Set-theoretic methods in robust detection and isolation of sensor faults

Feng Xu, Vicenç Puig, Carlos Ocampo-Martinez, Sorin Olaru & Florin Stoican

Taylor & Francis
Taylor & Francis Group

Set-theoretic methods in robust detection and isolation of sensor faults

Feng Xu[a], Vicenç Puig[a,*], Carlos Ocampo-Martinez[a], Sorin Olaru[b] and Florin Stoican[c]

[a]Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Technical University of Catalonia (UPC), Barcelona, Spain; [b]Automatic Control Department, E3S (SUPELEC Systems Sciences), Gif sur Yvette, France; [c]Department of Automatic Control and Systems Engineering, Faculty of Automatic Control and Computers 'Politehnica' University of Bucharest, Bucharest, Romania

This paper proposes a sensorfault detection and isolation (FDI) approach based on interval observers and invariant sets. In fault detection (FD), both interval observer-based and invariant set-based mechanisms are used to provide real-time fault alarms. In fault isolation (FI), the proposed approach also uses these two different mechanisms. The former, based on interval observers, aims to isolate faults during the transient-state operation induced by faults. If the former does not succeed, the latter, based on both interval observers and invariant sets, is started to guarantee FI after the system enters into steady state. Besides, a collection of invariant set-based FDI conditions are established by using all available system-operating information provided by all interval observers. In order to reduce computational complexity, a method to remove all available but redundant/unnecessary system-operating information is incorporated into this approach. If the considered faults satisfy the proposed FDI conditions, it can be guaranteed that they are detectable and isolable after their occurrences. This paper concludes with a case study based on a subsystem of a wind turbine benchmark, which can illustrate the effectiveness of this FDI technique.

Keywords: fault detection and isolation; interval observers; positively invariant sets; linear systems; zonotopes

1. Introduction

As the industry advances, the technical systems become more and more sophisticated. Generally, the more complex the system is, the more likely it is affected by faults. In order to protect the system once it becomes faulty, it is necessary to timely detect and isolate faults and take remedial measures, which requires effective fault detection and isolation (FDI) techniques (Blanke, Kinnaert, Lunze, & Staroswiecki, 2006). This paper focuses on the set-theoretic methods for detection and isolation of sensor faults (Blanchini & Miani, 2008). In the literature, the set-theoretic fault diagnosis, known as a class of robust techniques, is mainly involved in set-membership estimation, interval observers and invariant sets (Blesa, Puig, Romera, & Saludes, 2011; Puig, 2010; Seron, Zhuo, Doná, & Martínez, 2008).

Set-membership estimation and interval observers provide set-valued estimations for both states and outputs of a monitored system in real time. The fault detection (FD) principles of both techniques consist in testing the consistency between the measured outputs and their corresponding estimated sets. If the measured outputs are consistent with the estimated output sets, it is considered that the system is healthy, otherwise it implies that the system has become faulty. In the literature, set-membership estimation is used for FD and rarely for FI (Chai & Qiao, 2014), while interval observers have been extended for FI mainly in terms

of actuator faults in the previous works of the authors (Xu, Stoican, Puig, Ocampo-Martinez, & Olaru, 2013; Xu, Puig, Ocampo-Martinez, Stoican, & Olaru, 2014).

Instead of estimating state and output sets online, the invariant set-based approach relies on the invariant set characterisation of system behaviours at steady state (Kofman et al., 2007). It is known that a system always has several different modes, a healthy one and at least one faulty one. For each mode, the corresponding invariant set to confine the residual (a signal sensitive to faults) can be constructed. If the invariant sets of all modes are separate from each other, FD can be implemented by testing whether the residual goes out of its healthy invariant set (Seron et al., 2008). After FD, the fault can be isolated by searching which faulty invariant set the residual finally converges to.

Within the knowledge range of the authors, a brief summary of the set-theoretic methods related to this paper is presented. Set-membership state estimation based on zonotopes was discussed in Alamo et al. (2005), Combastel (2003) and Le et al. (2013). The application of set-membership estimation to fault diagnosis of a wind turbine benchmark can be found in Blesa et al. (2011), and interval observers were used for robust state estimation in Mazenc and Bernard (2011). In Guerra et al. (2008) and Raïssi et al. (2010), FD techniques based on interval observers were proposed. In Seron et al. (2008) and Olaru

---

et al. (2010), invariant sets were used to implement sensor FDI in a multi-sensor fault-tolerant control (FTC) scheme.

This paper has two main objectives. The first one is to summarise and collect the previous works of the authors (Xu et al., 2013, 2014) and makes the picture more systematic. The second one is to propose a set-based robust sensor FDI approach with no need of multi-sensor redundancy. It is stated that, to improve the set-based FDI approach, one has to propose new techniques with tighter guaranteed FDI conditions and less complexity. Because of different characteristics between actuator and sensor faults, one has to pay more efforts to implement set-based detection and isolation for sensor faults.

The contribution of this paper is threefold. First, the proposed sensor FDI approach combines the advantages of interval observers in the transient-state regime with the advantages of invariant sets describing the steady-state system operation. Second, the FDI decision is jointly made by both interval observer-based and invariant set-based FDI mechanisms. Third, the proposed approach uses the available system-operating information from all interval observers, selects all useful information for FDI conditions and FDI implementation, but discards all available but unnecessary/redundant system-operating information to reduce computational complexity. These mentioned points imply the obtaining of a balance between computational complexity and FDI conservatism.

In the remaining of this paper, Section 2 introduces the FDI scheme. Section 3 defines the residual signal and constructs relevant invariant sets. Interval observer-based and invariant set-based FDI mechanisms are briefly recapitulated in Section 4. In Section 5, FDI conditions for the proposed technique are given. An FDI algorithm is used to summarise this approach in Section 6. In Section 7, a case study based on a subsystem of a wind turbine benchmark is used to illustrate the effectiveness of the FDI approach. Finally, Section 8 draws the general conclusions.

Throughout this paper, the notation $\mathbb{B}^r$ denotes a box composed of r unitary intervals, $\mathbb{R}^+$ denotes the set of positively real numbers, $\mathbb{R}^{m \times n}$ represents the $m \times n$ real matrix space, the inequalities are understood elementwise, the symbol $\oplus$ denotes the Minkowski sum, I denotes the identity matrix with suitable dimensions, a^T denotes the transpose of a matrix or vector a, $[\overline{x}, \underline{x}]$ denotes an interval, $([\overline{x}_1, \underline{x}_1], [\overline{x}_2, \underline{x}_2], \ldots, [\overline{x}_n, \underline{x}_n])^T$ denotes an n-dimensional interval vector, $\text{diag}(\cdot)$ denotes the diagonal matrix and $\| \cdot \|_s$ is the Euclidean vector s-norm.

2. Problem formulation

2.1. Plant models

The FDI scheme in Figure 1 assumes that the plant under the effect of sensor faults is modelled as

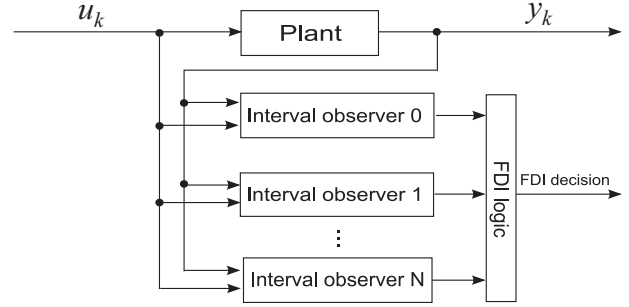$$x_{k+1} = Ax_k + Bu_k + \omega_k, \qquad (1a)$$

Figure 1.   Sensor FDI scheme.

$$y_k = G_i C x_k + \eta_k, \qquad (1b)$$

where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times p}$ and $C \in \mathbb{R}^{q \times n}$ are time-invariant matrices, $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^p$ and $y_k \in \mathbb{R}^q$ are states, inputs and outputs at time instant k, respectively, and ω_k and η_k are process disturbances and measurement noises, respectively.

In (1), G_i ($i \in \mathcal{I} = \{0, 1, \ldots, N\}$) is a $q \times q$ diagonal matrix which characterises the sensor functioning modes. The diagonal elements take values from an interval $[0, 1]$ (i.e., the ith diagonal value of G_i characterises the status of the corresponding sensor with 0, a value inside the interval $(0,1)$ and 1 describing a total outage, a partial outage and the healthy functioning, respectively). By convention, one assumes G_0 to be the identity matrix (denoting the healthy sensor mode) and G_i with $i \neq 0$ a-priori known faulty modes.

Remark 2.1: The number N of the considered faulty modes is different from that of the sensors in a system, because there may be several fault modes corresponding to one sensor or one fault mode corresponding to several sensors.

Assumption 2.1: *The disturbances and noises ω_k and η_k are bounded by the known sets,*

$$W = \{\omega \in \mathbb{R}^n : |\omega - \omega^c| \leq \bar{\omega}, \omega^c \in \mathbb{R}^n, \bar{\omega} \in \mathbb{R}^n\}, \quad (2a)$$

$$V = \{\eta \in \mathbb{R}^q : |\eta - \eta^c| \leq \bar{\eta}, \eta^c \in \mathbb{R}^q, \bar{\eta} \in \mathbb{R}^q\}, \quad (2b)$$

respectively, where the vectors ω^c, η^c, $\bar{\omega}$ and $\bar{\eta}$ have known and constant values.

Remark 2.2: The sets W and V are described by means of zonotopes $W = \omega^c \oplus H_{\bar{\omega}}\mathbb{B}^n$ and $V = \eta^c \oplus H_{\bar{\eta}}\mathbb{B}^q$, where $H_{\bar{\omega}} \in \mathbb{R}^{n \times n}$ and $H_{\bar{\eta}} \in \mathbb{R}^{q \times q}$ are diagonal matrices with their main diagonals composed of $\bar{\omega}$ and $\bar{\eta}$, respectively.

Assumption 2.2: *The pair (A, B) is stabilisable and the pairs (A, G_iC) are detectable for all $i \in \mathcal{I}$.*

Remark 2.3: Considering that the scheme shown in Figure 1 is an open-loop system, the first priority for this scheme is the system stability. Thus, under Assumption 2.2, the input u_k should stabilise the system or the system itself is stable (i.e., A is a Schur matrix).

Assumption 2.3: *The considered faults can persist sufficiently long time such that the FDI module has enough time to detect and isolate them.*

2.2. Interval observers

In this scheme, a bank of interval observers is designed to monitor the system. In accordance with (1), the interval observer corresponding to the *j*th mode ($j \in \mathcal{I}$) is designed as

$$\hat{X}_{k+1}^j = (A - L_j G_j C)\hat{X}_k^j \oplus \{Bu_k\} \oplus \{L_j y_k\}$$
$$\oplus (-L_j)V \oplus W, \tag{3a}$$

$$\hat{Y}_k^j = G_j C \hat{X}_k^j \oplus V, \tag{3b}$$

where $\hat{X}_k^j$ and $\hat{Y}_k^j$ are the state and output sets estimated by the *j*th interval observer, respectively.

Assumption 2.4: *The interval observer gain L_j is chosen to assure that $A - L_j G_j C$ is a Schur matrix (always possible under Assumption 2.2).*

Assumption 2.5: *The initial state of the plant and the initial state set of all interval observers are denoted as x_0 and $\hat{X}_0$, respectively, and $x_0 \in \hat{X}_0$ holds.*

Note that, for different interval observers, the initial sets could be different. However, without loss of generality, one uses one initial set $\hat{X}_0$ for all interval observers. Since W and V are zonotopes, if the initial set $\hat{X}_0$ is chosen as a zonotope, $\hat{X}_{k+1}^j$ and $\hat{Y}_k^j$ are also zonotopes as the Minkowski sum of zonotopes is a closed operation. By using zonotope operations introduced in appendices, (3) can be transformed into the equivalent centre-segment matrix form:

$$\hat{x}_{k+1}^{j,c} = (A - L_j G_j C)\hat{x}_k^{j,c} + Bu_k + L_j y_k - L_j \eta^c + w^c, \tag{4a}$$

$$\hat{H}_{k+1}^{j,x} = \left[(A - L_j G_j C)\hat{H}_k^{j,x} \quad -L_j H_{\tilde{\eta}} \quad H_{\tilde{\omega}}\right], \tag{4b}$$

$$\hat{y}_k^{j,c} = G_j C \hat{x}_k^{j,c} + \eta^c, \tag{4c}$$

$$\hat{H}_k^{j,y} = \left[G_j C \hat{H}_k^{j,x} \quad H_{\tilde{\eta}}\right], \tag{4d}$$

where $\hat{x}_k^{j,c}$, $\hat{y}_k^{j,c}$, $\hat{H}_k^{j,x}$ and $\hat{H}_k^{j,y}$ are the centres and segment matrices of $\hat{X}_k^j$ and $\hat{Y}_k^j$, respectively.

