Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Stackelberg Differential Game-based Optimal Fault Estimation and Accommodation for Continuous-time Linear Systems

Yuhang Xu [*,**] **Hao Yang** [*] **Bin Jiang** [*]
Marios M. Polycarpou [**]

[*] *The College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China. (e-mail: xuyuhang@nuaa.edu.cn; haoyang@nuaa.edu.cn; binjiang@nuaa.edu.cn).*
[**] *KIOS Research and Innovation Center of Excellence and the Dept. of Electrical and Computer Engineering, University of Cyprus, Nicosia, 1678, Cyprus (e-mail: mpolycar@ucy.ac.cy).*

Abstract: This paper deals with the optimal fault estimation and accommodation problem for a class of linear systems in the framework of Stackelberg differential game theory. In this framework, the observer plays the role of the follower, while the system plays the role of the leader in making sequential decisions. A dual controller approach is used to design an auxiliary controller for the observer such that it can non-cooperate with the controller of the system to achieve the Stackelberg equilibrium. To achieve the online updating of the fault-tolerant controller, an adaptive dynamic programming methodology is used by establishing two critic neural networks for the observer and system respectively. Finally, a simulation is presented to illustrate the efficiency and applicability of the theoretical results.

Keywords: Optimal fault estimation, optimal fault-tolerant control, Stackelberg differential game, adaptive dynamic programming.

1. INTRODUCTION

With growing demand for increased safety and reliability, fault-tolerant control has attracted significant attention from various fields, such as industrial production (Ding (2008)). Since an undetected incipient fault may result in a serious disaster, the development of more reliable fault-tolerant control (FTC) methods to guarantee the stability and performance of the system after faults occur is a significant and challenging issue (Zhang and Jiang (2008)).

Existing FTC methods can be classified into two categories: (i) passive FTC and (ii) active one. The main idea of passive FTC is employing only one controller with fixed structures and fixed parameters to deal with both healthy and faulty situations. Active FTC is based on reconfiguration of the controller in response to the detection of occurred faults (Wen et al. (2018)). In contrast to passive FTC, the active FTC approach has some key advantages such as being less conservative and more effective due to its flexible structure and adaptive parameters (Chen et al. (2016); Zhang et al. (2004)). It is worth mentioning that few of the existing results address the optimal performance of the system under the faulty case (Wang et al. (2018)), which motivates the research presented in this paper through the use of the game theory.

Game theory is an interdisciplinary subject initiated from mathematics, which has infiltrated into various disciplines, such as economics and control (Basar and Olsder (1999)). The players in a game theoretic framework may cooperate or non-cooperate with each other to reach the Nash equilibrium, which is a state that none of the players would like to leave. One of the typical games is the Stackelberg differential game, where there are two players: leader and follower. The key characteristic of the Stackelberg differential game is that the leader is one-step ahead of the follower. The follower takes an action to response the strategy of the leader after it gets the information of the leader. Meanwhile, the leader is able to know this situation, therefore, it can select its optimal control to deal with the follower's action. As a result of this characteristic, Stackelberg differential game takes effect in many areas of our daily life, such as the smart grid (Maharjan et al. (2013)). However, how to apply the Stackelberg differential game into solving the active FTC problem still an open and challenging problem.

[⋆] This work is supported by National Natural Science Foundation of China (61773201), Fundamental Research Funds for the Central Universities (NC2020002, NP2020103), Priority Academic Program Development of Jiangsu Higher Education Institutions, the 111 Project (B20007); European Union's Horizon 2020 Research and Innovation Program under the grant agreement No. 739551 (KIOS CoE); Funding for Outstanding Doctoral Dissertation in NUAA (BCXJ19-04), Nanjing University of Aeronautics and Astronautics PhD short-term visiting scholar project (190404DF03). *Corresponding author: Bin Jiang.*

Inspired by the above challenges, this paper deals with the optimal fault estimation (FE) and fault accommodation (FA) problem for a class of linear systems by making use of the Stackelberg differential game. To the best of our knowledge, there are no existing results on the use of Stackelberg differential game theory to solve this problem. Different from the previous results, we cast the optimal FE and FA problem into a Stackelberg differential game to achieve sequential decisions between the observer and system. The main contributions are summarized as follows.

- A Stackelberg differential game-based optimal FE and FA approach is proposed, in which the observer plays the role of the the follower and the system plays the role of the leader to obtain the optimal state/fault estimation and the optimal fault-tolerant controller, respectively;
- An auxiliary controller variable is designed for the observer such that it can non-cooperate with the controller of the system to construct the complete game situation and further to achieve the Stackelberg equilibrium;
- Utilizing the adaptive dynamic programming (ADP) approach, two new adaptive updating laws of the critic weights are designed for the system and observer respectively.

