Fault-Tolerant Control for the Formation of Multiple Unknown Nonlinear Quadrotors via Reinforcement Learning

Wanbing Zhao∗ Hao Liu∗∗ Frank L. Lewis∗∗∗

∗ Key Laboratory of Spacecraft Design Optimization and Dynamic Simulation Technologies of Ministry of Education, Beihang University, Beijing 100191, P.R. China (e-mail: zhaowb27@buaa.edu.cn).
∗∗ Key Laboratory of Spacecraft Design Optimization and Dynamic Simulation Technologies of Ministry of Education, Beihang University, Beijing 100191, P.R. China (e-mail: liuhao13@buaa.edu.cn)
∗∗∗ University of Texas at Arlington Research Institute, University of Texas at Arlington, Fort Worth, Texas 76118, USA (e-mail: lewis@uta.edu)

Abstract: In this paper, the fault-tolerant control problem for the formation of unknown quadrotor team with nonlinearities, couplings, and actuator faults in the dynamics is investigated. A distributed observer is designed to estimate the position references for each quadrotor. A hierarchical control scheme is constructed including a fault-tolerant position controller to achieve the desired formation and a fault-tolerant attitude controller to track the attitude references. Reinforcement learning algorithms are designed to learn the optimal control policies of the position and attitude controllers. Simulation results are given to illustrate the effectiveness of the proposed controller.

Keywords: Fault-tolerant control, formation control, reinforcement learning, model-free, quadrotor system.

1. INTRODUCTION

As a typical class of UAVs, quadrotors have received much research interest because of their capabilities of hovering, slow flight, and vertical take-off and landing. Furthermore, the cooperation of the multi-agent quadrotor system exhibits special superiority in the complicated task such as telecommunication relay, cooperative load transportation, and wilderness search (see, Dong et al. (2018)). The quadrotors are composed of a rigid frame with four individual rotors, which generate the control torques for the quadrotor by changing the spinning speeds of the four rotors. However, the actuating rotors are prone to faults because of the degradation of the motors or the damage of the propellers (see, e.g., Avram, Zhang, and Muse. (2017)). For the multi-agent quadrotor system, the probability of the system encountering actuator faults is growing due to the increasing team scale and the complexity of the multi-agent system (see, Qin et al. (2017)). The appearance of the actuator faults may result in undesirable trajectory tracking performance, instability of the multi-agent quadrotor system, and even catastrophic consequence. Therefore, it is necessary to address the fault-tolerant control (FTC) problem for the multi-agent quadrotor system under actuator faults.

Recently, the FTC methods have been implemented to the cooperative control of the multi-agent systems. In Deng and Yang. (2017), the distributed adaptive FTC problem was studied for the synchronization of linear multi-agent systems under the effects of actuator faults. In Hua et al. (2016), the distributed FTC problem was investigated to achieve the formation for second-order linear multi-agent systems in the presence of multiple actuator faults. In Yu, Qu, and Zhang. (2019), a distributed fault-tolerant controller was studied to achieve the desired synchronization for multiple cooperative UAVs with nonlinearities and actuator faults in the dynamics. In Shi et al. (2017), a fault-tolerant formation controller involving a decentralized state observer and an adaptive fault estimator was developed for nonlinear quadrotor team to drive each quadrotor to the desired formation pattern. However, in the fault-tolerant controller design process of Deng and Yang. (2017); Hua et al. (2016); Yu, Qu, and Zhang. (2019); Shi et al. (2017), the dynamics of the system was required, which is unpractical in the controller design of quadrotor team. In fact, the quadrotor is a nonlinear system subject to actuator faults and the dynamics of the quadrotor system is difficult to obtain in practical applications. Therefore, it is promising to design a fault-tolerant controller for multiple unknown quadrotors subject to nonlinearities and actuator faults in the quadrotor dynamics.

