Ultimately Bounded Filtering Subject to Impulsive Measurement Outliers

Lei Zou, Zidong Wang, Jun Hu, and Hongli Dong

Abstract—This paper is concerned with the ultimately bounded filtering problem for a class of linear time-delay systems subject to norm-bounded disturbances and impulsive measurement outliers (IMOs). The considered IMOs are modeled by a sequence of impulsive signals with certain known minimum norm (i.e. the minimum of the norms of all impulsive signals). In order to characterize the occasional occurrence of IMOs, a sequence of independently and identically distributed random variables is introduced to depict the interval lengths (i.e. the durations between two adjacent IMOs) of the outliers. In order to achieve satisfactory filtering performance, a novel parameter-dependent filtering approach is proposed to protect the filtering performance from IMOs by using a special outlier detection scheme, which is developed based on a particular input-output model. First, the ultimate boundedness (in mean square) of the filtering error is investigated by using the stochastic analysis technique and Lyapunov-functional-like method. Then, the desired filter gain matrix is derived through solving a constrained optimization problem. Furthermore, the designed filtering scheme is applied to the case where the statistical properties about the interval lengths of outliers are completely unknown. Finally, a simulation example is provided to demonstrate the effectiveness of our proposed filtering strategy.

Index Terms—Ultimately bounded filtering; Impulsive measurement outliers; Parameter-dependent filter; Matrix-fraction description; Time-delay systems.

I. INTRODUCTION

Filtering or state estimation problem is one of the mostly investigated topics in signal processing and control communities. Ever since the pioneering work of Luenberger in 1960s [27], there has been an ongoing research interest in the filtering or state estimation problems for various systems [23]. For example, in [32], the optimal filtering problem has been

Lei Zou and Zidong Wang are with the Department of Computer Science, Brunel University London, Uxbridge, Middlesex, UB8 3PH, United Kingdom. (Emails: Lei.Zou@brunel.ac.uk; Zidong.Wang@brunel.ac.uk).

Jun Hu is with the Department of Mathematics, Harbin University of Science and Technology, Harbin 150080, China. (Email: hujun2013@gmail.com)

Hongli Dong is with the Artificial Intelligence Energy Research Institute, Northeast Petroleum University, Daqing 163318, China; and is also with the Heilongjiang Provincial Key Laboratory of Networking and Intelligent Control, Northeast Petroleum University, Daqing 163318, China. (Email: shiningdhl@vip.126.com)

investigated for the linear time-invariant system with unknown covariances of process noises and measurement noises. In [22], a so-called Gaussian particle filter has been developed for nonlinear systems through approximating the posterior distributions by single Gaussian models. In [15], the robust $\mathcal{H}_\infty$ state estimation issue has been studied for uncertain systems with limited communication capacity based on a parameter-dependent approach. The moving horizon estimation problem has been studied in [48], [49] for networked systems with unknown inputs and communication constraints, respectively.

In order to evaluate the estimation accuracy of a filter, various filtering performance requirements have been reported in the literature, where the widely employed performance requirements include, but are not limited to, $\mathcal{H}_\infty$ requirement, minimum mean squared error (MMSE) index, ellipsoidal bound constraint, ultimate boundedness requirement and linear quadratic performance index. Among others, the ultimate boundedness requirement is particularly suitable for quantifying the steady-state filtering performance subject to norm-bounded noises [18], [26], [30], [33].

In numerous practical applications, the filtering performance might deteriorate when the underlying system suffers from model mismatch, unreliable communication, sensor failures, unknown inputs and measurement outliers [9], [28], [29], [37], [44]. In order to deal with the filtering task subject to these adverse effects, a number of filtering schemes have been developed [5], [6], [34], [36], [39]. For instance, in [31], several adaptive Kalman filtering approaches have been introduced for networked systems and sensor networks subject to mixed uncertainties of sensor delays, missing measurements and packet dropouts. The distributed $\mathcal{H}_\infty$ filtering issue has been studied in [16] for sensor networks with communication delays and event-triggered transmission scheme. In [35], the $\mathcal{H}_\infty$ filtering problem has been addressed for semi-Markov jump systems with randomly occurring uncertainties and sensor failures. When it comes to the *ultimately bounded filtering* problem for *time-delay systems* subject to *measurement outliers*, the corresponding results have been really scattered due mainly to the difficulties in handling the outlier-induced effects.

In general, measurement outliers refer to certain "abnormal" measurement disturbances that are quite distinguished from those widely studied noises. Compared with the noises, measurement outliers have their own distinctive features of 1) occasional/intermittent/probabilistic occurrences and 2) unexpectedly large amplitudes. So far, some initial results have been reported in the literature concerning the filtering problem subject to measurement outliers, see, e.g. [1], [8], [13], [14], [17], [38] and the references therein. Roughly speaking, the

main strategies dealing with measurement outliers can be categorized into groups, namely, the passive robustness-based schemes and the active detection-based schemes where, for the former, the filters are designed to be insensitive to outliers by selecting the suitable filter parameters through statistical tests or on-line adaptive adjustments [3], [8], [21]. It is worth mentioning that, in [3], a so-called stubborn state observer has been constructed by using the saturated innovation in the observer design, under which the effects induced by the possible outliers would be restrained. Nevertheless, the passive robustness-based schemes cannot guarantee the complete elimination of the effects induced by the measurement outliers in the filtering process.

In the active detection-based methods for tackling outliers, the filters are developed based on certain detection strategy with the hope to completely discard the measurements contaminated by outliers. Along this direction, a representative work has unarguably been [2], in which a leave-one-out detection strategy has been developed by removing the measurements possibly contaminated by outliers from the estimation process. Note that the results of [2] are still dependent on the upper bound of outliers, and this implies that the estimation performance might not be satisfactory if the amplitudes of outliers are out of normal range. To this end, a seemingly interesting research topic (which is also of clear engineering insight) is to develop an ultimately bounded filtering strategy such that the filtering results are *independent* of the upper bound of outliers, and this has motivated our present research.

Summarizing the discussions made so far, in this paper, we endeavor to investigate the ultimately bounded filtering problem for time-delay systems with measurement outliers by utilizing an active-detection based method. Subsequently, three fundamental challenges we are confronted with are identified as follows: 1) how to model the effects of measurement outliers according to practical engineering? 2) how to distinguish the outlier-corrupted-measurements with those normal measurements? and 3) how to design the filter such that the filtering error dynamics is ultimately bounded in mean square subject to the outliers without having to know the upper bound of outliers? The main purpose of this paper is to provide satisfactory answers to these three questions.

In response to the identified challenges, the contributions of this paper are highlighted as follows: 1) the filtering problem is, for the first time, investigated for time-delay systems with IMOs where a novel detection scheme is developed to identify whether the current measurement output is corrupted by an outlier; 2) a dedicatedly designed parameter-dependent (PD) filter is put forward under which the filtering error is ultimately bounded in mean square; and 3) the proposed filter design algorithm is applied to the case where the statistical properties about the interval lengths of IMOs (i.e. the durations between two adjacent IMOs) are completely unknown.

The remainder of this work is organized as follows. In Section II, some preliminaries are firstly introduced, and the time-delay systems with IMOs and the corresponding PD filtering structure are proposed. In Section III, the dependent parameter of the PD filter is designed. Furthermore, the desired filter gain matrix K is derived based on the statistical properties of

interval length for the measurement outliers. Then, the derived filtering strategy is applied to systems with unknown statistics of the interval length for the measurement outliers. Moreover, a simulation example is given in Section IV to demonstrate the correctness and effectiveness of our proposed PD filtering scheme. Finally, we present the conclusion of this work in Section V.

Notations: The notation used here is fairly standard except where otherwise stated. $\mathbb{R}^n$ and $\mathbb{R}^{n \times m}$ denote, respectively, the n dimensional Euclidean space and set of all $n \times m$ real matrices. $\mathbb{N}(\mathbb{N}^+, \mathbb{N}^-)$ denote, respectively, the set of integers (nonnegative integers, negative integers), and the set of all nonnegative real numbers is denoted by $\mathbb{R}^+$. The notation $X \geq Y$ $(X > Y)$, where X and Y are real symmetric matrices, means that $X - Y$ is positive semi-definite (positive definite). M^T represents the transpose of the matrix M. If A is a matrix, $\lambda_{\max}\{A\}$ $(\lambda_{\min}\{A\})$ stands for the maximum (minimum) eigenvalue of A, and $\text{tr}\{A\}$ represents the trace of A. For a matrix $P \in \mathbb{R}^{n \times n}$, $\det(P)$ denotes the determinant of the matrix P. 0 represents zero matrix of compatible dimensions. $\mathbf{1}_N$ represents an N dimensional row vector with all ones. The n-dimensional identity matrix is denoted as I_n or simply I, if no confusion is caused. The shorthand $\text{diag}\{\cdots\}$ stands for a block-diagonal matrix and the notation $\text{diag}_n\{\bullet\}$ is employed to stand for $\text{diag}\{\underbrace{\bullet, \cdots, \bullet}_n\}$. $\text{Prob}\{\cdot\}$ means the occurrence probability of the event "$\cdot$". $\mathbb{E}\{x\}$ and $\mathbb{E}\{x|y\}$ will, respectively, denote expectation of the stochastic variable x and expectation of x conditional on y. Given a generic vector x, $\|x\|$ describes the Euclidean norm of x. In symmetric block matrices, "$*$" is used as an ellipsis for terms induced by symmetry. The degree of a polynomial f is represented by $\deg(f)$, which is the highest degree of its monomials with nonzero coefficients. A polynomial f is said to be monic if the leading coefficient (the nonzero coefficient of highest degree) of this polynomial is equal to 1. Matrices, if they are not explicitly specified, are assumed to have compatible dimensions. The Kronecker delta function $\delta(a)$ is a binary function that equals 1 if $a = 0$ and equals 0 otherwise.

II. PROBLEM FORMULATION AND PRELIMINARIES

A. Preliminaries: matrix-fraction description

1) Polynomial matrix and rational matrix: Polynomial matrix is a matrix whose elements are polynomials, in other words, a polynomial matrix is a polynomial whose coefficients are matrices. A $m \times m$ polynomial matrix $P(z)$ is said to be singular if $\det(P(z)) = 0$ for all the z. $P(z)$ is said to be nonsingular if $\det(P(z)) \neq 0$ for certain z. A polynomial matrix $P(z)$ is said to be unimodular if its determinant (i.e. $\det(P(z))$) is a nonzero constant. Rational matrix $G(z)$ is a matrix whose elements are rational functions of the variable z.

2) Matrix-fraction description: A given rational matrix $G(z)$ can be written as a fraction of two polynomial matrices. As the product of matrices is not commutative, there are two different ways to proceed (e.g. the left matrix-fraction

description and right matrix-fraction description). More specifically, any rational matrix $G(z)$ admits a left matrix-fraction description (LMFD)

$$G(z) = D_L^{-1}(z) N_L(z) \qquad (1)$$

and a right matrix-fraction description (RMFD)

$$G(z) = N_R(z) D_R^{-1}(z) \qquad (2)$$

where $D_L(z)$, $D_R(z)$, $N_L(z)$ and $N_R(z)$ are polynomial matrices of suitable dimensions. LMFD (1) is irreducible if $D_L(z)$ and $N_L(z)$ are left coprime. Similarly, RMFD (2) is irreducible if $D_R(z)$ and $N_R(z)$ are right coprime.

3) How to compute the LMFD: Consider a rational $m \times r$ matrix $G(z)$ given as follows:

$$G(z) = \begin{bmatrix} \frac{e_{11}(z)}{d_{11}(z)} & \frac{e_{12}(z)}{d_{12}(z)} & \cdots & \frac{e_{1r}(z)}{d_{1r}(z)} \\ \frac{e_{21}(z)}{d_{21}(z)} & \frac{e_{22}(z)}{d_{22}(z)} & \cdots & \frac{e_{2r}(z)}{d_{2r}(z)} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{e_{m1}(z)}{d_{m1}(z)} & \frac{e_{m2}(z)}{d_{m2}(z)} & \cdots & \frac{e_{mr}(z)}{d_{mr}(z)} \end{bmatrix}.$$

where all the polynomials $d_{ij}(z)$ ($i = 1, 2, \cdots, m$; $j = 1, 2, \cdots, r$) are monic. Let the lowest common multiple (LCM) of $\{d_{ij}(z)\}_{j=1,2,\cdots,r}$ be $\hat{d}_i(z)$ and assume that $\hat{d}_i(z) = f_{ij}(z) d_{ij}(z)$. Then, a LMFD of the rational matrix $G(z)$ is given as follows:

$$G(z) = D_L^{-1}(z) N_L(z) \qquad (3)$$

where

$$\begin{cases} D_L(z) = \text{diag}\{\hat{d}_1(z), \hat{d}_2(z), \cdots, \hat{d}_n(z)\} \\ N_L(z) = \begin{bmatrix} e_{11}(z) f_{11}(z) & e_{12}(z) f_{12}(z) & \cdots & e_{1r}(z) f_{1r}(z) \\ e_{21}(z) f_{21}(z) & e_{22}(z) f_{22}(z) & \cdots & e_{2r}(z) f_{2r}(z) \\ \vdots & \vdots & \ddots & \vdots \\ e_{m1}(z) f_{m1}(z) & e_{m2}(z) f_{m2}(z) & \cdots & e_{mr}(z) f_{mr}(z) \end{bmatrix} \end{cases} \qquad (4)$$

Remark 1: Matrix-fraction description (MFD) of the transfer function matrix is a basic tool for characterizing the input-output relationships for linear time-invariant (LTI) systems. MFD-based technique provides a simple way to deal with the analysis and synthesis issues for LTI systems (e.g. the stability analysis). It is worth mentioning that the LMFD (3) is not unique to the rational matrix $G(z)$. Letting $\Lambda(z)$ be a nonsingular matrix, we have $G(z) = D_{L\Lambda}^{-1}(z) N_{L\Lambda}(z)$ where $D_{L\Lambda}(z) = \Lambda(z) D_L(z)$ and $N_{L\Lambda}(z) = \Lambda(z) N_L(z)$. Further introduction about MFD and the transfer function matrix can be found in [47].