Notation: x^T (or A^T) represents the transposition of a vector x (or a matrix A), $|\cdot|$ denotes the Euclidean norm, $0_{\bar{n}\times\bar{m}} \in R^{\bar{n}\times\bar{m}}$ expresses a matrix with all elements are zero, $I_{\bar{p}\times\bar{q}} \in R^{\bar{p}\times\bar{q}}$ means a matrix with diagonal elements are one and the other elements are zero, $I_{\bar{k}} \in R^{\bar{k}}$ expresses a column vector with all elements are one.

2. PROBLEM FORMULATION

2.1 System dynamics

Consider a class of continuous linear systems described as

$$\begin{aligned} \dot{z} &= \bar{A}z + \bar{B}u + \bar{E}f \\ y &= \bar{C}z \end{aligned} \tag{1}$$

where $z \in R^n$, $u \in R^q$, $f \in R^d$ and $\bar{y} \in R^p$ are the state, control input, fault function, and output of the system, respectively. The matrices $\bar{A}$, $\bar{B}$, $\bar{C}$, and $\bar{E}$ are known constant matrices of appropriate dimensions.

The fault signal $f(t)$ is generated from a dynamical exosystem:

$$\begin{aligned} \dot{x}_f &= \bar{A}_f x_f, \quad t \geq t_f \\ f &= \bar{C}_f x_f, \end{aligned} \tag{2}$$

where x_f is the fault-state with the initial fault value $x_f(t_f) = x_{f0}$, x_{f0} is an unknown value and t_f is the unknown time that the fault occurs. The matrix $\bar{A}_f \in R^{d\times d}$ is assumed to be stable, $\bar{C}_f \in R^{p\times d}$ is the output matrix with full row rank. This fault model investigated in Chen and Patton (2012) can describe various kinds of faults, such as actuator faults, sensor faults and process faults (Abfalg and Allgower (2006)).

For well-posedness, we make the following assumption.

Assumption 1. The pair $(\bar{A}, \bar{B})$ is controllable, the pairs $(\bar{A}, \bar{C})$, $(\bar{A}_f, \bar{C}_f)$ are observable, and all the eigenvalues λ_f of the matrix $\bar{A}_f$ make the matrix $\begin{bmatrix} \lambda_f I - \bar{A} & -\bar{E} \\ \bar{C} & 0 \end{bmatrix}$ have full column rank. □

2.2 The Stackelberg differential game

The definitions of the best response and the Stackelberg equilibrium are given.

Definition 1. In the game theory, the best response is the optimal strategy of a player, which takes account of the other players' strategies to minimize its own cost. □

Definition 2. Suppose P_1 and P_2 are the leader and follower in a Stackelberg differential game. The cost functions and the strategies for leader and follower are denoted as Γ_i and μ_i, $\forall i = 1, 2$, respectively. The pair $\{\mu_1^*, \mu_2^*(\mu_1^*)\}$ is defined as a Stackelberg equilibrium strategy if the following properties are satisfied:

(1) For each μ_1, there exists $\mu_2^*(\mu_1)$ for the follower such that
$$\Gamma_2\left(\mu_1, \mu_2{}^*(\mu_1)\right) \leq \Gamma_2\left(\mu_1, \mu_2\right), \forall \mu_1;$$
(2) There exists μ_1^* on the best responses of the leader such that
$$\Gamma_1\left(\mu_1^*, \mu_2^*\left(\mu_1^*\right)\right) \leq \Gamma_1\left(\mu_1, \mu_2\left(\mu_1\right)\right)$$
for any pair $(\mu_1, \mu_2(\mu_1))$ on the best responses of the leader. □

Based on the above definitions, our objective is to design an optimal active FTC scheme under the framework of the Stackelberg differential game such that the state z of the faulty system (1) converges to zero and the system's cost function is minimized.

To achieve this goal, the bridge connecting the Stackelberg differential game with FE/FA is established in the following section.

3. STACKELBERG DIFFERENTIAL GAME-BASED OPTIMAL FE/FA DESIGN

In this section, the link between the Stackelberg differential game and FE/FA is developed. Then, the designs of the optimal FE and FA are presented.

3.1 Stackelberg differential game-based optimal FE/FA design

To estimate the state and fault simultaneously, the system's dynamic model is augmented by setting $x \triangleq \left[z^T, x_f^T\right]^T$. The augmented system is given by

$$S: \begin{cases} \dot{x} = Ax + Bu \\ y = Cx \end{cases}, \tag{3}$$

where $A \triangleq \begin{bmatrix} \overline{A} & \overline{E}\bar{C}_f \\ 0_{d\times n} & \overline{A}_f \end{bmatrix}$, $B \triangleq \begin{bmatrix} \overline{B} \\ 0_{d\times q} \end{bmatrix}$, $C \triangleq \begin{bmatrix} \bar{C} & 0 \end{bmatrix}$.

