Automatica 153 (2023) 111020

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Exclusion tendency-based observer design framework for active fault diagnosis☆

Yidian Fan [a], Feng Xu [a,*], Xueqian Wang [a], Bin Liang [b]

[a] Tsinghua Shenzhen International Graduate School, Tsinghua University, 518055 Shenzhen, PR China
[b] Department of Automation, Tsinghua University, 10084 Beijing, PR China

ARTICLE INFO

Article history:
Received 11 March 2022
Received in revised form 22 November 2022
Accepted 5 February 2023
Available online 14 April 2023

Keywords:
Active fault diagnosis
Set-valued observers
Optimal observer design
Exclusion tendency

ABSTRACT

This paper proposes a new closed-loop observer-based active fault diagnosis (AFD) framework using a bank of set-valued observers (SVOs). Each SVO is designed for one healthy or faulty system mode to obtain the corresponding output estimation set (OES). The principle of fault diagnosis is to check whether the considered OES includes the system output. Compared with the existing works that mainly make use of SVOs to obtain accurate estimation results, our proposed method optimizes the observer gains to deform OESs such that the exclusion tendency of the system output in each OES is maximized. With this idea, inconsistent system modes can be removed rapidly, and faults can be diagnosed. The design logic is formulated as a bi-level max–min problem, which is transformed into a linear problem by Lagrangian duality and solved efficiently. At the end of this paper, a quadruple-tank example is used to compare the results of our proposed method with an existing fault diagnosis method to illustrate the effectiveness of the proposed method.

1. Introduction

Because of the increasing complexity of modern systems, fault diagnosis has become a crucial technique to improve system safety (Blanke et al., 2015; Chiang et al., 2001). In general, there are two classes of fault diagnosis methods: the passive/active methods. The passive methods collect the output information for given inputs to diagnose faults without interfering with system operations (Efimov et al., 2013; Zhang & Yang, 2020), which are non-invasive to the system but suffer from conservative performance. The AFD methods design appropriate inputs to excite the system such that richer output information can be obtained to realize fault diagnosis (Cao et al., 2022; Heirung & Mesbah, 2019; Xu, 2023). In practice, an effective AFD method is required to be robust against uncertainties in real systems like process disturbances and measurement noises. The set-based AFD methods, which model uncertainties by considering their bounds, can provide deterministic diagnosis results and have been extensively studied in the community (Puncochar & Skach, 2018).

The set-based AFD methods usually assume that the system has different dynamics under different faults and construct a bank of OESs accordingly. It was proposed in Nikoukhah (1998) that an input sequence could be designed offline such that different OESs were separated in some time steps. After injecting the separating inputs, faults were guaranteed to be diagnosed by checking the only consistent set to which the output eventually belonged. Scott et al. (2014) formulated a Mixed Integer Quadratic Programming (MIQP) problem to design the separating input sequence. Based on Scott et al. (2014), Raimondo et al. (2016) proposed a closed-loop framework that built SVOs and employed real-time system outputs to generate more accurate OESs, hence reducing the conservatism of input design. Although the AFD method in Raimondo et al. (2016) has demonstrated excellent performance, the computation complexity is considerable because of the following two factors:

- MIQP is a non-convex optimization problem involving multiple binary variables that requires sizable computational resources to solve, especially for high-dimensional systems.
- SVOs implemented in Raimondo et al. (2016) require set intersection to update OESs, complicating the set representation as time elapses.

There exist recent works addressing the computation complexity issue. Yang et al. (2020) resorted to invariant sets and, instead of

☆ This work was supported by the National Natural Science Foundation of China (62003186), the Natural Science Foundation of Guangdong, China (2021A1515012628), and the Shenzhen Science and Technology Program, China (JCYJ20210324132606015). The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Angelo Alessandri under the direction of Editor Thomas Parisini.

* Corresponding author.
E-mail addresses: fyd20@mails.tsinghua.edu.cn (Y. Fan), xu.feng@sz.tsinghua.edu.cn (F. Xu), wang.xq@sz.tsinghua.edu.cn (X. Wang), liangbin@tsinghua.edu.cn (B. Liang).

https://doi.org/10.1016/j.automatica.2023.111020
0005-1098/

solving an input sequence, designed the input step by step. Tan et al. (2022) formulated a fractional programming problem to design one-step input that maximized the separation tendency between OESs. Xu (2021) constructed a closed-loop framework to optimize both input and a bank of SVOs. The input was designed to maximize the center distances between OESs and solved in a reasonable time. Unlike in Raimondo et al. (2016), the Luenberger observers were implemented in Xu (2021) so that the set dynamics became mathematically efficient.

As an optimization variable, the observer gain can deform OES for a posed objective. Usually, the size of OES is minimized for a small set and accurate estimation (Combastel, 2015; Le et al., 2013; Xu, 2021). However, there exist works suggesting that observer gains also have the potential to improve fault diagnosis performance. Pourasghar, Combastel et al. (2019) came up with a fractional-programming problem to derive an observer gain that maximized system sensitivity to faults, and Tan et al. (2020) gave a rigorous solution. Xu et al. (2022) improved the design logic of Pourasghar, Combastel et al. (2019) by proposing a new notion named excluding degree to better characterize the set-based fault diagnosis performance. In this paper, inspired by the set-based fault diagnosis principle (i.e., excluding the system output from inconsistent OESs), we propose optimizing observer gains to maximize the exclusion tendency of the system output in each OES to accelerate the online diagnosis process. As a further improvement of the previous work (Xu, 2021), the new observer-based AFD framework in this paper has the following novelties:

- Compared with Xu (2021), the observer gains are further optimized by a new exclusion tendency-based principle to improve system diagnosability. As a result, our AFD method has a higher potential to achieve better diagnosis results.
- Compared with set separation-based AFD methods that need to solve MIQP problems, our proposed AFD framework solves optimal inputs and observer gains in a typical convex framework, which essentially reduces computational complexity.
- Compared with previous observer design methods that implicitly maximize system sensitivity to faults, we define a quantity called exclusion tendency that explicitly characterizes the inconsistency of the system output in each OES. The maximization of exclusion tendency is in accordance with the ultimate objective of fault diagnosis. The optimization is formulated as a max–min problem, then transformed into a linear problem by Lagrangian duality and solved efficiently.

The remainder of the paper is organized as follows: Section 2 provides preliminaries about notations. Section 3 introduces the considered system model, the principle of fault diagnosis, and the AFD method in Xu (2021). Section 4 presents the main results of our proposed method. In Section 5, an example is shown to compare the effectiveness of our approach with the previous work. Finally, conclusions are drawn in Section 6.

2. Preliminaries

The symbol I_n denotes the identity matrix of $(n \times n)$ dimension. $\mathbf{1}$ and $\mathbf{0}$ are vectors whose elements are all 1 and 0, respectively. $\mathrm{diag}(v)$ is a diagonal matrix whose diagonal elements are composed of a vector v. The index set $\mathbb{I}_n$ denotes $\{0, 1, \ldots, n\}$. $\geq$ and $\leq$ are component-wise inequalities when they are used on vectors or matrices. $\prec$ and $\succ$ denote negative and positive definiteness, respectively. $\lambda(A)$ and $\mathrm{tr}(A)$ indicate the eigenvalues and the trace of the matrix A, respectively. Given an interval matrix $\mathbf{A}$ as $\underline{A} \leq \mathbf{A} \leq \overline{A}$, we have $\mathrm{mid}(\mathbf{A}) = \frac{\underline{A}+\overline{A}}{2}$ and $\mathrm{rad}(\mathbf{A}) = \frac{\overline{A}-\underline{A}}{2}$. Given two n-dimensional sets X, Y and a matrix $K \in \mathbb{R}^{m\times n}$, the following set operations $X \oplus Y = \{x + y \mid x \in X, y \in Y\}$ and $KX = \{Kx \mid x \in X\}$ are defined.

An n-dimensional zonotope Z is described by its center $g \in \mathbb{R}^n$ and generator matrix $H \in \mathbb{R}^{n\times r}$ as $Z = \{g + H\xi \mid \xi \in \mathbb{R}^r, \|\xi\|_\infty \leq 1\}$, or $Z = \langle g, H\rangle$ for short, and its order is r (Alamo et al., 2005). Given two zonotopes $Z_1 = \langle g_1, H_1\rangle$ and $Z_2 = \langle g_2, H_2\rangle$, we have $Z_1 \oplus Z_2 = \langle g_1 + g_2, [H_1, H_2]\rangle$, $KZ_1 = \langle Kg_1, KH_1\rangle$. One commonly used measure of the size of Z is the Frobenius radius (F-radius), which is the Frobenius norm of H, i.e., $\|Z\|_F = \|H\|_F = \sqrt{\mathrm{tr}\left(H^T H\right)}$. Reordering the columns of H as $H = [h_1, h_2, \ldots, h_r]$, $\|h_j\|_2^2 \geq \|h_{j+1}\|_2^2$ $(j = 1, \ldots, r-1)$, a reduction operator $\downarrow_q$ is defined as $\downarrow_q (H) = [H_>, \square(H_<)] \in \mathbb{R}^{n\times q}$, in which $n \leq q \leq r$, $H_> = \left[h_1, \ldots, h_{q-n}\right]$ and $\square(H_<)_{ii} = \sum_{j=q-n+1}^{r} |H_{ij}|$ $(i = 1, \ldots, n)$. With the operator $\downarrow_q$, the inclusion property $\langle g, H\rangle \subseteq \langle g, \downarrow_q (H)\rangle$ is guaranteed, and the order of zonotope is reduced to q (Combastel, 2015).