Remark 2.4: Under Assumptions 2.2, 2.4 and 2.5, the state x_k is always contained inside the state set estimated by the interval observer matching the current mode, if there is no mode switching. The same results hold for the outputs and their corresponding sets (i.e., $x_k \in \hat{X}_k^j$ and $y_k \in \hat{Y}_k^j$).

3. Residual analysis

In this section, residuals are defined in terms of zonotopes and the corresponding bounding sets of residual zonotopes (named residual-bounding zonotopes) are derived.

3.1. Residual definition

If the system currently operates in the *i*th mode, the residual zonotope corresponding to the *j*th interval observer at time instant *k* is defined as

$$R_k^{ij} = \{y_k\} \oplus (-\hat{Y}_k^j)$$
$$= \{G_i C x_k + \eta_k\} \oplus \left\{\left(-G_j C \hat{X}_k^j\right) \oplus (-V)\right\}$$
$$= G_j C \left\{\{x_k\} \oplus \left(-\hat{X}_k^j\right)\right\} \oplus \{(G_i - G_j) C x_k\}$$
$$\oplus \{\eta_k\} \oplus (-V), \tag{5}$$

where, denoting by $\tilde{X}_k^{ij}$ the term $\{x_k\} \oplus (-\hat{X}_k^j)$, the residual zonotope can be rewritten as

$$R_k^{ij} = G_j C \tilde{X}_k^{ij} \oplus \{(G_i - G_j) C x_k\} \oplus \{\eta_k\} \oplus (-V). \tag{6}$$

Moreover, since $\tilde{X}_{k+1}^{ij}$ is a zonotope, $\tilde{x}_{k+1}^{ij,c}$ and $\tilde{H}_{k+1}^{ij,x}$ are used to denote its centre and segment matrix, respectively. Thus, by using (1) and (4), one can have

$$\tilde{x}_{k+1}^{ij,c} = x_{k+1} - \hat{x}_{k+1}^{j,c}, \tag{7a}$$

$$\tilde{H}_{k+1}^{ij,x} = \hat{H}_{k+1}^{j,x}, \tag{7b}$$

where, using (1), (4) and (7), $\tilde{x}_{k+1}^{ij,c}$ and $\tilde{H}_{k+1}^{ij,x}$ can be further derived as

$$\tilde{x}_{k+1}^{ij,c} = (A - L_j G_j C)\tilde{x}_k^{ij,c} + L_j (G_j - G_i) C x_k$$
$$- L_j (\eta_k - \eta^c) + \omega_k - \omega^c, \tag{8a}$$

$$\tilde{H}_{k+1}^{ij,x} = \hat{H}_{k+1}^{j,x} = \left[(A - L_j G_j C)\hat{H}_k^{j,x} \quad -L_j H_{\tilde{\eta}} \quad H_{\tilde{\omega}}\right]. \tag{8b}$$

3.2. Residual-bounding zonotopes

In (7), one cannot measure $\tilde{x}_k^{ij,c}$ since it involves unmeasurable quantities (η_k, ω_k and x_k). Thus, to precisely describe the bounds of residual zonotopes, one needs to consider residual-bounding zonotopes which are defined to contain the corresponding residual zonotopes. In order to obtain residual-bounding zonotopes, the following assumption is made for the plant inputs.

Assumption 3.1: *The input vector u_k of the plant is bounded by a known set $U = \{u \in \mathbb{R}^p : |u - u^c| \leq \bar{u}, u^c \in \mathbb{R}^p, \bar{u} \in \mathbb{R}^p\}$, where the vectors u^c and $\bar{u}$ are constant and known.*

Notice that, U can be rewritten as a zonotope $U = u^c \oplus H_{\bar{u}} \mathbb{B}^p$, where $H_{\bar{u}}$ is a $p \times p$ diagonal matrix with the main diagonal composed of $\bar{u}$. For construction of invariant sets, (1a) can be rewritten as

$$x_{k+1} = Ax_k + [B \ I] \begin{bmatrix} u_k \\ \omega_k \end{bmatrix}. \tag{9}$$

Considering $u_k \in U$ and $\omega_k \in W$, a *robust positively invariant* (RPI) set, denoted as X, can be constructed to confine the states of the dynamics (9) by using Theorem B.1 and Proposition B.1.

Remark 3.1: Any set $X \subset \mathbb{R}^n$ invariant with respect to the dynamics (9) is also invariant with respect to all modes (healthy or faulty). This statement holds since the dynamics (9) are not affected by these sensor faults directly or indirectly, as long as the inputs are bounded by the same set.

Since U and W are zonotopes, X is also a zonotope and is denoted as

$$X = x^c \oplus H_x \mathbb{B}^n,$$

where x^c and H_x are the centre and segment matrix, respectively. By substituting X, W and V into (8) to replace x_k, ω_k and η_k, respectively, a bounding zonotope (denoted as $\check{X}_{k+1}^{ij}$ with the centre $\check{x}_{k+1}^{ij,c}$ and segment matrix $\check{H}_{k+1}^{ij,x}$) to contain $\tilde{X}_{k+1}^{ij}$ can be characterised by its centre and segment matrix:

$$\check{x}_{k+1}^{ij,c} = (A - L_j G_j C)\check{x}_k^{ij,c} + L_j(G_j - G_i)Cx^c, \tag{10a}$$

$$\begin{aligned} \check{H}_{k+1}^{ij,x} = [(A - L_j G_j C)\check{H}_k^{j,x} \ \ L_j(G_j - G_i)CH_x \\ - L_j H_{\bar{\eta}} \ \ H_{\bar{\omega}} \ \ L_j H_{\bar{\eta}} \ \ -H_{\bar{\omega}}]. \end{aligned} \tag{10b}$$

Note that, comparing (8) with (10), it can be observed that (10) is the set-valued version of (8) by considering the bounds of states, disturbances and noises.

Remark 3.2: As per (8) and (10), if $\tilde{X}_{k^*}^{ij} \subseteq \check{X}_{k^*}^{ij}$ holds, $\tilde{X}_k^{ij}$ is always contained by $\check{X}_k^{ij}$ for all $k \geq k^*$.

Thus, as per (6) and Remark 3.2, one obtains a residual-bounding zonotope $\check{R}_k^{ij}$ to contain R_k^{ij}, i.e.,

$$\check{R}_k^{ij} = G_j C\check{X}_k^{ij} \oplus (G_i - G_j)CX \oplus V \oplus (-V). \tag{11}$$

For the centre-segment matrix description, an equivalent set description[1] of $\check{X}_{k+1}^{ij}$ is obtained as

$$\begin{aligned} \check{X}_{k+1}^{ij} = (A - L_j G_j C)\check{X}_k^{ij} \oplus L_j(G_j - G_i)CX \oplus L_j V \\ \oplus (-W) \oplus (-L_j V) \oplus W. \end{aligned} \tag{12}$$

In this paper, the use of residual-bounding zonotopes is twofold. First, they are used for establishing guaranteed FDI conditions. Second, they are used for the transient-state FI after a fault is detected.

4. Set-based FDI approaches

This section gives a brief introduction of the two set-based FDI approaches: interval observer-based and invariant set-based approaches.

4.1. Interval observer-based FDI

According to Guerra et al. (2008), the principle of the interval observer-based FD approach is summarised in Proposition 4.1.

Proposition 4.1: *During the operation of the ith mode, the interval observer-based FD approach consists in online checking whether or not*

$$\mathbf{0} \in R_k^{ii}, \tag{13}$$

where $\mathbf{0}$ is the origin and R_k^{ii} denotes the residual zonotope estimated by the ith interval observer in the ith mode at time instant k.

If the inclusion (13) is violated, it implies that a fault is detected at the current time instant. Otherwise, it is considered that the system still operates in the *i*th mode.

In Xu et al. (2013), an FI algorithm using a bank of interval observers has been proposed, where each interval observer is designed to match a considered mode. At each time instant, a collection of residual zonotopes can be obtained corresponding to all interval observers. With respect to Proposition 4.1, the FI algorithm in Xu et al. (2013) is summarised in Proposition 4.2.

Proposition 4.2: *During the operation of the ith mode, if a violation of* (13) *is detected, a waiting time is timed at the FD time. After the waiting time elapses, FI is implemented by searching an interval observer indexed by f that can estimate residual zonotopes containing the origin, i.e.,*

$$\mathbf{0} \in R_k^{ff}, \ f \in \mathcal{I} \setminus \{i\}. \tag{14}$$

Remark 4.1: The definition of the waiting time in Proposition 4.2 can be referred in Xu et al. (2013).

This paper proposes a new technique to implement sensor FDI by making full use of the system-operating information from all interval observers (3) and exploiting less conservative FDI conditions.

4.2. Invariant set-based FDI

It is known that a fault-affected system practically switches between several different modes including a healthy one

Table 1. Residual zonotopes.

	Interval observer 0	...	Interval observer i	...	Interval observer N
Mode 0	R_k^{00}	...	R_k^{0i}	...	R_k^{0N}
$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$
Mode i	R_k^{i0}	...	R_k^{ii}	...	R_k^{iN}
$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$
Mode N	R_k^{N0}	...	R_k^{Ni}	...	R_k^{NN}

and at least a faulty one. The principle of the invariant set-based FDI technique is to consider invariant sets characterising each of the system modes. For each mode, an RPI set characterising the residual can be constructed (Olaru et al., 2010; Seron et al., 2008; Stoican & Olaru, 2013). After the system enters into steady state, the residual must be confined by one of these invariant sets, and as long as all invariant sets are disjoint, FDI can be guaranteed.

The FDI technique proposed in this paper includes the invariant set-based FDI principle as a part. Different from the invariant set-based FDI approach, the proposed technique enhances FI performance by complementing its use with the system-operating information from all interval observers. Thus, theoretically, it is possible to loosen FI conditions and improve fault sensitivity and FDI reliability.

5. Guaranteed FDI conditions

This section proposes a novel set-based FDI strategy by combining both interval observer-based and invariant set-based mechanisms.

5.1. Collecting process information

At each time instant, the system can be monitored in terms of residual zonotopes by means of a bank of interval observers. For brevity, in the ith mode, one defines a vector[2] of residual zonotopes to collect residual zonotopes estimated by all interval observers at time instant k, i.e.,

$$\mathbf{R}_k^i = \left(R_k^{i0}, \ R_k^{i1}, \ \ldots, \ R_k^{iN} \right). \quad (15)$$

Remark 5.1: In this paper, the indices of rows and columns of tables and matrices start from 0. The index 0 corresponds to the healthy mode or the interval observer.

Furthermore, if considering residual zonotopes corresponding to all modes and interval observers, one can collect all available real-time system-operating information, which is presented in Table 1. Except the real-time process-operating information conveyed by residual zonotopes estimated by a bank of interval observers, there exists additional offline process information provided by the limit sets of all residual-bounding zonotopes. Table 2 collects the limit sets of all residual-bounding zonotopes, i.e., the smallest sets of residual zonotopes. Each row of Table 2 corresponds to a

mode. Thus, from Table 2, a matrix describing all modes can be extracted as

$$\mathbf{M} = \begin{bmatrix} \check{R}_\infty^{00} & \ldots & \check{R}_\infty^{0i} & \ldots & \check{R}_\infty^{0N} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \check{R}_\infty^{i0} & \ldots & \check{R}_\infty^{ii} & \ldots & \check{R}_\infty^{iN} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \check{R}_\infty^{N0} & \ldots & \check{R}_\infty^{Ni} & \ldots & \check{R}_\infty^{NN} \end{bmatrix}. \quad (16)$$

Remark 5.2: Comparing Table 1 with Table 2, it is known that, during the steady-state operation, each entry in Table 2 is the set of the corresponding entry in Table 1 such as $R_k^{i0} \subseteq \check{R}_\infty^{i0}$ when k is sufficiently large. Furthermore, one should have $\mathbf{R}_k^i \subseteq \mathbf{M}^i$, ($i \in \mathcal{I}$), where $\mathbf{M}^i$ is the ith row of $\mathbf{M}$ and $\subseteq$ should be understood elementwise.