In the latest years, Reinforcement learning (RL) has become a powerful approach to learn the optimal control
policy and emerged in the fault-tolerant controller design for single-agent systems (see, e.g., Liu, Wang, and Wang (2017); Wang et al. (2016); Zhang et al. (2018); Deptula et al. (2018); Ma, Xu, and Yang (2019)). In Liu, Wang, and Wang (2017); Wang et al. (2016), the RL-based fault-tolerant controllers were developed to tolerate actuator faults for nonlinear multiple-input multiple-output (MIMO) discrete-time systems. In Zhang et al. (2018), a RL-based FTC algorithm was designed by using online policy iteration approach to learn the optimal controller for nonlinear tracking systems. In the RL-based optimal controller design of Deptula et al. (2018), a data-based estimator was constructed to estimate the loss of effectiveness (LOE) actuator faults and then an approximate dynamic programming approach was implemented to learn the optimal control solutions for the system. In Ma, Xu, and Yang. (2019), a series of RL algorithms were implemented in the data-driven FTC design for a single quadrotor to estimate the actuator faults and then compensate the actuator faults. However, the FTC problem for the formation of multiple nonlinear unknown quadrotor systems subject to actuator faults has not been discussed in the literatures mentioned before.

Therefore, the fault-tolerant cooperative control problem for multiple unknown quadrotor systems with nonlinearities and actuator faults in the dynamics remains an open issue. In this paper, a distributed fault-tolerant cooperative controller is designed including a distributed observer, a fault-tolerant position controller, and a fault-tolerant attitude controller. The distributed observer is constructed to estimate the desired position for each quadrotor using the information of its neighbors and itself. Then, a position controller and an attitude controller are constructed. The optimal control policies for the position and attitude controller are learned by using RL approach. Then, the fault-tolerant controllers are designed for the quadrotor position subsystem and the quadrotor attitude subsystem to compensate the actuator faults.

The rest parts of this paper are given as follows. Section 2 gives the preliminaries on the graph theory, the quadrotor model, the actuator fault model, and the problem formulation. The fault-tolerant formation controller is designed in Section 3. In Section 4, the simulation results of six quadrotor are shown and the concluding remarks are given in Section 5.

Notations: Define $I_N$ as a $N \times N$ unit matrix, $\lambda_i(M)$ the $i$th eigenvalue of the matrix $M$, $0_k \times l$ a $k \times l$ zero matrix, and $c_{a,b}$ an $a \times 1$ vector with 1 in the $b$th element and 0s elsewhere.

2. PRELIMINARIES

2.1 Graph Theory

Consider a team of $N$ quadrotors. Let $\Phi = \{1, 2, \cdots, N\}$. The communication topology among the quadrotors is described with a time invariant graph $G = (V, E, W)$, where $V = \{v_i\} (i \in \Phi)$ is the node set, $E \subset V \times V$ the edge set, and $W = [w_{ij}] \in \mathbb{R}^{N \times N}$ the weighted adjacency matrix. For node $v_i$ and node $v_j$ ($i, j \in \Phi$), the weight $w_{ij}$ satisfies that $w_{ij} > 0$ if and only if $(v_i, v_j) \in E$, otherwise $w_{ij} = 0$. Denote the neighbors of $v_i$ as $N_i = \{v_j | (v_i, v_j) \in E\}$. Let $d_i$ be the in-degree of $v_i$ with $d_i = \sum_{j=1}^{N} w_{ij}$. Define $L = D - W$, where $D = \text{diag}(d_i) \in \mathbb{R}^{N \times N}$. The path from $v_i$ to $v_j$ is a sequence of ordered edges with \{$(v_i, v_k), (v_k, v_n), \cdots , (v_l, v_j)$\}. The graph $G$ contains a spanning tree if there is at least one node connected to all the other nodes.

