INTERNATIONAL JOURNAL OF ADAPTIVE CONTROL AND SIGNAL PROCESSING
Int. J. Adapt. Control Signal Process. 2011; **25**:671–686
Published online 9 February 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/acs.1230

On stable simultaneous input and state estimation for discrete-time linear systems

Huazhen Fang[1], Yang Shi[2,*,†] and Jingang Yi[3]

[1]*Department of Mechanical and Aerospace Engineering, University of California, San Diego, CA 92092, U.S.A.*
[2]*Department of Mechanical Engineering, University of Victoria, Victoria, BC, Canada V8W 3P6*
[3]*Department of Mechanical and Aerospace Engineering, Rutgers University, Piscataway, NJ 08854-8058, U.S.A.*

SUMMARY

This work is devoted to solving simultaneous input and state estimation (SISE) problem for discrete-time linear systems. Our aim is to develop stable SISE algorithms. By applying the minimum variance unbiased estimation technique, we derive two SISE algorithms in the presence or absence of direct feedthrough, respectively. Riccati-like equations are formulated and presented to analyze the stability conditions of the proposed algorithms. Simulation examples are provided to further illustrate the effectiveness of the proposed algorithms and support the theoretical findings. Copyright © 2011 John Wiley & Sons, Ltd.

Received 15 October 2009; Revised 8 December 2010; Accepted 10 December 2010

KEY WORDS: unknown input estimation; state estimation; unbiased estimation; minimum variance; convergence property

1. INTRODUCTION

The problem of simultaneous input and state estimation (SISE) for dynamic systems has a wide range of applications, such as fault detection and diagnosis [1], maneuvering target tracking [2], geophysics and environmentology [3], etc. In these applications, inputs and state variables are often unmeasurable or inaccessible. SISE can also find applications in networked control systems with unknown input package delays and even losses [4]. Because of its practical significance, SISE and related problems have received considerable attention during the past several decades.

State estimation under unknown inputs is directly related with the SISE problem. In [5], an unbiased minimum variance linear state filter is developed and the state estimation is designed independently with the unknown inputs. The design in [5] is extended to a more general filter structure in [6] and the convergence conditions are given for linear time-invariant systems. The optimization in [5] and [6] has been conducted by employing a linear recursive filter and minimum variance estimation. Moreover, the global optimality of the design of the recursive filters in [5] and [6] is presented in [7] and it is shown that the optimal solution over the class of all linear unbiased estimates can be written in the form of a linear recursive filter. An optimal filter design with stability conditions is proposed in [8] for systems with direct feedthrough. In addition to linear minimum variance estimation, some other methods are used for state estimation with unknown inputs. For instance, in [9–11] and the references therein, full-order and reduced-order state observers with unknown inputs are developed by using matrix manipulations. In [12], a sliding mode observer is

*Correspondence to: Yang Shi, Department of Mechanical Engineering, University of Victoria, Victoria, BC, Canada V8W 3P6.
†E-mail: yshi@uvic.ca

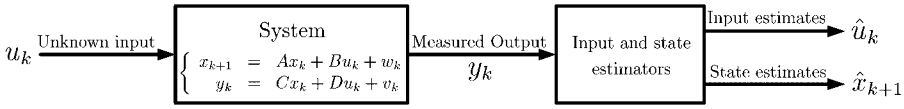

Figure 1. Block diagram of simultaneous input and state estimation.

proposed and the observer convergence is proven. In many applications, such as fault detection and diagnosis, it is desirable to estimate the unknown inputs of a dynamic system. There exist numerous works in this area, for example, see [13–16] and the references therein.

The above-discussed work focus on the estimation of either unknown inputs or system states, but not at the same time. It is much more challenging to simultaneously estimate both inputs and state variables because typically they are inherently interconnected and coupled. In [17], a two-stage Kalman filter and an input filtering technique are combined to achieve joint input and state estimation. An asymptotic input and state estimation scheme is developed in [18] for a class of uncertain systems. By employing a linear matrix inequality (LMI)-based technique, an SISE approach is designed in [19] for systems with a linear part plus a Lipschitz nonlinear term. Most recently, in [20, 21], a set of multi-step recursive filters are proposed to jointly estimate inputs and states by minimizing the error variance for discrete-time linear systems without and with the direct feedthrough, respectively. However, due to the complex structure, the stability analysis of the algorithms proposed in [20, 21] is not given.

In this paper, we develop algorithms to simultaneously estimate the inputs and states for discrete-time linear systems with and without direct feedthrough. The proposed algorithms are built upon the linear minimum variance unbiased estimation. Unlike the other existing approaches such as those in [20, 21], the proposed SISE algorithms have stability properties proven explicitly. Numerical simulation examples are illustrated to show the effectiveness of the theoretical findings. We also compare the performance of the proposed SISE algorithms with other reported algorithms. The results show comparable performance among these algorithms but the proposed ones in this paper have guaranteed estimation stability.

The remainder of the paper is organized as follows. In Section 2, we briefly formulate the problem. Section 3 presents the development of the SISE algorithm for systems with direct feedthrough, and the stability analysis of the algorithm is presented in Section 4. Section 5 extends the results to the systems without direct feedthrough. Section 6 gives illustrative examples to demonstrate the effectiveness of the proposed algorithms and performance comparison with existing SISE algorithms. Finally, we conclude the paper in Section 7.

2. PROBLEM FORMULATION

The notation used throughout the paper is as follows. Small case letters denote vectors and capital letters denote matrices. For matrix X, we use X^{T} and X^{-1} to indicate the transpose and inverse of X, respectively. For random variables, 'E' indicates expectation. We use $\det(X)$ to denote the determinant of X. For symmetric matrix, $X>0$ or $X\geqslant 0$ indicates that X is positive definite or nonnegative definite, respectively, and $X\geqslant Y$ indicates $X-Y\geqslant 0$. We use $\|\cdot\|_2$ and $\|\cdot\|_\infty$ to denote the 2-norm and ∞-norm, respectively.

We consider a discrete-time linear system with direct feedthrough as shown in Figure 1

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

where $x_k \in \mathbb{R}^n$ denotes the system state variable at time instant k, $u_k \in \mathbb{R}^m$ is the unknown input and $y_k \in \mathbb{R}^p$ the system measurement. A, B, C and D are known system matrices with compatible dimensions. The process noise w_k and measurement noise v_k are assumed to be mutually

DOI: 10.1002/acs

uncorrelated zero-mean white noises with known covariances, namely

$$\mathrm{E}\{w_k w_l^{\mathrm{T}}\} = R_w \delta_{k-l}, \quad \mathrm{E}\{v_k v_l^{\mathrm{T}}\} = R_v \delta_{k-l}, \quad \mathrm{E}\{w_k v_l^{\mathrm{T}}\} = 0,$$

where δ_k is the Kronecker delta function, $R_w > 0$ and $R_v > 0$ are variances of $\{w_k\}$ and $\{v_k\}$, respectively. Throughout the paper, it is assumed that D is of full column rank. As will be shown in Lemma 1, this is necessary to ensure unbiased estimation.

From the observer theory for deterministic linear systems [22], we design the input and state estimators, respectively, as follows:

$$\hat{u}_k = H_k(y_k - C\hat{x}_k), \tag{2}$$

$$\hat{x}_{k+1} = A\hat{x}_k + B\hat{u}_k + L_k(y_k - C\hat{x}_k - D\hat{u}_k), \tag{3}$$

where $\hat{u}_k$ represents the input estimate and $\hat{x}_k$ the state estimate. H_k and L_k are the gain matrices that will be determined later. As indicated in (1), y_k is the first measurement that contains information of u_k, so $\hat{u}_k$ can be reconstructed from the residual between y_k and $C\hat{x}_k$. In addition, the fact that x_{k+1} depends on x_k and u_k motivates $\hat{x}_{k+1}$ to be recovered by $\hat{x}_k$ and $\hat{u}_k$ with a correcting term.