Using the duality between the linear optimal observer and the linear quadratic tracking control (Besançon and Munteanu (2015)), an observer is constructed for the system (3) based on the design of an auxiliary controller variable v,

$$O: \begin{cases} \dot{\hat{x}} = A\hat{x} + Bu + Fv \\ \hat{y} = C\hat{x} \end{cases}, \tag{4}$$

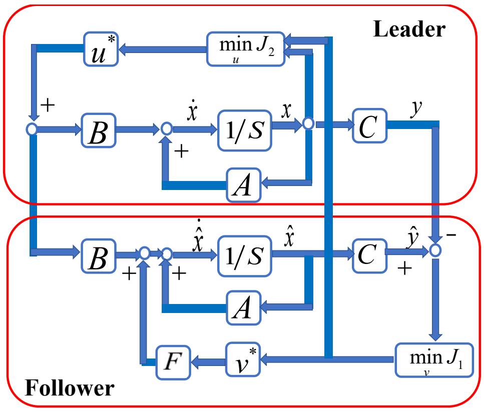

Fig. 1. The Stackelberg differential game between the system (leader) and observer (follower)

where $\hat{x}$ is the estimation of the state x, v is an auxiliary controller variable designed for the observer O, $F \in R^{(n+d)\times q}$ is an input designed matrix, and $\hat{y}$ is the output of the observer. Compared with the traditional observer design, the output y is included into the observer (4) through v.

The links between the augmented system (3) and the observer (4) is illustrated in Fig. 1. From the strategy-making perspective, the system has an initial strategy u_0 working before the auxiliary controller v of the observer works. After the observer receives the information of u and y from the system, the observer will make the best response to the system's current strategy, i.e., designing the optimal auxiliary controller v^* to minimize $e_y \triangleq y - \hat{y}$. In return, the system will adjust its own strategy based on the state and fault estimation from the observer, i.e., designing the matrix u^*, such that the goal of FTC can be achieved. Therefore, the system has the dominant position and has one-step advance than the observer.

Based on the above analysis, it is shown that the decisions between the observer and system can be sequential, in which the system first takes action and then the observer follows based on the strategy of the leader. Such sequential decision feature coincides with the core of the Stackelberg differential game. Thus, we set the system as the leader and the observer as the follower in the Stackelberg differential game to achieve the optimal FE and FA. The optimal fault estimation will be presented in the following section.

3.2 *The follower's design*

In this section, the optimal FE problem for the observer is formulated as the follower's problem in the Stackelberg differential game under an arbitrary strategy of the system, $u \in \mathcal{U}$, where $\mathcal{U}$ is the set of the admissible strategies of the system.

The cost function for the observer (4) is defined as

$$J_1(v) \triangleq \int_0^\infty \left\{ (\hat{y}-y)^T Q (\hat{y}-y) + v^T R v + u^T G v \right\} dt, \tag{5}$$

where Q, R, and G are symmetric positive definite matrices. The term $u^T G v$ represents the impact from the system S to the observer O.

According to the analysis of the sequential decisions between the system and observer, there exists controller signal u_0 of the system in operation which is used by the observer. Thus, we firstly investigate the problem of the follower under an arbitrary strategy of the leader $u \in \mathcal{U}$. More specifically, the problem of the follower is descried by

$$\begin{array}{ll} \min\limits_{v} \ J_1 & \\ s.t.\ (4) & \textit{for an arbitrary } u \in \mathcal{U}. \end{array}$$

To solve this problem, the associated Hamiltonian function is defined as

$$\begin{aligned} H_1 \triangleq\ & (\hat{y}-y)^T Q (\hat{y}-y) + v^T R v + u^T G v \\ & + \nabla J_1^T (A\hat{x} + Bu + Fv), \end{aligned} \tag{6}$$

where $\nabla J_1 \triangleq \partial J_1 / \partial \hat{x}$ denotes the partial derivative of the cost function J_1 with respect to the state estimation $\hat{x}$.

The necessary conditions are given by

$$\nabla \dot{J}_1 = -2C^T QC\hat{x} + 2C^T Qy - A^T \nabla J_1, \tag{7}$$

and

$$\frac{\partial H_1}{\partial v} = 0 \Rightarrow 2Rv + Gu + F^T \nabla J_1 = 0, \tag{8}$$

which implies that

$$v^* = -1/2R^{-1}\left(F^T \nabla J_1^* + Gu\right). \tag{9}$$

The optimal value of the cost function J_1^* satisfies the Hamilton-Jacobi (HJ) equation

$$0 = \min_v H_1(\hat{x}, \nabla J_1^*, v). \tag{10}$$

Substituting control policy (9) into (10) leads to the following coupled HJ equations:

$$\begin{aligned} & (\hat{y}-y)^T Q (\hat{y}-y) - \frac{1}{4}(\nabla J_1^*)^T F R^{-1} F^T \nabla J_1^* \\ & -\frac{1}{4} u^T G R^{-1} G u - \frac{1}{2} \nabla u^T G R^{-1} F^T \nabla J_1^* + (\nabla J_1^*)^T A\hat{x} \\ & + (\nabla J_1^*)^T Bu = 0 \end{aligned} \tag{11}$$

with $J_1^*(0) = 0$.