In order to simplify Table 2, one defines the following rules. First, if $\mathbf{0} \in \check{R}_\infty^{ij}$, the position of $\check{R}_\infty^{ij}$ is labelled as 1. Second, if $\mathbf{0} \notin \check{R}_\infty^{ij}$, the position of $\check{R}_\infty^{ij}$ is labelled as 0. After applying these rules for Table 2, Table 3 is generated, containing binary information in concordance with the above logical propositions. Notice that, as per the interval observer-based FDI principle in Proposition 4.1, $\mathbf{0} \in R_k^{ii} \subseteq \check{R}_\infty^{ii}$ (for all $i \in \mathcal{I}$) should always hold during the steady-state operation of the ith mode. Similarly, from Table 3, a matrix describing all collected offline mode information can be obtained as

$$\mathbf{I} = \begin{bmatrix} 1 & \ldots & 1\backslash 0 & \ldots & 1\backslash 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1\backslash 0 & \ldots & 1 & \ldots & 1\backslash 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1\backslash 0 & \ldots & 1\backslash 0 & \ldots & 1 \end{bmatrix}. \quad (17)$$

5.2. FDI conditions

After collecting all available information as in (17), one should analyse how much information is useful and how much is available but redundant/unnecessary for FDI. It is mentioned that all diagonal entries of (17) are 1 because $\mathbf{0} \in R_k^{ii}$ and $R_k^{ii} \subseteq \check{R}_\infty^{ii}$ always hold during the steady-state operation of the ith mode as in Proposition 4.1. For the non-diagonal entries of (17), one does not know theoretically

Table 2. Limit sets of residual-bounding zonotopes.

	Interval observer 0	. . .	Interval observer i	. . .	Interval observer N
Mode 0	$\check{R}^{00}_\infty$	. . .	$\check{R}^{0i}_\infty$	. . .	$\check{R}^{0N}_\infty$
$\vdots$		. . .	$\vdots$	. . .	$\vdots$
Mode i	$\check{R}^{i0}_\infty$	. . .	$\check{R}^{ii}_\infty$	. . .	$\check{R}^{iN}_\infty$
$\vdots$		. . .	$\vdots$	. . .	$\vdots$
Mode N	$\check{R}^{N0}_\infty$	. . .	$\check{R}^{Ni}_\infty$	. . .	$\check{R}^{NN}_\infty$

whether or not they can contain **0** in advance. In this case, one should consider two possibilities:

- For the non-diagonal entries with 0, one has $\mathbf{0} \notin R^{ij}_\infty$ ($i \neq j$), because $R^{ij}_\infty \subseteq \check{R}^{ij}_\infty$ and $\mathbf{0} \notin \check{R}^{ij}_\infty$ imply $\mathbf{0} \notin R^{ij}_\infty$.
- For the non-diagonal entries with 1, one cannot have $\mathbf{0} \in R^{ij}_\infty$ ($i \neq j$), because $R^{ij}_\infty \subseteq \check{R}^{ij}_\infty$ and $\mathbf{0} \in \check{R}^{ij}_\infty$ do not guarantee $\mathbf{0} \in R^{ii}_\infty$.

Since the non-diagonal entries with 1 in the matrix **I** cannot guarantee that their corresponding residual zonotopes in Table 1 contain **0**, these entries are not useful for the proposed FDI approach and the residual-bounding zonotopes corresponding to these entries should be discarded in order to reduce computational complexity. Comparatively, the non-diagonal entries with 0 can guarantee that the residual zonotopes corresponding to them do not contain **0**, which are useful for the proposed FDI approach.

In this paper, the FDI approach is based on the combined use of residual zonotopes and residual-bounding zonotopes. Residual zonotopes estimated by interval observers can always be obtained in real time. Residual-bounding zonotopes are generated especially for the transient-state FI after a fault is detected. The limit sets of residual-bounding zonotopes (in Table 2) are mainly used for establishing guaranteed FDI conditions, which are used for offline pre-checking whether or not the considered faults are detectable and isolable in advance. For the sake of explaining how to establish these FDI conditions, one takes the following matrix **I** as an example, i.e.,

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}. \tag{18}$$

The example (18) considers four modes corresponding to the four rows of the matrix, respectively. Moreover, four interval observers should be designed to monitor these modes. According to the aforementioned analysis, the system information of the example (18) useful for FDI can be described by

$$\mathbf{I}' = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \times & 1 & 0 & \times \\ \times & \times & 1 & \times \\ \times & 0 & 0 & 1 \end{bmatrix}. \tag{19}$$

In this example, if one wants to guarantee that all the four modes are detectable and isolable after their occurrences, it should be assured that any two rows of the matrix (19) are distinguishable. Notice that when verifying the distinguishability of the four modes, one does not consider the entries with $\times$ of the matrix (19). Instead, only the entries with 0 or 1 are considered. Moreover, one cares about the columns (only columns without containing entries with $\times$) of any two rows, as long as there exists one column whose two entries are different (i.e., one is 0 and the other one is 1), it is guaranteed that the two modes corresponding to these two rows are distinguishable.

Similarly, taking (19) as an example, when considering Rows 0 and 1, one can see that Columns 1 and 2 of the two rows should be considered to verify whether or not the two rows are distinguishable. It can be observed that, in Column 1 of the two rows, the two entries corresponding to Rows 0 and 1 are 0 and 1, respectively, which are different. This means that, only by using these two entries of Column 1, Rows 0 and 1 (i.e., Modes 0 and 1) can be distinguished. Based on the same principle, for any other rows, one can make the similar analysis to know whether or not they are distinguishable.

Table 3. Transformation of Table 2.

	Interval observer 0	. . .	Interval observer i	. . .	Interval observer N
Mode 0	1	. . .	1\0	. . .	1\0
$\vdots$	$\vdots$	. . .	$\vdots$	. . .	$\vdots$
Mode i	1\0	. . .	1	. . .	1\0
$\vdots$	$\vdots$	. . .	$\vdots$	. . .	$\vdots$
Mode N	1\0	. . .	1\0	. . .	1

Remark 5.3: For simplicity, one takes the example (19) to show the principle of the FDI conditions. Thus, for real applications, one should obtain the corresponding matrices $\mathbf{I}$ and $\mathbf{I}'$ of these applications and use the method shown in this example to analyse whether or not all considered modes (healthy or faulty) are detectable and isolable.

Based on the example (19), the FDI conditions for the proposed FDI approach can be established in Proposition 5.1.

Proposition 5.1: *For all modes (healthy or faulty), a matrix $\mathbf{I}$ as in (18) corresponding to residual-bounding zonotopes can be obtained. Furthermore, based on the matrix $\mathbf{I}$, a simplified matrix $\mathbf{I}'$ as in (19) can be obtained. If for any two rows in the matrix $\mathbf{I}'$, there exists one column without $\times$ entries, whose two column entries are different from each other, the two modes corresponding to these two rows are detectable and isolable after their occurrences.*

Proof: For any two rows of the matrix $\mathbf{I}'$, as long as there exist one column of these two rows whose two column entries are different from each other, it means that the two modes corresponding to these two rows are distinguishable from each other. When the two entries of this column are different from each other (one is 0 and the other one is 1), it means that as the time elapses, the residual zonotopes estimated by the interval observer corresponding to this column will either always contain the origin or not, which will only accord with one entry of this column (i.e., only one row out of the two rows). Thus, the switching between these two modes is always detectable and isolable. Furthermore, if any two rows can satisfy these conditions, all modes are detectable and then isolable. □

Notice that, if there exists an FDI algorithm that can identify these differences of rows in the matrix $\mathbf{I}'$, it implies that Proposition 5.1 can be used as guaranteed FDI conditions of this FDI algorithm.

6. FDI Algorithm

This section proposes an algorithm implementing the sensor FDI based on the FDI conditions stated in Proposition 5.1. The three parts of this algorithm, namely the FD, FI and initialisation procedure, will be detailed and the entire procedure is summed up in Section 6.4.

6.1. *Fault detection*

To explain the proposed FD principle, it is assumed that the system is in the ith mode. Thus, a vector $\mathbf{R}_k$ (i.e., $\mathbf{R}_k^i$ in the ith mode) composed of residual zonotopes estimated by all interval observers can be obtained in real time. If the system operates at steady state, residual zonotopes should be bounded by the limit sets of residual-bounding zonotopes corresponding to the ith mode, i.e.,

$$\mathbf{R}_k \subseteq \mathbf{M}^i. \tag{20}$$

Thus, whenever (20) is violated elementwise, it implies that a fault[3] has occurred. Additionally, after the system enters into the operation of the ith mode, residual zonotopes can be used to test whether or not

$$\mathbf{0} \in \mathbf{R}_k, \tag{21}$$

where (21) should be understood elementwise. After testing (21) at each time instant, one can obtain an $(N + 1)$-dimensional fault signature vector $\mathcal{F}_i$ (i denotes that, before FD, the system is in the ith mode) full of 0 and 1, where, as in (17), 0 represents that residual zonotopes estimated by the interval observer corresponding to the vector entry with 0 do not contain the origin, while 1 has the opposite explanation. Thus, for FD, one should compare this real-time vector with the ith row of the matrix $\mathbf{I}'$ (the elements with $\times$ of this row should be omitted). If this real-time vector matches the ith row of the matrix $\mathbf{I}'$, it is considered that the system is still in the ith mode. Otherwise, it means that a fault has occurred.

Remark 6.1: A criterion simpler than (21) can also be used to detect faults in the ith mode, i.e., testing whether or not

$$\mathbf{0} \in R_k^{ii} \tag{22}$$

holds. However, under the FDI conditions in Proposition 5.1, (22) is only sufficient but not necessary for FD. Thus, if (22) is violated, it means that the system becomes faulty. Otherwise, one cannot assure whether or not a mode switching has occurred. This implies that (22) may only be able to detect a part of the considered modes. Instead, under the FDI conditions in Proposition 5.1, (21) can detect all the considered modes. Thus, at each time instant, for simplicity, one can first test the ith component of (21) (i.e., (22)) to make a quick FD decision. If (22) is violated, it implies a mode switching. Otherwise, then test the rest of components of (21) to confirm whether or not a mode switching has occurred.

Proposition 6.1: *If the system is at steady-state regime of its ith mode, FD can be performed by simultaneously testing (20) and (21). As long as either of them detects anomaly, it implies that a fault has occurred. Otherwise, it is considered that the system is still in the ith mode.*

Proof: At steady state of the ith mode, once residual zonotopes enter into their sets, they remain inside. Only when the system mode is changed by faults, a violation of (20) becomes possible. Regarding (21), the FD principle is also similar by testing the consistency as explained. Thus, as long as (20) or (21) is violated, it implies that a fault has occurred. □

It is known that testing (20) or (21) allows to implement FD. But based on the FDI conditions proposed in Proposition 5.1, it is not known whether the FD strategy (20) can guarantee that all the considered modes are detectable or not. But, the FD strategy (21) can guarantee that. If one chooses to test both of them online for FD, this combination may be more sensitive to some faults. But, the combination has high computational complexity. Thus, an alternative is that only the FD strategy (21) is used in the proposed FD approach, which has less computational complexity.

6.2. Fault isolation

In this paper, the proposed FI algorithm is based on residual zonotopes and residual-bounding zonotopes, which is also composed of two different FI strategies (transient-state and steady-state). In this subsection, these two FI strategies are presented, respectively.

In order to explain the FI algorithm, it is assumed that the system is at steady state of the ith mode and a fault is detected at time instant k_d, which implies that the system switches to a new mode different from the ith one after the FD time k_d. At the FD time k_d, the residual zonotope $R_{k_d}^{fj}$ ($f \in \mathcal{I}_i = \mathcal{I} \setminus \{i\}$) can always be obtained, where f denotes the index of a new and unknown mode.

