Reinforcement Learning-Based Fractional-Order Adaptive Fault-Tolerant Formation Control of Networked Fixed-Wing UAVs With Prescribed Performance

Ziquan Yu, *Member, IEEE*, Jiaxu Li, Yiwei Xu, Youmin Zhang, *Fellow, IEEE*, Bin Jiang, *Fellow, IEEE*, and Chun-Yi Su, *Senior Member, IEEE*

***Abstract*— This article investigates the fault-tolerant formation control (FTFC) problem for networked fixed-wing unmanned aerial vehicles (UAVs) against faults. To constrain the distributed tracking errors of follower UAVs with respect to neighboring UAVs in the presence of faults, finite-time prescribed performance functions (PPFs) are developed to transform the distributed tracking errors into a new set of errors by incorporating user-specified transient and steady-state requirements. Then, the critic neural networks (NNs) are developed to learn the long-term performance indices, which are used to evaluate the distributed tracking performance. Based on the generated critic NNs, actor NNs are designed to learn the unknown nonlinear terms. Moreover, to compensate for the reinforcement learning errors of actor-critic NNs, nonlinear disturbance observers (DOs) with skillfully constructed auxiliary learning errors are developed to facilitate the FTFC design. Furthermore, by using the Lyapunov stability analysis, it is shown that all follower UAVs can track the leader UAV with predesigned offsets, and the distributed tracking errors are finite-time convergent. Finally, comparative simulation results are presented to show the effectiveness of the proposed control scheme.**

***Index Terms*— Disturbance observer (DO), fault-tolerant formation control (FTFC), finite-time prescribed performance, fixed-wing unmanned aerial vehicle (UAV), fractional-order (FO) control, reinforcement learning control.**

Manuscript received 27 November 2022; revised 23 February 2023; accepted 18 May 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 62003162, Grant 61833013, Grant 62020106003, and Grant 62233009; in part by the Natural Science Foundation of Jiangsu Province of China under Grant BK20200416 and Grant BK20222012; in part by the National Key Research and Development Program of China under Grant 2021YFB3301300; in part by the Science and Technology on Space Intelligent Control Laboratory under Grant HTKJ2022KL502015; in part by the Industry-University Research Innovation Foundation for the Chinese Ministry of Education under Grant 2021ZYA02005; in part by the 111 Project under Grant B20007; in part by the Postgraduate Research and Practice Innovation Program of Nanjing University of Aeronautics and Astronautics (NUAA) under Grant xcxjh20220325; in part by the Aeronautical Science Foundation of China under Grant 20200007018001; and in part by the Natural Sciences and Engineering Research Council of Canada. *(Corresponding author: Youmin Zhang.)*

Ziquan Yu, Jiaxu Li, Yiwei Xu, and Bin Jiang are with the College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China (e-mail: yuziquan@nuaa.edu.cn; ljx_nuaa@foxmail.com; xuyiwei@nuaa.edu.cn; binjiang@nuaa.edu.cn).

Youmin Zhang and Chun-Yi Su are with the Department of Mechanical, Industrial and Aerospace Engineering, Concordia University, Montreal, QC H3G 1M8, Canada (e-mail: youmin.zhang@concordia.ca; chun-yi.su@concordia.ca).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TNNLS.2023.3281403.

Digital Object Identifier 10.1109/TNNLS.2023.3281403

I. Introduction

RECENTLY, networked unmanned aerial vehicles (UAVs) have been widely used to increase the efficiency in cooperative search and rescue [1], forest fire monitoring [2], and wireless communication relay [3]. Compared with rotary-wing UAVs, fixed-wing UAVs can provide monitoring/observation of a wide area due to the long flight duration and high flight speed [4]. However, the high flight speed and complex aerodynamic characteristics are very challenging to the formation control design. To solve this problem, some effective formation control architectures have been investigated, such as leader–follower and virtual structure frameworks [5]. Moreover, sliding-mode control, adaptive control, backstepping control, and intelligent control methods are usually integrated into the aforementioned formation control architectures to facilitate the control design. It should be noted that unanticipated faults are rarely handled in the aforementioned results. Moreover, to ensure the formation flight safety against faults, user-specified formation control requirements in the presence of faults should be explicitly considered.