Theorem 1. Suppose Assumption 1 holds. The optimal auxiliary controller designed in (9) guarantees

$$\lim_{t\to\infty} (x(t) - \hat{x}(t)) = 0.$$

□

Proof. The proof is omitted due to space limitations.

3.3 *The leader's design*

In this section, the optimal FA problem for the system will be solved by designing the leader's problem of the Stackelberg differential game.

The cost function for the system (3) is defined as

$$J_2(u) \triangleq \int_0^\infty (x^T M x + u^T L u + \nabla J_1^T \pi \nabla J_1) dt, \tag{12}$$

where M, L and π are symmetric positive definite matrices.

Under the cost function (12), the problem for the leader is

$$\begin{array}{ll} \min\limits_{u} & J_2(u) \\ s.t. & (3),\quad (7). \end{array}$$

The goal is to minimize the cost function (12) under the constraints of both the system (3) and the observer

(7). The reason to regard (7) as the constraint from the observer is that ∇J_1 is the unique influencing factor of v^* in (9). By considering the constraint (7), the system's optimization contains the best response v^* of the observer, which means that the system adjusts its controller based on the state and fault estimation from the observer to achieve its optimal FTC goal.

To solve the leader's problem, the Hamiltonian function is defined as

$$\begin{aligned} H_2 \triangleq & x^T M x + u^T L u + \nabla J_1^T \pi \nabla J_1 + \nabla J_2^T (Ax + Bu) \\ & + \beta^T \left(-2C^T Q C \hat{x} + 2C^T Q y - A^T \nabla J_1\right), \end{aligned} \tag{13}$$

where $\nabla J_2 \triangleq \partial J_2/\partial x$ and $\beta \triangleq -\partial J_2/\partial(\nabla J_1)$.

The necessary conditions for the leader are given

$$\nabla \dot{J}_2 = -\frac{\partial H_2}{\partial x} \Rightarrow \nabla \dot{J}_2 = -2Mx - A^T \nabla J_2 - 2C^T Q C \beta, \tag{14}$$

$$\dot{\beta} = -\frac{\partial H_2}{\partial(\nabla J_1)} \Rightarrow \dot{\beta} = A^T \beta - 2\pi \nabla J_1, \tag{15}$$

and the optimal fault-tolerant controller can be obtained

$$\frac{\partial H_2}{\partial u} = 0 \Rightarrow u^* = -\frac{1}{2} L^{-1} B^T \nabla J_2^*. \tag{16}$$

Substituting (16) into (13), the coupled HJ equation for the leader is as follows

$$\begin{aligned} & x^T M x - \frac{1}{4} (\nabla J_2^*)^T B L^{-1} B^T \nabla J_2^* + (\nabla J_1^*)^T \pi \nabla J_1^* \\ & + (\nabla J_2^*)^T A x + \beta^T \left(-2C^T Q C \hat{x} + 2C^T Q y - A^T \nabla J_1^*\right) = 0 \end{aligned} \tag{17}$$

with $J_2^*(0) = 0$.

Also, under the controller u^* in (16), the coupled HJ equation (11) for the observer can be further expressed as

$$\begin{aligned} & (\hat{y} - y)^T Q (\hat{y} - y) - \frac{1}{4} (\nabla J_1^*)^T F R^{-1} F^T \nabla J_1^* \\ & - \frac{1}{16} (\nabla J_2^*)^T B L^{-1} G R^{-1} G L^{-1} B^T \nabla J_2^* \\ & + \frac{1}{4} (\nabla J_2^*)^T B L^{-1} G R^{-1} F^T \nabla J_1^* + (\nabla J_1^*)^T A \hat{x} \\ & - \frac{1}{2} (\nabla J_1^*)^T B L^{-1} B^T \nabla J_2^* = 0. \end{aligned} \tag{18}$$

Theorem 2. Under Assumption 1, the optimal fault-tolerant controller u^* designed in (16) can guarantee the faulty system (3) under the cost function (12) to be asymptotically stable, and the pair $\{u^*, v^*(u^*)\}$ reaches the Stackelberg equilibrium if the optimal costs J_1^* and J_2^* satisfy the coupled HJ equations in (17) and (18). □

The Stackelberg equilibrium $\{u^*, v^*(u^*)\}$ in Theorem 2 means that after several rounds of making the best response to each other, the competition between the system and observer arrives at a balance point. At this equilibrium, the observer achieves the optimal state and fault estimation with minimal cost J_1^*, and the system achieves optimal FTC with the minimal cost J_2^*. Moreover, even they are in the equilibrium, they still have interactions with each other such that one of them leaving this Stackelberg equilibrium would result in the loss of the other side. Thus, they have no intensive to leave this equilibrium point.

On the other hand, Theorem 2 reveals that the solutions of the coupled HJ equations in (17) and (18) are the key to achieve the Stackelberg equilibrium. However, the coupled HJ equations are difficult to be solved in practice. To address this problem, the Adaptive Dynamic Programming (ADP) approach (Wang et al. (2018)) will be employed in the following section.

