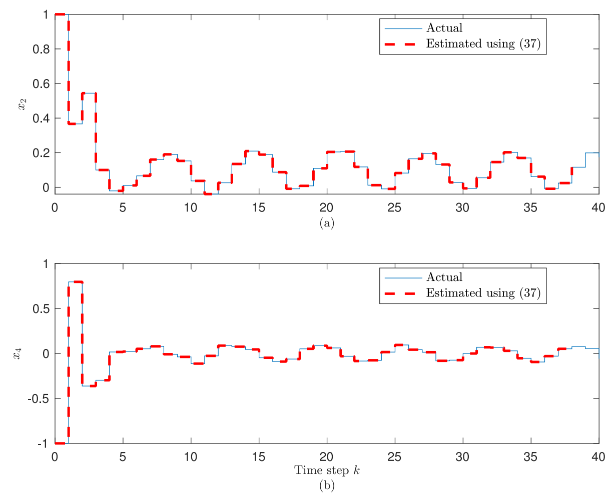

Fig. 2. Application of Theorem 6 to Example 3. For all $0 \leq k \leq 40 - \mu = 38$, the estimated state is equal to the actual state despite the presence of the unknown input, which confirms (37).

the measurements are the position x_1 and velocity x_3 of m_1. The system (A, B, C, D) has no invariant zeros, and thus we use Theorem 6 to estimate the position and velocity of m_2. Furthermore, $\mu = 2$. Let the unknown initial condition be $x(0) = [-6\ 1\ 4\ -1]^{\mathrm{T}}$, and let the unknown input be $u(k) = 1 + w(k) + \sin(kT_{\mathrm{s}})$, where w is zero-mean Gaussian white noise with variance 0.1. Furthermore, let the available measurement be $[y^{\mathrm{T}}(0)\ \cdots\ y^{\mathrm{T}}(40)]^{\mathrm{T}}$.

To apply (37), we choose $r = \mu = 2$. Fig. 2 shows that, for all $0 \leq k \leq 40 - \mu = 38$, the estimated state is equal to the actual state, which confirms (37). $\diamond$

6. Deadbeat input reconstruction for systems without invariant zeros

The following result shows that, if (A, B, C, D) has no invariant zeros and η is finite, then deadbeat input reconstruction is possible with a delay of η steps, whether or not $x(0)$ is known. The proof depends on Propositions 4, 5, Lemmas B and D.

Theorem 7. *Assume that (A, B, C, D) has no invariant zeros and η is finite. Then, for all $k \geq 0$ and $r \geq \max\{\mu, \eta\}$,*

$$\begin{bmatrix} x(k) \\ u(k) \\ \vdots \\ u(k+r-\eta) \end{bmatrix} = \begin{bmatrix} I_{n+(r-\eta+1)m} & 0_{[n+(r-\eta+1)m]\times\eta m} \end{bmatrix} \cdot \Psi_r^+ \begin{bmatrix} y(k) \\ y(k+1) \\ \vdots \\ y(k+r) \end{bmatrix}. \tag{43}$$

Proof. Let $k \geq 0$. Since, η exists, it follows from Proposition 4 that, for all $r \geq \eta$,

$$\mathcal{R}(N_{r,\eta}) \cap \mathcal{R}(Q_{r,\eta}) = \{0\}. \quad (44)$$

Noting $M_r = [N_{r,\eta}\ Q_{r,\eta}]$ and using (44) and Lemma B, it follows that, for all $r \geq \eta$,

$$\operatorname{rank} M_r = \operatorname{rank} N_{r,\eta} + \operatorname{rank} Q_{r,\eta}. \quad (45)$$

Since (A, B, C, D) has no invariant zeros, it follows from Proposition 5 that, for all $r \geq \mu$,

$$\mathcal{R}(\Gamma_r) \cap \mathcal{R}(M_r) = \{0\}. \quad (46)$$

Noting $\Psi_r = [\Gamma_r\ M_r]$ and using (46) and Lemma B, it follows that, for all $r \geq \mu$,

$$\operatorname{rank} \Psi_r = \operatorname{rank} \Gamma_r + \operatorname{rank} M_r. \quad (47)$$

