



# MMSE State Estimation Approach for Linear Discrete-Time Systems With Time-Delay and Multi-Error Measurements

Yao Sun, Fengshui Jing, Zize Liang, and Min Tan

**Abstract**—This technical note focuses on designing a novel optimal estimation approach for linear discrete-time systems with time-delay and multi-error measurements, inspired by the measurement demands of the Five-hundred-meter Aperture Spherical radio Telescope. Measurement errors from different measurement channels, which are considered equal in many previous estimation methods, are usually negatively correlated with time-delays. Measurements with higher time-delays usually have less errors than the ones with lower delays. Our approach improves the estimation accuracy by decreasing the usage rate of measurements with larger errors. To prove the optimality of the approach in the minimum mean square error sense, a derivation procedure is presented. A numerical example and comparison are also given to demonstrate the feasibility and advantage of the proposed approach.

**Index Terms**—Explicit time-delay, MMSE state estimation approach, multi-error measurements.

## I. INTRODUCTION

The Five-hundred-meter Aperture Spherical radio Telescope (FAST), which was proposed by astronomers from ten countries including China at the General Assembly of the International Union of Radio Science in 1993, is a Chinese mega-science project to build the most sensitive single dish radio telescope in the world. The active moving feed cabin of FAST requires high accuracy and real-time position estimation by multi-fold measurements [1]. GPS-RTK and total station, whose parameters are given in Table I, are the two selected measurement equipment for the feed cabin [2]–[5]. The required real-time measurement accuracy of position is  $\text{RMSE} \leq 3 \text{ mm}$  [6]. Neither of the two equipment can meet the demands: the measurement time-delay of GPS-RTK is approximately equal to zero, but the RMSE of its accuracy is overlarge; conversely, the measurement time-delay of total station is too great to ignore, although the RMSE is acceptable. Therefore, we describe the active moving feed cabin subsystem as a discrete-time system with time-delay and multi-error measurements, and our aim is to find an optimal estimation approach for it.

Optimal state estimation approaches for linear discrete-time systems have been widely studied and have found many practical applications in signal processing, control and communication systems [7]–[9].

Manuscript received March 30, 2016; accepted June 13, 2016. Date of publication June 15, 2016; date of current version February 24, 2017. This work was supported by the National R&D Infrastructure and Facility Development Program of China under Grant FAST/HT013-2014-CK-013 and by the National Natural Science Foundation of China under Grant 61573358. Recommended by Associate Editor W. X. Zheng.

The authors are with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, and also with University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: yao.sun@ia.ac.cn; fengshui.jing@ia.ac.cn; zize.liang@ia.ac.cn; min.tan@ia.ac.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TAC.2016.2581479

TABLE I  
DETAILED PARAMETERS OF GPS-RTK AND TOTAL STATION

Parameter	GPS-RTK	Total Station
Time-delay	$5 \times 10^{-5} \text{ ms}$	220ms
RMSE <sup>1</sup>	20mm	3mm

<sup>1</sup> Root-mean-square error.

Optimal state estimation is the minimum mean square error estimation, which is termed as MMSE state estimation [10]. A classical MMSE state estimation approach named Kalman filter has been well studied in the past decades [11]. With the development of multi-sensor data fusion technology and network control systems [12]–[16], time-delay and multi-error problems are becoming two focus issues in MMSE estimation. However, the standard Kalman filter cannot be applied to systems with time-delay and multi-error measurements.

Although several estimation methods including augmented state Kalman filter [17], [18] and re-organization observation Kalman filter [19]–[21] have been studied to solve the time-delay problem, previous optimal estimation methods for multi-measurement systems were studied based on a widely accepted assumption: measurement errors from multi-measurement channels with different delays were considered equal. This assumption leads to a heavy use of measurements with lower delay during the iterations. As shown in Table I, measurements with lower time-delay normally have larger measurement errors than the ones with higher delay, otherwise, there is little reason to use measurement equipment with higher delay. Nonobservance of the negative correlation between time-delay and measurement error results in accuracy degradation in the final results. Thus, the previous methods can hardly be used to solve the state estimation problem of FAST.

In this technical note, we design a novel MMSE estimation approach for linear discrete-time systems with time-delay and multi-error measurements. This approach uses innovation re-organization to deal with the explicit time-delay problem [19], [22]. Compared with the previous estimation methods, which use measurements with larger errors several times in each loop iteration, our approach uses this type of measurements only once in each loop iteration. By decreasing the usage rate of measurements with larger errors, our approach improves the final estimation accuracy and solves the multi-error problem.

The rest of this technical note is organized as follows. In Section II, we will give the problem formulation. The novel MMSE estimator is presented in Section III. In Section IV, we compare the proposed estimation approach with the traditional augmented state method and the re-organization observation method in [19], via a numerical example. Finally, some conclusions are drawn in Section V.

## II. PROBLEM FORMULATION

Inspired by the measurement demands of FAST, we consider the following linear discrete-time system:

$$x(t_{k+1}) = \Phi(t_k)x(t_k) + \Gamma(t_k)u(t_k) \quad (k = 0, 1, 2, \dots) \quad (1)$$

where  $t_k$  is an arbitrary discrete moment,  $\Phi(t_k) \in R^{q \times q}$  is the bounded time-varying transition matrix,  $\Gamma(t_k) \in R^{q \times r}$  is the

time-varying coefficient matrix,  $x(t_k) \in R^q$  is the system state vector and  $u(t_k) \in R^r$  is the system error. The system state is observed by two measurement channels with unequal time-delays, which are described as

$$y_0(t_k) = H_0(t_k)x(t_k^+) + v_0(t_k) \quad (t_k^+ = t_k - \lambda_0) \quad (2)$$

$$y_1(t_k) = H_1(t_k)x(t_k^-) + v_1(t_k) \quad (t_k^- = t_k - \lambda_1) \quad (3)$$

where  $y_0(t_k) \in R^{p_0}$  is the measurement vector with time-delay  $\lambda_0$ ,  $y_1(t_k) \in R^{p_1}$  is the measurement vector with time-delay  $\lambda_1$ ,  $H_0(t_k) \in R^{p_0 \times q}$  and  $H_1(t_k) \in R^{p_1 \times q}$  are bounded time-varying matrices,  $v_0(t_k) \in R^{p_0}$  and  $v_1(t_k) \in R^{p_1}$  are measurement errors. The above  $p_0, p_1, q, r \in N^+$ .

**Remark 1:** According to Table I,  $\lambda_0 \ll \lambda_1$  and  $\lambda_0 \ll t_{k+1} - t_k$ , thus we can ignore  $\lambda_0$ .

**Remark 2:** To simplify discussion, we note the time delta as  $t_{k+1} - t_k = 1$  and  $t_0 = 0$ .

**Assumption 1:** The noises  $u(t_k), v_0(t_k)$  and  $v_1(t_k)$  are mutually independent white noises with zero means and known covariance matrices

$$\text{Cov}[u(t_k), u(j)] = Q_u(t_k)\delta_{t_k,j}$$

$$\text{Cov}[v_0(t_k), v_0(j)] = Q_{v_0}(t_k)\delta_{t_k,j}$$

$$\text{Cov}[v_1(t_k), v_1(j)] = Q_{v_1}(t_k)\delta_{t_k,j}$$

where  $\delta_{t_k,j}$  is the Kronecker delta. Moreover, these noise variables and  $x(0)$  are mutually independent.

**Assumption 2:** The initial system state  $x(0)$  is subject to Gaussian distribution, i.e.,  $x(0) \sim N(\mu_x, \delta_x^2)$ . In this technical note, we assume  $\mu_x = 0$  for the purpose of mathematical tractability. The case of  $\mu_x \neq 0$  could be considered homoplastically.

**Assumption 3:**  $x(t)$  is the state that is to be estimated at  $t_k = t$ , we assume the given time  $t > \lambda_1$  in this technical note. The case of  $\lambda_0 < t \leq \lambda_1$  could be solved by the standard Kalman filter because there is only one available measurement channel without delay during that time. The case of  $t \leq \lambda_0$  cannot be solved because there is no available measurement during that time.

**Assumption 4:** As the measurements with lower time-delay have larger errors than the ones with higher delay, we assume  $Q_{v_0} > Q_{v_1}$ .