The covariance matrices of the estimation errors are defined as

$$P_k^u = \mathrm{E}\{\tilde{u}_k \tilde{u}_k^{\mathrm{T}}\}, \quad P_k^{ux} = \mathrm{E}\{\tilde{u}_k \tilde{x}_k^{\mathrm{T}}\}, \quad P_{k+1}^x = \mathrm{E}\{\tilde{x}_{k+1} \tilde{x}_{k+1}^{\mathrm{T}}\}, \tag{4}$$

where $\tilde{u}_k = u_k - \hat{u}_k$ and $\tilde{x}_k = x_k - \hat{x}_k$. It is straightforward to see that matrices P_k^u and P_{k+1}^x are symmetric positive definite (SPD).

The objectives of this paper are summarized as follows:

P1. To design the input and state estimators, that is, determining estimator gains H_k and L_k in (2) and (3).

P2. To analyze the stability properties of the proposed estimation algorithm.

P3. To extend the results obtained above to discrete-time linear systems without direct feedthrough.

3. THE ALGORITHM DESCRIPTION

This section focuses on the development of the SISE algorithm. We start with the minimum variance unbiased design, that is, minimizing P_k^u to derive L_k , and minimizing P_{k+1}^x to obtain H_k. Then we present a numerically stable algorithm to compute the design.

3.1. Preliminaries

To study the optimality property of the proposed SISE algorithm, it is important to ensure that the estimates are unbiased. We have the following results.

Lemma 1

For unbiased input and state estimates given in (2) and (3), matrix D of system (1) must be of full column rank, and the following initial condition must be satisfied:

$$\hat{x}_0 = \mathrm{E}(x_0). \tag{5}$$

Proof

Substituting the state equations (1) into (2) and (3), respectively, we obtain

$$\tilde{u}_k = -H_k(C\tilde{x}_k + v_k) + (I - H_k D)u_k, \tag{6}$$

$$\tilde{x}_{k+1} = (A - L_k C)\tilde{x}_k + (B - L_k D)\tilde{u}_k - L_k v_k + w_k, \tag{7}$$

where I is the identity matrix. Recursively applying the above dynamics until $k = 0$, it can be seen that the estimates are unbiased, namely, $\mathrm{E}(\tilde{u}_k) = 0$ and $\mathrm{E}(\tilde{x}_k) = 0$, if both (5) and the following input unbiasedness constraint are satisfied

$$H_k D = I. \tag{8}$$

From constraint (8), D should be of full column rank. Proof of Lemma 1 is completed. □

DOI: 10.1002/acs

Under the above unbiasedness condition in Lemma 1, (6) is then simplified as

$$\tilde{u}_k = -H_k(C\tilde{x}_k + v_k). \tag{9}$$

3.2. Input estimation

The following theorem provides the design of optimal H_k^* that leads to minimum variance of input estimation. That is, with the designed H_k^*, we will have the lowest possible P_k^u.

Theorem 1
Assume that the input estimation given in (2) is unbiased for system (1). If the optimal gain matrix H_k^* is

$$H_k^* = (D^{\mathrm{T}} Q_k^{-1} D)^{-1} D^{\mathrm{T}} Q_k^{-1}, \tag{10}$$

where $Q_k = C P_k^x C^{\mathrm{T}} + R_v$, then the covariance of input estimation error P_k^u is minimized.

Proof
Rewrite P_k^u by plugging (9) into (4) as follows:

$$P_k^u = H_k Q_k H_k^{\mathrm{T}}. \tag{11}$$

It is desirable to design H_k such that P_k^u is minimum under the constraint (8). Consider using the Lagrange multipliers with the constraint (8). Let us introduce the Lagrange multiplier Λ, which is a symmetric matrix of appropriate dimension. The Lagrange function $P_k^u(\Lambda)$ is defined as

$$\begin{aligned} P_k^u(\Lambda) &= H_k Q_k H_k^{\mathrm{T}} + (1 - H_k D)\Lambda + \Lambda^{\mathrm{T}}(1 - D^{\mathrm{T}} H_k^{\mathrm{T}}) \\ &= (H_k - \Lambda^{\mathrm{T}} D^{\mathrm{T}} Q_k^{-1}) Q_k (H_k - \Lambda^{\mathrm{T}} D^{\mathrm{T}} Q_k^{-1})^{\mathrm{T}} \\ &\quad -[\Lambda - (D^{\mathrm{T}} Q_k^{-1} D)^{-1}]^{\mathrm{T}} (D^{\mathrm{T}} Q_k^{-1} D)[\Lambda - (D^{\mathrm{T}} Q_k^{-1} D)^{-1}] + (D^{\mathrm{T}} Q_k^{-1} D)^{-1}. \end{aligned} \tag{12}$$

To achieve minimum $P_k^u(\Lambda)$, which is SPD, the first term on the right-hand side should make zero contributions (not negative contributions because Q_k is SPD). This occurs only if

$$H_k = \Lambda^{\mathrm{T}} D^{\mathrm{T}} Q_k^{-1}.$$

Combining (12) with (8) to solve for Λ, we have

$$\Lambda = (D^{\mathrm{T}} Q_k^{-1} D)^{-1}.$$

Then the optimal H_k^* is obtained as in (10), and the minimum P_k^u is given by

$$P_k^u = H_k^* Q_k H_k^{*\mathrm{T}} = (D^{\mathrm{T}} Q_k^{-1} D)^{-1}. \tag{13}$$

This proves the theorem. □

Theorem 1 achieves minimization of P_k^u by determining the optimal H_k^*. That is, for any $H_k \neq H_k^*$ satisfying the constraint (8) and with (11) and (13), we have

$$H_k Q_k H_k^{\mathrm{T}} \geqslant (D^{\mathrm{T}} Q_k^{-1} D)^{-1}.$$

In addition, consider two matrices $X \geqslant Y \geqslant 0$. Then the eigenvalues of X will be less than those of Y if arranged in the same order [23], implying the trace of X is larger than that of Y. Hence, if P_k^u is minimized, then its trace, i.e. the mean square estimation error, is also minimized [6].

With the optimal H_k^* and from (4) and (9), we have

$$P_k^{ux} = \mathrm{E}\{-H_k^*[C\tilde{x}_k + v_k]\tilde{x}_k^{\mathrm{T}}\} = -H_k^* C P_k^x. \tag{14}$$

3.3. State estimation

For state estimation, we similarly aim to design the optimal L_k^* by minimizing the covariance of state estimation error P_{k+1}^x.

Define the following matrices:

$$S_k = M O_k M^{\mathrm{T}}, \quad T_k = M O_k N^{\mathrm{T}} - B H_k^* R_v, \tag{15a}$$

$$U_k = N O_k N^{\mathrm{T}} + R_v - D H_k^* R_v - R_v H_k^{*\mathrm{T}} D^{\mathrm{T}}, \tag{15b}$$

where

$$M = [A \quad B], \quad N = [C \quad D], \quad O_k = \begin{bmatrix} P_k^x & (P_k^{ux})^{\mathrm{T}} \\ P_k^{ux} & P_k^u \end{bmatrix}. \tag{15c}$$

Before proceeding to obtain the optimal gain matrix L_k^*, we show the following property for U_k.

Lemma 2

For any k, U_k in (15b) is singular, that is, $\det U_k = 0$.