Substituting (45) into (47) yields, for all $r \geq \max\{\mu, \eta\}$,

$$\operatorname{rank} \Psi_r = \operatorname{rank} \Gamma_r + \operatorname{rank} N_{r,\eta} + \operatorname{rank} Q_{r,\eta}. \quad (48)$$

Next, noting $\Psi_r = [\Gamma_r\ N_{r,\eta}\ Q_{r,\eta}]$ and using Lemma B, it follows that, for all $r \geq 0$,

$$\begin{aligned}\operatorname{rank} \Psi_r &= \operatorname{rank}\, [\Gamma_r\ N_{r,\eta}] + \operatorname{rank} Q_{r,\eta} \\ &\quad - \dim[\mathcal{R}([\Gamma_r\ N_{r,\eta}]) \cap \mathcal{R}(Q_{r,\eta})] \\ &= \operatorname{rank} \Gamma_r + \operatorname{rank} N_{r,\eta} + \operatorname{rank} Q_{r,\eta} \\ &\quad - \dim[\mathcal{R}(\Gamma_r) \cap \mathcal{R}(N_{r,\eta})] \\ &\quad - \dim[\mathcal{R}([\Gamma_r\ N_{r,\eta}]) \cap \mathcal{R}(Q_{r,\eta})]. \end{aligned} \quad (49)$$

Subtracting (49) from (48) yields, for all $r \geq \max\{\mu, \eta\}$,

$$\dim[\mathcal{R}(\Gamma_r) \cap \mathcal{R}(N_{r,\eta})] + \dim[\mathcal{R}([\Gamma_r\ N_{r,\eta}]) \cap \mathcal{R}(Q_{r,\eta})] = 0. \quad (50)$$

Since both terms in (50) are nonnegative, it follows that, for all $r \geq \max\{\mu, \eta\}$,

$$\mathcal{R}(\Gamma_r) \cap \mathcal{R}(N_{r,\eta}) = \{0\}, \quad (51)$$

$$\mathcal{R}([\Gamma_r\ N_{r,\eta}]) \cap \mathcal{R}(Q_{r,\eta}) = \{0\}. \quad (52)$$

Using (52), it follows from Lemma D that, for all $r \geq \max\{\mu, \eta\}$,

$$\Psi_r^+ \Psi_r = \begin{bmatrix} [\Gamma_r\ N_{r,\eta}]^+ [\Gamma_r\ N_{r,\eta}] & 0 \\ 0 & Q_{r,\eta}^+ Q_{r,\eta} \end{bmatrix}. \quad (53)$$

Next, Proposition 5 implies that, for all $r \geq \mu$, Γ_r has full column rank. Furthermore, Proposition 4 implies that, for all $r \geq \eta$, $N_{r,\eta}$ has full column rank. Therefore, using (51) it follows that

$$\begin{aligned}\operatorname{rank}\, [\Gamma_r\ N_{r,\eta}] &= \operatorname{rank} \Gamma_r + \operatorname{rank} N_{r,\eta} - \dim[\mathcal{R}(\Gamma_r) \cap \mathcal{R}(N_{r,\eta})] \\ &= \operatorname{rank} \Gamma_r + \operatorname{rank} N_{r,\eta}. \end{aligned} \quad (54)$$

Therefore, for all $r \geq \max\{\mu, \eta\}$, $[\Gamma_r\ N_{r,\eta}]$ has full column rank, and thus $[\Gamma_r\ N_{r,\eta}]^+$ is a left inverse of $[\Gamma_r\ N_{r,\eta}]$. Hence, for all $r \geq \max\{\mu, \eta\}$,

$$\Psi_r^+ \Psi_r = \begin{bmatrix} I_{n+(r-\eta+1)m} & 0 \\ 0 & Q_{r,\eta}^+ Q_{r,\eta} \end{bmatrix}. \quad (55)$$

