Online fault compensation control based on policy iteration algorithm for a class of affine non-linear systems with actuator failures

Bo Zhao\textsuperscript{1}, Derong Liu\textsuperscript{2} and Yuanchun Li\textsuperscript{3}

\textsuperscript{1}State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, People's Republic of China
\textsuperscript{2}School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, People's Republic of China
\textsuperscript{3}Department of Control Science and Engineering, Changchun University of Technology, Changchun 130012, People's Republic of China
\textsuperscript{E-mail: derong@ustb.edu.cn}

Abstract: In this study, a novel online fault compensation control scheme based on policy iteration (PI) algorithm is developed for a class of affine non-linear systems with actuator failures. The control scheme consists of a PI algorithm and a fault compensator. For fault-free dynamic models, the PI algorithm is developed to solve the Hamilton–Jacobi–Bellman equation by constructing a critic neural network, and then the approximate optimal control policy can be derived directly. Alternatively, the actuator failure is reconstructed adaptively to achieve online fault compensation without the fault detection and isolation mechanism. The closed-loop system is proved to be asymptotically stable via Lyapunov's direct method. Two numerical simulation examples are given to demonstrate the effectiveness of the proposed fault compensation control scheme.

1 Introduction

With the fast development of science and technology, industrial applications are becoming increasingly complex and large scale. Consequently, the occurrence of failures is inevitable as the number of components increases. Malfunctions of a component may not only degrade control performance, but also result in the loss of system reliability and safety. It should be emphasised that about 60 \% of control performance degradation in industrial systems is caused by actuator and sensor failures and equipment fouling \cite{1}. To achieve higher reliability and better control performance, a large amount of research efforts on fault tolerant control (FTC) systems have been made during the past 30 years to ensure stability and maintain acceptable control. Among all kinds of possible failures, actuator failures are considered as one of the most critical challenges, mainly for the reason that the control performance can be deteriorated by unexpected and unknown actuator actions.

Many FTC approaches have been developed to deal with actuator failures, such as linear quadratic control, intelligent control, adaptive control and methods based on a combination of different strategies. In general, FTC approaches can be classified into two categories, namely passive approaches and active ones. The main difference between them lies in whether the fault tolerant controller depends on the fault detection and identification (FDI) unit or not. By using passive designs, Zhou and Ren \cite{2} proposed an architecture that included two parts, where the feedback control system was solely controlled by the performance controller, while the model uncertainties and external disturbances were handled by the robustness of the controller. Xiao et al. \cite{3} derived an adaptive sliding mode controller by estimating the bound of actuator faults with an online updating law. Wang et al. \cite{4} investigated a robust fault-tolerant \( H_{\infty} \) control of active suspension systems with finite-frequency constraints. The passive designs achieved insensitivity of systems to certain possible failures by their robustness, so they have the drawback that the designed controllers are often constrained to handling large failures. On the contrary, active FTC, which possesses stronger fault tolerant capability, achieves stability and required performance by tuning control strategies under the decision of an FDI unit. Zhang and Jiang \cite{5}, Hwang et al. \cite{6} gave some excellent reviews on fault reconfiguration methods. Different methods for handling the reconfiguration problem have been reported, such as multiple-model approach \cite{7, 8}, adaptive control approach \cite{9, 10}, linear quadratic control \cite{11, 12}, pseudo-inverse \cite{13, 14}, artificial intelligence \cite{15}, model predictive control \cite{16, 17}, linear matrix inequality \cite{18, 19}, variable structure control \cite{20, 21} and so on. Nazari et al. \cite{22} reconfigured the sensory faulty system by using a virtual sensor which is adapted to the FDI unit. Fault accommodation strategy is another way to achieve the goal of active FTC. Yang and Wang \cite{23} employed a bank of T-S fuzzy model-based FDI observers to describe particular faults, such that one of them can track the current system state, and the corresponding observer estimated error can converge exponentially to zero. Yoo \cite{24} investigated a time-delay independent fault detection and accommodation scheme, where an approximation-based fault accommodation design is activated to compensate for multiple time-delay faults after the fault being detected. Based on the fault compensation technique, Wang and Wen \cite{25} proposed an adaptive failure compensation control scheme with the non-linear damping and parameter projection techniques for parametric strict feedback non-linear systems.