2.2 Quadrotor Model

Define $E^I = \{e_i^I, e_j^I, e_k^I\}$ as the inertial frame and $E^B = \{e_i^B, e_j^B, e_k^B\}$ the body-fixed frame. Let $p_i = [p_{xi} \ p_{yi} \ p_{zi}]^T \in \mathbb{R}^{3 \times 1}$ be the position of the $i$th quadrotor and $\Theta_i = [\phi_i \ \theta_i \ \psi_i]^T \in \mathbb{R}^{3 \times 1}$ the Euler angle of the $i$th quadrotor. From Liu et al. (2019), one can obtain the quadrotor dynamic model as

$$m \ddot{p}_i = R_{fi} F_i,$$

$$J_{\Theta_i} \ddot{\Theta}_i = -C(\Theta_i, \dot{\Theta}_i, \Theta_i) \dot{\Theta}_i + \tau_i,$$

where $m_i$ is the mass of the $i$th quadrotor, $J_{\Theta_i} = \text{diag}(J_{\phi_i}, J_{\theta_i}, J_{\psi_i}) \in \mathbb{R}^{3 \times 3}$ the inertial matrix, $C(\Theta_i, \dot{\Theta}_i, \Theta_i) \in \mathbb{R}^{3 \times 3}$ the nonlinear Coriolis term described in Raffo, Ortega, and Rubio. (2010), and $R_{fi} \in \mathbb{R}^{3 \times 3}$ the orientation matrix from $E^B$ to $E^I$ shown in Liu et al. (2019). $F_i \in \mathbb{R}^{3 \times 1}$ and $\tau_i \in \mathbb{R}^{3 \times 1}$ are the external force and the torque generated by the four rotors with

$$F_i = c_{3,3} k_{wi} \sum_{k=1}^{4} w_{i,k}^2 R_{fi}^T c_{3,3} m_{i} g$$

and

$$\tau_i = [k_{ti,k_{wi}}(\omega_{i,3}^2 - \omega_{i,2}^2) \ k_{ti,k_{wi}}(\omega_{i,2}^2 - \omega_{i,1}^2) \ l_{r} \sum_{k=1}^{4} (\omega_{i,k}^2)^2]_{i},$$

respectively, where $g$ indicates the gravity constant, $\omega_{i,j}$ ($j = 1, 2, 3, 4$) is the rotational speed of the $j$th rotor of the $i$th quadrotor, and $k_{ti}, k_{wi}, l_{r}$ are positive scaling factors for the $i$th quadrotor. Let $\omega_p^2 = [\omega_{i,1}^2 \ \omega_{i,2}^2 \ \omega_{i,3}^2 \ \omega_{i,4}^2]^T$. Due to the existence of a power distribution board, the control inputs can be distributed as:

$$\begin{bmatrix} u_{gi} \\ u_{gi} \\ u_{gi} \\ u_{zi} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \omega_{i,1}^2 \\ \omega_{i,2}^2 \\ \omega_{i,3}^2 \\ \omega_{i,4}^2 \end{bmatrix} = G_m \omega_p^2.$$  

Define $u_{pi} = [u_{pzi} \ u_{pzi} \ u_{pzi}]^T \in \mathbb{R}^{3 \times 1}$ as the virtual position control input with

$$u_{pi} = R_{fi} c_{3,3} u_{zi}.$$ 

Then, one can rewrite the quadrotor model (1) as

$$\ddot{p}_i = b_{pi} u_{pi} - c_{3,3}, g,$$

$$\ddot{\Theta}_i = -J_{\Theta_i}^{-1} (C(\Theta_i, \dot{\Theta}_i, \Theta_i)) + b_{r_i} \tau_{ri},$$

where $u_{r_i} = [u_{ri} \ u_{ri} \ u_{ri}]^T$, $b_{pi} = m_i^{-1} k_{wi} I_{3 \times 3}$, and $b_{r_i} = J_{\Theta_i}^{-1} \text{diag}(k_{ti,k_{wi}}, k_{ti,k_{wi}}, l_{r})$.
2.3 Actuator Fault Model