3. Problem formulation

3.1. System models and SVOs

Consider the following discrete-time linear time-invariant (LTI) system affected by multiplicative actuator faults:

$$x_{k+1} = Ax_k + BG_i u_k + E\omega_k, \tag{1a}$$

$$y_k = Cx_k + F\eta_k, \tag{1b}$$

where $A \in \mathbb{R}^{n_x\times n_x}$, $B \in \mathbb{R}^{n_x\times n_u}$, $C \in \mathbb{R}^{n_y\times n_x}$, $E \in \mathbb{R}^{n_x\times n_\omega}$ and $F \in \mathbb{R}^{n_y\times n_\eta}$ are parametric matrices. u_k is the input constrained by a convex polytope U. $\omega_k \in \mathbb{R}^{n_\omega}$ is the unknown input vector like disturbances and modeling errors, and $\eta_k \in \mathbb{R}^{n_\eta}$ represents measurement noises. ω_k and η_k are contained in zonotopes $\mathcal{W} = \langle \omega^c, H_\omega\rangle$ and $\mathcal{V} = \langle \eta^c, H_\eta\rangle$, respectively, where ω^c and η^c are centers, and H_ω and H_η are generator matrices with appropriate dimensions. $G_i \in \mathbb{R}^{n_u\times n_u}$ $(i \in \mathbb{I}_{n_u})$ model different actuator modes. G_0 is the identity matrix modeling the healthy mode. $G_i = \mathrm{diag}(1, \ldots, 1, f_i, 1, \ldots, 1)$ $(i \in \mathbb{I}_{n_u}\backslash\{0\})$ models faults that occur in the ith actuator, whose ith component satisfies $0 \leq f_i < 1$. Here, we allow the fault magnitude f_i to vary in the given interval $[\underline{f}_i, \overline{f}_i]$ as the system evolves.

Assumption 3.1. For the system (1), the pair (A, C) in (1) is detectable. The matrix A is Schur and the matrix C has full row rank.

Assumption 3.2. The system (1) operates under one particular mode during the diagnosis process.

Assumption 3.1 is a usual requirement for the open-loop system stability and the existence of SVOs. Assumption 3.2 implies that the system mode remains unchanged while implementing the diagnosis algorithm, which is also necessary for successful fault diagnosis. According to (1), an SVO matching the ith actuator mode is designed as

$$\hat{X}_{k+1}^i = (A - L_k^i C)\hat{X}_k^i \oplus L_k^i y_k \oplus (-L_k^i F\mathcal{V}) \oplus B\mathbf{G}_i u_k \oplus E\mathcal{W}, \tag{2a}$$

$$\hat{Y}_{k+1}^i = C\hat{X}_{k+1}^i \oplus F\mathcal{V}, \tag{2b}$$

in which L_k^i is the gain of the ith SVO. $\mathbf{G}_0 = I_{n_u}$ models the healthy mode. $\mathbf{G}_i = \mathrm{diag}(1, \ldots, 1, [\underline{f}_i, \overline{f}_i], 1, \ldots, 1)$ $(i \in \mathbb{I}_{n_u}\backslash\{0\})$ are interval matrices modeling faults in the ith actuator. $\hat{X}_{k+1}^i$ and $\hat{Y}_{k+1}^i$ are the state and output estimation sets corresponding to the ith system mode at the $(k+1)$th time step, respectively. When the system is operating at the ith mode and $x_0 \in \hat{X}_0^i$ is given, the inclusions $x_{k+1} \in \hat{X}_{k+1}^i$ and $y_{k+1} \in \hat{Y}_{k+1}^i$ hold for all $k \geq 0$.

Lemma 3.1 (Xu, 2021). *For each interval matrix $\mathbf{G}_i$, $B\mathbf{G}_i u_k$ is included in a zonotope, i.e., $B\mathbf{G}_i u_k \subseteq \langle \text{mid}(B\mathbf{G}_i)u_k, \text{diag}\,(\text{rad}(B\mathbf{G}_i)u_k)\rangle$.*

Based on zonotopic operations in Section 2 and Lemma 3.1, the dynamics (2) can be rewritten into a center-generator matrix form:

$$\hat{x}^{i,c}_{k+1} = (A - L^i_k C)\hat{x}^{i,c}_k + L^i_k(y_k - F\eta^c) + \text{mid}(B\mathbf{G}_i)u_k + E\omega^c, \tag{3a}$$
$$\hat{H}^{i,x}_{k+1} = [(A - L^i_k C)\hat{H}^{i,x}_k, -L^i_k F H_\eta, \text{diag}(\text{rad}(B\mathbf{G}_i)u_k), EH_\omega], \tag{3b}$$
$$\hat{y}^{i,c}_{k+1} = C\hat{x}^{i,c}_{k+1} + F\eta^c, \tag{3c}$$
$$\hat{H}^{i,y}_{k+1} = [C\hat{H}^{i,x}_{k+1}, FH_\eta], \tag{3d}$$

where $\hat{x}^{i,c}_{k+1}$ and $\hat{H}^{i,x}_{k+1}$, $\hat{y}^{i,c}_{k+1}$ and $\hat{H}^{i,y}_{k+1}$ are centers and generator matrices of $\hat{X}^i_{k+1}$ and $\hat{Y}^i_{k+1}$, respectively. It should be mentioned that as system evolves, the order of $\hat{X}^i_{k+1}$ will increase rapidly. To avoid computational burden involving high-order zonotopes, the order reduction technique addressed in Section 2 should be utilized to control the order.

3.2. Asymptotic AFD framework

The principle of fault diagnosis is to check whether

$$y_{k+1} \notin \hat{Y}^i_{k+1},\ i \in \mathbb{I}_{n_u}, k \geq 0, \tag{4}$$

holds or not. The inconsistency (4) indicates that the system is not at the ith mode. Therefore, a fault can be diagnosed if there exists only one consistent OES eventually. Xu (2021) computed the input and a bank of time-varying SVOs for fault diagnosis at each time step. The input was optimized to maximize the center distances of OESs. Motivated by the idea of Zonotopic Kalman Filter (Combastel, 2015), the observer gains were designed to minimize the F-radius of each OES. The optimization problems were formulated as

$$u^*_k = \arg\max_{u_k \in U} \sum_{i,j \in \mathbb{I}_{n_u}, i \neq j} \lambda_{ij} \|\hat{y}^{i,c}_{k+1} - \hat{y}^{j,c}_{k+1}\|^2, \tag{5a}$$
$$L^{i,KF}_k = \arg\min_{L^i_k} \|\hat{Y}^i_{k+1}\|^2_F, \tag{5b}$$

where u^*_k and $L^{i,KF}_k$ were the optimal input and observer gains designed at the kth time step, and λ_{ij} $(i, j \in \mathbb{I}_{n_u}, i \neq j)$ were weighting coefficients. The key idea is that the separation tendency of different OESs is enlarged such that their intersection region decreases. Therefore, the system output is more likely to stay in the non-intersecting part of the consistent set and indicates the fault. See Fig. 1 for a demonstration of the design logic (in Fig. 1, the system has the second actuator fault). Since diagnosis is realized asymptotically, this method is called asymptotic active fault diagnosis (AAFD).

Remark 3.1. The optimization problem (5) is a simplified version of that in Xu (2021). In fact, u^*_k and $L^{i,KF}_k$ were coupled in the original problem, and Xu (2021) constructed a two-layer framework to solve it. Here the key design logic (i.e., (5)) is presented, based on which we propose a novel observer gain design framework in this paper.

However, the principle of set-based AFD, according to (4), is to ensure that the system output leaves inconsistent sets as quickly as possible. The set separation requirement is an effective but indirect way to realize the goal. In Section 4, we derive a new optimization objective for observer gains based on the principle (4), demonstrating a higher fault diagnosis potential.