It is well-known that adaptive dynamic programming (ADP) is a powerful approximation tool to solve Hamilton–Jacobi–Bellman (HJB) equations in non-linear systems \cite{26}. In recent years, ADP algorithms were further developed to solve the control problem of continuous-time and discrete-time systems with time delays \cite{27}, external disturbances \cite{28}, and control constraints \cite{29}, as well as for trajectory tracking \cite{30}, coordination control \cite{31} and so on. These algorithms are mainly classified into heuristic dynamic programming (HDP), dual HDP (DHP), action-dependent HDP (ADHDP), ADDHP, globalised HDP (GDHP) and ADGDHP. Iterative methods that can be classified into value iteration (VI) algorithms and policy iteration (PI) algorithms are used in ADP to solve the HJB equation indirectly. Al-Tamimi et al. \cite{32} proved that the iterative performance index function is a non-decreasing sequence and with upper bound, and it converges to the optimal performance index function, which satisfies the HJB equation. Liu et al. \cite{33} investigated a neuro-optimal control scheme for a class of unknown discrete-time non-linear systems with a discount factor in the cost function and GDHP technique. Zhang et al. \cite{34} addressed the infinite-time optimal tracking control problem by using the greedy HDP iteration algorithm. We can conclude from these studies that VI can remove the requirement of the initial stabilising control, but it cannot guarantee the stability of the
system. Actually, only the converged optimal control law can be used to control non-linear systems [35]. In contrast to VI algorithms, the iterative performance index function of PI algorithms converge to the optimum non-increasingly and each of the iterative controls stabilizes the non-linear systems [36, 37]. Abu-Khalaf and Lewis [38] proposed a PI algorithm for continuous-time non-linear systems with control constraints. Liu and Wei [35] proposed a discrete-time PI ADP method for solving the infinite horizon optimal control problem of non-linear systems. Zhang et al. [31] addressed the optimal coordination control for multiagent differential games by solving the coupled Hamilton–Jacobi equations via a PI algorithm.

Several papers considered the fault tolerance problem by using reinforcement learning and ADP strategies. Wang et al. [39] developed a robust state feedback reliable control scheme integrated with an iterative learning. By solving linear matrix inequalities (LMIs), the developed scheme was explicitly formulated together with an adjustable robust $H_{\infty}$ performance level for batch process systems with unknown actuator failures. Feng et al. [40] proposed a reconfigurable fault tolerant deflection routing algorithm based on reinforcement learning for network on chip. An optimised routing algorithm-based hierarchical Q-learning was proposed to reduce the routing table size. He and Shayman [41] developed a reinforcement-learning based fast algorithm for proactive network fault management. The proactive diagnosis information was considered to produce effective monitoring and control policies for intelligent managers or agents. Zhu and Yuan [42] presented a novel approach to automate recovery policy generation with reinforcement learning techniques. It could learn a new and locally optimal policy that outperformed the original one based on the recovery history of the original user-defined policy. Yen and DeLima [43, 44] proposed a supervisor making use of two quality indices to perform FDI and isolation based on GDHP. Although it could reduce the reconfiguration time of the controller, the strategy was implemented under the condition that a priori knowledge was stored in a dynamic model bank.

In this paper, an online fault compensation control scheme-based PI algorithm is established to obtain the optimal control of affine non-linear systems with actuator failures. Due to the occurrence of actuator failures, the PI algorithm may be biased or fail to achieve the optimal control. In order to reduce the degradation caused by faults, a redesigned fault compensation based PI controller is provided. The weight errors of the critic neural network are proved to be uniformly ultimately bounded (UUB), and the stability of the closed-loop system with actuator failures is guaranteed via Lyapunov’s approach. Different from classic ADP algorithms, the action neural network is no longer required in this algorithm, which reduces the computational burden effectively. Meanwhile, the proposed FTC strategy consists of two parts, namely the PI-based optimal control part and the online fault compensation part. In this sense, it can be conveniently implemented to handle fault tolerant problems.