Let  $\hat{x}(t|t, t)$  denote the MMSE state estimation of  $x(t)$ , given the two measurement sequences  $\{y_0(0), y_0(1), \dots, y_0(t)\}$  and  $\{y_1(\lambda_1), y_1(\lambda_1 + 1), \dots, y_1(t)\}$  (denoted as  $\{\bar{y}_0(t)\}$  and  $\{\bar{y}_1(t)\}$ , respectively, for convenience). Our aim is to obtain  $\hat{x}(t|t, t)$  at any given time  $t$ , i.e., for any time  $t \in (\lambda_1, \infty)$ .

### III. CONSTRUCTION OF STATE ESTIMATOR

In this section, we design a novel MMSE state estimation approach for linear discrete-time system (1) with time-delay and multi-error measurements from measurement channels (2) and (3).

#### A. Observation Re-Organization

Let  $y(t)$  denote a consequential measurement from systems (2) and (3) at the given time  $t$  and  $v(t)$  denote the related measurement noise. Then we have

$$y(t_k) = \begin{cases} \left[ \begin{array}{c} y_0^T(t_k), \underbrace{0, \dots, 0}_{p_1} \end{array} \right]^T & (0 \leq t_k < \lambda_1) \\ \left[ \begin{array}{c} y_0^T(t_k), y_1^T(t_k) \end{array} \right]^T & (t_k \geq \lambda_1) \end{cases}$$

$$v(t_k) = \begin{cases} \left[ \begin{array}{c} v_0^T(t_k), \underbrace{0, \dots, 0}_{p_1} \end{array} \right]^T & (0 \leq t_k < \lambda_1) \\ \left[ \begin{array}{c} v_0^T(t_k), v_1^T(t_k) \end{array} \right]^T & (t_k \geq \lambda_1) \end{cases}$$

The MMSE state estimation of  $x(t)$ , given the consequential observation sequence  $\{y(t_k)|_{0 \leq t_k \leq t}\}$ , is  $\hat{x}(t|t, t)$ .

By re-organizing measurements, we want to eliminate the explicit time-delay item. At the given time  $t$ , our aim of re-organization is to produce a measurement sequence

$$\{\mathcal{Y}_2(t_k)|_{0 \leq t_k \leq t^-}; \mathcal{Y}_1(t_k)|_{t^- < t_k \leq t}\}$$

$$= \{\mathcal{Y}_2(0), \mathcal{Y}_2(1), \dots, \mathcal{Y}_2(t^-); \mathcal{Y}_1(t^- + 1), \dots, \mathcal{Y}_1(t)\} \quad (4)$$

without the explicit time-delay item.  $\forall t_k \in [0, t^-]$ ,  $\mathcal{Y}_2(t_k)$  denotes the re-organization measurement at this moment,  $\mathcal{H}_2(t_k)$  and  $\mathcal{V}_2(t_k)$  denote the re-organization measurement matrix and the re-organization noise, respectively.  $\forall t_k \in (t^-, t]$ ,  $\mathcal{Y}_1(t_k)$ ,  $\mathcal{H}_1(t_k)$  and  $\mathcal{V}_1(t_k)$  denote the analogous re-organization variables at this moment. The above variables are shown in (5)–(7)

$$\begin{cases} \mathcal{Y}_2(t_k) \triangleq \begin{bmatrix} y_0(t_k) \\ y_1(t_k + \lambda_1) \end{bmatrix} & (t_k \in [0, t^-]) \\ \mathcal{Y}_1(t_k) \triangleq [y_0(t_k)] & (t_k \in (t^-, t]) \end{cases} \quad (5)$$

$$\begin{cases} \mathcal{H}_2(t_k) \triangleq \begin{bmatrix} H_0(t_k) \\ H_1(t_k + \lambda_1) \end{bmatrix} & (t_k \in [0, t^-]) \\ \mathcal{H}_1(t_k) \triangleq [H_0(t_k)] & (t_k \in (t^-, t]) \end{cases} \quad (6)$$

$$\begin{cases} \mathcal{V}_2(t_k) \triangleq \begin{bmatrix} v_0(t_k) \\ v_1(t_k + \lambda_1) \end{bmatrix} & (t_k \in [0, t^-]) \\ \mathcal{V}_1(t_k) \triangleq [v_0(t_k)] & (t_k \in (t^-, t]) \end{cases} \quad (7)$$

where  $\mathcal{V}_2(t_k)$  and  $\mathcal{V}_1(t_k)$  are mutually uncorrelated white noises with zero means and known covariance matrices

$$\begin{cases} \mathcal{Q}_{v_2}(t_k) \triangleq \begin{bmatrix} Q_{v_0}(t_k) & 0 \\ 0 & Q_{v_1}(t_k + \lambda_1) \end{bmatrix} \\ \mathcal{Q}_{v_1}(t_k) \triangleq [Q_{v_0}(t_k)]. \end{cases}$$

**Definition 1:** Let  $\hat{x}_2(t^-|t^-)$  denote the MMSE state estimation of  $x(t^-)$  given the re-organization measurement sequence  $\{\mathcal{Y}_2(t_k)|_{0 \leq t_k \leq t^-}\}$  and  $\hat{x}_1(t|t)$  denote the MMSE state estimation of  $x(t)$  given the re-organization measurement sequence  $\{\mathcal{Y}_2(t_k)|_{0 \leq t_k \leq t^-}; \mathcal{Y}_1(t_k)|_{t^- < t_k \leq t}\}$ .

**Lemma 1:** The following equation holds:

$$\hat{x}(t|t, t) = \hat{x}_1(t|t).$$

*Proof:* Note that

$$\mathcal{Y}_2(t_k) = \begin{bmatrix} I_0 & O_1 \\ O_0 & 0 \cdot I_1 \end{bmatrix} y(t_k) + \begin{bmatrix} 0 \cdot I_0 & O_1 \\ O_0 & I_1 \end{bmatrix} y(t_k + \lambda_1) \quad (8)$$

$$\mathcal{Y}_1(t_k) = [I_0 \quad O_1] y(t_k) \quad (9)$$

where  $I_0 \in R^{p_0 \times p_0}$  and  $I_1 \in R^{p_1 \times p_1}$  are identity matrices,  $O_0 \in R^{p_1 \times p_0}$  and  $O_1 \in R^{p_0 \times p_1}$  are null matrices,  $t_k \in [0, t^-]$  in (8) and  $t_k \in (t^-, t]$  in (9). Correspondingly,  $y(t_k)$  can also be expressed by  $\mathcal{Y}_2(t_k)$  and  $\mathcal{Y}_1(t_k)$ . Formulas (8) and (9) show that the linear space spanned by sequence  $\{y(t_k)|_{0 \leq t_k \leq t}\}$  contains the same measurement information with the linear space spanned by sequence  $\{\mathcal{Y}_2(t_k)|_{0 \leq t_k \leq t^-}; \mathcal{Y}_1(t_k)|_{t^- < t_k \leq t}\}$ . According to the orthogonal projection theorem, the MMSE state estimation  $\hat{x}(t|t, t)$  is the projection of  $x(t)$  onto the linear space spanned by sequence  $\{y(t_k)|_{0 \leq t_k \leq t}\}$  at the given time  $t$  [23], [24]. Thus the MMSE estimation given  $\{\mathcal{Y}_2(t_k)|_{0 \leq t_k \leq t^-}; \mathcal{Y}_1(t_k)|_{t^- < t_k \leq t}\}$  equals the MMSE estimation given  $\{y(t_k)|_{0 \leq t_k \leq t}\}$ , i.e.,  $\hat{x}(t|t, t) = \hat{x}_1(t|t)$ .

The derivation of lemma 1 is complete. ■

### B. Innovation Re-Organisation

**Lemma 2:**  $\hat{\mathcal{X}}_1(t|t)$  can be expressed by formula (10), and  $\hat{\mathcal{X}}_2(t^-|t^-)$  can be expressed by formula (11)

$$\hat{\mathcal{X}}_1(t|t) = E[x(t)|\bar{\mathcal{Y}}_2(t^-); \bar{\mathcal{Y}}_1(t)] \quad (10)$$

$$\hat{\mathcal{X}}_2(t^-|t^-) = E[x(t^-)|\bar{\mathcal{Y}}_2(t^-)] \quad (11)$$

where  $E(\cdot)$  denotes the mathematical expectation of random vector,  $\bar{\mathcal{Y}}_1(t)$  and  $\bar{\mathcal{Y}}_2(t^-)$  have similar definitions to  $\bar{y}_0(t)$  and  $\bar{y}_1(t)$ .

