A robust state estimator against constant measurement delay based on the sensitivity penalisation of model-parameter errors for systems with no exogenous inputs

Qiunong He1,2,3 | Huabo Liu4,5 | Qianwen Duan1,2,3 | Yao Mao1,2,3

1Key Laboratory of Optical Engineering, Chinese Academy of Sciences, Chengdu, China
2Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu, China
3University of Chinese Academy of Science, Beijing, China
4Automation, Qingdao University, Qingdao, China
5Institute for Future, Qingdao University, Qingdao, China

Correspondence
Yao Mao, Institute of Optics and Electronics, Chinese Academy of Sciences, Mailbox no. 350, No. 1 Guangguang Ave, Shuangliu District, Chengdu, 610209 Sichuan, China. Email: maoyao@ioe.ac.cn

Abstract
In this study, a class of linear system, which is with no exogenous input and suffered from constant measurement delay and uncertain model-parameter errors, is under consideration. To combat both the parametric uncertainties and constant measurement delay, a novel robust state estimator is proposed. Accounting for the constant measurement delay, a clever approach is utilised to expand the state vector and the system model is converted into an augmented delay-free model. Considering the deterioration of estimation performance caused by stochastic model-parameter errors, the sensitivity penalisation function of model-parameter errors is defined and introduced into the objective function of the regularised least-squares (RLS) problem, whose solution is the standard Kalman filter. Furthermore, by restricting the range of introduced parameter, the objective function of the modified RLS problem is converted into a strict convex function. Then, the recursive procedure of the proposed estimator is derived. The asymptotic stability conditions of the proposed estimator and the conditions for boundness of the estimation error matrix are given. Numerical simulations show the effectiveness of the estimator proposed in this paper.

1 | INTRODUCTION

State estimation is an indispensable part of cybernetics. Since the model-parameter errors and time delays are usually unavoidable in actual systems, the automatic control and signal processing of the systems encounter great difficulties and challenges [1–4]. It is often difficult to reduce the model-parameter errors and time delays directly. Therefore, an estimation approach, which can tackle both model-parameter errors and time delays, is needed. In this research, the estimator, which has the above mentioned attribute, is designated by the name of the anti-time delay robust state estimator (ATDRSE).

Throughout the literatures, the studies of ATDRSE have attracted so many researches. As seen in [5, 6], the robust Kalman filtering for known state-delay is researched from the discrete and continuous domains, respectively. Certainly, there is also no lack of research on the ATDRSE against unknown state delays [7–9]. In order to distinguish, the type of ATDRSE for state delays is called anti-state delay robust state estimator (ASDRSE) in this paper. For example, based on the published results of H-infinity filtering against systems without time delays, Mahmoud [8] proposed an ASDRSE scheme for linear uncertain systems suffered by unknown state delay. Focussed on making the estimation error stochastic exponentially stable,
Wang et al. [7] designed an ASDRSE for a sort of non-linear systems suffered by unknown state delays, modelling errors, and unknown external non-linear interferences. For a class of non-linear systems with state delays, Wei et al. [9] combined the fuzzy theory and H-infinity filtering method to design the ASDRSE.

For measurement delays, the published documents normally design estimators based on certain specific scenarios. Few documents provide a universal method such as [5, 6]. In the present research, the class of ATRDSE for measurement delays is called anti-measurement delay robust state estimator (AMDRSE). For example, the scenario of [10] assumes that the measurement delay of the system is a one-step random delay; then, the robust Kalman filtering is employed and modified in [10] to handle this situation. Similarly, [11] focuses on a type of discrete non-linear systems with unknown measurement delays. The measurement delays are treated as disturbances, and an enhanced estimator is presented in [11] to estimate and compensate the perturbed terms. The AMDRSE designing problem for a class of non-linear uncertain systems is investigated in [12]. In addition to considering the random measurement delays, the effects of packet loss and related noise are also considered at the same time. The state estimation problem for a special type of system is studied in [13] whose dimension is limited to be two, model parameters are time varying, and parametric uncertainties are norm-bounded. The delay and loss of measurements are considered at the same time. Besides, for time-delay systems expressed by complex networks, the results of [14–17] can provide some reference for state estimation design.

Based on the above analysis, the system time delay mainly includes the measurement and state delay. Measurement delay can be divided into constant delay and time-varying delay. In digital information processing systems with sensors, constant measurement delay is a common phenomenon. In systems with severe network congestion and packet loss, the time-varying delay occurs universally. However, the sensors’ acquisition environment, acquisition algorithm and transmission environment of most systems are fixed, so the measurement delay only changes slightly. Considering from the discrete domain, these subtle changes can be ignored. Therefore, it is reasonable to treat the measurement delay as approximately constant. For example, the target detector in the photoelectric tracking system has a considerable constant measurement delay. The photoelectric tracking systems are extensively applied in target tracking, aerospace, astronomical observations, and quantum laser communications that have made major breakthroughs in recent years. The constant measurement delay is one of the most important factors hindering the improvement of the system tracking performance. In the past few years, many research studies have been devoted to solving the problems caused by constant measurement delay from a cybernetics perspective, which reflects the constant measurement delay as an attractive issue rather than a conservative assumption.

The present study explores an AMDRSE solution for discrete-time linear systems suffered from parametric uncertainties, constant measurement delay and external disturbances. First, the state-space definition of the systems concerned in this paper is given. Then, the state vector is expanded based on the constant delay and the measurement equation. Combined with the process equation, a model without time delay and equivalent to the original system is obtained. Furthermore, considering the deterioration of estimation performance caused by random model errors, sensitivity penalisation function of the modelling errors is given, which is utilised to enhance the objective function of the regularised least square problem corresponding to the standard Kalman filter (SKF). To derive the recursive procedure, the improved cost function is further transformed into a strictly convex function by limiting the range of a multiplicative coefficient of the sensitivity penalty function. A rudiment of the filter is obtained by solving the partial derivative of the cost function. Founded on the rudiment, a complete robust filtering recursive programme is deduced. Moreover, some important attributes of the proposed filter are given and proven. The main contributions of this research are listed below.

1. In this study, an AMDRSE solution has been researched for the discrete-time linear systems that suffered from constant measurement delay, random parametric uncertainties, and external disturbances. This solution is essentially an improvement of Kalman filtering. One of the improvements is the clever conversion of the system model, and the other is the introduction of the sensitivity penalisation for modelling errors in the cost function.

2. This study gives a detailed derivation of the recursive procedure, conditions for boundness of the estimation error matrix and asymptotic stability conditions for the proposed AMDRSE.

3. Numerical simulations are appended to manifest the effectuality of the proposed approach.

Here is a summary of the rest of this article. The problem statement and the design of AMDRSE are presented in Section 2. Section 3 focusses on the derivation of recursive programs and some important properties of the proposed AMDRSE. The numerical simulations are discussed in Section 4. Section 5 presents the conclusion. The appendix exhibits the detailed derivation of the proposed AMDRSE’s recursive procedures and detailed proof of some theories proposed in this study.