Proof

Expanding U_k gives

$$U_k = (I - D H_k^*)(C P_k^x C^{\mathrm{T}} + R_v)(I - D H_k^*)^{\mathrm{T}}$$

from which it follows that

$$\det U_k = \det(I - D H_k^*) \det(C P_k^x C^{\mathrm{T}} + R_v) \det(I - D H_k^*)^{\mathrm{T}}.$$

Because $\det(I - D H_k^*) = \det(I - H_k^* D) = 0$, we have $\det U_k = 0$. □

From (4) and (7), we have

$$P_{k+1}^x = S_k - L_k T_k^{\mathrm{T}} - T_k L_k^{\mathrm{T}} + L_k U_k L_k^{\mathrm{T}} + R_w. \tag{16}$$

Note that, however, the matrix U_k is singular by Lemma 2. Therefore, there exists no unique L_k^* that achieves minimization of P_{k+1}^x. However, we can assume that U_k is nonsingular to see some quick results. Then

$$P_{k+1}^x = (L_k - T_k U_k^{-1}) U_k (L_k - T_k U_k^{-1})^{\mathrm{T}} + S_k - T_k U_k^{-1} T_k^{\mathrm{T}} + R_w.$$

The optimal L_k^* that yields lowest possible P_{k+1}^x is given by

$$L_k^* = T_k U_k^{-1}.$$

To avoid the singularity problem to obtain a unique gain matrix with guaranteed stability of the SISE estimation, we seek for an alternative construction of related matrices as follows:

$$\hat{P}_{k+1}^x = \hat{S}_k - \hat{L}_k \hat{T}_k^{\mathrm{T}} - \hat{T}_k \hat{L}_k^{\mathrm{T}} + \hat{L}_k \hat{U}_k \hat{L}_k^{\mathrm{T}} + R_w, \tag{17a}$$

$$\hat{S}_k = M \hat{O}_k M^{\mathrm{T}}, \quad \hat{T}_k = M \hat{O}_k N^{\mathrm{T}}, \quad \hat{U}_k = N \hat{O}_k N^{\mathrm{T}} + R_v, \quad \hat{H}_k^* = (D^{\mathrm{T}} \hat{Q}_k^{-1} D)^{-1} D^{\mathrm{T}} \hat{Q}_k^{-1}, \tag{17b}$$

$$\hat{Q}_k = C \hat{P}_k^x C^{\mathrm{T}} + R_v, \quad \hat{O}_k = \begin{bmatrix} \hat{P}_k^x & (\hat{P}_k^{ux})^{\mathrm{T}} \\ \hat{P}_k^{ux} & \hat{P}_k^u \end{bmatrix}, \quad \hat{P}_k^{ux} = -\hat{H}_k^* C \hat{P}_k^x, \quad \hat{P}_k^u = \hat{H}_k^* \hat{Q}_k \hat{H}_k^{*\mathrm{T}}. \tag{17c}$$

Note that

$$\hat{P}_{k+1}^x = (\hat{L}_k - \hat{T}_k \hat{U}_k^{-1}) \hat{U}_k (\hat{L}_k - \hat{T}_k \hat{U}_k^{-1})^{\mathrm{T}} + \hat{S}_k - \hat{T}_k \hat{U}_k^{-1} \hat{T}_k^{\mathrm{T}} + R_w.$$

If the near-optimal matrix gain L_k^* is chosen as

$$\hat{L}_k^* = \hat{T}_k \hat{U}_k^{-1} \tag{18}$$

then $\hat{P}_{k+1}^x$ will achieve its minimum

$$\hat{P}_{k+1}^x = \hat{S}_k - \hat{T}_k \hat{U}_k^{-1} \hat{T}_k^{\mathrm{T}} + R_w = \hat{S}_k - \hat{L}_k^* \hat{T}_k^{\mathrm{T}} + R_w. \tag{19}$$

3.4. Algorithm summary

A summary of the proposed procedures is given in Algorithm 1:

Algorithm 1: The SISE algorithm for discrete-time linear systems with direct feedthrough.

Initialize: $\hat{x}_0 = \mathrm{E}(x_0)$, $\hat{P}_0^x = p_0 I$, where p_0 is a large positive value;
for $k = 0$ *to* $N-1$ **do**
  $\hat{Q}_k = C\hat{P}_k^x C^{\mathrm{T}} + R_v$;
  $\hat{H}_k^* = (D^{\mathrm{T}}\hat{Q}_k^{-1}D)^{-1}D^{\mathrm{T}}\hat{Q}_k^{-1}, \quad \hat{u}_k = \hat{H}_k^*(y_k - C\hat{x}_k)$;
  $\hat{P}_k^u = \hat{H}_k^*\hat{Q}_k\hat{H}_k^{*\mathrm{T}}$;
  if $k < N-1$ **then**
    $\hat{P}_k^{ux} = -\hat{H}_k^* C\hat{P}_k^x$;
    $\hat{O}_k = \begin{bmatrix} \hat{P}_k^x & (\hat{P}_k^{ux})^{\mathrm{T}} \\ \hat{P}_k^{ux} & \hat{P}_k^u \end{bmatrix}$;
    $\hat{S}_k = [A\ B]\hat{O}_k[A\ B]^{\mathrm{T}}, \quad \hat{T}_k = [A\ B]\hat{O}_k[C\ D]^{\mathrm{T}}, \quad \hat{U}_k = [C\ D]\hat{O}_k[C\ D]^{\mathrm{T}} + R_v$;
    $\hat{L}_k^* = \hat{T}_k\hat{U}_k^{-1}, \quad \hat{x}_{k+1} = A\hat{x}_k + B\hat{u}_k + \hat{L}_k^*[y_k - C\hat{x}_k - D\hat{u}_k]$;
    $\hat{P}_{k+1}^x = \hat{S}_k - \hat{L}_k^*\hat{T}_k^{\mathrm{T}} + R_w$
  end
end

Remark 1

First, since U_k is singular, we will obtain no numerically feasible algorithm if continuing to follow the standard procedures to use $L_k^* = T_k U_k^{-1}$. As a result, it is imperative to restructure the matrices to ensure numerical feasibility. Second, further study on the singularity of U_k shows that it is caused by the correlation between $\tilde{u}_k$ and v_k. If the correlation is zero, then the singularity problem will not occur. That is actually how we restructure the matrices. To see this, if the correlation is ignored, the matrices in (17a) are equivalent to their original counterparts in (14) and (15c). Finally, we believe that the reconstruction is appropriate because of the guaranteed stability of the algorithm that will be proven in Section 4.

Remark 2

The development of Algorithm 1 is based on unbiased minimum variance estimation (MVUE), which has been commonly used to address state estimation problems, e.g. [5, 6, 20, 21, 24]. However, in order to obtain an SISE algorithm with guaranteed stability, we design Algorithm 1 by taking a tradeoff between optimality (minimum variance) and stability due to an aforementioned structural singularity property of the matrix U_k. Specifically, optimality is sacrificed slightly, but at the same time, the algorithm will gain stability properties that can be proven rigorously. We will note in Section 6 that Algorithm 1, though not being optimal, produces satisfactory and comparable performance with other existing SISE algorithms without guaranteed stability in design.

Remark 3

Consider the case of a linear time-varying discrete-time system

$$\begin{aligned} x_{k+1} &= A_k x_k + B_k u_k + w_k, \\ y_k &= C_k x_k + D_k u_k + v_k, \end{aligned} \tag{20}$$

where A_k, B_k, C_k and D_k may change with time and the properties of the noises w_k and v_k follow the assumptions in Section 2. It is straightforward to extend Algorithm 1 to such systems by replacing A, B, C and D with their time-varying counterparts A_k, B_k, C_k and D_k. The derivation process is analogous to the discussion in this section and we omit it here.

4. STABILITY ANALYSIS

We define the stability of the SISE algorithms in this paper as follows.

Definition 1

The sequences $\{P_k^x\}$ and $\{P_k^u\}$ are said to be stable if there exists $\overline{P^x}<\infty$ and $\overline{P^u}<\infty$ such that

$$\lim_{k\to\infty} P_k^x=\overline{P^x}, \quad \lim_{k\to\infty} P_k^u=\overline{P^u}.$$

Accordingly, the SISE algorithm is stable if both covariance matrices $\{P_k^x\}$ and $\{P_k^u\}$ are stable.