Similarly to Avram, Zhang, and Muse. (2018), the actuator fault of the ith quadrotor in the jth rotor is described as 
\((\omega_{ij}^*)^2 = \rho_i \omega_{ij}^2, i \in \Phi, j = 1, 2, 3, 4\), where \(\omega_{ij}^*\) indicates the actual spinning speed and \(\rho_i\) an unknown positive LOE gain with \(0 < \rho_i \leq 1\). In this paper, the actuator faults of the rotors are described with a diagonal matrix \(K_{pi} \in \mathbb{R}^{4 \times 4}\) satisfying that \(K_{pi} = diag(\rho_{i1}, \rho_{i2}, \rho_{i3}, \rho_{i4})\). Then, the quadrotor model with actuator faults can be written as

\[
\begin{align*}
\dot{p}_i &= b_i (u_{pi} - \mu(t - ts)u_{\Delta p_i}) - c_3 \sigma_1, \\
\dot{\Theta}_i &= -J_i^{-1} \left(C \left(\Theta_i, \Theta_i \right) \Theta_i \right) + b_r (u_{ri} - \mu(t - ts)u_{\Delta \Theta_i}),
\end{align*}
\]

\(7\)

where \(\mu(t - ts)\) is a step function with an unknown fault occurrence time \(t_s\) satisfying that \(\mu(t - ts) = 0\) if \(t < t_s\) and \(\mu(t - t_s) = 1\). \(u_{\Delta p_i} \in \mathbb{R}^{4 \times 1}\) and \(u_{\Delta \Theta_i} \in \mathbb{R}^{3 \times 1}\) in \(7\) are the input uncertainties resulted from the actuator faults of the rotors. From \((3), (4), (5), (6), \)\(7\) are the input uncertainties resulted from the actuator faults.

The dynamics of the virtual leader is described as 
\(\dot{x}_{\pi} = [\hat{\rho}_{1i} \hat{p}_{i1}] \in \mathbb{R}^{6 \times 1}\) as the state of the ith observer for the jth quadrotor. Let \(\hat{\sigma}_{ij} = [\hat{\delta}_{ij}, \hat{\delta}_{ij}^*] \in \mathbb{R}^{4 \times 1}\) and \(\hat{\delta}_{ij} = [0_{1 \times 3} \hat{\delta}_{ij}]\). The distributed observer is designed as

\[
\hat{\zeta}_{pi} = A_{pi} x_{\pi} + \gamma_i \sum_{j \in N_i} \left(\|u_{\pi j} - \hat{\alpha}_{pi j}\|^2 + \gamma_i (\hat{x}_{pi} - \hat{x}_{pi j})\right),
\]

\(9\)

where \(\gamma_i\) is a positive constant to be determined and \(\hat{\alpha}_{pi}\) will converge to \(0\) i.e., \(\hat{\alpha}_{pi}\) will converge to \(\alpha_{pi}\) by using the proposed distributed observer. The position tracking error \(e_{pi}\) satisfies \(e_{pi} = [C_{pi} - C_{pi}] X_{pi} \) and \(C_{pi} = [c_{1i}, c_{2i}, c_{3i}]\). Define the performance function for the augmented system as

\[
V_{pi}(e_{pi}, u_{pi}) = \int_{t_i}^{\infty} e^{-\gamma_1(t - \tau)} (e_{pi}^T Q_{pi} e_{pi} + u_{pi}^T R_{pi} u_{pi}) \, d\tau,
\]

\(11\)

where \(Q_{pi} > 0\), \(R_{pi} > 0\), and \(\gamma_1\) is a positive constant to guarantee that \(V_{pi}\) is bounded for any given \(u_{ri}\) (see, Modares and Lewis. (2014)). From the optimal control theory (see, Lewis and Syrmos. (1995)), one can obtain the optimal position control policy \(u_{pi}^*\) as

\[
u_{pi}^* = -\frac{1}{2} R_{pi}^{-1} B_{pi}^T V_{pi}^*.
\]

\(12\)

3. CONTROL PROTOCOL DESIGN

3.1 Distributed Observer

Let \(x_{\pi 0} = [\hat{\rho}_{1i} \hat{p}_{i1}^T]^T \in \mathbb{R}^{6 \times 1}\) be the state of the leader. The dynamics of the virtual leader is described as \(\dot{x}_{\pi 0} = A_{pi} x_{\pi 0}\), where \(A_{pi} \in \mathbb{R}^{6 \times 6}\). Denote \(\hat{\alpha}_{pi} = [\hat{\rho}_{pi}, \hat{p}_{pi}] \in \mathbb{R}^{6 \times 1}\) as the state of the ith observer for the jth quadrotor.
Algorithm 3.2: RL algorithm for position controller

Initiation: Apply a stabilizing control input \( u_{pi} \) with bounded \( u_{pei} \) to the quadrotor system and collect the data. Determine the stopping criterion \( \vartheta_{t1} > 0 \).

