IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received March 4, 2021, accepted March 10, 2021, date of publication March 15, 2021, date of current version March 23, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3066109

Distributed Fault Detection for Linear Time-Varying Multi-Agent Systems With Relative Output Information

PEILU ZOU, PING WANG, AND CHENGPU YU
School of Automation, Beijing Institute of Technology, Beijing 100081, China
Chongqing Innovation Center, Beijing Institute of Technology, Chongqing 401147, China

Corresponding author: Ping Wang (pingwang0906@126.com)

This work was supported in part by the Beijing Natural Science Foundation under Grant L201006; in part by the National Natural Science Foundation of China under Grant 61991414, Grant 61873301, and Grant 62088101; and in part by the National Key Research and Development Project under Grant 2020YFC1512500.

ABSTRACT This paper investigates the distributed fault detection problem for linear discrete time-varying heterogeneous multi-agent systems under relative output information. Due to the lack of absolute outputs, an augmented model is built by stacking all local relative output information. Then, the fault detection problem consisting of residual-generation and residual-evaluation is handled using the H_∞ filtering framework. The residual-generation problem is actually a minimization problem of an indefinite quadratic form, and the Krein space-Kalman filtering theory is applied, which results in a low computational burden despite the time-varying characteristic. Using the Krein space theory, a necessary and sufficient condition for the minimum is derived, and a residual-generation algorithm is developed. Further, a residual-evaluation mechanism is designed by constructing an evaluation function and detecting faults by comparing it with a threshold. Finally, two illustrative examples are given to demonstrate the effectiveness of the proposed fault detection approach.

INDEX TERMS Discrete-time systems, distributed algorithms, fault detection, fault diagnosis, filtering theory, heterogeneous networks, linear systems, minimization, multi-agent systems, time-varying systems.

I. INTRODUCTION

With the increasing complexity of work environments and task scales, it is difficult for traditional research on a single controlled object to meet the actual demand. Accordingly, multi-agent systems (MASs) have received considerable attention due to their characteristics of autonomy [1], [2], distribution [3], [4] and coordination [5], [6]. Control problems of MASs have been extensively studied such as adaptive control [7], [8] and event-triggered control [9]–[11], and many results have been obtained for communication delay [12], [13]. Moreover, the fault-tolerant control problems for MASs has been recently studied by many researchers [14]–[16] because MASs are vulnerable to faults due to their structural complexity. However, prior to employing the fault-tolerant control technique, it is necessary to confirm the occurrence of faults. If the fault occurring on one agent node is not detected and dealt with in time, it will spread to the other nodes through the network, giving rise to a huge threat to the security of the whole MAS. Therefore, the investigation of fault detection (FD) for MASs is highly urgent.

During the past few decades, distributed FD approaches have been developed for networked systems using local information only. [17] solved the fault isolation problem for discrete-time fuzzy interconnected systems with unknown interconnections. In [18], an observer-based fault estimation methodology was proposed for leader-following linear multi-agent systems subject to actuator faults. [19] dealt with the FD problem for the discrete-time Markovian jump linear system with a stochastic packet dropping effect. In [20] and [21], the H_∞ optimization was used for FD of linear time-invariant systems where the generated residuals were sensitive to faults while robust against disturbances and noises. In [22] and [23], unknown input observers were used to solve the FD problems for second-order MASs. In [24] and [25], linear matrix inequalities were applied to obtain sufficient conditions for

The associate editor coordinating the review of this manuscript and approving it for publication was Gerard-Andre Capolino.

the solvability of FD problems. In [26], an adaptive threshold method and a sliding mode observer were used to solve the FD problem for the physical layer network in cyber-physical systems, and [27] also designed an adaptive threshold according to the dynamics of residual vectors to detect faults in networked robots. Moreover, the time delay phenomenon has attracted much attention in FD problems for MASs, see [28]–[32] for examples.

So far, the FD of time-varying MASs has attracted much attention, and a number of results have been reported, particularly for linear discrete time-varying (LDTV) systems. In [33], the Kalman filter was used to solve the FD problem for a class of LDTV networked sensing systems where the disturbances and noises were assumed to satisfy the known probability distributions. In [34], the FD problem for time-varying sensor networks with multiplicative noises was investigated based on the least-squares approach. In [35], the distributed fault estimation problem of MASs with sensor faults and partially decoupled disturbances was treated using the unknown input observer and recursive linear matrix inequalities. Moreover, [36] investigated the finite-horizon distributed H_∞ fault estimation problem for a class of LDTV sensor networks based on the Krein space theory, and [37] developed a distributed Krein space-based attack detection algorithm over sensor networks under deception attacks.

All the above results rely on the absolute measurement outputs, making them unsuitable for application to MASs where only relative output information is available. As mentioned in [38]–[40], in typical situations, a single node of the MAS may lack absolute state observation and only be capable of measuring relative values against its neighboring nodes. Furthermore, according to [41], the transmission of absolute information requires communication channels, which may suffer from potential network attacks. In [42]–[44], it was assumed that each agent was equipped with sensors for relative output measurements. To summarize, it is crucial to investigate the FD problem for time-varying MASs with only relative outputs.

In this paper, we deal with the distributed fault detection problem for a class of LDTV heterogeneous MASs with relative output information. We construct an augmented model for each agent by stacking all locally obtained relative output information, and apply the Krein space-based H_∞ filtering theory to yield a distributed residual-generator and a distributed residual-evaluation mechanism.

The contributions of this paper are twofold.

First, pure relative information is used for the fault detection of time-varying MASs. For MAS fault detection with relative information, some results have been obtained. Relative output information was considered in [42], [43] and [44], however, the absolute measurements were also used for FD in these works; on the other hand, [39] and [40] investigated FD problems for multi-agent networks with only relative state measurements; however, the systems considered in these studies were all time-invariant, making the obtained results unsuitable for our problem. To deal with the lack of absolute output information, we construct an augmented model for each agent by stacking all locally obtained relative output information, based on which each agent carries on the fault detection for its neighbors and itself.

Second, linear discrete time-varying heterogeneous MASs are considered in this paper. For time-varying systems, some FD results have also been obtained [33]–[37]. However, in [33] and [34], the employment of the least-squares method required the disturbance and noise to satisfy known probability distributions, making the obtained results unsuitable for our problem. Additionally, [36] and [37] investigated the FD problems for LDTV sensor networks, where there was only one system dynamic equation and only absolute output information was considered. Moreover, [35] studied the distributed fault estimation problem for MASs, but the considered MASs were homogeneous and absolute measurements were used in the result. To deal with the time-varying characteristic of the coefficient matrices in the considered problem, we apply the Krein space projection theory to yield a distributed residual-generation recursive algorithm, which has a low computational burden and is suitable for application.

The rest of this paper is organized as follows. Some preliminary information about the Krein space is presented in section II, and the problem is formulated in section III. Moreover, the design procedure of the distributed residual-generator is presented in section IV, and a distributed residual-evaluation mechanism is provided in section V. Finally, two illustrative examples are given in section VI to verify the effectiveness of the proposed FD method.

Notation: $\mathbb{R}^n$ denotes the n-dimensional Euclidean space. The script letters $\mathcal{V}, \mathcal{E}, \cdots$ denote sets. $|\mathcal{V}|$ denotes the cardinality of $\mathcal{V}$. $\|\cdot\|$ denotes the Euclidean norm of a vector. $col\{\cdots\}$ represents a column vector. $diag\{\cdots\}$ denotes a block-diagonal matrix. I is an identity matrix with an appropriate size. $w(k) \in l_2$ means that $w(k)$ is l_2-norm bounded for $\forall k \in N$. The elements in the Euclidean space are denoted by normal letters such as x, y, w, f, v, while the elements in the Krein space are denoted by bold letters, such as $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{f}, \boldsymbol{v}$.

II. PRELIMINARY

Definition 1: An abstract vector space $\mathcal{K}$, $\langle \cdot, \cdot \rangle$ that satisfies the following requirements is called a Krein space:

1) *$\mathcal{K}$ is a linear space over the complex field $\mathcal{C}$.*
2) *There exists a bilinear form $\langle \cdot, \cdot \rangle \in \mathcal{C}$ on $\mathcal{K}$ such that*
   a) $\langle \boldsymbol{y}, \boldsymbol{x} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle^*$.
   b) $\langle a\boldsymbol{x} + b\boldsymbol{y}, \boldsymbol{z} \rangle = a\langle \boldsymbol{x}, \boldsymbol{z} \rangle + b\langle \boldsymbol{y}, \boldsymbol{z} \rangle$

   for any $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathcal{K}$, $a, b \in \mathcal{C}$, and where $*$ denotes complex conjugation.
3) *$\mathcal{K}$ admits a direct orthogonal sum decomposition*
$$\mathcal{K} = \mathcal{K}_+ \oplus \mathcal{K}_-,$$
   such that $\mathcal{K}_+$, $\langle \cdot, \cdot \rangle$ and $\mathcal{K}_-$, $\langle \cdot, \cdot \rangle$ are Hilbert spaces, and
$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$$
   for any $\boldsymbol{x} \in \mathcal{K}_+$ and $\boldsymbol{x} \in \mathcal{K}_-$.

Remark 1: Infinite-dimensional Hibert spaces and finite-dimensional Hibert spaces (often called Euclidean spaces) are well known. Although infinite-dimensional Krein spaces and finite-dimensional Krein spaces (often called Minkowski spaces) share many properties with the corresponding Hilbert spaces such as requirements 1) and 2), they also differ from Hilbert spaces in some aspects.

In Hibert spaces,

$$\langle x, x\rangle > 0 \quad when \quad x \neq 0.$$

However, in a Krein space $\mathcal{K}$, the fundamental decomposition defines two projection operators $\mathcal{P}_+$ and $\mathcal{P}_-$ such that

$$\mathcal{P}_+\mathcal{K} = \mathcal{K}_+ \quad and \quad \mathcal{P}_-\mathcal{K} = \mathcal{K}_-,$$

and every $\boldsymbol{x} \in \mathcal{K}$ can be written as

$$\boldsymbol{x} = \mathcal{P}_+\boldsymbol{x} + \mathcal{P}_-\boldsymbol{x} = \boldsymbol{x}_+ + \boldsymbol{x}_-, \boldsymbol{x}_\pm \in \mathcal{K}_\pm.$$

For $\forall \boldsymbol{x} \in \mathcal{K}_+$, we have $\langle \boldsymbol{x}, \boldsymbol{x}\rangle \geq 0$; however, for an arbitrary vector $\boldsymbol{x} \in \mathcal{K}$, it may not be satisfied that $\langle \boldsymbol{x}, \boldsymbol{x}\rangle \geq 0$. Moreover, a vector $\boldsymbol{x} \in \mathcal{K}$ is said to be positive if $\langle \boldsymbol{x}, \boldsymbol{x}\rangle > 0$, neutral if $\langle \boldsymbol{x}, \boldsymbol{x}\rangle = 0$, or negative if $\langle \boldsymbol{x}, \boldsymbol{x}\rangle < 0$.

Definition 2: Assume that the elements $\boldsymbol{y}_0, \ldots, \boldsymbol{y}_T$ are in $\mathcal{K}$. The Gramian of the collection of the elements $\{\boldsymbol{y}_0, \ldots, \boldsymbol{y}_T\}$ is defined as a $(T+1) \times (T+1)$ matrix

$$R_y \triangleq [\langle \boldsymbol{y}_i, \boldsymbol{y}_j\rangle]_{i,j=0:T}.$$

The reflexivity property $\langle \boldsymbol{y}_j, \boldsymbol{y}_i\rangle = \langle \boldsymbol{y}_i, \boldsymbol{y}_j\rangle^*$, indicates that the Gramian is a Hermitian matrix.

In 1990s, Babak Hassibi developed a self-contained Krein space-Kalman filtering theory for linear estimation [45]. Based on simple concepts such as projections and matrix factorizations, a relation between the Krein space projection and the recursive computation of the stationary points of certain quadratic forms was discussed in [45], providing an approach to solve the H_∞ filtering problems for LDTV systems.

III. PROBLEM FORMULATION

In practice, many multi-agent nonlinear systems, such as unmanned vehicles, power networks and chemical processes, can be approximated by linear time-varying models using trajectory linearization techniques, and continuous processes are usually discretized for online implementation. Hence, the study of linear discrete time-varying (LDTV) system is practically meaningful. In addition, many practical MAS nodes cannot acquire absolute state observation and can only obtain relative values against its neighboring nodes, such as unmanned vehicle systems where each vehicle can only obtain the disturbances between itself and its neighbors rather than its absolute position. Motivated by the above discussion, in this paper we investigate the FD problem for LDTV heterogeneous MASs with relative output information.