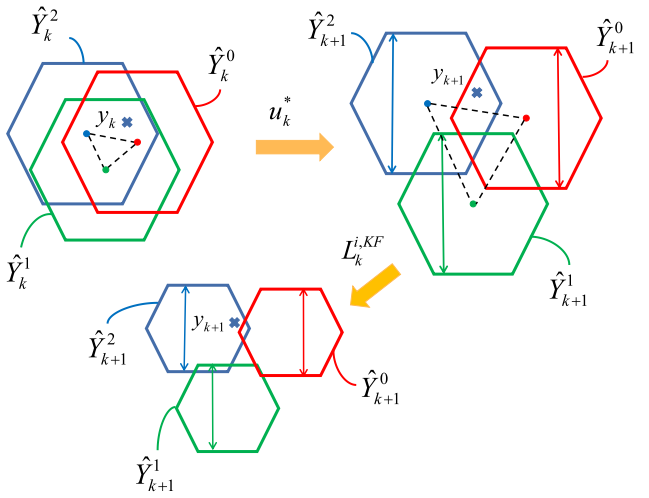


Fig. 1. Design logic of the AAFD method in Xu (2021). The input u^*_k maximizes the center distances, while $L^{i,KF}_k$ minimizes the size of $\hat{Y}^i_{k+1}$. In this figure, u^*_k and $L^{i,KF}_k$ are implemented separately for a clear illustration. In practice, u^*_k and $L^{i,KF}_k$ work together at the kth time step.

4. Main results

4.1. Optimal observer gains for fault diagnosis

Given the zonotopic representation $\hat{Y}^i_{k+1} = \langle \hat{y}^{i,c}_{k+1}, \hat{H}^{i,y}_{k+1}\rangle$, Scott et al. (2014) checks the relation between y_{k+1} and $\hat{Y}^i_{k+1}$ by the following optimization problem:

$$\hat{\delta}^i_{k+1} := \min_{\delta^i_{k+1}, \xi} \ \delta^i_{k+1}, \tag{6a}$$
$$\text{s.t.}\ y_{k+1} = \hat{y}^{i,c}_{k+1} + \hat{H}^{i,y}_{k+1}\xi, \tag{6b}$$
$$\|\xi\|_\infty \leq \delta^i_{k+1}. \tag{6c}$$

The optimal objective value $\hat{\delta}^i_{k+1}$ is the scaling factor of $\hat{Y}^i_{k+1}$ around its center $\hat{y}^{i,c}_{k+1}$ such that y_{k+1} is right on the surface of the scaled zonotope. The inclusion $y_{k+1} \in \hat{Y}^i_{k+1}$ holds if $\hat{\delta}^i_{k+1} \leq 1$. An illustrative explanation of (6) is given in Fig. 2. Since $y^I_{k+1} \in \hat{Y}^i_{k+1}$, $\hat{Y}^i_{k+1}$ needs to shrink around $\hat{y}^{i,c}_{k+1}$, leading to $\hat{\delta}^i_{k+1} < 1$. Since $y^E_{k+1} \notin \hat{Y}^i_{k+1}$, $\hat{Y}^i_{k+1}$ has to enlarge itself and $\hat{\delta}^i_{k+1} > 1$. Moreover, $\hat{\delta}^i_{k+1} = 1$ serves as an exact boundary between inclusion and exclusion. We formally define $\hat{\delta}^i_{k+1}$ as the exclusion tendency in Definition 4.1.

Definition 4.1. The optimal objective value $\hat{\delta}^i_{k+1}$ of the problem (6) measures the exclusion tendency of y_{k+1} in an output estimation zonotope $\hat{Y}^i_{k+1}$.

Based on (4) and the above analysis, the increase of exclusion tendency of the system output in each inconsistent set from $\hat{\delta} < 1$ to $\hat{\delta} > 1$ is the ultimate objective of set-based AFD methods. According to the set propagation dynamics (3), L^i_k can change the center $\hat{y}^{i,c}_{k+1}$ and generator matrix $\hat{H}^{i,y}_{k+1}$ of $\hat{Y}^i_{k+1}$, and thus affects the exclusion tendency $\hat{\delta}^i_{k+1}$. In this paper, we propose to optimize L^i_k such that $\hat{\delta}^i_{k+1}$ is maximized. The optimization problem is formulated as follows:

$$\hat{\delta}^{i,*}_{k+1} := \max_{L^i_k} \hat{\delta}^i_{k+1} = \max_{L^i_k} \min_{\delta^i_{k+1}, \xi} \delta^i_{k+1}. \tag{7}$$

The optimal fault diagnosis observer gain in (7) is denoted as $L^{i,FD}_k$. We take Fig. 3 as an example to demonstrate the significance of

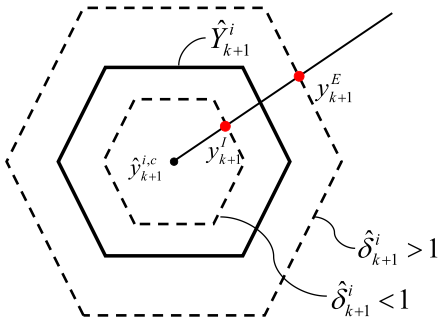


Fig. 2. The intuitive explanation of the optimization problem (6). $\hat{\delta}^i_{k+1}$ is the scaling factor of $\hat{Y}^i_{k+1}$ around $\hat{y}^{i,c}_{k+1}$.

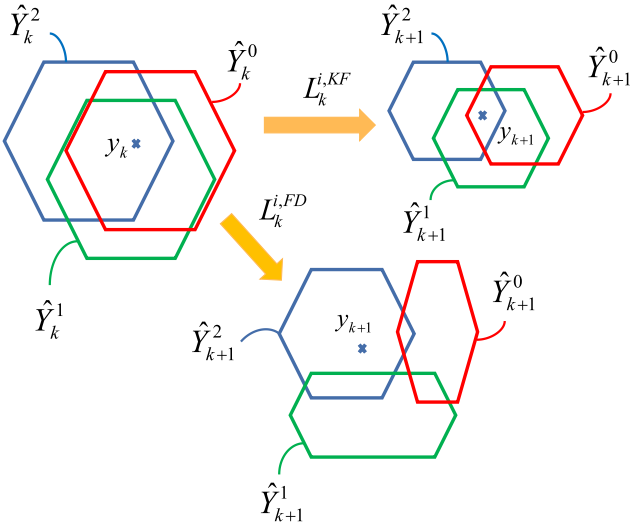


Fig. 3. An illustration of the significance of $L^{i,FD}_k$ compared with $L^{i,KF}_k$.

$L^{i,FD}_k$. For inconsistent OES, $L^{i,FD}_k$ that maximizes $\hat{\delta}^i_{k+1}$ can boost the exclusion of y_{k+1} from $\hat{Y}^i_{k+1}$, which corresponds to the bottom part of Fig. 3. Since the only consistent set is $\hat{Y}^2_{k+1}$, the current fault is identified in the second actuator. The top part of Fig. 3 shows OESs obtained from $L^{i,KF}_k$ designed in (5b). Although the size of each $\hat{Y}^i_{k+1}$ is minimized, y_{k+1} is still included in the intersection region, and we cannot determine the current system mode.

It should be mentioned that our proposed method designs $L^{i,FD}_k$ at the $(k+1)$th time step because y_{k+1} should be known as a prior. Since u_k has already been determined by using the method in Xu (2021) and injected into the system at the kth time step to obtain y_{k+1}, the exclusion tendency-based framework does not repeat the input design process for brevity and directly uses the same input design method as the AAFD framework (i.e., (5a)). The procedure of our proposed AFD framework is summarized as follows:

- At the kth time step, u^*_k is injected to maximize the separation tendency between OESs at the $(k+1)$th time step;
- At the $(k+1)$th time step, each OES is further deformed by $L^{i,FD}_k$ to exclude the system output y_{k+1} from $\hat{Y}^i_{k+1}$.

In general, u^*_k excites the system to generate more useful output information for fault inference, while $L^{i,FD}_k$ corrects results to enhance system diagnosability. Both design principles work in tandem and constitute a powerful AFD paradigm.

4.2. Dual transformation of the optimization problem

The max–min problem (7) is difficult to solve directly. Due to the strong duality of the linear problem, (7) can be transformed into a single-level max problem.