The rest of this paper is organised as follows. In Section 2, we present the problem statement. In Section 3, the PI algorithm for fault-free systems is presented. A fault compensator is developed with the adaptive technique to redesign the FTC, and then stability analysis is given. In Section 4, two examples are provided to demonstrate the effectiveness of the present scheme. In Section 5, the conclusion is drawn.

## 2 Problem statement

Consider the following affine non-linear system with actuator failures:

$$x(t) = f(x(t)) + g(x(t))u(x(t)) - f_{a}(t)$$

(1)

where $x \in \mathbb{R}^{n}$ is the system state vector, $u \in \mathbb{R}^{m}$ is the control input vector, $f(\cdot)$ and $g(\cdot)$ are locally Lipschitz and differentiable in their arguments with $f(0) = 0$, and $f_{a}(t) \in \mathbb{R}^{m}$ is an unknown additive actuator failure. Here, let $x(0) = x_{0}$ be the initial state.

For the system (1) with $f_{a}(t) = 0$ (i.e. the system is fault-free), the performance index function can be defined as

$$J(x_{0}) = \int_{0}^{\infty} U(x(\tau), u(\tau))d\tau$$

(2)

where $U(x, u) = x^{T}Qx + u^{T}Ru$ is the utility function, $U(0, 0) = 0$, and $U(x, u) \geq 0$ for all $x$ and $u$, in which $Q \in \mathbb{R}^{n \times n}$ and $R \in \mathbb{R}^{m \times m}$ are positive definite matrices.

Remark 1: Possible failures that occur on actuators may present many scenarios, such as partial loss of effectiveness, locked in place, saturation and free-swing. They affect the efficiency of actuators, i.e. execution capability of the actuators will change. Thus, the term $u(x)$ in (1) should be changed directly to affect the considered system. In this case, actuator failures can be seen as matched disturbances. However, there exist different physical meanings between them. Generally speaking, disturbances are assumed to be known norm-bounded and inevitable in real applications. On the other hand, actuator failures occur stochastically and are assumed to be unknown bounded functions. Indeed, disturbances will lead the system performance to imprecision, rather than badly destroy the system which may suffer from actuator faults.

To handle the optimal control problem, the designed feedback control must be admissible. Before the algorithm is presented, the definition of admissible control is introduced.

**Definition 1:** For system (1) with $f_{a}(t) = 0$, a control policy $u(x)$ is said to be admissible, if $u(x)$ is continuous on a set $\Omega \subseteq \mathbb{R}^{n}$, $u(0) = 0$, $u(x)$ stabilises the system, and $J(x_{0})$ in (2) is finite for all $x \in \Omega$.

For any admissible control policy $u(x) \in \Psi(\Omega)$, where $\Psi(\Omega)$ denotes the set of admissible control, if the performance index function

$$V(x_{0}) = \int_{0}^{\infty} U(x(\tau), u(\tau))d\tau$$

(3)

is continuously differentiable, then the infinitesimal version of (3) is the Lyapunov equation

$$0 = U(x, \mu) + (\nabla V(x))^T f(x) + g(x)\mu$$

(4)

with $V(0) = 0$, and the term $\nabla V(x)$ denotes the partial derivative of $V(x)$ with respect to $x$, i.e. $\nabla V(x) = \partial V(x)/\partial x$.

Define the Hamiltonian function of the problem and the optimal performance index function as

$$H(x, \mu, \nabla V(x)) = U(x, \mu) + (\nabla V(x))^T f(x) + g(x)\mu$$

and

$$f'(x) = \min_{\mu \in \Psi(\Omega)} \int_{0}^{\infty} U(x(\tau), \mu(\tau))d\tau.$$  

(5)

Let $f'(x)$ be the optimal performance index function, then

$$0 = \min_{\mu \in \Psi(\Omega)} H(x, \mu, \nabla f'(x))$$

(6)

where $\nabla f'(x) = \partial f'(x)/\partial x$. If the solution $f'(x)$ exists and is continuously differentiable, the optimal control can be expressed as

$$u^{*}(x) = -\frac{1}{2}R^{-1}g^{T}(x)\nabla f'(x).$$

(7)

In general, if the system is fault-free (i.e. $f_{a}(t) = 0$), the solution of (6) can be approximated by using the PI technique (see Algorithm 1).
Thus, the Hamiltonian function can be expressed as

\[ H(x, \mu, W_c) = U(x, \mu) + W_c^T \sigma(x) \dot{x} = -\nabla \epsilon_c(x) \dot{x} \pm e_{ch} \]  

where \( e_{ch} \) is the residual error caused by neural network approximation.