*Proof:* First, define  $P_2(t^-|t^-)$  as the mean square error matrix between  $x(t^-)$  and  $\hat{\mathcal{X}}_2(t^-|t^-)$

$$P_2(t^-|t^-) = E \left[ \left( x(t^-) - \hat{\mathcal{X}}_2(t^-|t^-) \right) \cdot \left( x(t^-) - \hat{\mathcal{X}}_2(t^-|t^-) \right)^T \right].$$

Then it can be transformed into

$$\begin{aligned} P_2(t^-|t^-) &= E \left[ \left( x(t^-) - \hat{\mathcal{X}}_2(t^-|t^-) - E[x(t^-)|\bar{\mathcal{Y}}_2(t^-)] + E[x(t^-)|\bar{\mathcal{Y}}_2(t^-)] \right) \right. \\ &\quad \cdot \left. \left( x(t^-) - \hat{\mathcal{X}}_2(t^-|t^-) - E[x(t^-)|\bar{\mathcal{Y}}_2(t^-)] + E[x(t^-)|\bar{\mathcal{Y}}_2(t^-)] \right)^T \right]. \end{aligned}$$

In this way,  $P_2(t^-|t^-)$  becomes the expectation of joint distribution of random vectors  $\bar{\mathcal{Y}}_2(t^-), x(t^-)$ . Based on the Bayes formula,  $E_{\bar{\mathcal{Y}}_2, x}[\cdot] = E_{\bar{\mathcal{Y}}_2}[E_x|\bar{\mathcal{Y}}_2[\cdot]]$  [23]. Thus,  $P_2(t^-|t^-)$  satisfies

$$\begin{aligned} P_2(t^-) &= E_{\bar{\mathcal{Y}}_2} \left\{ E_{x|\bar{\mathcal{Y}}_2} \left\{ \left[ x(t^-) - E[x(t^-)|\bar{\mathcal{Y}}_2(t^-)] \right] \right. \right. \\ &\quad \cdot \left. \left[ x(t^-) - E[x(t^-)|\bar{\mathcal{Y}}_2(t^-)] \right]^T \right\} \\ &\quad + E_{x|\bar{\mathcal{Y}}_2} \left\{ \left[ x(t^-) - E[x(t^-)|\bar{\mathcal{Y}}_2(t^-)] \right] \right. \\ &\quad \cdot \left. \left[ E[x(t^-)|\bar{\mathcal{Y}}_2(t^-)] - \hat{\mathcal{X}}_2(t^-|t^-) \right]^T \right\} \\ &\quad + E_{x|\bar{\mathcal{Y}}_2} \left\{ \left[ E[x(t^-)|\bar{\mathcal{Y}}_2(t^-)] - \hat{\mathcal{X}}_2(t^-|t^-) \right] \right. \\ &\quad \cdot \left. \left[ x(t^-) - E[x(t^-)|\bar{\mathcal{Y}}_2(t^-)] \right]^T \right\} \\ &\quad + E_{x|\bar{\mathcal{Y}}_2} \left\{ \left[ E[x(t^-)|\bar{\mathcal{Y}}_2(t^-)] - \hat{\mathcal{X}}_2(t^-|t^-) \right] \right. \\ &\quad \cdot \left. \left[ E[x(t^-)|\bar{\mathcal{Y}}_2(t^-)] - \hat{\mathcal{X}}_2(t^-|t^-) \right]^T \right\}. \end{aligned}$$

Second, there are four items  $E_{x|\bar{\mathcal{Y}}_2} + E_{x|\bar{\mathcal{Y}}_2} + E_{x|\bar{\mathcal{Y}}_2} + E_{x|\bar{\mathcal{Y}}_2}$  in the above expression  $E_{\bar{\mathcal{Y}}_2}[\cdot]$ . The first item is a constant non-negative definite matrix [25]. The second item can be rewritten as

$$\begin{aligned} E_{x|\bar{\mathcal{Y}}_2} \left\{ \left[ x(t^-) - E[x(t^-)|\bar{\mathcal{Y}}_2(t^-)] \right] \right. \\ \cdot \left. \left[ E[x(t^-)|\bar{\mathcal{Y}}_2(t^-)] - \hat{\mathcal{X}}_2(t^-|t^-) \right]^T \right\}. \end{aligned}$$

Because of  $E_{x|\bar{\mathcal{Y}}_2} \{ x(t^-) - E[x(t^-)|\bar{\mathcal{Y}}_2(t^-)] \} = 0$ , the second item is 0. For the same reason, the third item is 0. As  $E[x(t^-)|\bar{\mathcal{Y}}_2(t^-)]$  is a constant vector, the fourth item is a non-negative definite matrix and a function matrix of  $\hat{\mathcal{X}}_2(t^-|t^-)$ . For this reason,  $P_2(t^-|t^-)$  is a non-negative definite matrix and the lower bound depends on  $\hat{\mathcal{X}}_2(t^-|t^-)$ . To minimize  $P_2(t^-|t^-)$ , we have to make the following equation hold:

$$\begin{aligned} E_{x|\bar{\mathcal{Y}}_2} \left[ \left( E[x(t^-)|\bar{\mathcal{Y}}_2(t^-)] - \hat{\mathcal{X}}_2(t^-|t^-) \right) \right. \\ \cdot \left. \left( E[x(t^-)|\bar{\mathcal{Y}}_2(t^-)] - \hat{\mathcal{X}}_2(t^-|t^-) \right)^T \right] = 0 \end{aligned}$$

which means  $\hat{\mathcal{X}}_2(t^-|t^-) = E[x(t^-)|\mathcal{Y}_2(0), \mathcal{Y}_2(1), \dots, \mathcal{Y}_2(t^-)]$ .

Here, formula (11) has been derived. In a similar way, (10) can be derived.

Thus the derivation of lemma 2 is complete.  $\blacksquare$

Next, we will give theorem 1 to calculate  $\hat{\mathcal{X}}_2(t^-|t^-)$  and  $\hat{\mathcal{X}}_1(t|t)$  at any time  $t$ . Theorem 1 is derived by reorganizing the innovations of  $\mathcal{Y}_2(t^-)$  and  $\mathcal{Y}_1(t)$ .

**Definition 2:** The innovation of re-organization measurement  $\mathcal{Y}_2(t^-)$  is defined by

$$\tilde{\mathcal{Y}}_2(t^-) \triangleq \mathcal{Y}_2(t^-) - \hat{\mathcal{Y}}_2(t^-|t^- - 1)$$

where  $\hat{\mathcal{Y}}_2(t^-|t^- - 1)$  is the MMSE prediction of  $\mathcal{Y}_2(t^-)$ , given the measurement subsequence  $\{\mathcal{Y}_2(0), \mathcal{Y}_2(1), \dots, \mathcal{Y}_2(t^- - 1)\}$  [17].  $\tilde{\mathcal{Y}}_1(t)$  has a similar definition.

**Theorem 1:**  $\hat{\mathcal{X}}_1(t|t)$  and  $\hat{\mathcal{X}}_2(t^-|t^-)$  follow by lemma 2. At the given time  $t$ ,  $\hat{\mathcal{X}}_1(t|t)$  can be calculated via formulas (12)–(15) in the case of  $\varepsilon = 1$  and  $\tau = t$ , the corresponding intermediate estimation result  $\hat{\mathcal{X}}_2(t^-|t^-)$  can be calculated via formulas (12)–(15) in the case of  $\varepsilon = 2$  and  $\tau = t^-$ :

$$\hat{\mathcal{X}}_\varepsilon(\tau|\tau) = \Phi(\tau-1)\hat{\mathcal{X}}_\varepsilon(\tau-1|\tau-1) + E[x(\tau)|\tilde{\mathcal{Y}}_\varepsilon(\tau)] \quad (12)$$

$$\begin{aligned} E[x(\tau)|\tilde{\mathcal{Y}}_\varepsilon(\tau)] &= [\Phi(\tau-1)P_\varepsilon(\tau-1|\tau-1)\Phi^T(\tau-1) \\ &\quad + \Gamma(\tau-1)Q_u(\tau-1)\Gamma^T(\tau-1)] \\ &\quad \cdot \mathcal{H}_\varepsilon^T(\tau)M_\varepsilon^{-1}(\tau) \\ &\quad \cdot [\mathcal{Y}_\varepsilon(\tau) - \mathcal{H}_\varepsilon(\tau)\Phi(\tau-1)\hat{\mathcal{X}}_\varepsilon(\tau-1|\tau-1)] \quad (13) \end{aligned}$$