Notations: Define \( \|x\| = \sqrt{x^T x} \) and \( \|x\|_W = \sqrt{x^T W x} \), in which \( x \) is a column vector and \( W \) is a positive definite matrix. \( E(y) \) and \( \text{cov}(y) \) represent the mathematical expectation and covariance matrix of stochastic vector \( y \), respectively. Define

\[
\text{col}\{z_1, z_2, z_3, \ldots\} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \end{bmatrix},
\]

where \( z_i, (i = 1,2,3,\ldots) \) is a vector or a matrix.
2 | PROBLEM STATEMENT AND THE AMDRSE DESIGN

The state-space model definition of the linear system with constant measurement delay and structural parameter uncertainties is as follows:

\[
\begin{align*}
\begin{cases}
    x_{k+1} &= A_k(e_k)x_k + B_k(e_k)w_k \\
y_k &= C_k(e_k)x_{k-d} + v_k
\end{cases}
\end{align*}
\]

(1)

In (1), \(x_k \in \mathbb{R}^n\) and \(y_k \in \mathbb{R}^m\) represent the state vector and measurement vector at moment \(k\), respectively. \(e_k\) expresses the modelling errors of the system, and consists of \(L\) real-valued scalars \(e_i, i = 1, \ldots, L\). \(d \in \mathbb{N}^+\) is known and constant, which implies the frames of measurement-delay. \(A_k(e_k), B_k(e_k), C_k(e_k)\) are the system matrices of size \(n \times n, n \times n, m \times n\), respectively. \(x_0, w_k\) and \(v_k\) are uncorrelated zero mean random variables. That is,

\[
E\left(\begin{bmatrix} x_0 \\
w_k \\
v_k \end{bmatrix} \begin{bmatrix} x_0 \\
w_k \\
v_k \end{bmatrix}^T\right) = \begin{bmatrix} \Pi_0 & Q_k \delta_{kj} \\
R_k \delta_{kj} & R_k \end{bmatrix}
\]

(2)

In (2)

\[
\delta_{kj} = \begin{cases} 
1, & k = j \\
0, & k \neq j
\end{cases}
\]

\(\Pi_0, Q_k\) and \(R_k\) represent the covariance matrix of \(x_0, w_k\), and \(v_k\), respectively.

Remark 1 In many actual systems, once the acquisition environment, acquisition algorithm and transmission environment of the sensors are determined, the measurement delay of the system will only make small fluctuations around a certain constant value. Considering this problem from the discrete domain, the measurement delay can be approximated as a constant value. In the measurement equation of system (1), \(y_k\) represents the measurement vector of system at time \(k\). But due to the measurement delay, \(y_k\) is actually the measurement corresponding to the state at time \(k - d\).

Since \(d\) is known and time invariant, the system (1) can be equivalently reconstructed as a delay-free system (3).

\[
\begin{align*}
\begin{cases}
\tilde{x}_{k+1} &= \tilde{A}_k(e_k)\tilde{x}_k + \tilde{B}_k(e_k)w_k \\
y_k &= \tilde{C}_k(e_k)\tilde{x}_k + v_k
\end{cases}
\end{align*}
\]

(3)

In (3), \(\tilde{x}_k = [x_k^T, \Delta_{k-1}^T, \ldots, \Delta_{k-d+1}^T]^T, \Delta_k = C_{k-1}(e_{k-1})x_k - C_k(e_k)x_k\)

When \(e_k \equiv 0, d \equiv 0\) and \(x_0, w_k\) and \(v_k\) all follow a normal distribution, the SKF is the optimal filter under the mean square error evaluation standard. The cost function of SKF can be expressed as (4) [27].

\[
\begin{align*}
\hat{x}_{k|k+1}(\tilde{x}_{k|k+1}) &= \arg\min_{\tilde{x}_{k|k+1}} \left[ \|x_k - \tilde{x}_{k|k+1}\|_P^{-1} \right. \\
&+ \left. \|w_k\|_Q^{-1} + \left\|y_{k+1} - C\tilde{x}_{k+1}\right\|^2_{P_{k+1}^{-1}} \right]
\end{align*}
\]

(4)

As the modelling errors and measurement delay caused the deterioration of the state estimation, the cost function (4) is changed in this paper.

\[
J(\alpha_k) = y_k\left[ \|\alpha_k\|_\Phi_k + \|H_k(0,0)\alpha_k - \beta_k(0,0)\|^2_{\psi_k} \right]
\]

\[
+ (1 - y_k) \sum_{i=1}^L \left( \left\| \frac{\partial e_k(0,0)}{\partial \alpha_{k,i}} \right\|^2 + \left\| \frac{\partial \beta_k(0,0)}{\partial \alpha_{k,i}} \right\|^2 \right)
\]

(5)

In (5),

\[
\begin{align*}
\Psi_k &= R_k^{-1}, \\
H_k(0,0) &= \tilde{C}_k(e_k), \\
\beta_k(0,0) &= y_k - \tilde{C}_{k+1}(e_{k+1})\tilde{A}_k(e_k)\tilde{x}_{k|k}, \\
\Phi_k &= \text{diag}\left\{ P_{k|k}^{-1}, Q_k^{-1} \right\}, \\
\alpha_k &= \text{col}\{\tilde{x}_k - \tilde{x}_{k|k}, w_k\}, \\
e_k(0,0) &= y_{k+1} - \tilde{C}_{k+1}(e_{k+1})\tilde{A}_k(e_k)\tilde{x}_{k|k}
\end{align*}
\]

(6)
The sensitivity function [28] \( e_k(e_k, e_{k+1}) \) is derived from \( y_{k+1} - C_{k+1}(0)[A_k(0)\hat{x}_{k|k} + B_k(0)\hat{w}_{k|k+1}] \). It reflects the impact of model-parameter errors on the prediction error of \( y_{k+1} \). Therefore, to abate the impact of model-parameter errors on the performance of the estimator, the sensitivity penalisation function of \( e_k(e_k, e_{k+1}) \) is added in the cost function of SKF. Here, \( y_k \in (0, 1] \). The effect of the parameter \( \gamma_k \) is to assign the weight of two factors that affect the overall performance of the estimator. One factor is the importance of the nominal performance, and the other is the importance of reducing the adverse effects caused by model-parameter errors. To utilise the cost function designed in this paper for further derivation, define two matrices as follows.

\[
S_k = \begin{bmatrix} S_{k,1}(0,0) & \vdots & S_{k,L}(0,0) \end{bmatrix}, \quad T_k = \begin{bmatrix} T_{k,1}(0,0) & \vdots & T_{k,L}(0,0) \end{bmatrix}
\]

(7)

In every \( i = 1, 2, \ldots, L, \)

\[
S_{k,i}(e_k, e_{k+1}) = \begin{bmatrix} \frac{\partial C_{k+1}(e_{k+1})}{\partial e_{k+1,i}} A_k(e_k) \\ \frac{\partial C_{k+1}(e_{k+1})}{\partial \epsilon_{k,i}} H_k(e_k) \end{bmatrix}
\]

(8)

\[
T_{k,i}(e_k, e_{k+1}) = \begin{bmatrix} \frac{\partial C_{k+1}(e_{k+1})}{\partial e_{k+1,i}} B_k(e_k) \\ \frac{\partial C_{k+1}(e_{k+1})}{\partial \epsilon_{k,i}} \beta_k(e_k) \end{bmatrix}
\]

(9)

According to (7) and (8), (9) is obtained.

\[
\sum_{i=1}^{L}\left( \| \frac{\partial e_k(0,0)}{\partial \epsilon_{k,i}} \| ^2 + \| \frac{\partial e_{k+1}(0,0)}{\partial \epsilon_{k,i}} \| ^2 \right) = \| [S_k \ T_k] \alpha_k + S_k \hat{x}_{k|k} \|^2
\]

(10)

Substituting (9) into (5) and finding the partial derivative of \( S_k \) for \( f(\alpha_k) \), (10) can be obtained.

\[
\frac{\partial f(\alpha_k)}{\partial \alpha_k} = 2\gamma_k \left\{ \left( \Phi_k + H_k(0,0)^T \Psi_k H_k(0,0) \right) + \frac{1 - \gamma_k}{\gamma_k} [S_k \ T_k] [S_k \ T_k]^T \alpha_k - H_k(0,0)^T \Psi_k \beta_k(0,0) + \frac{1 - \gamma_k}{T_k} [S_k \ T_k]^T S_k \hat{x}_{k|k} \right\}
\]

(11)

According to the definition given by formula (6), \( \Phi_k \) is positive definite and \( H_k(0,0)^T \Psi_k H_k(0,0) \) is at least semipositive definite. It can be seen from the expression in (5) that if \( \gamma_k \in (0, 1] \), \( f(\alpha_k) \) is strictly convex. When \( \frac{\partial f(\alpha_k)}{\partial \alpha_k} = 0 \), \( f(\alpha_k) \) obtains the global optimal value. The optimal \( \alpha_k \) is denoted by \( \hat{\alpha}_{k|k+1} \) which is uniquely determined by (11).