Proposition 4.1. *For the system (1), the bi-level max–min problem (7) is equivalent to the following single-level max problem:*

$$\hat{\delta}^{i,*}_{k+1} = \max_{L^i_k, \lambda^i_{n,k}, v^i_k} \; {v^i_k}^T CL^i_k \left(y_k - \hat{y}^{i,c}_k\right) + {v^i_k}^T c^{i,\Delta}_k \tag{8a}$$

$$s.t.\ {\lambda^i_{1,k}}^T - {\lambda^i_{2,k}}^T = -{v^i_k}^T C(A - L^i_k C)\hat{H}^{i,x}_k, \tag{8b}$$

$${\lambda^i_{3,k}}^T - {\lambda^i_{4,k}}^T = {v^i_k}^T CL^i_k FH_\eta, \tag{8c}$$

$${\lambda^i_{5,k}}^T - {\lambda^i_{6,k}}^T = -{v^i_k}^T H^{i,\Delta}_k, \tag{8d}$$

$$\lambda^i_{n,k} \geq \mathbf{0},\ n = 1, \ldots, 6, \tag{8e}$$

$$\sum_{n=1}^{6} {\lambda^i_{n,k}}^T \mathbf{1} = 1,\ \ \forall i \in \mathbb{I}_{n_u}, \tag{8f}$$

where v^i_k and $\lambda^i_{n,k}$ $(n = 1, \ldots, 6)$ are Lagrange multipliers derived from the dual transformation of the problem (6), $c^{i,\Delta}_k = C(A\hat{x}^{i,c}_k + \text{mid}(B\mathbf{G}_i)u_k + E\mathbf{w}^c) + F\eta^c - y_{k+1}$ and $H^{i,\Delta}_k = [C\ \text{diag}\,(\text{rad}(B\mathbf{G}_i)u_k)\,, CEH_\omega, FH_\eta]$.

Proof. Based on the expressions of $\hat{y}^{i,c}_{k+1}$ and $\hat{H}^{i,y}_{k+1}$ in (3), the inner problem (6) is reformulated as follows:

$$\min_{\delta^i_{k+1}, \xi} \ \delta^i_{k+1} \tag{9}$$

$$\begin{aligned}
s.t.\ \mathbf{0} &= CL^i_k(y_k - \hat{y}^{i,c}_k) + c^{i,\Delta}_k \\
&\quad + C(A - L^i_k C)\hat{H}^{i,x}_k \xi_1 - CL^i_k FH_\eta \xi_2 + H^{i,\Delta}_k \xi_3, \\
&\xi_1 - \delta^i_{k+1}\mathbf{1} \leq \mathbf{0}, -\xi_1 - \delta^i_{k+1}\mathbf{1} \leq \mathbf{0}, \\
&\xi_2 - \delta^i_{k+1}\mathbf{1} \leq \mathbf{0}, -\xi_2 - \delta^i_{k+1}\mathbf{1} \leq \mathbf{0}, \\
&\xi_3 - \delta^i_{k+1}\mathbf{1} \leq \mathbf{0}, -\xi_3 - \delta^i_{k+1}\mathbf{1} \leq \mathbf{0}, \\
&\xi = \left[\xi_1^T, \xi_2^T, \xi_3^T\right]^T.
\end{aligned}$$

The Lagrangian of (9) is

$$\begin{aligned}
&L(v^i_k, \lambda^i_{n,k}, \delta^i_{k+1}, \xi) = {v^i_k}^T CL^i_k(y_k - \hat{y}^{i,c}_k) + {v^i_k}^T c^{i,\Delta}_k \\
&+ (1 - \sum_{n=1}^{6} {\lambda^i_{n,k}}^T \mathbf{1})\delta^i_{k+1} + ({\lambda^i_{1,k}}^T - {\lambda^i_{2,k}}^T + {v^i_k}^T C(A \\
&- L^i_k C)\hat{H}^{i,x}_k)\xi_1 + ({\lambda^i_{3,k}}^T - {\lambda^i_{4,k}}^T - {v^i_k}^T CL^i_k FH_\eta)\xi_2 \\
&+ ({\lambda^i_{5,k}}^T - {\lambda^i_{6,k}}^T + {v^i_k}^T H^{i,\Delta}_k)\xi_3,
\end{aligned} \tag{10}$$

in which v^i_k and $\lambda^i_{n,k}$ $(n = 1, \ldots, 6)$ are Lagrange multipliers associated with equality and inequality constraints of the ith SVO at the kth time step, respectively. According to (10), the Lagrange dual problem is further derived as

$$\max_{\lambda^i_{n,k}, v^i_k} \ {v^i_k}^T CL^i_k \left(y_k - \hat{y}^{i,c}_k\right) + {v^i_k}^T c^{i,\Delta}_k \tag{11a}$$

$$s.t.\ \ (8b) - (8f), \forall i \in \mathbb{I}_{n_u}. \tag{11b}$$

It is obvious that the inner problem (6) is linear. Therefore, according to Boyd and Vandenberghe (2004), the optimal objective values of the problems (6) and (11) are the same. Based

on the formulation (7), the bi-level max–min problem can be transformed into a single-level max problem by setting L_k^i in (11) as a new optimization variable. The final form is given in (8). □

Remark 4.1. In fact, the formulation of the optimization problem (8) is not limited to actuator faults. One intuitive generalization is to replace A and C by A_i and C_i, in which A_i and C_i model system faults and sensor faults of the ith system mode, respectively. However, the AAFD method in Xu (2021) only considers actuator faults. To realize a fair comparison with the previous work, only actuator faults are evaluated in this work.

We can get the optimal fault diagnosis observer gain $L_k^{i,FD}$ by solving (8). However, the optimization variables v_k^i and L_k^i are multiplied in (8), making the original problem unsolvable in a typical convex framework.

4.3. Problem solutions

In Proposition 4.2, we show that (8) can be reformulated as a linear problem. Before that, Lemma 4.1 is needed.

Lemma 4.1 (*Golub & Van Loan, 2013*)**.** *Consider linear equations*

$$AX = Y, \tag{12}$$

where $A \in \mathbb{R}^{m\times n}$, $X \in \mathbb{R}^{n\times r}$, $Y \in \mathbb{R}^{m\times r}$. If $\mathrm{rank}(A) = m$, then the general solution to (12) is

$$X^* = A^\dagger Y + (I_n - A^\dagger A)R, \tag{13}$$

in which $A^\dagger = A^T(AA^T)^{-1}$ is the pseudo-inverse of A, $R \in \mathbb{R}^{n\times r}$ is an arbitrary matrix.

Proposition 4.2. *Consider the following linear optimization problem:*

$$\overline{\delta}_{k+1}^{i,*} := \max_{\gamma_k^i, \lambda_{n,k}^i, v_k^i} \ {\gamma_k^i}^T \left(y_k - \hat{y}_k^{i,c}\right) + {v_k^i}^T c_k^{i,\Delta} \tag{14a}$$

$$s.t.\ {\lambda_{1,k}^i}^T - {\lambda_{2,k}^i}^T = -{v_k^i}^T CA\hat{H}_k^{i,x} + {\gamma_k^i}^T C\hat{H}_k^{i,x}, \tag{14b}$$

$${\lambda_{3,k}^i}^T - {\lambda_{4,k}^i}^T = {\gamma_k^i}^T FH_\eta, \tag{14c}$$

$$(8d),\ (8e),\ (8f),\ \forall i \in \mathbb{I}_{n_u}. \tag{14d}$$

If the optimal solution $v_k^{i,*}$ of (14) is not $\mathbf{0}$, then $\overline{\delta}_{k+1}^{i,*} = \hat{\delta}_{k+1}^{i,*}$, i.e., the optimal objective values of (14) and (8) are the same. With the optimal solutions $\gamma_k^{i,*}$ and $v_k^{i,*}$, the optimal fault diagnosis observer gain $L_k^{i,FD}$ in the problem (8) has the form

$$L_k^{i,FD} = L_k^{i,0} + L_k^{i,1} R_k^i, \tag{15}$$

where $R_k^i \in \mathbb{R}^{n_x\times n_y}$ is an arbitrary matrix,

$$L_k^{i,0} = \frac{C^T v_k^{i,*} {\gamma_k^{i,*}}^T}{{v_k^{i,*}}^T CC^T v_k^{i,*}}, \quad L_k^{i,1} = I_{n_x} - \frac{C^T v_k^{i,*} {v_k^{i,*}}^T C}{{v_k^{i,*}}^T CC^T v_k^{i,*}}.$$

Proof. It is obvious that (8) is equal to the linear problem (14) with an additional constraint ${v_k^i}^T CL_k^i = {\gamma_k^i}^T$. As a new optimization variable γ_k^i is introduced, the relation $\overline{\delta}_{k+1}^{i,*} \geq \hat{\delta}_{k+1}^{i,*}$ holds. After solving (14), the equality $\overline{\delta}_{k+1}^{i,*} = \hat{\delta}_{k+1}^{i,*}$ is satisfied if there exists $L_k^{i,FD}$ such that ${v_k^{i,*}}^T CL_k^{i,FD} = {\gamma_k^{i,*}}^T$. Lemma 4.1 indicates that $L_k^{i,FD}$ has the general form (15) if $\mathrm{rank}({v_k^{i,*}}^T C) = 1$. Since Assumption 3.1 says that C has full row rank, $\mathrm{rank}({v_k^{i,*}}^T C) = 1$ holds if $v_k^{i,*} \neq \mathbf{0}$. In conclusion, $\overline{\delta}_{k+1}^{i,*} = \hat{\delta}_{k+1}^{i,*}$ if $v_k^{i,*} \neq \mathbf{0}$, and the optimal fault diagnosis observer gain $L_k^{i,FD}$ of the problem (8) has the general form (15). □

From Proposition 4.2, we know that the optimal solution $L_k^{i,FD}$ of (8) has the general form (15), in which $v_k^{i,*}$ and $\gamma_k^{i,*}$ are obtained from solving the linear problem (14), and R_k^i is a free variable remained to be determined. In principle, one essential purpose of observers is to generate an accurate estimation result. Based on the parametric form (15), R_k^i can be further optimized to increase estimation accuracy.