Step 1: Solve the equation (16) for \( V_n^{\pi} \) and \( u_{n+i}^{\pi} \), simultaneously.

Step 2: Stop if: \( |u_{n+i}^{\pi} - u_{n+i}^{\pi}| < \vartheta_{t1} \), otherwise, let \( n = n + 1 \) and go to Step 1.

\[
d\hat{u}_{\Delta p}/dt = -k_{s1} \sigma_1(t) \hat{u}_{\Delta p} + 2k_{s1} \| R_{pi} u_{pei}(x) \| , \quad (15)
\]

where \( k_{s1} \) is a positive constant. It can be seen from (14) that the proposed fault-tolerant controller rely on the optimal control policy \( u_{pei}^{\pi} \). But the optimal control approach (12), (13) requires the accurate dynamic information of the quadrotor team, which is impractical for quadrotor team in the formation flight applications. In the following, a RL algorithm is designed to obtain \( u_{pei}^{\pi} \) without knowledge of the quadrotor dynamic information.

For a given bounded signal \( u_{pei} \in \mathbb{R}^{3\times1} \), one can obtain from (12) that

\[
e^{-\alpha_1 \Delta t} V_n^{\pi}(X_{pi} (t + \Delta t)) - V_n^{\pi}(X_{pi} (t)) = \int_t^{t+\Delta t} e^{-\alpha_1 (\tau - t)} \left[ -e^{T} P_{pi} e_{pi} - (u_{n+i}^{\pi})^T R_{pi} u_{pei}^{\pi} \right] d\tau \quad (16)
\]

where \( V_n^{\pi}(X_{pi}) \) and \( u_{n+i}^{\pi} \) indicate the updated value of \( V_n^{\pi}(X_{pi}) \) and \( u_{n+i}^{\pi} \) in the \( n \)th iteration. It can be seen from (16) that the \( V_n^{\pi} \) and \( u_{n+i}^{\pi} \) can be updated using (16), which yields Algorithm 3.2.

In Algorithm 3.2, the performance function \( V_n^{\pi} \) and the control policy \( u_{n+i}^{\pi} \) in the \( n \)th iteration can be approximated by a critic neural network (NN) and an action NN. The weights of the NNs can be updated by using the least-square (LS) method under a persistence excitation (PE) condition (see, Lewis and Syrmos. (1995)). Define \( \Theta_{ri} = [\Theta_{ri} \Theta_{ri} \Theta_{ri}]^T \) as the attitude reference for the attitude controller of the \( i \)th quadrotor. Once the fault-tolerant position control policy \( u_{pi} \) in (14) is determined, from (5), one can derive the desired control input \( u_{zri} \), the desired roll angle \( \phi_{ri} \), and the desired pitch angle \( \theta_{ri} \) for the attitude controller with

\[
\begin{align*}
    u_{zri} & = u_{zpi}/\cos \phi_i/\cos \phi_i, \\
    \phi_{ri} = \arcsin((\cos \phi_i \cos \phi_i - u_{pei}/u_{zri})/\cos \phi_i), \\
    \theta_{ri} = \arcsin((u_{pei}/u_{zri} - \sin \phi_i \sin \phi_i)/\cos \phi_i/\cos \phi_i).
\end{align*}
\]