As per Proposition B.1 in Appendix B, using a starting (or initial) set to initialise (12), a set sequence can be generated by iterating the dynamics. As k tends to infinity, the set sequence finally converges to a fixed set (i.e., the minimal robust positively invariant (mRPI) set $\check{X}_\infty^{ij}$ of $\tilde{X}_k^{ij}$ indicated in (8) for the jth interval observer in the ith mode).

Thus, at the FD time k_d, by initialising all the dynamics (12) corresponding to each interval observer ($j \in \mathcal{I}_i$) with their respective starting sets, a group of set sequences can be generated and their corresponding residual-bounding zonotope sequences can be simultaneously obtained by using (11). In this paper, a starting set for the jth interval observer in the fth mode is denoted as $\bar{X}_{k_d}^{fj}$ that should satisfy

$$\bar{X}_{k_d}^{fj} \supseteq \tilde{X}_{k_d}^{fj}.$$

Thus, a collection of starting sets should be constructed to initialise the corresponding set-based dynamics (12), each of which corresponds to one interval observer.

Remark 6.2: Although the residual-bounding zonotope sequences corresponding to all interval observers under all modes can be generated by using their corresponding starting sets at the FD time k_d, this FI approach only uses the N residual-bounding zonotope sequences $\check{R}_k^{jj}$ ($k \geq k_d$ and $j \in \mathcal{I}_i$) for less computational complexity.

By means of the dynamics (12), with the corresponding starting sets at the FD time k_d, the set sequences $\check{X}_k^{jj}$ ($k \geq k_d$ and $j \in \mathcal{I}_i$) can be generated, which will always contain the state estimation error sets $\tilde{X}_k^{jj}$ ($k \geq k_d$) ac-

cording to Remark 3.2 if the current mode is also the jth one. Furthermore, by using the expression (11), the corresponding residual-bounding zonotope sequences $\check{R}_k^{jj}$ can be computed, respectively. Similarly, the generated residual-bounding zonotope sequence $\check{R}_k^{jj}$ will always contain the residual zonotopes estimated by the jth interval observer if the current system is also in the jth mode. However, if the current system is not in the jth mode, the sequence $\check{R}_k^{jj}$ ($k \geq k_d$) may not contain the residual zonotopes R_k^{fj} even at the FD time k_d. This can be explained by the fact that residual-bounding zonotopes depend on the system modes (see (11)). At time instant k_d, even though the starting sets of (12) for all interval observers $j \in \mathcal{I}_i$, respectively, satisfy

$$\bar{X}_{k_d}^{fj} \supseteq \tilde{X}_{k_d}^{fj}, \tag{23}$$

it cannot still guarantee $\check{R}_{k_d}^{jj} \supseteq R_{k_d}^{fj}$, where $\check{R}_{k_d}^{jj}$ is computed as

$$\check{R}_{k_d}^{jj} = G_j C \check{X}_{k_d}^{jj} \oplus V \oplus (-V) \tag{24}$$

with $\check{X}_{k_d}^{jj} = \bar{X}_{k_d}^{fj}$ to initialise the corresponding dynamics (12) of $\check{X}_k^{jj}$.

Thus, whenever a fault is detected, the proposed transient FI strategy generates N residual-bounding zonotope sequences $\check{R}_k^{jj}$ ($k \geq k_d$ and $j \in \mathcal{I}_i$), each of which corresponds to one candidate mode. This implies that, among the N residual-bounding zonotope sequences, there exists at least one (i.e., the one matching the current after-fault mode) that can always contain the residual zonotopes estimated by its corresponding interval observer for all $k \geq k_d$. Based on this fact, the transient-state FI strategy is summarised in Proposition 6.2.

Proposition 6.2: *In the ith mode, when a fault is detected at the FD time k_d, N residual-bounding zonotope sequences described by the candidate-mode set $\mathcal{I}_i$ can be generated by initialising (12) with their corresponding starting sets and using (11), the transient-state FI strategy consists in searching a mode by testing whether or not*

$$R_k^{fj} \subseteq \check{R}_k^{jj}, \ \ j \in \mathcal{I}_i, \ \text{for all } k \geq k_d, \tag{25}$$

holds in real time. If a violation of (25) corresponding to the jth interval observer is detected, one removes the index j from the candidate-mode set $\mathcal{I}_i$ until the time instant when the set $\mathcal{I}_i$ remains only one element or the time window used to describe the transition completely elapses. The time instant and unique element indicate the FI time and fault, respectively.

Remark 6.3: Since sensor faults can immediately affect the system outputs, a sensor fault is possible to be directly isolated by this FI strategy with the initialisation at the FD time k_d (see (23) and (24)).

Remark 6.4: It is possible that, even though the proposed transient-state FI strategy is persistently executed, there always exist at least two elements in $\mathcal{I}_i$ during the corresponding time window. This implies that an accurate FI decision may not be obtained by this transient-state FI strategy during the transition. Thus, in order to describe the transition induced by faults and the persistent time of the use of this transient-state FI strategy, a proper time window should be defined in advance.

In order to avoid the situation indicated in Remark 6.4, one also proposes a steady-state FI strategy to complement the transient-state FI strategy proposed in Proposition 6.2 by using residual zonotopes and testing the inclusions between residual zonotopes and the origin after the time window. By testing these inclusions, one can obtain the fault signature vector $\mathcal{F}_i$. Finally, by matching $\mathcal{F}_i$ with the rows of $\mathbf{I}'$, if one row of $\mathbf{I}'$ can match $\mathcal{F}_i$, then the index of this row indicates the fault. This steady-state FI strategy is summarised in Proposition 6.3.

Proposition 6.3: *After applying the transient-state FI strategy over a defined time window, if there still exist at least two elements in $\mathcal{I}_i$, FI can still be guaranteed by searching an unique row in $\mathbf{I}'$ that can match the fault signature vector $\mathcal{F}_i$ and the FI decision is indicated by the index of this row.*

Proof: Under Proposition 5.1, as k tends to infinity, all residual-bounding zonotope sequences converge to their corresponding fixed sets (i.e., the elements of the matrix $\mathbf{M}$). Since if residual-bounding zonotope sequences do not contain $\mathbf{0}$, it is guaranteed that its corresponding residual zonotopes in Table 1 at steady state do not contain $\mathbf{0}$ too. Moreover, in the case that an interval observer matches the current system mode, the residual zonotopes estimated by this interval observer can contain $\mathbf{0}$ at steady state. This implies that, under Proposition 5.1, the fault signature vector is different in different modes and can match one and only one row of the matrix $\mathbf{I}'$, which can indicate the fault. $\square$

To explain how to obtain $\mathcal{F}_i$, one still uses the example given by (18) and (19). First, one assumes that the system is in the healthy mode at the beginning, after a fault is detected, one knows that all the faults 1, 2 and 3 are candidates. With the matrix (19), it can be observed that the faults 1 and 2 can be distinguished by using residual zonotopes estimated by the observer 2, the faults 1 and 3 by the observer 1 and the faults 2 and 3 also by the observer 2. This means that, after a fault is detected, one only needs to use residual zonotopes estimated by the interval observers 1 and 2, while the other two interval observers 0 and 3 are not necessary. From the computational complexity point of view, $\mathcal{F}_0$ should be a two-dimensional vector whose elements are composed of the binary information by testing the inclusions between residual zonotopes estimated by the interval observers 1 and 2 and the origin. But, for simplic-

ity, one can also define $\mathcal{F}_0$ as a four-dimensional vector, whose four elements are obtained by testing the inclusions between residual zonotopes estimated by the four interval observers and the origin. Finally, only the first and second elements of $\mathcal{F}_0$ are used to compare with the first and second elements of Rows 1, 2 and 3 of the matrix (19) for FI, while the zeroth and third elements of $\mathcal{F}_0$ are not useful.

Remark 6.5: The whole FI algorithm simultaneously includes these two FI strategies in Propositions 6.2 and 6.3. The former may be able to isolate faults during the transition induced by faults but without FI guarantees, while the latter with FI guarantees requires longer FI time.

6.3. Initialisation procedure: set construction

In Section 6.2, it can be observed that the starting sets for residual-bounding zonotope sequences are crucial for the transient-state FI strategy. Here, an idea to construct these starting sets is given. First, by using a zonotope $V_0 = H_{\tilde{\eta}}\mathbb{B}^q$, in the unknown mode f, for the jth interval observer, the corresponding residual zonotope indicated in (5) can be transformed into

$$R_{k_d}^{fj} = G_j C \tilde{X}_{k_d}^{fj} \oplus \{(G_f - G_j) C x_{k_d}\}$$
$$\oplus \{\eta_{k_d} - \eta^c\} \oplus (-V_0), \qquad (26)$$

By adding $-(\eta_{k_d} - \eta^c)$ and $-(G_f - G_j)C x_{k_d}$ to both sides, the previous equation turns into

$$R_{k_d}^{fj} \oplus \{-(\eta_{k_d} - \eta^c)\} \oplus \{(G_j - G_f)C x_{k_d}\}$$
$$= G_j C \tilde{X}_{k_d}^{fj} \oplus (-V_0). \qquad (27)$$

Considering $\eta_{k_d} - \eta^c \in V_0$ and $x_{k_d} \in X$, one further has

$$G_j C \tilde{X}_{k_d}^{fj} \oplus (-V_0) \subseteq R_{k_d}^{fj} \oplus (-V_0) \oplus \{(G_j - G_f)CX\}. \qquad (28)$$

Since V_0 can be removed from both sides of the previous equation, one can obtain

$$G_j C \tilde{X}_{k_d}^{fj} \subseteq R_{k_d}^{fj} \oplus \{(G_j - G_f)CX\}. \qquad (29)$$

It can be observed that the right side of (29) is dependent of modes. But, before FI, it is not possible to know the new mode. Thus, one should construct the starting sets without being affected by modes. Here, one should consider three cases:

- If $j \neq f$ and $j = 0$, one always has

$$G_j - G_f = \text{diag}(0, \ldots, 0, 1 - g_f, 0, \ldots, 0),$$

where g_f is the fth diagonal element of G_f which models the fth sensor fault.

- If $j \neq f$ and $j \neq 0$, one always has

$$G_j - G_f = \operatorname{diag}(0, \ldots, 0, g_j - 1, 0, \ldots,$$
$$0, 1 - g_f, 0, \ldots, 0),$$

where g_j is the jth diagonal element of G_j which models the jth sensor fault.
- If $j = f$, one always has

$$G_j - G_f = 0.$$

By summarising the aforementioned three cases, one can obtain that, if $j = 0$, one always has

$$G_j - G_f \in G_{fj} = \operatorname{diag}([0, 1 - g_1],$$
$$[0, 1 - g_2], \ldots, [0, 1 - g_q]),$$

while if $j \neq 0$, one always has

$$G_j - G_f \in G_{fj} = \operatorname{diag}([0, 1 - g_1], \ldots, [0, 1 - g_{j-1}],$$
$$g_j - 1, [0, 1 - g_{j+1}], \ldots, [0, 1 - g_q]),$$

where G_{fj} is a diagonal interval matrix that can always include $G_j - G_f$ inside its interval as long as the mode switching from the jth mode occurs. Thus, one can further have

$$(G_j - G_f)CX \subseteq G_{fj}CX, \tag{30}$$

where, considering that X is a zonotope and G_{fj} is an interval matrix, with Properties A.6 and A.7, $G_{fj}CX$ can be over-approximated by a zonotope denoted as

$$Z_{fj} = z^{fj,c} \oplus H_z^{fj} \mathbb{B}_z^{s^{fj}},$$

where s^{fj} is the order of Z_{fj}. Thus, in order to remove the effect of modes, a solution is to further transform (29) into

$$G_j C \tilde{X}_{k_d}^{fj} \subseteq R_{k_d}^{fj} \oplus Z_{fj}. \tag{31}$$

Since $R_{k_d}^{fj}$ is a zonotope, the term $R_{k_d}^{fj} \oplus Z_{fj}$ can be rewritten into the zonotopic form as

$$R_{k_d}^{fj} \oplus Z_{fj} = r_{k_d}^{fj,c} \oplus H_{k_d}^{fj,r} \mathbb{B}^{s_{k_d}^{fj,r}}, \tag{32}$$

where $s_{k_d}^{fj,r}$ is the corresponding zonotope order, and $r_{k_d}^{fj,c}$ and $H_{k_d}^{fj,r}$ can be, respectively, derived as

$$r_{k_d}^{fj,c} = y_{k_d} - \hat{y}_{k_d}^{j,c} + z^{fj,c}, \tag{33a}$$

$$H_{k_d}^{fj,r} = \begin{bmatrix} \hat{H}_{k_d}^{j,y} & H_z^{fj} \end{bmatrix}. \tag{33b}$$

With the help of the zonotopic form of $R_{k_d}^{fj} \oplus Z_{fj}$, (31) is rewritten as

$$G_j C \tilde{X}_{k_d}^{fj} \oplus \{-r_{k_d}^{fj,c}\} \subseteq H_{k_d}^{fj,r} \mathbb{B}_{k_d}^{s^{fj,r}}. \tag{34}$$

Using $F(l)$, $H_{k_d}^{fj,r}(l)$ and $r_{k_d}^{fj,c}(l)$ to denote the lth rows of $G_j C$ and $H_{k_d}^{fj,r}$ and the lth component of $r_{k_d}^{fj,c}$, respectively, one can obtain a group of inequalities corresponding to (34), where the lth inequality can be written as

$$\mid F(l)\tilde{x}_k^{fj} - r_{k_d}^{fj,c}(l) \mid \leq \parallel H_{k_d}^{fj,r}(l) \parallel_1, \ l = 1, 2, \ldots, q, \tag{35}$$

where $\tilde{x}_k^{fj}$ represents the elements that satisfy (35).