For the high-speed fixed-wing UAVs, faults may significantly threaten flight safety or even cause crashes [6], [7]. To guarantee reliable flight, fault-tolerant control (FTC) concept has been developed by using control reconfiguration [8], [9], [10]. For the FTC of fixed-wing UAV, [11] studied a nonlinear adaptive control strategy with the involvement of sliding-mode online learning. With respect to networked UAVs, some FTC methods have been explored for enhancing reliability, such as the finite-time FTC method and the adaptive estimator-based FTC strategy for multiple quadrotor UAVs [12], [13]. Recently, some investigations have been made to develop FTC strategies for fixed-wing UAVs. In [14], an FTC method was developed for longitudinal motions of multiple fixed-wing UAVs by integrating the disturbance observer (DO) into the dynamic surface control architecture. Then, by further introducing the finite-time characteristics into

the dynamic surface control framework, a finite-time FTC scheme was investigated to regulate the longitudinal motions for networked fixed-wing UAVs [15]. More recently, several FTC schemes were studied for achieving the attitude synchronization tracking control of networked fixed-wing UAVs in a distributed communication network [16]. Moreover, to achieve refined adjustments of the transient performance at the fault stage and the steady-state performance at the post-fault stage, fractional-order (FO) calculus has been incorporated into the FTC design for the attitude control of networked UAVs, leading to the FO FTC schemes [17]. However, these results are focused on the FTC of longitudinal or attitude motions for UAVs. To date, the FTC methods for the position motions of networked fixed-wing UAVs are very rare due to the challenge caused by the high flight speeds and fast fault effect spread, which motivates this research.

To satisfy the constrained performance requirements, prescribed performance functions (PPFs) with user-assigned convergence rates and error bounds have been widely employed to achieve the prescribed performance control (PPC) of UAVs [18], manipulators [19], unmanned surface vehicles [20], [21], unmanned underwater vehicles [22], tractor–trailer convoys [23], and spacecrafts [24]. Regarding the FTC design, it is highly desirable if the transient and steady-state performances stay within the prescribed error bounds even when the system is suddenly encountered by faults [25]. In [26], an FTC method was studied for UAV against faults by incorporating prescribed transient performance requirements. Recently, finite-time PPFs are preliminarily investigated for intending to achieve appointed-time convergence, which is different from the exponential convergence in the traditional PPC schemes [27], [28]. In [29], a finite-time fuzzy adaptive tracking control method was developed for a nonlinear system via a finite-time PPF. More recently, a finite-time PPF-based FTC scheme was developed for networked unmanned airships in [30]. Despite the fact that there exist numerous PPC methods [18], [19], [20], [21], [24], a few PPC schemes with fault-tolerant capabilities [25], [26], several finite-time PPC strategies [27], [28], [29], and only a finite-time PPC framework with fault-tolerant competence [30], the effective FTC design for networked fixed-wing UAVs with finite-time prescribed performance is still an open topic and should be further investigated.

To address the unknown characteristics inherent in the systems, neural networks (NNs) are usually utilized to approximate unknown terms in the control design [31], [32]. Recently, reinforcement learning is innovatively investigated to design control schemes, which can handle the unknown features in linear and nonlinear systems, such as dual-arm robots [33], quadrotor UAVs [34], and helicopters [35]. In addition to the approximation capabilities of traditional NNs, the reinforcement learning strategy also has the optimal control performance [36]. Different from the critic learning strategy [37], [38], the actor-critic learning mechanism integrates the advantages of the value-based and policy-based reinforcement learning structures, such that the learning framework can be used online for achieving good control performance in continuous-time systems [35]. For the actor-critic network, a compensated control signal is generated by the actor NN, which is fed into the control unit, and the performance index function is evaluated by the critic NN, which is then used to update the actor NN [39]. By considering model uncertainties, nonlinearities, and external disturbances, a reinforcement learning-based flight controller was developed in [40] for quadrotor UAVs with prescribed transient and steady performance. Recently, the actor-critic reinforcement learning network has been used in the FTC design. In [41], a reinforcement learning-based FTC scheme was investigated for a multiple-input–multiple-output (MIMO) system with less learning parameters in the actor-critic learning framework. In [42], an FTC scheme was studied for a discrete-time multiple agent system by using the reinforcement learning-based policy iteration mechanism. However, to the best of the authors' knowledge, very few results have been reported for the reinforcement learning-based FTC of single fixed-wing UAV, let alone the FTC of networked fixed-wing UAVs with nonlinear dynamics, which needs to be investigated for providing intelligent FTC schemes. Moreover, the reinforcement learning error was rarely compensated, which may weaken the control performance and further motivate this research.