Since the ideal weight vector \( W_c \) is unknown, the critic neural network can be approximated by

\[ \hat{V}(x) = W_c^T \sigma(x) \]  

Then, we have the gradient of \( \hat{V}(x) \) as

\[ \nabla \hat{V}(x) = (\nabla \sigma(x))^T W_c + \nabla \epsilon_c(x) \]  

Thus, the approximate Hamiltonian function can be obtained as

\[ \hat{H}(x, \mu, \hat{W}_c) = U(x, \mu) + \hat{W}_c^T \sigma(x) \dot{x} \pm \epsilon_c. \]  

Let \( \theta = \nabla \sigma(x) \dot{x} \), and define the weight approximation error as \( \tilde{W}_c = W_c - \hat{W}_c \). Then, by (11) and (12), we have

\[ e_c = e_{ch} - \tilde{W}_c^T \dot{\theta}. \]  

The weight approximation error can be updated as

\[ \dot{\tilde{W}}_c = l_e (e_{ch} - \tilde{W}_c^T \dot{\theta}). \]  

In order to tune the critic neural network weight vector \( \tilde{W}_c \), the objective function \( E_c = (1/2) e_{ch}^T e_c \) should be minimised with the steepest descent algorithm. \( \tilde{W}_c \) should be updated as

\[ \dot{\tilde{W}}_c = -\tilde{W}_c = -l_e \theta. \]  

where \( l_e > 0 \) is the learning rate of the critic neural network.

Hence, by (7) and (9), the ideal control policy can be obtained as

\[ \mu(x) = -\frac{1}{2} R^{-1} g(x) \left( (\nabla \sigma(x))^T W_c + \nabla \epsilon_c(x) \right). \]  

and it can be approximated as

\[ \hat{\mu}(x) = -\frac{1}{2} R^{-1} g(x) (\nabla \sigma(x))^T \hat{W}_c. \]  

From (14), one can see that the control policy can be derived depending only on the critic neural network, and the training of the action neural network is no longer required.

Theorem 1: For a fault-free system, if the weights of the critic neural network are updated by (13), then the weight approximate error can be guaranteed to be UUB.

Proof: Choose the Lyapunov function candidate as

\[ \frac{1}{2} \nabla \epsilon_c(x) \dot{x} = -\frac{1}{2} R^{-1} g(x) \left( (\nabla \sigma(x))^T W_c + \nabla \epsilon_c(x) \right), \]  

where

\[ \frac{1}{2} \left[ \begin{array}{c} \nabla \epsilon_c(x) \dot{x} \\ -\frac{1}{2} R^{-1} g(x) (\nabla \sigma(x))^T \hat{W}_c \end{array} \right] = -\frac{1}{2} R^{-1} g(x) (\nabla \sigma(x))^T \hat{W}_c. \]  

Therefore, the approximate Hamiltonian function can be expressed as

\[ \hat{H}(x, \mu, \hat{W}_c) = U(x, \mu) + \hat{W}_c^T \sigma(x) \dot{x} \pm \epsilon_c. \]  

where \( \epsilon_c \) is the residual error caused by neural network approximation.
\[ \Sigma_i = \frac{1}{2\tau_i} \hat{W}_i^T \hat{W}_i. \]

Its time derivative along the solution of (13) is

\[ \Sigma_i = \frac{1}{\tau_i} \hat{W}_i^T \dot{\hat{W}}_i \]
\[ = \hat{W}_i^T (e_{iu} - \hat{W}_i \theta) \theta \]
\[ = \hat{W}_i^T e_{iu} \theta - \| \hat{W}_i \theta \|^2 \]
\[ \leq \frac{1}{2} \varepsilon^2 \varepsilon - \frac{1}{2} \| \hat{W}_i \theta \|^2. \]