\[
\begin{aligned}
\Phi_k + H_k(0,0)^T \Psi_k H_k(0,0) + \frac{1 - \gamma_k}{\gamma_k} [S_k \ T_k] [S_k \ T_k]^T S_k \hat{x}_{k|k} \\
= H_k(0,0)^T \Psi_k \beta_k(0,0) + \frac{1 - \gamma_k}{\gamma_k} [S_k \ T_k] [S_k \ T_k]^T S_k \hat{x}_{k|k}
\end{aligned}
\]

(12)

There is no process noise at moment 0. To obtain an accurate initial state estimate, it is necessary to minimise \( J_0 \) described by (13).

\[
J_0 = \gamma_0 \left[ \| \hat{x}_0 \|_{\Pi_0}^2 + \| y_0 - \tilde{C}_0(\epsilon_0) \hat{x}_0 \|_{\tilde{R}_0}^2 \right] + (1 - \gamma_0) \sum_{i=1}^{L} \left( \| \frac{\partial \epsilon_0(0,0)}{\partial \epsilon_{0,i}} \| \right)^2
\]

(13)

In (12), \( \epsilon_0(\epsilon_0) = y_0 - \tilde{C}_0(\epsilon_0) \hat{x}_0 \). Obviously, when \( 0 < \gamma_k \leq 1 \), \( J_0 \) is also strictly convex. Let \( \frac{\partial f_0}{\partial \hat{x}_0} = 0 \), \( (13) \) is obtained.

\[
\hat{x}_{0|0} = \left( \hat{\Pi}_0^{-1} + \tilde{C}_0^T(\epsilon_0) R_0^{-1} \tilde{C}_0(\epsilon_0) \right)^{-1} \tilde{C}_0(\epsilon_0) R_0^{-1} y_0,
\]

(14)

where \( \hat{\Pi}_0^{-1} = \left( \hat{\Pi}_0^{-1} + \frac{1 - \gamma_k}{\gamma_k} \sum_{i=1}^{L} \left( \frac{\partial \epsilon_0(0,0)}{\partial \epsilon_{0,i}} \right) \right) \). In addition, it can be directly proved that when modelling error is equal to zero, \( \text{cov} \left( \hat{x}_0 - \hat{x}_{0|0} \right) = \left( \hat{\Pi}_0^{-1} + \tilde{C}_0(\epsilon_0) R_0^{-1} \tilde{C}_0(\epsilon_0) \right)^{-1} \). According to (11), the complete recursive process of the estimator can be further obtained. See Appendix A for its detailed derivation.

First, the initialisation of the recursive process is shown in (14) and (15).

\[
P_{0|0} = \left( \hat{\Pi}_0^{-1} + \tilde{C}_0^T(\epsilon_0) R_0^{-1} \tilde{C}_0(\epsilon_0) \right)^{-1}.
\]

(14)

\[
\hat{x}_{0|0} = P_{0|0} \hat{C}_0(\epsilon_0) R_0^{-1} y_0.
\]

(15)

Then, (16)–(19) are the renewal equations of some intermediate matrices used in state estimation.
\[
\hat{\Delta}_k(0) = \left[ A_k(0) - \frac{1 - \gamma_k}{\gamma_k} \hat{B}_k(0) \hat{Q}_k S_k \right] \times \left( I - \frac{1 - \gamma_k}{\gamma_k} \hat{P}_{k|k} S_k^T S_k \right)
\]

(16)

\[
\hat{B}_k(0) = \hat{B}_k(0) - \frac{1 - \gamma_k}{\gamma_k} \hat{A}_k(0) \hat{P}_{k|k} S_k^T S_k
\]

(17)

\[
\hat{P}_{k|k} = \left( P_{k|k}^{-1} - \frac{1 - \gamma_k}{\gamma_k} S_k^T S_k \right)^{-1}
\]

(18)

\[
\hat{Q}_k = \left[ Q_k^{-1} + \frac{1 - \gamma_k}{\gamma_k} T_k^T \left( I - \frac{1 - \gamma_k}{\gamma_k} S_k P_{k|k} S_k^T \right)^{-1} T_k \right]^{-1}
\]

(19)

Finally, (20)–(23) give the state estimation equation and updates of related covariance matrices.

\[
\hat{x}_{k+1|k+1} = \hat{A}_k(0) \hat{x}_{k|k} + P_{k+1|k+1} \hat{C}_{k+1}(0) R_{k+1}^{-1}
\]

\[
\times \left[ \gamma_{k+1} - C_{k+1}(0) \hat{A}_k(0) \hat{x}_{k|k} \right]
\]

(20)

\[
P_{k+1|k} = \hat{A}_k(0) \hat{P}_{k|k} \hat{A}_k(0)^T + \hat{B}_k(0) \hat{Q}_k \hat{B}_k(0)^T + C_k(0) P_{k+1|k} \hat{C}_k(0)^T
\]

(21)

\[
R_{k+1} = R_{k+1} + C_k(0) P_{k+1|k} \hat{C}_k(0)^T
\]

(22)

\[
P_{k+1|k+1} = P_{k+1|k}
\]

\[
- P_{k+1|k} \hat{C}_k(0)^T R_{k+1}^{-1} P_{k+1|k} \hat{C}_k(0)^T
\]

(23)

\section{Convergence and Boundedness of the AMDRSE}

This section will discuss the asymptotic stability conditions for the proposed AMDRSE and conditions for boundness of the estimation error matrix. Before starting these discussions, \( \epsilon_{k,i} \) is assumed to be contractive. This assumption can be satisfied by normalising the model-parameter errors. A set is used to include these \( \epsilon_{k,i} \), that is, \( \mathbb{S} = \{ \epsilon | \epsilon_{k,i} \leq 1, i = 1, 2, \ldots, L \} \). In addition, an extensively applied extended definition about quadratic stability [29, 30] is introduced.

**Definition 1** If the following condition, if there is a constant \( V > 0 \) such that \( \lim_{k \to \infty} | V - \Omega_k(\sigma_k) V \Omega_k^T(\sigma_k) | > 0, \forall \sigma_k \in \Lambda' \), is satisfied, then the time-varying uncertain square matrix \( \Omega_k(\sigma_k) \) is called to be asymptotically quadratically stable (AQS) relative to the set \( \Lambda \).

To discuss the convergence of the proposed AMSRE, its estimation error covariance matrix (EECM) should be considered first. Note that the system’s dynamics description in this article includes uncertain parameters. The EECM of AMSRE cannot be accurately described. Therefore, this article will discuss the convergence of the proposed AMSRE from the upper bound of EECM [30, 31]. For the sake of brevity, \( (1 - \gamma_k)/\gamma_k \) is denoted by \( \lambda_k^\gamma \) in the following discussion. Three assumptions are given below, which will be used many times in the following discussion.

A1) \( \hat{A}_k(0), \hat{B}_k(0), \hat{C}_k(0), R_k, \hat{Q}_k, S_k, T_k, \gamma_k \) are time-invariant.

A2) \( \hat{A}_k(\epsilon_k) \) is AQS, it can also be said that system (3) is AQS. \( \hat{A}_k(\epsilon_k), \hat{B}_k(\epsilon_k), \hat{C}_k(\epsilon_k) \) are bounded when \( k \geq 0 \) and \( \epsilon_k \in \mathbb{S} \).

A3) \( T_k^T S_k = 0, k = 1, 2, \ldots \).