B. Impulsive measurement outliers

In this subsection, we consider the mechanism of modeling the IMOs. In practical engineering, an outlier occurs at one of the sampling instants and thus the number of occurred outliers should be accountable. Without loss of generality, let the sequence of occurrence moments for the outliers be

$$t(0) < t(1) < \cdots < t(j) < \cdots$$

where $t(j)$ ($j \in \mathbb{N}^+$) represents the occurrence moment of the j-th measurement outlier. Based on such an occurrence

moment sequence, the measurement outlier o_k (i.e. the outlier occurring at the sampling instant k) is modeled by the following form of impulsive signals

$$o_k = \sum_{i=0}^{\infty} \delta(k - t(i)) \hat{o}_i$$

where $\hat{o}_i$ represents the amplitude (a vector to be defined later) of the i-th measurement outlier. Moreover, by defining the interval length T_i as

$$T_i = t(i) - t(i-1)$$

for $i \in \mathbb{N}^+$ with initial value $T_0 = t(0)$, the occurrence moment $t(i)$ can be rewritten as $t(i) = \sum_{j=0}^{i} T_i$, which implies that

$$o_k = \sum_{i=0}^{\infty} \delta\left(k - \sum_{j=0}^{i} T_j\right) \hat{o}_i. \qquad (5)$$

Next, let us introduce some necessary assumptions about the proposed IMO.

Assumption 1: The interval length T_i satisfies the following condition:

$$T_i \in \Upsilon \triangleq \{\underline{T}, \underline{T}+1, \cdots, \underline{T}+s\}$$

where $\underline{T}$ and s are two known positive constants. Furthermore, the amplitude $\hat{o}_i$ satisfies the condition $\|\hat{o}_i\| \geq \underline{o}$ where $\underline{o}$ is a known constant.

Assumption 2: The interval lengths $\{T_i\}_{i \geq 0}$ are a sequence of independently and identically distributed (i.i.d.) random variables and the occurrence probability of $T_i = \underline{T} + j$ ($\forall j \in \{0, 1, \cdots, s\}$) is given by

$$\text{Prob}\{T_i = \underline{T} + j\} = p_j$$

where $p_j > 0$ ($\forall j \in \{0, 1, \cdots, s\}$) and $\sum_{j=0}^{s} p_j = 1$.

Remark 2: In this paper, the measurement outliers represent the anomalous signals that might result from sensor malfunctions, wrong replacement of measures, cyber-attacks or large non-Gaussian noises [2]. These anomalous signals typically take place on an occasional basis with relatively large amplitudes as compared with the conventional disturbances. In order to capture such an occasional nature, a set of special impulsive signals has been employed to model the measurement outliers. In our previous work [50], we have considered the set-membership filtering problem of time-varying systems with IMOs, which are modeled by a sequence of impulsive signals whose interval lengths and norms are larger than certain fixed thresholds. In the current investigation, the interval lengths of IMOs are modeled by a sequence of i.i.d. random variables $\{T_i\}_{i \in \mathbb{N}^+}$. The statistical properties of these i.i.d. random variables are identifiable from engineering practice. It is worth noting that the interval length T_i is an important index which is commonly used in practical applications to characterize the occurrence frequency of outliers. A typical example of such an index can be found in the failure model of a repairable system, where an important index named the time between failures (TBFs) is adopted to model the failure frequency [7], [10]. Obviously, the so-called TBFs can be regarded as the

interval lengths of outliers. Furthermore, it is worth mentioning that the IMO model adopted in this paper is more general as compared with our previous work [50] since, when the scalar s and the statistical properties of $\{T_i\}_{i \in \mathbb{N}^+}$ are completely unknown, the IMO model utilized in this paper reduces to the one established in [50].

C. Problem formulation: plant and filter structure

Consider a discrete-time linear system of the form

$$\begin{cases} x_{k+1} = Ax_k + Ex_{k-\tau} + B\omega_k \\ y_k = Cx_k + F\nu_k + o_k \\ z_k = Mx_k \end{cases} \tag{6}$$

where $x_k \in \mathbb{R}^n$, $y_k \in \mathbb{R}^m$ and $z_k \in \mathbb{R}^l$ are the system state, the measurement output and the output vector to be estimated, respectively; $\omega_k \in \mathcal{W} \triangleq \{\omega : \|\omega\| \leq \bar{\omega}; \omega \in \mathbb{R}^r\}$ and $\nu_k \in \mathcal{V} \triangleq \{\nu : \|\nu\| \leq \bar{\nu}; \nu \in \mathbb{R}^s\}$ denote, respectively, the process and measurement noises where $\bar{\omega}$ and $\bar{\nu}$ are known positive constants; the parameters A, B, C, E, F and M are real-valued matrices of appropriate dimensions; τ is the known constant; and the vector $o_k \in \mathbb{R}^m$ is the IMO of the form (5). Here, y_k and z_k are the system output and the controlled output, respectively, where y_k is the actual measurement output that is obtained from the sensors and corrupted by the noise ν_k and possibly the outlier o_k. The matrix C is the output matrix. The controlled output z_k denotes the signal to be estimated, which is actually the signal we are interested in. In other words, z_k is a noise-free signal that describe the "real information" that we want to know, and the matrix M is problem-dependent that can be determined by engineering practice.

Remark 3: Note that the model (6) includes the time-delay $Ex_{k-\tau}$, which is justified as follows. As is well known, time-delays commonly occur in many dynamical systems including microwave oscillators, electronics, biological systems and hydraulic systems. For example, in a robotic teleoperation system, a time-delay might take place during the information transmission if the teleoperation is performed over a great distance [4]. So far, there has been a rich body of research results available in the literature on various control and filtering problems for time-delay systems. The existence of time delays could lead to the deterioration of system performance or even the instability of the system. As can be seen in the sequel, the inclusion of the time-delay $Ex_{k-\tau}$ in (6) will give rise to additional difficulty in the filter analysis.

It is supposed that the occurrence moment sequence of the outliers is completely unknown and Assumption 1 holds for any $i \in \mathbb{N}^+$. Without loss of generally, in this paper, we assume that $x_k = 0$, $\nu_k = 0$ and $o_k = 0$ for all $k < 0$. Actually, this assumption is not restrictive at all. For example, suppose that the initial condition of (6) is $x_i = \phi$ $(i < 0)$. Obviously, we have $y_i = y_\phi \triangleq C\phi$ $(i < 0)$. Letting $\tilde{x}_k \triangleq x_k - \phi$, $\tilde{y}_k \triangleq y_k - y_\phi$ and $\tilde{\omega}_k \triangleq B\omega_k + A\phi + E\phi - \phi$, we can derive the following linear system

$$\begin{cases} \tilde{x}_{k+1} = A\tilde{x}_k + E\tilde{x}_{k-\tau} + \tilde{\omega}_k \\ \tilde{y}_k = C\tilde{x}_k + F\nu_k + o_k \\ \tilde{x}_i = 0, \quad \forall i < 0 \end{cases}$$

and it is easy to see that the new process noise vector $\tilde{\omega}_k$ is still a norm-bounded signal. Hence, the above reformulated system falls into the plant form (6).

Next, let us consider the filter structure for the system (6). In conventional filtering scheme, the filter is often designed based on the structure of Luenberger observer and the filter gain matrix is selected such that the filtering error dynamics is *input-to-state stable* subject to the external inputs. The filtering performance corresponding to this strategy is largely dependent on the external inputs (e.g. the process noise, measurement noise and IMOs). Clearly, such kind of conventional Luenberger-observer-based filters would fail to achieve a satisfactory filtering performance if the amplitudes of outliers are quite large. To specifically tackle the outlier issue, in this work, we aim to design a dedicated parameter-dependent (PD) filter which is capable of "protecting" the filtering performance from the effects of IMOs. In order to achieve such a purpose, we adopt the following PD filter:

$$\begin{cases} \hat{x}_{k+1} = A\hat{x}_k + E\hat{x}_{k-\tau} + \theta_k K(y_k - C\hat{x}_k) \\ \hat{z}_k = M\hat{x}_k \end{cases} \tag{7}$$

where $\hat{x}_k$ and $\hat{z}_k$ denote, respectively, the estimates of x_k and z_k; θ_k is a binary function to be designed which takes values of 1 or 0; and the PD filter gain matrix K is the parameter to be designed.

In filter (7), the binary function θ_k characterizes the outlier detection result at time instant k. The rationale behind the proposed structure of the PD filter (7) is to remove certain outlier-related innovation from the filtering process according to the outlier detection result θ_k. Based on the outlier detection results, the "rejection" of the outliers can be ensured in our filtering scheme if the outlier detection technology is reliable (i.e. $\theta_k = 0$ holds if $o_k \neq 0$). In other words, the binary function θ_k should be designed such that $\theta_k = 0$ if there exists an integer $i \in \mathbb{N}^+$ satisfying the condition $t(i) = k$.

Before presenting the problem addressed in this paper, let us introduce the following definition about exponentially ultimate boundedness (EUB) in mean square.

Definition 1: [41] The dynamics of the filtering error $e_k \triangleq x_k - \hat{x}_k$ is said to be exponentially ultimately bounded in mean square if there exist constants $\varsigma > 0$, $\beta > 0$, $\epsilon > 0$ such that

$$\mathbb{E}\{\|e_k\|^2\} \leq \varsigma^k \beta + \epsilon \tag{8}$$

where $\varsigma \in [0, 1)$, $\beta > 0$ and $\epsilon > 0$, with ς and ϵ denoting the decay rate and the asymptotic upper bound of $\mathbb{E}\{\|e_k\|^2\}$, respectively.

Now, we are ready to introduce the problem addressed in this work as follows. The objective of this paper is twofold.

a) Firstly, design the binary function θ_k such that the condition $\theta_k = 0$ holds if and only if $o_k \neq 0$ (i.e. there is an outlier occurring at time instant k). In other words, the designed PD filter (7) is capable of removing the innovations contaminated by IMOs.

b) Secondly, based on the developed binary function θ_k, design the filter parameter K such that the filtering error e_k is exponentially ultimately bounded in mean square subject to the process noise ω_k, measurement noise ν_k and the IMO o_k.

III. MAIN RESULTS

A. Design of the function θ_k

In this subsection, we intend to develop a detection method capable of identifying whether the received measurement contains an outlier. Before proceeding further, let us first establish the input-output model of the plant (6) which will be used in developing the outlier detection scheme. According to the system dynamics (6), we have the following delay-free linear system:

$$\begin{cases} \bar{x}_{k+1} = \bar{A}\bar{x}_k + \bar{B}\omega_k \\ y_k = \bar{C}\bar{x}_k + F\nu_k + o_k \end{cases} \tag{9}$$

where

$$\bar{x}_k \triangleq \begin{bmatrix} x_k^T & x_{k-1}^T & \cdots & x_{k-\tau}^T \end{bmatrix}^T, \quad \bar{C} \triangleq \begin{bmatrix} C & 0 & \cdots & 0 \end{bmatrix},$$

$$\bar{A} \triangleq \begin{bmatrix} A & 0 & \cdots & 0 & E \\ I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix}, \quad \bar{B} \triangleq \begin{bmatrix} B \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

The discrete-time transfer function matrices of (9) are given as follows:

$$\begin{cases} G_{y\omega}(z) = \bar{C}(zI - \bar{A})^{-1}\bar{B} \\ G_{y\nu}(z) = F \\ G_{yo}(z) = I \end{cases} \tag{10}$$

Since the initial condition of (9) is zero, it is concluded from (10) that

$$y(z) = G_{y\omega}(z)\omega(z) + G_{y\nu}(z)\nu(z) + G_{yo}(z)o(z) \tag{11}$$

where $y(z)$, $\omega(z)$, $\nu(z)$ and $o(z)$ represent the Z-transforms of the discrete-time signals y_k, ω_k. ν_k and o_k, respectively.