Next, multiplying (55) by $[x^{\mathrm{T}}(k)\ u^{\mathrm{T}}(k)\ \cdots\ u^{\mathrm{T}}(k+r)]^{\mathrm{T}}$ implies that, for all $r \geq \max\{\mu, \eta\}$,

$$\begin{bmatrix} I_{n+(r-\eta+1)m} & 0 \\ 0 & Q_{r,\eta}^+ Q_{r,\eta} \end{bmatrix} \begin{bmatrix} x(k) \\ u(k) \\ \vdots \\ u(k+r) \end{bmatrix} = \Psi_r^+ \begin{bmatrix} y(k) \\ y(k+1) \\ \vdots \\ y(k+r) \end{bmatrix}. \quad (56)$$

Finally, multiplying (56) by

$[I_{n+(r-\eta+1)m}\ 0_{[n+(r-\eta+1)m]\times\eta m}]$

implies that, for all $r \geq \max\{\mu, \eta\}$, (43) holds. □

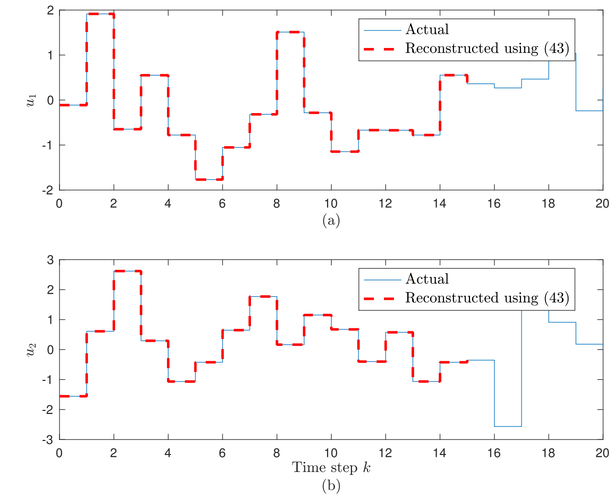

Fig. 3. Application of Theorem 7 to Example 4. For all $0 \leq k \leq r - \eta = 15$, the reconstructed input is equal to the actual input, which confirms (43).

The following example illustrates Theorem 7 for a system with rank-deficient Markov parameters and with no invariant zeros.

Example 4. Consider $G(\mathbf{z}) = C(\mathbf{z}I - A)^{-1}B + D$ given by

$$G(\mathbf{z}) = \frac{1}{\mathbf{z}^5}(H_5 + H_4\mathbf{z} + H_3\mathbf{z}^2 + H_2\mathbf{z}^3 + H_1\mathbf{z}^4), \quad (57)$$

where

$$H_1 = H_2 = H_3 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},\ H_4 = \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 1 & 1 \end{bmatrix},\ H_5 = \begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}. \quad (58)$$

Note that H_0 is zero, (A, B, C, D) has no invariant zeros, $\mu = 4$, and $\eta = 5$. To apply Theorem 7 using (43), we choose $k = 0$ and $r = 20 \geq \max\{\mu, \eta\} = 5$. For all $k \geq 0$, let the unknown input $u(k) = [u_{(1)}(k)\ u_{(2)}(k)]^{\mathrm{T}}$, where $u_{(1)}(k)$ and $u_{(2)}(k)$ are sampled Gaussian white noise with variance 1. Fig. 3 shows that, for all $0 \leq k \leq r - \eta = 15$, the reconstructed input is equal to the actual input, which confirms (43). Furthermore, the reconstructed initial state (not shown in Fig. 3) is equal to $x(0)$. ◇

Example 5. Reconsider Example 3 but now with the objective of reconstructing the unknown input u. Note that $\eta = 1$. To apply (43), we choose $k = 0$ and $r = 40 \geq \max\{\mu, \eta\} = 2$. Fig. 4 shows that, for all $0 \leq k \leq 40 - \eta = 39$, the reconstructed input is equal to the actual input, which confirms (43). ◇

7. Conclusions and future research

Using the generalized inverse of a block-Toeplitz matrix, this paper presented unified algorithms for deadbeat unknown-input state estimation and simultaneous input reconstruction and state estimation for MIMO systems that are d-delay invertible, that is, invertible with a delay of d steps. These algorithms do not assume the existence of a full-column-rank Markov parameter.