$$\begin{aligned} P_\varepsilon(\tau|\tau) &= [\Phi(\tau-1)P_\varepsilon(\tau-1|\tau-1)\Phi^T(\tau-1) \\ &\quad + \Gamma(\tau-1)Q_u(\tau-1)\Gamma^T(\tau-1)] \\ &\quad \cdot \{I - \mathcal{H}_\varepsilon^T(\tau)M_\varepsilon^{-1}(\tau)\mathcal{H}_\varepsilon(\tau) \\ &\quad \cdot [\Phi(\tau-1)P_\varepsilon(\tau-1|\tau-1)\Phi^T(\tau-1) \\ &\quad + \Gamma(\tau-1)Q_u(\tau-1)\Gamma^T(\tau-1)]\} \quad (14) \end{aligned}$$

$$\begin{aligned} M_\varepsilon(\tau) &= \mathcal{H}_\varepsilon(\tau) \cdot [\Phi(\tau-1)P_\varepsilon(\tau-1|\tau-1)\Phi^T(\tau-1) \\ &\quad + \Gamma(\tau-1)Q_u(\tau-1)\Gamma^T(\tau-1)] \\ &\quad \cdot \mathcal{H}_\varepsilon^T(\tau) + \mathcal{Q}_{v_\varepsilon}(\tau) \quad (15) \end{aligned}$$

where  $P_\varepsilon(\tau|\tau)$  is the mean square error matrix of estimation as we have defined before, satisfying  $\hat{\mathcal{X}}_2(0|0) = \mu_x$  and  $P_2(0|0) = \delta_x^2$ .

*Proof:* First, because  $\tilde{\mathcal{Y}}_2(t^-|t^- - 1)$  is linearly correlated with the re-organization measurement sequence  $\{\bar{\mathcal{Y}}_2(t^- - 1)\}$  [23], it can be expressed by  $\hat{\mathcal{Y}}_2(t^-|t^- - 1) = \sum_{\varphi=0}^{t^- - 1} d_\varphi \cdot \mathcal{Y}_2(\varphi)$ , where  $\varphi \in N$ ,  $d_\varphi$  is the optimal weighting coefficient. So definition 2 is written as

$$\mathcal{Y}_2(t^-) = \tilde{\mathcal{Y}}_2(t^-) + \sum_{\varphi=0}^{t^- - 1} d_\varphi \cdot \mathcal{Y}_2(\varphi). \quad (16)$$

This shows that measurement information of  $\mathcal{Y}_2(t^-)$  can be restored via formula (16), i.e.,  $\{\bar{\mathcal{Y}}_2(t^-)\}$  and  $\{\bar{\mathcal{Y}}_2(t^- - 1), \tilde{\mathcal{Y}}_2(t^-)\}$  contain the same measurement information, where  $\tilde{\mathcal{Y}}_2(t^-)$  and  $\{\bar{\mathcal{Y}}_2(t^- - 1)\}$  are uncorrelated [23]. Thus formula (11) is transformed into  $\hat{\mathcal{X}}_2(t^-|t^-) = E[x(t^-)|\bar{\mathcal{Y}}_2(t^- - 1)] + E[x(t^-)|\tilde{\mathcal{Y}}_2(t^-)]$ . Based on assumption 1,  $u(t^- - 1)$  is uncorrelated with  $\{\mathcal{Y}_2(0), \dots, \mathcal{Y}_2(t^- - 1)\}$ . Thus we can get  $E[x(t^-)|\bar{\mathcal{Y}}_2(t^- - 1)] = \Phi(t^- - 1)\hat{\mathcal{X}}_2(t^- - 1|t^- - 1)$ .

The derivation of formula (12) is complete. Meanwhile, we define  $\hat{\mathcal{X}}_2(t^-|t^- - 1) \triangleq E[x(t^-)|\bar{\mathcal{Y}}_2(t^- - 1)]$  for convenience.

Second, according to the multidimensional Gaussian probability density function theorem:  $E(r_1|r_2) = E(r_1) + \text{Cov}(r_1, r_2) \cdot \text{Var}^{-1}(r_2, r_2) \cdot [r_2 - E(r_2)]$  [23], where random vectors  $r_1$  and  $r_2$  are

subject to joint Gaussian distribution,  $\text{Cov}()$  is the covariance matrix,  $\text{Var}()$  is the variance matrix, the following holds:

$$\begin{aligned} E[x(t^-)|\tilde{\mathcal{Y}}_2(t^-)] &= E[x(t^-)] + \text{Cov}[x(t^-), \tilde{\mathcal{Y}}_2(t^-)] \\ &\quad \cdot \text{Var}^{-1}[\tilde{\mathcal{Y}}_2(t^-), \tilde{\mathcal{Y}}_2(t^-)] \cdot \{\tilde{\mathcal{Y}}_2(t^-) - E[\tilde{\mathcal{Y}}_2(t^-)]\}. \end{aligned}$$

Because of  $E[x(t^-)] = 0$  and  $E[\tilde{\mathcal{Y}}_2(t^-)] = 0$ , the above expression is transformed into

$$\begin{aligned} E[x(t^-)|\tilde{\mathcal{Y}}_2(t^-)] &= \text{Cov}[x(t^-), \tilde{\mathcal{Y}}_2(t^-)] \\ &\quad \cdot \text{Var}^{-1}[\tilde{\mathcal{Y}}_2(t^-), \tilde{\mathcal{Y}}_2(t^-)] \cdot [\mathcal{Y}_2(t^-) - \hat{\mathcal{Y}}_2(t^-|t^- - 1)] \quad (17) \end{aligned}$$

where

$$\begin{aligned} \text{Cov}[x(t^-), \tilde{\mathcal{Y}}_2(t^-)] &= E[x(t^-)\tilde{\mathcal{Y}}_2^T(t^-)] \\ &= \left\{ \Phi(t^- - 1)E\left[\left(x(t^- - 1) - \hat{\mathcal{X}}_2(t^- - 1|t^- - 1)\right)\right.\right. \\ &\quad \cdot \left.\left(x(t^- - 1) - \hat{\mathcal{X}}_2(t^- - 1|t^- - 1)\right)^T\right] \\ &\quad \cdot \Phi^T(t^- - 1) + \Gamma(t^- - 1)Q_u(t^- - 1)\Gamma^T(t^- - 1) \Big\} \\ &\quad \cdot \mathcal{H}_2^T(t^-) \end{aligned}$$

$$\begin{aligned} \text{Var}[\tilde{\mathcal{Y}}_2(t^-), \tilde{\mathcal{Y}}_2(t^-)] &= E[\tilde{\mathcal{Y}}_2(t^-)\tilde{\mathcal{Y}}_2^T(t^-)] \\ &= \mathcal{H}_2(t^-) \left\{ \Phi(t^- - 1)E\left[\left(x(t^- - 1) - \hat{\mathcal{X}}_2(t^- - 1|t^- - 1)\right)\right.\right. \\ &\quad \cdot \left.\left(x(t^- - 1) - \hat{\mathcal{X}}_2(t^- - 1|t^- - 1)\right)^T\right] \\ &\quad \cdot \Phi^T(t^- - 1) + \Gamma(t^- - 1)Q_u(t^- - 1)\Gamma^T(t^- - 1) \Big\} \\ &\quad \times \mathcal{H}_2^T(t^-) + \mathcal{Q}_{v_2}(t^-). \end{aligned}$$