The major idea to prove the stability of the proposed SISE Algorithm 1 is sketched as follows. By inspecting Algorithm 1, it is found that its stability depends on the stability of P_k^x since P_k^u is a monotonically increasing function of P_k^x. Thus the stability analysis of Algorithm 1 is reduced to the analysis of the upper boundedness of P_k^x as $k\to\infty$. To analyze the upper boundedness of P_k^x, it is necessary to analyze the convergence properties of $\hat{P}_{k+1}^x$. We first formulate the matrix $\hat{P}_{k+1}^x$ calculation as a solution of a Riccati-like matrix equation. We then utilize and extend some results in convergence analysis of solutions of the Riccati equation and its variants in [25–28] to obtain the stability conditions for our algorithm. We now present such a development in details.

From (17a) and (19), $\hat{P}_{k+1}^x$ is expressed as

$$\hat{P}_{k+1}^x=M\hat{O}_kM^{\mathrm{T}}-M\hat{O}_kN^{\mathrm{T}}(N\hat{O}_kN^{\mathrm{T}}+R_v)^{-1}N\hat{O}_kM^{\mathrm{T}}+R_w. \tag{21}$$

We define a generalized algebraic Riccati (GAR) function $g(X)$ as follows:

$$g(X)=M\hat{O}(X)M^{\mathrm{T}}-M\hat{O}(X)N^{\mathrm{T}}(N\hat{O}(X)N^{\mathrm{T}}+R_v)^{-1}N\hat{O}(X)M^{\mathrm{T}}+R_w, \tag{22}$$

where $\hat{O}(X)$ has the same structure as $\hat{O}_k$ in (17a) with $\hat{P}_k^x$ replaced by X, that is

$$\hat{O}(X)=\begin{bmatrix} X & (\hat{P}_{uX})^{\mathrm{T}} \\ \hat{P}_{uX} & \hat{P}_u \end{bmatrix}$$

with

$$\hat{P}_{uX}=-[D^{\mathrm{T}}(CXC^{\mathrm{T}}+R_v)^{-1}D]^{-1}D^{\mathrm{T}}(CXC^{\mathrm{T}}+R_v)^{-1}CX,$$
$$\hat{P}_u=[D^{\mathrm{T}}(CXC^{\mathrm{T}}+R_v)^{-1}D]^{-1}.$$

Similar to the SPD requirement of $\hat{P}_k^x$ in $\hat{O}_k$, we assume that X is also SPD in $\hat{O}(X)$. By (22), (21) can be written as an iterative equation

$$\hat{P}_{k+1}^x=g(\hat{P}_k^x). \tag{23}$$

We define a Riccati operator

$$\phi(K,X)=F\hat{O}(X)F^{\mathrm{T}}+V, \tag{24}$$

where $F=M+KN$, $V=KR_vK^{\mathrm{T}}+R_w$ for matrix K with compatible dimensions. The following theorem establishes the convergence properties of the GAR function $g(X)$.

Theorem 2

Assume that

(A1) there exit a $\widetilde{K}$ and a $\widetilde{P}>0$ such that $\widetilde{P}>\phi(\widetilde{K},\widetilde{P})$.

Then, for any $\hat{P}_0\geqslant 0$, the sequence $\{\hat{P}_k^x\}$ generated by the iterative equation $\hat{P}_{k+1}^x=g(\hat{P}_k^x)$ converges, namely

$$\lim_{k\to\infty}\hat{P}_k^x=\overline{\hat{P}},$$

where $\overline{\hat{P}}$ is the fixed point of the GAR function $g(X)$, that is, $\overline{\hat{P}}=g(\overline{\hat{P}})$.

Proofs of Theorems 2 and 3 with related lemmas are given in Appendix A.

DOI: 10.1002/acs

Remark 4
The convergence properties of Algorithm 1 can be quickly concluded from Theorem 2, which presents a sufficient condition (A1) that guarantees the convergence of $\hat{P}_k^x$. If (A1) is satisfied, then $\hat{P}_k^x$ converges to the fixed point of the GAR function $g(X)$, regardless of the initial $\hat{P}_0$. However, with the convergence properties of $\hat{P}_k^x$ proven, we still need to explore the upper boundedness of the true covariances, P_k^x and P_k^u.

Suppose $\hat{P}_{k+1}^x \to \overline{\hat{P}}$ as $k \to \infty$ by Theorem 2. As a consequence, $\hat{L}_k^*$ and $\hat{H}_k^*$ converge to fixed values, denoted as $\overline{\hat{L}}$ and $\overline{\hat{H}}$, respectively. The following theorem can be established.

Theorem 3
Assume that
(A2) there exists $0 < \widetilde{X} < \infty$ such that $\widetilde{X} > \varphi(\widetilde{X})$, where

$$\varphi(\widetilde{X}) = (\overline{\hat{L}}N - M)\hat{O}(\widetilde{X})(\overline{\hat{L}}N - M)^{\mathrm{T}}.$$

Then, for any initial value $P_0 \geqslant 0$, the sequences $\{P_k^x\}$ and $\{P_k^u\}$ are upper bounded as $k \to \infty$.

Remark 5
Based on results in Theorem 2, Theorem 3 shows that when (A2) is satisfied, the true error covariances, P_k^x and P_k^u have upper bounds. In simulation, it is observed that P_k^x is also convergent to a fixed value. Analysis of the exact convergence properties of P_k^x is one of the on-going research directions and we will report new development in the future.

Remark 6
Stability properties of the classical estimation techniques the Kalman filter are often formulated in terms of some stabilizability/detectability conditions. However, considered in this work is the more complicated SISE problem, which leads to the analysis of a non-standard Riccati equation. We found that stabilizability/detectability are not sufficient to guarantee stability in this case. It is also inferred intuitively that such conditions are implicitly included as part of Theorems 2 and 3. Therefore, it is left as an on-going research topic.

Remark 7
It should be noted that the stability analysis developed for linear time-invariant systems here is not considered generalizable to linear time-varying systems, although Remark 3 claims that Algorithm 1 can be extended to the time-varying case.

5. EXTENSION TO DISCRETE-TIME LINEAR SYSTEMS WITHOUT DIRECT FEEDTHROUGH

The previous sections focused on SISE for discrete-time linear systems with direct feedthrough. In this section, we will obtain some analogous results by applying the techniques to systems with no direct feedthrough.

Consider the discrete-time linear system without direct feedthrough

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k, \\ y_k &= Cx_k + v_k. \end{aligned} \tag{25}$$

Obviously, (25) is a special case (1) by setting $D=0$. Since no feedthrough is applied here, the input and state estimators are designed, respectively, as follows:

$$\hat{u}_k = H_k(y_{k+1} - CA\hat{x}_k), \tag{26}$$

$$\hat{x}_{k+1} = A\hat{x}_k + B\hat{u}_k + L_k(y_k - C\hat{x}_k). \tag{27}$$

Following similar lines in Section 3, we can readily derive the SISE algorithm for the system (25), which is summarized in Algorithm 2. To ensure that Algorithm 2 has stability properties proven explicitly, the correlation between $\tilde{u}_k$ and w_k is assumed to be zero in the derivation.

Algorithm 2: The SISE algorithm for discrete-time linear systems without direct feedthrough.