3.3 Attitude Controller Design

Let \( x_{\Theta ri} = [\Theta_{ri}^T \Theta_{ri}^T]^T \in \mathbb{R}^{6\times1} \) be the state of the attitude reference and \( \dot{x}_{\Theta ri} = [\dot{\Theta}_{ri}^T \dot{\Theta}_{ri}^T]^T \in \mathbb{R}^{6\times1} \) the attitude state of the \( i \)th quadrotor. From (7), an attitude augmented system can be constructed as

\[
\dot{X}_{\Theta ri} = \dot{F}_\Theta (X_{\Theta ri}) + \dot{B}_{ri} (u_{zri} - \mu(t - t_s) u_{\Delta \Theta ri}) \quad (18)
\]

Algorithm 3.3: RL algorithm for attitude controller

Initialization: Apply a stabilizing policy \( u_{rri} \) with bounded \( u_{rri} \) to the quadrotor and collect the system data.

Step 1: Solve for \( V_n^{\Theta ri} \) and \( u_{n+i}^{\Theta ri} \) simultaneously:

\[
\begin{align*}
    e^{-\alpha_2 \Delta t} V_n^{\Theta ri}(X_{\Theta ri} (t + \Delta t)) - V_n^{\Theta ri}(X_{\Theta ri} (t)) &= \int_t^{t+\Delta t} e^{-\alpha_2 (\tau - t)} \left[ -e^{T} Q_{ri} e_{ri} - (u_{n+i}^{\Theta ri})^T R_{ri} u_{rei}^{\Theta ri} \right] d\tau, \\
    -\int_t^{t+\Delta t} e^{-\alpha_2 (\tau - t)} 2(u_{n+i}^{\Theta ri} - \Theta_{ri}^T)^T R_{ri} u_{rei} d\tau.
\end{align*}
\]

Step 2: Set \( u_{rei}^{\Theta ri} = u_{n+i}^{\Theta ri} \) and \( V_n^{\Theta ri} = V_{n+i}^{\Theta ri} \) and go to Step 1 until a stopping criterion is reached.

Define \( e_{ri} = [e_{r0} e_{r0} e_{r0}]^T \in \mathbb{R}^{3\times1} \) as the attitude tracking error with \( e_{r0} = [C_{r0} - C_{r}]X_{\Theta ri} \) and \( C_r = [c_0 \beta_{r1} c_0 \beta_{r2} c_0 \beta_{r3}] \). Then, a performance function for the attitude system is defined as

\[
V_{\Theta ri}(e_{rri}, u_{rri}) = \int_0^\infty e^{-\alpha_2 (\tau - t)} (e_{rri}^T Q_{ri} e_{rri} + u_{rri}^T R_{rei} u_{rri}) d\tau. \quad (20)
\]

where \( Q_{ri} = Q_{ri}^T > 0, R_{rei} = R_{rei}^T > 0, \) and \( \alpha_2 > 0 \). Then, one can obtain an optimal attitude control policy \( u_{rri}^* \), as

\[
\begin{align*}
    &\dot{x}_{\Theta ri} = \left( \frac{1}{4} (\Delta V_{\Theta ri}^*) \right)^T B_{ri} \dot{\Theta}_{ri} \Delta V_{\Theta ri}^* - \sigma_2 u_{rei} + (\Delta V_{\Theta ri}^*)^T F_\Theta (X_{\Theta ri}) = 0, \quad (21)
\end{align*}
\]

It can be seen that the equation (21) is nonlinear with respect to \( V_{\Theta ri}^* \) and requires the dynamic model of the quadrotor team. To track the desired attitude reference for quadrotor team with actuator faults, similarly to the position controller design, a fault-tolerant attitude controller is designed as

\[
\begin{align*}
    u_{rri}(t) &= \frac{1}{2} u_{rri}(t) + \frac{\dot{\theta}_{rri}^2}{2 \Theta_{ri} u_{rei}^*} (x_{rei} + \sigma_2(t)),
\end{align*}
\]

where \( \sigma_2(t) \) is a positive bounded function with \( \int_0^\infty \sigma_2(t) d\tau \leq \sigma_2 \leq \infty \). \( \dot{\theta}_{rri} \) is the estimated value of \( \dot{\theta}_{\Delta \Theta ri} \) with

\[
\begin{align*}
    \dot{\theta}_{\Delta \Theta ri} &= -k_{s2} \sigma_2(t) \theta_{\Delta \Theta ri} + 2k_{s2} \| R_{rei} u_{rei}(x)_r \|, \quad (23)
\end{align*}
\]

where \( k_{s2} \) is a positive constant. Therefore, similarly to the position controller design, Algorithm 3.3 is designed to learn the optimal attitude control policy \( u_{rri}^* \).