4.4. Improvements in estimation performance

To balance the performances of fault diagnosis and state estimation, we propose three design principles that exhibit respective privileges in optimizing observer gains.

(I) Minimization of F-radius: F-radius is a metric that characterizes the size of zonotope (Combastel, 2015). We can optimize R_k^i to minimize the F-radius of the state estimation set for more accurate results. According to (3b), the optimization problem is given by

$$\min_{R_k^i} \ \|\hat{X}_{k+1}^i\|_F^2 = \mathrm{tr}\left({\hat{H}_{k+1}^{i,x}}^T \hat{H}_{k+1}^{i,x}\right), \forall i \in \mathbb{I}_{n_u}, \tag{16}$$

in which

$$\hat{H}_{k+1}^{i,x} = \left[(A - L_k^{i,FD}C)\hat{H}_k^{i,x} \ \ -L_k^{i,FD}FH_\eta \ \ H_\Xi\right],$$
$$H_\Xi = \left[\mathrm{diag}(\mathrm{rad}(B\mathbf{G}_i)u_k) \ \ EH_\omega\right],$$

and $L_k^{i,FD}$ has the form (15). According to results in matrix calculus (Combastel, 2015), the derivative of $\|\hat{X}_{k+1}^i\|_F^2$ with respect to R_k^i is

$$\frac{\partial \|\hat{X}_{k+1}^i\|_F^2}{\partial R_k^i} = 2\left({L_k^{i,1}}^T L_k^{i,1}\right) R_k^i P - 2Q, \tag{17}$$

in which $P = FH_\eta H_\eta^T F^T + CH_k^{i,x} {H_k^{i,x}}^T C^T$, $Q = {L_k^{i,1}}^T (A - L_k^{i,0}C) H_k^{i,x} {H_k^{i,x}}^T C^T - {L_k^{i,1}}^T L_k^{i,0} FH_\eta H_\eta^T F^T$. Since (16) is an unconstrained convex quadratic problem, the optimal $R_k^{i,*}$ is given by setting (17) to zero, i.e.,

$$R_k^{i,*} = \left({L_k^{i,1}}^T L_k^{i,1}\right)^\dagger QP^{-1}. \tag{18}$$

(II) Stabilization of state estimation dynamics: One critical design principle for observer gain is to ensure the convergence of estimation error. The SVO (1) is called robustly stable if all eigenvalues of the matrix $(A - L_k^i C)$ have magnitudes less than one at each time step k (Combastel, 2015). The design of $L_k^{i,FD}$ that ensures the stable dynamics can also be addressed by resorting to the Lyapunov stability in the following proposition.

Proposition 4.3. *Consider the following optimization problem:*

$$\min_{R_k^i, \alpha} \ \alpha \tag{19a}$$

$$s.t. \begin{bmatrix} -W & \left(A - L_k^{i,FD}C\right)^T W \\ \star & -W \end{bmatrix} \preceq \alpha I_{2n_x}, \forall i \in \mathbb{I}_{n_u}, \tag{19b}$$

where $W \succ 0$ is a positive definite matrix, $\star$ denotes the symmetric term, and $L_k^{i,FD}$ has the form (15). Then the dynamics of SVO is stable at the kth time step if the optimal solution $\alpha^* < 0$.

Proof. Choose a Lyapunov function $V_k = x_k^T W x_k$. For the dynamics $x_{k+1} = (A - L_k^{i,FD}C)x_k$, $V_{k+1} < V_k$ is satisfied if $(A - L_k^{i,FD}C)^T W (A - L_k^{i,FD}C) - W \prec 0$, which is equivalent to the left-hand side of (19b) $\prec 0$ through Schur complement. If $\alpha^* < 0$, then it is obvious that the Lyapunov stability is guaranteed. □

By solving the optimization problem (19), we can derive an optimal $R_k^{i,*}$ that stabilizes the SVO dynamics if $\alpha^* < 0$. It should be mentioned that α^* is not warranted to be negative. Nevertheless, the optimization (19) still has the potential to suppress significant distortion.

(III) Stability guarantee: The most robust approach is to solve the optimization problem (8) directly with the Lyapunov stability condition added as a new constraint, i.e.,

$$\hat{\delta}_{k+1}^{i,L} = \max_{L_k^i, \lambda_{n,k}^i, v_k^i} \quad {v_k^i}^T C L_k^i \left(y_k - \hat{y}_k^{i,c}\right) + {v_k^i}^T c_k^{i,\Delta} \tag{20a}$$

$$\text{s.t. (8b)} - \text{(8f)}, \tag{20b}$$

$$\begin{bmatrix} -W & \left(A - L_k^i C\right)^T W \\ \star & -W \end{bmatrix} \prec 0, \forall i \in \mathbb{I}_{n_u}. \tag{20c}$$

The optimal observer gain given by solving (20) guarantees stable state estimation dynamics. Since v_k^i and L_k^i are decoupled in the constraint (20c), we cannot substitute the bi-linear term ${v_k^i}^T C L_k^i$ to transform this problem into a linear version as Proposition 4.2 does. One effective approach to solve (20) is branch-and-bound (BNB) (Liberti & Pantelides, 2006). The general idea is that the search space is branched into smaller regions in each iteration until the optimal solution is detected. In each region, the lower bound of the optimal objective value $\hat{\delta}_{k+1}^{i,L}$ can be given by a local solver, while the upper bound is obtained by replacing the bi-linear term with its convex relaxation (McCormick, 1976) and then solving the relaxed problem. Upper and lower bounds help the search algorithm to eliminate the regions that do not contain the optimal solution. The BNB algorithm will terminate if the gap between the global upper and lower bounds is below a given threshold. The software toolkit YALMIP (Lofberg, 2004) provides an efficient solver to solve (20), which is shown in Section 5.

The design principle (I) can give an analytic expression of the optimal $R_k^{i,*}$. The design principle (II) can generate observer gains that suppress the unstable dynamics of SVOs. The design principle (III) guarantees stable dynamics during the diagnosis process. However, the BNB method requires significant computation time to implement. The choice of one specific design principle depends on the application scenarios, such as the stability requirement, the available computation resources, etc. In Section 5, we give a numerical performance comparison of three design principles. With a bit of abuse of notations, we still denote the final observer gain that is further optimized by one of the above principles as $L_k^{i,FD}$.

4.5. Analysis of the singular solution

After solving (14), if the optimal solution $v_k^{i,*} = \mathbf{0}$ is obtained, then ${v_k^{i,*}}^T C L_k^i$ is always equal to $\mathbf{0}$, leading to no solution of the optimal $L_k^{i,FD}$. We call $v_k^{i,*} = \mathbf{0}$ as the singular solution. In this subsection, the property of the singular solution is analyzed.

Theorem 4.1. *For the optimization problem (14), if $v_k^{i,*} = \mathbf{0}$, then $\bar{\delta}_{k+1}^{i,*} = \hat{\delta}_k^i$.*

Proof. Substituting $v_k^i = \mathbf{0}$ into (14), we have

$$\bar{\delta}_{k+1}^{i,*} = \max_{\gamma_k^i, \lambda_{n,k}^i} \quad {\gamma_k^i}^T \left(y_k - \hat{y}_k^{i,c}\right) \tag{21a}$$

$$\text{s.t. } {\lambda_{1,k}^i}^T - {\lambda_{2,k}^i}^T = {\gamma_k^i}^T C \hat{H}_k^{i,x}, \tag{21b}$$

$${\lambda_{3,k}^i}^T - {\lambda_{4,k}^i}^T = {\gamma_k^i}^T F H_\eta, \tag{21c}$$

$$\lambda_{5,k}^i = \lambda_{6,k}^i,\ \lambda_{n,k}^i \geq \mathbf{0},\ n = 1, \ldots, 6, \tag{21d}$$

$$\sum_{n=1}^{6} {\lambda_{n,k}^i}^T \mathbf{1} = 1. \tag{21e}$$

Using the dual transformation again, the Lagrange dual problem of (21) is

$$\min_{\delta_k^i, \xi} \quad \delta_k^i \tag{22a}$$

$$\text{s.t. } y_k = \hat{y}_k^{i,c} + \hat{H}_k^{i,y} \xi,\ \ \|\xi\|_\infty \leq \delta_k^i. \tag{22b}$$

The optimization problem (22) computes the exclusion tendency of y_k in $\hat{Y}_k^i$, i.e., $\hat{\delta}_k^i$. The result indicates that $\bar{\delta}_{k+1}^{i,*} = \hat{\delta}_k^i$. □

It can be derived from Theorem 4.1 that $\forall L_k^i \in \mathbb{R}^{n_x \times n_y}$, we have $\hat{\delta}_{k+1}^i \leq \bar{\delta}_{k+1}^{i,*} = \hat{\delta}_k^i$. As a result, the exclusion tendency at the $(k+1)$th time step cannot exceed $\hat{\delta}_k^i$ by the design of L_k^i. Under this circumstance, an optimal fault diagnosis observer gain $L_k^{i,FD}$ cannot be computed. Nevertheless, an appropriate observer gain is needed for subsequent AFD. In our framework, we choose to implement $L_k^{i,KF}$ in (5b) as a compromise, and our proposed method will be reduced to the AAFD method in Xu (2021) when the singular solution happens. One essential issue worth addressing is that the computation of $L_k^{i,FD}$ will not be required for diagnosing faults when $\bar{\delta}_{k+1}^{i,*} \geq 1$ holds, thus saving the computation resources.