The assumption that the system is d-delay invertible is equivalent to the finiteness of the index η, which is the smallest delay d such that the system is d-delay invertible. Numerical examples suggest that the existence of at least one Markov parameter with full column rank implies that η is finite; however, (18) shows that this condition is not necessary.

Extensions for future research include a stochastic treatment of input estimation that accounts for sensor noise as well as disturbances whose presence violates the requirement $m \leq p$.

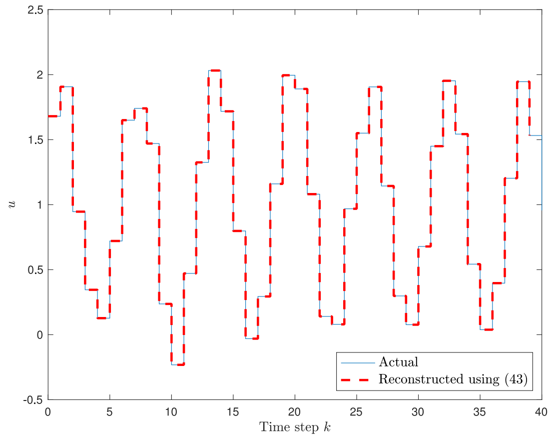

Fig. 4. Application of Theorem 7 to Example 5. For all $0 \le k \le 40 - \eta = 39$, the reconstructed input is equal to the actual input, which confirms (43).

Acknowledgment

The authors thank Götz Trenkler for helpful discussions.

Appendix. Partitioned matrices

The following result is used in the proofs of Lemma C and Proposition 5.

Lemma A. *Let $A \in \mathbb{R}^{n\times m}$, $B \in \mathbb{R}^{l\times m}$, $C \in \mathbb{R}^{n\times p}$, $D \in \mathbb{R}^{l\times p}$, and $E \in \mathbb{R}^{l\times q}$. Assume that A has full column rank, and $\mathcal{R}(A)\cap\mathcal{R}(C) = \{0\}$. Then*

$$\mathcal{R}\left(\begin{bmatrix} A \\ B \end{bmatrix}\right) \cap \mathcal{R}\left(\begin{bmatrix} C & 0 \\ D & E \end{bmatrix}\right) = \{0\}. \tag{59}$$

Proof. Let

$$\begin{bmatrix} x \\ y \end{bmatrix} \in \mathcal{R}\left(\begin{bmatrix} A \\ B \end{bmatrix}\right) \cap \mathcal{R}\left(\begin{bmatrix} C & 0 \\ D & E \end{bmatrix}\right).$$

Therefore, $x \in \mathcal{R}(A)\cap\mathcal{R}(C) = \{0\}$, and thus $x = 0$. Furthermore, there exists $z \in \mathbb{R}^m$ such that $\left[\begin{smallmatrix} 0 \\ y \end{smallmatrix}\right] = \left[\begin{smallmatrix} A \\ B \end{smallmatrix}\right] z$, and thus $Az = 0$ and $y = Bz$. Since A has full column rank, it follows that $z = 0$, and thus $y = 0$. □

Lemma B. *Let $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{n\times l}$. Then,*

$$\operatorname{rank}\,[A\ \ B] = \operatorname{rank} A + \operatorname{rank} B - \dim(\mathcal{R}(A)\cap\mathcal{R}(B)).$$

Proof. See Fact 3.14.15 in Bernstein (2018, p. 322). □

The following result is used in the proof of Proposition 4.