Now, we convert the discrete-time transfer function matrix $G_{y\omega}(z)$ into the LMFD (according to the introduction in Subsection II-A3):

$$G_{y\omega}(z) = D_L^{-1}(z)N_L(z) \tag{12}$$

where $D_L(z)$ and $N_L(z)$ are polynomial matrices and $D_L(z)$ is a diagonal matrix with the form $D_L(z) = \text{diag}\{\hat{d}_1(z), \hat{d}_2(z), \cdots, \hat{d}_m(z)\}$. Then, it follows from (11) and (12) that

$$D_L(z)y(z) = N_L(z)\omega(z) + D_L(z)F\nu(z) + D_L(z)o(z).$$

Let the degree of $\hat{d}_i(z)$ be $\vec{d}_i$ (i.e. $\deg(\hat{d}_i(z)) = \vec{d}(i)$). Define a diagonal polynomial matrix as follows:

$$W(z) = \text{diag}\{z^{\bar{d}-\vec{d}(1)}, z^{\bar{d}-\vec{d}(2)}, \cdots, z^{\bar{d}-\vec{d}(m)}\} \tag{13}$$

where $\bar{d} \triangleq \max_{i=1,2,\cdots,m}\{\vec{d}(i)\}$. Then, we have

$$\bar{D}_L(z)y(z) = \bar{N}_L(z)\omega(z) + \bar{D}_L(z)F\nu(z) + \bar{D}_L(z)o(z) \tag{14}$$

in which $\bar{D}_L(z) \triangleq W(z)D_L(z)$ and $\bar{N}_L(z) = W(z)N_L(z)$. Obviously, $\bar{D}_L(z)$ is a diagonal matrix whose diagonal element is monic and the degree of any diagonal element is $\bar{d}$.

As such, the polynomial matrices $\bar{D}_L(z)$ and $\bar{N}_L(z)$ can be rewritten as follows:

$$\begin{cases} \bar{D}_L(z) = z^{\bar{d}}I + z^{\bar{d}-1}\bar{D}_{L1} + \cdots + \bar{D}_{L\bar{d}} \\ \bar{N}_L(z) = z^{\bar{d}-1}\bar{N}_{L1} + z^{\bar{d}-2}\bar{N}_{L2} + \cdots + \bar{N}_{L\bar{d}} \end{cases} \tag{15}$$

where $\bar{D}_{Li}$ $(i = 1, 2, \cdots, \bar{d})$ and $\bar{N}_{Lj}$ $(j = 1, 2, \cdots, \bar{d})$ are real constant matrices. Furthermore, it can be concluded that $\{\bar{D}_{Li}\}_{i=1,2,\cdots,\bar{d}}$ are diagonal matrices.

Without loss of generally, we let $\omega_k = 0$ for $k < 0$ and then the input-output model of the plant (6) can be described as follows:

$$\mathcal{D}_L y_{k-\bar{d}}^k = \mathcal{N}_L \omega_{k-\bar{d}}^{k-1} + \mathcal{D}_L \mathcal{F} \nu_{k-\bar{d}}^k + \mathcal{D}_L o_{k-\bar{d}}^k \tag{16}$$

where

$$y_{k-\bar{d}}^k \triangleq \begin{bmatrix} y_k^T & y_{k-1}^T & \cdots & y_{k-\bar{d}}^T \end{bmatrix}^T,$$
$$\omega_{k-\bar{d}}^{k-1} \triangleq \begin{bmatrix} \omega_{k-1}^T & \omega_{k-2}^T & \cdots & \omega_{k-\bar{d}}^T \end{bmatrix}^T,$$
$$\nu_{k-\bar{d}}^k \triangleq \begin{bmatrix} \nu_k^T & \nu_{k-1}^T & \cdots & \nu_{k-\bar{d}}^T \end{bmatrix}^T,$$
$$o_{k-\bar{d}}^k \triangleq \begin{bmatrix} o_k^T & o_{k-1}^T & \cdots & o_{k-\bar{d}}^T \end{bmatrix}^T,$$
$$\mathcal{D}_L \triangleq \begin{bmatrix} I & \bar{D}_{L1} & \cdots & \bar{D}_{L\bar{d}} \end{bmatrix},$$
$$\mathcal{N}_L \triangleq \begin{bmatrix} \bar{N}_{L1} & \bar{N}_{L2} & \cdots & \bar{N}_{L\bar{d}} \end{bmatrix}, \quad \mathcal{F} \triangleq \text{diag}_{\bar{d}+1}\{F\}.$$

Based on the above input-output model (16), we have the following proposition.

Proposition 1: Under Assumption 1, suppose that $\underline{T} > \bar{d}$. Define the sequence of $\{\hat{t}(i)\}_{i \geq 0}$ as follows:

$$\hat{t}(i) = \begin{cases} \min\{k | k \geq \underline{T}, f(k) > \bar{f}\}, & \text{if } i = 0 \\ \min\{k | k \geq \hat{t}(i-1) + \underline{T}, f(k) > \bar{f}\}, & \text{otherwise} \end{cases} \tag{17}$$

where

$$f(k) \triangleq \|\mathcal{D}_L y_{k-\bar{d}}^k\|, \quad \bar{f} \triangleq \|\mathcal{N}_L\|\bar{d}\bar{\omega} + \|\mathcal{D}_L\mathcal{F}\|(\bar{d}+1)\bar{\nu}.$$

Then, the condition $\hat{t}(i) = t(i)$ holds for all $i \geq 0$ if $\underline{o} > 2\bar{f}$.

Proof: The proof of this proposition is performed by *mathematical induction*.

Initial step. Considering the function $f(k)$, for any $\bar{d} < \underline{T} \leq k < t(0)$, it follows that

$$f(k) = \|\mathcal{D}_L y_{k-\bar{d}}^k\| = \|\mathcal{N}_L \omega_{k-\bar{d}}^{k-1} + \mathcal{D}_L \mathcal{F} \nu_{k-\bar{d}}^k\| \leq \bar{f} \tag{18}$$

On the other hand, under Assumption 1, it is deduced that

$$o_{t(0)-\bar{d}}^{t(0)} = \begin{bmatrix} o_{t(0)} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Then, it follows from the condition $\underline{o} > 2\bar{f}$ that

$$\begin{aligned} f(t(0)) &= \|\mathcal{D}_L y_{t(0)-\bar{d}}^{t(0)}\| \\ &= \|\mathcal{N}_L \omega_{t(0)-\bar{d}}^{t(0)-1} + \mathcal{D}_L \mathcal{F} \nu_{t(0)-\bar{d}}^{t(0)} + \mathcal{D}_L o_{t(0)-\bar{d}}^{t(0)}\| \\ &\geq \|o_{t(0)}\| - \|\mathcal{N}_L \omega_{t(0)-\bar{d}}^{t(0)-1} + \mathcal{D}_L \mathcal{F} \nu_{t(0)-\bar{d}}^{t(0)}\| \\ &\geq \|o_{t(0)}\| - \bar{f} > \bar{f}. \end{aligned} \tag{19}$$

As such, it is immediately known from (18) and (19) that

$$\hat{t}(0) = \min\{k | k \geq \underline{T}, f(k) > \bar{f}\} = t(0), \qquad (20)$$

which implies that the assertion $\hat{t}(i) = t(i)$ holds for $i = 0$.

Inductive step. Now that the assertion $\hat{t}(i) = t(i)$ is true for $i = 0$. Next, given that the assertion is true for $i = j$ (i.e. $\hat{t}(j) = t(j)$), we aim to show that the same assertion is true for $i = j + 1$.

Along the similar lines of the Initial step, we have for any $t(j) + \underline{T} \leq k < t(j+1)$ that

$$f(k) = \left\|\mathcal{D}_L y_{k-\bar{d}}^k\right\| = \left\|\mathcal{N}_L \omega_{k-\bar{d}}^{k-1} + \mathcal{D}_L \mathcal{F} \nu_{k-\bar{d}}^k + \mathcal{D}_L \mathcal{F} o_{k-\bar{d}}^k\right\|. \tag{21}$$

Note that $k - \bar{d} > k - \underline{T} \geq t(j)$, which implies that $o_{k-\bar{d}}^k = 0$. Hence, it follows from (21) that

$$f(k) = \left\|\mathcal{N}_L \omega_{k-\bar{d}}^{k-1} + \mathcal{D}_L \mathcal{F} \nu_{k-\bar{d}}^k\right\| \leq \bar{f}. \tag{22}$$

Moreover, we have

$$o_{t(j+1)-\bar{d}}^{t(j+1)} = \begin{bmatrix} o_{t(j+1)} \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

which means that

$$\begin{aligned} f(t(j+1)) &= \left\|\mathcal{D}_L y_{t(j+1)-\bar{d}}^{t(j+1)}\right\| \\ &= \left\|\mathcal{N}_L \omega_{t(j+1)-\bar{d}}^{t(j+1)-1} + \mathcal{D}_L \mathcal{F} \nu_{t(j+1)-\bar{d}}^{t(j+1)} + \mathcal{D}_L o_{t(j+1)-\bar{d}}^{t(j+1)}\right\| \\ &\geq \left\|o_{t(j+1)}\right\| - \left\|\mathcal{N}_L \omega_{t(j+1)-\bar{d}}^{t(j+1)-1} + \mathcal{D}_L \mathcal{F} \nu_{t(j+1)-\bar{d}}^{t(j+1)}\right\| \\ &\geq \left\|o_{t(j+1)}\right\| - \bar{f} > \bar{f}. \end{aligned} \tag{23}$$

As such, it follows readily from (22) and (23) that

$$\begin{aligned} \hat{t}(j+1) &= \min\{k | k \geq \hat{t}(j) + \underline{T}, f(k) > \bar{f}\} \\ &= \min\{k | k \geq t(j) + \underline{T}, f(k) > \bar{f}\} = t(j+1). \end{aligned} \tag{24}$$

Hence, by the induction, it is concluded that the assertion $\hat{t}(i) = t(i)$ is true for $i = 0, 1, \cdots$. The proof is now complete. ∎

Let $\underline{T} > \bar{d}$ and $\underline{\varrho} > 2\bar{f}$. By means of Proposition 1, we propose an algorithm (Algorithm 1) to compute the occurrence moment sequence of outliers (i.e. $\{\hat{t}(i)\}_{i \geq 0}$) as follows:

Algorithm 1:

Step 1. Initialization: let $i = 0$. According to Proposition 1, compute the value of $\bar{f}$. If the current time instant $k < \underline{T}$, wait for the next time instant. Otherwise, go to the next step.

Step 2. Based on the received measurement outputs $\{y_j\}_{k-\bar{d} \leq j \leq k}$, compute the value of the function $f(k)$.

Step 3. If $f(k) > \bar{f}$, set $\hat{t}(i) = k$. Let $i = i + 1$ and go to the next step.

Step 4. If the current time instant $k < \hat{t}(i-1) + \underline{T}$, wait for the next time instant. Otherwise, go to *Step 3*.

Algorithm 1 is an on-line outlier detection algorithm which is very easy to implement in practical applications. At each time instant, the computational cost depends on the calculation of $f(k)$, which has a polynomial time complexity. More specifically, the corresponding computational complexity about the

calculation of $f(k)$ is bounded by $O(m^2(\bar{d}+1))$. According to the above algorithm, we can recursively calculate the values of $\{\hat{t}(i)\}_{i \geq 0}$, based on which the value of the binary function θ_k is designed as follows:

$$\theta_k = \begin{cases} 0, & \text{if } \{i \geq 0 | k = \hat{t}(i)\} \neq \emptyset \\ 1, & \text{otherwise} \end{cases}. \tag{25}$$

Obviously, the outlier detection result θ_k effectively characterizes the occurrence of outliers at time instant k and, more specifically, the condition $\theta_k = 0$ holds if and only if $o_k \neq 0$ (i.e. there is an outlier occurring at time instant k).

Note that the binary function θ_k can be rewritten by

$$\theta_k = 1 - \sum_{i=0}^{\infty} \delta(k - \hat{t}(i)) = 1 - \sum_{i=0}^{\infty} \delta(k - t(i)),$$

which implies that

$$\begin{aligned} \theta_k o_k &= \left(1 - \sum_{i=0}^{\infty} \delta(k - t(i))\right) \sum_{i=0}^{\infty} \delta(k - t(i))\hat{o}_i \\ &= \sum_{i=0}^{\infty} \delta(k - t(i))\hat{o}_i - \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \delta(k - t(i))\delta(k - t(j))\hat{o}_j \\ &= \sum_{i=0}^{\infty} \delta(k - t(i))\hat{o}_i - \sum_{i=0}^{\infty} \delta(k - t(i))\delta(k - t(i))\hat{o}_i = 0. \end{aligned}$$

Hence, we have $\theta_k K(y_k - C\hat{x}_k) = \theta_k K(Cx_k + F\nu_k - C\hat{x}_k)$. Obviously, according to the PD filter (7), it is easy to find that the rejection of the outliers is ensured by the designed binary function θ_k.

Remark 4: It is seen that $\bar{D}_L^{-1}(z)\bar{N}_L(z)$ serves as another LMFD of the discrete-time transfer function matrix $G_{y\omega}(z)$. The main reason to adopt this new LMFD of $\bar{D}_L^{-1}(z)\bar{N}_L(z)$ (rather than $D_L^{-1}(z)N_L(z)$) is that the leading coefficient of $\bar{D}_L(z)$ is I which is nonsingular matrix. Based on this new LMFD, we are able to examine the effects of outliers from the measurement outputs exactly, that is, $\left\|\mathcal{D}_L o_{t(i)-\bar{d}}^{t(i)}\right\| > 0$ is ensured for any $i \in \mathbb{N}^+$.