Let $\mathcal{V} = \{1, 2, \ldots, N\}$ denote the node index set of a multi-agent system consisting of N agents. For an arbitrary agent $i \in \mathcal{V}$, the corresponding system model is described as

$$\begin{aligned} x_i(k+1) &= A_i(k)x_i(k) + B_{w_i}(k)w_i(k) + B_{f_i}(k)f_i(k), \\ y_i(k) &= C_i(k)x_i(k) + v_i(k) + D_{f_i}(k)f_i(k) \end{aligned} \quad (1)$$

where $x_i(k) \in \mathbb{R}^{n_{x_i}}$ is the system state, $w_i(k) \in \mathbb{R}^{n_{w_i}}$ belonging to l_2 is the unknown external disturbance, $f_i(k) \in \mathbb{R}^{n_{f_i}}$ is the fault of agent i, $y_i(k) \in \mathbb{R}^{n_{y_i}}$ denotes the absolute measurement output of agent i (n_{y_i} is the same for all agents), and $v_i(k) \in \mathbb{R}^{n_{v_i}}$ belonging to l_2 denotes the measurement noise. $A_i(k), B_{w_i}(k), B_{f_i}(k), C_i(k), D_{f_i}(k)$ are known time-varying matrices with appropriate dimensions. It should be noted that the fault vector is composed as $f_i(k) = [f_{ip}^T(k), f_{is}^T(k)]^T$, with $f_{ip}(k)$ denoting the fault occurring on the agent dynamics and $f_{is}(k)$ denoting the fault occurring on the absolute measurement sensor. Accordingly, the coefficient matrices $B_{f_i}(k)$ and $D_{f_i}(k)$ are composed as $B_{f_i}(k) = [B_{f_{ip}}(k), 0]$ and $D_{f_i}(k) = [D_{f_{ip}}(k), D_{f_{is}}(k)]$, respectively.

Let $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ be the edge set. The edge $(i,j) \in \mathcal{E}$ represents that node i is a neighbor of node j, which means that agent j can obtain the relative output $y_{ji}(k) = y_j(k) - y_i(k)$. Let $\mathcal{A} = [a_{ij}] \in \mathbb{R}^{N\times N}$ denote the adjacency matrix, where $a_{ij} = 1$ means $(i,j) \in \mathcal{E}$ and $a_{ij} = 0$ means $(i,j) \notin \mathcal{E}$, $\forall i,j \in \mathcal{V}$. Let $\mathcal{N}_i = \{j|a_{ji} = 1\}$ denote the neighbor set of agent i.

Note that agent $i \in \mathcal{V}$ only acquires the relative outputs against its neighbors. Therefore, let $y_{\mathcal{N}_i}(k) = col\{y_{ii_1}(k), y_{ii_2}(k), \cdots, y_{ii_{|\mathcal{N}_i|}}(k)\}$ denote all output information that agent i obtains locally, where $i_1, i_2, \cdots, i_{|\mathcal{N}_i|} \in \mathcal{N}_i$ represent all neighbors of agent i. Similarly, we define $v_{ij}(k) = v_i(k) - v_j(k), \forall (i,j) \in \mathcal{E}$.

Thus, we build the following augmented model for agent i that contains $y_{\mathcal{N}_i}(k)$:

$$\begin{aligned} x_{\mathcal{N}_i}(k+1) &= A_{\mathcal{N}_i}(k)x_{\mathcal{N}_i}(k) + B_{w\mathcal{N}_i}(k)w_{\mathcal{N}_i}(k) \\ &\quad + B_{f\mathcal{N}_i}(k)f_{\mathcal{N}_i}(k), \\ y_{\mathcal{N}_i}(k) &= C_{\mathcal{N}_i}(k)x_{\mathcal{N}_i}(k) + v_{\mathcal{N}_i}(k) + D_{f\mathcal{N}_i}(k)f_{\mathcal{N}_i}(k), \end{aligned} \quad (2)$$

where the augmented vectors $x_{\mathcal{N}_i}(k), w_{\mathcal{N}_i}(k), f_{\mathcal{N}_i}(k)$ are defined by $b_{\mathcal{N}_i}(k) = col\{b_i(k), b_{i_1}(k), \cdots, b_{i_{|\mathcal{N}_i|}}(k)\}$ with $b \in \{x, w, f\}$, $v_{\mathcal{N}_i}(k) = col\{v_{ii_1}, \cdots, v_{ii_{|\mathcal{N}_i|}}\}$, and the coefficient block diagonal matrices $A_{\mathcal{N}_i}(k), B_{w\mathcal{N}_i}(k), B_{f\mathcal{N}_i}(k)$ are defined in the form

$$M_{\mathcal{N}_i}(k) = \begin{bmatrix} M_i(k) & 0 & \cdots & 0 \\ 0 & M_{i_1}(k) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & M_{i_{|\mathcal{N}_i|}}(k) \end{bmatrix},$$

and the coefficient block diagonal matrices $C_{\mathcal{N}_i}(k), D_{f\mathcal{N}_i}(k)$ are defined in the form

$$E_{\mathcal{N}_i}(k) = \begin{bmatrix} E_i(k) & -E_{i_1}(k) & 0 & \cdots & 0 \\ E_i(k) & 0 & -E_{i_2}(k) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ E_i(k) & 0 & 0 & \cdots & -E_{i_{|\mathcal{N}_i|}}(k) \end{bmatrix},$$

Based on the augmented model (2), we will solve the FD problem in two steps:

(a) Design a distributed residual-generator for each agent $i \in \mathcal{V}$ using all locally obtained information $y_{\mathcal{N}_i}(k)$.

(b) Develop an effective residual-evaluation mechanism so that the alarm can be raised after the fault occurrence. The design of this mechanism includes building an evaluation function of residuals and determining a threshold function, so that the fault can be detected by comparing the values of the two functions using a simple "yes" or "no" logic.

IV. DISTRIBUTED RESIDUAL-GENERATOR DESIGN

To deal with the time-varying characteristic of the coefficient matrices, we adopt the H_∞ filtering scheme to solve this FD problem. According to the augmented model (2), the residual of agent $i \in \mathcal{V}$ is expected to satisfy the following performance index under the H_∞ filtering scheme:

$$\sup_{\phi_i \neq 0} \frac{\sum_{k=0}^{W} \|r_i(k) - f_{\mathcal{N}_i}(k)\|^2}{\varphi_i + \sum_{k=0}^{W}(\|w_{\mathcal{N}_i}(k)\|^2 + \|f_{\mathcal{N}_i}(k)\|^2 + \|v_{\mathcal{N}_i}(k)\|^2)} < \gamma_i^2, \tag{3}$$

where $r_i(k)$ is the residual of agent i to be designed; $\gamma_i > 0$ is a given performance level scalar measuring the robustness and sensitivity quality of the residual, $\phi_i = (x_{\mathcal{N}_i}(0), w_{\mathcal{N}_i}^W, f_{\mathcal{N}_i}^W, v_{\mathcal{N}_i}^W)$, where $w_{\mathcal{N}_i}^W = col\{w_{\mathcal{N}_i}(0), \ldots, w_{\mathcal{N}_i}(W)\}$, $f_{\mathcal{N}_i}^W = col\{f_{\mathcal{N}_i}(0), \ldots, f_{\mathcal{N}_i}(W)\}$, $v_{\mathcal{N}_i}^W = col\{v_{\mathcal{N}_i}(0), \ldots, v_{\mathcal{N}_i}(W)\}$, and $\varphi_i = x_{\mathcal{N}_i}^T(0)P_{0i}^{-1}x_{\mathcal{N}_i}(0)$, with $P_{0i} > 0$ being a weighting matrix.

Next, we design a distributed residual-generator for $\forall i \in \mathcal{V}$ such that the inequality (3) is satisfied.

In the remaining part of this section, the residual design problem is converted into a minimization problem of an indefinite quadratic form, and then an auxiliary Krein space model is introduced, based on which we propose a recursive residual-generation algorithm.

A. A MINIMIZATION PROBLEM OF AN INDEFINITE QUADRATIC FORM

In this subsection, we convert the H_∞ inequality (3) into a minimization problem of an indefinite quadratic form, which is instrumental for developing the residual-generator as described below.

We rewrite the augmented system model (2) as

$$\begin{aligned} x_{\mathcal{N}_i}(k+1) &= A_{\mathcal{N}_i}(k)x_{\mathcal{N}_i}(k) + B_{\mathcal{N}_i}(k)d_i(k), \\ y_{\mathcal{N}_i}(k) &= C_{\mathcal{N}_i}(k)x_{\mathcal{N}_i}(k) + v_{\mathcal{N}_i}(k) + D_{\mathcal{N}_i}(k)d_i(k), \end{aligned} \tag{4}$$

where the disturbance and fault are combined as $d_i(k) = col\{w_{\mathcal{N}_i}(k), f_{\mathcal{N}_i}(k)\}$, and the coefficient matrices are correspondingly combined as $B_{\mathcal{N}_i}(k) = [B_{w_{\mathcal{N}_i}}(k), B_{f_{\mathcal{N}_i}}(k)]$, $D_{\mathcal{N}_i}(k) = [0, D_{f_{\mathcal{N}_i}}(k)]$.

We denote the error between the residual and the augmented fault vector as

$$e_i(k) = r_i(k) - f_{\mathcal{N}_i}(k) = r_i(k) - H_i d_i(k), \tag{5}$$

where $H_i = [0, I]$. According to the inequality (3), we define the following indefinite quadratic form

$$\begin{aligned} J_i^W &= x_{\mathcal{N}_i}^T(0)P_{0i}^{-1}x_{\mathcal{N}_i}(0) + \sum_{k=0}^{W}\Big(\|d_i(k)\|^2 + \|v_{\mathcal{N}_i}(k)\|^2 \\ &\quad - \gamma_i^{-2}e_i^T(k)e_i(k)\Big). \end{aligned} \tag{6}$$

It is obvious that the H_∞ inequality (3) is satisfied if and only if $J_i^W > 0$ for $\phi_i \neq 0$.

To motivate the subsequent discussion, we define a new output vector

$$\bar{y}_i(k) = \bar{C}_i(k)x_{\mathcal{N}_i}(k) + \bar{v}_i(k) + \bar{D}_i(k)d_i(k), \tag{7}$$

where $\bar{y}_i(k) = col\{y_{\mathcal{N}_i}(k), r_i(k)\}$, $\bar{C}_i(k) = [C_{\mathcal{N}_i}^T(k), 0]^T$, $\bar{v}_i(k) = col\{v_{\mathcal{N}_i}(k), e_i(k)\}$, $\bar{D}_i(k) = [D_{\mathcal{N}_i}^T(k), H_i^T]^T$.

We denote

$$\begin{aligned} d_i^W &= col\{d_i(0), d_i(1), \ldots, d_i(W)\}, \\ \bar{v}_i^W &= col\{\bar{v}_i(0), \bar{v}_i(1), \ldots, \bar{v}_i(W)\}, \\ \bar{y}_i^W &= col\{\bar{y}_i(0), \bar{y}_i(1), \ldots, \bar{y}_i(W)\}, \\ Q_i^W &= diag\{\underbrace{Q_i, Q_i, \ldots, Q_i}_{(W+1)\text{ block entries}}\}, \end{aligned}$$

where $Q_i = diag\{I, -\gamma_i^2 I\}$. Thus, we write the indefinite quadratic form in the matrix form as

$$J_i^W = \begin{bmatrix} x_{\mathcal{N}_i}(0) \\ d_i^W \\ \bar{v}_i^W \end{bmatrix}^T \begin{bmatrix} P_{0i} & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & Q_i^W \end{bmatrix}^{-1} \begin{bmatrix} x_{\mathcal{N}_i}(0) \\ d_i^W \\ \bar{v}_i^W \end{bmatrix}. \tag{8}$$

Because $\bar{v}_i^W$ can be written as a linear combination of $x_{\mathcal{N}_i}(0)$, d_i^W and $\bar{y}_i^W$, we can convert J_i^W into an indefinite quadratic form of $x_{\mathcal{N}_i}(0)$, d_i^W and $\bar{y}_i^W$. Thus, the residual-generator design problem can be interpreted as follows: (i) ensuring the indefinite quadratic form J_i^W to have a minimum w.r.t. $\{x_{\mathcal{N}_i}(0), d_i^W\}$ for a given $\bar{y}_i^W$; (ii) determining the residual $r_i(k)$ to ensure this minimum to be positive.

For the minimization problem, the global expressions of both the stationary point of J_i^W over $\{x_{\mathcal{N}_i}(0), d_i^W\}$ and the value of J_i^W at this stationary point can be obtained directly according to [46], as well as the global expression of the condition for this value to be the minimum. However, the direct computation using these global expressions will incur a huge computational burden. Therefore, in the next subsection we introduce the Krein space projection theory to recursively solve the minimization problem.

Remark 2: As shown in [47], the H_∞ filtering problem can be cast into a minimization problem of a certain quadratic form. By considering the appropriate state space model and Gramians, we can use the Krein space projection theory to calculate the value at the stationary point and study its properties.