Proposition 4.4. *For the linear problem (14), if the optimal objective value $\bar{\delta}_{k+1}^{i,*} \geq 1$, then $\exists L_k^i \in \mathbb{R}^{n_x \times n_y}$ such that $y_{k+1} \notin \hat{Y}_{k+1}^i$.*

Proof. First, we prove that if $\bar{\delta}_{k+1}^{i,*} > 1$, then $v_k^{i,*} \neq \mathbf{0}$, i.e., the singular solution does not happen at the current step. If $v_k^{i,*} = \mathbf{0}$, it can be known from Theorem 4.1 that $\bar{\delta}_{k+1}^{i,*} = \hat{\delta}_k^i > 1$ holds, so $y_k \notin \hat{Y}_k^i$. Under this circumstance, the mode i should be removed from alternative current system modes at the previous step, which induces a contradiction. Since $v_k^{i,*} \neq \mathbf{0}$, Proposition 4.2 indicates that $\bar{\delta}_{k+1}^{i,*} = \hat{\delta}_{k+1}^{i,*} > 1$ holds. Thus there exists a certain L_k^i ensuring that the maximal exclusion tendency $\hat{\delta}_{k+1}^{i,*}$ exceeds one. Under this circumstance, the computation of $L_k^{i,FD}$ is unnecessary. □

Remark 4.2. For the SVO of the actual system mode i, $y_k \in \hat{Y}_k^i$ always holds, so the exclusion tendency $\hat{\delta}_k^i < 1$ holds for all $k \geq 0$. In this case, it is unrealistic to keep increasing $\hat{\delta}^i$. As a result, the singular solution is not a rare case in practice. However, according to Proposition 4.4, when $y_{k+1} \notin \hat{Y}_{k+1}^i$ holds for a certain L_k^i, the singular solution will not occur to interfere with the diagnosis. Moreover, since OESs keep correcting themselves to exclude the outputs, inconsistent SVOs are becoming increasingly sensitive to faults. Although the singular solution exists, our proposed work still has a higher fault diagnosis potential than the AAFD method in Xu (2021).

The entire design procedure of our proposed framework is summarized as Algorithm 1. The diagnosis is successful if there exists only one element in the final index set $\mathbb{I}_{n_u}$.

5. Illustrative examples

In this section, we compare our proposed method with the method in Xu (2021) on a quadruple-tank system (Pourasghar, Puig et al., 2019). The system is linearized around the equilibrium point $\bar{x} = [12.4, 12.7, 1.8, 1.4]^T$. Euler discretization with the sampling time $T_s = 1$ s is implemented to obtain an LTI system form (1). For a better comparison, we assume that all system states can be measured with appropriate sensors, so the matrix

Algorithm 1 Online implementation of the proposed AFD framework

Input: The system (1), initial state sets $\hat{X}_0^i$, $i \in \mathbb{I}_{n_u}$;
Initialization: $k = 0$;

while length($\mathbb{I}_{n_u}$) > 1 **and** $k \leq k_{\max}$ **do**
  Compute u_k^* and $L_k^{i,KF}$ ($i \in \mathbb{I}_{n_u}$) by (5);
  Iterate the system (1) and obtain y_{k+1};
  for $i \in \mathbb{I}_{n_u}$ **do**
    Solve (14) to get $v_k^{i,*}$ and $\bar{\delta}_{k+1}^{i,*}$;
    if $\bar{\delta}_{k+1}^{i,*} > 1$ **then**
      $\mathbb{I}_{n_u} \leftarrow \mathbb{I}_{n_u} \backslash \{i\}$;
      break
    end if
    if $v_k^{i,*} = \mathbf{0}$ **then**
      $L_k^{i,*} \leftarrow L_k^{i,KF}$;
    else
      Compute $L_k^{i,FD}$ by a specific principle;
      $L_k^{i,*} \leftarrow L_k^{i,FD}$;
    end if
    Use u_k^* and $L_k^{i,*}$ to compute $\hat{X}_{k+1}^i$ and $\hat{Y}_{k+1}^i$;
  end for
  $k \leftarrow k + 1$;
end while

Output: $\mathbb{I}_{n_u}$.

C is slightly modified. The parametric matrices of the system are given as

$$A = \begin{bmatrix} 0.9842 & 0 & 0.0419 & 0 \\ 0 & 0.9890 & 0 & 0.0333 \\ 0 & 0 & 0.9581 & 0 \\ 0 & 0 & 0 & 0.9672 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.2102 & 0 \\ 0 & 0.0628 \\ 0 & 0.0479 \\ 0.0094 & 0 \end{bmatrix},$$

$$E = \begin{bmatrix} 0.05 & 0.01 & 0 & 0 & 0 & 0 & 0 \\ 0.05 & 0 & 0.01 & 0 & 0 & 0 & 0 \\ 0.05 & 0 & 0 & 0.01 & 0 & 0 & 0 \\ 0.05 & 0 & 0 & 0 & 0.01 & 0 & 0 \end{bmatrix},$$

$$C = \begin{bmatrix} 0.5 & 0 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \end{bmatrix},$$

$$F = \begin{bmatrix} 0.01 & 0 & 0 & 0 & 0 & 0.01 & 0 \\ 0 & 0.01 & 0 & 0 & 0 & 0 & 0.01 \end{bmatrix}. \quad (23)$$

The system is equipped with two actuators, so three system modes are considered, i.e., $\mathbb{I}_{n_u} = \{0, 1, 2\}$. The input constraint set is $U = \langle \mathbf{0}, 0.6I_2 \rangle$. The initial conditions and bounds of uncertainties are given by

$$x_0 = [0, 0, 0, 0]^T, \hat{X}_0^i = \langle x_0, 0.1I_4 \rangle \quad (i \in \mathbb{I}_{n_u}),$$
$$\mathcal{W} = \langle \mathbf{0}, 0.1I_7 \rangle, \mathcal{V} = \langle \mathbf{0}, 0.1I_7 \rangle.$$

The positive definite matrix in (19) is chosen as $W = I_{n_x}$. For the balance of efficiency and complexity, the order of $\hat{X}_k^i$ is kept as 60, and that of $\hat{Y}_k^i$ is 10. The maximal diagnosis step $k_{\max} = 10$ in Algorithm 1 is set. The actuator fault matrices $\mathbf{G}_i$ are

$$\text{mid}(\mathbf{G}_1) = \text{diag}([f_1^c, 1]), \quad \text{mid}(\mathbf{G}_2) = \text{diag}([1, f_2^c]),$$
$$\text{rad}(\mathbf{G}_1) = \text{diag}([0.05, 0]), \quad \text{rad}(\mathbf{G}_2) = \text{diag}([0, 0.05]),$$

where f_1^c and f_2^c are centers of the fault intervals. The radius of the fault intervals is 0.05. All programs are evaluated on an Intel(R) Core(TM) i7-9700 CPU @ 3.00 GHz machine with MATLAB 2020a equipped with single-thread execution. The optimization problems have been solved using the CVX toolbox except for the BNB algorithm involved in the design principle (III), which is conducted by the YALMIP toolbox using the BMIBNB solver.

Take $(f_1^c, f_2^c) = (0.8, 0.85)$ as an example. Suppose the first actuator is disrupted by a constant fault $f_1 = 0.8$. We use our proposed method to design inputs and observer gains step by step to diagnose the fault. Moreover, both the design principles (I) and (II) are implemented to show their diagnosis performances. The results are shown in Fig. 4. It can be seen that, for both design principles, $y_3 \in \hat{Y}_3^1$ while $y_3 \notin \hat{Y}_3^0$ and $y_3 \notin \hat{Y}_3^2$ are observed, indicating the first actuator fault at $k = 3$. As a comparison, we adopt the AAFD method in Xu (2021) to diagnose the fault, and the results are shown in Fig. 5. At $k = 3$, $y_3 \in \hat{Y}_3^0 \cap \hat{Y}_3^1 \cap \hat{Y}_3^2$, so the system mode cannot be identified at $k = 3$. Further simulation shows that the AAFD method fails to diagnose within $k_{\max} = 10$ evolution steps. The comparison suggests that our proposed method outperforms the AAFD method in this example.