Initialize: $\hat{x}_0 = \mathrm{E}(x_0)$, $P_0^x = p_0 I$, where p_0 is a large positive value ;
for $k = 0$ *to* $N-1$ **do**
  $\hat{Q}_k = C\hat{P}_k^x C^{\mathrm{T}} + CR_wC^{\mathrm{T}} + R_v$;
  $\hat{H}_k^* = (B^{\mathrm{T}}C^{\mathrm{T}}\hat{Q}_k^{-1}CB)^{-1}B^{\mathrm{T}}C^{\mathrm{T}}\hat{Q}_k^{-1}$, $\hat{u}_k = \hat{H}_k^*(y_{k+1} - CA\hat{x}_k)$;
  $\hat{P}_k^u = \hat{H}_k^*\hat{Q}_k\hat{H}_k^{*\mathrm{T}}$;
  if $k < N-1$ **then**
    $\hat{P}_k^{ux} = -\hat{H}_k^* CA\hat{P}_k^x$;
    $\hat{O}_k = \begin{bmatrix} \hat{P}_k^x & (\hat{P}_k^{ux})^{\mathrm{T}} \\ \hat{P}_k^{ux} & \hat{P}_k^u \end{bmatrix}$;
    $\hat{S}_k = [A\ B]\hat{O}_k[A\ B]^{\mathrm{T}}$, $\hat{T}_k = [A\ B]\hat{O}_k[C\ 0]^{\mathrm{T}}$, $\hat{U}_k = [C\ 0]\hat{O}_k[C\ 0]^{\mathrm{T}} + R_v$;
    $\hat{L}_k^* = \hat{T}_k\hat{U}_k^{-1}$, $\hat{x}_{k+1} = A\hat{x}_k + B\hat{u}_k + \hat{L}_k^*[y_k - C\hat{x}_k]$;
    $\hat{P}_{k+1}^x = \hat{S}_k - \hat{L}_k^*\hat{T}_k^{\mathrm{T}} + R_w$
  end
end

Remark 8
To ensure the unbiasedness of Algorithm 2, two assumptions must be established: (1) the matrix CB is of full column rank, and (2) the initial condition $\hat{x}_0 = \mathrm{E}(x_0)$ is satisfied. The rationale for these assumptions is quite similar to the proof of Lemma 1. Note that an equivalent statement of (1) is $rank(CB) = rank(B) = m$, which is common in the literature, e.g. see [5, 6, 21].

Not surprisingly, P_k^x again determines the stability of Algorithm 2. Let us consider $\hat{P}_k^x$ in the first place. The propagation of $\hat{P}_k^x$ is governed by

$$\hat{P}_{k+1}^x = M\hat{O}_kM^{\mathrm{T}} - M\hat{O}_kN^{\mathrm{T}}(N\hat{O}_kN^{\mathrm{T}} + R_v)^{-1}N\hat{O}_kM^{\mathrm{T}} + R_w, \tag{28}$$

where $M = [A\ \ B]$ and $N = [C\ \ 0]$. The GAR function $g(X)$ for Algorithm 2 is defined as

$$g(X) = M\hat{O}(X)M^{\mathrm{T}} - M\hat{O}(X)N^{\mathrm{T}}(N\hat{O}(X)N^{\mathrm{T}} + R_v)^{-1}N\hat{O}(X)M^{\mathrm{T}} + R_w, \tag{29}$$

where $\hat{O}(X)$ evolves from $\hat{O}_k$ by replacing $\hat{P}_k^x$ with X.

It is clear that the GAR functions defined in (29) and (22) have common characteristics. The properties of (22) have been studied extensively in Section 4. In similar fashion, we study (29) for Algorithm 2 to deduce the stability results. Omitting the intermediate proofs, let us present Theorem 4 for the stability analysis of Algorithm 2.

Theorem 4
Assume that
(A3) there exists a $\widetilde{K}$ and a $\widetilde{P} > 0$ such that $\widetilde{P} > \phi(\widetilde{K}, \widetilde{P})$, where

$$\phi(K, X) = (M + KN)\hat{O}(X)(M + KN)^{\mathrm{T}} + KR_vK^{\mathrm{T}} + R_w.$$

Then, for any $P_0 \geqslant 0$, the sequence $\{\hat{P}_k^x\}$ generated by the iterative equation $\hat{P}_{k+1}^x = g(\hat{P}_k^x)$ converges, namely

$$\lim_{k\to\infty} \hat{P}_k^x = \overline{\hat{P}},$$

DOI: 10.1002/acs

where $\overline{\hat{P}}$ is the fixed point of the GAR function $g(X)$, that is, $\overline{\hat{P}} = g(\overline{\hat{P}})$. Define $\overline{\hat{L}} = \lim_{k\to\infty} \hat{L}_k^*$ and $\overline{\hat{H}} = \lim_{k\to\infty} \hat{H}_k^*$. Further, assume that

(A4) there exists $0 < \widetilde{X} < \infty$ such that $\widetilde{X} > \varphi(\widetilde{X})$, where

$$\varphi(X) = (\overline{\hat{L}}N - M)\hat{O}(X)(\overline{\hat{L}}N - M)^{\mathrm{T}}.$$

Then, for any initial value $P_0 \geqslant 0$, the sequences of true covariances, $\{P_k^x\}$ and $\{P_k^u\}$, are upper bounded as $k \to \infty$.

Remark 9

Theorem 4 indicates that, when Algorithm 2 is applied to the system (25), the estimation error covariances P_k^x and P_k^u are upper bounded when (A3) and (A4) are satisfied. Note that (A3) guarantees the convergence of $\hat{P}_k^x$ to a fixed point, and that (A4) further ensures the upper boundedness of P_k^x and P_k^u.

6. NUMERICAL EXAMPLES

In this section, three numerical examples are presented to illustrate the effectiveness of the proposed algorithm.

Example 1

Consider an LTI system described by

$$A = \begin{bmatrix} 0.67 & 0 \\ 0 & 0.53 \end{bmatrix}, \quad B = \begin{bmatrix} 1.00 \\ 0.53 \end{bmatrix}, \quad C = \begin{bmatrix} 0.95 & 0.01 \\ 0.03 & 1.39 \end{bmatrix}, \quad D = \begin{bmatrix} 1.05 \\ 1.20 \end{bmatrix},$$

$$R_w = R_v = \begin{bmatrix} 0.08 & 0 \\ 0 & 0.08 \end{bmatrix}.$$

In this example, the input $\{u_k\}$ is taken as a Gaussian white noise with zero mean and unit variance. Suppose that no information about $\{u_k\}$ is available for use. For initialization, set $\hat{P}_0^x = 10^6 I$. In the simulation, only the output $\{y_k\}$ is used and Algorithm 1 is applied. The observed results are summarized as follows:

(1) The input estimation and state estimation results are shown in Figure 2. It is observed that input estimates are close to the actual input values. Meanwhile, estimation of the two states also exhibits good performance as there are only small differences between the estimates and their true values.

(2) In the simulation, we find that $\hat{P}_k^x$ quickly converges to $\overline{\hat{P}}$:

$$\overline{\hat{P}} = \begin{bmatrix} 0.1133 & 0.0027 \\ 0.0027 & 0.0885 \end{bmatrix}.$$

It is easy to check that $\overline{\hat{P}} = g(\overline{\hat{P}})$ holds true, which indicates that $\overline{\hat{P}}$ is the fixed point of the GAR function $g(X)$. This result confirms the convergence analysis in Section 4. Meanwhile, we note that the true error variance P_k^x also converges to a fixed point $\overline{P}$:

$$\overline{P} = \begin{bmatrix} 0.1321 & 0.0113 \\ 0.0113 & 0.0924 \end{bmatrix}.$$

The trace of P_k^x shown in Figure 3 (solid line) demonstrates the monotonically converging trend. This suggests that, in addition to being upper bounded, P_k^x is likely to have certain convergence property.