**Remark 2** Among the above assumptions, the condition restricted by the last assumption is the narrowest. According to (9) and (6), it is not difficult to understand the physical meaning of A3). That is, the combined influences of the model-parameter errors and state estimation on the proposed AMDRSE’s performance are constantly uncorrelated with that of the model-parameter errors and process noise. There is a sufficient and unnecessary condition that is easier to be accepted. When there is no model-parameter errors in the process noise coefficient matrix \( (\hat{B}_k(\epsilon_k)) \) and the output coefficient matrix \( (\hat{C}_k(\epsilon_k)) \) in system (3), A3) can be satisfied.

From \( P_{k+1|k+1} \) and \( \hat{Q}_k \) in the recursive programme, (24) and (25) are obtained by direct algebraic operations.

\[
P_{k+1|k} = P_{k+1|k} - P_{k+1|k} \hat{C}_k(0)^T R_{k+1}^{-1} \hat{C}_k(0) P_{k+1|k}
\]

(24)

\[
\hat{Q}_k = \left[ Q_k^{-1} + \lambda_k T_k^T T_k \right]^{-1}
\]

(25)

In (25),

\[
\begin{align*}
S_k &= \left( I + \lambda_k T_k \hat{Q}_k T_k^T \right)^{-1/2} S_k \\
T_k &= \left( I + \lambda_k T_k \hat{Q}_k T_k^T \right)^{-1/2} T_k
\end{align*}
\]

(26)
Based on the recursive formula of $P_{k+1|k}$, (30) can be further obtained

$$P_{k+1|k} = \tilde{A}_k(0)P_{k|k-1}\tilde{A}_k^T(0)$$

$$+ \tilde{A}_k(0)P_{k|k-1}Z_k^T(\tilde{I} + Z_kP_{k|k-1}Z_k^T)^{-1}$$

$$\times \left(\tilde{A}_k(0)P_{k|k-1}Z_k^T\right)^T$$

$$+ B_k(0)\left(Q_k^{-1} + \lambda_kT_k^TT_k\right)^{-1}B_k^T(0).$$

On the basis of the relationship in (30) and the fact that $Q_k^{-1} + \lambda_kT_k^TT_k$ is reversible, the following conclusions can be obtained. Since it is very similar to the well-known argument in Kalman filtering, a detailed proof is omitted [29, 30].

**Lemma 1** Assuming that both A1) and A3) are true, $(\tilde{A}_k(0), Z_k)$ is detectable and $(\tilde{A}_k(0), \tilde{B}_k(0))$

$\left(Q_k^{-1} + \lambda_kT_k^TT_k\right)^{-1/2}$ is stabilisable. The Riccati of (30) recursively converges to a positive semidefinite matrix $P$ by stabilising

$$A_f = \tilde{A}_k(0)\left[I - PZ_k^T(\tilde{I} + Z_kPZ_k^T)^{-1}Z_k\right].$$

Based on Lemma 1, some stability conditions of the proposed AMDRSE can be further built.

**Theorem 1** Assuming that the two assumptions and two conditions used in 1 are true, the proposed AMDRSE converges to a linear time-invariant stable system as the time variable $k$ increases.

Proof According to the recursive process of the proposed AMDRSE, (20) can be transformed into (31) through simple algebraic operations.

$$\hat{x}_{k+1|k+1} = \left[I - P_{k+1|k+1}\tilde{C}_{k+1}(0)R_k^{-1}\tilde{C}_{k+1}(0)\right]$$

$$\times \tilde{A}_k(0)\hat{x}_{k|k} + P_{k+1|k+1}\tilde{C}_{k+1}(0)R_k^{-1}\tilde{y}_{k+1}.$$ (31)

In addition, if $T_k^TS_k = 0$, it is obvious that

$$\left[I - P_{k+1|k+1}\tilde{C}_{k+1}(0)R_k^{-1}\tilde{C}_{k+1}(0)\right]^{-1}P_{k|k-1}\tilde{Z}_k$$

$$\times \left[I - P_{k|k-1}\tilde{Z}_k^T(\tilde{I} + Z_kP_{k|k-1}\tilde{Z}_k^T)^{-1}Z_k\right]$$

$$\times \left[I + P_{k|k-1}\tilde{C}_k^T(0)R_k^{-1}\tilde{C}_k(0)\right].$$

According to Lemma 1, if all its conditions are true, then $\lim_{k\to\infty}P_{k|k-1} = P$ and $A_f$ is stable. The proof is complete. □

To simplify the mathematical expression for analysing the boundedness of the proposed AMDRSE, define

$$X_k = [I + \Gamma_k(0)]\hat{x}_k$$

$$\tilde{X}_{k|k} = [I + \Gamma_k(0)]\hat{x}_{k|k}$$

(33)

$$\tilde{X}_{k|k} = X_k - \hat{X}_k.$$ In (33), $\Gamma_k(\varepsilon_k) = P_{k|k-1}\tilde{C}_k^T(0)R_k^{-1}\tilde{C}_k(\varepsilon_k)$. According to these definitions, direct algebraic operations can get $X_{k|k} = [I + \Gamma_k(0)](x_k - \hat{x}_{k|k})$. Because matrix $I + \Gamma_k(0)$ is always invertible, the covariance matrix of $(x_k - \hat{x}_{k|k})$ is bounded if and only if $X_{k|k}$ has a bounded covariance matrix. Define matrices,

$$F_k(\varepsilon_k) = [I + \Gamma_k(0)]\tilde{A}_k(\varepsilon_k)[I + \Gamma_k(0)]^{-1}$$

$$G_k(\varepsilon_k) = [I + \Gamma_k(0)]\tilde{B}_k(\varepsilon_k)$$

$$A_f(\varepsilon_k) = \tilde{A}_k(0)\left[I - P_{k|k-1}\tilde{Z}_k^T(\tilde{I} + Z_kP_{k|k-1}\tilde{Z}_k^T)^{-1}Z_k\right]$$

$$\tilde{A}_k(\varepsilon_k) = \begin{bmatrix} \tilde{A}_k^T(0) & \tilde{A}_k^T(0) \\ \tilde{A}_k^T(0) & \tilde{A}_k^T(0) \end{bmatrix}$$

$$\tilde{B}_k(\varepsilon_k) = \begin{bmatrix} \tilde{B}_k^T(0) & \tilde{B}_k^T(0) \\ \tilde{B}_k^T(0) & \tilde{B}_k^T(0) \end{bmatrix}$$

(34)
In (34),
\[
A_k(e_k)_{11} = F_k(e_k) - \Gamma_k(e_k)[I + \Gamma_k(0)]^{-1}F_k(e_k)
\]
\[
A_k(e_k)_{12} = F_k(e_k) - A_k - [I + \Gamma_k(0)]^{-1}F_k(e_k)
\]
\[
\tilde{A}_k(e_k)_{21} = A_k \Gamma_k(e_k)[I + \Gamma_k(0)]^{-1}F_k(e_k)
\]
\[
\tilde{A}_k(e_k)_{22} = A_k \Gamma_k(e_k)[I + \Gamma_k(0)]^{-1}F_k(e_k)
\]
\[
\tilde{B}_k(e_k)_{11} = \{I - \Gamma_k(e_k)[I + \Gamma_k(0)]^{-1}\}G_k(e_k)
\]
\[
\tilde{B}_k(e_k)_{12} = -P_{k|k-1}C_k(0)R_k
\]
\[
\tilde{B}_k(e_k)_{21} = \Gamma_k(e_k)[I + \Gamma_k(0)]^{-1}G_k(e_k)
\]
\[
\tilde{B}_k(e_k)_{22} = P_{k|k-1}C_k(0)R_k
\]

If all assumptions and conditions used in Lemma 1 are true, when \(k \to \infty\), then
\[
\begin{bmatrix}
\hat{X}_{k+1|k+1} \\
\hat{X}_{k+1|k+1}
\end{bmatrix} = \begin{bmatrix}
\hat{A}_k(e_k) & \hat{B}_k(e_k)
\end{bmatrix} \begin{bmatrix}
\hat{X}_{k|k} \\
\hat{X}_{k|k}
\end{bmatrix} + \begin{bmatrix}
\hat{B}_k(e_k)
\end{bmatrix} \begin{bmatrix}
\hat{w}_k \\
\hat{v}_{k+1}
\end{bmatrix}
\]

(35)

According to the above relationship and the stability of \(\hat{A}_k(e_k)\), the conditions about the boundedness of the proposed AMDRSE’s EECM can be obtained. See Appendices B and C for details.