Remark 5: So far, we have accomplished the design of the function θ_k. By setting the values of $\{\theta_k\}_{k \geq 0}$ according to Proposition 1, we can examine whether the received measurement output contains an outlier or not. Note that Proposition 1 is derived based on Assumption 1 and the condition $\underline{\varrho} > 2\bar{f}$, which have ensured the *shortest* occurrence period of the outliers and the *lower bound* of the norm of the outlier. These findings are particularly reasonable for real-world applications since the occurrence frequency of outliers is typically quite low and the amplitude of outliers is typically very large as compared with the noises. Based on the designed θ_k, we will deal with the design of the filter parameter K in the next subsection to ensure the EUB in mean square of the filtering error.

Remark 6: The key point of removing the "harmful" innovations (corrupted by outliers) is to "identify" the occurrence moment for each IMO. Such a removal is, however, difficult to accomplish by adopting the traditional model-based fault detection (MbFD) method that relies on the so-called residual (at each time instant) which is further affected by the historical external inputs (including the historical outliers). As such,

it might be the case that the generated residual is quite small even if the received measurement is corrupted by an outlier. Different from the MbFD method, the outlier detection approach proposed in this paper is capable of accurately distinguishing the IMO-corrupted measurements from the normal measurements under the condition $\varrho > 2\bar{f}$. In this sense, our proposed detection scheme is superior to the MbFD scheme when detecting the IMOs considered in this paper.

B. Design of the parameter K

Now, let us consider the estimation error dynamics. Let the filtering error be $e_k \triangleq x_k - \hat{x}_k$. According to the proposed PD filter (7), the proposed binary function θ_k (25) and Proposition 1, the dynamics of the filtering error could be derived as follows:

$$e_{k+1} = (A - \theta_k KC)e_k + Ee_{k-\tau} + B\omega_k - \theta_k KF\nu_k. \quad (26)$$

By defining two parameter-dependent matrices as follows:

$$\mathcal{A}(\theta_k) \triangleq A - \theta_k KC, \quad \mathcal{F}(\theta_k) \triangleq -\theta_k KF,$$

the filtering error dynamics (26) could be rewritten as follows:

$$\begin{cases} e_{k+1} = \mathcal{A}(\theta_k)e_k + Ee_{k-\tau} + B\omega_k + \mathcal{F}(\theta_k)\nu_k \\ \tilde{z}_k \triangleq z_k - \hat{z}_k = Me_k \end{cases} \quad (27)$$

In light of Assumption 2 and the filtering error dynamics (27), sufficient conditions guaranteeing the EUB in mean square of the filtering error are given in the following theorem.

Theorem 1: Consider the filtering error dynamics (27). Suppose that Assumptions 1-2 hold and let the filter parameter K be given. Assume that there exist two positive definite matrices $P_1, P_2 \in \mathbb{R}^{n \times n}$ and five positive scalars λ_i ($i = 1, 2, 3$), μ_1 and μ_2 satisfying

$$\bar{\Upsilon}_1 = \begin{bmatrix} \bar{\Upsilon}_1^{11} & \bar{\Upsilon}_1^{12} & \bar{\Upsilon}_1^{13} & \bar{\Upsilon}_1^{14} \\ * & \bar{\Upsilon}_1^{22} & \bar{\Upsilon}_1^{23} & \bar{\Upsilon}_1^{24} \\ * & * & \bar{\Upsilon}_1^{33} & \bar{\Upsilon}_1^{34} \\ * & * & * & \bar{\Upsilon}_1^{44} \end{bmatrix} < 0 \quad (28)$$

$$\bar{\Upsilon}_2 = \begin{bmatrix} \bar{\Upsilon}_2^{11} & \bar{\Upsilon}_2^{12} & \bar{\Upsilon}_2^{13} \\ * & \bar{\Upsilon}_2^{22} & \bar{\Upsilon}_2^{23} \\ * & * & \bar{\Upsilon}_2^{33} \end{bmatrix} < 0 \quad (29)$$

$$\gamma = (1 + \mu_2) \sum_{i=0}^{s} p_i (1 - \mu_1)^{\underline{T}+i-1} < 1 \quad (30)$$

where

$\bar{\Upsilon}_1^{11} \triangleq (A - KC)^T P_1 (A - KC) - (1 - \mu_1)P_1 + P_2,$
$\bar{\Upsilon}_1^{12} \triangleq (A - KC)^T P_1 E, \quad \bar{\Upsilon}_1^{13} \triangleq (A - KC)^T P_1 B,$
$\bar{\Upsilon}_1^{14} \triangleq -(A - KC)^T P_1 KF, \ \bar{\Upsilon}_1^{22} \triangleq E^T P_1 E - (1 - \mu_1)^\tau P_2,$
$\bar{\Upsilon}_1^{23} \triangleq E^T P_1 B, \ \bar{\Upsilon}_1^{24} \triangleq -E^T P_1 KF, \ \bar{\Upsilon}_1^{33} \triangleq B^T P_1 B - \lambda_1 I,$
$\bar{\Upsilon}_1^{34} \triangleq -B^T P_1 KF, \quad \bar{\Upsilon}_1^{44} \triangleq F^T K^T P_1 KF - \lambda_2 I,$
$\bar{\Upsilon}_2^{11} \triangleq A^T P_1 A - (1 + \mu_2)P_1 + P_2, \quad \bar{\Upsilon}_2^{12} \triangleq A^T P_1 E,$
$\bar{\Upsilon}_2^{13} \triangleq A^T P_1 B, \quad \bar{\Upsilon}_2^{22} \triangleq E^T P_1 E - (1 - \mu_1)^\tau P_2,$
$\bar{\Upsilon}_2^{23} \triangleq E^T P_1 B, \quad \bar{\Upsilon}_2^{33} \triangleq B^T P_1 B - \lambda_3 I.$

Then, the dynamics of the filtering error is ultimately bounded in mean square subject to the disturbances ω_k, ν_k and the measurement outlier o_k.

Proof: To analyze the EUB in mean square of the filtering error e_k, we first consider the following two cases.

Case 1: $\theta_k = 1$.

In this case, we select the following Lyapunov-like functional:

$$V_k = V_{1,k} + V_{2,k} \quad (31)$$

where

$$V_{1,k} \triangleq e_k^T P_1 e_k, \quad V_{2,k} \triangleq \sum_{i=k-\tau}^{k-1} (1 - \mu_1)^{k-i-1} e_i^T P_2 e_i.$$

Obviously, $\theta_k = 0$ means that $0 \leq k < t(0)$ or there exists a nonnegative integer i satisfying $t(i) < k < t(i+1)$. Calculating the difference of V_k along the trajectories of (27), one has

$$\begin{aligned} &V_{1,k+1} - (1 - \mu_1)V_{1,k} \\ &= e_{k+1}^T P_1 e_{k+1} - (1 - \mu_1)e_k^T P_1 e_k \\ &= \big(\mathcal{A}(\theta_k)e_k + Ee_{k-\tau} + B\omega_k + \mathcal{F}(\theta_k)\nu_k\big)^T P_1 \big(\mathcal{A}(\theta_k)e_k \\ &\quad + Ee_{k-\tau} + B\omega_k + \mathcal{F}(\theta_k)\nu_k\big) - (1 - \mu_1)e_k^T P_1 e_k \\ &= \begin{bmatrix} e_k \\ e_{k-\tau} \\ \omega_k \\ \nu_k \end{bmatrix}^T \Omega \begin{bmatrix} e_k \\ e_{k-\tau} \\ \omega_k \\ \nu_k \end{bmatrix} \end{aligned} \quad (32)$$

where

$$\Omega \triangleq \begin{bmatrix} \Omega_{11} & \Omega_{12} & \mathcal{A}^T(\theta_k)P_1 B & \mathcal{A}^T(\theta_k)P_1 \mathcal{F}(\theta_k) \\ * & E^T P_1 E & E^T P_1 B & E^T P_1 \mathcal{F}(\theta_k) \\ * & * & B^T P_1 B & B^T P_1 \mathcal{F}(\theta_k) \\ * & * & * & \mathcal{F}^T(\theta_k)P_1 \mathcal{F}(\theta_k) \end{bmatrix},$$

$\Omega_{11} \triangleq \mathcal{A}^T(\theta_k)P_1 \mathcal{A}(\theta_k) - (1 - \mu_1)P_1$, $\Omega_{12} \triangleq \mathcal{A}^T(\theta_k)P_1 E$,

and

$$\begin{aligned} &V_{2,k+1} - (1 - \mu_1)V_{2,k} \\ &= \sum_{i=k-\tau+1}^{k} (1 - \mu_1)^{k-i} e_i^T P_2 e_i - \sum_{i=k-\tau}^{k-1} (1 - \mu_1)^{k-i} e_i^T P_2 e_i \\ &= e_k^T P_2 e_k - (1 - \mu_1)^\tau e_{k-\tau}^T P_2 e_{k-\tau}. \end{aligned} \quad (33)$$

Noting that $\|\omega_k\| \leq \bar{\omega}$ and $\|\nu_k\| \leq \bar{\nu}$, it follows from (32) and (33) that

$$\begin{aligned} &\Delta V_k + \mu_1 V_k = V_{k+1} - V_k + \mu_1 V_k \\ &\leq \begin{bmatrix} e_k \\ e_{k-\tau} \\ \omega_k \\ \nu_k \end{bmatrix}^T \Upsilon \begin{bmatrix} e_k \\ e_{k-\tau} \\ \omega_k \\ \nu_k \end{bmatrix} + \lambda_1 \bar{\omega}^2 + \lambda_2 \bar{\nu}^2 \end{aligned} \quad (34)$$

where

$$\Upsilon \triangleq \begin{bmatrix} \Upsilon_{11} & \Upsilon_{12} & \Upsilon_{13} & \mathcal{A}^T(\theta_k)P_1 \mathcal{F}(\theta_k) \\ * & \Upsilon_{22} & \Upsilon_{23} & E^T P_1 \mathcal{F}(\theta_k) \\ * & * & \Upsilon_{33} & B^T P_1 \mathcal{F}(\theta_k) \\ * & * & * & \mathcal{F}^T(\theta_k)P_1 \mathcal{F}(\theta_k) - \lambda_2 I \end{bmatrix},$$

$$\Upsilon_{11} \triangleq \mathcal{A}^T(\theta_k)P_1\mathcal{A}(\theta_k) - (1-\mu_1)P_1 + P_2,$$
$$\Upsilon_{12} \triangleq \mathcal{A}^T(\theta_k)P_1 E, \quad \Upsilon_{22} \triangleq E^T P_1 E - (1-\mu_1)^\tau P_2,$$
$$\Upsilon_{13} \triangleq \mathcal{A}^T(\theta_k)P_1 B, \ \Upsilon_{23} \triangleq E^T P_1 B, \ \Upsilon_{33} \triangleq B^T P_1 B - \lambda_1 I.$$

Since $\mathcal{A}(1) = A - KC$ and $\mathcal{F}(1) = -KF$, it can be observed from (28) that $\Upsilon = \bar{\Upsilon}_1 < 0$, which implies that

$$\Delta V_k \leq -\mu_1 V_k + \lambda_1 \bar{\omega}^2 + \lambda_2 \bar{\nu}^2, \quad \text{if } \theta_k = 0 \qquad (35)$$

As such, for any $t(i) < k < t(i+1)$ and scalar $\delta > 0$, it is derived from (35) that

$$\delta^{k+1} V_{k+1} - \delta^k V_k = \delta^{k+1}\Delta V_k + \delta^k(\delta - 1)V_k$$
$$\leq \delta^k(\delta - \delta\mu_1 - 1)V_k + \delta^{k+1}(\lambda_1\bar{\omega}^2 + \lambda_2\bar{\nu}^2). \qquad (36)$$

Letting $\bar{\delta} \triangleq \frac{1}{1-\mu_1}$ and summing up both sides of (36) from $t(i)+1$ to $t(i+1)$ with respect to k, we obtain

$$\bar{\delta}^{t(i+1)} V_{t(i+1)} - \bar{\delta}^{t(i)+1} V_{t(i)+1} \leq (\lambda_1\bar{\omega}^2 + \lambda_2\bar{\nu}^2) \sum_{i=t(i)+2}^{t(i+1)} \bar{\delta}^i$$
$$= (\lambda_1\bar{\omega}^2 + \lambda_2\bar{\nu}^2)\frac{\bar{\delta}^{t(i)+2} - \bar{\delta}^{t(i+1)+1}}{1 - \bar{\delta}}$$

which implies that

$$V_{t(i+1)} \leq (\lambda_1\bar{\omega}^2 + \lambda_2\bar{\nu}^2)\frac{\bar{\delta}^{2-T_{i+1}} - \bar{\delta}}{1-\bar{\delta}} + \bar{\delta}^{1-T_{i+1}}V_{t(i)+1} \qquad (37)$$

Note that

$$\mathbb{E}\{\bar{\delta}^{-T_{i+1}}\} = \sum_{j=0}^{s}\text{Prob}\{T_{i+1} = \underline{T} + j\}\bar{\delta}^{-(\underline{T}+j)}$$
$$= \sum_{j=0}^{s}p_j\bar{\delta}^{-\underline{T}-j}.$$