Lemma C. *Let $r \ge 2$, for all $i \in \{0, 1, \ldots, r\}$, let $H_i \in \mathbb{R}^{n\times m}$, and define the block-Toeplitz matrix*

$$T_r = \begin{bmatrix} H_0 & 0 & \cdots & 0 & 0 \\ H_1 & H_0 & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ H_{r-1} & H_{r-2} & \ddots & H_0 & 0 \\ H_r & H_{r-1} & \cdots & H_1 & H_0 \end{bmatrix} = \begin{bmatrix} C_r & C_{r-1} & \cdots & C_1 & C_0 \end{bmatrix}, \tag{60}$$

where, for all $i \in \{0, \ldots, r\}$, C_i denotes the $(i+1)$th block column of T_r labeled from right to left. Furthermore, let $l \in \{1, \ldots, r\}$, and assume that C_l has full column rank and

$$\mathcal{R}(C_l) \cap \mathcal{R}([C_{l-1}\ \cdots\ C_0]) = \{0\}. \tag{61}$$

Then, $[C_r\ \cdots\ C_l]$ has full column rank, and

$$\mathcal{R}([C_r\ \cdots\ C_l]) \cap \mathcal{R}([C_{l-1}\ \cdots\ C_0]) = \{0\}. \tag{62}$$

Proof. Noting

$$T_l = \left[\begin{array}{c|c} C_l & \begin{matrix} 0 \\ T_{l-1} \end{matrix} \end{array}\right] = \left[\begin{array}{ccc|c} & T_{l-1} & & 0 \\ \hline H_l & \cdots & H_1 & H_0 \end{array}\right] \tag{63}$$

and using $\operatorname{rank} C_l = m$, (61), and Lemma B, it follows that

$$\begin{aligned} \operatorname{rank} T_l &= \operatorname{rank} C_l + \operatorname{rank} \left[\begin{smallmatrix} 0 \\ T_{l-1} \end{smallmatrix}\right] - \dim(\mathcal{R}(C_l)\cap\mathcal{R}(\left[\begin{smallmatrix} 0 \\ T_{l-1} \end{smallmatrix}\right])) \\ &= m + \operatorname{rank} T_{l-1} - \dim(\mathcal{R}(C_l)\cap\mathcal{R}([C_{l-1}\ \cdots\ C_0])) \\ &= m + \operatorname{rank} T_{l-1}. \end{aligned} \tag{64}$$

Similarly, since

$$T_{l+1} = \left[\begin{array}{c|c} C_{l+1} & \begin{matrix} 0 \\ T_l \end{matrix} \end{array}\right] = \left[\begin{array}{c|ccc|c} C_l & & \begin{matrix} 0 \\ T_{l-1} \end{matrix} & & 0 \\ \hline H_{l+1} & H_l & \cdots & H_1 & H_0 \end{array}\right], \tag{65}$$

it follows from $\operatorname{rank} C_{l+1} = m$, (64), and (65) that

$$\begin{aligned} \operatorname{rank} T_{l+1} &= \operatorname{rank} C_{l+1} + \operatorname{rank} T_l \\ &\qquad - \dim\left(\mathcal{R}(C_{l+1})\cap\mathcal{R}\left(\begin{bmatrix} 0 \\ T_l \end{bmatrix}\right)\right) \\ &= 2m + \operatorname{rank} T_{l-1} \\ &- \dim\left(\mathcal{R}\left(\begin{bmatrix} C_l \\ H_{l+1} \end{bmatrix}\right)\cap\mathcal{R}\left(\left[\begin{array}{ccc|c} & \begin{matrix} 0 \\ T_{l-1} \end{matrix} & & \begin{matrix} 0 \\ 0 \end{matrix} \\ \hline H_l & \cdots & H_1 & H_0 \end{array}\right]\right)\right). \end{aligned} \tag{66}$$

Since $\mathcal{R}(C_l)\cap\mathcal{R}(\left[\begin{smallmatrix} 0 \\ T_{l-1} \end{smallmatrix}\right]) = \{0\}$ and C_l has full column rank, it follows from Lemma A that

$$\mathcal{R}\left(\begin{bmatrix} C_l \\ H_{l+1} \end{bmatrix}\right)\cap\mathcal{R}\left(\left[\begin{array}{ccc|c} & \begin{matrix} 0 \\ T_{l-1} \end{matrix} & & \begin{matrix} 0 \\ 0 \end{matrix} \\ \hline H_l & \cdots & H_1 & H_0 \end{array}\right]\right) = \{0\}. \tag{67}$$