**Theorem 2** When all the conditions required by Lemma 1 are met, and A2) is also true, then the EECM of the proposed AMDRSE is bounded at each time \(k\).

The following introduces another property of the proposed AMDRSE. Here, continue to use the conditions of Theorem 2 and the definitions in Appendix C. First, assuming
\[
M = \begin{bmatrix}
M_{11} & M_{12} \\
M_{21} & M_{22}
\end{bmatrix},
\]
and the set of \(M_{11}\) is represented by \(M_{11}\), the following inequality is obtained.
\[
\text{cov}(\tilde{x}_k - \hat{x}_{k|k}) \leq \inf_{M_{11}} \left\{ \left[I + P_{k|k-1}C_k(0)R_k^{-1}C_k(0)^T\right]M_{11} \right\} \times \left[I + P_{k|k-1}C_k(0)R_k^{-1}C_k(0)^T\right]^T
\]

Further, according to (35), the following equation is obtained.

\[
E \left\{ V^{-12} \begin{bmatrix}
\hat{X}_{k+1|k+1} \\
\hat{X}_{k+1|k+1}
\end{bmatrix} \right\} = \left[\prod_{j=0}^{k} \left(V^{-12} A_j(e) V^{-12}\right)\right] \left(V^{-12} \begin{bmatrix}
\hat{X}_{0|0} \\
\hat{X}_{0|0}
\end{bmatrix}\right)
\]

Consequently,
\[
\lim_{k \to \infty} E \left\{ V^{-12} \begin{bmatrix}
\hat{X}_{k+1|k+1} \\
\hat{X}_{k+1|k+1}
\end{bmatrix} \right\} = 0.
\]

This further indicates that \(\lim_{k \to \infty} E[\tilde{x}_k - \hat{x}_{k|k}] = 0\). In other words, the AMDRSE proposed in this paper is asymptotically unbiased.

### 4 | NUMERICAL SIMULATIONS

A comparative approach is used to demonstrate the effectiveness of the proposed AMDRSE in this study. This comparative experiment performed 500 random simulations. Each simulation generates 1000 frames of data, each frame contains the corresponding real state vector, process noise, measurement vector and measurement noise for the estimated plant. In each simulation, the first state vector is initialised as a zero vector. The number of measurement-delay frames is set to 3. The process noise \(\hat{w}_k\) and measurement noise \(\hat{v}_k\) are assumed to obey a normal distribution. The results of this \(5 \times 10^5\) random simulations are used to calculate the ensemble-average estimated error variance at each sampling time. That is,
\[
E\left\|x_k - \hat{x}_{k|k}\right\|^2 \approx \frac{1}{500} \sum_{i=1}^{500} \left\|x_k - \hat{x}_{k|k}^{(i)}\right\|
\]

where \(i\) is the serial number of the random simulations.

The proposed AMDRSE is compared with the other two methods. One is a Kalman filter estimation method founded on a conjunction of the nominal parameters and the proposed augmented delay-free model (KFND). The other is founded on a conjunction of the actual parameters and the proposed augmented delay-free model (KFAD). This comparison can still demonstrate the effectiveness of the proposed AMDRSE. Referencing to the nominal parameters of [27, 29] the system parameters selected in this example are as follows.

\[
A_k(e_k) = \begin{bmatrix}
0.9802 & 0.0196 + p \cdot e_k \\
0.0000 & 0.9802
\end{bmatrix},
\]
\[
B_k(e_k) = \begin{bmatrix}
1.0000 \\
0.0000
\end{bmatrix},
\]
\[
C_k(e_k) = \begin{bmatrix}
1.0000 \\
-1.0000
\end{bmatrix}, R_k = 1.0000,
\]
\[
Q_k = \begin{bmatrix}
1.9608 & 0.0195 \\
0.0195 & 1.9605
\end{bmatrix}, \Pi_0 = \begin{bmatrix}
1.0000 & 0.0000 \\
0.0000 & 1.0000
\end{bmatrix}.
\]

(36)
Suppose the sampling period of the system is $T = 0.01$. In this paper, three sets of experiments are made. The differences among these three sets of experiments are the change of $p$. The purpose is to change the 'size' of uncertainty. $p = 0.1$ is in the first group, $p = 0.5$ and $p = 1.0$ are in the second and third group, respectively. At the same time, in these $5 \times 10^2$ simulations, the modeling errors $\varepsilon_k$ follow a truncated normal distribution ($\varepsilon_k \sim N(0, 1)$, $\tilde{S} = \{\varepsilon_k|\varepsilon_k \leq 1, i = 1, \ldots, L\}$). In addition, each set of experiments select two different $\gamma_k$ ($\gamma_k = 0.7$ or $\gamma_k = 0.8$). When the appropriate $\gamma_k$ is selected, the accuracy of AMDRSE proposed in this paper will be better [29, 31].

Before starting the formal experiment, we first need to verify the convergence properties of the proposed AMDRSE. Obviously, $A(k(0), B_k(0), C_k(0), R_k, Q_k, \gamma_k)$ are time invariant. Since the term containing $\varepsilon_k$ in the model (36) is first order, $S_k$ and $T_k$ are time invariant according to (8) and (9). Consequently, Assumption A1 is true. Substituting (36) into system (3), (8) and (9), $A_k(\epsilon_k), B_k(\epsilon_k), C_k(\epsilon_k)$ can be obtained. Because $B_k(\epsilon_k)$ and $C_k(\epsilon_k)$ do not contain the uncertain factor $\varepsilon_k$, $T_k = 0$. Therefore, Assumption A3 is true. According to the results in [32], an equivalent condition for $\tilde{A}_k(0), Z_k$ to be detectable is that if

\[
\text{rank}\left( \frac{\tilde{A}_k(0) - \lambda I_{n + dm}}{Z_k} \right) = n + dm,
\]

where $\lambda_i$ is the eigenvalue of $\tilde{A}_k(0)$. After simple algebraic operations, $\lambda_1 = \lambda_2 = \lambda_3 = 0$, $\lambda_4 = \lambda_5 = 0.9802$ is obtained. Substituting $\lambda_i = 0$ and $\lambda_i = 0.9802$ into matrix $\left( \frac{\tilde{A}_k(0) - \lambda I_{n + dm}}{Z_k} \right)$, rank $\left( \frac{\tilde{A}_k(0) - \lambda I_{n + dm}}{Z_k} \right) = n + dm$ is finally obtained. That is, $(\tilde{A}_k(0), Z_k)$ is detectable. On the other hand, the stabilisation of $(\tilde{A}_k(0), B_k(0)\left( Q_k^{-1} + \lambda_k T_k^T T_k \right)^{-1/2})$ has similar equivalent conditions. That is, if $\text{rank}\left( \frac{\tilde{A}_k(0) - \lambda I_{n + dm}}{Z_k} \right) = n + dm$

where $\lambda_i$ is the eigenvalue of $\tilde{A}_k(0)$ and

\[
\begin{align*}
TM1 &= \tilde{A}_k(0) - \lambda I_{n + dm} \\
TM2 &= B_k(0)\left( Q_k^{-1} + \lambda_k T_k^T T_k \right)^{-1/2}
\end{align*}
\]

The eigenvalues of $\tilde{A}_k(0)$ are known in the above analysis, and now, substituting $\lambda_k = 0.4286 (\gamma_k = 0.7)$ and $\lambda_k = 0.2299 (\gamma_k = 0.8)$ into matrix $\left( \frac{\tilde{A}_k(0) - \lambda I_{n + dm}}{Z_k} \right)$, $\tilde{A}_k(0) - \lambda I_{n + dm}$ is stabilisable whenever $\gamma_k = 0.7$ or $\gamma_k = 0.8$. According to the results of Theorem 1, the proposed AMDRSE, corresponding to the example of this section, converges to a linear time-invariant stable system as the time variable $k$ increases. Next, the EECM’s boundness of the proposed AMDRSE is verified. It is obvious that $A_k(\epsilon_k), B_k(\epsilon_k), C_k(\epsilon_k)$ are all bounded for $k > 0$ and $\epsilon_k \in \tilde{S}$. Simple algebraic operations show that the eigenvalues of $\tilde{A}_k(\epsilon_k)$ can be obtained as $\lambda_1 = \lambda_2 = \lambda_3 = 0, \lambda_4 = \lambda_5 = 0.9802$. They are all inside the unit circle. On referring to the results of [32], it can be observed that System 3 is AQs According to the results of Theorem 2, the EECM of the proposed AMDRSE is bounded at each time $k$.