Remark 1. It is worth mentioning that the Stackelberg game considered in this paper is in an open-loop structure. One of our contributions is the development of the open-loop Stackelberg equilibrium solutions to ensure that the closed-loop system can be asymptotically stable. Moreover, this open-loop Stackelberg equilibrium solutions can be written as a state-feedback form, more details can refer to Moon and Başar (2018).

4. ADP-BASED STACKELBERG DIFFERENTIAL GAME STRATEGY

In this section, the Stackelberg differential game strategies are carried out online by using the ADP technology, in which two new adaptive learning laws of critic weights are proposed.

Two critic networks are designed for the system and observer to approximate the cost functions J_1^* and J_2^*, respectively

$$J_1^* = W_{c1}^T \sigma_{c1}(\bar{x}) + \varepsilon_{c1}(\bar{x}), \tag{19}$$

$$J_2^* = W_{c2}^T \sigma_{c2}(x) + \varepsilon_{c2}(x), \tag{20}$$

where $\bar{x} \triangleq \left[\, \hat{x}^T \; y^T \,\right]^T$, $W_{ci} \in R^{\Gamma_i}$, $\forall i = 1, 2$, represents the ideal critic weights, Γ_i is the number of neurons belonging to the hidden layers, σ_{c_i} is the actuation function, and ε_{c_i} denotes the error of the approximation, which converges to zero as $\Gamma_i \to \infty$.

Note that the dynamics of the system may not available in practice, the state estimation is employed into the system. Therefore, a hidden layer is added into the critic network of the system (20), which is expressed as

$$J_2^* = W_{c2}^T \sigma_{c2}\left(w_{c2}^T \hat{x}\right) + \varepsilon_{c2}(\hat{x}), \tag{21}$$

where w_{c2} is the ideal weight of the hidden layer, which assumes to be fixed during the training process.

The derivatives of J_1^* and J_2^* are obtained

$$\nabla J_1^* = \nabla \sigma_{c1}^T W_{c1} + \nabla \varepsilon_{c1}, \tag{22}$$

$$\nabla J_2^* = \nabla \sigma_{c2}^T W_{c2} + \nabla \varepsilon_{c2}, \tag{23}$$

where $\nabla J_1^* \triangleq \partial J_1^*/\partial \hat{x}$, $\nabla J_2^* \triangleq \partial J_2^*/\partial \hat{x}$, $\nabla \sigma_{c1} \triangleq \partial \sigma_{c1}/\partial \hat{x}$, $\nabla \sigma_{c1} \triangleq w_{c2}(\partial \sigma_{c2}(\alpha)/\partial \hat{x})$, $\alpha \triangleq w_{c2}^T \hat{x}$, $\nabla \varepsilon_{c1} \triangleq \partial \varepsilon_{c1}/\partial \hat{x}$ and $\nabla \varepsilon_{c2} \triangleq \partial \varepsilon_{c2}/\partial \hat{x}$.

Using (22) and (23), the Stackelberg equilibrium solutions (u^*, v^*) in (9) and (16) can be expressed as

$$v^* = -\frac{1}{2} R^{-1} \left(F^T \left(\nabla \sigma_{c1}^T W_{c1} + \nabla \varepsilon_{c1}\right) + G u^*\right), \tag{24}$$

$$u^* = -\frac{1}{2} L^{-1} B^T \left(\nabla \sigma_{c2}^T W_{c2} + \nabla \varepsilon_{c2}\right). \tag{25}$$

Applying (24) and (25) into the Hamiltonian functions (6) and (13), it can obtained that

$$
\begin{aligned}
&(\hat{y}-y)^T Q(\hat{y}-y)-\frac{1}{4}W_{c1}^T\nabla\sigma_{c1}FR^{-1}F^T\nabla\sigma_{c1}^TW_{c1}\\
&-\frac{1}{16}W_{c2}^T\nabla\sigma_{c2}BL^{-1}GR^{-1}GL^{-1}B^T\nabla\sigma_{c2}^TW_{c2}\\
&+\frac{1}{4}W_{c2}^T\nabla\sigma_{c2}BL^{-1}GR^{-1}F^T\nabla\sigma_{c1}^TW_{c1}\\
&+W_{c1}^T\nabla\sigma_{c1}A\hat{x}-\frac{1}{2}W_{c1}^T\nabla\sigma_{c1}BL^{-1}B^T\nabla\sigma_{c2}^TW_{c2}\\
=&\xi_{c1},
\end{aligned}
\tag{26}
$$