(3) To illustrate the performance of the proposed algorithm, we compare the results by Algorithm 1 with the Gillijns–De Moor algorithm in [20]. In Figure 3, it is shown that $\mathrm{tr}(P_k^x)$ yielded by

DOI: 10.1002/acs

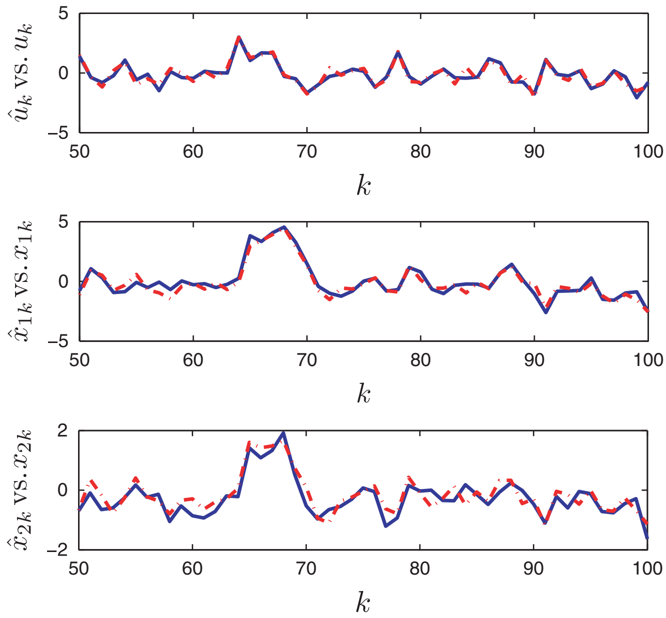

Figure 2. Example 1: Actual values (solid line) vs estimates (dashed line). From top to bottom: input estimation, the first state estimation, and the second state estimation.

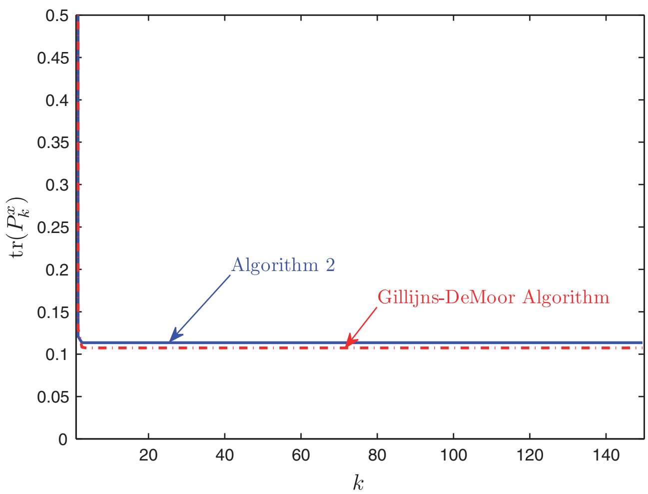

Figure 3. Example 1: Trace of P_k^x vs k.

the Gillijns–De Moor algorithm (red-dashed line) is slightly smaller than that by Algorithm 1. We also compare the input estimation errors $\tilde{u}_k$ and the state estimation errors $\tilde{x}_k$ under both algorithms. The corresponding 2-norm and ∞-norm values of the estimation errors in one implementation are shown in Table I. From Figure 3 and Table I, we conclude that both algorithms produce comparable estimation performance.

Example 2

In this example, we study the performance of Algorithm 2 and its comparison with the Gillijns–De Moor algorithm. Now consider the system

$$A=\begin{bmatrix}0.67 & 0\\ 0 & 0.53\end{bmatrix},\quad B=\begin{bmatrix}1.00\\ 0.53\end{bmatrix},\quad C=\begin{bmatrix}0.95 & 0.01\\ 0.03 & 1.39\end{bmatrix},\quad R_w=R_v=\begin{bmatrix}0.08 & 0\\ 0 & 0.08\end{bmatrix},$$

DOI: 10.1002/acs

Table I. Example 1: Comparisons between the norms of input and state estimation errors by Algorithm 1 and the Gillijns–De Moor algorithm.

		Algorithm 1	Gillijns–De Moor algorithm
Input estimation	$\|\tilde{u}_k\|_2$	4.1234	4.0540
	$\|\tilde{u}_k\|_\infty$	1.5504	1.5504
State estimation	$\|\tilde{x}_{1k}\|_2$	4.5928	4.5624
	$\|\tilde{x}_{1k}\|_\infty$	1.0000	2.0814
	$\|\tilde{x}_{2k}\|_2$	3.7564	3.5121
	$\|\tilde{x}_{2k}\|_\infty$	1.0000	1.0452

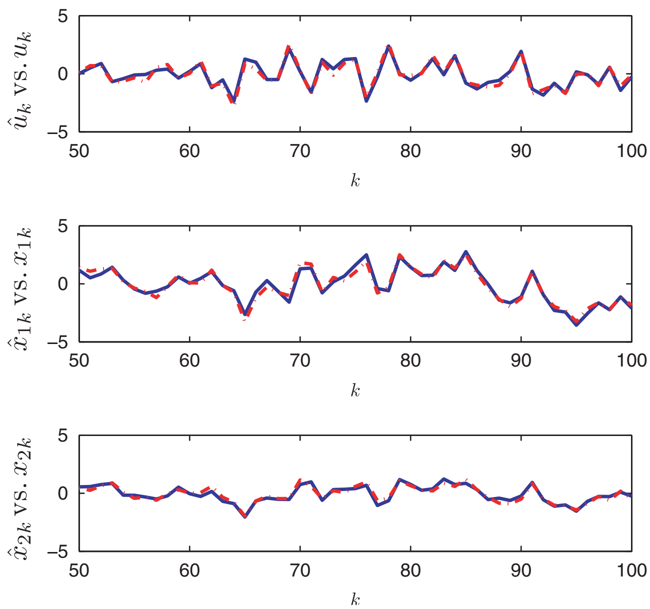

Figure 4. Example 2: Actual values (solid line) vs estimates (dashed line). From top to bottom: input estimation, the first state estimation, and the second state estimation.

where no direct feedthrough is present. The estimation results by Algorithm 2 are given in Figure 4; the trace and norm comparison results between Algorithm 2 and the Gillijns–De Moor algorithm are shown in Figure 5 and Table II, respectively.

In fact, through numerous simulations for the purpose of comparison, we consistently observe that the performances of both Algorithm 1 in this paper and the Gillijns–De Moor algorithm are comparable. However, an attractive advantage when applying Algorithms 1 and 2 is that the stability can be verified and ensured in advance, while in [20, 21] the stability and convergence analysis are not always guaranteed.

7. CONCLUSION

This paper considered SISE for systems with and without direct feedthrough. The challenge of the SISE problem lies in limited information such as the output measurements and system structure and coupling between input and state estimations. Our major contribution is to obtain SISE algorithms with proven stability properties. We first analyzed the optimality conditions for minimum variance. Then we presented optimal design procedures for the SISE problem. The stability of the proposed algorithms were analyzed and guaranteed under the proposed SISE algorithms. We have shown

DOI: 10.1002/acs

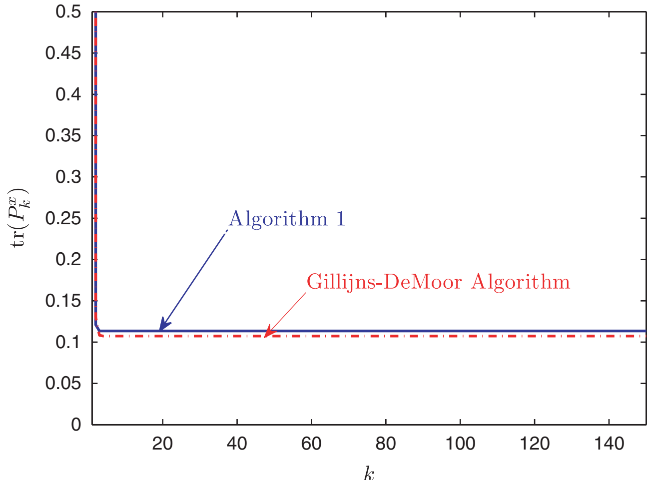

Figure 5. Example 2: Trace of P_k^x vs k.

Table II. Example 2: Comparisons between the norms of input and state estimation errors by Algorithm 2 and the Gillijns–De Moor algorithm.