The experimental results are shown in Figure 1. Figure 1 presents the experimental results of the first, second and third groups. Experimental results show that with the increase of uncertainty, the robust state estimation method proposed in this paper has better estimation performance than the KFND. The performance of the KFAD has always been the best of the three methods. On the contrary, as the uncertainty decreases, the estimation performance of three methods eventually tends to fair level. As the uncertainty increases, the method proposed in this paper, such as the KFND and KFAD, is finally stable. At the same time, these three groups of experiments also verify the proposed Theorems 1 and 2 from the perspective of experiments. In addition, the experimental results of each set of experiments show that the choice of $\gamma_k$ will affect the performance of the proposed filter.

As explained by Remark 1, in many actual systems, once the acquisition environment, acquisition algorithm and transmission environment of the sensors are determined, the measurement delay of the system will only make small fluctuations around a certain constant value. Considering this problem from the discrete domain, this study treats the measurement delay as a constant value. Considering that the measurement delay has a large range of variation in some application systems, an experiment is designed to analyse the performance of the proposed estimator, in which, the measurement delay varies randomly between two and four frames. In this experiment, we set $p = 0.5$ and $\epsilon_k \sim N(0, 1)$. The four situations of $(\gamma_k = 0.7, d = 3), (\gamma_k = 0.8, d = 3), (\gamma_k = 0.7, d = 2 \sim 4), (\gamma_k = 0.8, d = 2 \sim 4)$ are compared, respectively (in this experiment, when $d$ varies randomly in the range of $2-4$, our estimator still treats $d$ as a constant $3$ when it is implemented). According to the experimental results shown in Figure 2, when the measurement delay $(d)$ varies randomly within a range $(2–4)$, the performance of the proposed estimator will decrease but not severely. This implies that the proposed estimator is still effective when the measurement delay varies randomly within a small range.

Since the state augmentation method used by the proposed estimator to deal with the measurement delay will add additional computational burden, the proposed estimator is more suitable for the case where $d$ is small (suggest $d = 2–5$). Therefore, this study will not compare the performance of the proposed estimator when $d$ varies randomly in a large range.

An application-oriented example is provided below to further illustrate the effectiveness of this research in actual systems. Assume that the current scene such that a
photoelectric tracking system is tracking a target. The image sensor detecting the target has a time delay of three frames. Without loss of generality, we assume that the target is moving at a constant velocity. However, the target's motion model suffers from the uncertain parameters with known statistical characteristics. The specific model parameters are as follows:

FIGURE 1 The experimental results
In (37), we assume $\xi \sim N(0, 1)$. In this simulation experiment, we only aimed at the movement of the target in a certain dimension, but it is still sufficient to verify the effectiveness of the estimator proposed in this paper.

Figure 3 shows the simulated target trajectory, which is affected by model parametric uncertainties and the time-delayed measurements. These measurements are used as the input of the KFND, KFAD and the proposed AMDRSE method. Figure 4 shows the estimation errors of the three methods. Since the actual parameters are used, the KFAD method has the best estimation performance among the three methods. Compared with the KFND method using nominal parameters, the proposed robust design effectively reduces the estimation error when the state is affected by model parametric uncertainties. The results of this experiment illustrate the effectiveness of the proposed estimator in practical applications.

5 | CONCLUSION

Aiming at the state estimation problem of discrete-time linear systems with constant measurement-delay and random parametric uncertainties, this article proposes a new robust state estimator. First, the system's state space model is converted into an augmented delay-free model by expanding the dimension of the state vector in a clever way. Then, an AMDRSE is designed, that is founded on sensitivity penalisation of the modelling errors, and an analytical scheme is derived which can be implemented recursively. Furthermore, the conditions for the asymptotic stability of the estimator as well as the conditions for the boundness of the EECM are explicitly given. Numerical simulations manifest that the proposed AMDRSE performs excellent especially when the system encounters large uncertainty. At the same time, the choice of parameter $\gamma_k$ also has a certain impact on the performance of the AMDRSE.

Even though this article provides proof of some properties of the proposed AMDRSE, more efforts are still needed to alleviate the restrictive conditions of these properties in the future. For example, A3) is usually unsatisfied. Furthermore, there are two issues that will be the challenges faced by the
proposed AMDRSE. One is the reasonable selection range of parameter $\gamma_k$, and the other one is how to analyse the tolerance for the model parameters error and measurement delay.

ACKNOWLEDGEMENTS

This work was supported in part by the National Natural Science Foundation of China under Grant 61873138, in part by the Project of Shandong Province Higher Educational Science and Technology Program under Grant J18KA355, in part by the Shandong Provincial Natural Science Foundation under Grant ZR2019MF063, and in part by the Key Research and Development Plan of Shandong Province under Grant 2017GGX10115.

ORCID

Yao Mao https://orcid.org/0000-0003-1785-2018

REFERENCES

1. Rajan, M.R.B.B., Mohanty, A.R.: Time delay estimation using wavelet denoising maximum likelihood method for underwater reverberant environment. IET Radar Sonar Navig. 14(8), 1183–1191 (2020)
2. Liu, H., et al.: Delay-distribution-dependent h-infinity state estimation for discrete-time memristive neural networks with mixed time-delays and fading measurements. IEEE Trans. Cybern. 50(2), 440–451 (2020)
3. Ling, Y., TianShuang, Q., ShengYang, L.: Robust joint estimation for time delay and Doppler frequency shift based on generalised sigmoid cross-ambiguity function. IET Radar Sonar Navig. 11(5), 721–728 (2017)
4. Chen, T.C., Sheu, T.T.: Model reference robust speed control for induction-motor drive with time delay based on neural network. IEEE Trans. Syst. Man Cybern. Syst. Hum. 31(6), 746–753 (2001)
5. Sandberg, I.W.: Robust Kalman filter design for discrete time-delay systems. Circ. Syst. Signal Process. 21(3), 319–335 (2002)
6. Mahmoud, M.S., Muthairi, N.F.A., Bingulac, S.: Robust Kalman filtering for continuous time-lag systems. Syst. Contr. Lett. 38(4-5), 309–319 (1999)
7. Wang, Z., Burnham, K.J.: Robust filtering for a class of stochastic uncertain nonlinear time-delay systems via exponential state estimation. IEEE Trans. Signal Process. 49(4), 794–804 (2001)
8. Mahmoud, M.S.: Robust stability and H-infinity-estimation for uncertain discrete systems with state-delay. Math. Probl. Eng. 7(5), 393–412 (2001)
9. Wei, Y., et al.: A new design of H-infinity piecewise filtering for discrete-time nonlinear time-varying delay systems via T-S fuzzy affine models. IEEE Trans. Syst. Man Cybern. 47(8), 2034–2047 (2017)
APPENDIX A

A. Derivation of the recursive procedure

For simplicity, $\lambda_k$ is used to express $(1 - \gamma_k)/\gamma_k$. Direct matrix operations can get (A.1).