$$
\begin{aligned}
&\hat{x}^TM\hat{x}-\frac{1}{4}W_{c2}^T\nabla\sigma_{c2}BL^{-1}B^T\nabla\sigma_{c2}^TW_{c2}\\
&+W_{c2}^T\nabla\sigma_{c2}A\hat{x}+\beta^T\left(-2C^TQC\hat{x}+2C^TQy\right.\\
&\left.-A^T\left(\nabla\sigma_{c1}^TW_{c1}\right)\right)+W_{c1}^T\nabla\sigma_{c1}\pi\nabla\sigma_{c1}^TW_{c1}\\
=&\xi_{c2}.
\end{aligned}
\tag{27}
$$

where $\xi_{c1} \triangleq -\nabla\varepsilon_{c1}^T(Ax+Bu^*+Fv^*)$, $\xi_{c2} \triangleq -\nabla\varepsilon_{c2}^T(Ax+Bu^*)+\beta^TA^T\nabla\varepsilon_{c1}$ represent the residual errors. With the growth of the number of hidden layers, it has $|\xi_{ci}| \le l_{\xi ci}$.

To solve the problem of the unknown ideal weight W_{ci}, two adaptive critic networks are employed to estimate J_i as $\hat{J}_i = \hat{W}_{ci}^T\sigma_{ci}$, where $\hat{W}_{c1}$, $\hat{W}_{c2}$ are the estimations of W_{c1}, W_{c2}, respectively. Then, the estimations of u^* and v^* can be expressed as

$$\hat{v} = -\frac{1}{2}R^{-1}\left(F^T\left(\nabla\sigma_{c1}^T\hat{W}_{c1}\right)+G\hat{u}\right), \tag{28}$$

$$\hat{u} = -\frac{1}{2}L^{-1}B^T\nabla\sigma_{c2}^T\hat{W}_{c2}. \tag{29}$$

Hence, the estimated Hamiltonian functions can be represented as

$$
\begin{aligned}
&H_1\left(\hat{x},y,\nabla\hat{J}_1,\hat{v}\right)\\
=&(\hat{y}-y)^TQ(\hat{y}-y)-\frac{1}{4}\hat{W}_{c1}^T\nabla\sigma_{c1}FR^{-1}F^T\nabla\sigma_{c1}^T\hat{W}_{c1}\\
&-\frac{1}{16}\hat{W}_{c2}^T\nabla\sigma_{c2}BL^{-1}GR^{-1}GL^{-1}B^T\nabla\sigma_{c2}^T\hat{W}_{c2}\\
&+\frac{1}{4}\hat{W}_{c2}^T\nabla\sigma_{c2}BL^{-1}GR^{-1}F^T\nabla\sigma_{c1}^T\hat{W}_{c1}\\
&+\hat{W}_{c1}^T\nabla\sigma_{c1}A\hat{x}-\frac{1}{2}\hat{W}_{c1}^T\nabla\sigma_{c1}BL^{-1}B^T\nabla\sigma_{c2}^T\hat{W}_{c2}\\
=&e_{c1},
\end{aligned}
\tag{30}
$$

$$
\begin{aligned}
&H_2\left(\hat{x},\nabla\hat{J}_2,\hat{u}\right)\\
=&\hat{x}^TM\hat{x}-\frac{1}{4}\hat{W}_{c2}^T\nabla\sigma_{c2}BL^{-1}B^T\nabla\sigma_{c2}^T\hat{W}_{c2}\\
&+\hat{W}_{c2}^T\nabla\sigma_{c2}A\hat{x}+\beta^T\left(-2C^TQC\hat{x}+2C^TQy\right.\\
&\left.-A^T\left(\nabla\sigma_{c1}^T\hat{W}_{c1}\right)\right)+\hat{W}_{c1}^T\nabla\sigma_{c1}\pi\nabla\sigma_{c1}^T\hat{W}_{c1}\\
=&e_{c2}.
\end{aligned}
\tag{31}
$$

In order to minimize the errors in (30) and (31), the following functions are introduced, $\forall i = 1,2$,

$$E_{ci} \triangleq \frac{1}{2}e_{ci}^Te_{ci}.$$

Taking advantage of the the gradient descent method, two adaptive weights of the critic networks are proposed

$$
\begin{aligned}
\dot{\hat{W}}_{c1} =& -a_1\frac{\psi_1}{k_1}e_{c1}+a_1\left[\frac{\psi_1}{4k_1}\hat{W}_{c1}^T\nabla\sigma_{c1}FR^{-1}F^T\nabla\sigma_{c1}^T\hat{W}_{c1}\right.\\
&-\frac{\psi_1}{16k_1}\hat{W}_{c2}^T\nabla\sigma_{c2}BL^{-1}GR^{-1}GL^{-1}B^T\nabla\sigma_{c2}^T\hat{W}_{c2}\\
&\left.-\left(F_{12}\hat{W}_{c1}-F_{11}\psi_1^T\hat{W}_{c1}\right)\right],
\end{aligned}
\tag{32}
$$