Remark 6.6: It can be observed that the description (35) is more conservative than (34), which means, if the q inequalities of (35) can determine a closed set, the closed set can fully contain $\tilde{X}_{k_d}^{fj}$ in (34).

Assumption 6.1: *For the jth interval observer, $\tilde{X}^j$ denotes a given zonotope determined by the physical constraints of the system and can always bound $\tilde{X}_k^{ij}$ for all $i \in \mathcal{I}$ and $k \geq 0$.*

As per Property A.4 in Appendix A, each inequality out of the q inequalities of (35) generally determines a strip and the q inequalities together can form a closed set that contains $\tilde{X}_{k_d}^{fj}$. However, there exist two possible cases that depend on the system dynamics:

- If the q inequalities themselves can form a closed set, then the formed set can contain $\tilde{X}_{k_d}^{fj}$.
- If the q inequalities cannot form a closed set (i.e., there are not enough strips such that their intersection can lead to a closed set), then $\tilde{X}^j$ indicated in Assumption 6.1 can be further used to construct a closed set that can contain $\tilde{X}_{k_d}^{fj}$.

However, because Property A.4 can only compute a zonotope approximation of the intersection of a zonotope and a strip, for the jth interval observer, $\tilde{X}^j$ has to be used as the initial zonotope of Property A.4 for both cases. As seen in Section 6.2, for the jth interval observer, the constructed starting zonotope for the residual-bounding zonotope sequence is denoted as $\bar{X}_{k_d}^{fj}$.

Remark 6.7: In a particular case, if the matrix $G_j C$ is invertible, (31) can be directly transformed into

$$\tilde{X}_{k_d}^{fj} \subseteq (G_j C)^{-1}\left(R_{k_d}^{fj} \oplus Z_{fj}\right), \tag{36}$$

where $(G_j C)^{-1}$ denotes the inverse of $G_j C$, and $(G_j C)^{-1}(R_{k_d}^{fj} \oplus Z_{fj})$ is directly used as $\bar{X}_{k_d}^{fj}$.

At the FD time k_d, the proposed FI approach constructs a group of starting sets for generating N corresponding residual-bounding set sequences $\check{R}_k^{jj}$ ($j \in \mathcal{I}_i$). As derived before, for the jth interval observer, its residual-bounding set sequence is generated by initialising its corresponding dynamics (12) with the starting set

$$\check{X}_{k_d}^{jj} = \bar{X}_{k_d}^{fj}, \ j \in \mathcal{I}_i. \qquad (37)$$

Remark 6.8: In order to construct the starting sets at the FD time k_d, this subsection proposes a method by using Property A.4 or Property A.4 with $\tilde{X}^j$ in Assumption 6.1. However, one can still introduce another method based on Property A.5 to construct the starting sets, where Property A.5 computes a zonotopic approximation of the intersection of a polytope and a zonotope. Using Property A.5, one can use all q strips (35) as a whole (i.e., a polytope) and compute the zonotopic approximation of the intersection of this polytope and a zonotope to construct the starting zonotopes.

6.4. Practical FDI algorithm

According to the aforementioned discussions, based on the FDI conditions in Proposition 5.1, faults can be detected with Proposition 6.1. Whenever a fault is detected, after applying the proposed FI strategies in Propositions 6.2 and 6.3, theoretically, the worst case is that the fault can only be isolated at infinity because the FDI conditions are built by means of the limit sets of residual-bounding zonotopes. However, from the practical point of view, it is impossible to obtain residual-bounding zonotopes at infinity. Thus, in order to establish the FDI conditions and implement the FDI approach, one has to consider the approximations of the limit sets of residual-bounding zonotopes.

According to Theorem B.1 in Appendix B, for the set-based dynamics (12), the mRPI set $\check{X}_\infty^{ij}$ can be approximated by an RPI set denoted as $\mathring{X}^{ij}$ with an arbitrarily approximate precision to $\check{X}_\infty^{ij}$. Thus, the set $\check{R}_\infty^{ij}$ can be approximated by the corresponding set

$$\mathring{R}^{ij} = G_j C \mathring{X}^{ij} \oplus (G_i - G_j) CX \oplus V \oplus (-V). \quad (38)$$

Notice that whenever $\tilde{X}_k^{ij}$ goes into and stays inside $\mathring{X}^{ij}$, R_k^{ij} also goes into and stays inside $\mathring{R}^{ij}$. Since all sets (i.e., $\mathring{R}^{ij}$) can be computed offline, all entries of Table 2 are over-approximated by Table 4.

According to the definition of the mRPI set in Definition B.2 in Appendix B, $\check{X}_\infty^{ij} \subseteq \mathring{X}^{ij}$ and $\check{R}_\infty^{ij} \subseteq \mathring{R}^{ij}$ hold. This implies that after initialising (12), the set sequence of the corresponding residual-bounding zonotopes will finally converge to its corresponding set (i.e., $\mathring{R}^{ij}$) as k increases.

Thus, as long as the given approximate precision is sufficiently high, Table 4 can be used to replace Table 2 for

---

Algorithm 1 FD algorithm

Require: $\hat{X}_0$, $\mathcal{I}$ and current mode index $i \in \mathcal{I}$;
Ensure: f;
 1: Interval observer initialisation: $\hat{X}_0^{ij} = \hat{X}_0$ (for all $j \in \mathcal{I}$);
 2: At time instant k: no fault alarm and $f \leftarrow$ FAULT;
 3: **While** $f \neq$ TRUE do
 4:     $k \leftarrow k + 1$;
 5:     Obtain $\mathbf{R}_k$;
 6:     **if** (20) or (21) makes a fault alarm **then**
 7:        $f \leftarrow$ TRUE;
 8:     **end if**
 9: **end while**
10: **return** f;

---

verifying the FDI conditions in Proposition 5.1. Moreover, Table 3 can be derived from Table 4. Furthermore, based on Table 4, for the ith mode, one defines a vector

$$\mathbf{\mathring{R}}^i = (\mathring{R}^{i0}, \ \mathring{R}^{i1}, \ \ldots, \ \mathring{R}^{iN}), \qquad (39)$$

which is used for Algorithm 1 to carry out the proposed FD strategy. The proposed FI algorithm is a combination of the two FI strategies in Propositions 6.2 and 6.3. Practically, once a fault is detected, this FI algorithm first starts up the FI strategy in Proposition 6.2. If the FI strategy in Proposition 6.2 cannot isolate the fault within a defined time window, the FI algorithm will terminate it and then the FI strategy in Proposition 6.3 is started to guarantee FI at steady state. Thus, in the proposed FI algorithm, the defined time window for the transient-state FI strategy is used as a switching mechanism between the two FI strategies in Propositions 6.2 and 6.3.

According to (12), after the initialisation required to obtain residual-bounding zonotope sequences online, the differences between the set values of residual-bounding zonotopes at different time instants are dependent of the term $(A - L_j G_j C) \check{X}_k^{ij}$, i.e., the eigenvalues of $A - L_j G_j C$ and the starting set of the initialisation. Because all eigenvalues of $A - L_j G_j C$ are inside the unit circle (i.e., the dynamics (12) of residual-bounding zonotopes are stable), after the transition induced by a fault, all residual-bounding zonotope sequences finally enter into steady state and can sufficiently and asymptotically approximate their corresponding limit sets (i.e., $\check{R}_\infty^{ij}$). This implies that, at steady state, the difference between the set values of a residual-bounding zonotope sequence at different time instants will gradually decrease. This fact allows to define a proper time window mentioned in Remark 6.4 as the switching mechanism between the two FI strategies.

Definition 6.1: The time window T starting from the FD time instant k_d is defined at least as the maximal settling time of the dynamics of all interval observers such that

Table 4. Sets of residual zonotopes.

	Interval observer 0	...	Interval observer i	...	Interval observer N
Mode 0	$\mathring{R}^{00}$	...	$\mathring{R}^{0i}$	...	$\mathring{R}^{0N}$
$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$
Mode i	$\mathring{R}^{i0}$	...	$\mathring{R}^{ii}$	...	$\mathring{R}^{iN}$
$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$
Mode N	$\mathring{R}^{N0}$	...	$\mathring{R}^{Ni}$	...	$\mathring{R}^{NN}$

the proposed FI strategy in Proposition 6.3 can guarantee reliable FI.

Remark 6.9: The transitions between different modes are determined by the eigenvalues of the corresponding dynamics. Theoretically, it can also be assessed by the settling time of interval observers. Most importantly, a proper time window can always be selected by a sufficient number of simulations.

As per Definition 6.1 and Remark 6.9, a mechanism of switching between the proposed FI strategies in Propositions 6.2 and 6.3 is introduced by the time window T and the proposed FI algorithm is summarised as follows:

(1) Once a fault is detected by Proposition 6.1, the transient-state FI strategy in Proposition 6.2 is first started to isolate the fault within the time window T.
(2) After T, if the fault is still not isolated, then the first FI strategy is terminated and the FI algorithm starts up the second FI strategy in Proposition 6.3 for FI at steady state.
(3) After the second FI strategy enters in operation, at each time instant, the inclusions between residual zonotopes and $\mathbf{0}$ are tested, and the testing results are used to compare with the offline inclusion information contained in each row of the matrix $\mathbf{I}'$.
(4) If at a time instant, the obtained results (i.e., $\mathcal{F}_i$) match one row of the matrix $\mathbf{I}'$, the index of this row indicates the fault (this index is assumed as f).
(5) In order to improve reliability of the FI decision given in Step 4, $\mathbf{R}_k \subseteq \mathring{\mathbf{R}}^f$ is also tested. If $\mathbf{R}_k \subseteq \mathring{\mathbf{R}}^f$ holds, the FI decision of Step 4 can be confirmed. Otherwise, the second FI strategy in Proposition 6.3 is repeated again.

Remark 6.10: Both Steps 4 and 5, respectively, considering the invariant set-based and the interval observer-based FI principles, are used to make FI decisions. However, the core step is Step 4, while Step 5 is an assistant step. For simplicity, one can omit Step 5 directly and only use Step 4.

Notice that, the proposed FDI approach is based on the combination of invariant sets and interval observers. By

this combination, the conservatism of the FDI conditions should be reduced in some sense because of the use of all available system-operating information from all the corresponding interval observers. Comparing with the purely interval observer-based or the purely invariant set-based approach, it can be observed that the FDI effectiveness and reliability are possible to be improved.