To comprehensively compare our proposed method with the AAFD method in Xu (2021), we test their diagnosis performances on this system with different fault settings. Fix the radius of two actuator faults as 0.05. We choose different fault centers $(f_1^c, f_2^c) \in \mathbb{F} \times \mathbb{F}$, where $\mathbb{F} = \{0.1, 0.2, \ldots, 0.9\}$. Therefore, 9×9 different fault settings in total are considered. Under each fault setting, the AAFD method, our proposed method with the design principle (I), and the design principle (II) are performed to diagnose the fault in $k_{\max} = 10$ time steps. To reduce the influence of uncertainties, 100 tests with ω_k, η_k, and f_i randomly drawn from their given bounded sets are conducted under each setting. Fig. 6 shows the number of successful cases among 100 tests when f_1 is injected, and Fig. 7 demonstrates the results when f_2 is injected. The numerical results show that

- Our method can realize fault diagnosis in more cases than the AAFD method, which can be visualized by the colored dots Figs. 6 and 7.
- The diagnosis performance of the design principle (II) is slightly better than that of the design principle (I) at the expense of increased computation complexity.

In conclusion, our method is more efficient than the AAFD method in diagnosing actuator faults that occur in this system. At the same time, the design principle (I) is preferred since it provides excellent performance with a shorter average computation time than the design principle (II).

In the final part, we use a numerical example to illustrate the performance of the design principle (III). Considering one example in the fault setting $(f_1^c, f_2^c) = (0.5, 0.6)$ and the fault f_2 occurs, we use the design principle (II) to diagnose the fault first. The fault can be identified at the $k = 8$ time step. However, the optimal fault diagnosis observer gains $L_3^{2,FD}$ and $L_5^{2,FD}$ do not ensure stable estimation dynamics, with $\lambda(A - L_3^{2,FD}C) = [-1.0174, 0.9502, 0, 0.9615]$ and $\lambda(A - L_5^{2,FD}C) = [-1.1730, 0.9502, 0, 0.9615]$. To guarantee stable estimation dynamics, the design principle (III) is implemented. The problems (20) have been successfully solved by the BMIBNB solver at each time step, and the fault has been diagnosed at $k = 4$ time step. However, the solving time varies from 2.1 s to 85 s, depending on the required branching iterations in the BNB algorithm. The high computation time makes the design principle (III) inappropriate for online fault diagnosis. Nevertheless, the BNB algorithm fulfills the task of guaranteeing stable estimation dynamics and enhancing system diagnosability.

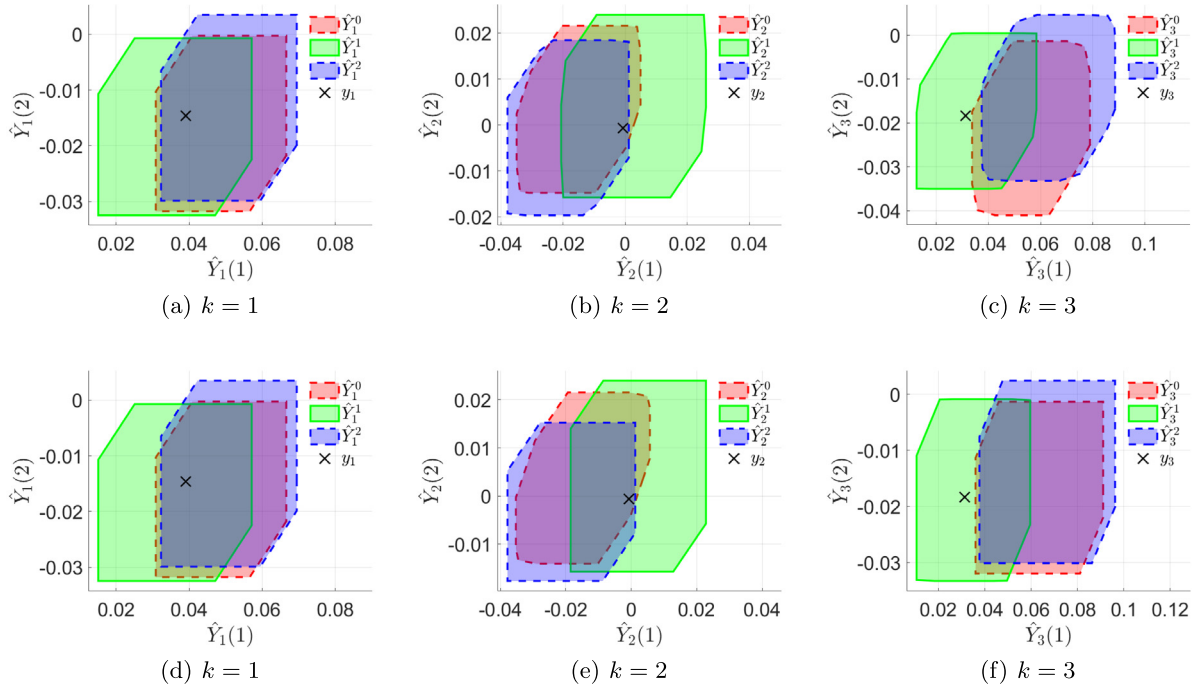


Fig. 4. AFD results using our proposed method. (a)–(c) illustrate the results using the design principle (I) to optimize $L_k^{i,FD}$, while (d)–(f) show the case when the design principle (II) is implemented.

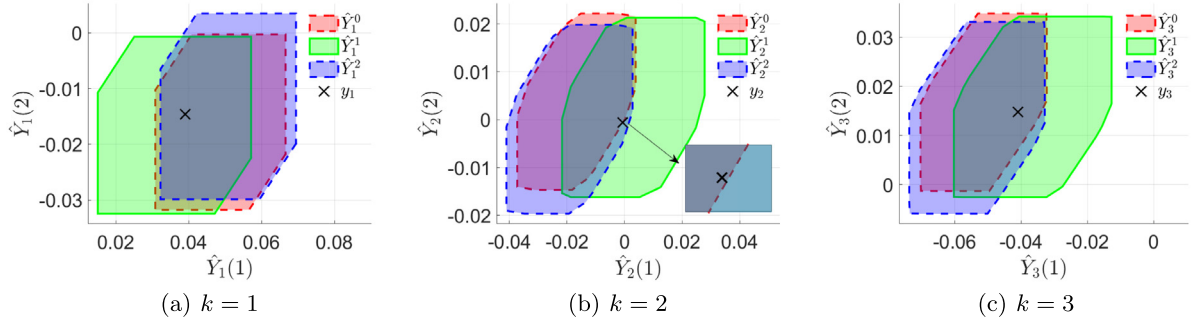


Fig. 5. AFD results using the AAFD method in Xu (2021). At $k = 3$, the output y_k still belongs to the intersection of three OESs.

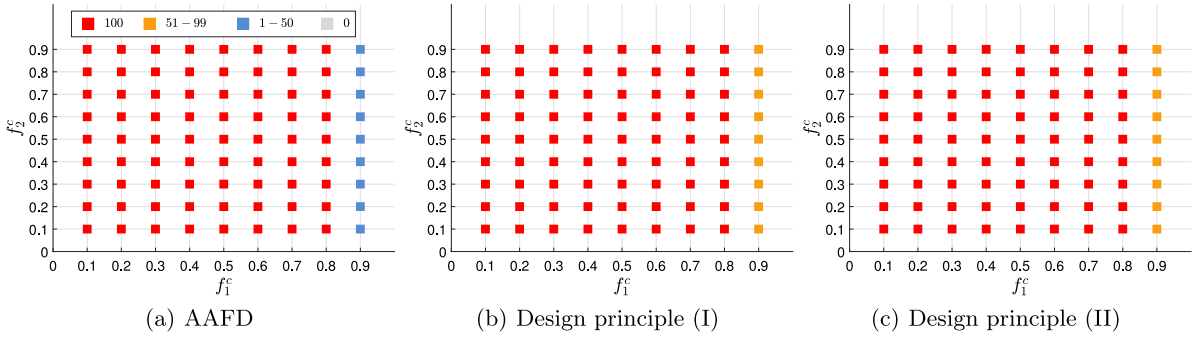
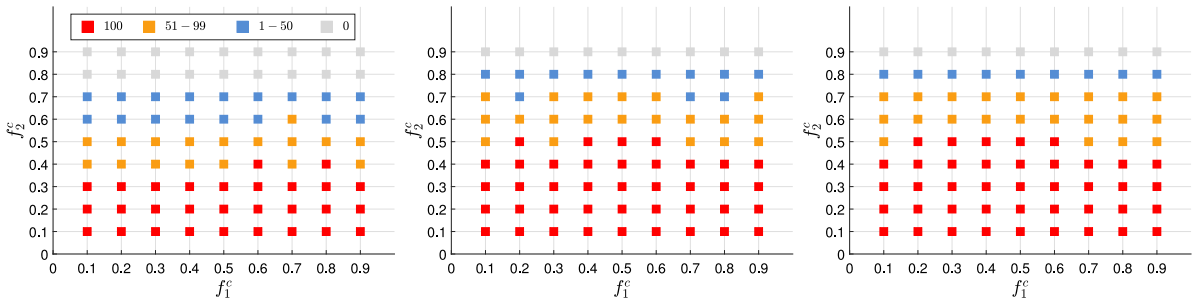