B. INTRODUCTION OF THE KREIN SPACE

In this subsection, the following auxiliary Krein space state-space model is introduced to recursively solve the minimization problem of the indefinite quadratic form J_i^W:

$$\begin{aligned} \boldsymbol{x}_{\mathcal{N}_i}(k+1) &= A_{\mathcal{N}_i}(k)\boldsymbol{x}_{\mathcal{N}_i}(k) + B_{\mathcal{N}_i}(k)\boldsymbol{d}_i(k), \\ \bar{\boldsymbol{y}}_i(k) &= \bar{C}_i(k)\boldsymbol{x}_{\mathcal{N}_i}(k) + \bar{D}_i(k)\boldsymbol{d}_i(k) + \bar{\boldsymbol{v}}_i(k), \end{aligned} \tag{9}$$

where $\boldsymbol{x}_{\mathcal{N}_i}(k)$ is the state vector, $\boldsymbol{d}_i(k)$ is the external input vector, $\bar{\boldsymbol{v}}_i(k) = col\{\boldsymbol{v}_{\mathcal{N}_i}(k), \boldsymbol{e}_i(k)\}$ is the noise vector, $\bar{\boldsymbol{y}}_i(k) = col\{\boldsymbol{y}_{\mathcal{N}_i}(k), \boldsymbol{r}_i(k)\}$ is the output vector. Let $\boldsymbol{x}_{\mathcal{N}_i}(0)$, $\boldsymbol{d}_i(k)$ and $\bar{\boldsymbol{v}}_i(k)$ take the following Gramian

$$\left\langle \begin{bmatrix} \boldsymbol{x}_{\mathcal{N}_i}(0) \\ \boldsymbol{d}_i(k) \\ \bar{\boldsymbol{v}}_i(k) \end{bmatrix}, \begin{bmatrix} \boldsymbol{x}_{\mathcal{N}_i}(0) \\ \boldsymbol{d}_i(l) \\ \bar{\boldsymbol{v}}_i(l) \end{bmatrix} \right\rangle = \begin{bmatrix} P_{0i} & 0 & 0 \\ 0 & I\delta_{kl} & 0 \\ 0 & 0 & Q_i\delta_{kl} \end{bmatrix}, \tag{10}$$

where $\delta_{kl} = 0$ for $k \neq l$ and $\delta_{kl} = 1$ for $k = l$.

The linear space generated by the output vectors $\{\bar{\boldsymbol{y}}_i(q)\}_{q=0}^{k}$ is denoted by $\mathscr{L}\{\bar{\boldsymbol{y}}_i(q)\}_{q=0}^{k}$, and the projection of a vector $\boldsymbol{g}(k)$ onto $\mathscr{L}\{\bar{\boldsymbol{y}}_i(q)\}_{q=0}^{k-1}$ in the Krein space is denoted by $\hat{\boldsymbol{g}}(k|k-1)$. Defining $\tilde{\boldsymbol{g}}(k) = \boldsymbol{g}(k) - \hat{\boldsymbol{g}}(k|k-1)$, we obtain the innovation denoted as

$$\tilde{\bar{\boldsymbol{y}}}_i(k) = \bar{\boldsymbol{y}}_i(k) - \hat{\bar{\boldsymbol{y}}}(k|k-1). \tag{11}$$

Further, the value of J_i^W at its stationary point is given by

$$J_i^{W*} = \sum_{k=0}^{W} \tilde{\bar{\boldsymbol{y}}}_i(k) R_{\tilde{\bar{y}}_i}^{-1}(k) \tilde{\bar{\boldsymbol{y}}}_i(k), \tag{12}$$

where the Gramian $R_{\tilde{\bar{y}}_i}(k) = \langle \tilde{\bar{\boldsymbol{y}}}_i(k), \tilde{\bar{\boldsymbol{y}}}_i(k) \rangle$, see [45] for details.

C. A NECESSARY AND SUFFICIENT CONDITION FOR THE EXISTENCE OF THE QUADRATIC FORM's MINIMUM

According to equation (12), the value of the quadratic form at its stationary point has been expressed in terms of the innovations. Next, a necessary and sufficient condition for the minimum of this quadratic form is given below.

Lemma 1: [45] The minimum of J_i^W over $\{x_{\mathcal{N}_i}(0), d_i^W\}$ is equal to J_i^{W*} if and only if $R_{\tilde{\bar{y}}_i}(k)$ has the same inertia as Q_i's for all $k = 0, 1, \ldots, W$.

Note that it is impossible to directly check the inertia condition in Lemma 1. Therefore, a feasible numerical approach for checking this minimum condition is given in Lemma 2 as below.

Lemma 2: The minimum of J_i^W over $\{x_{\mathcal{N}_i}(0), d_i^W\}$ is equal to J_i^{W*} if and only if

$$\Psi_i(k) > 0 \quad and \quad \Phi_i(k) < 0 \tag{13}$$

for all $k = 0, 1, \ldots, W$, where

$$\Psi_i(k) = C_{\mathcal{N}_i}(k)P_i(k)C_{\mathcal{N}_i}^T(k) + D_{\mathcal{N}_i}(k)D_{\mathcal{N}_i}^T(k) + I, \tag{14}$$

$$\Phi_i(k) = -\gamma_i^2 I + H_i H_i^T - H_i D_{\mathcal{N}_i}^T(k)\Psi_i^{-1}(k)D_{\mathcal{N}_i}(k)H_i^T, \tag{15}$$

$$\begin{aligned} P_i(k) &= \langle \tilde{\boldsymbol{x}}_{\mathcal{N}_i}(k), \tilde{\boldsymbol{x}}_{\mathcal{N}_i}(k) \rangle \\ &= \langle \boldsymbol{x}_{\mathcal{N}_i}(k) - \hat{\boldsymbol{x}}_{\mathcal{N}_i}(k), \boldsymbol{x}_{\mathcal{N}_i}(k) - \hat{\boldsymbol{x}}_{\mathcal{N}_i}(k) \rangle. \end{aligned} \tag{16}$$

Proof: According to equations (9), (10) and (11), we have

$$\tilde{\bar{\boldsymbol{y}}}_i(k) = \bar{C}_i(k)\tilde{\boldsymbol{x}}_{\mathcal{N}_i}(k) + \bar{\boldsymbol{v}}_i(k) + \bar{D}_i(k)\boldsymbol{d}_i(k). \tag{17}$$

Thus, the Gramian in Lemma 1 can be calculated as

$$\begin{aligned} R_{\tilde{\bar{y}}_i}(k) &= \bar{C}_i(k)P_i(k)\bar{C}_i^T(k) + Q_i + \bar{D}_i(k)\bar{D}_i^T(k) \\ &= \begin{bmatrix} \Psi_i(k) & D_{\mathcal{N}_i}(k)H_i^T \\ H_i D_{\mathcal{N}_i}^T(k) & -\gamma_i^2 I + H_i H_i^T \end{bmatrix}, \end{aligned} \tag{18}$$

where

$$\begin{aligned} P_i(k) &= \langle \tilde{\boldsymbol{x}}_{\mathcal{N}_i}(k), \tilde{\boldsymbol{x}}_{\mathcal{N}_i}(k) \rangle \\ &= \langle \boldsymbol{x}_{\mathcal{N}_i}(k) - \hat{\boldsymbol{x}}_{\mathcal{N}_i}(k), \boldsymbol{x}_{\mathcal{N}_i}(k) - \hat{\boldsymbol{x}}_{\mathcal{N}_i}(k) \rangle, \\ \Psi_i(k) &= C_{\mathcal{N}_i}(k)P_i(k)C_{\mathcal{N}_i}^T(k) + D_{\mathcal{N}_i}(k)D_{\mathcal{N}_i}^T(k) + I. \end{aligned}$$

To further investigate the inertia of the Gramian, we apply the triangular factorization to equation (18), yielding

$$R_{\tilde{\bar{y}}_i}(k) = \begin{bmatrix} I & 0 \\ \Omega_i(k) & I \end{bmatrix} \begin{bmatrix} \Psi_i(k) & 0 \\ 0 & \Phi_i(k) \end{bmatrix} \begin{bmatrix} I & 0 \\ \Omega_i(k) & I \end{bmatrix}^T, \tag{19}$$

where

$$\begin{aligned} \Phi_i(k) &= -\gamma_i^2 I + H_i H_i^T - H_i D_{\mathcal{N}_i}^T(k)\Psi_i^{-1}(k)D_{\mathcal{N}_i}(k)H_i^T, \\ \Omega_i(k) &= H_i D_{\mathcal{N}_i}^T(k)\Psi_i^{-1}(k). \end{aligned}$$

The factorization result in equation (19) indicates the congruent relationship

$$R_{\tilde{\bar{y}}_i}(k) \simeq \begin{bmatrix} \Psi_i(k) & 0 \\ 0 & \Phi_i(k) \end{bmatrix}, \tag{20}$$

which means that the two matrices have the same inertia. Further, recalling $Q_i = diag\{I, -\gamma_i^2 I\}$ and applying the result of Lemma 1, Lemma 2 is then proved. □

D. DISTRIBUTED RESIDUAL-GENERATOR

Lemma 2 presents a necessary and sufficient condition for the value of the indefinite quadratic form J_i^W at the stationary point to be the minimum w.r.t. $\{x_{\mathcal{N}_i}(0), d_i^W\}$ for a given $\bar{y}_i^W$. We note that the residual $r_i(k)$ is included in the defined output vector $\bar{y}_i(k)$ according to equation (7). Therefore, in this subsection, we recursively determine the residuals in $\bar{y}_i^W$ to guarantee positive J_i^{W*}, which satisfies the H_∞ performance index (3) under the minimum condition.

Theorem 1: Suppose the minimum condition (13) *holds. Then, the residual $r_i(k)$ can be given as*

$$r_i(k) = \Omega_i(k)\left(y_{\mathcal{N}_i}(k) - \hat{y}_{\mathcal{N}_i}(k|k-1)\right) \tag{21}$$

to guarantee that $\min(J_i^W) = J_i^{W*} > 0$ for $\phi_i \neq 0$, where $\hat{y}_{\mathcal{N}_i}(k|k-1) = C_{\mathcal{N}_i}(k)\hat{x}_{\mathcal{N}_i}(k|k-1)$ and $\hat{x}_{\mathcal{N}_i}(k|k-1)$ can be recursively calculated as

$$\begin{aligned} \hat{x}_{\mathcal{N}_i}(k+1|k) &= A_{\mathcal{N}_i}(k)\hat{x}_{\mathcal{N}_i}(k|k-1) \\ &\quad + \Theta_i(k)\Psi_i^{-1}(k)\left(y_{\mathcal{N}_i}(k) - \hat{y}_{\mathcal{N}_i}(k|k-1)\right), \end{aligned} \tag{22}$$

with $\Theta_i(k) = A_{\mathcal{N}_i}(k)P_i(k)C_{\mathcal{N}_i}^T(k) + B_{\mathcal{N}_i}(k)D_{\mathcal{N}_i}^T(k)$, $\hat{x}_{\mathcal{N}_i}(0|-1) = 0$.

Moreover, the involved Gramian $P_i(k)$ can be recursively calculated as

$$\begin{aligned}P_i(k+1) = & A_{\mathcal{N}_i}(k)P_i(k)A^T_{\mathcal{N}_i}(k) + B_{\mathcal{N}_i}(k)B^T_{\mathcal{N}_i}(k)\\ & - \Theta_i(k)\Psi_i^{-1}(k)\Theta_i^T(k)\\ & - \Theta_i(k)\Omega_i^T(k)\Phi_i^{-1}(k)\Omega_i(k)\Theta_i^T(k)\\ & - B_{\mathcal{N}_i}(k)H_i^T(k)\Phi_i^{-1}(k)H_i(k)B^T_{\mathcal{N}_i}(k)\\ & + \Theta_i(k)\Omega_i^T(k)\Phi_i^{-1}(k)H_i(k)B^T_{\mathcal{N}_i}(k)\\ & + B_{\mathcal{N}_i}(k)H_i^T(k)\Phi_i^{-1}(k)\Omega_i(k)\Theta_i^T(k), \end{aligned} \quad (23)$$

where $P_i(0) = P_{0i}$.