7. Illustrative example

In this paper, the second pitch system of a wind turbine benchmark proposed in Odgaard et al. (2009) is used as the case study. The continuous-time dynamics of this subsystem are given as

$$\dot{x}(t) = Ax(t) + Bu(t), \tag{40a}$$

$$y(t) = Cx(t). \tag{40b}$$

In this subsystem, two sensors are used to measure the pitch position $y(t)$, whose measurements are

$$y_1(t) = G_1(Cx(t) + \eta_1(t)), \tag{41a}$$

$$y_2(t) = G_2(Cx(t) + \eta_2(t)), \tag{41b}$$

where $y_1(t)$ and $y_2(t)$ denote the measurements of the first and second sensors, respectively, $\eta_1(t)$ and $\eta_2(t)$ are the corresponding measurement noises, G_1 and G_2 model the fault in the first and second sensors, respectively. If G_1 (or G_2) is the identity matrix, it means that the corresponding sensor is healthy. Otherwise, it implies that the corresponding sensor is faulty. Besides, the control action of the subsystem is designed as

$$u(t) = u^{\text{ref}}(t) + u^f(t), \tag{42}$$

where $u^{ref}(t)$ is the given reference input and $u^f(t)$ is the feedback signal[4] with a form $u^f(t) = y(t) - 0.5(y_1(t) + y_2(t))$. Furthermore, one can equivalently reformulate (40) and (41) of the pitch system into a compact form, which entirely describes the pitch system, i.e.,

$$\dot{x}(t) = Ax(t) + Bu(t), \tag{43a}$$

$$y_{12}(t) = G_{12}Cx(t) + \eta(t), \tag{43b}$$

where

$$y_{12}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}, \ G_{12} = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix},$$

$$\eta(t) = \begin{bmatrix} G_1 & 0 \\ 0 & G_2 \end{bmatrix} \begin{bmatrix} \eta_1(t) \\ \eta_2(t) \end{bmatrix}.$$

The sets of the noises of the sensors are described by $\bar{\eta}_1 = 0.8$, $\eta_1{}^c = 0$, $\bar{\eta}_2 = 0.8$ and $\eta_2{}^c = 0$, which follow the form as in (2). The parameters of the second pitch system are given as

$$A = \begin{bmatrix} -13.33 & -123.43 \\ 1 & 0 \end{bmatrix}, \ B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ C = \begin{bmatrix} 0 & 123.43 \end{bmatrix}.$$

Remark 7.1: Since the noises are Gaussian, the aforementioned bounds are empirical values according to the fact that usual choices for Gaussian distribution are the band $[-3\sigma, 3\sigma]$ with probability of 99%, or band $[-6\sigma, 6\sigma]$ with probability of 99.99%.

The sampling time of the pitch system is 0.01s. After discretisation, the system parameters are

$$A_d = \begin{bmatrix} 0.867 & -1.234 \\ 0.01 & 1 \end{bmatrix}, \ B_d = \begin{bmatrix} 0.01 \\ 0 \end{bmatrix}, \ C_d = \begin{bmatrix} 0 & 123.43 \end{bmatrix}.$$

In this case study, one considers two faults in the two sensors, respectively, and it is assumed that one and only one sensor becomes faulty at any given time (note that one can also consider that two sensors become faulty simultaneously). Thus, the system should have three different sensor modes: healthy, fault in the first sensor and fault in the second sensor, which are, respectively, defined as

$$G_{12}^0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ G_{12}^1 = \begin{bmatrix} 0.1 \\ 1 \end{bmatrix}, \ G_{12}^2 = \begin{bmatrix} 1 \\ 0.1 \end{bmatrix}.$$

In this paper, one assumes that the reference input[5] of the pitch system varies in the operating range $u^{ref}(t) \in [10°, 30°]$. Furthermore, as per (42), by simulating the pitch system with a time span of 10^6s, an empirical bound of $u(t)$ is obtained as $u(t) \in [8.561, 52.2314]$. Thus, as per the discretised model, Theorem B.1 and Proposition B.1 in Appendix B, an RPI approximation of the mRPI set of states is constructed by iterating 150 times from an initial-state RPI set, which is denoted as

$$X = \begin{bmatrix} 0 \\ 0.2463 \end{bmatrix} \oplus \begin{bmatrix} 2.3934 & 0 \\ 0 & 0.2252 \end{bmatrix} \mathbb{B}^2.$$

Table 5. Available offline system information.

	Interval observer 0	Interval observer 1	Interval observer 2
Mode 0	1	0	0
Mode 1	1	1	0
Mode 2	1	0	1

Based on the discrete-time model from (43), three interval observers indicated as in (3), a healthy interval observer and two faulty interval observers, are, respectively, designed to monitor the system. The initial state of the system and the initial zonotope of interval observers are, respectively, given as

$$x_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \ \hat{X}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \oplus \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbb{B}^2.$$

In order to obtain FDI guarantees, one checks the FDI conditions in Proposition 5.1. As per Theorem B.1 and Proposition B.1 in Appendix B, for each mode, the approximations of the limit sets (i.e., $\check{R}_\infty^{ij}$) of the relevant residual-bounding zonotopes, indicated in Table 4, are obtained by iterating the dynamics (12) 120 times with initial RPI sets of $\tilde{X}_k^{ij}$. These approximations are presented as follows:

- For the healthy mode,

$$\mathring{R}^{00} = ([-3.6507, \ 3.6507], [-3.6507, \ 3.6507])^T,$$
$$\mathring{R}^{01} = ([0.7196, \ 53.5725], [-9.0443, \ 4.7407])^T,$$
$$\mathring{R}^{02} = ([-9.0443, \ 4.7407], [0.7196, \ 53.5725])^T.$$

- For the fault in the first sensor,

$$\mathring{R}^{10} = ([-59.8823, \ 26.8929], [-8.2239, \ 29.9569])^T,$$
$$\mathring{R}^{11} = ([-0.3651, \ 0.3651], [-3.6507, \ 3.6507])^T,$$
$$\mathring{R}^{12} = ([-63.4498, \ 47.4594], [0.4448, \ 58.1507])^T.$$

- For the fault in the second sensor,

$$\mathring{R}^{20} = ([-8.2239, \ 29.9569], [-59.8823, \ 26.8929])^T,$$
$$\mathring{R}^{21} = ([0.4448, \ 58.1507], [-63.4498, \ 47.4594])^T,$$
$$\mathring{R}^{22} = ([-3.6507, \ 3.6507], [-0.3651, \ 0.3651])^T.$$

By analysing these approximations, one can obtain Table 5, which collects the offline system-operating information corresponding to all modes and interval observers. Furthermore, based on Table 5, one can obtain the matrix $\mathbf{I}'$, which is shown as

$$\mathbf{I}' = \begin{bmatrix} 1 & 0 & 0 \\ \times & \mathbf{1} & 0 \\ \times & \mathbf{0} & \mathbf{1} \end{bmatrix}. \tag{44}$$

It can be checked that the matrix $\mathbf{I}'$ in (44) satisfies the FDI conditions in Proposition 5.1. As observed in (44), all the non-diagonal entries with 1 are omitted. After omitting these entries, it is observed that the three rows of the matrix $\mathbf{I}'$ are different from each other, which means that the three modes are distinguishable from each other. Thus, if a considered mode has occurred, it can be guaranteed that the occurrence of the mode can be detected and then isolated by the FDI approach.

With loss of generality, this example only considers the dynamical process of the system from healthy to faulty. Thus, it is not necessary to take Row 0 corresponding to the healthy mode into account in terms of FI of the two sensor faults. According to the FD strategy in Proposition 6.1, at most, R_k^0, R_k^1, R_k^2, $\mathring{R}^{00}$, $\mathring{R}^{01}$ and $\mathring{R}^{02}$ are used for FD implementation. Moreover, according to the proposed FI strategies in Propositions 6.2 and 6.3, one only needs to obtain R_k^1 or R_k^2 to distinguish the two sensor-fault modes and to test the inclusions between the origin and them, respectively. Eventually, FI can be done by comparing the tested inclusion results with the rows of $\mathbf{I}'$, and a matched row of $\mathbf{I}'$ will indicate the fault.

According to the proposed FI approach in Proposition 6.2, by initialising (12), one can obtain the corresponding residual-bounding zonotope sequences $\check{R}_k^{11}$ and $\check{R}_k^{22}$ for isolating the faults 1 and 2 during the transient-state operation. However, one should notice that the transient-state FI strategy in Proposition 6.2 cannot be guaranteed by the FDI conditions in Proposition 5.1. This implies that the transient-state FI strategy may be able to isolate faults during the transition or not, which depends on the faults themselves. Besides, for the two considered scenarios, the set for the transient-state FI strategy in Assumption 6.1 is empirically given by simulations as

$$\tilde{X}^0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \oplus \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \mathbb{B}^2.$$

The time window in Definition 6.1 is given as five steps, and in order to compute the starting zonotopes for the initialisation for residual-bounding zonotope sequences, λ in Property A.4 in Appendix A is given as $\lambda = \begin{bmatrix} 0.005 & 0.005 \end{bmatrix}^T$.

Remark 7.2: The selection of λ is important for the proposed FI approach, which affects the volume of the starting zonotopes obtained by Property A.4. Please refer to Alamo et al. (2005) for details.

In this example, one considers the same scenario for both sensor faults: from time instants 1–40, the system is healthy, then a fault occurs at time instant 41 and the system is in the faulty operation from time instants 41–80. Thus, the simulations for the two sensor faults are done, respectively, and the diagnostic results of the faults in the first and second sensors are shown in Figures 2–5.
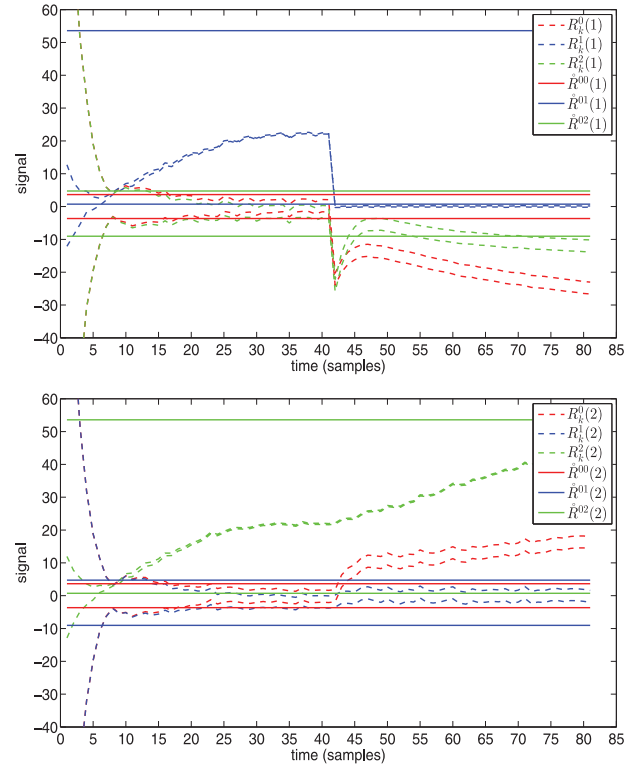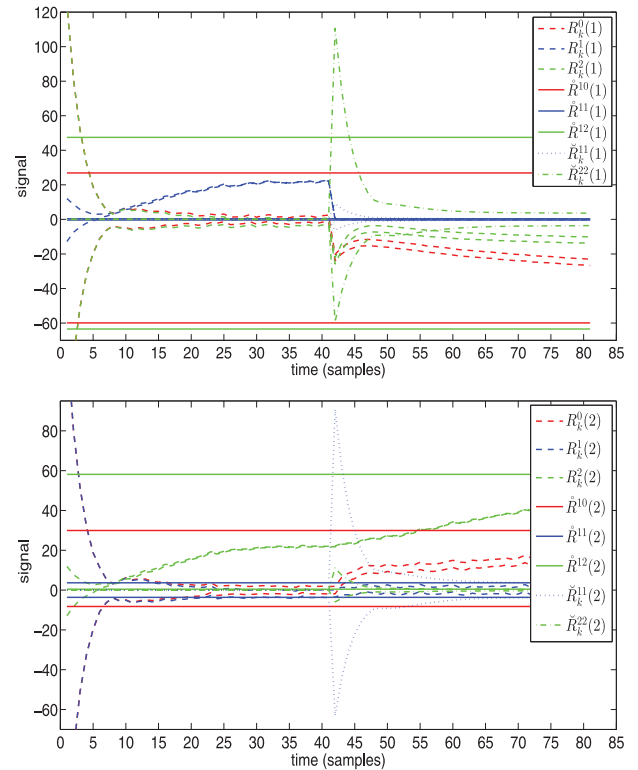