Similarly to the lines of reasoning of the position controller design, the performance function \( V_{\Theta ri}^* \) and the control
Then, the fault-tolerant formation controller including the position controller and the weight of the NNs can be updated using the LS technique under the PE condition.

4. SIMULATION RESULTS

In this section, three quadrotor systems are modeled as (7) with the parameters being $J_{bi} = diag(4, 4, 8) \times 10^{-3}$ kg · m$^2$, $b_{bi} = diag\{41, 41, 113\}$, $g_i = 9.81$ m/s$^2$, and $b_{pi} = diag\{1, 1, 1\}$ ($i = 1, 2, 3$). The communication graph between the quadrotor team is described with $w_{13} = w_{21} = w_{32} = 1$ and $\rho_1 = 1$. The quadrotor team is required to form a triangle formation and track a virtual leader, simultaneously. The dynamic of the leader is set as $\dot{v}_0(0) = 0_{3 \times 1}$ m, and the states of the quadrotors are initialized as: $p_1(0) = [2.25 2.25 1.00]^T$ m, $p_2(0) = [-2.25 3.38 0]^T$ m, $p_3(0) = [0 -3.38 0.50]^T$ m, $\dot{p}_i = 0_{3 \times 1}$, the position deviations are set as $\delta_{13} = [1.5 3.75 0]^T$ m, $\delta_{21} = [-3 0.75 0]^T$ m, and $\delta_{32} = [1.50 -4.5 0]^T$ m. The distributed observer is implemented for each quadrotor with the parameter as $\gamma_i = 80$ and the position estimation errors are shown in Fig. 1. The solid lines colored with red, black, and blue represent the 1-3th quadrotor. After the estimation errors of the observers converge to 0s, the RL algorithms are applied to learn the optimal control policies $u^*_{bi}, u^*_{ri}$ for the position subsystem and the attitude subsystem. The base functions of the critic NNs and the action NNs are chosen to be multiple polynomials. The time interval $\Delta t$ is chosen to be $\Delta t = 0.05s$. Then, the fault-tolerant formation controller including the position controller and
the attitude controllers is constructed. The parameters of the fault-tolerant controller are set as: \( R_{pi} = R_{\Theta i} = \ldots \) continuous-time nonlinear tracking systems with actuator failures. Journal of the Franklin Institute, 355(15):6947–6968.

It can be seen from Figs. 4 that the traditional optimal control policies can maintain a considerable formation performance without actuator faults, but the formation performance turns undesirable when the actuator faults appear. When the actuator faults occur at 4 s, it can be observed from Figs. 3 that the formation performance of the quadrotor team is greatly improved by using the proposed fault-tolerant formation controller.

5. CONCLUSION

In this paper, the fault-tolerant formation control problem is addressed for multiple cooperative quadrotors with unknown dynamics. The quadrotor is modeled as a nonlinear and coupled system being subject to actuator faults. A fault-tolerant formation controller is proposed including a distributed observer to estimate the position references for each quadrotor, a fault-tolerant position controller to achieve the desired position formation, and a fault-tolerant attitude controller to track the desired attitude. Reinforcement learning algorithms are designed to learn the optimal control policies of the position and attitude controller without knowledge of the quadrotor dynamic information. Simulation results are provided to show the effectiveness of the proposed fault-tolerant formation controller.

REFERENCES

Avram, R. C., Zhang, X., and Muse, J. (2018) Nonlinear adaptive fault-tolerant controller altitude and attitude tracking with multiple actuator faults. IEEE Transactions on Control Systems Technology, 26(2):701–707.