Motivated by the aforementioned analyses, a reinforcement learning-based fault-tolerant formation control (FTFC) scheme is developed for networked fixed-wing UAVs by involving user-specified transient and steady-state requirements. To address the unknown nonlinear terms in the fixed-wing UAVs, the reinforcement learning framework with actor-critic NNs is designed for providing the compensatory signal to the unknown nonlinearities. Nonlinear DOs with the consideration of learning errors are skillfully developed to enhance the reinforcement learning capabilities. Finite-time PPFs are developed to strictly constrain distributed tracking errors. Moreover, FO calculus is utilized to achieve refined performance by adjusting the FO operator. The contributions of this article can be summarized as follows.

1) Different from the FTC schemes for single UAV [6], [7] and the fault-tolerant synchronization tracking control strategies for multiple fixed-wing UAVs [16], an FTFC method is developed to achieve the outer-loop position motions of networked fixed-wing UAVs by involving finite-time PPFs to impose the user-specified requirements on the distributed tracking errors, such that the FTFC capability is significantly enhanced in the presence of faults.
2) Compared with integer-order FTC strategies [8], [9], the FO calculus is integrated into the developed FTFC methods for achieving refined adjustment of the FTFC performance, since the FO operator can be chosen as a noninteger. By using such a strategy, a flexible FTFC method is obtained to finely adjust the transient and steady-state performances.
3) In contrast to widely investigated neural/fuzzy adaptive FTC methods [31], [32], this article adopts the reinforcement learning architecture with actor-critic NNs to handle the unknown nonlinear terms and long-term performance indices. To further compensate for the

reinforcement learning errors, nonlinear DOs with auxiliary learning errors are skillfully developed to enhance the FTFC capabilities. To the best of the authors' knowledge, this may be the first result to address the FTFC problem for networked fixed-wing UAVs via reinforcement learning mechanism with simultaneous consideration of learning error compensation.

The remainder of this article is organized as follows. Section II presents the preliminaries and states the investigated problem. Section III shows the error transformations, reinforcement learning-based FO FTFC scheme, learning error compensation, and stability analysis. The comparative simulation results are shown in Section IV, followed by the conclusions and future works in Section V.

II. Preliminaries and Problem Statement

A. Faulty Fixed-Wing UAV Model

In this article, a formation team consisting of N follower fixed-wing UAVs (labeled as $1, 2, \ldots, N$) and one leader fixed-wing UAV (labeled as 0) is considered, and the outer-loop position model of the ith follower UAV is expressed as follows [43]:

$$\begin{cases} \dot{x}_i = V_i \cos\gamma_i \cos\chi_i \\ \dot{y}_i = V_i \cos\gamma_i \sin\chi_i \\ \dot{h}_i = V_i \sin\gamma_i \end{cases} \tag{1}$$

$$\begin{cases} \dot{V}_i = u_{i1} - g\sin\gamma_i + d_{i1} \\ \dot{\chi}_i = \dfrac{u_{i2}}{V_i\cos\chi_i} + \dfrac{d_{i2}}{V_i\cos\chi_i} \\ \dot{\gamma}_i = \dfrac{u_{i3}}{V_i} - \dfrac{g\cos\gamma_i}{V_i} + \dfrac{d_{i3}}{V_i} \end{cases} \tag{2}$$

where $i \in \{1, \ldots, N\}$. $\boldsymbol{q}_i = [x_i, y_i, h_i]^T$ is the position vector. V_i, χ_i, and γ_i represent the velocity, heading angle, and flight path angle, respectively. $\boldsymbol{u}_i = [u_{i1}, u_{i2}, u_{i3}]^T$ denotes the control input vector. $\boldsymbol{d}_i = [d_{i1}, d_{i2}, d_{i3}]^T$ is the disturbance vector.