Figure 2.   FD of fault 1.
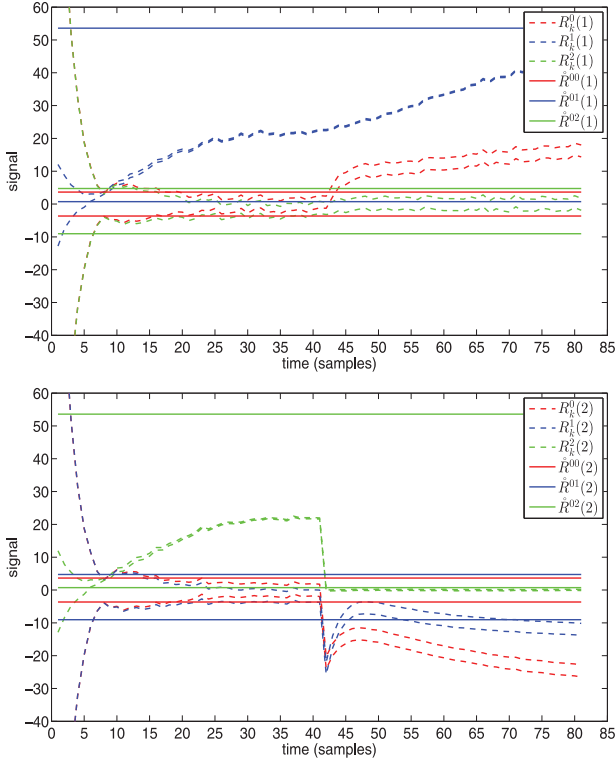

Figure 3.   FI of fault 1.
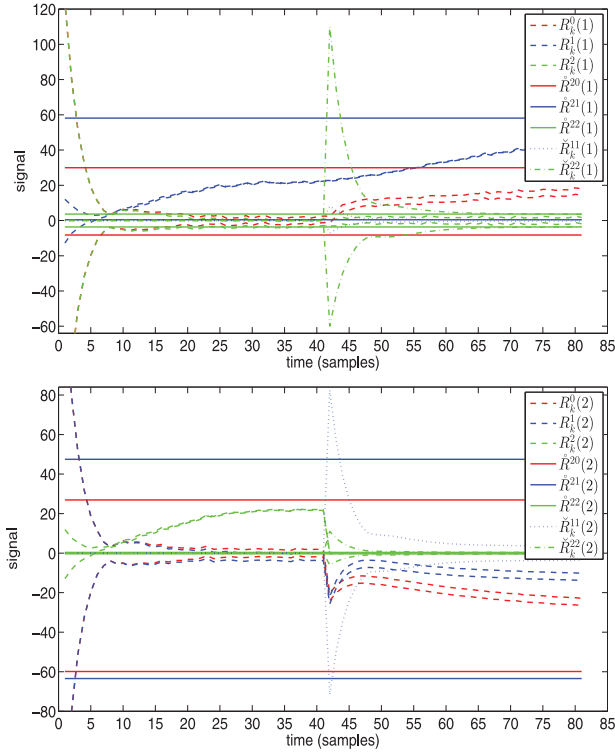
Figure 4. FD of fault 2.

Figure 5. FI of fault 2.

Remark 7.3: In the figures, $R_k^i(1)$ and $R_k^i(2)$ denote the first and second components of the residual zonotope R_k^i estimated by the ith interval observer at time instant k. Similarly, $\mathring{R}^{ij}(1)$ and $\mathring{R}^{ij}(2)$ denote the first and second components of the approximation $\mathring{R}^{ij}$ of R_∞^{ij} corresponding to the jth interval observer in the ith mode.

According to the FD principle in Proposition 6.1 and the simulation results shown in Figure 2, it can be observed that a fault is detected at time instant 42 (i.e, $\mathbf{0} \notin R_{42}^0$, $R_{42}^0 \not\subseteq \mathring{R}^{00}$, $R_{42}^1 \not\subseteq \mathring{R}^{01}$ and $R_{42}^2 \not\subseteq \mathring{R}^{02}$). Furthermore, according to the proposed transient-state FI approach in Proposition 6.2 and the simulation results shown in Figure 3, the first fault is isolated at time instant 42 because of $R_{42}^1 \subseteq \check{R}_{42}^{11}$ and $R_{42}^2 \not\subseteq \check{R}_{42}^{22}$. This indicates that the first fault is isolated at the same time when it is detected, which means no time delay between FD and FI (see Remark 6.3). Additionally, in order to show the steady-state FI strategy proposed in Proposition 6.3, one should wait a defined five-step time window. Thus, one should test whether or not $\mathbf{0} \in R_{47}^1$ (or/and $\mathbf{0} \in R_{47}^2$) holds. As shown in Figure 3, it is seen that $\mathbf{0} \in R_{47}^1$ (or/and $\mathbf{0} \notin R_{47}^2$) holds, which matches Row 1 of the matrix $\mathbf{I}'$. This implies that the fault is in the first sensor, which provides the same FI decision with the transient-state FI strategy.

Remark 7.4: The size of $\bar{X}_{k_d}^{fj}$ affects the quickness of the transient-state FI strategy in Proposition 6.2, but the effect is not decisive. Because, even though the transient-state FI strategy cannot isolate faults within the time window, the proposed FI algorithm can still use the steady-state FI strategy in Proposition 6.3 to guarantee to isolate faults after the time window.

Similarly, FDI of the second fault is illustrated in Figures 4 and 5. In Figure 4, it is seen that a fault is detected at time instant 42 (i.e., $\mathbf{0} \notin R_{42}^0$, $R_{42}^0 \not\subseteq \mathring{R}^{00}$, $R_{42}^1 \not\subseteq \mathring{R}^{01}$ and $R_{42}^2 \not\subseteq \mathring{R}^{02}$). In Figure 5, it can be observed that $R_{42}^1 \not\subseteq \check{R}_{42}^{11}$ and $R_{42}^2 \subseteq \check{R}_{42}^{22}$ hold, which means that the second sensor fault has occurred by the transient-state FI strategy. Moreover, if considering the steady-state FI strategy in Proposition 6.3, one can observe that $\mathbf{0} \notin R_{47}^{11}$ (or/and $\mathbf{0} \in R_{47}^{21}$) holds, which is in accordance with Row 2 of the matrix $\mathbf{I}'$ in (44), which means that at time instant 47, the fault can also be isolated by the steady-state FI strategy.

Remark 7.5: According to the proposed FI algorithm, after a fault is detected, the fault should be either in the first or second sensor. Thus, as seen in the matrix $\mathbf{I}'$, one only needs to use the residual zonotopes estimated by either the interval observer 1 or 2 (or both of these two interval observers), which corresponds to the bold and italic columns of Rows 1 and 2 of the matrix $\mathbf{I}'$, respectively. This can reduce computational complexity of FI.

8. Conclusion

This paper proposes a sensor FDI approach based on set-theoretic approaches. In this approach, two different set-theoretic FDI mechanisms are simultaneously used, i.e., invariant set-based and interval observer-based mechanisms. This approach implements the combination of invariant sets and interval observers for sensor FDI, and both FD and FI decisions are based on the results of the two FDI mechanisms. A bank of interval observers is used to monitor the system operation in real time and collect system-operating information for FDI, and invariant sets describing the steady-state behaviours of the system are used to establish guaranteed FDI conditions. Furthermore, in order to reduce the computational complexity, the available but redundant system-operating information is discarded by a new strategy. By means of these strategies, the objective is to use as much available information from all interval observers as possible to improve fault sensitivity but simultaneously discard all redundant information to simplify the FDI algorithm. Besides, another objective of the proposed approach is to isolate faults as soon as possible. For sensor faults, this approach can isolate them during the transition induced by faults and the fastest FI case is that sensor faults are isolated at the same time instant when they are detected. Due to more system-operating information used, this approach should have less conservative guaranteed FDI conditions comparing with the existing invariant set-based or interval observer-based approaches. Additionally, for simplicity, although this paper only considers the linear discrete time-invariant plant and the given magnitudes of sensor faults, it could be extended to the system with parametric uncertainties and unknown but bounded faults in principle.

Notes

1. The equivalence of (10) and (12) can be verified by applying zonotope operations into (12) to obtain its centre-segment matrix equivalent form, which is the same as (10).
2. Because, in any mode, residual zonotopes are obtainable, without ambiguity, the notation $\mathbf{R}_k^i$ corresponding to the ith mode can be generally replaced by the notation $\mathbf{R}_k$.
3. Notice that, for the proposed FDI approach, if it is used for detecting recovery process, then one still uses the terms such as *fault detection and isolation* and *fault occurrence* to describe the general mode switching.
4. The form of this feedback signal completely respects the original structure of the second pitch system in the wind turbine benchmark proposed in Odgaard et al. (2009).
5. In this case study, as an example, the reference input $u^{ref}(t)$ is chosen as a sinusoidal signal.

Notes on contributors

Feng Xu was born on 4 January 1988 (or 15 November 1987 in the lunar calender) in Tianmen, Hubei, P.R. China. He received his BSc degree in measurement and control technology & instruments from the Northwestern Polytechnical University (NWPU), Xi'an, Shaanxi, P.R.China, in July 2010. In November 2014, he received his Ph.D. degree in automatic control from the Institut de Robótica i Informática Industrial (CSIC-UPC), Technical University of Catalonia (UPC), Barcelona, Spain. His research interests are fault diagnosis and fault-tolerant control.

Vicenç Puig was born in Girona, Spain, on 6 November 1969. He received his Ph.D. degree in control engineering in 1999 and his Telecommunications Engineering degree in 1993, both from the Technical University of Catalonia (UPC), Barcelona, Spain. He is currently an associate professor of automatic control and leader of the Advanced Control Systems (SAC) research group of the Research Center for Supervision, Safety and Automatic Control (CS2AC) at UPC. His main research interests are fault detection and isolation of fault-tolerant control of dynamic systems. He has remained involved in several European projects and networks and has published many papers in international conference proceedings and scientific journals.
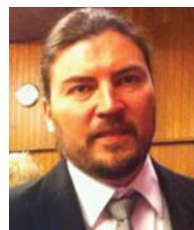
Carlos Ocampo-Martinez received his Electronic Engineering degree and MSc degree in industrial automation from the National University of Colombia, Campus Manizales, in 2001 and 2003, respectively. In 2007, he received his Ph.D. degree in control engineering from the Technical University of Catalonia (UPC), Barcelona, Spain. After an year as a research fellow of the ARC Centre of Complex Dynamic Systems and Control (University of Newcastle, Australia), he remained associated with Spanish National Research Council (CSIC) at the Institut de Robótica i Informàtica Industrial (IRI) in Barcelona as a Juan de la Cierva associate researcher. Since 2011 he has been working as a assistant professor at the Technical University of Catalonia, Automatic Control Department (ESAII). His main research interests are in the areas of constrained MPC, large-scale systems management, nonlinear dynamics, and industrial applications.

Sorin Olaru received his Ph.D. degrees from the University Paris XI, France, and the University Politehnica in Bucharest, Romania, and holds the French Research Habilitation. Since 2001 he has held various positions at INRIA and SUPELEC in France and has been visiting appointments at the University of Newcastle, Australia, and NTNU Trondheim, Norway. Currently

he is a professor in SUPELEC and a member of the INRIA Disco team. His research interests include optimization-based control design and set theoretic characterization of constrained dynamical systems.

Florin Stoican is an assistant professor at the University Politehnica Bucharest (UPB), Department of Automatic Control and Systems Engineering. He received his BSc degree in electrical in 2008 from the Faculty of Automatic Control and Computers, UPB. He obtained his Ph.D. in control engineering in 2011 from the Department of Automatic Control, SUPELEC, France, and was a postdoctoral research fellow in the Department of Engineering Cybernetics, NTNU, Trondheim, Norway, with an ERCIM Alain Bensoussan Fellowship Programme. His interests are in constrained optimization, fault-tolerant control and mixed integer programming.

References

Alamo, T., Bravo, J., & Camacho, E. (2005). Guaranteed state estimation by zonotopes. *Automatica, 41*(6), 1035–1043.