Combining (67) with (66) yields

$$\operatorname{rank} T_{l+1} = 2m + \operatorname{rank} T_{l-1}. \tag{68}$$

Similarly, since

$$T_{l+2} = \left[\begin{array}{c|c} C_{l+2} & \begin{matrix} 0 \\ T_{l+1} \end{matrix} \end{array}\right] = \left[\begin{array}{c|ccc|c} C_{l+1} & & \begin{matrix} 0 \\ T_l \end{matrix} & & 0 \\ \hline H_{l+2} & H_{l+1} & \cdots & H_1 & H_0 \end{array}\right], \tag{69}$$

it follows from $\operatorname{rank} C_{l+2} = m$ and (68) that

$$\begin{aligned} \operatorname{rank} T_{l+2} &= \operatorname{rank} C_{l+2} + \operatorname{rank} T_{l+1} \\ &\qquad - \dim\left(\mathcal{R}(C_{l+2})\cap\mathcal{R}\left(\begin{bmatrix} 0 \\ T_{l+1} \end{bmatrix}\right)\right) \\ &= 3m + \operatorname{rank} T_{l-1} \\ &- \dim\left(\mathcal{R}\left(\begin{bmatrix} C_{l+1} \\ H_{l+2} \end{bmatrix}\right)\cap\mathcal{R}\left(\left[\begin{array}{ccc|c} & \begin{matrix} 0 \\ T_l \end{matrix} & & \begin{matrix} 0 \\ 0 \end{matrix} \\ \hline H_{l+1} & \cdots & H_1 & H_0 \end{array}\right]\right)\right). \end{aligned} \tag{70}$$

It follows from (65) and (67) that $\mathcal{R}(C_{l+1}) \cap \mathcal{R}(\left[\begin{smallmatrix} 0 \\ T_l \end{smallmatrix}\right]) = \{0\}$, and, since C_{l+1} has full column rank, it follows from Lemma A that

$$\mathcal{R}\left(\begin{bmatrix} C_{l+1} \\ H_l \end{bmatrix}\right) \cap \mathcal{R}\left(\left[\begin{array}{c|c} 0 & 0 \\ T_l & 0 \\ \hline H_{l+1} \quad \cdots \quad H_1 & H_0 \end{array}\right]\right) = \{0\}. \tag{71}$$

Combining (71) with (70) yields

$$\operatorname{rank} T_{l+2} = 3m + \operatorname{rank} T_{l-1}. \tag{72}$$

By similar arguments, it follows that, for all $k \geq 1$,

$$\operatorname{rank} T_{l+k} = (k+1)m + \operatorname{rank} T_{l-1}, \tag{73}$$

which, with $k = r - l$, yields

$$\operatorname{rank} T_r = (r - l + 1)m + \operatorname{rank} T_{l-1}. \tag{74}$$

Noting

$$T_r = \left[\begin{array}{c|c} C_r \ \cdots \ C_l & \begin{matrix} 0 \\ T_{l-1} \end{matrix} \end{array}\right] = \left[\begin{array}{ccc|ccc} C_r & \cdots & C_l & C_{l-1} & \cdots & C_0 \end{array}\right], \tag{75}$$

it follows that

$$\operatorname{rank} T_r = \operatorname{rank}\,[C_r \ \cdots \ C_l] + \operatorname{rank} T_{l-1} \\ - \dim\left(\mathcal{R}([C_r \ \cdots \ C_l]) \cap \mathcal{R}([C_{l-1} \ \cdots \ C_0])\right). \tag{76}$$

Combining (76) with (74) yields

$$0 \leq \dim\left(\mathcal{R}([C_r \ \cdots \ C_l]) \cap \mathcal{R}([C_{l-1} \ \cdots \ C_0])\right) = \\ \operatorname{rank}\,[C_r \ \cdots \ C_l] - (r - l + 1)m \leq 0,$$

which implies that $[C_r \ \cdots \ C_l]$ has full column rank and (62) holds. □

The following result is used in the proofs of Theorem 6 and Theorem 7.