		Algorithm 2	Gillijns–De Moor algorithm
Input estimation	$\|\tilde{u}_k\|_2$	4.0272	4.0343
	$\|\tilde{u}_k\|_\infty$	0.9031	0.9050
State estimation	$\|\tilde{x}_{1k}\|_2$	3.8715	3.6163
	$\|\tilde{x}_{1k}\|_\infty$	1.0000	1.4749
	$\|\tilde{x}_{2k}\|_2$	3.0387	2.1247
	$\|\tilde{x}_{2k}\|_\infty$	1.0000	0.8100

and demonstrated that the estimation error covariances are upper bounded and thus stable. Several numerical examples were given to illustrate the effectiveness of the proposed algorithm. We also compared the performance of the proposed algorithms with other existing SISE algorithms, achieving comparable performance with guaranteed stability. Unknown inputs often appear in networked control systems, which have been an active research area in the past decade [28–31]. It is worthwhile extending the SISE algorithm to networked dynamic systems subject to random delays and packet losses [32].

APPENDIX A

We need following lemmas to prove Theorems 2 and 3.

Lemma A1
$\hat{O}(X)$ is SPD and monotonically increasing with X.

Proof
The proof is straightforward and thus is omitted. □

Lemma A2
The following facts hold true for functions $g(X)$ and $\phi(K,X)$:

(a) $g(X)=\phi(K_X, X)$ with $K_X=-M\hat{O}(X)N^{\mathrm{T}}(N\hat{O}(X)N^{\mathrm{T}}+R_v)^{-1}$;
(b) $g(X)=\min_K \phi(K,X)$; and
(c) for $0<X\leqslant Y$, $g(X)\leqslant g(Y)$.

Proof
The fact (a) can be readily obtained by substituting $K = K_X$ into (24). For (b), it is straightforward to obtain

$$\phi(K, X) = (M + KN)\hat{O}(X)(M + KN)^{\mathrm{T}} + KR_v K^{\mathrm{T}} + R_w.$$

We can solve for K that minimizes $\phi(K, X)$ by letting

$$\frac{\partial \phi(K, X)}{\partial K} = 0.$$

It is found that $K = K_X$ is the solution of the above equation. For (c), as $\hat{O}(X)$ monotonically increases with X, so does $\phi(X)$. Then for $0 < X \leqslant Y$, we have

$$g(X) = \phi(K_X, X) \leqslant \phi(K_Y, X) \leqslant \phi(K_Y, Y) = g(Y),$$

which proves the fact. □

Lemma A3 (*Sinopoli et al.* [*28*])
Assume that $h(\cdot)$ is a monotonically increasing function. Given two sequences $X_{k+1} = h(X_k)$ and $Y_{k+1} = h(Y_k)$, if $X_1 \geqslant X_0$, then $X_{k+1} \geqslant X_k$; if $X_1 \leqslant X_0$, then $X_{k+1} \leqslant X_k$; and if $X_0 \leqslant Y_0$, then $X_k \leqslant Y_k$, for any $k \geqslant 1$.

Lemma A4
Let $\psi(X) = F\hat{O}(X)F^{\mathrm{T}}$. Assume that there exists $0 < \widetilde{X} < \infty$ such that $\widetilde{X} > \psi(\widetilde{X})$. Consider $X_{k+1} \leqslant \psi(X_k) + \Delta$ with $\Delta \geqslant 0$ and initial value $X_0 \geqslant 0$. Then the sequence X_k is upper bounded.

Proof
It is noted that $\psi(X)$ is linear, positive definite and monotonically increasing with X. For any $0 < X < \infty$, there exist $m \geqslant 0$ and $0 \leqslant r < 1$ such that $X \leqslant m\widetilde{X}$ and $\psi(\widetilde{X}) < r\widetilde{X}$. Then it follows from Lemma A3 that

$$0 \leqslant \psi^k(X) \leqslant mr^{k-1}\psi(\widetilde{X}) \leqslant mr^k\widetilde{X},$$

where

$$\psi^k(X) = \underbrace{\psi(\cdots\psi(\psi(X))\cdots)}_{k} \quad \text{and} \quad \psi^0(X) = X.$$

Thus as $k \to \infty$, $mr^k\widetilde{X} \to 0$, and further, $\psi^k(X) \to 0$. This conclusion indicates that there exists $0 \leqslant m_{X_0}, m_\Delta < \infty$ such that

$$X_k \leqslant \psi^k(X_0) + \sum_{i=0}^{k-1} \psi^i(\Delta) \leqslant \left(m_{X_0} r^k + \sum_{i=0}^{k-1} m_\Delta r^i \right) \widetilde{X} \leqslant \left(m_{X_0} + \frac{m_\Delta}{1 - r} \right) \widetilde{X} < \infty.$$

The lemma is proven. □

Proof of Theorem 2
Note that

$$\widetilde{P} > \phi(\widetilde{K}, \widetilde{P}) = \psi(\widetilde{P}) + \widetilde{K}R_v\widetilde{K}^{\mathrm{T}} + R_w \geqslant \psi(\widetilde{P}).$$

Then from Lemma A2, we obtain

$$\hat{P}^x_{k+1} = g(\hat{P}^x_k) \leqslant \phi(\widetilde{K}, \hat{P}^x_k) = \psi(\hat{P}^x_k) + \widetilde{K}R_v\widetilde{K}^{\mathrm{T}} + R_w = \psi(\hat{P}^x_k) + \hat{\Delta},$$

where $\hat{\Delta} = \widetilde{K}R_v\widetilde{K}^{\mathrm{T}} + R_w \geqslant 0$. Using Lemma A4, we conclude that $\{\hat{P}^x_k\}$ is upper bounded.

Next, we shall show the sequence $\{\hat{P}^x_k\}$ converges to $\overline{\hat{P}}$ from any $\hat{P}_0 \geqslant 0$. First, consider the extreme case when $\hat{P}_0 = 0$. Then

$$0 = \hat{P}_0 \leqslant g(\hat{P}_0) = \hat{P}_1,$$

which implies by Lemmas A2(c) and A3 that

$$\hat{P}_0 \leqslant \hat{P}_1 \leqslant \hat{P}_2 \leqslant \cdots.$$

Since this sequence is proven upper bounded, its upper bound, $\overline{P}$, must be the solution of $\overline{\hat{P}}=g(\overline{\hat{P}})$, namely, the fixed point of function $g(X)$.

We now consider the case when $\hat{P}_0>\overline{P}$. Define

$$\hat{K}=-A\overline{\hat{P}}X(C\overline{\hat{P}}C^{\mathrm{T}}+R_v)^{-1}, \quad \hat{F}=A+\hat{K}C, \quad \hat{\psi}(X)=\hat{F}X\hat{F}.$$

We obtain

$$\overline{\hat{P}}=g(\overline{\hat{P}})=\phi(\hat{K},\overline{\hat{P}})>\hat{\psi}(\overline{\hat{P}}).$$

Thus from Lemma A4, we obtain

$$\lim_{k\to\infty}\hat{\psi}^k(X)=0 \quad \text{for any } X>0. \tag{A1}$$

Since $\hat{P}_0>\overline{\hat{P}}$, it holds that $\hat{P}_1=g(\hat{P}_0)\geqslant g(\overline{\hat{P}})=\overline{\hat{P}}$, which shows $\hat{P}_k^x\geqslant\overline{\hat{P}}$. Let us consider the sequence $\{\hat{P}_k^x-\overline{\hat{P}}\}$. We have

$$0\leqslant(\hat{P}_{k+1}^x-\overline{\hat{P}})=g(\hat{P}_k^x)-g(\overline{\hat{P}})=\phi(K_{\hat{P}_k^x},\hat{P}_k^x)-\phi(\hat{K},\overline{\hat{P}})\leqslant\phi(\hat{K},\hat{P}_k^x)-\phi(\hat{K},\overline{\hat{P}})=\hat{\psi}(\hat{P}_k^x-\overline{\hat{P}}).$$

The above inequality, together with (A1), implies that $\lim_{k\to\infty}\hat{P}_{k+1}^x=\overline{\hat{P}}$.