Taking the conditional expectation on both sides of (37), one has

$$\mathbb{E}\{V_{t(i+1)}|t(i)\} \leq \rho + \bar{\delta}\vec{\delta}\mathbb{E}\{V_{t(i)+1}|t(i)\} \qquad (38)$$

where $\vec{\delta} \triangleq \sum_{j=0}^{s}p_j\bar{\delta}^{-\underline{T}-j}$ and

$$\rho \triangleq (\lambda_1\bar{\omega}^2 + \lambda_2\bar{\nu}^2)\frac{\bar{\delta}^2\vec{\delta} - \bar{\delta}}{1-\bar{\delta}}.$$

Taking the mathematical expectation on the inequality (38), we have

$$\mathbb{E}\{V_{t(i+1)}\} \leq \rho + \bar{\delta}\vec{\delta}\mathbb{E}\{V_{t(i)+1}\} \qquad (39)$$

Case 2: $\theta_k = 0$.
Obviously, the equality $\theta_k = 0$ means that there exists a nonnegative integer i satisfying $k = t(i)$. In this case, one has

$$\Delta V_k - \mu_2 V_k$$
$$= V_{1,k+1} - (1+\mu_2)V_{1,k} + V_{2,k+1} - (1+\mu_2)V_{2,k}$$
$$< V_{1,k+1} - (1+\mu_2)V_{1,k} + V_{2,k+1} - (1-\mu_1)V_{2,k}$$
$$\leq \begin{bmatrix} e_k \\ e_{k-\tau} \\ \omega_k \end{bmatrix}^T \bar{\Upsilon}_2 \begin{bmatrix} e_k \\ e_{k-\tau} \\ \omega_k \end{bmatrix} + \lambda_3\bar{\omega}^2 \leq \lambda_3\bar{\omega}^2$$

which implies that

$$V_{t(i)+1} \leq (1+\mu_2)V_{t(i)} + \lambda_3\bar{\omega}^2 \qquad (40)$$

Then, it follows from (39) and (40) that

$$\mathbb{E}\{V_{t(i+1)}\} \leq \bar{\rho} + \bar{\delta}\vec{\delta}(1+\mu_2)\mathbb{E}\{V_{t(i)}\} \qquad (41)$$

where $\bar{\rho} \triangleq \bar{\delta}\vec{\delta}\lambda_3\bar{\omega}^2 + \rho$. Furthermore, it is concluded from (30) that $\mathbb{E}\{V_{t(i+1)}\} \leq \gamma\mathbb{E}\{V_{t(i)}\} + \bar{\rho}$ where $\gamma < 1$, which implies that

$$\sigma^{i+1}\mathbb{E}\{V_{t(i+1)}\} - \sigma^i\mathbb{E}\{V_{t(i)}\}$$
$$\leq \sigma^i(\sigma - \sigma(1-\gamma) - 1)\mathbb{E}\{V_{t(i)}\} + \sigma^{i+1}\bar{\rho} \qquad (42)$$

Letting $\bar{\sigma} \triangleq \frac{1}{\gamma}$ and summing up both sides of (42) from $t(0)$ to $t(j)$ with respect to i, we obtain

$$\sigma^j\mathbb{E}\{V_{t(j)}\} - \mathbb{E}\{V_{t(0)}\} \leq \frac{\bar{\rho}\sigma(1-\sigma^j)}{1-\sigma} = \frac{\bar{\rho}(1-\gamma^j)}{\gamma^j(1-\gamma)}.$$

Moreover, it is found from (35) that

$$V_{t(0)} < \bar{\delta}^{-\bar{d}}V_0 + \frac{\bar{\delta}(\bar{\delta}^{-\bar{d}} - 1)}{1-\bar{\delta}}(\lambda_1\bar{\omega}^2 + \lambda_2\bar{\nu}^2) \triangleq \bar{V}_0$$

Hence, we arrive at

$$\mathbb{E}\{V_{t(j)}\} < \gamma^j\bar{V}_0 + \frac{\bar{\rho}(1-\gamma^j)}{1-\gamma} = \gamma^j\left(\bar{V}_0 - \frac{\bar{\rho}}{1-\gamma}\right) + \frac{\bar{\rho}}{1-\gamma} \qquad (43)$$

Therefore, the dynamics of $\mathbb{E}\{V_{t(j)}\}$ is ultimately bounded, that is,

$$\lim_{j\to+\infty} = \frac{\bar{\rho}}{1-\gamma} < +\infty.$$

Next, for any $t(j) \leq k < t(j+1)$, we have

$$\lambda_{\min}\{P_1\}\mathbb{E}\{\|e_k\|^2\} \leq \mathbb{E}\{V_k\}$$
$$\leq (\lambda_1\bar{\omega}^2 + \lambda_2\bar{\nu}^2)\frac{\bar{\delta}(\bar{\delta}^{1-k+t(j)})}{1-\bar{\delta}} + \bar{\delta}^{1-k+t(j)}\mathbb{E}\{V_{t(j)+1}\}$$
$$\leq (\lambda_1\bar{\omega}^2 + \lambda_2\bar{\nu}^2)\frac{\bar{\delta}(\bar{\delta}^{1-k+t(j)})}{1-\bar{\delta}} + \bar{\delta}^{1-k+t(j)}\lambda_3\bar{\omega}^2$$
$$+ \bar{\delta}^{1-k+t(j)}(1+\mu_2)\mathbb{E}\{V_{t(j)}\}$$
$$\leq (\lambda_1\bar{\omega}^2 + \lambda_2\bar{\nu}^2)\frac{\bar{\delta}^2}{1-\bar{\delta}} + \bar{\delta}\lambda_3\bar{\omega}^2 + \bar{\delta}(1+\mu_2)\mathbb{E}\{V_{t(j)}\}$$

which implies that the dynamics of the filtering error is ultimately bounded in mean square. Finally, we conclude that

$$\lim_{k\to+\infty}\mathbb{E}\{\|e_k\|^2\} \leq (\lambda_1\bar{\omega}^2 + \lambda_2\bar{\nu}^2)\frac{\bar{\delta}^2}{1-\bar{\delta}}$$
$$+ \bar{\delta}\lambda_3\bar{\omega}^2 + \bar{\delta}(1+\mu_2)\frac{\bar{\rho}}{1-\gamma}.$$

The proof is now complete. ∎

Remark 7: So far, we have designed the binary function θ_k and derived sufficient conditions to ensure the EUB (in mean square) of the filtering error. By employing a LMFD-based technique, we are able to examine whether the received measurement contains an outlier through a fixed number of past measurements. Based on the analysis results in Theorem 1, we have the following observations.

1) To ensure that the matrix inequality (28) is feasible, the scalar μ_1 is required to satisfy the condition $0 < \mu_1 < 1$.

2) The decay rate of $\mathbb{E}\{\|e_k\|^2\}$ is mainly determined by γ. Obviously, increasing the value of the lower bound of T_i (i.e. $\underline{T}$) would improve the feasibility of the matrix constraints (28)-(30), which indicates that a long minimum interval length of outliers would contribute to the improvement of the filtering performance.

3) Both the design of θ_k and the sufficient conditions derived in Theorem 1 are independent of the upper bound of the outlier o_k. Hence, our developed filtering strategy is capable of dealing with unbounded IMOs.

In Theorem 1, sufficient conditions have been proposed to ensure that the dynamics of the filtering error is ultimately bounded in mean square subject to the disturbances and the IMOs. Note that the constraints (28)-(30) are nonlinear matrix inequalities that are quite difficult to solve. In what follows, we aim to present a computationally appealing algorithm to determine the filter parameter K. Before doing so, we first give the following corollary concerning the design of the filter parameter K based on given scalars μ_1 and μ_2.

Corollary 1: Consider the filtering error dynamics given by (27). Suppose that Assumptions 1-2 hold. Let two scalars $0 < \mu_1 < 1$ and $\mu_2 > 0$ satisfying the constraint (30) be given. Assume that there exist two positive definite matrices $P_1, P_2 \in \mathbb{R}^{n \times n}$, three positive scalar λ_i ($i = 1, 2, 3$) and a matrix $\bar{K} \in \mathbb{R}^{n \times m}$ satisfying the constraint (29) and the following two matrix inequalities

$$\Phi_1 = \begin{bmatrix} \Phi_1^{11} & \Phi_1^{12} \\ * & \Phi_1^{22} \end{bmatrix} < 0 \qquad (44)$$

$$P_1 \geq M^T M \qquad (45)$$

where

$$\Phi_1^{11} \triangleq -\mathrm{diag}\big\{(1-\mu_1)P_1 - P_2, (1-\mu_1)^\tau P_2, \lambda_1 I, \lambda_2 I\big\},$$
$$\Phi_1^{12} \triangleq \begin{bmatrix} P_1 A - \bar{K} C & P_1 E & P_1 B & \bar{K} F \end{bmatrix}^T, \ \Phi_1^{22} \triangleq -P_1.$$

Then, the dynamics of the filtering error is ultimately bounded in mean square subject to the disturbances ω_k, ν_k and the measurement outlier o_k. Furthermore, the minimum of the asymptotic upper bound of $\mathbb{E}\{\|\tilde{z}_k\|^2\}$ can be derived by solving the following minimization problem:

$$\min\left\{(\lambda_1\bar{\omega}^2 + \lambda_2\bar{\nu}^2)\frac{\bar{\bar{\delta}}^2}{1-\bar{\delta}} + \bar{\delta}\lambda_3\bar{\omega}^2 + \bar{\delta}(1+\mu_2)\frac{\bar{\rho}}{1-\gamma}\right\} \qquad (46)$$

subject to the matrix inequality constraints (29), (44) and (45), where

$$\bar{\delta} \triangleq \frac{1}{1-\mu_1}, \ \vec{\delta} \triangleq \sum_{i=0}^{s} p_i \bar{\delta}^{-\underline{T}-i}, \ \rho \triangleq (\lambda_1\bar{\omega}^2 + \lambda_2\bar{\nu}^2)\frac{\bar{\delta}^2\vec{\delta} - \bar{\delta}}{1-\bar{\delta}},$$
$$\bar{\rho} \triangleq \bar{\delta}\vec{\delta}\lambda_3\bar{\omega}^2 + \rho,$$

and γ is defined in Theorem 1. Also, the desired filter parameter can be computed by $K = P_1^{-1}\bar{K}$.

Proof: The proof is straightforward based on Theorem 1 and Schur complement lemma, and is therefore omitted here for space saving. ∎

Remark 8: It is well known that the feasibility of a filter algorithm is largely dependent on the observability (or detectability) of the underlying system. Furthermore, the observability of the system is affected by the rank deficiency of the corresponding system matrix [11], [12]. Compared with the observability-based analysis technique, the matrix-inequality-based technique is more appropriate for handling the filter design issue addressed in this paper due mainly to the consideration of the impulsive measurement outliers. In fact, there should be some relationship between the feasibility of our proposed matrix inequalities and the rank deficiency of the system matrix, and this deserves further investigation in our future research.

By means of Corollary 1, we propose an algorithm (Algorithm 2) to design the filter parameter based on a particle-swarm-optimization (PSO) method as follows. Obviously, the computation of Algorithm 2 is quite costly due mainly to the solving process of the constrained minimization problem (46). Such a constrained minimization problem is solved by using the well-known linear matrix inequality (LMI) technique, which has a polynomial time complexity. In particular, the number $\mathscr{N}(\varepsilon)$ of flops needed to compute an ε-accurate solution is bounded by $O(\mathscr{M}\mathscr{N}^3 \log(\mathscr{V}/\varepsilon))$, where $\mathscr{M}$ is the total row size of the LMI system, $\mathscr{N}$ is the total number of scalar decision variables, $\mathscr{V}$ is a data-dependent scaling factor, and ε is the relative accuracy set for algorithm. For the LMIs (29), (44) and (45), one has both $\mathscr{M} = 8n + 2r + s + 3$ and $\mathscr{N} = n^2 + mn + 3$. Therefore, the computational complexity of the established LMIs is $O((8n + 2r + s + 3)(n^2 + mn + 3)^3 \log(\mathscr{V}/\varepsilon))$. On the other hand, it should be noted that the computational complexity of our proposed PSO-based algorithm is bounded by $O(2\bar{r}k_{\max}\mathscr{S})$, where $\bar{r}$ is the number of particles, $k_{\max}$ is the maximum number of iterations, $\mathscr{S}$ is the computational complexity of updating the particle and the global best location. As such, the whole computational complexity of Algorithm 2 is $O(2\bar{r}k_{\max}(8n+2r+s+3)(n^2+mn+3)^3 \log(\mathscr{V}/\varepsilon))$. Note that the computation of Algorithm 2 is executed off-line for the subsequent filtering task. The computational complexity of the filtering task will not be affected by such a high computational complexity.

Remark 9: It should be pointed out that the observability/detectability of the system is not required in outlier detection (via the design of θ_k) or the filter parameter design (via the design of K). In the proposed outlier detection scheme, the detection result is based on the input-output model (16) by using the LMFD of the transfer function matrix. The computation of such a LMFD is *not* dependent on the observability of the delay-free linear system (9). On the other hand, it is easy to see that the filter design task is accomplished by solving certain matrix inequalities, and such a procedure is quite different from the traditional methods (involving extensive observability analysis), and is therefore more suitable for handling optimization problems. In fact, by solving a constrained optimization problem outlined in Algorithm 2, the derived filter would help suppress the asymptotic upper bound of $\mathbb{E}\{\|\tilde{z}_k\|^2\}$.