Lemma D. *Let $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{n\times l}$, define $C \triangleq (I - AA^+)B$ and $D \triangleq (I - BB^+)A$, and assume that $\mathcal{R}(A) \cap \mathcal{R}(B) = \{0\}$. Then, $C^+A = 0$, $D^+B = 0$, $C^+B = B^+B$, $D^+A = A^+A$,*

$$[A\ B]^+ = \begin{bmatrix} D^+ \\ C^+ \end{bmatrix}, \quad [A\ B]^+[A\ B] = \begin{bmatrix} A^+A & 0 \\ 0 & B^+B \end{bmatrix}. \tag{77}$$

Proof. The result follows from Theorem 1, line 6 on page 21, and line 7 on page 22 of Baksalary and Baksalary (2007). □

References

Ansari, A., & Bernstein, D. S. (2018). Input estimation for nonminimum-phase systems with application to acceleration estimation for a maneuvering vehicle. *IEEE Transactions on Control Systems Technology*, (99), 1–12.

Baksalary, J. K., & Baksalary, O. M. (2007). particular formulae for the moore–penrose inverse of a columnwise partitioned matrix. *Linear Algebra and its Applications*, *421*(1), 16–23.

Bernstein, D. S. (2018). *Scalar, vector, and matrix mathematics: Theory, fact, and formulas* (revised and expanded). Princeton: Princeton University Press.

Chakrabarty, A., Ayoub, R., Zak, S. H., & Sundaram, S. (2017). Delayed unknown input observers for discrete-time linear systems with guaranteed performance. *Systems & Control Letters*, *103*, 9–15.

Chen, M.-S., & Chen, C.-C. (2010). Unknown input observer for linear non-minimum phase systems. *Journal of the Franklin Institute*, *347*(2), 577–588.

Corless, M., & Tu, J. (1998). State and input estimation for a class of uncertain systems. *Automatica*, *34*(6), 757–764.

Crassidis, J. L., & Junkins, J. L. (2011). *Optimal estimation of dynamic systems*. CRC Press.

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

Fang, H., de Callafon, R. A., & Cortes, J. (2013). Simultaneous input and state estimation for nonlinear systems with applications to flow field estimation. *Automatica*, *49*(9), 2805–2812.

Fang, H., Shi, Y., & Yi, J. (2011). On Stable Simultaneous Input and State Estimation for Discrete-Time Linear Systems. *International Journal of Adaptive Control and Signal Processing*, *25*(8), 671–686.

Floquet, T., & Barbot, J. P. (2006). State and Unknown Input Estimation for Linear Discrete-Time Systems. *Automatica*, *42*, 1883–1889.

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

Glover, J. D. (1969). The linear estimation of completely unknown signals. *IEEE Transactions on Automatic Control*, *14*(6), 766–767.

Ho, C.-C., & Ma, C.-K. (2007). Active vibration control of structural systems by a combination of the linear quadratic Gaussian and input estimation approaches. *Journal of Sound and Vibration*, *301*(3), 429–449.

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

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

Hsieh, C.-S. (2000). Robust two-stage kalman filters for systems with unknown inputs. *IEEE Transactions on Automatic Control*, *45*(12), 2374–2378.

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

Hsieh, C.-S. (2017). Unbiased minimum-variance input and state estimation for systems with unknown inputs: A system reformation approach. *Automatica*, *84*, 236–240.

Khaloozadeh, H., & Karsaz, A. (2009). Modified input estimation technique for tracking manoeuvring targets. *Radar, Sonar & Navigation, IET*, *3*(1), 30–41.

Kim, P. S. (2010). An alternative FIR filter for state estimation in discrete-time systems. *Digital Signal Processing* , *20*(3), 935–943.