Proof: The value of the indefinite quadratic form at the stationary point can be written as

$$\begin{aligned}J_i^{W*} &= \sum_{k=0}^{W} \tilde{\bar{\mathbf{y}}}_i^T(k)R^{-1}_{\tilde{\bar{\mathbf{y}}}_i}(k)\tilde{\bar{\mathbf{y}}}_i(k)\\ &= \sum_{k=0}^{W} \begin{bmatrix}\tilde{\mathbf{y}}_{\mathcal{N}_i}(k)\\ \tilde{\mathbf{r}}_i(k)\end{bmatrix}^T R^{-1}_{\tilde{\bar{\mathbf{y}}}_i}(k)\begin{bmatrix}\tilde{\mathbf{y}}_{\mathcal{N}_i}(k)\\ \tilde{\mathbf{r}}_i(k)\end{bmatrix}. \end{aligned} \quad (24)$$

Noting $\mathbf{r}_i(k) = \mathbf{e}_i(k) + H_i\mathbf{d}_i(k)$ and according to the Gramian in equation (10), we have

$$\hat{\mathbf{r}}_i(k|k-1) = 0. \quad (25)$$

Furthermore, according to equation (19), we have

$$\begin{aligned}&R^{-1}_{\tilde{\bar{\mathbf{y}}}_i}(k)\\ &= \begin{bmatrix} I & 0\\ -\Omega_i(k) & I\end{bmatrix}^T \begin{bmatrix}\Psi_i^{-1}(k) & 0\\ 0 & \Phi_i^{-1}(k)\end{bmatrix}\begin{bmatrix} I & 0\\ -\Omega_i(k) & I\end{bmatrix}\\ &= \begin{bmatrix}\Psi_i^{-1}(k) + \Omega_i^T(k)\Phi_i^{-1}(k)\Omega_i(k) & -\Omega_i^T(k)\Phi_i^{-1}(k)\\ -\Phi_i^{-1}(k)\Omega_i(k) & \Phi_i^{-1}(k)\end{bmatrix}. \end{aligned} \quad (26)$$

According to equations (25) and (26), equation (24) can be rewritten as

$$\begin{aligned}J_i^{W*} &= \sum_{k=0}^{W}\begin{bmatrix}\tilde{\mathbf{y}}_{\mathcal{N}_i}(k)\\ \mathbf{r}_i(k)\end{bmatrix}^T\\ &\quad\times\begin{bmatrix}\Psi_i^{-1}(k) + \Omega_i^T(k)\Phi_i^{-1}(k)\Omega_i(k) & -\Omega_i^T(k)\Phi_i^{-1}(k)\\ -\Phi_i^{-1}(k)\Omega_i(k) & \Phi_i^{-1}(k)\end{bmatrix}\\ &\quad\times\begin{bmatrix}\tilde{\mathbf{y}}_{\mathcal{N}_i}(k)\\ \mathbf{r}_i(k)\end{bmatrix}\\ &= \sum_{k=0}^{W}\Big(\tilde{y}^T_{\mathcal{N}_i}(k)\Psi_i^{-1}(k)\tilde{y}_{\mathcal{N}_i}(k) + (r_i(k) - \Omega_i(k)\tilde{y}_{\mathcal{N}_i}(k))^T\\ &\quad\times\Phi_i^{-1}(k)(r_i(k) - \Omega_i(k)\tilde{y}_{\mathcal{N}_i}(k))\Big). \end{aligned} \quad (27)$$

Recalling $\Psi_i(k) > 0$ in Lemma 2, we set

$$r_i(k) = \Omega_i(k)\tilde{y}_{\mathcal{N}_i}(k), \quad (28)$$

so that

$$J_i^{W*} = \sum_{k=0}^{W}\tilde{y}^T_{\mathcal{N}_i}(k)\Psi_i^{-1}(k)\tilde{y}_{\mathcal{N}_i}(k) > 0. \quad (29)$$

Next, we show how to recursively compute the state projection $\hat{\mathbf{x}}_{\mathcal{N}_i}(k|k-1)$ and $\mathbf{x}_{\mathcal{N}_i}(k)$'s Gramian $P_i(k)$.

According to the Krein space projection theory, we have

$$\begin{aligned}&\hat{\mathbf{x}}_{\mathcal{N}_i}(k+1|k)\\ &= \sum_{h=0}^{k}\langle\mathbf{x}_{\mathcal{N}_i}(k+1), \tilde{\bar{\mathbf{y}}}_i(h)\rangle R^{-1}_{\tilde{\bar{\mathbf{y}}}_i}(h)\tilde{\bar{\mathbf{y}}}_i(h)\\ &= \sum_{h=0}^{k-1}\langle\mathbf{x}_{\mathcal{N}_i}(k+1), \tilde{\bar{\mathbf{y}}}_i(h)\rangle R^{-1}_{\tilde{\bar{\mathbf{y}}}_i}(h)\tilde{\bar{\mathbf{y}}}_i(h)\\ &\quad+\langle\mathbf{x}_{\mathcal{N}_i}(k+1), \tilde{\bar{\mathbf{y}}}_i(k)\rangle R^{-1}_{\tilde{\bar{\mathbf{y}}}_i}(k)\tilde{\bar{\mathbf{y}}}_i(k). \end{aligned} \quad (30)$$

The first summation of equation (30) can be calculated as

$$\begin{aligned}&\sum_{h=0}^{k-1}\langle\mathbf{x}_{\mathcal{N}_i}(k+1), \tilde{\bar{\mathbf{y}}}_i(h)\rangle R^{-1}_{\tilde{\bar{\mathbf{y}}}_i}(h)\tilde{\bar{\mathbf{y}}}_i(h)\\ &= A_{\mathcal{N}_i}(k)\sum_{h=0}^{k-1}\langle\mathbf{x}_{\mathcal{N}_i}(k), \tilde{\bar{\mathbf{y}}}_i(h)\rangle R^{-1}_{\tilde{\bar{\mathbf{y}}}_i}(h)\tilde{\bar{\mathbf{y}}}_i(h)\\ &\quad+B_{\mathcal{N}_i}(k)\sum_{h=0}^{k-1}\langle\mathbf{d}_i(k), \tilde{\bar{\mathbf{y}}}_i(h)\rangle R^{-1}_{\tilde{\bar{\mathbf{y}}}_i}(h)\tilde{\bar{\mathbf{y}}}_i(h)\\ &= A_{\mathcal{N}_i}(k)\hat{\mathbf{x}}_{\mathcal{N}_i}(k|k-1). \end{aligned} \quad (31)$$

The cross-Gramian $\langle\mathbf{x}_{\mathcal{N}_i}(k+1), \tilde{\bar{\mathbf{y}}}_i(k)\rangle$ in the second summation of equation (30) can be calculated as

$$\begin{aligned}&\langle\mathbf{x}_{\mathcal{N}_i}(k+1), \tilde{\bar{\mathbf{y}}}_i(k)\rangle\\ &= A_{\mathcal{N}_i}(k)\langle\mathbf{x}_{\mathcal{N}_i}(k), \tilde{\bar{\mathbf{y}}}_i(k)\rangle + B_{\mathcal{N}_i}(k)\langle\mathbf{d}_i(k), \tilde{\bar{\mathbf{y}}}_i(k)\rangle\\ &= A_{\mathcal{N}_i}(k)\langle\mathbf{x}_{\mathcal{N}_i}(k), \bar{C}_i(k)\tilde{\mathbf{x}}_{\mathcal{N}_i}(k) + \bar{\mathbf{v}}_i(k) + \bar{D}_i(k)\mathbf{d}_i(k)\rangle\\ &\quad+B_{\mathcal{N}_i}(k)\langle\mathbf{d}_i(k), \bar{C}_i(k)\tilde{\mathbf{x}}_{\mathcal{N}_i}(k) + \bar{\mathbf{v}}_i(k) + \bar{D}_i(k)\mathbf{d}_i(k)\rangle\\ &= A_{\mathcal{N}_i}(k)\langle\mathbf{x}_{\mathcal{N}_i}(k), \bar{C}_i(k)\tilde{\mathbf{x}}_{\mathcal{N}_i}(k)\rangle\\ &\quad+B_{\mathcal{N}_i}(k)\langle\mathbf{d}_i(k), \bar{D}_i(k)\mathbf{d}_i(k)\rangle\\ &= A_{\mathcal{N}_i}(k)P_i(k)\bar{C}_i^T(k) + B_{\mathcal{N}_i}(k)\bar{D}_i^T(k). \end{aligned} \quad (32)$$

Moreover, recalling $r_i(k) = \Omega_i(k)\tilde{y}_{\mathcal{N}_i}(k)$ in equation (28), we have

$$\begin{aligned}&R^{-1}_{\tilde{\bar{\mathbf{y}}}_i}(k)\tilde{\bar{y}}_i(k)\\ &= \begin{bmatrix}\Psi_i^{-1}(k) + \Omega_i^T(k)\Phi_i^{-1}(k)\Omega_i(k) & -\Omega_i^T(k)\Phi_i^{-1}(k)\\ -\Phi_i^{-1}(k)\Omega_i(k) & \Phi_i^{-1}(k)\end{bmatrix}\\ &\quad\times\begin{bmatrix}I\\ \Omega_i(k)\end{bmatrix}\tilde{y}_{\mathcal{N}_i}(k)\\ &= \begin{bmatrix}\Psi_i^{-1}(k)\\ 0\end{bmatrix}\tilde{y}_{\mathcal{N}_i}(k). \end{aligned} \quad (33)$$

According to the above three equations, the state projection is given below in a recursive form:

$$\hat{x}_{\mathcal{N}_i}(k+1|k) = A_{\mathcal{N}_i}(k)\hat{x}_{\mathcal{N}_i}(k|k-1) + \Theta_i(k)\Psi_i^{-1}(k)\tilde{y}_{\mathcal{N}_i}(k), \quad (34)$$

where $\Theta_i(k) = A_{\mathcal{N}_i}(k)P_i(k)C^T_{\mathcal{N}_i}(k) + B_{\mathcal{N}_i}(k)D^T_{\mathcal{N}_i}(k)$ and the initial projection $\hat{\mathbf{x}}_{\mathcal{N}_i}(0|-1) = 0$.

On the other hand, applying the projection theory to equation (16) yields that

$$P_i(k) = \langle \boldsymbol{x}_{\mathcal{N}_i}(k), \boldsymbol{x}_{\mathcal{N}_i}(k)\rangle - \langle \hat{\boldsymbol{x}}_{\mathcal{N}_i}(k|k-1), \hat{\boldsymbol{x}}_{\mathcal{N}_i}(k|k-1)\rangle. \tag{35}$$

We note that the two Gramians in equation (35) can be calculated recursively as

$$\begin{aligned}
&\langle \boldsymbol{x}_{\mathcal{N}_i}(k+1), \boldsymbol{x}_{\mathcal{N}_i}(k+1)\rangle \\
&\quad = \langle A_{\mathcal{N}_i}(k)\boldsymbol{x}_{\mathcal{N}_i}(k) + B_{\mathcal{N}_i}(k)\boldsymbol{d}_i(k), A_{\mathcal{N}_i}(k)\boldsymbol{x}_{\mathcal{N}_i}(k) \\
&\qquad + B_{\mathcal{N}_i}(k)\boldsymbol{d}_i(k)\rangle \\
&\quad = A_{\mathcal{N}_i}(k)\langle \boldsymbol{x}_{\mathcal{N}_i}(k), \boldsymbol{x}_{\mathcal{N}_i}(k)\rangle A^T_{\mathcal{N}_i}(k) + B_{\mathcal{N}_i}(k)B^T_{\mathcal{N}_i}(k)
\end{aligned} \tag{36}$$

and

$$\begin{aligned}
&\langle \hat{\boldsymbol{x}}_{\mathcal{N}_i}(k+1|k), \hat{\boldsymbol{x}}_{\mathcal{N}_i}(k+1|k)\rangle \\
&\quad = \langle A_{\mathcal{N}_i}(k)\hat{\boldsymbol{x}}_{\mathcal{N}_i}(k|k-1) + \langle \boldsymbol{x}_{\mathcal{N}_i}(k+1), \tilde{\bar{\boldsymbol{y}}}_i(k)\rangle R^{-1}_{\tilde{\bar{\boldsymbol{y}}}_i}(k)\tilde{\bar{\boldsymbol{y}}}_i(k), \\
&A_{\mathcal{N}_i}(k)\hat{\boldsymbol{x}}_{\mathcal{N}_i}(k|k-1) + \langle \boldsymbol{x}_{\mathcal{N}_i}(k+1), \tilde{\bar{\boldsymbol{y}}}_i(k)\rangle R^{-1}_{\tilde{\bar{\boldsymbol{y}}}_i}(k)\tilde{\bar{\boldsymbol{y}}}_i(k)\rangle \\
&\quad = A_{\mathcal{N}_i}(k)\langle \hat{\boldsymbol{x}}_{\mathcal{N}_i}(k|k-1), \hat{\boldsymbol{x}}_{\mathcal{N}_i}(k|k-1)\rangle A^T_{\mathcal{N}_i}(k) \\
&\qquad + \langle \boldsymbol{x}_{\mathcal{N}_i}(k+1), \tilde{\bar{\boldsymbol{y}}}_i(k)\rangle M^{-1}_{\tilde{\bar{\boldsymbol{y}}}_i}(k)\langle \boldsymbol{x}_{\mathcal{N}_i}(k+1), \tilde{\bar{\boldsymbol{y}}}_i(k)\rangle^T.
\end{aligned} \tag{37}$$

Therefore, the Riccati recursion of $P_i(k)$ can be deduced as

$$\begin{aligned}
P_i(k+1) &= A_{\mathcal{N}_i}(k)P_i(k)A^T_{\mathcal{N}_i}(k) + B_{\mathcal{N}_i}(k)B^T_{\mathcal{N}_i}(k) \\
&\quad - \langle \boldsymbol{x}_{\mathcal{N}_i}(k+1), \tilde{\bar{\boldsymbol{y}}}_i(k)\rangle R^{-1}_{\tilde{\bar{\boldsymbol{y}}}_i}(k)\langle \boldsymbol{x}_{\mathcal{N}_i}(k+1), \tilde{\bar{\boldsymbol{y}}}_i(k)\rangle^T \\
&= A_{\mathcal{N}_i}(k)P_i(k)A^T_{\mathcal{N}_i}(k) + B_{\mathcal{N}_i}(k)B^T_{\mathcal{N}_i}(k) \\
&\quad - \Theta_i(k)\Psi_i^{-1}(k)\Theta_i^T(k) \\
&\quad - \Theta_i(k)\Omega_i^T(k)\Phi_i^{-1}(k)\Omega_i(k)\Theta_i^T(k) \\
&\quad - B_{\mathcal{N}_i}(k)H_i^T(k)\Phi_i^{-1}(k)H_i(k)B^T_{\mathcal{N}_i}(k) \\
&\quad + \Theta_i(k)\Omega_i^T(k)\Phi_i^{-1}(k)H_i(k)B^T_{\mathcal{N}_i}(k) \\
&\quad + B_{\mathcal{N}_i}(k)H_i^T(k)\Phi_i^{-1}(k)\Omega_i(k)\Theta_i^T(k),
\end{aligned} \tag{38}$$

where $P_i(0) = \langle \tilde{\boldsymbol{x}}_{\mathcal{N}_i}(0), \tilde{\boldsymbol{x}}_{\mathcal{N}_i}(0)\rangle = P_{0i}$. □

For convenient reference, the calculation of the residual $r_i(k)$ is summarized in the following algorithm.