Note the following three useful equalities:

$$\begin{aligned} E\left\{\left[x(t^-) - \hat{\mathcal{X}}_2(t^-|t^- - 1)\right]\tilde{\mathcal{Y}}_2^T(t^-)\right\} &= E[x(t^-)\tilde{\mathcal{Y}}_2^T(t^-)] \\ E\left\{\mathcal{V}_2(t^-)\left[x(t^-) - \hat{\mathcal{X}}_2(t^-|t^- - 1)\right]^T\right\} &= 0 \\ \hat{\mathcal{Y}}_2(t^-|t^- - 1) &= \mathcal{H}_2(t^-)\hat{\mathcal{X}}_2(t^-|t^- - 1) \end{aligned}$$

as  $\tilde{\mathcal{Y}}_2(t^-)$  is uncorrelated with  $\hat{\mathcal{X}}_2(t^-|t^- - 1)$ ,  $\mathcal{V}_2(t^-)$  is uncorrelated with  $\hat{\mathcal{X}}_2(t^-|t^- - 1)$  and  $x(t^-)$ ,  $\mathcal{V}_2(t^-)$  is uncorrelated with  $\{\mathcal{Y}_2(0), \dots, \mathcal{Y}_2(t^- - 1)\}$ . Then we define  $M_2(t^-)$  as the variance matrix of  $\tilde{\mathcal{Y}}_2(t^-)$  and get

$$\begin{aligned} E[x(t^-)|\tilde{\mathcal{Y}}_2(t^-)] &= [\Phi(t^- - 1)P_2(t^- - 1|t^- - 1)\Phi^T(t^- - 1) \\ &\quad + \Gamma(t^- - 1)Q_u(t^- - 1)\Gamma^T(t^- - 1)] \cdot \mathcal{H}_2^T(t^-)M_2^{-1}(t^-) \\ &\quad \cdot [\mathcal{Y}_2(t^-) - \mathcal{H}_2(t^-)\Phi(t^- - 1)\hat{\mathcal{X}}_2(t^- - 1|t^- - 1)] \quad (18) \end{aligned}$$

where

$$M_2(t^-) = \mathcal{H}_2(t^-)[\Phi(t^- - 1)P_2(t^- - 1|t^- - 1)\Phi^T(t^- - 1) \\ + \Gamma(t^- - 1)Q_u(t^- - 1)\Gamma^T(t^- - 1)]\mathcal{H}_2^T(t^-) + \mathcal{Q}_{v_2}(t^-). \quad (19)$$

Here, the derivation of formula (13) and (15) is complete.

Finally, we need to know the iteration expression of  $P_2(t^-|t^-)$ . According to the definition of  $P_2(t^-|t^-)$ :  $P_2(t^-|t^-) = E[(x(t^-) - \hat{\mathcal{X}}_2(t^-|t^-))(x(t^-) - \hat{\mathcal{X}}_2(t^-|t^-))^T]$ , we can get

$$\begin{aligned} P_2(t^-|t^-) &= E\left[\left(x(t^-) - \hat{\mathcal{X}}_2(t^-|t^- - 1)\right)\left(x(t^-) - \hat{\mathcal{X}}_2(t^-|t^- - 1)\right)^T\right] \\ &\quad - E\left[G_2(t^-)\left(\mathcal{Y}_2(t^-) - \mathcal{H}_2(t^-)\hat{\mathcal{X}}_2(t^-|t^- - 1)\right)\right. \\ &\quad \cdot \left.\left(x(t^-) - \hat{\mathcal{X}}_2(t^-|t^- - 1)\right)^T\right] \\ &\quad - E\left[\left(x(t^-) - \hat{\mathcal{X}}_2(t^-|t^- - 1)\right)\right. \\ &\quad \cdot \left.\left(\mathcal{Y}_2(t^-) - \mathcal{H}_2(t^-)\hat{\mathcal{X}}_2(t^-|t^- - 1)\right)^T G_2^T(t^-)\right] \\ &\quad + E\left[G_2(t^-)\left(\mathcal{Y}_2(t^-) - \mathcal{H}_2(t^-)\hat{\mathcal{X}}_2(t^-|t^- - 1)\right)\right. \\ &\quad \cdot \left.\left(\mathcal{Y}_2(t^-) - \mathcal{H}_2(t^-)\hat{\mathcal{X}}_2(t^-|t^- - 1)\right)^T G_2^T(t^-)\right] \quad (20) \end{aligned}$$

where  $G_2(t^-) \triangleq E[x(t^-)\tilde{\mathcal{Y}}_2^T(t^-)] \cdot E^{-1}[\tilde{\mathcal{Y}}_2(t^-)\tilde{\mathcal{Y}}_2^T(t^-)]$ . Because  $G_2(t^-)$  is the product of 2 mathematical expectations, it is also a mathematical expectation. Therefore, the following three hold:

$$\begin{aligned} E\left[G_2(t^-) \cdot \left(\mathcal{Y}_2(t^-) - \mathcal{H}_2(t^-)\hat{\mathcal{X}}_2(t^-|t^- - 1)\right)\right. \\ \cdot \left.\left(x(t^-) - \hat{\mathcal{X}}_2(t^-|t^- - 1)\right)^T\right] \\ = E\left[\left(x(t^-) - \hat{\mathcal{X}}_2(t^-|t^- - 1)\right)\right. \\ \cdot \left.\left(x(t^-) - \hat{\mathcal{X}}_2(t^-|t^- - 1)\right)^T\right] \cdot \mathcal{H}_2^T(t^-) \cdot G_2^T(t^-) \quad (21) \end{aligned}$$

$$\begin{aligned} E\left[\left(x(t^-) - \hat{\mathcal{X}}_2(t^-|t^- - 1)\right)\right. \\ \cdot \left.\left(\mathcal{Y}_2(t^-) - \mathcal{H}_2(t^-)\hat{\mathcal{X}}_2(t^-|t^- - 1)\right)^T \cdot G_2^T(t^-)\right] \\ = E\left[\left(x(t^-) - \hat{\mathcal{X}}_2(t^-|t^- - 1)\right)\right. \\ \cdot \left.\left(x(t^-) - \hat{\mathcal{X}}_2(t^-|t^- - 1)\right)^T\right] \cdot \mathcal{H}_2^T(t^-) \cdot G_2^T(t^-) \quad (22) \end{aligned}$$

$$\begin{aligned} E\left[G_2(t^-) \cdot \left(\mathcal{Y}_2(t^-) - \mathcal{H}_2(t^-)\hat{\mathcal{X}}_2(t^-|t^- - 1)\right)\right. \\ \cdot \left.\left(\mathcal{Y}_2(t^-) - \mathcal{H}_2(t^-)\hat{\mathcal{X}}_2(t^-|t^- - 1)\right)^T \cdot G_2^T(t^-)\right] \\ = E\left[\left(x(t^-) - \hat{\mathcal{X}}_2(t^-|t^- - 1)\right)\right. \\ \cdot \left.\left(x(t^-) - \hat{\mathcal{X}}_2(t^-|t^- - 1)\right)^T\right] \cdot \mathcal{H}_2^T(t^-) \cdot G_2^T(t^-). \quad (23) \end{aligned}$$

Also, according to formula (17) and the definition of  $G_2(t^-)$ , it is known  $G_2(t^-) = E[x(t^-)|\tilde{\mathcal{Y}}_2(t^-)] \cdot [\mathcal{Y}_2(t^-) - \hat{\mathcal{Y}}_2(t^-|t^- - 1)]^{-1}$ . So we have a very simple expression of  $G_2(t^-)$ :

$$\begin{aligned} G_2(t^-) &= [\Phi(t^- - 1)P_2(t^- - 1|t^- - 1)\Phi^T(t^- - 1) \\ &\quad + \Gamma(t^- - 1)Q_u(t^- - 1)\Gamma^T(t^- - 1)] \cdot \mathcal{H}_2^T(t^-)M_2^{-1}(t^-). \quad (24) \end{aligned}$$

Next, in view of that  $M_2(t^-)$ ,  $P_2(t^- - 1|t^- - 1)$ ,  $Q_u(t^- - 1)$  are all variance matrices, i.e., they are all symmetrical matrices, we can substitute formulas (21)–(24) into (20)

$$\begin{aligned} P_2(t^-|t^-) &= E \left[ \left( x(t^-) - \hat{x}_2(t^-|t^- - 1) \right) \right. \\ &\quad \cdot \left. \left( x(t^-) - \hat{x}_2(t^-|t^- - 1) \right)^T \right] \\ &\quad \cdot (I - \mathcal{H}_2^T(t^-) \cdot G_2^T(t^-)) \\ &= [\Phi(t^- - 1)P_2(t^- - 1|t^- - 1)\Phi^T(t^- - 1) \\ &\quad + \Gamma(t^- - 1)Q_u(t^- - 1)\Gamma^T(t^- - 1)] \\ &\quad \cdot \{I - \mathcal{H}_2^T(t^-)M_2^{-1}(t^-)\mathcal{H}_2(t^-) \\ &\quad \cdot [\Phi(t^- - 1)P_2(t^- - 1|t^- - 1)\Phi^T(t^- - 1) \\ &\quad + \Gamma(t^- - 1)Q_u(t^- - 1)\Gamma^T(t^- - 1)]\}. \quad (25) \end{aligned}$$

Now, the derivation of formula (14) is complete. And the result occurs as a recursion relation.