$$
\begin{aligned}
\dot{\hat{W}}_{c2} =& -a_2\frac{\psi_2}{k_2}e_{c2}+a_2\left[\frac{\psi_2}{4k_2}\hat{W}_{c2}^T\nabla\sigma_{c2}BL^{-1}B^T\nabla\sigma_{c2}^T\hat{W}_{c2}\right.\\
&+\frac{\psi_2}{k_2}\hat{W}_{c1}^T\nabla\sigma_{c1}\pi\nabla\sigma_{c1}^T\hat{W}_{c1}\\
&\left.-\left(F_{22}\hat{W}_{c2}-F_{21}\psi_2^T\hat{W}_{c2}\right)\right].
\end{aligned}
\tag{33}
$$

where a_i represents the learning rates, $k_i \triangleq 1+\theta_i^T\theta_i$ and $\psi_i \triangleq \frac{\theta_i}{k_i}$, $\theta_1 \triangleq \nabla\sigma_{c1}(A\hat{x}+B\hat{u}+F\hat{v})$, $\theta_2 \triangleq \nabla\sigma_{c2}(A\hat{x}+B\hat{u})$. The terms F_{i1} and F_{i2} are two matrices to be designed.

The following general assumption in the neural network is given.

Assumption 2. The critic weights W_{ci}, $\forall i = 1,2$ are bounded, i.e., $|W_{ci}| \le W_{i\max}$ for $W_{i\max} > 0$. Similarly, $\nabla\varepsilon_{ci}$, $\nabla\sigma_{ci}$, β, A, B, G and F are upper bounded by positive constants, i.e., $|\nabla\varepsilon_{ci}| \le l_{\varepsilon i}$, $|\nabla\sigma_{ci}| \le l_{\sigma i}$, $|A| \le l_A$, $|B| \le l_B$, $|G| \le l_G$, and $|F| \le l_F$. □

Theorem 3. Under Assumptions 1-2, the approximate optimal auxiliary controller (28) with the adaptive weight updating law (32), and the approximate optimal fault-tolerant controller (29) with adaptive weight updating law (33) can achieve that the states of the faulty system (3), the states of the observer (4), and the estimation errors of the critic weights are all uniformly ultimately bounded. □

Remark 2. The advantages of the proposed method are concluded: By using the proposed Stackelberg differential game, the relation between the system and observer is revealed from a sequential decision perspective, which also helps to achieve the optimal FE and fault-tolerant controller for the system. Moreover, the solutions are approximated by using the ADP method, which avoids the problem of dimension disaster.

5. SIMULATION

In this section, an example of continuous-time linear system is given to verify the effectiveness of the proposed method. Consider the following faulty system, in which only parts of the states are measureable,

$$
\begin{aligned}
\dot{z} &= \begin{bmatrix} -0.5 & -1.5 \\ 2 & -3 \end{bmatrix} z + \begin{bmatrix} 5 \\ 1 \end{bmatrix} u + \begin{bmatrix} 5 & 3 \\ 0 & 5 \end{bmatrix} f,\\
\bar{y} &= \begin{bmatrix} 1 & 0 \end{bmatrix} z,
\end{aligned}
$$

where $z \triangleq [z_1, z_2]^T$ is the state of system.

The fault is supposed to be

$$
\begin{aligned}
\dot{x}_f &= \begin{bmatrix} -1 & 0.5 \\ 0 & -1.5 \end{bmatrix} x_f,\\
f &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x_f,
\end{aligned}
$$

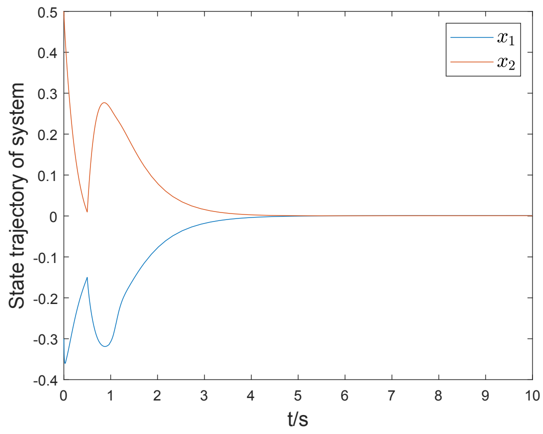

Fig. 2. The state trajectory of system

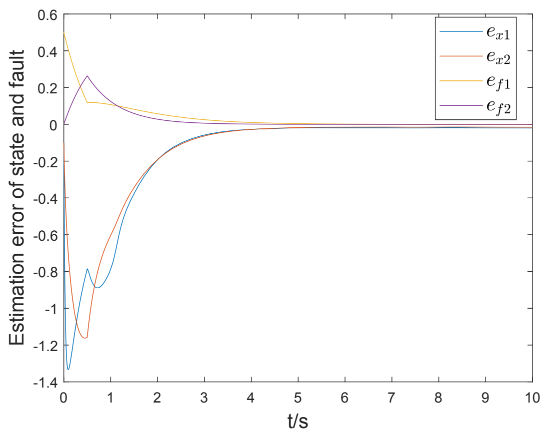

Fig. 3. The estimation errors of state and fault

where $x_f \triangleq [x_{f1}, x_{f2}]^T$. Assume that the fault takes place at $t_f = 0.5s$. Choose the parameters in the cost function as