TABLE 1. Distributed residual-generation algorithm.

Input:
The relative output information
$y_{\mathcal{N}_i}(k) = col\{y_{ii_1}(k), y_{ii_2}(k), \cdots, y_{ii_{|\mathcal{N}_i|}}(k)\}$,
The coefficient matrices
$A_{\mathcal{N}_i}(k), B_{\mathcal{N}_i}(k), C_{\mathcal{N}_i}(k), D_{\mathcal{N}_i}(k)$.
Output:
The residual $r_i(k)$.

step 1	Select an appropriate performance level scalar γ_i and a weighting matrix P_{0i} of the H_∞ performance index (3). Set $k = 0$.
step 2	Calculate $\Psi_i(k)$ and $\Phi_i(k)$ by equations (14) and (15) respectively. If the minimum condition (13) in Lemma 2 holds, go to step 3; else, exit.
step 3	Calculate the residual $r_i(k)$ by equation (21).
step 4	Set $k = k + 1$ and go to step 5.
step 5	Calculate $\hat{x}_{\mathcal{N}_i}(k\|k-1)$ by equation (22) and $P_i(k)$ by the Riccati recursion (23), respectively. Turn to step 2.

V. DISTRIBUTED RESIDUAL-EVALUATION MECHANISM

In this section, we construct a residual-evaluation mechanism by building a residual-evaluation function and determining a threshold function.

According to the H_∞ performance index (3), in the fault-free case, we have

$$\begin{aligned}
\sum_{k=0}^{W}\|r_i(k)\|^2 &\le \gamma_i^2\big(x_{\mathcal{N}_i}(0)^T P_{0i}^{-1}x_{\mathcal{N}_i}(0) + \sum_{k=0}^{W}\|w_{\mathcal{N}_i}(k)\|^2 \\
&\quad + \sum_{k=0}^{W}\|v_{\mathcal{N}_i}(k)\|^2\big).
\end{aligned} \tag{39}$$

Under the assumption that $w_{\mathcal{N}_i}(k)$ and $v_{\mathcal{N}_i}(k)$ are l_2-norm bounded with $\|w_{\mathcal{N}_i}(k)\| \le \sigma_{w_{\mathcal{N}_i}}$ and $\|v_{\mathcal{N}_i}(k)\| \le \sigma_{v_{\mathcal{N}_i}}$ for $\forall k \in N$, the threshold function for the fault detection can be set as follows:

$$\begin{aligned}
Th_i^W &= \gamma_i^2\big(x_{\mathcal{N}_i}(0)^T P_{0i}^{-1}x_{\mathcal{N}_i}(0) + (W+1)\sigma^2_{w_{\mathcal{N}_i}} \\
&\quad + (W+1)\sigma^2_{v_{\mathcal{N}_i}}\big).
\end{aligned} \tag{40}$$

Defining the residual-evaluation function as

$$V_i^W = \sum_{k=0}^{W}\|r_i(k)\|^2,$$

fault detection can be accomplished according to the following logic:

$$\begin{cases} V_i^W > Th_i^W \Rightarrow \text{An alarm for } f_{\mathcal{N}_i} \\ V_i^W \le Th_i^W \Rightarrow \text{No alarm} \end{cases} \tag{41}$$

Remark 3: Note that if the H_∞ performance level scalar γ_i is small enough, the residual $r_i(k)$ is exactly an H_∞ estimate of the augmented fault vector $f_{\mathcal{N}_i}(k)$. However, in many cases we must set $\gamma_i > 1$ to satisfy the minimum condition in Lemma 2, leading to a distinct error between the residual $r_i(k)$ and the augmented fault vector $f_{\mathcal{N}_i}(k)$. Thus, in many cases the H_∞ framework in (3) can only be used to carry on the detection of the fault rather than the estimation. Furthermore, for fault detection problems, there are always trade-offs between false alarms and missed detections. Note that the threshold function Th_i^W shows the worst-case tolerant limit decided by the l_2-norm bounds of the unknown disturbance vector $w_{\mathcal{N}_i}(k)$ and the measurement noise vector $v_{\mathcal{N}_i}(k)$. Therefore, the proposed residual-evaluation mechanism guarantees zero false alarm rate, at the cost of possible missed detections. Furthermore, because the residual-evaluation function V_i^W and the threshold function Th_i^W are accumulations, the FD efficiency is influenced by the fault start time, and it is also possible that the fault appears too late in the fixed FD time interval to be detected. Moreover, the l_2-norm bounds of disturbance vectors and noise vectors also influence the FD efficiency, and lower bounds will shorten the time delay from the fault occurrence to the alarm-raising.

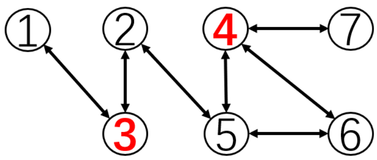


FIGURE 1. Relative output topology of the MAS in EXAMPLE 1, where faulty agents 3 and 4 are marked in red.

VI. TWO ILLUSTRATIVE EXAMPLES

In this section, we use two simulation examples to illustrate the efficiency of the proposed FD method. Matlab R2016a is used in the simulation.

A. EXAMPLE 1

We consider a multi-agent system consisting of 7 agents, where the coefficient matrices in (1) are set as

$$A_i(k) = \begin{bmatrix} 0.52 + 0.1sin(k) & 0.12 \\ 0.05 & 0.37 + 0.1cos(k) \end{bmatrix},$$
$$B_{w_i}(k) = \begin{bmatrix} 0.05cos(k) & 0.12 \\ -0.04 & 0.07 \end{bmatrix},$$
$$C_i(k) = \begin{bmatrix} 0.5sin(k) & 0.6cos(k) \end{bmatrix},$$

$\forall i \in \mathcal{V} = \{1, 2, 3, 4, 5, 6, 7\}$.

In this example, only sensor faults are considered, and the coefficient matrices of $f_i(k)$, $i \in \mathcal{V}$ are set as

$$B_{f_i}(k) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad D_{f_i}(k) = 2.$$

The relative output topology is shown in Fig. 1, with the adjacency matrix defined as

$$\mathcal{A} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

We set the initial states $x_1(0) = [0, 0]^T$, $x_2(0) = [1, 0.5]^T$, $x_3(0) = [-0.3, 0.5]^T$, $x_4(0) = [0, 0.1]^T$, $x_5(0) = [0.1, 0.1]^T$, $x_6(0) = [0.5, 0.2]^T$, $x_7(0) = [0, -0.3]^T$, and choose the weighting matrix $P_{0i} = I$, $\forall i \in \mathcal{V}$. We set the disturbance as $w_i(k) = 0.1[sin(ik), cos(ik)]^T$ and the measurement noise as $v_i(k) = 0.01cos(i+k)$, $\forall i \in \mathcal{V}$. Assuming that the l_2-norm bounds of disturbance vectors and noise vectors are known as $\|w_i(k)\|^2 \leq 0.1^2$ and $\|v_i(k)\|^2 \leq (0.01)^2$, $\forall k \in N$, we have

$$\begin{aligned}
\|w_{\mathcal{N}_1}(k)\|^2 &= \|w_1(k)\|^2 + \|w_3(k)\|^2 \\
&\leq 0.1^2 \times 2, \\
\|w_{\mathcal{N}_2}(k)\|^2 &= \|w_2(k)\|^2 + \|w_3(k)\|^2 + \|w_5(k)\|^2 \\
&\leq 0.1^2 \times 3, \\
\|w_{\mathcal{N}_3}(k)\|^2 &= \|w_3(k)\|^2 + \|w_1(k)\|^2 + \|w_2(k)\|^2 \\
&\leq 0.1^2 \times 3, \\
\|w_{\mathcal{N}_4}(k)\|^2 &= \|w_4(k)\|^2 + \|w_5(k)\|^2 + \|w_6(k)\|^2 + \|w_7(k)\|^2 \\
&\leq 0.1^2 \times 4, \\
\|w_{\mathcal{N}_5}(k)\|^2 &= \|w_5(k)\|^2 + \|w_2(k)\|^2 + \|w_4(k)\|^2 + \|w_6(k)\|^2 \\
&\leq 0.1^2 \times 4, \\
\|w_{\mathcal{N}_6}(k)\|^2 &= \|w_6(k)\|^2 + \|w_4(k)\|^2 + \|w_5(k)\|^2 \\
&\leq 0.1^2 \times 3, \\
\|w_{\mathcal{N}_7}(k)\|^2 &= \|w_7(k)\|^2 + \|w_4(k)\|^2 \\
&\leq 0.1^2 \times 2, \\
\|v_{\mathcal{N}_1}(k)\|^2 &= \|v_1(k) - v_3(k)\|^2 \\
&\leq (0.01 - (-0.01))^2, \\
\|v_{\mathcal{N}_2}(k)\|^2 &= \|v_2(k) - v_3(k)\|^2 + \|v_2(k) - v_5(k)\|^2 \\
&\leq (0.01 - (-0.01))^2 \times 2, \\
\|v_{\mathcal{N}_3}(k)\|^2 &= \|v_3(k) - v_1(k)\|^2 + \|v_3(k) - v_2(k)\|^2 \\
&\leq (0.01 - (-0.01))^2 \times 2, \\
\|v_{\mathcal{N}_4}(k)\|^2 &= \|v_4(k) - v_5(k)\|^2 + \|v_4(k) - v_6(k)\|^2 \\
&\quad + \|v_4(k) - v_7(k)\|^2 \\
&\leq (0.01 - (-0.01))^2 \times 3, \\
\|v_{\mathcal{N}_5}(k)\|^2 &= \|v_5(k) - v_2(k)\|^2 + \|v_5(k) - v_4(k)\|^2 \\
&\quad + \|v_5(k) - v_6(k)\|^2 \\
&\leq (0.01 - (-0.01))^2 \times 3, \\
\|v_{\mathcal{N}_6}(k)\|^2 &= \|v_6(k) - v_4(k)\|^2 + \|v_6(k) - v_5(k)\|^2 \\
&\leq (0.01 - (-0.01))^2 \times 2, \\
\|v_{\mathcal{N}_7}(k)\|^2 &= \|v_7(k) - v_4(k)\|^2 \\
&\leq (0.01 - (-0.01))^2.
\end{aligned}$$

Further, the l_2-norm bounds of augmented disturbance vectors and augmented relative noise vectors can be set as $\sigma_{w_{\mathcal{N}_1}} = \sqrt{0.02}$, $\sigma_{w_{\mathcal{N}_2}} = \sqrt{0.03}$, $\sigma_{w_{\mathcal{N}_3}} = \sqrt{0.03}$, $\sigma_{w_{\mathcal{N}_4}} = \sqrt{0.04}$, $\sigma_{w_{\mathcal{N}_5}} = \sqrt{0.04}$, $\sigma_{w_{\mathcal{N}_6}} = \sqrt{0.03}$, $\sigma_{w_{\mathcal{N}_7}} = \sqrt{0.02}$, $\sigma_{v_{\mathcal{N}_1}} = \sqrt{0.0001}$, $\sigma_{v_{\mathcal{N}_2}} = \sqrt{0.0002}$, $\sigma_{v_{\mathcal{N}_3}} = \sqrt{0.0002}$, $\sigma_{v_{\mathcal{N}_4}} = \sqrt{0.0003}$, $\sigma_{v_{\mathcal{N}_5}} = \sqrt{0.0003}$, $\sigma_{v_{\mathcal{N}_6}} = \sqrt{0.0002}$, $\sigma_{v_{\mathcal{N}_7}} = \sqrt{0.0001}$,

In the simulation, we set the absolute measurement faults of agents 3 and 4 as

$$f_3(k) = \begin{cases} 0, & 0 \leq k < 35, \\ 1, & 35 \leq k \leq 100, \end{cases} \qquad f_4(k) = \begin{cases} 0, & 0 \leq k < 20, \\ 1, & 20 \leq k \leq 100. \end{cases}$$

Additionally, we set $f_1(k) = f_2(k) = f_5(k) = f_6(k) = f_7(k) = 0, \quad k = 0, 1, \cdots, 100$.