In the case of  $\varepsilon = 2$  and  $\tau = t^-$ , formulas (12)–(15) have been derived. Also, in the case of  $\varepsilon = 1$  and  $\tau = t$ , formulas (12)–(15) can be derived in a similar way.  $\blacksquare$

Therefore, the derivation of theorem 1 is complete.  $\blacksquare$

### C. MMSE State Estimator

At any given time  $t$ ,  $t = \lambda_1 + 1, \lambda_1 + 2, \dots$  according to remark 2 and assumption 3, we obtain two re-organization measurements  $\mathcal{Y}_2(t^-)$  and  $\mathcal{Y}_1(t)$ . Consider the system (1)–(3). The optimal state estimation  $\hat{x}(t|t, t)$  is given via lemma 1

$$\hat{x}(t|t, t) = \hat{x}_1(t|t)$$

where  $\hat{x}_1(t|t)$  is calculated recursively via theorem 1 based on the current measurement  $\mathcal{Y}_1(t)$  and  $\hat{x}_1(t-1|t-1)$ ,  $\hat{x}_1(t-1|t-1)$  is calculated by the following steps:

- Step 1: Calculate  $\hat{x}_2(t^-|t^-)$  via theorem 1 based on the current measurement  $\mathcal{Y}_2(t^-)$  and  $\hat{x}_2(t^- - 1|t^- - 1)$ .
- Step 2: Calculate  $\hat{x}_1(t^- + \alpha|t^- + \alpha)$  via formulas (12) and (13),  $\alpha = 1, 2, \dots, \lambda_1 - 1$ :

$$\begin{aligned} \hat{x}_1(t^- + 1|t^- + 1) &= \mathcal{F}[\hat{x}_1(t^-|t^-), \mathcal{Y}_1(t^- + 1)] \\ &\vdots \\ \hat{x}_1(t-1|t-1) &= \mathcal{F}[\hat{x}_1(t^- + \lambda_1 - 2|t^- + \lambda_1 - 2), \mathcal{Y}_1(t-1)] \end{aligned}$$

where  $t-1=t^-+(\lambda_1-1)$ ,  $\hat{x}_1(t^-|t^-)=\hat{x}_2(t^-|t^-)$ ,  $\mathcal{Y}_1(t^-+\alpha)$  is the past measurement, and  $\mathcal{F}[\hat{x}_1(t^-|t^-), \mathcal{Y}_1(t^- + 1)]$  means  $\hat{x}_1(t^-|t^-)$  and  $\mathcal{Y}_1(t^- + 1)$  are used to calculate the result.

In step 2, measurements with larger errors, i.e.,  $\mathcal{Y}_1(\cdot)$ , are used  $\lambda_1 - 1$  times to calculate the intermediate estimations similar to some previous estimators. This reduces the estimation accuracy of the final result. As shown in assumption 4, measurements with lower time-delay usually have larger errors than the ones with higher delay; otherwise, there is little reason to use a measurement system with higher time-delay. To improve the estimation accuracy, we design a new iterator by decreasing the usage rate of measurements with larger errors, i.e., modifying the equalities in step 2.

$\forall \alpha \in [1, \lambda_1 - 1]$ , consider the following equality:

$$\begin{aligned} \hat{\mathcal{Y}}_1(t^- + \alpha|t^- + \alpha) &= E\{[\mathcal{H}_1(t^- + \alpha)x(t^- + \alpha) + \mathcal{Y}_1(t^- + \alpha)]|\overline{\mathcal{Y}_2}(t^-); \overline{\mathcal{Y}_1}(t^- + \alpha)\} \\ &= \mathcal{H}_1(t^- + \alpha)E[x(t^- + \alpha)|\overline{\mathcal{Y}_2}(t^-); \overline{\mathcal{Y}_1}(t^- + \alpha)] \\ &\quad + E[\mathcal{Y}_1(t^- + \alpha)|\overline{\mathcal{Y}_2}(t^-); \overline{\mathcal{Y}_1}(t^- + \alpha)]. \end{aligned}$$

Because  $\mathcal{V}_1(t^- + \alpha)$  is uncorrelated with  $\mathcal{Y}_2(t_k)|_{t_k=0, \dots, t^-}$  and  $\mathcal{Y}_1(t_k)|_{t_k=t^-, \dots, t^- + \alpha - 1}$ , the above equality can be rewritten as

$$\begin{aligned} &\mathcal{H}_1(t^- + \alpha)E[x(t^- + \alpha)|\overline{\mathcal{Y}_2}(t^-); \overline{\mathcal{Y}_1}(t^- + \alpha)] \\ &\quad + E[\mathcal{V}_1(t^- + \alpha)|\overline{\mathcal{Y}_2}(t^-); \overline{\mathcal{Y}_1}(t^- + \alpha)] \\ &= \mathcal{H}_1(t^- + \alpha)\hat{x}_1(t^- + \alpha|t^- + \alpha) + E[\mathcal{V}_1(t^- + \alpha)|\widetilde{\mathcal{Y}_1}(t^- + \alpha)]. \end{aligned}$$

$E[\mathcal{V}_1(t^- + \alpha)|\widetilde{\mathcal{Y}_1}(t^- + \alpha)]$  is the conditional mathematical expectation, it is a constant value in ideal conditions. Therefore, a new iterator has been obtained.  $\hat{x}_1(t^- + \alpha|t^- + \alpha)$  can be calculated recursively by this iterator, and thus the usage rate of the measurements with larger errors is decreased

$$\begin{aligned} &\hat{x}_1(t^- + \alpha|t^- + \alpha) \\ &= \{I - [\Phi(t^- + \alpha - 1)P_1(t^- + \alpha - 1|t^- + \alpha - 1)\Phi^T(t^- + \alpha - 1) \\ &\quad + \Gamma(t^- + \alpha - 1)Q_u(t^- + \alpha - 1)\Gamma^T(t^- + \alpha - 1)] \\ &\quad \cdot \mathcal{H}_1^T(t^- + \alpha)M_1^{-1}(t^- + \alpha)\mathcal{H}_1(t^- + \alpha)\}^{-1} \\ &\quad \cdot \{\Phi(t^- + \alpha - 1)\hat{x}_1(t^- + \alpha - 1|t^- + \alpha - 1) \\ &\quad - [\Phi(t^- + \alpha - 1)P_1(t^- + \alpha - 1|t^- + \alpha - 1)\Phi^T(t^- + \alpha - 1) \\ &\quad + \Gamma(t^- + \alpha - 1)Q_u(t^- + \alpha - 1)\Gamma^T(t^- + \alpha - 1)] \\ &\quad \cdot \mathcal{H}_1^T(t^- + \alpha)M_1^{-1}(t^- + \alpha)\mathcal{H}_1(t^- + \alpha) \\ &\quad \cdot \Phi(t^- + \alpha - 1)\hat{x}_1(t^- + \alpha - 1|t^- + \alpha - 1)\} \quad (26) \end{aligned}$$

where  $P_1(t^- + \alpha - 1|t^- + \alpha - 1)$  and  $M_1(t^- + \alpha)$  can be calculated via formulas (14) and (15). Then we will obtain the MMSE state estimation  $\hat{x}_1(t-1|t-1)$  during the iterations. Iterators in previous methods use measurements with larger error in each loop iteration because of the nonobservance of the negative correlation between time-delay and measurement error. Compared with these iterators, we estimate  $\hat{x}_1(t^- + \alpha|t^- + \alpha)$  without using any measurements with larger errors, i.e.,  $\mathcal{Y}_1(t_k)$ , during the iterations. This contributes to improving the final estimation accuracy of our estimator.