Algorithm 2:

Step 1. Initialization: let $X_0^i \triangleq \begin{bmatrix} \mu_1 & \mu_2 \end{bmatrix}^T$ be the location of the at the i-th particle initial step. Generate $\bar{r}$ particles under which the the matrix inequalities (29), (30), (44) and (45) are feasible. The velocity of the i-th particle is set to be S_0^i. Let the maximum number of iterations be $k_{\max}$.

Step 2. Let the local best location of the i-th particle and the global best location at step k be $P_L^i = X_0^i$ and $P_G^k = 0$, respectively. Let the iteration step k be 0.

Step 3. Update the values of P_L^i and P_G^k: for the i-th particle, solve the minimization problem (46) subject to the matrix inequality constraints (29), (44), (45), and let the value to (46) be $\eta(X_k^i)$. Then, update the values of P_L^i and P_G^k by $P_L^i = \min_k \left\{ \eta(X_k^i) \right\}$ and $P_G^k = \min_i \left\{ P_L^i \right\}$, respectively. If $\|P_G^k - P_G^{k-1}\| < \xi$ where ξ is a given small positive scalar representing the computation accuracy, go to *Step 7*.

Step 4. Update the values of S_k^i and X_k^i as follows:
$S_{k+1}^i = \varphi S_k^i + \upsilon_1 r_i (P_L^i - X_k^i) + \upsilon_2 r_g (P_G^k - X_k^i)$
$X_{k+1}^i = X_k^i + S_{k+1}^i$
where φ, υ_1 and υ_2 are the given inertia parameter and two momentum parameters, respectively. r_i and r_g are the random numbers between $(0, 1)$. Let $k = k + 1$.

Step 5. Based on the values of μ_1^i and μ_2^i according to the i-th particle X_k^i, if the matrix inequalities (29), (30), (44) and (45) are infeasible, let $X_k^i = X_{k-1}^i$ and $S_k^i = S_{k-1}^i$.

Step 6. If $k < k_{\max}$, go back to *Step 3*, else update the values of P_L^i and P_G^k by $P_L^i = \min_k \left\{ \eta(X_k^i) \right\}$ and $P_G^k = \min_i \left\{ P_L^i \right\}$, respectively, and turn to the next step.

Step 7. Solve the minimization problem (46) subject to the matrix inequality constraints (29), (44) and (45) according to P_G^k. Calculate the desired filter parameter by $K = P_1^{-1} \bar{K}$. Stop.

C. The case of unknown statistical properties of interval length for outliers

So far, we have completed the filter design for the time-delay system (6) with random occurring IMOs, where the statistical properties of the outlier interval length are *known*. However, sometimes it is quite difficult to obtain these statistical properties in an exact way. In what follows, we consider the case where the statistical properties of the interval length for the impulsive outliers are completely unknown. In this case, the filter parameter K is designed under the worst situation (i.e. each interval length equals to $\underline{T}$) in order to ensure that the filtering error is exponentially ultimately bounded for *all possible* interval length. In other words, we aim to deal with the filtering issue *without* the condition proposed in Assumption 2.

Corollary 2: Consider the filtering error dynamics given by (27). Let Assumptions 1-2 hold and two scalars ($0 < \mu_1 < 1$ and $\mu_2 > 0$) satisfy the following constraint:

$$\tilde{\gamma} = (1 + \mu_2)(1 - \mu_1)^{\underline{T}-1} < 1. \tag{47}$$

If there exist two positive definite matrices P_1, $P_2 \in \mathbb{R}^{n \times n}$, three positive scalar λ_i ($i = 1, 2, 3$) and a matrix $\bar{K} \in \mathbb{R}^{n \times m}$ satisfying the matrix constraints (29), (44) and (45), then the dynamics of the filtering error is ultimately bounded subject to the disturbances ω_k, ν_k and the measurement outlier o_k. Furthermore, the minimum of the asymptotic upper bound of $\|\tilde{z}_k\|^2$ can be derived by solving the following minimization

problem:

$$\min \left\{ (\lambda_1 \bar{\omega}^2 + \lambda_2 \bar{\nu}^2) \frac{\bar{\delta}^2}{1 - \bar{\delta}} + \bar{\delta} \lambda_3 \bar{\omega}^2 + \bar{\delta}(1 + \mu_2) \frac{\tilde{\rho}}{1 - \tilde{\gamma}} \right\} \tag{48}$$

subject to the matrix inequality constraints (29), (44) and (45), where

$$\bar{\delta} \triangleq \frac{1}{1 - \mu_1}, \; \check{\rho} \triangleq (\lambda_1 \bar{\omega}^2 + \lambda_2 \bar{\nu}^2) \frac{\bar{\delta}^{2-\underline{T}} - \bar{\delta}}{1 - \bar{\delta}},$$
$$\tilde{\rho} \triangleq \bar{\delta}^{1-\underline{T}} \lambda_3 \bar{\omega}^2 + \check{\rho}.$$

The desired filter parameter can be computed by $K = P_1^{-1} \bar{K}$.

Proof: The proof is straightforward and omitted here for space saving. ∎

Similar to Algorithm 2, we can derive the desired filter parameter K by means of Corollary 2 (by employing the PSO algorithm) and the corresponding algorithm is omitted here for conciseness.

Remark 10: It is easy to conclude from Corollaries 1 and 2 that the filtering error results from the process and measurement noises. Obviously, the main purpose of the filter design is just to reduce the effects of noises on the state estimates according to certain performance specifications. When the underling system is undergoing unknown (but deterministic) noises, an effective way of estimating the system state is to employ the well-known unknown-input-observer-based (UIO-based) approach whose main idea is to decouple the estimation error from the unknown (but external) inputs [25]. One of our future research topics is to consider the state estimation problem subject IMOs by using the UIO-based strategy.

Remark 11: It is observed from Corollary 2 that 1) the existence of the filter is determined by the condition (47), whose feasibility depends on the parameter $\underline{T}$, 2) increasing the value of $\underline{T}$ would contribute to the feasibility of the condition (47) and the suppression of the minimum asymptotic upper bound $\|\tilde{z}_k\|^2$, and 3) the condition (47) is independent of the scalar s. In fact, the results in Corollary 2 can be easily extended to the case that $s = +\infty$. Moreover, it should be pointed out that the condition (47) is stricter than the condition (30), which means that the minimum asymptotic upper derived in Corollary 2 would be larger than that in Corollary 1. In other words, the statistical properties of interval length for outliers would contribute to the resulted filtering performance. The distinctive novelty of this paper lies on the follow three aspects: 1) a new outlier detection technique is developed to distinguish the measurements corrupted by outliers from those normal measurements; 2) a new PD filter is designed to guarantee that the filtering error is ultimately bounded in mean square; 3) the desired filter parameter is achieved by combing the PSO algorithm and LMI techniques; and 4) the proposed filter design strategy is applied to deal with the scenario with unknown statistical properties of interval length for the IMOs.

Remark 12: In this paper, we have dedicatedly designed an "outlier-resistant" filter for a class of linear time-delay systems subject to IMOs, and such kind of filters is fundamentally different from the existing "attack-resilient" filters. For example, in [19], [20], the distributed filtering problems have been studied for linear time-invariant systems subject to sensor attacks,

where a saturation-like scheme has been adopted to construct the filters by assuming that the maximal number of attacked sensors is known, and this maximal number is employed to indicate the "level of attacks". In this paper, instead of making any assumption on the number of the "problematic" sensors, we exploit the information about interval lengths of outliers in order to characterize the "occurrence frequency of outliers". Moreover, different from the saturation-like scheme adopted in [19], [20], the PD filtering scheme proposed in this paper is developed based on the idea of "decoupling-like" scheme, which has proven to be particularly suitable for IMOs.

IV. AN ILLUSTRATIVE EXAMPLE

In this section, we provide two numerical example to illustrate the effectiveness of the proposed filter design scheme.

Example 1: To make our simulation nontrivial, we consider an *unstable* linear time-delay system of the form (6) with

$$A = \begin{bmatrix} 0.65 & 0.38 \\ 0.32 & -0.53 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ -0.7 \end{bmatrix}, \quad E = \begin{bmatrix} 0.3 & 0.2 \\ -0.1 & 0.2 \end{bmatrix},$$
$$C = \begin{bmatrix} 1 & 0.5 \end{bmatrix}, \quad F = 0.5, \quad M = I, \quad \tau = 1.$$

The process noise and measurement noise are chosen as follows:

$$\omega_k = 0.4 \sin(0.4k^{1.2}), \quad \nu_k = 0.3 r_m(k)$$

where $r_m(k)$ is a random number at time instant k satisfying $r_m(k) \in [0,1]$. Applying the augmentation method and the LMFD technology introduced in Subsection III-A, we have

$$\mathcal{D}_L = \begin{bmatrix} 1 & -0.12 & -0.966 & -0.055 & 0.08 \end{bmatrix},$$
$$\mathcal{N}_L = \begin{bmatrix} 0.65 & 0.6515 & -0.285 & 0 \end{bmatrix},$$

and $\bar{d} = 4$. In this illustrative example, we assume that $\underline{T} = 6$ and $s = 4$. Then, the threshold $\bar{f}$ is computed as follows:

$$\bar{f} = \|\mathcal{N}_L\|\bar{d}\bar{\omega} + \|\mathcal{D}_L\mathcal{F}\|(\bar{d}+1)\bar{\nu} = 2.5907.$$

Let the lower bound of the IMOs be $\underline{o} = 6.4768$. Obviously, we have $\underline{o} = 2.5\bar{f} > 2\bar{f}$. Based on the design of the binary function θ_k in Proposition 1, the IMOs and the values of $\{1 - \theta_k\}_{k \geq 0}$ are shown in Fig. 1. It can be observed from Fig. 1 that our developed $\{\theta_k\}_{k \geq 0}$ is capable of identifying whether the received measurement contains an outlier ($1 - \theta_k = 1$ holds when $o_k \neq 0$).

Next, let us design the filter parameter K by applying the developed PSO-based Algorithm 2. Set the inertia parameter and two momentum parameters be 0.6, 0.7 and 0.7, respectively. The statistical properties of the intervals T_i are given as follows:

$$\text{Prob}\{T_i = \underline{T}\} = 0.1, \quad \text{Prob}\{T_i = \underline{T} + 1\} = 0.1,$$
$$\text{Prob}\{T_i = \underline{T} + 2\} = 0.2, \quad \text{Prob}\{T_i = \underline{T} + 3\} = 0.4,$$
$$\text{Prob}\{T_i = \underline{T} + 4\} = 0.2.$$

Then, according the derived the global best location, we have $\mu_1 = 0.1753$ and $\mu_2 = 0.5331$. Furthermore, based on the derived values of μ_1 and μ_2, the minimum of the asymptotic upper bound of can be derived by solving the minimization
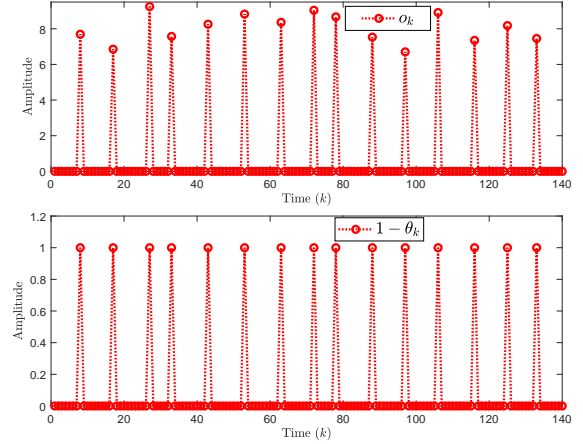

Fig. 1: The impulsive measurement outliers and the values of the sequence $\{1 - \theta_k\}_{k \geq 0}$

problem in Corollary 1. The corresponding filter parameter K is given as follows:

$$K = \begin{bmatrix} 0.36594 \\ 0.02054 \end{bmatrix}.$$

Based on the derived filter parameter and the developed function θ_k, we can design the PD filter according to (7). Numerical simulation results are given in Figs. 2-3, which depict the state trajectories and their corresponding estimates. All the simulation results confirm that the filtering performance is well achieved.
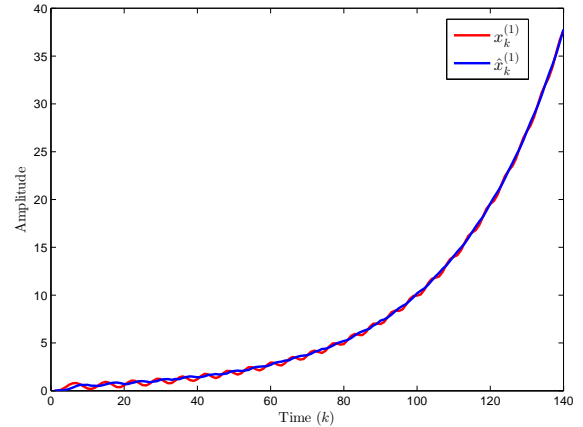

Fig. 2: The state trajectories of $x_k^{(1)}$ and $\hat{x}_k^{(1)}$

In order to show the superiority of our designed filtering scheme, we would like to conduct a comparative simulation between our developed PD filtering scheme and the traditional Luenberger-observer-based filter. Let the filter gain matrix of the traditional Luenberger-observer-based filter be K, which is exactly the same parameter of our PD filter when $\theta_k = 0$. The trajectories of $\|\tilde{z}_k\|$ under different filtering methods are shown in Fig. 4. Table I shows the values of $\sum_{k=0}^{420} \|\tilde{z}_k\|$ of our developed PD filtering and the traditional Luenberger-observer-based filtering. It can be easily observed from Fig. 4 and Table I that our developed PD filtering scheme performs
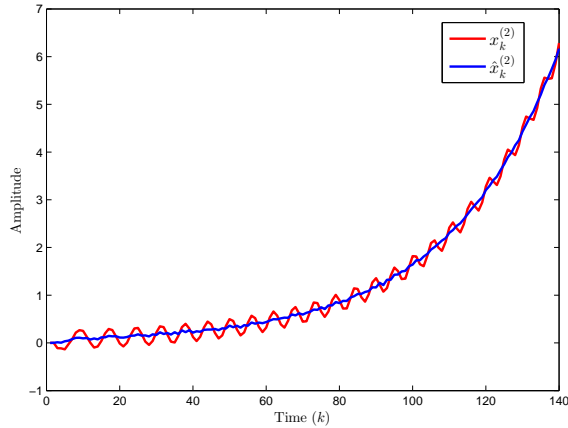
Fig. 3: The state trajectories of $x_k^{(2)}$ and $\hat{x}_k^{(2)}$

much better than the traditional Luenberger-observer-based filter. The main reason of such a performance superiority lies in the fact that our developed filtering could ensure the "rejection" of the impulsive outliers in the filtering process.