Avram, R. C., Zhang, X., and Muse, J. (2017). Quadrotor attitude controller fault diagnosis and accommodation using nonlinear adaptive estimators. IEEE Transactions on Control Systems Technology, 25(6):2219–2226.

Deng, C., Yang, G. H. (2017). Adaptive fault-tolerant control for a class of nonlinear multi-agent systems with actuator faults. Journal of the Franklin Institute, 354(12):4784–4800.

Deptula, P., Bell, Z. I., Doucette, E. A., Curtis, J. W., and Dixon, W. E. (2018). Data-based reinforcement learning approximate optimal control for an uncertain nonlinear system with partial loss of control effectiveness. 2018 Annual American Control Conference, Wisconsin, U.S.A., 6947–6968.

Dong, X., Hua, Y., Zhou, Y., Ren, Z., and Zhong, Y. (2018). Theory and experiment on formation-containment control of multiple multirotor unmanned aerial vehicle systems. IEEE Transactions on Automation Science and Engineering, 16(1):229–240.

Hu, Q., Zhang, Y., Huo, X., and Bing, X. (2011). Adaptive integral-type sliding mode control for spacecraft attitude maneuvering under actuator stuck failures. Chinese Journal of Aeronautics, 24(1):32–45.

Hua, Y., Dong, X., Li, Q., and Ren, Z. (2016). Distributed fault-tolerant time-varying formation control for second-order multi-agent systems with actuator failures and directed topologies. IEEE Transactions on Circuits and Systems II: Express Briefs, 63(6):774–778.

Liu, H., Tian, Y., Lewis, F. L., Wan, Y., and Valavanis, K. P. (2019). Robust formation tracking control for multiple quadrotors under aggressive maneuvers. Automatica, 105:179–185.

Liu, L., Wang, Z., and Wang, H. (2017). Adaptive fault-tolerant tracking control for MIMO discrete-time systems via reinforcement learning algorithm with less learning parameters. IEEE Transactions on Automation Science and Engineering, 14(1):299–313.

Modares, H. and Lewis, F. L. (1995). Optimal control. NJ, Hoboken: Wiley.

Ma, H. J., Xu, L. X., and Yang, G. H. (2019). Multiple environment integral reinforcement learning-based fault-tolerant control for affine nonlinear systems. IEEE transactions on cybernetics, unpublished.

Modares, H. and Lewis, F. L. (2014). Linear quadratic tracking control of partially-unknown continuous-time systems using reinforcement learning. IEEE Transactions on Automatic Control, 59(11):3051–3056.

Qin, J., Ma, Q., Shi, Y., and Wang, L. (2017). Recent advances in consensus of multi-agent systems: A brief survey. IEEE Transactions on Industrial Electronics, 64(6):4972–4983.

Raffo, G. V., Ortega, M. G., and Rubio, F. R. (2010). An integral predictive/nonlinear \( H_\infty \) control structure for a quadrotor helicopter. Automatica, 46(1):29–39.

Shi, J., Yang, Y., Sun, J., He, X., Zhou, D., and Zhong, Y. (2017). Fault-tolerant formation control of non-linear multi-vehicle systems with application to quadrotors. IET Control Theory & Applications, 11(17):3179–3190.

Slotine, J. and Li, W. (1991). Applied Nonlinear Control. Englewood Cliffs, NJ, USA: Prentice Hall.

Wang, Z., Liu, L., Zhang, H., and Xiao, G. (2016). Fault-tolerant controller design for a class of nonlinear MIMO discrete-time systems via online reinforcement learning algorithm. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 46(5):611–622.

Yu, Z., Qu, Y., and Zhang, Y. (2019). Distributed fault-tolerant cooperative control for multi-UAVs under actuator fault and input saturation. IEEE Transactions on Control Systems Technology, unpublished.

Zhang, K., Zhang, H., Gao, Z., and Su, H. (2018). Online adaptive policy iteration based fault-tolerant control algorithm for continuous-time nonlinear tracking systems with actuator failures. Journal of the Franklin Institute, 355(15):6947–6968.