Assume \( \| \theta \| \leq \theta^M \), where \( \theta^M \) is a positive constant. Hence, \( \Sigma_i < 0 \) whenever \( \hat{W}_i \) lies outside the compact set

\[ \Omega_{W_i} = \left\{ \hat{W}_i : \| \hat{W}_i \| \leq \varepsilon \frac{\theta^M}{\theta^M} \right\} \]

Therefore, according to Lyapunov’s direct method, the weight approximation error is UUB. This completes the proof. \( \Box \)

**Remark 2:** This paper provides an FTC strategy consists of two parts, namely the PI-based control part for the fault-free system and the online fault compensation part for handling the effects of actuator failure. For the PI-based control part, it is developed with the help of the performance index function (3), which is approximated by the critic neural network (9), so Theorem 1 is provided for a fault-free system and it does not need to prove the convergence of the weights of the critic neural network in case of fault.

### 3.3 Online fault compensation design

Based on Section 3.1, we will analyse the system stability in the case of the actuator failure \( f_a \neq 0 \).

Let \( u = \mu^0 \), the system with an actuator failure should be rewritten as

\[ \dot{x} = f(x) + g(x)(\mu^0 - f_a). \]

By using (5), \( f^0(x) > 0 \) for any \( x \neq 0 \) and \( f^0(x) = 0 \) for \( x = 0 \), it implies that \( f^0(x) \) is a positive definite function. Moreover, its time derivative is

\[ f^0(x) = (\nabla f^0(x))^T \dot{x}. \]

Considering the dynamic model of the system with an actuator failure (1), one can obtain

\[ f^0(x) = (\nabla f^0(x))^T (f(x) + g(x)(\mu^0 - f_a)) \]
\[ = (\nabla f^0(x))^T (f(x) + g(x)(\mu^0) - (\nabla f^0(x))^T g(x)f_a). \]

From (6), we have

\[ (\nabla f^0(x))^T (f(x) + g(x)(\mu^0)) = - U(x, \mu^0). \]

Substituting (8) and (16) into (15), one has

\[ f^0(x) = - U(x, \mu^0) + 2(\mu^0)^T R f_a \]
\[ \leq - x^T Q x - (\mu^0)^T R \mu^0 + (\mu^0)^T R \mu^0 + f_a^T R f_a \]
\[ = - x^T Q x + f_a^T R f_a. \]

This implies that \( f^0(x) \) is biased by a term depending on the unknown actuator failure \( f_a \). Therefore, the PI control \( \mu^0 \) cannot guarantee the stability of the closed-loop system due to the unknown actuator failure \( f_a \). Next, a fault compensation based PI algorithm will be established to keep the closed-loop system stable with an actuator failure.

Design the fault compensation-based control law as

\[ u = \mu + \tilde{f}_a \]

where \( \tilde{f}_a \) is the fault compensation term proposed to overcome the performance degradation, and it can be derived from the following adaptive law:

\[ \tilde{f}_a = l_s (2\mu^0 R - x^T g(x))^T. \]

**Remark 3:** As previously mentioned, the FTC approaches are classified into passive approaches and active ones, according to whether an FDD mechanism exists. In the designed fault compensation-based FTC law (17), the fault estimation can be seen as an FDD mechanism. So the proposed FTC scheme can be considered as an active FTC approach.

### 3.4 Stability analysis

**Theorem 2:** Consider the affine non-linear system with an actuator failure (1) and the performance index function (2), and the adaptive law of the actuator failure (18). The fault compensation based PI control law (17) can guarantee the closed-loop system to be asymptotically stable.