TABLE I: Comparison with the traditional filtering algorithm

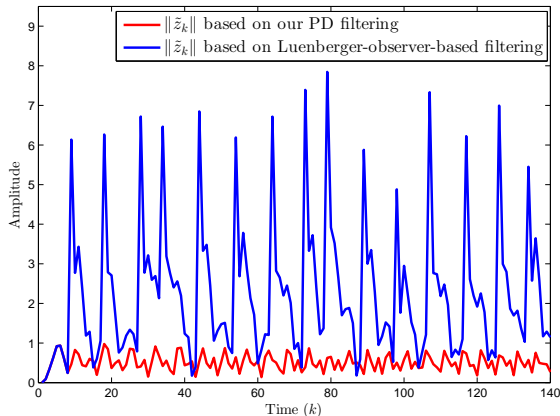|	$\sum_{k=0}^{420} \|\tilde{z}_k\|^2$
Our developed PD filtering	125.7173
Luenberger-observer-based filtering	3230.4203

Fig. 4: The trajectories of $\|\tilde{z}_k\|$ under different filtering approach

In what follows, we would like to consider the case where the statistical properties of interval length for the outliers are *completely unknown*. In this case, by employing Corollary 2, we can obtain the corresponding values of μ_1, μ_2 and K as follows:

$$\mu_1 = 0.46117, \quad \mu_2 = 18.98175, \quad K = \begin{bmatrix} 0.70198 \\ -0.33229 \end{bmatrix}.$$

The trajectories of $\|\tilde{z}_k\|^2$ with known and unknown statistical properties are shown in Fig. 5, from which we can find that the filtering with known statistical properties could achieve a smaller upper bound of the filtering error. Table II shows the values of $\sum_{k=0}^{420} \|\tilde{z}_k\|^2$ of our developed PD filtering schemes with known and unknown statistical properties, respectively. Obviously, a better filtering performance can be achieved if

the statistical properties of the interval length for the outliers are known.
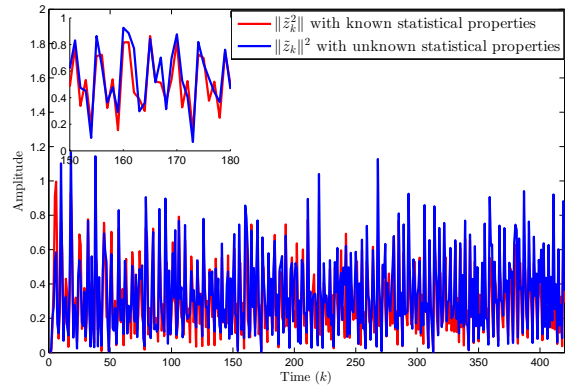

Fig. 5: The trajectories of $\|\tilde{z}_k\|^2$ under different filtering approach

TABLE II: Filtering performance with known and unknown statistical properties

	$\sum_{k=0}^{420} \|\tilde{z}_k\|^2$
Known statistical properties	125.7173
Unknown statistical properties	145.0910

According to the main results proposed in Theorem 1, it is concluded that the existence of the desired filter is affected by the time-delay parameter τ. In order to show the effect of time-delay on the filtering performance, we would like to provide a simulation to show the filtering results subject to different values of the time-delay parameter (i.e. $\tau = 1, 2, 3$). Let $\underline{T} = 9$ and $s = 4$. The corresponding trajectories of $\|\tilde{z}_k\|$ are shown in Fig. 6. Table III shows the corresponding values of $\sum_{k=0}^{420} \|\tilde{z}_k\|$. It is easy to find that reducing the value of τ would contribute to the improvement of the filtering performance.
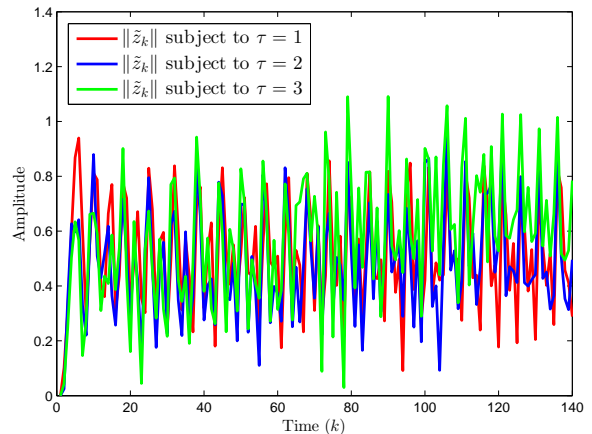

Fig. 6: The trajectories of $\|\tilde{z}_k\|$ with different values of τ

Example 2: To show the applicability of our developed filtering scheme in practical systems, let us consider the following two-degree-of-freedom quarter-car suspension systems

TABLE III: Filtering results subject to different values of τ

τ	$\sqrt{\sum_{k=0}^{420} \|\tilde{z}_k\|^2}$
$\tau = 1$	11.1775
$\tau = 2$	12.8546
$\tau = 3$	18.1281

(2DOF-QSSs) [45]:

$$A = \begin{bmatrix} -0.1055 & 0.0085 & -0.0022 & 0.0217 \\ -0.2260 & 0.1669 & 0.0028 & 0.1123 \\ 8.2196 & -2.7734 & -0.0139 & -0.0831 \\ -0.4540 & -6.5121 & -0.0109 & -0.0482 \end{bmatrix},$$

$$B = \begin{bmatrix} -1.0970 \\ -1.0591 \\ 5.4462 \\ -6.9661 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad F = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix},$$

$$E = 0, \quad \tau = 0, \quad M = I.$$

The process noise and measurement noise are chosen as follows:

$$\omega_k = 0.3\cos(0.6k), \quad \nu_k = 0.2\sin(0.4k^{1.1}).$$

Then, by using our developed PSO-based Algorithm 2, the desired filter gain matrix K is obtained as follows:

$$K = \begin{bmatrix} -0.2125 & -0.1182 \\ 0.1344 & 0.0469 \\ 4.1958 & -4.9949 \\ 2.7638 & -4.3444 \end{bmatrix}.$$

The trajectories of $\|\tilde{z}_k\|$ under different filtering schemes are shown in Fig. 7, which has demonstrated the effectiveness of our developed PD filtering scheme.
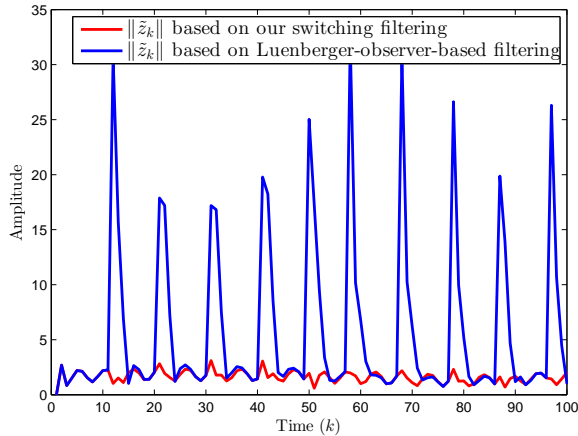

Fig. 7: The trajectories of $\|\tilde{z}_k\|$ under different filtering schemes

V. CONCLUSION

In this paper, the ultimately bounded filtering problem has been studied for a class of linear time-delay systems subject to IMOs. The measurement outliers under consideration are modeled by a sequence of impulsive signals whose interval lengths (i.e. the durations between two adjacent outliers) are assumed to be a sequence of independent and identically distributed (i.i.d.) random variables. In order to restrain the filtering performance from being degraded by IMOs, a special PD filter has been designed based on an active detection-based framework to discard the "harmful" measurements (i.e. the measurements corrupted by outliers). A novel outlier detection approach has been developed to distinguish these "harmful" measurements from other normal measurements. Then, the EUB of the filtering error in mean square has been carried out by combining the stochastic analysis technique and Lyapunov-like approach. A PSO-based algorithm has been developed to acquire the desired filter parameter. Furthermore, the filter design scheme has been extended to the scenario with unknown statistical properties for the measurement outliers (i.e. the occurrence probabilities of the outlier interval lengths are completely unknown). Finally, an illustrative example has been provided to demonstrate the effectiveness of our developed PD filtering scheme.

It is worth noted that, our developed PD filtering scheme is typical model-based method dealing with the linear time-delay systems with IMOs. The key technique is to obtain the input-output model based on the transfer function matrices of the plat. Accordingly, such a technique is not applicable to some complex systems (e.g. nonlinear system, time-varying systems, uncertain systems). Furthermore, it should be noted that the desired filter parameter is designed by using the matrix-inequality-based technique. The corresponding computational cost would be very high if the system dimension n is large. Such two deficiencies limit the application of our developed PD filtering scheme. Further research topics include 1) the extension of the main results to some complex systems such as time-varying systems, networked systems and uncertain systems [40], [42], [43], [46]; 2) the reduction of the computational cost for the filter parameter design issue; and 3) the filtering problem of networked systems subject to cyber-attacks [24].

REFERENCES

[1] A. D. Akkaya and M. L. Tiku, Robust estimation in multiple linear regression model with non-Gaussian noise, *Automatica*, vol. 44, no. 2, pp. 407-417, Feb. 2008.

[2] A. Alessandri and M. Awawdeh, Moving-horizon estimation with guaranteed robustness for discrete-time linear systems and measurements subject to outliers, *Automatica*, vol. 67, pp. 85-93, May 2016.

[3] A. Alessandri and L. Zaccarian, Stubborn state observers for linear time-invariant systems, *Automatica*, vol. 88, pp. 1-9, Feb. 2018.

[4] R. J. Anderson and M. W. Spong, Bilateral control of teleoperators with time delay, *IEEE Transactions on Automatic Control*, vol. 34, no. 5, pp. 494-501, May 1989.

[5] M. V. Basin, A. G. Loukianov and M. Hernandez-Gonzalez, Joint state and parameter estimation for uncertain stochastic nonlinear polynomial systems, *International Journal of Systems Science*, vol. 44, no. 7, pp. 1200-1208, 2013.

[6] R. Caballero-Águila, I. García-Garrido and J. Linares-Pérez, Information fusion algorithms for state estimation in multi-sensor systems with correlated missing measurements, *Applied Mathematics and Computation*, vol. 226, pp. 548-563, 2014.

[7] S. Cabuk, Simple test of hypotheses on system availability and mean time to repair, *IEEE Transactions on Reliability*, vol. 35, no. 5, pp. 581-583, Dec. 1986.

[8] L. Chang and K. Li, Unified form for the robust Gaussian information filtering based on M-estimate, *IEEE Signal Processing Letters*, vol. 24, no. 4, pp. 412-416, Feb. 2017.

[9] D. Ciuonzo, A. Aubry and V. Carotenuto, Rician MIMO channel- and jamming-aware decision fusion, *IEEE Transactions on Signal Processing*, vol. 65, no. 15, pp. 3866-3880, Mar. 2017.

[10] Y. Dai, Y. Zhou and Y. Jia, Distribution of time between failures of machining center based on type I censored data, *Reliability Engineering and System Safety*, vol. 79, no. 3, pp. 377-379, Mar. 2003.

[11] M. Doostmohammadian and U. A. Khan, On the distributed estimation of rank-deficient dynamical systems: A generic approach, in *Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver, BC, Canada, pp. 4618-4622, May 2013.

[12] M. Doostmohammadian, U. A. Khan, Measurement partitioning and observational equivalence in state estimation, in *Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, China, pp. 4855-4859, Mar. 2016.

[13] A. Gabriel, J. I. Nieto and E. M. Nebot, An outlier-robust Kalman filter, In *Proceedings of the 2011 IEEE International Conference on Robotics and Automation*, Shanghai, China, pp. 1151-1158, May 2011.