$$
\begin{bmatrix}
P^{-1}_{k|k} & 0 \\
0 & Q^{-1}_k \\
\end{bmatrix}
+ \lambda_k
\begin{bmatrix}
S_k^T S_k & S_k^T T_k \\
T_k S_k & T_k T_k \\
\end{bmatrix}
= \begin{bmatrix}
I & 0 \\
\lambda_k T_k^T S_k \hat{P}_{k|k} & I \\
\end{bmatrix}
\begin{bmatrix}
\hat{P}_{k|k} & 0 \\
0 & \hat{Q}^{-1}_{k|k} \\
\end{bmatrix}
\begin{bmatrix}
I & 0 \\
\lambda_k \hat{P}_{k|k} S_k^T T_k & I \\
\end{bmatrix}
$$

(A.1)

To simplify the symbol, define

$$
\begin{align*}
\hat{H}_k &= \tilde{C}_{k+1}(0) \begin{bmatrix} A_k(0) & B_k(0) \end{bmatrix} \\
\hat{T}_k &= T_k - \lambda_k S_k \hat{P}_{k|k} S_k^T T_k \\
\check{x}_{k|k+1} &= \lambda_k \hat{P}_{k|k} S_k^T T_k \tilde{w}_{k|k+1}
\end{align*}
$$

(A.2)

Substituting (A.1) into (11) results in (A.3).

$$
\begin{bmatrix}
\tilde{P}_{k|k}^{-1} & 0 \\
0 & \tilde{Q}_k^{-1} \\
\end{bmatrix}
+ \tilde{H}_k^T R_{k+1}^{-1} \tilde{H}_k
\begin{bmatrix}
\check{x}_{k|k+1} - \tilde{x}_{k|k} \tilde{w}_{k|k+1} \\
\end{bmatrix}
= \tilde{H}_k^T R_{k+1}^{-1} \begin{bmatrix} y_{k+1} - \tilde{C}_{k+1}(0) \tilde{A}_k(0) \tilde{x}_{k|k} & - \lambda_k S_k \tilde{T}_k \tilde{S}_{k|k} \\
\end{bmatrix}
$$

(A.3)

To avoid using the inverse of

$$
\begin{bmatrix}
\tilde{P}_{k|k}^{-1} & 0 \\
0 & \tilde{Q}_k^{-1} \\
\end{bmatrix}
+ \tilde{H}_k^T R_{k+1}^{-1} \tilde{H}_k
$$

directly when calculating the state estimate $\tilde{x}_{k+1|k+1}$, define the variable.

$$
\tilde{x}_{k+1|k+1} = \tilde{A}_k(0) \tilde{x}_{k|k+1} + \tilde{B}_k(0) \tilde{w}_{k|k+1}
$$

(A.4)
Then, according to the definition in A.4, A.3 can be equally split into the following equations.

\[
\begin{align*}
\hat{x}_{k|k+1} &= \hat{x}_{k|k} + \hat{P}_{k|k}C_{k+1}^T(0)R_{k+1}^{-1} \\
&\quad \times \left( y_{k+1} - C_{k+1}(0)\hat{x}_{k+1|k+1} \right) \\
&\quad - \lambda_k \hat{P}_{k|k}S_k^T S_k \hat{x}_{k|k} \\
\end{align*}
\tag{A.5}
\]

\[
\begin{align*}
\hat{w}_{k|k} &= \hat{Q}_k \hat{B}_k^T(0)C_{k+1}^T(0)R_{k+1}^{-1} \\
&\quad \times \left( y_{k+1} - C_{k+1}(0)\hat{x}_{k+1|k+1} \right) - \lambda_k \hat{Q}_k \hat{B}_k^T S_k \hat{x}_{k|k} \\
\end{align*}
\tag{A.6}
\]

After arranging A.5 and A.6, A.7 is obtained.

\[
\begin{align*}
\hat{x}_{k+1|k+1} &= \hat{A}_k(0)\hat{x}_{k|k} + P_{k+1|k}C_{k+1}^T(0)R_{k+1}^{-1} y_{k+1} \\
&\quad - P_{k+1|k}C_{k+1}^T(0)R_{k+1}^{-1} \hat{C}_{k+1}(0)\hat{x}_{k+1|k+1} \\
\end{align*}
\tag{A.7}
\]

That is,

\[
\begin{align*}
\left[ I + P_{k+1|k}C_{k+1}^T(0)R_{k+1}^{-1} \hat{C}_{k+1}(0) \right] \hat{x}_{k+1|k+1} &= \left[ I + P_{k+1|k}C_{k+1}^T(0)R_{k+1}^{-1} \hat{C}_{k+1}(0) \right] y_{k+1} \\
&\quad - \hat{A}_k(0)\hat{x}_{k|k} \\
&\quad - \left( y_{k+1} - \hat{C}_{k+1}(0)\hat{A}_k(0)\hat{x}_{k|k} \right) \\
\end{align*}
\tag{A.8}
\]

According to the well-known matrix inversion formula, it can be proved that

\[
\begin{align*}
\left[ I + P_{k+1|k}C_{k+1}^T(0)R_{k+1}^{-1} \hat{C}_{k+1}(0) \right]^{-1} P_{k+1|k} &= P_{k+1|k+1} \\
\end{align*}
\tag{A.9}
\]

These relationships can be further proved

\[
\begin{align*}
\hat{x}_{k+1|k+1} &= \hat{A}_k(0)\hat{x}_{k|k} + P_{k+1|k+1}C_{k+1}^T(0)R_{k+1}^{-1} \\
&\quad \times \left( y_{k+1} - \hat{C}_{k+1}(0)\hat{A}_k(0)\hat{x}_{k|k} \right) \\
\end{align*}
\tag{A.10}
\]

Note that \([29]\), eqs. 36 and 37 have exactly same format as (A.5) and (A.6). This shows that \(\hat{x}_{k+1|k+1}\) can be specified as \(\hat{x}_{k+1|k+1}\) in robust filtering. Proof is complete.

**B. Proof of Theorem 2: Part 1**

Before starting to formally prove Theorem 2, the following lemma will be used.

**Lemma 1** Suppose that when \(k > 0\) and \(e_k \in \mathbb{S}\), matrix \(\hat{C}_k(e_k)\) is bounded. Then, under the conditions of 1, if and only if \(\hat{A}_k(e_k)\) is AQS, \(\hat{A}_k(e_k)\) is AQS relative to \(\mathbb{S}\).

**Proof** To simplify, define the matrices

\[
D = \begin{bmatrix} I & I \\ 0 & I \end{bmatrix}
\]

\[
\Lambda_k(e_k) = \Gamma_k(e_k)[I + \Gamma_k(0)]^{-1}F_k(e_k)
\]

Assume that \(\hat{A}_k(e_k)\) is AQS relative to \(\mathbb{S}\). According to Definition 1, \(V - \hat{A}_k(e_k)V\hat{A}_k^T(e_k) > 0\), \(\forall k \geq N_0\), \(\forall e_k \in \mathbb{S}\), in which \(N_0 > 0\) and \(V\) is a time-invariant positive definite matrix (PDM). Multiply the left and right sides of this inequality by \(D\) and \(D^T\). Divide \(DVD^T\) into \(\begin{bmatrix} \tilde{V}_{11} & \tilde{V}_{12} \\ \tilde{V}_{12} & \tilde{V}_{22} \end{bmatrix}\), where \(\tilde{V}_{11}\) has the same dimension with \(F_k(e_k)\). According to the aforementioned inequality, it can be further derived that \(\tilde{V}_{11} > 0\) and \(\tilde{V}_{11} - F_k(e_k)\tilde{V}_{11}F_k^T(e_k) > 0\). It is easy to know that if all conditions mentioned in Lemma 1 are satisfied, \(\Gamma_k(0)\) converges to a constant matrix. According to the definition of \(F_k(e_k)\), it can be directly obtained that \(\hat{A}_k(e_k)\) is AQS relative to \(\mathbb{S}\).