Kirtikar, S., Palanthandalam-Madapusi, H., Zattoni, E., & 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.

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

Lewis, F. L., Xie, L., & Popa, D. (2007). *Optimal and robust estimation: With an introduction to stochastic control theory*. CRC Press.

Lu, P., van Kampen, E.-J., de Visser, C. C., & Chu, Q. (2016). Framework for state and unknown input estimation of linear time-varying systems. *Automatica*, *73*, 145–154.

Massey, J. L., & Sain, M. (1968). Inverses of linear sequential circuits. *IEEE Transactions on Computers*, *17*, 330–337.

Orjuela, R., Marx, B., Ragot, J., & Maquin, D. (2009). On the simultaneous state and unknown input estimation of complex systems via a multiple model strategy. *IET Control Theory & Applications*, *3*(7), 877–890.

Palanthandalam-Madapusi, H., & Bernstein, D. S. (2009). A subspace algorithm for simultaneous identification and input reconstruction. *International Journal of Adaptive Control and Signal Processing*, *23*, 1053–1069.

Park, S. H., Kim, P. S., Kwon, O. K., & Kwon, W. H. (2000). Estimation and detection of unknown inputs using optimal FIR filter. *Automatica*, *36*(10), 1481–1488.

Rajamani, R., Wang, Y., Nelson, G. D., Madson, R., & Zemouche, A. (2017). Observers with dual spatially separated sensors for enhanced estimation: Industrial, automotive, and biomedical applications. *IEEE Control Systems Magazine*, *37*(3), 42–58.

Sain, M., & Massey, J. L. (1969). Invertibility of linear time-invariant dynamical systems. *IEEE Transactions on Automatic Control*, *14*, 141–149.

Sanchez, J., & Benaroya, H. (2014). Review of force reconstruction techniques. *Journal of Sound and Vibration*, *333*(14), 2999–3018.

Shmaliy, Y. S., Zhao, S., & Ahn, C. K. (2017). Unbiased finite impulse response filtering. *IEEE Control Systems Magazine*, *37*(5), 70–89.

Silverman, L. (1969). Inversion of multivariable linear systems. *IEEE Transactions on Automatic Control*, *14*, 270–276.

Simon, D. (2006). *Optimal state estimation: Kalman, H_∞, and nonlinear approaches*. John Wiley & Sons.

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

Valcher, M. E. (1999). State observers for discrete-time linear systems with unknown inputs. *IEEE Transactions on Automatic Control*, *44*(2), 397–401.

Willsky, A. (1974). On the invertibility of linear systems. *IEEE Transactions on Automatic Control*, *19*, 272–274.

Xiong, Y., & Saif, M. (2003). Unknown disturbance inputs estimation based on a state functional observer design. *Automatica*, *39*, 1389–1398.

Yang, J., Zhu, F., & Sun, X. (2013). State estimation and simultaneous unknown input and measurement noise reconstruction based on associated observers. *International Journal of Adaptive Control and Signal Processing*, *27*(10), 846–858.

Yong, S. Z., Zhu, M., & Frazzoli, E. (2015). Simultaneous input and state estimation with a delay. In *Decision and control (CDC), 2015 IEEE 54th annual conference on* (pp. 468–475). IEEE.

Yong, S. Z., Zhu, M., & Frazzoli, E. (2016). A unified filter for simultaneous input and state estimation of linear discrete-time stochastic systems. *Automatica*, *63*, 321–329.

Ahmad Ansari received the B.S. degree in aerospace engineering from The Institute of Space Technology, the M.S. degree in electrical engineering from The University of Michigan, and the Ph.D. degree in aerospace engineering from The University of Michigan. His research has focused on flight dynamics and controls of aircraft, as well as input estimation for sensor fault diagnostics.

Dennis S. Bernstein received the Sc.B. degree in applied mathematics from Brown University and the Ph.D. degree in control engineering from the University of Michigan in Ann Arbor, Michigan, where he is currently a professor in the Aerospace Engineering Department. His interests are in estimation and control for aerospace applications.