$$F = B, Q = 5, R = 2, G = 1.5, L = 2, M = 3I_{4\times4}.$$

The activation functions are selected as

$$\begin{aligned}\sigma_{c1}(\hat{y}, y) =& [y \quad y^2 \quad y\hat{y} \quad \hat{y}^2 \quad \hat{y}],\\ \sigma_{c1}(z, x_f) =& [z_1^2 \quad z_1 z_2 \quad z_1 x_{f1} \quad z_1 x_{f2} \quad z_2^2 \quad z_2 x_{f1} \quad z_2 x_{f2}\\ & x_{f1}^2 x_{f1} x_{f2} \quad x_{f2}^2 \quad z_1^4 \quad z_2^4 \quad x_{f1}^4 \quad x_{f2}^4 \quad z_1^3 z_2\\ & z_1^3 x_{f2} z_2^3 z_1 \quad z_2^3 x_{f1} \quad z_2^3 x_{f2} \quad x_{f2}^3 z_1 \quad x_{f2}^3 z_2\\ & x_{f2}^3 x_{f1} \quad z_1^2 z_2^2 \quad z_1^2 x_{f1}^2 \quad z_2^2 x_{f1} \quad z_1^3 x_{f1}].\end{aligned}$$

The learning rates are set as $a_1 = 0.1$, $a_2 = 0.8$. The parameters $F_{11} = 190I_5$, $F_{21} = 20I_5$, $F_{12} = 0.01I_{5\times5}$, $F_{22} = 0.001I_{5\times5}$. In addition, the persistent excitation condition is guaranteed by adding a small probing input signal $r_i(t) = 0.02sin^5(t)cos(t)$, for $t < 3s$.

The simulation results are displayed in Figs. 2-3, which show that the state trajectories for the system and the estimation errors of states/ faults all converge to a small neighborhood of zero. Thus, the effectiveness of the developed Stackelberg differential game-based FTC methods is verified.

6. CONCLUSIONS

In this paper, the Stackelberg differential game is extended to solve the optimal FE and FA problem, in which the observer plays the role of the follower and the system plays the role of the leader to achieve the Stackelberg equilibrium. Moreover, the ADP approach is used to make the fault-tolerant controller update online, in which two new adaptive laws of the critic weights are proposed. The simulation results have been presented to verify the effectiveness of the proposed methods. Futuer work will focus on how to extend these results into more complex systems, such as nonlinear systems.

REFERENCES

Abfalg, J. and Allgower, F. (2006). Fault diagnosis with structured augmented state models: modeling, analysis, and design. In *Proceedings of the 45th IEEE Conference on Decision and Control*, 1165–1170. IEEE.

Basar, T. and Olsder, G.J. (1999). *Dynamic noncooperative game theory*, volume 23. Siam.

Besançon, G. and Munteanu, I. (2015). Control strategy for state and input observer design. *Systems & Control Letters*, 85, 118–122.

Chen, G., Song, Y., and Lewis, F.L. (2016). Distributed fault-tolerant control of networked uncertain euler–lagrange systems under actuator faults. *IEEE Transactions on Cybernetics*, 47(7), 1706–1718.

Chen, J. and Patton, R.J. (2012). *Robust model-based fault diagnosis for dynamic systems*, volume 3. Springer Science & Business Media.

Ding, S.X. (2008). *Model-based fault diagnosis techniques: design schemes, algorithms, and tools*. Springer Science & Business Media.

Maharjan, S., Zhu, Q., Zhang, Y., Gjessing, S., and Basar, T. (2013). Dependable demand response management in the smart grid: A stackelberg game approach. *IEEE Transactions on Smart Grid*, 4(1), 120–132.

Moon, J. and Başar, T. (2018). Linear quadratic mean field stackelberg differential games. *Automatica*, 97, 200–213.

Wang, Z., Liu, L., Wu, Y., and Zhang, H. (2018). Optimal fault-tolerant control for discrete-time nonlinear strict-feedback systems based on adaptive critic design. *IEEE Transactions on Neural Networks and Learning Systems*, 29(6), 2179–2191.

Wen, L., Tao, G., Yang, H., and Jiang, B. (2018). Adaptive actuator failure compensation for possibly nonminimum-phase systems using control separation based lq design. *IEEE Transactions on Automatic Control*, 64(1), 143–158.

Zhang, X., Parisini, T., and Polycarpou, M.M. (2004). Adaptive fault-tolerant control of nonlinear uncertain systems: an information-based diagnostic approach. *IEEE Transactions on Automatic Control*, 49(8), 1259–1274.

Zhang, Y. and Jiang, J. (2008). Bibliographical review on reconfigurable fault-tolerant control systems. *Annual Reviews in Control*, 32(2), 229–252.