Fig. 6. Illustrations of the diagnosis results under different fault settings when f_1 is injected. The number of successful diagnosis cases in 100 tests is differentiated by colors. (a)–(c) correspond to the AAFD method, our proposed method with the design principle (I) and the design principle (II), respectively. The total numbers of successful diagnosis tests using the AAFD method, our method with the design principles (I) and (II), are 7584, 8017, and 8049, respectively. The average solving times at each time step are 0.00018 s, 0.5283 s, and 0.7354 s, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



(a) AAFD (b) Design principle (I) (c) Design principle (II)

Fig. 7. Illustrations of the diagnosis results under different fault settings when f_2 is injected. The total numbers of successful diagnosis tests using the AAFD method, our method with the design principles (I) and (II), are 4822, 5818, and 5942, respectively. The average solving times at each time step are 0.00017 s, 0.4498 s, and 0.6900 s, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

6. Conclusion

This paper proposes a novel observer-based AFD framework that optimizes the observer gains at each time step to deform OESs so that the system output can be excluded from inconsistent OESs rapidly. The proposed method has three main benefits. First, compared with the AAFD method, the observer gains are designed based on the exclusion tendency and enhance system diagnosability. Second, the computational complexity is acceptable since the original non-convex optimization problem is transformed into a linear form that can be solved efficiently. Third, compared with existing works optimizing observer gains to promote fault diagnosis, the state estimation performance is also improved as a secondary objective. In the future, we will focus on applying our method to more complex systems.

References

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

Blanke, M., Kinnaert, M., Lunze, J., & Staroswiecki, M. (2015). *Diagnosis and fault-tolerant control*. Berlin, Germany: Springer-Verlag.

Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge, U.K.: Cambridge University Press.

Cao, F., Zhang, Z., & He, X. (2022). Active fault isolation of over-actuated systems based on a control allocation approach. *IEEE Transactions on Instrumentation and Measurement*, *71*, 1–10. http://dx.doi.org/10.1109/TIM.2022.3173612.

Chiang, L. H., Russell, E. L., & Braatz, R. D. (2001). *Fault detection and diagnosis in industrial systems*. London, U.K.: Springer London.

Combastel, C. (2015). Zonotopes and Kalman observers: Gain optimality under distinct uncertainty paradigms and robust convergence. *Automatica*, *55*, 265–273.

Efimov, D., Perruquetti, W., Raïssi, T., & Zolghadri, A. (2013). Interval observers for time-varying discrete-time systems. *IEEE Transactions on Automatic Control*, *58*(12), 3218–3224.

Golub, G. H., & Van Loan, C. F. (2013). *Matrix computations*. JHU Press.

Heirung, T. A. N., & Mesbah, A. (2019). Input design for active fault diagnosis. *Annual Reviews in Control*, *47*, 35–50.

Le, V. T. H., Stoica, C., Alamo, T., Camacho, E. F., & Dumur, D. (2013). Zonotopic guaranteed state estimation for uncertain systems. *Automatica*, *49*(11), 3418–3424.

Liberti, L., & Pantelides, C. C. (2006). An exact reformulation algorithm for large nonconvex NLPs involving bilinear terms. *Journal of Global Optimization*, *36*(2), 161–189.

Lofberg, J. (2004). YALMIP: A toolbox for modeling and optimization in MATLAB. In *2004 IEEE international conference on robotics and automation (IEEE cat. no. 04CH37508)* (pp. 284–289). New Orleans, USA: IEEE.

McCormick, G. P. (1976). Computability of global solutions to factorable nonconvex programs: Part I—Convex underestimating problems. *Mathematical Programming*, *10*(1), 147–175.

Nikoukhah, R. (1998). Guaranteed active failure detection and isolation for linear dynamical systems. *Automatica*, *34*(11), 1345–1358.

Pourasghar, M., Combastel, C., Puig, V., & Ocampo-Martinez, C. (2019). FD-ZKF: A zonotopic Kalman filter optimizing fault detection rather than state estimation. *Journal of Process Control*, *73*, 89–102.

Pourasghar, M., Puig, V., & Ocampo-Martinez, C. (2019). Interval observer-based fault detectability analysis using mixed set-invariance theory and sensitivity analysis approach. *International Journal of Systems Science*, *50*(3), 495–516.

Puncochar, I., & Skach, J. (2018). A survey of active fault diagnosis methods. *IFAC-PapersOnLine*, *51*(24), 1091–1098.

Raimondo, D. M., Roberto Marseglia, G., Braatz, R. D., & Scott, J. K. (2016). Closed-loop input design for guaranteed fault diagnosis using set-valued observers. *Automatica*, *74*, 107–117.

Scott, J. K., Findeisen, R., Braatz, R. D., & Raimondo, D. M. (2014). Input design for guaranteed fault diagnosis using zonotopes. *Automatica*, *50*(6), 1580–1589.

Tan, J., Olaru, S., Xu, F., & Wang, X. (2022). Towards a convex design framework for online active fault diagnosis of LPV systems. *IEEE Transactions on Automatic Control*, *67*(8), 4154–4161.

Tan, J., Olaru, S., Xu, F., Wang, X., & Liang, B. (2020). Optimal robust fault detection of discrete-time LPV systems with measurement error-affected scheduling variables combining ZKF and pQP. *International Journal of Robust and Nonlinear Control*, *30*(16), 6782–6802.

Xu, F. (2021). Observer-based asymptotic active fault diagnosis: A two-layer optimization framework. *Automatica*, *128*, Article 109558.

Xu, F. (2023). Minimal detectable and isolable faults of active fault diagnosis. *IEEE Transactions on Automatic Control*, *68*(2), 1138–1145.

Xu, F., Wan, Y., & Wang, Y. (2022). Optimal fault detection observer design using excluding degree. In *Proceedings of the 61st IEEE conference on decision and control* (pp. 1560–1567). Cancún, Mexico.

Yang, S., Xu, F., Wang, X., & Liang, B. (2020). A novel online active fault diagnosis method based on invariant sets. *IEEE Control Systems Letters*, *5*(2), 457–462.

Zhang, Z., & Yang, G. (2020). Distributed fault detection and isolation for multiagent systems: An interval observer approach. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, *50*(6), 2220–2230.

Yidian Fan received his Bachelor's degree in Automation from Beihang University, Beijing, P.R. China, in July 2020. Currently, he is a master student in electronic information engineering at Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen, P.R. China. His research interests include fault diagnosis and state estimation.

Feng Xu received his Bachelor's degree in Measurement and Control Technology & Instrumentation from Northwestern Polytechnical University, Xi'an, P.R. China, in July 2010. In December 2014, he obtained his Ph.D. degree with honor in Automática, Robótica y Visión from Universitat Politècnica de Catalunya (UPC), Barcelona, Spain. During his Ph.D. period, he was also a jointly-trained Ph.D. student at Centrale-Supélec, Paris, France. From June 2015 to July 2017, he was a Postdoctoral Researcher in Control Science & Engineering, Tsinghua University. Currently, he is an Assistant Professor with Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen, P.R. China. His research interests include fault diagnosis, state estimation and fault-tolerant control.

Xueqian Wang received his Master's and Ph.D. degrees in Control Science and Engineering both from Harbin Institute of Technology (HIT), Harbin, P.R. China, in 2005 and 2010, respectively. From June 2010 to February 2014, he was a Postdoctoral Researcher at HIT. Since March 2014, he is currently a Professor with Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen, P.R. China. His research interests include Modeling and Teleoperation of Robotic Systems.

Bin Liang received his Bachelor's degree and Master's degree both from the Honors College, Northwestern Polytechnical University, Xi'an, P.R. China, in 1989 and 1991, respectively, and his Ph.D. degree from the Department of Precision Instrument, Tsinghua University, Beijing, P.R. China, in 1994. Since 2007, he is currently a Professor in the Department of Automation, Tsinghua University. His research interests include Design and Modeling of Robotic Systems.