Thus at any given time  $t$ ,  $t = \lambda_1 + 1, \lambda_1 + 2, \dots$  according to remark 2 and assumption 3, the state estimation result  $\hat{x}(t|t, t)$  of system (1)–(3) calculated by our new MMSE estimator is given as the following algorithm:

---

#### Algorithm 1 Procedure of MMSE state estimator

---

- 1: Set  $t = \lambda_1 + 1$  as the initial time,  $\mu_x$  as the initial state estimation, and  $\delta_x^2$  as the initial mean square error matrix.
  - 2: At the time  $t$ , obtain two current re-organization measurements  $\mathcal{Y}_2(t^-)$  and  $\mathcal{Y}_1(t)$  via formula (5).
  - 3: Calculate  $\hat{x}_2(t^-|t^-)$  via theorem 1 in the case of  $\varepsilon = 2, \tau = t^-$ , based on  $\hat{x}_2(t^- - 1|t^- - 1)$  and  $\mathcal{Y}_2(t^-)$ .
  - 4: Set  $\hat{x}_1(t^-|t^-) = \hat{x}_2(t^-|t^-)$  and  $P_1(t^-|t^-) = P_2(t^-|t^-)$ .
  - 5: Estimate  $\hat{x}_1(t_k|t_k)|_{t_k=t^- + \alpha}$  recursively via iterator (26) until  $t_k = t - 1$ , where  $\alpha = 1, 2, \dots, \lambda_1 - 1$ . Meanwhile, calculate  $P_1(t_k|t_k)|_{t_k=t^- + \alpha}$  during the iterations via formula (14).
  - 6: Calculate  $\hat{x}_1(t|t)$  via theorem 1 in the case of  $\varepsilon = 1, \tau = t$ , based on  $\hat{x}_1(t-1|t-1)$  and  $\mathcal{Y}_1(t)$ .
  - 7: Obtain  $\hat{x}(t|t, t)$  because  $\hat{x}(t|t, t) = \hat{x}_1(t|t)$ .
  - 8: Set  $t = t + 1$ , and return to step 2.
- 

In contrast with the augmented state Kalman filter in [17], [18] and the method in [19], our novel approach improves the estimation iterators and reduces the usage rate of measurements with larger errors to enhance the final estimation accuracy.

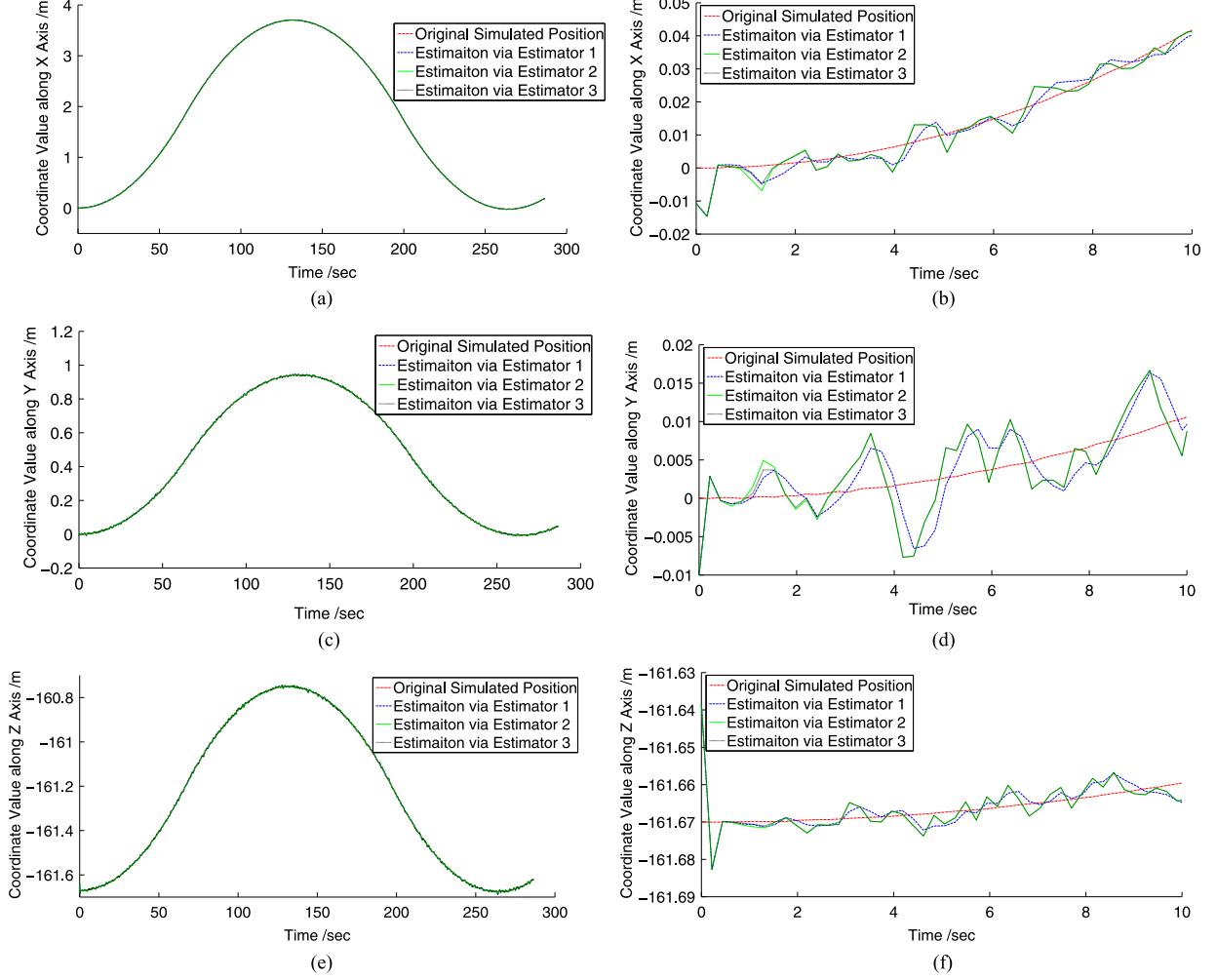


Fig. 1. Position estimation curves along three coordinate axes. (a), (c), (e): complete estimation curves along  $X$ ,  $Y$ ,  $Z$  coordinate axes, respectively. (b), (d), (f): partial estimation curves along three coordinate axes during the time  $[0, 10]$  sec (full size images are shown in supplementary material).

#### IV. NUMERICAL EXAMPLE AND ANALYSIS

In this section, a numerical example is given to simulate a motion of the feed cabin of FAST, and the position of the feed cabin is estimated by our novel method. In addition, to show the efficiency of our method, it is compared with the traditional augmented state Kalman filter and the method proposed in [19]. The original simulated positions of the feed cabin of FAST with the estimated positions along  $X$ ,  $Y$  and  $Z$  coordinate axes are depicted in Fig. 1(a), (c) and (e), respectively. In the rest of the section, we will show the application process of our method.

**Remark 3:** For convenience, Estimator 1–3 represent our MMSE estimation approach, the augmented state Kalman filter and the approach in [19], respectively.

At any given time  $t$ , the position of feed cabin can be expressed by

$$\begin{aligned} p(t) = p(t - \Delta t) + v(t - \Delta t) \cdot \Delta t + \frac{a(t - \Delta t) \cdot \Delta t^2}{2} \\ + \frac{u(t - \Delta t) \cdot \Delta t^3}{6} \end{aligned}$$

where  $p()$ ,  $v()$ , and  $a()$  are the position vector, velocity vector, and acceleration vector respectively,  $u()$  is the Gaussian white noise vector reflecting the result of random system disturbance combined with instantaneous jerk, and  $\Delta t = t_{k+1} - t_k = 220$  ms based on the design

requirements of FAST. Then we transform the kinematical equation into a state space representation, just as formula (1), by defining the system state and the parameter matrices

$$x(t_k)$$

$$= [x_1(t_k), x_2(t_k), x_3(t_k), x_4(t_k), x_5(t_k), x_6(t_k), \\ x_7(t_k), x_8(t_k), x_9(t_k)]^T$$

$$\Phi(t_k)$$

$$= \begin{bmatrix} 1 & 0 & 0 & \Delta t & 0 & 0 & \frac{\Delta t^2}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 & \Delta t & 0 & 0 & \frac{\Delta t^2}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & \Delta t & 0 & 0 & \frac{\Delta t^2}{2} \\ 0 & 0 & 0 & 1 & 0 & 0 & \Delta t & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \Delta t \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\Gamma(t_k)$$

$$= \begin{bmatrix} \frac{\Delta t^3}{6} & \frac{\Delta t^3}{6} & \frac{\Delta t^3}{6} & \frac{\Delta t^2}{2} & \frac{\Delta t^2}{2} & \frac{\Delta t^2}{2} & \Delta t & \Delta t & \Delta t \end{bmatrix}^T$$