Moreover, the augmented fault vectors are defined as

$$\begin{aligned}
f_{\mathcal{N}_1}(k) &= col\{f_1(k), f_3(k)\}, \\
f_{\mathcal{N}_2}(k) &= col\{f_2(k), f_3(k), f_5(k)\}, \\
f_{\mathcal{N}_3}(k) &= col\{f_3(k), f_1(k), f_2(k)\}, \\
f_{\mathcal{N}_4}(k) &= col\{f_4(k), f_5(k), f_6(k), f_7(k)\}, \\
f_{\mathcal{N}_5}(k) &= col\{f_5(k), f_2(k), f_4(k), f_6(k)\}, \\
f_{\mathcal{N}_6}(k) &= col\{f_6(k), f_4(k), f_5(k)\}, \\
f_{\mathcal{N}_7}(k) &= col\{f_7(k), f_4(k)\}.
\end{aligned}$$

To satisfy the minimum condition in Lemma 2, we set the parameter γ_i in (3) as $\gamma_i = 1.01$, $\forall i \in \mathcal{V}$.

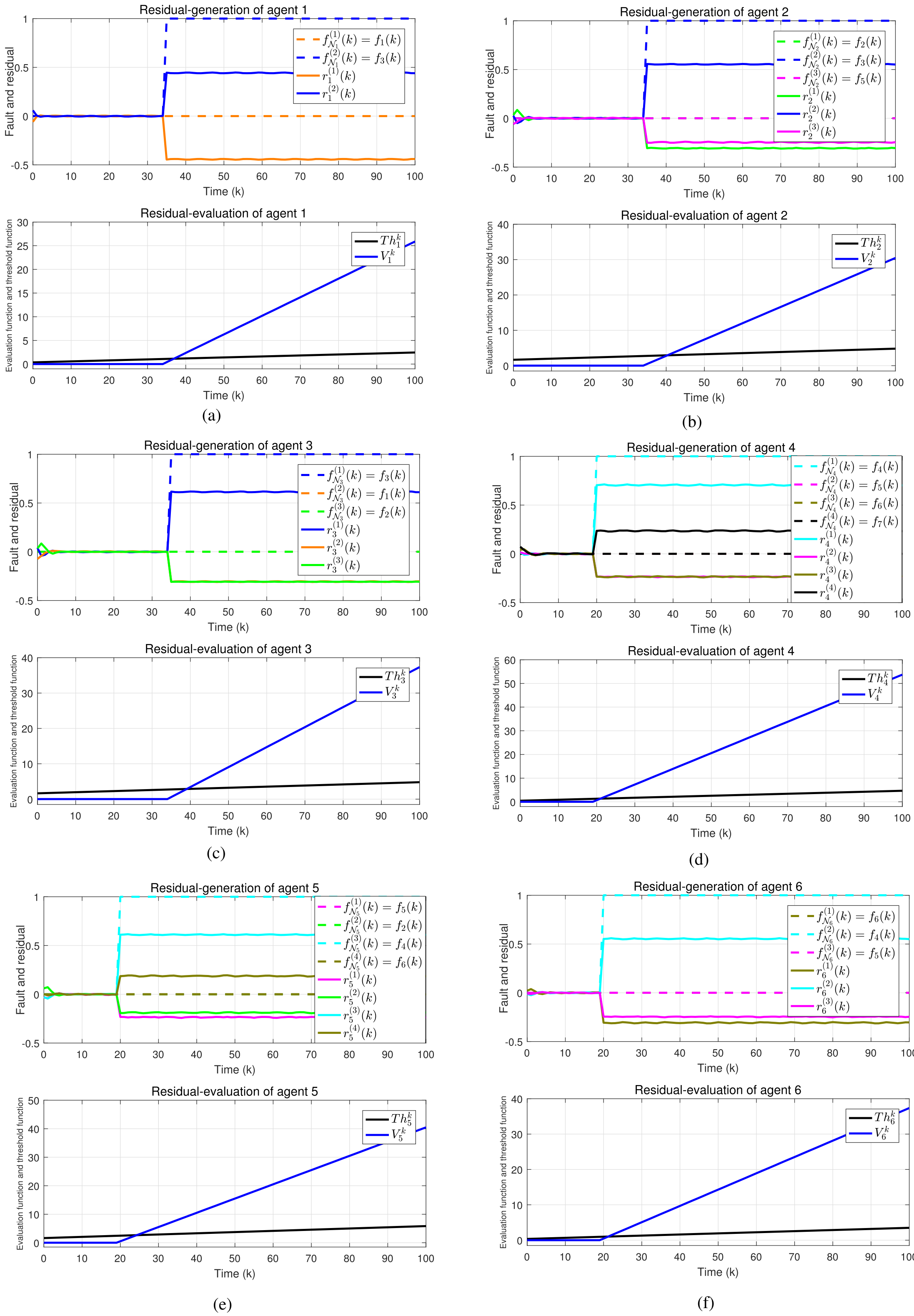
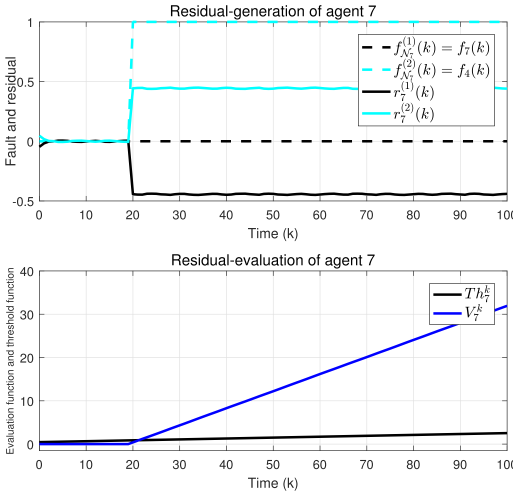


FIGURE 2. Residual-generation stages with the augmented fault $f_{\mathcal{N}_i}(k) = [f_i(k), f_{i_1}(k), \cdots, f_{i_{|\mathcal{N}_i|}}(k)]^T$ and the residual $r_i(k) = [r_i^{(1)}(k), r_i^{(2)}(k), \cdots, r_i^{(|\mathcal{N}_i|+1)}(k)]^T$ for $i = 1, 2, \cdots, 7$; Residual-evaluation stages with the evaluation function V_i^k and the threshold function Th_i^k for $i = 1, 2, \cdots, 7$.

The residual-generation stages of the 7 agents are shown on the upper part of each subfigure in Fig. 2 respectively, where the fault entries and residual entries corresponding to the same positions have the same colors. Additionally,



(g)

FIGURE 2. ***(Continued.)*** **Residual-generation stages with the augmented fault $f_{\mathcal{N}_i}(k) = [f_i(k), f_{i_1}(k), \cdots, f_{i_{|\mathcal{N}_i|}}(k)]^T$ and the residual $r_i(k) = [r_i^{(1)}(k), r_i^{(2)}(k), \cdots, r_i^{(|\mathcal{N}_i|+1)}(k)]^T$ for $i = 1, 2, \cdots, 7$; Residual-evaluation stages with the evaluation function V_i^k and the threshold function Th_i^k for $i = 1, 2, \cdots, 7$.**

the evaluation-stages of the 7 agents are shown on the lower part of each subfigure in Fig. 2 respectively, where the residual-evaluation function V_i^k is compared with the threshold function Th_i^k, $\forall i \in \mathcal{V}$.

For the sake of simplicity and without loss of generality, the performance of agent 2 shown in Fig. 2. (b) will be analyzed as an example. The residual-generation stage of agent 2 shows that the residual $r_2(k) = [r_2^{(1)}(k), r_2^{(2)}(k), r_2^{(3)}(k)]^T$ changes sharply immediately after the sensor fault $f_3(k)$ occurs at its neighbor agent 3. The corresponding residual-evaluation stage shows that the fault alarm is raised at $k = 41$ because at this time the evaluation function V_2^k exceeds the threshold Th_2^k, indicating the fault occurrence in the set $\{2\} \cup \mathcal{N}_2$. On the other hand, we note that there is a 6 s delay between the fault occurrence and the alarm-raising because of the accumulation characteristic of V_2^k and Th_2^k, which is a drawback of the proposed method as mentioned in Remark 3. Furthermore, it is observed that the positive fault signal $f_3(k)$ have a clear influence on each entry of the residual $r_2(k)$, and the entry $r_2^{(2)}(k)$ at the corresponding position is also positive due to the setting of the coefficient matrices.

According to the topology shown in Fig. 1, each agent is either faulty or adjacent to one faulty agent. Therefore, it is observed from the residual-evaluation stages in Fig. 2 that agents 1-7 all raise alarms, indicating that each of these agents detects at least one fault in the union of itself and their neighbors, which verifies the effectiveness of the proposed FD method.

Further, if it is stipulated that there are at most two faulty agents in this multi-agent system, the faulty agents 3 and 4 can be easily isolated according to the relative output topology in Fig. 1.

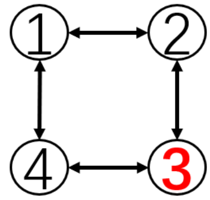


FIGURE 3. Relative output topology of the MAS in EXAMPLE 2, where faulty agent 3 is marked in red.

B. EXAMPLE 2

In this example, the GE F404 gas turbine engine system model, which was borrowed by [37] from [48], is investigated. We consider that the MAS consists of 4 F404 engines and that each engine is described by the continuous-time model $\dot{x}_i(t) = \begin{bmatrix} -1.46 & 0 & 0.2480 \\ 0.1643 & -0.4 & -0.3788 \\ 0.3107 & 0 & -2.23 \end{bmatrix} x_i(t) + \begin{bmatrix} 0.2 & 0.2 \\ 0.8 & 0.8 \\ -0.2 & -0.2 \end{bmatrix} w_i(t)$, $\forall i \in \mathcal{V} = \{1, 2, 3, 4\}$, where $x_i(t) = [x_i^{(1)}(t), x_i^{(2)}(t), x_i^{(3)}(t)]^T$ is the engine state, with $x_i^{(1)}(t)$ and $x_i^{(2)}(t)$ representing the horizontal positions and $x_i^{(3)}(t)$ denoting the altitude of the aircraft. Taking the sample period as 0.3 s and noting that the coefficient matrices may be time-varying due to environmental changes, we consider the LDTV models of the 4 engines with $A_i(k) = \begin{bmatrix} 0.6474 + 0.1isin(k) & 0 & 0.0429 \\ 0.0339 & 0.8869 & -0.0764 \\ 0.0538 & 0 & 0.5141 + 0.1icos(k) \end{bmatrix}$ and $B_{w_i}(k) = \begin{bmatrix} 0.0471 & 0.0471 \\ 0.2299 & 0.2299 \\ -0.0418 & -0.0418 \end{bmatrix}$, $\forall i \in \mathcal{V}$.

To observe engine conditions, on-board engine monitoring systems (EMSs) are usually employed in gas turbines, see [48]. We assume that the EMS on engine i, $\forall i \in \mathcal{V}$ is equipped with one sensor for the absolute measurement $y_i(k)$ and other sensors for the relative output measurements such as $y_{ij}(k)$ with engine j being the neighbor of engine i. The fault vector is composed as $f_i(k) = [f_{ip}(k), f_{is}(k)]^T$, where $f_{ip}(k)$ denotes the fault occurring on the engine dynamics and $f_{is}(k)$ denotes the fault occurring on the absolute measurement sensor. Other coefficient matrices for the absolute measurement sensor in (1) are given as $C_1(k) = [-0.5, 0.8, 0.5sin(k)]$, $C_2(k) = [-0.4, 0.6, 0.6cos(k)]$, $C_3(k) = [-0.9, 0.5, 0.6sin(k)]$, $C_4(k) = [-0.7, 0.4, 0.5cos(k)]$, $B_{f_i}(k) = \begin{bmatrix} 1 & 0 \\ 1.8 & 0 \\ 1.2 & 0 \end{bmatrix}$, $D_{f_i}(k) = [4i, 6i]$, $\forall i \in \mathcal{V}$.