[14] M. A. Gandhi and L. Mili, Robust Kalman filter based on a generalized maximum-likelihood-type estimator, *IEEE Transactions on Signal Processing*, vol. 58, no. 5, pp. 2509-2520, May 2010.

[15] H. Gao and T. Chen, $\mathcal{H}_\infty$ estimation for uncertain systems with limited communication capacity, *IEEE Transactions on Automatic Control*, vol. 52, no. 11, pp. 2070-2084, Nov. 2007.

[16] X. Ge and Q.-L. Han, Distributed event-triggered $\mathcal{H}_\infty$ filtering over sensor networks with communication delays, *Information Sciences*, vol. 291, pp. 128-142, Jan. 2015.

[17] R. G. Gibbs, New Kalman filter and smoother consistency tests, *Automatica*, vol. 49, no. 10, pp. 3141-3144, Oct. 2013.

[18] A. Gutiérrez-Giles and M. A. Arteaga-Pérez, GPI based velocity/force observer design for robot manipulators, *ISA Transactions*, vol. 53, no. 4, pp. 929-938, Jul. 2014.

[19] X. He, X. Ren, H. Sandberg and K. H. Johansson, How to secure distributed filters under sensor attacks? *arXiv preprint*, arXiv:2004.05409, 2020.

[20] X. He, X. Ren, H. Sandberg and K. H. Johansson, Secure distributed filtering for unstable dynamics under compromised observations, In *Proceedings of 2019 IEEE 58th Conference on Decision and Control (CDC)*, Nice, France, pp. 5344-5349, Dec. 2019.

[21] Y. Huang, Y. Zhang, N. Li and J. Chambers, Robust Student's t based nonlinear filter and smoother, *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 5, pp. 2586-2596, Oct. 2016.

[22] J. H. Kotecha and P. M. Djurić, Gaussian particle filtering, *IEEE Transactions on Signal Processing*, vol. 51, no. 10, pp. 2592-2601, Oct. 2003.

[23] H. Liu, L. Ma, Z. Wang, Y. Liu and F. E. Alsaadi, An overview of stability analysis and state estimation for memristive neural networks, *Neurocomputing*, vol. 391, pp. 1-12, May 2020.

[24] L. Liu, L. Ma, J. Zhang and Y. Bo, Sliding mode control for nonlinear Markovian jump systems with Denial-of-Service attacks, *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 6, pp. 1638-1648, Nov. 2020.

[25] X. Liu, Z. Gao and A. Zhang, Observer-based fault estimation and tolerant control for stochastic Takagi-Sugeno fuzzy systems with Brownian parameter perturbations, *Automatica*, vol. 102, pp. 137-149, Apr. 2019.

[26] N. Lotfi and M. Namvar, Global adaptive estimation of joint velocities in robotic manipulators, *IET Control Theory & Applications*, vol. 4, no. 12, 2672-2681, Dec. 2010.

[27] D. G. Luenberger, Observing the state of a linear system, *IEEE Transactions on Military Electronics*, vol. 8, no. 2, pp. 74-80, Apr. 1964.

[28] L. Ma, Z. Wang, J. Hu and Q-L. Han, Probability-guaranteed envelope-constrained filtering for nonlinear systems subject to measurement outliers, *IEEE Transactions on Automatic Control*, in press, DOI: 10.1109/TAC.2020.3016767.

[29] J. Mao, D. Ding, G. Wei and H. Liu, Networked recursive filtering for time-delayed nonlinear stochastic systems with uniform quantisation under Round-Robin protocol, *International Journal of Systems Science*, vol. 50, no. 4, pp. 871-884, Mar. 2019.

[30] R. Marino, G. L. Santosuosso and P. Tomei, Robust adaptive observers for nonlinear systems with bounded disturbances, *IEEE Transactions on Automatic Control*, vol. 46, no. 6, pp. 967-972, Jun. 2001.

[31] M. Moayedi, Y. K. Foo and Y. C. Soh, Adaptive Kalman filtering in networked systems with random sensor delays, multiple packet dropouts and missing measurements, *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1577-1588, Mar. 2010.

[32] R. Mehra, On the identification of variances and adaptive Kalman filtering, *IEEE Transactions on Automatic Control*, vol. 15, no. 2, pp. 175-184, Apr. 1970.

[33] A. A. Prasov and H. K. Khalil, A nonlinear high-gain observer for systems with measurement noise in a feedback control framework, *IEEE Transactions on Automatic Control*, vol. 58, no. 3, pp. 569-580, Mar. 2013.

[34] D. E. Quevedo, A. Ahlén, A. S. Leong and S. Dey, On Kalman filtering over fading wireless channels with controlled transmission powers, *Automatica*, vol. 48, no. 7, pp. 1306-1316, Jul. 2012.

[35] H. Shen, Z.-G. Wu and J. H. Park, Reliable mixed passive and $\mathcal{H}_\infty$ filtering for semi-Markov jump systems with randomly occurring uncertainties and sensor failures, *International Journal of Robust and Nonlinear Control*, vol. 25, no. 17, pp. 3231-3251, Nov. 2015.

[36] Y. Shen, Z. Wang, B. Shen, F. E. Alsaadi and F. E. Alsaadi, Fusion estimation for multi-rate linear repetitive processes under weighted Try-Once-Discard protocol, *Information Fusion*, vol. 55, pp. 281-291, Mar. 2020.

[37] S. Sun, T. Tan and H. Lin, Optimal linear estimators for systems with finite-step correlated noises and packet dropout compensations, *IEEE Transactions on Signal Processing*, vol. 64, no. 21, pp. 5672-5681, 2016.

[38] C. Xu, S. Zhao, B. Huang and F. Liu, Distributed Student's t filtering algorithm for heavy-tailed noises, *International Journal of Adaptive Control and Signal Processing*, vol. 32, no. 6, pp. 875-890, Jun. 2018.

[39] X. Wan, Z. Wang, M. Wu and X. Liu, $\mathcal{H}_\infty$ state estimation for discrete-time nonlinear singularly perturbed complex networks under the Round-Robin protocol, *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 2, pp. 415-426, Feb. 2019.

[40] F. Wang and J. Liang, Constrained $\mathcal{H}_\infty$ estimation for time-varying networks with hybrid incomplete information, *International Journal of Robust and Nonlinear Control*, vol. 28, no. 2, pp. 699-715, 2018.

[41] Z. Wang, J. Lam and X. Liu, Filtering for a class of nonlinear discrete-time stochastic systems with state delays, *Journal of Computational and Applied Mathematics*, vol. 201, no. 1, pp. 153-163, Sep. 2007.

[42] G. Wei, S. Liu, Y. Song and Y. Liu, Probability-guaranteed set-membership filtering for systems with incomplete measurements, *Automatica*, vol. 60, pp. 12-16, Oct. 2015.

[43] Y. Xu, R. Lu, P. Shi, H. Li and S. Xie, Finite-time distributed state estimation over sensor networks with round-robin protocol and fading channels, *IEEE Transactions on Cybernetics*, vol. 48, no. 1, pp. 336-345, Jan. 2018.

[44] F. Yang and Y. Li, Set-membership fuzzy filtering for nonlinear discrete-time systems, *IEEE Transactions on Systems, Man, and Cybernetics, Part B-Cybernetics*, vol. 40, no. 1, pp. 116-124, Feb. 2010.

[45] H. Zhang, Q. Hong, H. Yan, F. Yang and G. Guo, Event-based distributed H_∞ filtering networks of 2-DOF quarter-car suspension systems, *IEEE Transactions on Industrial Informatics*, vol. 13, no. 1, pp. 312-321, Feb. 2017.

[46] S. Zhang, D. Ding, G. Wei, Y. Liu and F. E. Alsaadi, $\mathcal{H}_\infty$ state estimation for artificial neural networks over redundant channels, *Neurocomputing*, vol. 226, pp. 117-125, Feb. 2017.

[47] D. Zheng, *Linear System Theory (Second Edition)*. Peking, P. R. China: Tsinghua University Press, 2002.

[48] L. Zou, Z. Wang, J. Hu and D. Zhou, Moving horizon estimation with unknown inputs under dynamic quantization effects, *IEEE Transactions on Automatic Control*, vol. 65, no. 12, pp. 5368-5375, Dec. 2020.

[49] L. Zou, Z. Wang, Q.-L. Han and D. Zhou, Moving horizon estimation for networked time-delay systems under Round-Robin protocol, *IEEE Transactions on Automatic Control*, vol. 64, no. 12, pp. 5191-5198, Dec. 2019.

[50] L. Zou and Z. Wang, Set-membership filtering subject to impulsive measurement outliers: a recursive algorithm, *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 2, pp. 377-288, Jan. 2021.

Lei Zou received the B.Sc. degree in automation from Beijing Institute of Petrochemical Technology, Beijing, China, in 2008, the M.Sc. degree in control science and engineering from China University of Petroleum (Beijing Campus), Beijing, China, in 2011 and the Ph.D degree in control science and engineering in 2016 from Harbin Institute of Technology, Harbin, China. From October 2013 to October 2015, he was a visiting Ph.D. student with the Department of Computer Science, Brunel University London, Uxbridge, U.K.

Since 2019, he has been working as a Research Fellow with the Department of Computer Science, Brunel University London, Uxbridge, Middlesex, U.K. His research interests include control and filtering of networked systems, moving-horizon estimation, and state estimation subject to outliers.

Dr. Zou is currently serving as a reviewer for *Mathematical Reviews*, an Associate Editor for *Neurocomputing*, *International Journal of Systems Science* and *International Journal of Control, Automation and Systems*; a Senior Member of IEEE and a Member of Chinese Association of Automation; and a very active reviewer for many international journals.

Hongli Dong received the Ph.D. degree in control science and engineering from the Harbin Institute of Technology, Harbin, China, in 2012.

From 2009 to 2010, she was a Research Assistant with the Department of Applied Mathematics, City University of Hong Kong, Hong Kong. From 2010 to 2011, she was a Research Assistant with the Department of Mechanical Engineering, The University of Hong Kong, Hong Kong. From 2011 to 2012, she was a Visiting Scholar with the Department of Information Systems and Computing, Brunel University London, London, U.K. From 2012 to 2014, she was an Alexander von Humboldt Research Fellow with the University of Duisburg–Essen, Duisburg, Germany. She is currently a Professor with the Artificial Intelligence Energy Research Institute, Northeast Petroleum University, Daqing, China. She is also the Director of the Heilongjiang Provincial Key Laboratory of Networking and Intelligent Control, Daqing. Her current research interests include robust control and networked control systems.

Dr. Dong is a very active reviewer for many international journals.

Zidong Wang (SM'03-F'14) was born in Jiangsu, China, in 1966. He received the B.Sc. degree in mathematics in 1986 from Suzhou University, Suzhou, China, and the M.Sc. degree in applied mathematics in 1990 and the Ph.D. degree in electrical engineering in 1994, both from Nanjing University of Science and Technology, Nanjing, China.

He is currently Professor of Dynamical Systems and Computing in the Department of Computer Science, Brunel University London, U.K. From 1990 to 2002, he held teaching and research appointments in universities in China, Germany and the UK. Prof. Wang's research interests include dynamical systems, signal processing, bioinformatics, control theory and applications. He has published more than 600 papers in international journals. He is a holder of the Alexander von Humboldt Research Fellowship of Germany, the JSPS Research Fellowship of Japan, William Mong Visiting Research Fellowship of Hong Kong.

Prof. Wang serves (or has served) as the Editor-in-Chief for *International Journal of Systems Science*, the Editor-in-Chief for *Neurocomputing*, and an Associate Editor for 12 international journals including *IEEE Transactions on Automatic Control*, *IEEE Transactions on Control Systems Technology*, *IEEE Transactions on Neural Networks*, *IEEE Transactions on Signal Processing*, and *IEEE Transactions on Systems, Man, and Cybernetics-Part C*. He is a Member of the Academia Europaea, a Fellow of the IEEE, a Fellow of the Royal Statistical Society and a member of program committee for many international conferences.

Jun Hu received the B.Sc. degree in information and computation science and M.Sc. degree in applied mathematics from Harbin University of Science and Technology, Harbin, China, in 2006 and 2009, respectively, and the Ph.D. degree in control science and engineering from Harbin Institute of Technology, Harbin, China, in 2013.

From September 2010 to September 2012, he was a Visiting Ph.D. Student in the Department of Information Systems and Computing, Brunel University, U.K. From May 2014 to April 2016, he was an Alexander von Humboldt research fellow at the University of Kaiserslautern, Kaiserslautern, Germany. He is with the Department of Mathematics, Harbin University of Science and Technology, Harbin 150080, China, and also with the School of Engineering, University of South Wales, Pontypridd CF37 1DL, United Kingdom. His research interests include nonlinear control, filtering and fault estimation, time-varying systems and complex networks. He has published more than 60 papers in refereed international journals.

Dr. Hu serves as a reviewer for *Mathematical Reviews*, as an editor for *Neurocomputing*, *Journal of Intelligent and Fuzzy Systems*, *Neural Processing Letters*, *Systems Science and Control Engineering*, and as a guest editor for *International Journal of General Systems* and *Information Fusion*.