**TABLE II**  
RMSEs OF STATE ESTIMATION  $\hat{x}_1(t|t, t)$

Experiment Number	Estimator 1 (/mm)	Estimator 2 (/mm)	Estimator 3 (/mm)
1	2.72828951	3.07855629	3.07730187
2	2.93343563	3.21315002	3.21137965
3	2.85697174	3.14172254	3.13981763
Average Upgrade Ratio	—	9.697%	9.649%

**TABLE III**  
RMSEs OF STATE ESTIMATION  $\hat{x}_2(t|t, t)$

Experiment Number	Estimator 1 (/mm)	Estimator 2 (/mm)	Estimator 3 (/mm)
1	2.91113948	3.20477291	3.20455662
2	2.89740383	3.14167405	3.14155346
3	2.82950762	3.15698516	3.15667038
Average Upgrade Ratio	—	9.106%	9.100%

**TABLE IV**  
RMSEs OF STATE ESTIMATION  $\hat{x}_3(t|t, t)$

Experiment Number	Estimator 1 (/mm)	Estimator 2 (/mm)	Estimator 3 (/mm)
1	2.88224021	3.16159472	3.15980223
2	2.79863518	3.13968695	3.13909065
3	2.89299165	3.20521602	3.20485288
Average Upgrade Ratio	—	9.810%	9.784%

where  $x_1(t_k), x_2(t_k), x_3(t_k)$  denote the position coordinates, with units of meters,  $x_4(t_k), x_5(t_k), x_6(t_k)$  denote the velocities along the three coordinate axes, with units of meters per second,  $x_7(t_k), x_8(t_k), x_9(t_k)$  denote the accelerations along the three coordinate axes, with units of meters per second squared. According to Table I, we set  $\lambda_1 = 220$  ms,  $v_0(t_k) \sim N(0, 0.02^2)$ , and  $v_1(t_k) \sim N(0, 0.003^2)$ , with units of meters. Finally, the parameter matrices of formulas (2) and (3) are

$$H_0(t_k) = H_1(t_k) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

because only position information can be obtained by the measurement equipment of FAST. This motion trajectory is composed of 1304 discrete points, i.e.,  $k_{\max} = 1304$ .

We completed a series of simulations with different system errors  $u(t_k)$ . The estimation errors of Estimator 1 in these experiments are all smaller than the other two. In this technical note, we set  $u(t_k) \sim (0, 0.032^2)$  as an example, with units of meters. The estimation curves are depicted in Fig. 1. To obtain the average results, we repeat the experiment 3 times. Detailed results of the numerical example are recorded in Tables II–IV. The RMSEs of Estimator 1 are significantly smaller than those of the other two. Finally, the average upgrade ratios of the Estimator 1 along the three coordinate axes, in comparison with the other two estimators, are listed in Tables II–IV.

## V. CONCLUSION

In this technical note, a novel MMSE estimation approach was proposed for linear discrete-time systems with time-delay and multi-error measurements, such as the feed cabin suspension subsystem of FAST. Compared with the augmented state Kalman filter and the method in [19], the proposed approach improves the estimation accuracy in the minimum mean square error sense.

Note that most systems are non-linear in realistic projects. Recognizing this problem, our future study will focus on the MMSE state

estimation approach for nonlinear discrete-time systems with time-delay and multi-error measurements.

## REFERENCES

- [1] R. D. Nan, D. Li, C. J. Jin, Q. M. Wang, L. C. Zhu, W. B. Zhu, H. Y. Zhang, Y. L. Yue, and L. Qian, "The five-hundred-meter aperture spherical radio telescope (fast) project," *Int. J. Modern Phys. D*, vol. 20, no. 6, pp. 989–1024, 2011.
- [2] J. W. Hu and L. C. Zhu, "Study on the combination of laser and inertial measurement technology for fast," *Sci. Technol. Eng.*, vol. 9, no. 11, pp. 2884–2889, 2009.
- [3] F. He, Y. Zheng, Y. B. Luo, and F. H. Ren, "Precision inspection and analysis of gps dynamic surveying," *Geotechn. Investigation Surveying*, vol. 37, no. 3, pp. 77–82, 2009.
- [4] B. F. Xie, M. Yu, T. Bai, and J. J. Guo, "Automatic gps-based surveying and control system for fast radio telescope," *J. Tsinghua Univ.: Sci. Technol.*, vol. 44, no. 3, pp. 379–382, 2004.
- [5] F. He, Y. Zheng, Y. B. Luo, and Y. Deng, "Precision inspection and time-delay analysis for dynamic surveying of total station," *Bulletin Surveying Mapping*, vol. 5, pp. 68–71, 2008.
- [6] H. Li, J. Sun, W. Zhu, and P. Gao Feng, "Simulation on flexible feed support system of five-hundred-meter aperture spherical radio telescope," *Comput. Aided Eng.*, vol. 20, no. 1, pp. 106–112, 2011.
- [7] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Automat. Control*, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.
- [8] M. Moayedi, Y. C. Soh, and Y. K. Foo, "Optimal kalman filtering with random sensor delays, packet dropouts and missing measurements," in *Proc. American Control Conf.*, 2009, pp. 3405–3410.
- [9] V. Dragan, "Optimal filtering for discrete-time linear systems with multiplicative white noise perturbations and periodic coefficients," *IEEE Trans. Automat. Control*, vol. 58, no. 4, pp. 1029–1034, Apr. 2013.
- [10] H. Kwakernaak, "Optimal filtering in linear systems with time delays," *IEEE Trans. Automat. Control*, vol. 12, no. 2, pp. 169–173, Feb. 1967.
- [11] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Fluids Eng.*, vol. 82, no. 1, pp. 35–45, 1960.
- [12] L. Schenato, "Optimal estimation in networked control systems subject to random delay and packet drop," *IEEE Trans. Automat. Control*, vol. 53, no. 5, pp. 1311–1317, May 2008.
- [13] D. Quevedo, A. Ahlen, and K. Johansson, "State estimation over sensor networks with correlated wireless fading channels," *IEEE Trans. Automat. Control*, vol. 58, no. 3, pp. 581–593, Mar. 2013.
- [14] A. S. Matveev and A. V. Savkin, "The problem of state estimation via asynchronous communication channels with irregular transmission times," *IEEE Trans. Automat. Control*, vol. 48, no. 4, pp. 670–676, Apr. 2003.
- [15] Y. G. Sun and S. Y. Qin, "Stability analysis and controller design for networked control systems with random time delay," *Int. J. Syst. Sci.*, vol. 42, no. 3, pp. 359–367, 2011.
- [16] S. Sun, "Optimal linear filters for discrete-time systems with randomly delayed and lost measurements with/without time stamps," *IEEE Trans. Automat. Control*, vol. 58, no. 6, pp. 1551–1556, Jun. 2013.
- [17] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Upper Saddle River, NJ, USA: Prentice Hall, 2000.
- [18] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Mineola, NY, USA: Courier Dover, 2012.
- [19] X. Lu, H. S. Zhang, W. Wang, and K. L. Teo, "Kalman filtering for multiple time-delay systems," *Automatica*, vol. 41, no. 8, pp. 1455–1461, 2005.
- [20] C. Y. Han and H. S. Zhang, "Optimal state estimation for discrete-time systems with random observation delays," *Acta Automatica Sinica*, vol. 35, no. 11, pp. 1446–1451, 2009.
- [21] B. Chen, L. Yu, and W. A. Zhang, "Robust kalman filtering for uncertain state delay systems with random observation delays and missing measurements," *Control Theory & Applications, IET*, vol. 5, no. 17, pp. 1945–1954, 2011.
- [22] Y. Sun, F. S. Jing, and Z. Z. Liang, "Iterated extended kalman filter for time-delay systems with multi-sample-rate measurements," in *Proc. World Congress Intelligent Control Automation*, 2014, pp. 4523–4527.
- [23] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. New York, NY, USA: Pearson Education, 1993.
- [24] D. C. Lay, *Linear Algebra and Its Applications*. Boston, MA, USA: Addison Wesley, 1997.
- [25] C. K. Chui and G. R. Chen, *Kalman Filtering With Real Time Applications*. Berlin, Germany: Springer, 1999.