On the contrary, suppose \(\hat{A}_k(e_k)\) is AQS relative to \(\mathbb{S}\). \(V_a - \hat{A}_k(e_k)\tilde{V}_a\tilde{A}_k^T(e_k) > 0\), \(\forall k \geq N_0\), \(\forall e_k \in \mathbb{S}\) can be obtained. In the above inequality, \(N_0 > 0\), \(V_a\) is also a constant PDM. \(\rho(\hat{A}_k(e_k)) < 1/\rho(\tilde{V}_a^{1/2}) \rho(\tilde{V}_a^{1/2})\). In this inequality, \(\rho(*)\) means the maximum singular value of a matrix. Conversely, \(\rho(*)\) indicates the minimum one. Furthermore, \(\hat{A}_k(e_k)\) and \(F_k(e_k)\) are bounded. Define matrix and scalars,

\[
\tilde{V}_a = [I + F_k(0)]V_a[I + F_k(0)]^T,
\]

\[
\mu_0 = \inf_{|\rho_a| \leq \rho_{a_{\min}}} \{ V_a - F_k(e_k)V_aF_k^T(e_k) \}
\]

\[
\mu_1 = \tilde{\rho}(\tilde{V}_a)
\]

\[
\mu_2 = \sup_{|\rho_a| \leq \tilde{\rho}(F_k(e_k))} \rho(\hat{A}_k(e_k))
\]

\[
\mu_3 = \sup_{|\rho_a| \leq \tilde{\rho}(\hat{A}_k(e_k))} \rho(\hat{A}_k(e_k))
\]

\[
\mu = \mu_1\mu_2\mu_3/\mu_0
\]

Note that

\[
V_a - \hat{A}_k(e_k)\tilde{V}_a\hat{A}_k^T(e_k)
\]

\[
= [I + \Gamma_k(0)]\left[ V_a - \hat{A}_k(e_k)\tilde{V}_a\hat{A}_k^T(e_k) \right] [I + \Gamma_k(0)]^T
\]

Obviously, \(0 \leq \mu \leq \infty\) can be further obtained if \(\mu_0 > 0\). Moreover, \(\tilde{\rho}\left( \left( V_a - F_k(e_k)V_aF_k^T(e_k) \right)^{-1/2} F_k(e_k) \tilde{V}_a \tilde{V}_a^T(e_k) \right) \leq \mu\).
Besides, according to the stability of $\Lambda_0$, there always be a PDF $V_1$ that satisfies $V_f - A_f V_f A_f^T > \mu^2 I$. Define matrix $V = D^{-1} \left[ \bar{V}_a \ V_f \right] (D^{-1})^T$.

Then, it is obvious that $V > 0$, and

$$D \left[ V - \bar{A}_k(e_k) V \bar{A}_k^T(e_k) \right] D^T = \left[ \bar{V}_a \ V_f \right] \left[ \begin{array}{cc} F_k(e_k) & 0 \\ \Lambda_k(e_k) & A_f \end{array} \right] \left[ \begin{array}{c} F_k(e_k) \\ \Lambda_k(e_k) \end{array} \right]^T.$$

Since $V_a$, $V_f$ and $V$ are all independent of $e_k$, the matrix $\bar{A}_k(e_k)$ is AQS relative to $\mathcal{S}$. This proof is complete. $\square$

**C. Proof of Theorem 2: Part 2**

Based on $\bar{A}_k(e_k)$ is AQS relative to $\mathcal{S}$ and $\bar{C}_k(e_k)$ is bounded. Then, according to Lemma A.1, there is a positive bounded number $N_0$ and a time-invariant PDM $V_1$, which results in

$$V - \bar{A}_k(e_k) V \bar{A}_k^T(e_k) > 0, \forall k > N_0, \forall e_k > \mathcal{S}.$$ Define scalars

$$\mu_0 = \inf_{k \geq N_0} \{ \inf_{|x_k| \leq \lambda_{\min}} \{ V - \bar{A}_k(e_k) V \bar{A}_k^T(e_k) \} \},$$

$$\mu_1 = \sup_{k \geq N_0} \{ \sup_{|x_k| \leq \lambda} \bar{p} \left( B_k(e_k) \right) \},$$

$$\mu_2 = \max_{k \geq N_0} \{ \bar{p}(Q_k) \bar{p}(R_{k+1}) \},$$

in which $\lambda_{\min}(\cdot)$ stands as the minimum eigenvalue of a matrix.

According to the boundedness of $B_k(e_k)$, $Q_k$, $R_{k+1}$, it is obvious that $-\infty < \mu^2 \mu_2 / \mu_0 V < \infty$. Define matrix $\bar{V} = \mu^2 \mu_2 / \mu_0 V$, it can be straightforwardly proved that each $e_k \in \mathcal{S}$, $M - \bar{A}_k(e_k) M \bar{A}_k^T(e_k) - \bar{B}_k(e_k) \left[ \begin{array}{c} Q_k \\ R_{k+1} \end{array} \right] \bar{B}_k^T(e_k)$$

\geq 0$. That is, there exists $M \geq 0$, so that $M - \bar{A}_k(e_k)$

\begin{align*}
M_k &= E \left( \left[ \bar{X}_{k|k} \ ar{X}_{k|k}^T \right] \left[ \bar{X}_{k|k} \ ar{X}_{k|k}^T \right]^T \right). 
\end{align*}

Recall that $w_{k|k}^{\infty}$ and $v_{k|k}^{\infty}$ are uncorrelated white sequences, then according to (35),

$$M_{k+1} = \bar{A}_k(e_k) M_k \bar{A}_k^T(e_k) + \bar{B}_k(e_k) \left[ \begin{array}{c} Q_k \\ R_{k+1} \end{array} \right] \bar{B}_k^T(e_k).$$

Therefore,

$$M - M_{k+1} \geq \bar{A}_k(e_k) (M - M_k) \bar{A}_k^T(e_k).$$

Founded on the above inequality, the following inequality can be further derived.

$$V^{-1/2} [M - M_{k+1}] V^{-1/2} \geq \left( \prod_{i=N_0}^{k} V^{-1/2} \bar{A}_k(e_k) V^{-1/2} \right)^T V^{-1/2} [M - M_{N_0}] V^{-1/2}.$$ 

Besides, define

$$\beta = \sup_{k \geq N_0} \{ \sup_{|x_k| \leq \lambda} \bar{p} \left( V^{-1/2} \bar{A}_k(e_k) V^{-1/2} \right) \},$$

Then, it is obvious that $\beta < 1$ and

$$\bar{p} \left( \prod_{i=N_0}^{k} \left( V^{-1/2} \bar{A}_i(e_i) V^{1/2} \right) \right) \leq \beta^{k+1}.$$ 

So, $\lim_{k \to \infty} \prod_{i=N_0}^{k} \left( V^{-1/2} \bar{A}_i(e_i) V^{1/2} \right) = 0$. Furthermore, according to the equivalence relationship of the systems (1) and (3), $\bar{A}_k(e_k)$, $\bar{B}_k(e_k)$ are bounded, and the boundedness of $\bar{A}_k(e_k)$, $\bar{B}_k(e_k)$ can be further derived. Since the $N_0$ is bounded, $M_{N_0}$ is bounded. Because all elements of $M$ are bounded, $M_{N_0}$ is bounded. Because all elements of $M$ are bounded, $\lim_{k \to \infty} V^{-1/2} [M - M_{k+1}] V^{1/2} \geq 0$ can be obtained. That is, $\lim_{k \to \infty} M_k \leq M$. Note that, $X_{k|k} = [I + \Gamma_k(0)](\bar{X}_{k|k} - \bar{X}_{k|k})$ and a symmetric matrix is positive semi-definite if and only if all its principal and sub-forms are positive semi-definite. The proof is completed.