The relative output topology is shown in Fig. 3, with the adjacency matrix given as

$$\mathcal{A} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$

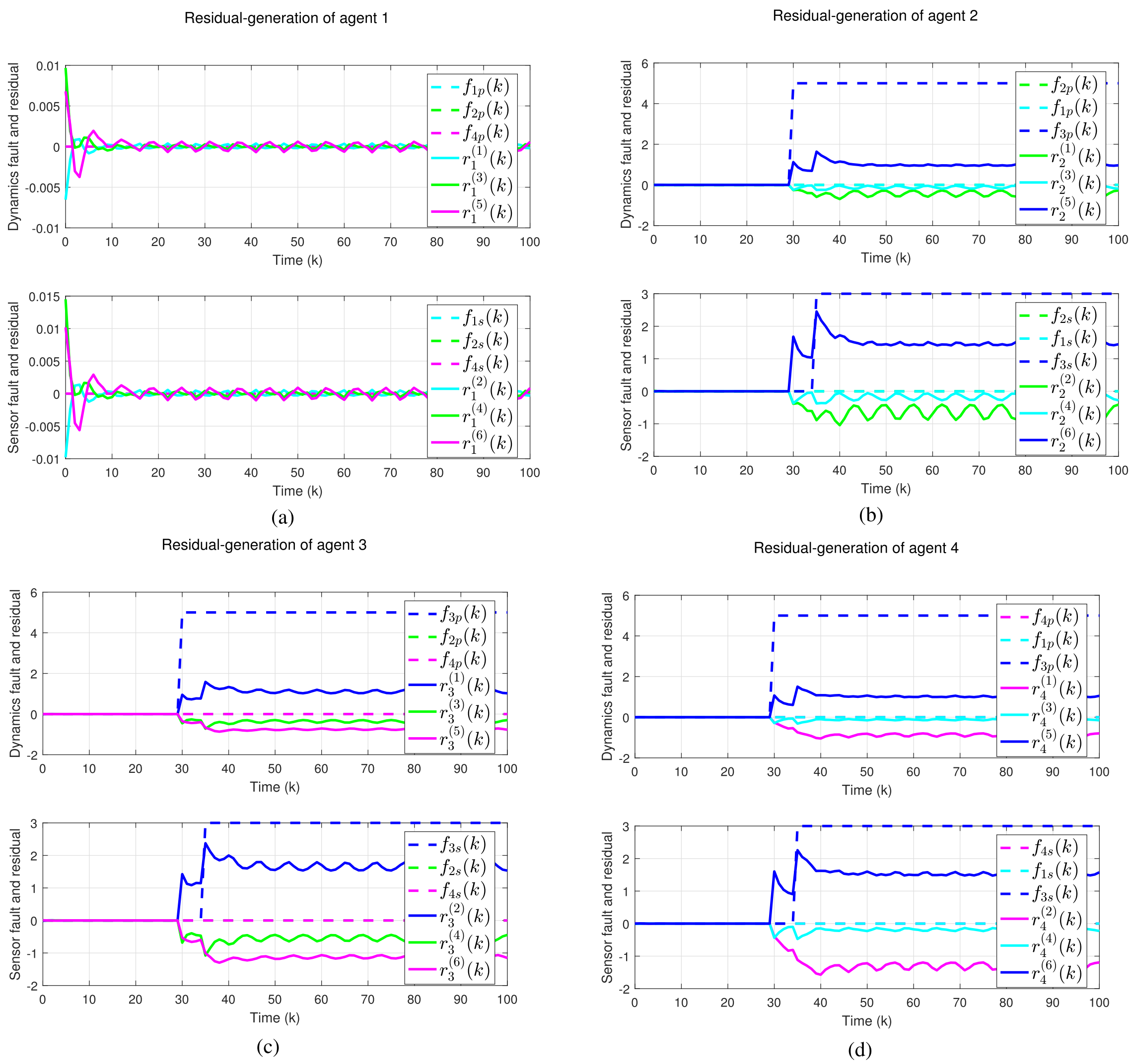


FIGURE 4. **Residual-generation stages with the augmented fault $f_{\mathcal{N}_i}(k) = [f_{ip}(k), f_{is}(k), f_{i_1p}(k), f_{i_1s}(k), \cdots, f_{i_{|\mathcal{N}_i|}p}(k), f_{i_{|\mathcal{N}_i|}s}(k)]^T$ and the residual $r_i(k) = [r_i^{(1)}(k), r_i^{(2)}(k), \cdots, r_i^{(2|\mathcal{N}_i|+2)}(k)]^T$ for $i = 1, 2, 3, 4$.**

The initial engine states are given as $x_i(0) = [0.1i, 0.3i, 0.2i]^T$, $\forall i \in \mathcal{V}$. Furthermore, the external disturbance is given as $w_i(k) = [0.2isin(2k), 0.1icos(2k)]^T$, and the measurement noise is given as $v_i(k) = 0.01icos(k)$, $\forall i \in \mathcal{V}$. Assuming that the l_2-norm bounds of disturbance vectors and noise vectors are known as $\|w_i(k)\|^2 \leq 0.01i^2 + 0.03i^2$ and $\|v_i(k)\|^2 \leq (0.01i)^2$, $\forall k \in N$, we have

$$\begin{aligned}
\|w_{\mathcal{N}_1}(k)\|^2 &= \|w_1(k)\|^2 + \|w_2(k)\|^2 + \|w_4(k)\|^2 \\
&\leq 0.04 \times 1^2 + 0.04 \times 2^2 + 0.04 \times 4^2, \\
\|w_{\mathcal{N}_2}(k)\|^2 &= \|w_2(k)\|^2 + \|w_1(k)\|^2 + \|w_3(k)\|^2 \\
&\leq 0.04 \times 2^2 + 0.04 \times 1^2 + 0.04 \times 3^2, \\
\|w_{\mathcal{N}_3}(k)\|^2 &= \|w_3(k)\|^2 + \|w_2(k)\|^2 + \|w_4(k)\|^2 \\
&\leq 0.04 \times 3^2 + 0.04 \times 2^2 + 0.04 \times 4^2, \\
\|w_{\mathcal{N}_4}(k)\|^2 &= \|w_4(k)\|^2 + \|w_1(k)\|^2 + \|w_3(k)\|^2 \\
&\leq 0.04 \times 4^2 + 0.04 \times 1^2 + 0.04 \times 3^2,
\end{aligned}$$

$$\begin{aligned}
\|v_{\mathcal{N}_1}(k)\|^2 &= \|v_1(k) - v_2(k)\|^2 + \|v_1(k) - v_4(k)\|^2 \\
&\leq (0.01 - (-0.02))^2 + (0.01 - (-0.04))^2, \\
\|v_{\mathcal{N}_2}(k)\|^2 &= \|v_2(k) - v_1(k)\|^2 + \|v_2(k) - v_3(k)\|^2 \\
&\leq (0.02 - (-0.01))^2 + (0.02 - (-0.03))^2, \\
\|v_{\mathcal{N}_3}(k)\|^2 &= \|v_3(k) - v_2(k)\|^2 + \|v_3(k) - v_4(k)\|^2 \\
&\leq (0.03 - (-0.02))^2 + (0.03 - (-0.04))^2, \\
\|v_{\mathcal{N}_4}(k)\|^2 &= \|v_4(k) - v_1(k)\|^2 + \|v_4(k) - v_3(k)\|^2 \\
&\leq (0.04 - (-0.01))^2 + (0.04 - (-0.03))^2,
\end{aligned}$$

and the l_2-norm bounds of the augmented disturbance vectors and augmented relative noise vectors can be set as $\sigma_{w_{\mathcal{N}_1}} = 0.9165$, $\sigma_{w_{\mathcal{N}_2}} = 0.7483$, $\sigma_{w_{\mathcal{N}_3}} = 1.0770$, $\sigma_{w_{\mathcal{N}_4}} = 1.0198$, $\sigma_{v_{\mathcal{N}_1}} = 0.0583$, $\sigma_{v_{\mathcal{N}_2}} = 0.0583$, $\sigma_{v_{\mathcal{N}_3}} = 0.0860$, $\sigma_{v_{\mathcal{N}_4}} = 0.0860$.

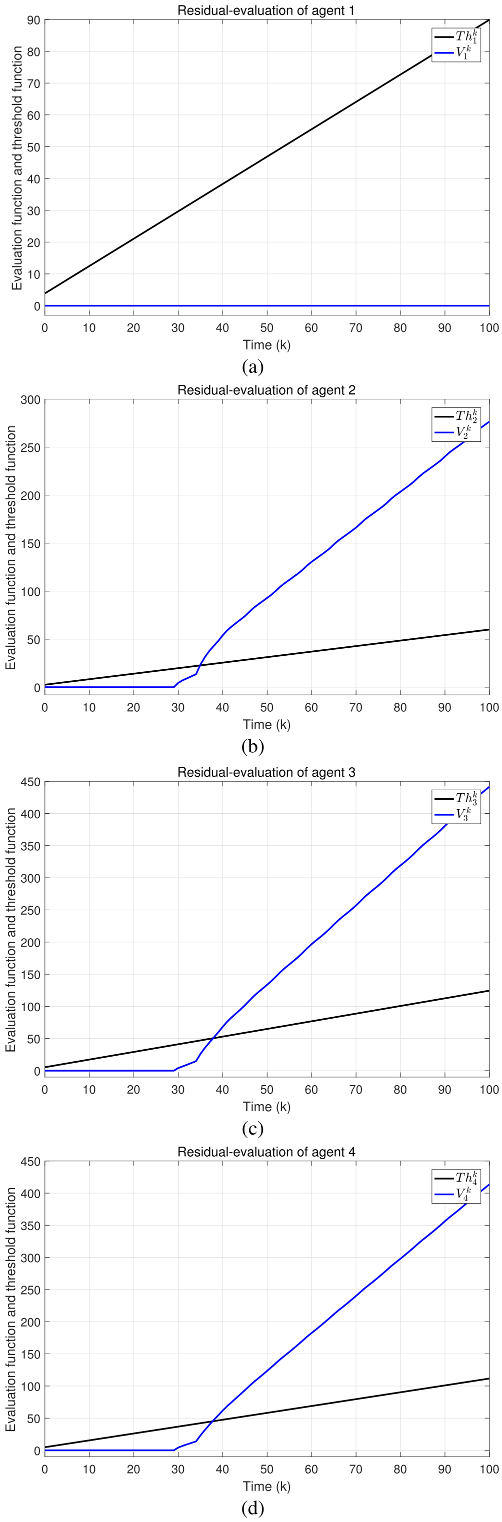


FIGURE 5. Residual-evaluation stages with the evaluation function V_i^k and the threshold function Th_i^k for $i = 1, 2, 3, 4$.

In the simulation, we set

$$f_{3p}(k) = \begin{cases} 0, & 0 \leq k < 30, \\ 5, & 30 \leq k \leq 100, \end{cases} \quad f_{3s}(k) = \begin{cases} 0, & 0 \leq k < 35, \\ 3, & 35 \leq k \leq 100. \end{cases}$$

Additionally, we set $f_{1p}(k) = f_{1s}(k) = f_{2p}(k) = f_{2s}(k) = f_{4p}(k) = f_{4s}(k) = 0, \ k = 0, 1, \cdots, 100$.

Moreover, the augmented fault vectors are defined as

$$\begin{aligned} f_{\mathcal{N}_1}(k) &= col\{f_1(k), f_2(k), f_4(k)\} \\ &= [f_{1p}(k), f_{1s}(k), f_{2p}(k), f_{2s}(k), f_{4p}(k), f_{4s}(k)]^T, \\ f_{\mathcal{N}_2}(k) &= col\{f_2(k), f_1(k), f_3(k)\} \\ &= [f_{2p}(k), f_{2s}(k), f_{1p}(k), f_{1s}(k), f_{3p}(k), f_{3s}(k)]^T, \\ f_{\mathcal{N}_3}(k) &= col\{f_3(k), f_2(k), f_4(k)\} \\ &= [f_{3p}(k), f_{3s}(k), f_{2p}(k), f_{2s}(k), f_{4p}(k), f_{4s}(k)]^T, \\ f_{\mathcal{N}_4}(k) &= col\{f_4(k), f_1(k), f_3(k)\} \\ &= [f_{4p}(k), f_{4s}(k), f_{1p}(k), f_{1s}(k), f_{3p}(k), f_{3s}(k)]^T. \end{aligned}$$

To satisfy the minimum condition in Lemma 2, we set the parameter γ_i in (3) as $\gamma_i = 1.01, \forall i \in \mathcal{V}$. The residual-generation stages of the 4 engines are shown in Fig. 4, where the fault entries and residual entries corresponding to the same positions have the same colors. Additionally, the evaluation-stages of the 4 engines are shown in Fig. 5.

Fig. 5. (a) indicates that no fault was detected in the set $\{1\} \cup \mathcal{N}_1 = \{1, 2, 4\}$, Fig. 5. (b) indicates the fault occurrence in the set $\{2\} \cup \mathcal{N}_2 = \{2, 1, 3\}$, Fig. 5. (c) indicates the fault occurrence in the set $\{3\} \cup \mathcal{N}_3 = \{3, 2, 4\}$, and Fig. 5. (d) indicates the fault occurrence in the set $\{4\} \cup \mathcal{N}_4 = \{4, 1, 3\}$. The effectiveness of the FD method is verified by these subfigures.

Furthermore, if it is known that there is at most one faulty agent, the faulty agent 3 can be easily isolated because only it satisfies the condition that any other raising-alarm agent is adjacent to it.

VII. CONCLUSION

In this paper, we have investigated the distributed fault detection problem for linear discrete time-varying heterogeneous multi-agent systems with relative output information. Under the H_∞ filtering framework, the distributed residual-generator has been designed using all locally obtained relative output information, with the Krein space projection theory introduced to reduce the computational burden. Further, a distributed residual-evaluation mechanism has been given that detects the faults by comparing the evaluation function with the threshold function and guarantees a zero false alarm rate. The effectiveness of the proposed fault detection approach has been verified through two illustrative examples, where the faulty agents can be isolated under certain conditions based on a fault isolation criterion.

To overcome the limitations of this current proposed scheme such as the lack of a balance between the false alarm rate and missed detection rate, possible future research may include developing more efficient FD methods to reduce the missed detection rate and investigating more robust fault isolation methods.

REFERENCES

[1] A. Freedy, O. Sert, E. Freedy, J. McDonough, G. Weltman, M. Tambe, T. Gupta, W. Grayson, and P. Cabrera, "Multiagent adjustable autonomy framework (MAAF) for multi-robot, multi-human teams," in *Proc. Int. Symp. Collaborative Technol. Syst.*, May 2008, pp. 498–505.