Last, if $0\leqslant\hat{P}_0\leqslant\overline{\hat{P}}$, then $g^k(0)<\hat{P}_{k+1}^x=g^k(\hat{P}_0)<g(\overline{\hat{P}})=\overline{\hat{P}}$ and, thus, if $k\to\infty$, we have $\lim_{k\to\infty}\hat{P}_{k+1}^x=\overline{\hat{P}}$, which shows that the sequence $\hat{P}_{k+1}^x=g(\hat{P}_k^x)$ is monotonically convergent to $\overline{\hat{P}}$, for any $\hat{P}_0\geqslant 0$. The proof is complete. □

Proof of Theorem 3:
From (16) it follows that $P_{k+1}^x=S_k-\hat{L}_k^*T_k^{\mathrm{T}}-T_k\hat{L}_k^{*\mathrm{T}}+\hat{L}_k^*U_k\hat{L}_k^{*\mathrm{T}}+R_w$. As $k\to\infty$, we have

$$\begin{aligned}\lim_{k\to\infty}P_{k+1}^x&=\lim_{k\to\infty}[MO_kM^{\mathrm{T}}-\overline{\hat{L}}(MO_kN^{\mathrm{T}}-B\overline{\hat{H}}R_v)^{\mathrm{T}}-(MO_kN^{\mathrm{T}}-B\overline{\hat{H}}R_v)\overline{\hat{L}}^{\mathrm{T}}\\&\quad+\overline{\hat{L}}(NO_kN^{\mathrm{T}}+R_v-D\overline{\hat{H}}R_v-R_v(D\overline{\hat{H}})^{\mathrm{T}})\overline{\hat{L}}^{\mathrm{T}}+R_w]\\&=\lim_{k\to\infty}[(\overline{\hat{L}}N-M)O_k(\overline{\hat{L}}N-M)^{\mathrm{T}}+\Omega],\end{aligned}$$

where $\Omega=\overline{\hat{L}}(B\overline{\hat{H}}R_v)^{\mathrm{T}}+(B\overline{\hat{H}}R_v)\overline{\hat{L}}^{\mathrm{T}}+\overline{\hat{L}}[R_v-D\overline{\hat{H}}R_v-R_v(D\overline{\hat{H}})^{\mathrm{T}}]\overline{\hat{L}}^{\mathrm{T}}+R_w$. Note that there always exists a $\Xi\geqslant 0$ such that $\Omega\leqslant\Xi$. Then we have $\lim_{k\to\infty}P_{k+1}^x\leqslant\lim_{k\to\infty}\varphi(P_k^x)+\Xi$. Using Lemma A4, it can be concluded that $\lim_{k\to\infty}P_{k+1}^x$ is upper bounded. Since P_k^u is a monotonically increasing function of P_k^x, $\lim_{k\to\infty}P_k^u$ is upper bounded as well. □

REFERENCES

1. Patton R, Frank P, Clark R (eds). *Fault Diagnosis in Dynamic Systems: Theory and Application*. Prentice-Hall: Upper Saddle River, 1989.
2. Li X, Jilkov V. A survey of maneuvering target tracking–Part IV: decision-based methods. *Proceedings of the SPIE Conference on Signal and Data Processing of Small Targets*, Orlando, FL, 2–4 April 2002; 4728–4760.
3. Sharan M, Raman S. *Weather and Climate: The M.P. Singh Volume Part II*. Birkhäuser: Basel, 2005.
4. Zhang W, Branicky M, Phillips S. Stability of networked control systems. *IEEE Control Systems Magazine* 2004; **21**:84–99.
5. Kitanidis P. Unbiased minimum-variance linear state estimation. *Automatica* 1987; **23**:775–778.
6. Darouach M, Zasadzinski M. Unbiased minimum variance estimation for systems with unknown exogenous inputs. *Automatica* 1997; **33**:717–719.
7. Kerwin W, Prince JL. On the optimality of recursive unbiased state estimation with unknown inputs. *Automatica* 2000; **36**:1381–1383.
8. Darouach M, Zasadzinski M, Boutayeb M. Extension of minimum variance estimation for systems with unknown inputs. *Automatica* 2003; **39**:867–876.
9. Yang F, Wilde R. Observers for linear systems with unknown inputs. *IEEE Transactions on Automatic Control* 1988; **33**:677–681.

10. Darouach M, Zasadzinski M, Xu S. Full-order observers for linear systems with unknown inputs. *IEEE Transactions on Automatic Control* 1994; **39**:606–609.
11. Sundaram S, Hadjicostis CN. On delayed observers for linear systems with unknown inputs. *44th IEEE Conference on Decision and Control and 2005 European Control Conference*, *CDC-ECC'05*, December 2005; 7210–7215.
12. Floquet T, Edwards C, Spurgeon S. On sliding mode observers for systems with unknown inputs. *International Journal of Adaptive Control and Signal Processing* 2007; **21**:638–656.
13. Hou M, Patton R. Input observability and input reconstruction. *Automatica* 1998; **34**:789–794.
14. Wang D, Lum K. Adaptive unknown input observer approach for aircraft actuator fault detection and isolation. *International Journal of Adaptive Control and Signal Processing* 2007; **21**:31–48.
15. Duan GR, Patton R. Robust fault detection using Luenberger-type unknown input observers—a parametric approach. *International Journal of Systems Science* 2001; **32**:533–540.
16. Xiong Y, Saif M. Unknown disturbance inputs estimation based on a state functional observer design. *International Journal of Systems Science* 2003; **39**:1389–1398.
17. Hsieh C. Robust two-stage Kalman filters for systems with unknown inputs. *IEEE Transactions on Automatic Control* 2000; **45**:2374–2378.
18. Corless M, Tu J. State and input estimation for a class of uncertain systems. *Automatica* 1998; **34**:757–764.
19. Ha Q, Trinh H. State and input simultaneous estimation for a class of nonlinear systems. *Automatica* 2004; **40**:1779–1785.
20. Gillijns S, De Moor B. Unbiased minimum-variance input and state estimation for linear discrete-time systems with direct feedthrough. *Automatica* 2007; **43**:934–937.
21. Gillijns S, De Moor B. Unbiased minimum-variance input and state estimation for linear discrete-time systems. *Automatica* 2007; **43**:111–116.
22. Chen C. *Linear System Theory and Design*. Oxford University Press: New York, 1999.
23. Horn R, Johnson C. *Matrix Analysis*. Cambridge University Press: Cambridge, 1985.
24. Simon D. *Optimal State Estimation*: *Kalman*, $\mathscr{H}_\infty$, *and Nonlinear Approaches*. Wiley: New York, 2006.
25. Anderson B, Moore J. Convergence properties of Riccati equation solutions. *IEEE Transactions on Automatic Control* 1968; **13**:732–733.
26. Wonham W. On a matrix Riccati equation of stochastic control. *SIAM Journal on Control* 1968; **6**:681–698.
27. Chan S, Goodwin G, Sin K. Convergence properties of the Riccati difference in optimal filtering and nonstabilizable systems. *IEEE Transactions on Automatic Control* 1984; **29**:110–118.
28. Sinopoli B, Schenato L, Franceschetti M, Poolla K, Jordan M, Sastry S. Kalman filtering with intermittent observations. *IEEE Transactions on Automatic Control* 2004; **49**:1453–1464.
29. Shi Y, Yu B. Output feedback stabilization of networked control systems with random delays modeled by Markov chains. *IEEE Transactions on Automatic Control* 2009; **54**:1668–1674.
30. Shi Y, Fang H, Yan M. Kalman filter based adaptive control for networked systems with unknown parameters and randomly missing outputs. *International Journal of Robust and Nonlinear Control* 2009; **19**:1976–1992.
31. Shi Y, Fang H. Kalman filter based identification for systems with randomly missing measurements in a network environment. *International Journal of Control* 2010; **83**:538–551.
32. Fang H, Wu J, Shi Y. Genetic adaptive state estimation with missing input/output data. *Proceedings of the Institution of Mechanical Engineers*, *Part I*, *Journal of Systems and Control Engineering* 2010; **224**:611–617.

DOI: 10.1002/acs