**Proof:** Choose the Lyapunov function candidate as

\[ \Sigma = \frac{1}{2} x^T x + J^0(x) + \frac{1}{2\tau_i} \tilde{f}_a^T \tilde{f}_a. \]

Define the fault compensation error \( \tilde{f}_a = f_a - f_a \). Then, by (17), we have (see (19) and (20)) As \( f(x) \) is locally Lipschitz, a positive constant \( D_f \) exists such that \( \| f(x) \| \leq D_f \| x \| \). Assume that \( \| g(x) \| \leq D_g \). Thus, (19) becomes (see (20)) Combining (8) and (16), (20) becomes

\[ \dot{\Sigma} = D_f \| x \|^2 + \frac{1}{2} D_g \| \mu \|^2 + \frac{1}{2} \| x \|^2 - U(x, \mu) \]
\[ + \left( 2\mu^0 R - x^T g(x) - \frac{1}{l_s} \tilde{f}_a \right) \tilde{f}_a. \]

Substituting (18) into (21), one can obtain

\[ \dot{\Sigma} = \left( D_f + \frac{1}{2} \right) \| x \|^2 + \frac{1}{2} D_g \| \mu \|^2 - x^T Q x - \mu^T R \mu \]
\[ \leq - \left( \lambda_{min}(Q) - D_f - \frac{1}{2} \right) \| x \|^2 - \left( \lambda_{min}(R) - \frac{1}{2} D_g \right) \| \mu \|^2. \]
Consider the following affine non-linear system:

\[
\Sigma = \begin{bmatrix}
0 & 0 & \cos(2x_2) + 2 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
-x_1^2 & -0.5x_1 + 0.5x_2 & x_1 & x_2 \\
\end{bmatrix}
\]

(22)

where \( x = [x_1, x_2, x_3, x_4] \in \mathbb{R}^4 \) and \( u = [u_1, u_2] \in \mathbb{R}^2 \) are the state and control input variables, respectively. The term \( f_\alpha = [f_\alpha^1, f_\alpha^2]^T \in \mathbb{R}^2 \) reflects an unknown additive actuator failure. We choose

\[
f_\alpha = \begin{cases} 
[0,0]^T, & 0 \leq t \leq 30 \text{s} \\
[2 + 5\sin(t/2\pi), 0]^T, & 30 \text{s} < t \leq 60 \text{s}
\end{cases}
\]

for the purpose of simulation.

Let the initial state be \( x_0 = [1, 1, -1, -1]^T \), and the initial admissible control policy be \( u = [-0.2, -0.4, -0.6, -0.8]^T \). We employ a critic neural network to approximate the performance index function, and its weight vector is denoted as \( W_c = [W_{c1}, W_{c2}, \ldots, W_{c10}]^T \). The initial weight of critic neural network is \( W_{c0} = [0.4, 0.5, 1.8, 0.2, 1.8, 0.3, 1.2, 0.7, 1.5]^T \).

The activation function of the critic network is chosen as \( \sigma_c(x) = [x_1 \cdot x_2, x_1 \cdot x_3, x_2 \cdot x_3, x_1, x_2, x_3, x_1^2, x_2^2, x_3^2, x_1 \cdot x_2, x_1 \cdot x_3, x_2 \cdot x_3] \). Let \( Q = 3 \times I_6 \), \( R = 3 \times I_2 \), and the learning rate of the critic network and actuator failure be \( \eta_1 = 0.0005 \) and \( \eta_2 = 40 \), respectively. The initial value of the actuator failure vector is chosen as \( f_\alpha^1 = [0, 0]^T \).

The simulation results are displayed in Figs. 2–5. From Fig. 2, we can find that the weight vector of the critic neural network converges to \([0.627751, 0.761803, 1.599625, -0.381745, 2.072987, -0.105430, \ldots, 0.825454, 1.320895, 1.780736, 1.400185]^T \). To demonstrate the effectiveness of the proposed fault compensation controller, the system states under the input without the compensation are shown in Fig. 3. As the unknown actuator fault occurs, the PI control law \( \mu^{(i)} \) cannot guarantee the stability of the system (22). In this sense, it fails to achieve satisfactory control performance. In order to compensate the failure online, the failure should be estimated by (18) adaptively, as shown in Fig. 4. While in Fig. 5, the system states under the online fault compensation based PI control (17) are illustrated. We can see that the system states converge to the equilibrium a short time after the occurrence of the fault, which learns the failure online.