Differentiating (1) yields

$$\ddot{\boldsymbol{q}}_i = \boldsymbol{f}_i + \boldsymbol{g}_i \boldsymbol{u}_i \tag{3}$$

where $\boldsymbol{f}_i$ and $\boldsymbol{g}_i$ have the following expressions:

$$\boldsymbol{f}_i = \begin{bmatrix} d_{i1}\cos\gamma_i\cos\chi_i - d_{i2}\sin\chi_i - d_{i3}\sin\gamma_i\cos\chi_i \\ d_{i1}\cos\gamma_i\sin\chi_i + d_{i2}\cos\chi_i - d_{i3}\sin\gamma_i\sin\chi_i \\ d_{i1}\sin\gamma_i + d_{i3}\cos\gamma_i - g \end{bmatrix} \tag{4}$$

$$\boldsymbol{g}_i = \begin{bmatrix} \cos\gamma_i\cos\chi_i & -\sin\chi_i & -\sin\gamma_i\cos\chi_i \\ \cos\gamma_i\sin\chi_i & \cos\chi_i & \sin\gamma_i\sin\chi_i \\ \sin\gamma_i & 0 & \cos\gamma_i \end{bmatrix}. \tag{5}$$

By considering loss-of-effectiveness and bias faults, the following faulty UAV model can be obtained:

$$\ddot{\boldsymbol{q}}_i = \boldsymbol{f}_i + \boldsymbol{g}_i \boldsymbol{\Lambda}_i \boldsymbol{u}_{i0} + \boldsymbol{g}_i \boldsymbol{b}_{if} \tag{6}$$

where $\boldsymbol{u}_{i0} = [u_{i01}, u_{i02}, u_{i03}]^T$ is the signal generated from the control scheme. $\boldsymbol{\Lambda}_i = \mathrm{diag}\{\Lambda_{i1}, \Lambda_{i2}, \Lambda_{i3}\}$ is the unknown remaining effectiveness matrix with $\Lambda_{iv} \in [\underline{\lambda}_i, 1]$, $v = 1$–3. $\underline{\lambda}_i$ is the lowest effectiveness factor, which is predetermined by the designer. $\boldsymbol{b}_{if} = [b_{if1}, b_{if2}, b_{if3}]^T$ is the bounded bias fault vector.

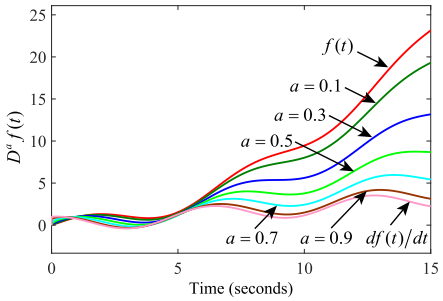


Fig. 1. Illustrative difference between IO and FO derivatives of $f(t) = \sin(t) + 0.1t^2$.

Remark 1: Different from the extensively studied FTC methods for the inner-loop attitudes of fixed-wing UAVs, this article focuses on the outer-loop FTFC design for networked UAVs. For the faults encountered by the outer-loop position model of fixed-wing UAV, Λ_{i1} and b_{if1} are related to the engine faults, Λ_{i2} and b_{if2} are mainly caused by the aileron and rudder faults, and Λ_{i3} and b_{if3} are primarily caused by the elevator faults.

B. Fractional Calculus

Since this article mainly focuses on the FO FTFC development of networked fixed-wing UAVs, the following definition and property are used.