Blanchini, F., & Miani, S. (2008). *Set-theoretic methods in control*. Boston, MA: Birkhäuser.

Blanke, M., Kinnaert, M., Lunze, J., & Staroswiecki, M. (2006). *Diagnosis and fault-tolerant control*. Berlin: Springer-Verlag.

Blesa, J., Puig, V., Romera, J., & Saludes, J. (2011, August–September). Fault diagnosis of wind turbines using a set-membership approach. In *Proceedings of the 18th IFAC World Congress* (pp. 8316–8321). Milano.

Chai, W., & Qiao, J. (2014). Passive robust fault detection using RBF neural modeling based on set membership identification. *Engineering Applications of Artificial Intelligence, 28*(0), 1–12.

Combastel, C. (2003). A state bounding observer based on zonotopes. In *Proceedings of the 2003 European Control Conference*. Cambridge.

Guerra, P., Puig, V., & Witczak, M. (2008, July). Robust fault detection with unknown-input interval observers using zonotopes. In *Proceedings of the 17th IFAC World Congress* (pp. 5557–5562). Seoul.

Kofman, E., Haimovich, H., & Seron, M. (2007). A systematic method to obtain ultimate bounds for perturbed systems. *International Journal of Control, 80*(2), 167–178.

Kühn, W. (1998). Rigorously computed orbits of dynamical systems without the wrapping effect. *Computing, 61*(1), 47–67.

Le, V., Stoica, C., Alamo, T., Camacho, E., & Dumur, D. (2013, August). Zonotope-based set-membership estimation for multi-output uncertain systems. In *Proceedings of 2013 IEEE international Symposium on Intelligent Control (ISIC), Part of 2013 IEEE Multi-Conference on Systems and Control* (pp. 212–217). Hyderabad.

Mazenc, F., & Bernard, O. (2011). Interval observers for linear time-invariant systems with disturbances. *Automatica, 47*(1), 140–147.

Odgaard, P., Stoustrup, J., & Kinnaert, M. (2009, July). Fault tolerant control of wind turbines-a benchmark model. In *Proceedings of the 7th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Process* (pp. 1168–1182). Barcelona.

Olaru, S., Doná, J.D., Seron, M., & Stoican, F. (2010). Positive invariant sets for fault tolerant multisensor control schemes. *International Journal of Control, 83*(12), 2622–2640.

Puig, V. (2010). Fault diagnosis and fault tolerant control using set-membership approaches: Application to real case studies. *International Journal of Applied Mathematics and Computer Science, 20*(4), 619–635.

Raïssi, T., Videau, G., & Zolghadri, A. (2010). Interval observer design for consistency checks of nonlinear continuous-time systems. *Automatica, 46*(3), 518–527.

Seron, M., Zhuo, X., Doná, J.D., & Martínez, J. (2008). Multisensor switching control strategy with fault tolerance guarantees. *Automatica, 44*(1), 88–97.

Stoican, F., & Olaru, S. (2013). *Set-theoretic fault tolerant control in multisensor systems*. Engineering & Materials Science. London: Wiley - ISTE.

Xu, F., Puig, V., Ocampo-Martinez, C., Stoican, F., & Olaru, S. (2014, August). Improved fault detection and isolation strategy using a bank of interval observers. In *Proceedings of the 19th IFAC World Congress* (pp. 8024–8029). Cape Town.

Xu, F., Stoican, F., Puig, V., Ocampo-Martinez, C., & Olaru, S. (2013, June). Fault detection and isolation based on the combination of a bank of interval observers and invariant sets. In *Proceedings of the 21st Mediterranean Conference on Control and Automation* (pp. 807–813). Chania.

Appendix 1. Zonotopes

According to Alamo et al. (2005), Le et al. (2013) and Kühn (1998), one summarises the relevant results of zonotopes used in this paper.

Definition A.1: An r-order zonotope Z is defined as $Z = g \oplus H\mathbb{B}^r$, where g and H are the centre and segment matrix (also generator matrix), respectively.

Definition A.2: The interval hull $\square Z$ of a zonotope $Z = g \oplus H\mathbb{B}^r \subset \mathbb{R}^n$ is the smallest box containing X and the interval hull is denoted as $\square Z = \{x \mid |x_i - g_i| \leq \| H_i \|_1\}$, where H_i is the ith row of H, and x_i and g_i are the ith components of x and g, respectively.

Property A.1: Given two zonotopes $Z_1 = g_1 \oplus H_1\mathbb{B}^{r_1} \subset \mathbb{R}^n$ and $Z_2 = g_2 \oplus H_2\mathbb{B}^{r_2} \subset \mathbb{R}^n$, the Minkowski sum of them is $Z_1 \oplus Z_2 = \{g_1 + g_2\} \oplus [H_1 \quad H_2]\mathbb{B}^{r_1+r_2}$.

Property A.2: Given a zonotope $Z = g \oplus H\mathbb{B}^r \subset \mathbb{R}^n$ and a compatible matrix K, $KZ = Kg \oplus KH\mathbb{B}^r$.

Property A.3: Given a zonotope $Z = g \oplus H\mathbb{B}^r \subset \mathbb{R}^n$ and an integer s (with $n < s < r$), denote by $\hat{H}$ the matrix resulting from the reordering of the columns of H in decreasing Euclidean norm. $Z \subseteq g \oplus [\hat{H}_T \quad Q]\mathbb{B}^s$ where $\hat{H}_T$ is obtained from the first $s - n$ columns of $\hat{H}$ and $Q \in \mathbb{R}^{n \times n}$ is a diagonal matrix whose elements satisfy $Q_{ii} = \sum_{j=s-n+1}^{r} | \hat{H}_{ij} |$, $i = 1, \ldots, n$.

Property A.4: Given a zonotope $Z = g \oplus H\mathbb{B}^r \subset \mathbb{R}^n$, a strip $S = \{x \in \mathbb{R}^n \mid |cx - d| \leq \sigma\}$ and a vector $\lambda \in \mathbb{R}^n$, then $Z \cap S \subseteq \hat{Z}(\lambda) = \hat{g}(\lambda) \oplus \hat{H}(\lambda)\mathbb{B}^{r+1}$ where $\hat{g}(\lambda) = g + \lambda(d - cg)$ and $\hat{H}(\lambda) = [(I - \lambda c)H \quad \sigma\lambda]$.

Property A.5: Given a matrix $\Lambda \in \mathbb{R}^{n \times m}$, a zonotope $Z = g \oplus H\mathbb{B}^r$, and an H-polytope $P = \{x \in \mathbb{R}^n : |Cx - d| \leq [\phi_1, \phi_2, ..., \phi_m]^T\}$, with $C \in \mathbb{R}^{m \times n}$, $d \in \mathbb{R}^m$, $\phi_i \in \mathbb{R}_+$ ($i = 1, 2, ..., m$), define a vector $\hat{g}(\Lambda) = g + \Lambda(d - Cg)$ and a matrix $\hat{H}(\Lambda) = [(I - \Lambda C)H \quad \Lambda\phi]$, with a diagonal matrix $\Lambda = \text{diag}(\phi_1, \phi_2, ..., \phi_m)$. Then, a family of zonotopes (parameterised by the matrix Λ) that contains the intersection of the zonotope Z, and the polytope P is obtained such as $Z \cap P \subseteq \hat{Z}(\Lambda) = \hat{g} \oplus \hat{H}\mathbb{B}^{r+m}$.

Property A.6: Given a family of zonotopes denoted by $Z = g \oplus \mathbf{H}\mathbb{B}^m$, where $g \in \mathbb{R}^n$ is a real vector and $\mathbf{H} \in \mathbb{R}^{n \times m}$ is an

interval matrix, a zonotope inclusion $\diamond(Z)$ is defined by

$$\diamond(Z) = g \oplus [\text{mid}(\mathbf{H}) \quad H] \mathbb{B}^{m+n},$$

where the matrix H is a diagonal matrix with

$$H_{ii} = \sum_{j=1}^{m} \frac{\text{diam}(\mathbf{H})_{ij}}{2}, \ i = 1, 2, \dots, n,$$

where mid($\cdot$) and diam($\cdot$) compute the centre and diameter of interval matrices, respectively.

Property A.7: Given $Z_{k+1} = \mathbf{A}Z_k \oplus \mathbf{B}u_k$, where $\mathbf{A}$ and $\mathbf{B}$ are interval matrices and u_k is the input at time instant k, if Z_k is a zonotope with the centre g_k, and segment matrix H_k, Z_{k+1} can be bounded by

$$Z_{k+1}^e = g_{k+1} \oplus H_{k+1}\mathbb{B}^r,$$

with

$$\begin{aligned}
g_{k+1} &= \text{mid}(\mathbf{A})g_k + \text{mid}(\mathbf{B}))u_k, \\
H_{k+1} &= \quad [J_1 \quad J_2 \quad J_3], \\
J_1 &= \quad \text{seg}(\diamond(\mathbf{A}H_k)), \\
J_2 &= \quad \frac{\text{diam}(\mathbf{A})}{2}g_k, \\
J_3 &= \quad \frac{\text{diam}(\mathbf{B})}{2}u_k,
\end{aligned}$$

where seg($\cdot$) computes the segment matrix of a zonotope.

Appendix 2. Invariant sets

According to Kofman et al. (2007) and Olaru et al. (2010), the linear discrete time-invariant dynamics are used to present the invariant set notions, i.e.,

$$x_{k+1} = A_\circ x_k + B_\circ \delta_k, \tag{B1}$$

where $A_\circ$ and $B_\circ$ are constant matrices and $A_\circ$ is a Schur matrix, δ_k belongs to $\Delta = \{\delta : |\delta - \delta^\circ| \le \bar{\delta}\}$ with δ° and $\bar{\delta}$ constant and the matrices in (B1) have suitable dimensions.

Definition B.1: A set $X \subset \mathbb{R}^n$ is called an RPI set for (B1), if and only if $A_\circ X \oplus B_\circ \Delta \subseteq X$.

Definition B.2: The mRPI set of (B1) is defined as an RPI set contained in any closed RPI set and the mRPI set is unique and compact.

Theorem B.1: *Considering* (B1) *and letting* $A_\circ = V\Lambda V^{-1}$ *be the Jordan decomposition of* $A_\circ$, *the set*

$$\Phi(\theta) = \{x \in \mathbb{R}^n : \left|V^{-1}x\right| \le (I - |\Lambda|)^{-1}\left|V^{-1}B_\circ\right|\bar{\delta} + \theta\} \oplus \xi^\circ \tag{B2}$$

is RPI and attractive for the trajectories of (B1), *with* θ *any (arbitrarily small) vector with positive components, where* $\xi^\circ = (I - A_\circ)^{-1}B_\circ\delta^\circ$:

(1) *For any* θ, *the set* $\Phi(\theta)$ *is (positively) invariant, that is, if* $x_0 \in \Phi(\theta)$, *then* $x_k \in \Phi(\theta)$ *for all* $k \ge 0$.
(2) *Given* $\theta \in \mathbb{R}^n$, $\theta > 0$, *and* $x_0 \in \mathbb{R}^n$, *there exists* $k^* \ge 0$ *such that* $x_k \in \Phi(\theta)$ *for all* $k \ge k^*$.

Proposition B.1: *Considering the dynamics* (B1), *the set sequence*

$$X_{j+1} = A_\circ X_j \oplus B_\circ \Delta, \ j = 0, 1, 2, \dots \tag{B3}$$

converges to the mRPI set of the dynamics, where if X_0 *is RPI, each iteration of the set sequence is an RPI approximation of the mRPI set of the dynamics.*

Definition B.3: Given a scalar $\epsilon > 0$ and a set $\Omega \subset \mathbb{R}^n$, the set $\Phi \subset \mathbb{R}^n$ is an outer ϵ-approximation of Ω if $\Omega \subseteq \Phi \subseteq \Omega \oplus \mathbb{B}_s^n(\epsilon)$ and it is an inner ϵ-approximation of Ω if $\Phi \subseteq \Omega \subseteq \Phi \oplus \mathbb{B}_s^n(\epsilon)$, where $\mathbb{B}_s^n(\epsilon) = \{x \in \mathbb{R}^n : \| x \|_s \le \epsilon\}$.