**Example 2:** Consider a 2-degree of freedom industrial manipulator system with an actuator failure in joint space coordination, whose dynamic model can be expressed as

\[
M(q) \ddot{q} + C(q, \dot{q})\dot{q} + G(q) = u - f_\alpha
\]

where \( q \in \mathbb{R}^2 \) is the vector of joint displacements, \( M(q) \in \mathbb{R}^{4 \times 4} \) is the inertia matrix, \( C(q, \dot{q}) \in \mathbb{R}^2 \) is the Coriolis and centripetal force, \( G(q) \in \mathbb{R}^2 \) is the gravity term, \( u \in \mathbb{R}^2 \) is the applied joint torque, \( f_\alpha \) is the additive actuator failure.
Let \( \mathbf{\alpha} = \begin{bmatrix} \alpha_1, \alpha_2, \alpha_3, \alpha_4 \end{bmatrix} \)
\( \mathbf{\dot{\alpha}} = \begin{bmatrix} \dot{\alpha}_1, \dot{\alpha}_2, \dot{\alpha}_3, \dot{\alpha}_4 \end{bmatrix} \in \mathbb{R}^4. \) By simple transformation, the system dynamic model can be expressed as (1) with \( f(x) = [0; 0; \xi(x)] \) and \( g(x) = [0, 0; 0; \eta(x)] \), in which \( \xi(x) = -M^{-1}(q)(C(q, q)\dot{q} + G(q)) \) and \( \eta(x) = M^{-1}(q). \) Choose
\[
\hat{f}_a = \begin{cases} 
[0, 0]^T, & 0 \leq t \leq 30 \\
[0, 8]^T, & 30 < t \leq 60 
\end{cases} 
\]
and other parameters are selected the same as Example 1.

Let \( x = [x_1, x_2, x_3, x_4]^T = [q_1, q_2, q_3, q_4]^T \in \mathbb{R}^4. \) By simple transformation, the system dynamic model can be expressed as (1) with \( f(x) = [0; 0; \xi(x)] \) and \( g(x) = [0, 0; 0; \eta(x)] \), in which \( \xi(x) = -M^{-1}(q)(C(q, q)\dot{q} + G(q)) \) and \( \eta(x) = M^{-1}(q). \) Choose
\[
\hat{f}_a = \begin{cases} 
[0, 0]^T, & 0 \leq t \leq 30 \\
[0, 8]^T, & 30 < t \leq 60 
\end{cases} 
\]
and other parameters are selected the same as Example 1.

Let \( Q = 12 \times I_4, R = 25 \times I_4, \) and other control parameters are same as Example 1. We employ the similar critic neural network to approximate the performance index function, and its initial weight vector is chosen as \( [0.4, 0.5, 1.2, 0.6, 1.8, 0.3, 1.1, 1, 0.7, 1.5]^T \). By using the proposed algorithm, the weights of the critic neural network converge to \( [0.441090, 0.541566, 1.535788, 1.245143, 1.841842, 0.640492, 1.743332, 0.729727, 0.259550, 0.930982]^T \). The simulation results are shown as Figs. 6–9, and from these figures, we can conclude the similar results as Example 1. Therefore, we can declare the effectiveness of the online fault compensation based PI algorithm developed in this paper.

### 5 Conclusion

An online fault compensation control scheme based on PI is developed to solve the FTC problem of a class of affine non-linear systems with actuator failures. The critic neural network, whose weights are updated adaptively, is employed to approximately obtain the performance index function. Based on actuator failures estimated by the adaptive law, the PI controller can be compensated to reduce the influence of the actuator failures. With the Lyapunov’s direct method, the convergence of the closed-loop system with actuator failures is guaranteed. Finally, two examples are given to demonstrate the effectiveness of the developed algorithm.
This work was supported in part by the National Natural Science Foundation of China under Grants 61233001, 61273140, 61304086, 61374105, 61374051 and 61533017, in part by the Early Career Development Award of SKLMCCS, and in part by the Scientific and Technological Development Plan Project in Jilin Province of China under Grants 20150520112JH and 20160414033GH.

References


IET Control Theory Appl., 2016, Vol. 10 Iss. 15, pp. 1816-1823
© The Institution of Engineering and Technology 2016