[2] T. D. Ngo, H. Raposo, and H. Schiøler, "Multiagent robotics: Toward energy autonomy," *Artif. Life Robot.*, vol. 12, nos. 1–2, pp. 47–52, Mar. 2008.

[3] D. Yu and C. L. P. Chen, "Distributed generation and control of persistent formation for multi-agent systems," in *Proc. 4th Int. Conf. Inf., Cybern. Comput. Social Syst. (ICCSS)*, Jul. 2017, pp. 142–147.

[4] Z. Peng, J. Hu, R. Luo, and B. K. Ghosh, "Distributed multi-agent temporal-difference learning with full neighbor information," *Control Theory Technol.*, vol. 18, no. 4, pp. 379–389, Dec. 2020, doi: 10.1007/s11768-020-00016-w.

[5] M. Tao, X. Li, X. Guan, Z. Liu, Y. Feng, and R. Qu, "The research of coordination control on distributed multi-agent," in *Proc. 33rd Youth Acad. Annu. Conf. Chin. Assoc. Automat. (YAC)*, 2018, pp. 328–333.

[6] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.

[7] X. Fan, P. Bai, H. Li, X. Deng, and M. Lv, "Adaptive fuzzy finite-time tracking control of uncertain non-affine multi-agent systems with input quantization," *IEEE Access*, vol. 8, pp. 187623–187633, 2020.

[8] H. Zhang, Y. Liu, J. Dai, and Y. Wang, "Command filter based adaptive fuzzy finite-time control for a class of uncertain nonlinear systems with hysteresis," *IEEE Trans. Fuzzy Syst.*, early access, Jun. 18, 2020, doi: 10.1109/TFUZZ.2020.3003499.

[9] H. Liang, X. Guo, Y. Pan, and T. Huang, "Event-triggered fuzzy bipartite tracking control for network systems based on distributed reduced-order observers(revised manuscript of TFS-2019-1049)," *IEEE Trans. Fuzzy Syst.*, early access, Mar. 23, 2020, doi: 10.1109/TFUZZ.2020.2982618.

[10] H. Liang, G. Liu, H. Zhang, and T. Huang, "Neural-network-based event-triggered adaptive control of nonaffine nonlinear multiagent systems with dynamic uncertainties," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jul. 14, 2020, doi: 10.1109/TNNLS.2020.3003950.

[11] J. Cai, R. Yu, B. Wang, C. Mei, and L. Shen, "Decentralized event-triggered control for interconnected systems with unknown disturbances," *J. Franklin Inst.*, vol. 357, no. 3, pp. 1494–1515, Feb. 2020.

[12] B. Kaviarasan, O.-M. Kwon, M. J. Park, and R. Sakthivel, "Stochastic faulty estimator-based non-fragile tracking controller for multi-agent systems with communication delay," *Appl. Math. Comput.*, vol. 392, Mar. 2021, Art. no. 125704.

[13] Z. Ahmed, M. M. Khan, M. A. Saeed, and W. Zhang, "Consensus control of multi-agent systems with input and communication delay: A frequency domain perspective," *ISA Trans.*, vol. 101, pp. 69–77, Jun. 2020.

[14] C. Deng, W.-W. Che, and P. Shi, "Cooperative fault-tolerant output regulation for multiagent systems by distributed learning control approach," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 11, pp. 4831–4841, Nov. 2020.

[15] C. Deng and C. Wen, "Distributed resilient observer-based fault-tolerant control for heterogeneous multiagent systems under actuator faults and DoS attacks," *IEEE Trans. Control Netw. Syst.*, vol. 7, no. 3, pp. 1308–1318, Sep. 2020.

[16] B. Wang, B. Zhang, and R. Su, "Optimal tracking cooperative control for cyber-physical systems: Dynamic fault-tolerant control and resilient management," *IEEE Trans. Ind. Informat.*, vol. 17, no. 1, pp. 158–167, Jan. 2021.

[17] Z.-H. Zhang and G.-H. Yang, "Interval observer-based fault isolation for discrete-time fuzzy interconnected systems with unknown interconnections," *IEEE Trans. Cybern.*, vol. 47, no. 9, pp. 2413–2424, Sep. 2017.

[18] B. Farrera, F.-R. Lṕez-Estrada, M. Chadli, G. Valencia-Palomo, and S. Gómez-Peñate, "Distributed fault estimation of multi–agent systems using a proportional–integral observer: A leader–following application," *Int. J. Appl. Math. Comput. Sci.*, vol. 30, no. 3, pp. 551–560, 2020.

[19] N. Meskin and K. Khorasani, "Fault detection and isolation of discrete-time Markovian jump linear systems with application to a network of multi-agent systems having imperfect communication channels," *Automatica*, vol. 45, no. 9, pp. 2032–2040, Sep. 2009.

[20] Y. Li, H. Fang, J. Chen, and C. Yu, "Distributed cooperative fault detection for multiagent systems: A mixed H_∞/H_2 optimization approach," *IEEE Trans. Ind. Electron.*, vol. 65, no. 8, pp. 6468–6477, Aug. 2018.

[21] Y. Li, H. Fang, and J. Chen, "Anomaly detection and identification for multiagent systems subjected to physical faults and cyberattacks," *IEEE Trans. Ind. Electron.*, vol. 67, no. 11, pp. 9724–9733, Nov. 2020.

[22] I. Shames, A. M. H. Teixeira, H. Sandberg, and K. H. Johansson, "Distributed fault detection for interconnected second-order systems," *Automatica*, vol. 47, no. 12, pp. 2757–2764, Dec. 2011.

[23] W. Jia and J. Wang, "Partial-nodes-based distributed fault detection and isolation for second-order multiagent systems with exogenous disturbances," *IEEE Trans. Cybern.*, early access, Jul. 28, 2020, doi: 10.1109/TCYB.2020.3007655.

[24] M. R. Davoodi, K. Khorasani, H. A. Talebi, and H. R. Momeni, "A novel distributed robust fault detection and isolation filter design for a network of nonhomogeneous multi-agent systems," in *Proc. 51st IEEE Conf. Decis. Control (CDC)*, Dec. 2012, pp. 592–599.

[25] Y. Quan, W. Chen, Z. Wu, and L. Peng, "Observer-based distributed fault detection and isolation for heterogeneous discrete-time multi-agent systems with disturbances," *IEEE Access*, vol. 4, pp. 4652–4658, 2016.

[26] D. Xu, F. Zhu, Z. Zhou, and X. Yan, "Distributed fault detection and estimation in cyber–physical systems subject to actuator faults," *ISA Trans.*, vol. 104, pp. 162–174, Sep. 2020.

[27] F. Arrichiello, A. Marino, and F. Pierri, "Observer-based decentralized fault detection and isolation strategy for networked multirobot systems," *IEEE Trans. Control Syst. Technol.*, vol. 23, no. 4, pp. 1465–1476, Jan. 2015.

[28] F. Boem, R. M. G. Ferrari, C. Keliris, T. Parisini, and M. M. Polycarpou, "A distributed networked approach for fault detection of large-scale systems," *IEEE Trans. Autom. Control*, vol. 62, no. 1, pp. 18–33, Jan. 2017.

[29] X. He, Z. Wang, and D. H. Zhou, "Robust fault detection for networked systems with communication delay and data missing," *Automatica*, vol. 45, no. 11, pp. 2634–2639, Nov. 2009.

[30] Y. Wang, S. X. Ding, H. Ye, and G. Wang, "A new fault detection scheme for networked control systems subject to uncertain time-varying delay," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5258–5268, Oct. 2008.

[31] Q. Zong, F. Zeng, W. Liu, Y. Ji, and Y. Tao, "Sliding mode observer-based fault detection of distributed networked control systems with time delay," *Circuits, Syst., Signal Process.*, vol. 31, no. 1, pp. 203–222, Feb. 2012.

[32] Y. Quan, W. Chen, Z. Wu, and L. Peng, "Distributed fault detection for second-order delayed multi-agent systems with adversaries," *IEEE Access*, vol. 5, pp. 16478–16483, 2017.

[33] X. He, Z. Wang, Y. Liu, and D. H. Zhou, "Least-squares fault detection and diagnosis for networked sensing systems using a direct state estimation approach," *IEEE Trans. Ind. Informat.*, vol. 9, no. 3, pp. 1670–1679, Aug. 2013.

[34] M. Gao, S. Yang, L. Sheng, and D. Zhou, "Fault diagnosis for time-varying systems with multiplicative noises over sensor networks subject to round-robin protocol," *Neurocomputing*, vol. 346, pp. 65–72, Jun. 2019.

[35] M. Gao, S. Yang, and L. Sheng, "Distributed fault estimation for time-varying multi-agent systems with sensor faults and partially decoupled disturbances," *IEEE Access*, vol. 7, pp. 147905–147913, 2019.

[36] B. Shen, Z. Wang, H. Dong, and S. Zhang, "Finite-horizon distributed H_∞ fault estimation for time-varying systems in sensor networks: A Krein-space approach," *IFAC PapersOnLine*, vol. 48, no. 21, pp. 48–53, 2015.

[37] X. Ge, Q.-L. Han, M. Zhong, and X.-M. Zhang, "Distributed Krein space-based attack detection over sensor networks under deception attacks," *Automatica*, vol. 109, Nov. 2019, Art. no. 108557.

[38] P. Barooah and J. P. Hespanha, "Estimation on graphs from relative measurements," *IEEE Control Syst.*, vol. 27, no. 4, pp. 57–74, Aug. 2007.

[39] N. Meskin and K. Khorasani, "Actuator fault detection and isolation for a network of unmanned vehicles," *IEEE Trans. Autom. Control*, vol. 54, no. 4, pp. 835–840, Apr. 2009.

[40] P. P. Menon and C. Edwards, "Robust fault estimation using relative information in linear multi-agent networks," *IEEE Trans. Autom. Control*, vol. 59, no. 2, pp. 477–482, Feb. 2014.

[41] Y. Lv, G. Wen, T. Huang, and Z. Duan, "Adaptive attack-free protocol for consensus tracking with pure relative output information," *Automatica*, vol. 117, Jul. 2020, Art. no. 108998.

[42] M. Chadli, M. Davoodi, and N. Meskin, "Distributed fault detection and isolation filter design for heterogeneous multi-agent LPV systems," in *Proc. Amer. Control Conf. (ACC)*, May 2017, pp. 1610–1615.

[43] M. Chadli, M. Davoodi, and N. Meskin, "Distributed state estimation, fault detection and isolation filter design for heterogeneous multi-agent linear parameter-varying systems," *IET Control Theory Appl.*, vol. 11, no. 2, pp. 254–262, Jan. 2017.

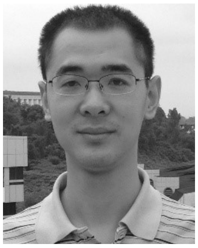
[44] M. R. Davoodi, K. Khorasani, H. A. Talebi, and H. R. Momeni, "Distributed fault detection and isolation filter design for a network of heterogeneous multiagent systems," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 3, pp. 1061–1069, May 2014.

[45] B. Hassibi, A. H. Sayed, and T. Kailath, *Indefinite-Quadratic Estimation and Control: A Unified Approach to H_2 and H_∞ Theories*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1999.

[46] B. Hassibi, A. H. Sayed, and T. Kailath, "Linear estimation in Krein spaces. I. Theory," *IEEE Trans. Autom. Control*, vol. 41, no. 1, pp. 18–33, Jan. 1996.
[47] B. Hassibi, A. H. Sayed, and T. Kailath, "Linear estimation in Krein spaces. II. Applications," *IEEE Trans. Autom. Control*, vol. 41, no. 1, pp. 34–49, Jan. 1996.
[48] R. W. Eustace, B. A. Woodyatt, G. L. Merrington, and A. Runacres, "Fault signatures obtained from fault implant tests on an F404 engine," *J. Eng. Gas Turbines Power*, vol. 116, no. 1, pp. 178–183, Jan. 1994.

PEILU ZOU received the B.S. degree from the School of Mechanical Electronic and Information Engineering, China University of Mining and Technology-Beijing, in 2019. She is currently pursuing the M.S. degree with the School of Automation, Beijing Institute of Technology, Beijing, China. Her current research interests include fault diagnosis of multi-agent systems and fault estimation of time-varying systems.

PING WANG received the M.S. degree from the Institute of Automation, Qufu Normal University, Qufu, China, in 2018. She is currently pursuing the Ph.D. degree with the School of Automation, Beijing Institute of Technology, Beijing, China. Her current research interests include nonlinear system adaptive control, fault detection, and fault-tolerant control of networked systems.



CHENGPU YU received the B.E. and M.E. degrees in electrical engineering from the University of Electronic Science and Technology of China, in 2006 and 2009, respectively, and the Ph.D. degree in electrical engineering from Nanyang Technological University, Singapore, in 2014. From 2013 to 2014, he was with the Internet of Things Laboratory, Nanyang Technological University, as a Research Associate. From 2014 to 2017, he held a postdoctoral position with the Delft Center for Systems and Control. Since 2018, he has been with the Beijing Institute of Technology, as a Full Professor. His research interests include system identification, distributed optimization, and optical imaging.