Definition 1: Based on the Gamma function $\Gamma(\cdot)$, the fractional integral and derivative with the Reimann–Liouville (RL) definition are, respectively, defined as follows [44]:

$${}_{t_0}D_t^{-a} f(t) = \frac{1}{\Gamma(a)} \int_{t_0}^{t} \frac{f(\tau)}{(t-\tau)^{1-a}} d\tau \tag{7}$$

$${}_{t_0}D_t^{a} f(t) = \frac{1}{\Gamma(n-a)} \frac{d^n}{dt^n} \int_{t_0}^{t} \frac{f(\tau)}{(t-\tau)^{a-n+1}} d\tau \tag{8}$$

where $n-1 < a \le n$ represents the positive FO operator and n is a positive integer.

Property 1: Under Definition 1, the following property is satisfied [45], [46]:

$$\begin{cases} D^{1-a}\left(D^a f(t)\right) = \dot{f}(t) \\ D^a\left(D^{-a} f(t)\right) = f(t), \end{cases} \quad a \in (0, 1]. \tag{9}$$

Remark 2: From Fig. 1, it is observed that refined adjustments can be achieved by varying FO operator a from 0 to 1, which is significantly different from the IO derivative. With this strategy, the FTFC performance can be finely adjusted.

C. Radial Basis Function NN

As a typical manner to handle unknown nonlinear functions, radial basis function (RBF) NN has been widely used to provide approximations of unknown terms, which are then fed into the control structure [47]. Regarding a nonlinear function vector $\boldsymbol{F}(\boldsymbol{r}) \in R^{l_1}$, the RBF NN is denoted by

$$\boldsymbol{F}(\boldsymbol{r}) = \boldsymbol{W}^{*T} \boldsymbol{\varphi}(\boldsymbol{r}) + \boldsymbol{\varepsilon}(\boldsymbol{r}) \tag{10}$$

where $\boldsymbol{r}$ is the input vector. $\boldsymbol{W}^* \in R^{l_2 \times l_1}$ represents the ideal weighting matrix, and l_2 is the number of neural nodes. $\boldsymbol{\varepsilon}(\boldsymbol{r}) \in R^{l_1 \times 1}$ is the bounded approximation error vector. $\boldsymbol{\varphi}(\boldsymbol{r}) \in R^{l_2 \times 1}$ denotes the basis function vector, and the corresponding element is defined as follows:

$$\varphi_n(\boldsymbol{r}) = \exp\left(-\frac{\|\boldsymbol{r} - \boldsymbol{c}_n\|^2}{m_n^2}\right) \tag{11}$$

where $n = 1, 2, \ldots, l_2$, $\boldsymbol{c}_n$ and m_n are the center vector and width of the nth neural node, respectively.

D. Basic Graph Theory

By viewing each fixed-wing UAV as a node and the communication link between two neighboring UAVs as an edge, then a undirected graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \boldsymbol{\mathcal{A}}\}$ can be utilized to describe the information flows among the N follower UAVs, where $\mathcal{V} = \{\text{UAV \#1}, \text{UAV \#2}, \ldots, \text{UAV \#}N\}$ is the set of follower UAVs, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ represents the communication edge set, and $\boldsymbol{\mathcal{A}} = [a_{ij}] \in R^{N \times N}$ denotes the adjacency matrix. The element a_{ij} of $\boldsymbol{\mathcal{A}}$ is set as a positive value if the ith UAV can obtain the information from the jth UAV, i.e., $(\text{UAV \#}i, \text{UAV \#}j) \in \mathcal{E}$; otherwise, $a_{ij} = 0$. Since the undirected information flows are considered, $a_{ij} = a_{ji}$; that is, $\boldsymbol{\mathcal{A}}$ is a symmetric matrix. The neighboring UAV set of the ith UAV is defined as $N_i = \{\text{UAV \#}j | (\text{UAV \#}i, \text{UAV \#}j) \in \mathcal{E}\}$. Define the degree matrix $\boldsymbol{\mathcal{D}} = \text{diag}\{D_1, D_2, \ldots, D_N\}$ with $D_i = \sum_{j=1}^{N} a_{ij}$; then, the symmetric Laplacian matrix $\boldsymbol{\mathcal{L}}$ can be further defined as follows:

$$\boldsymbol{\mathcal{L}} = \boldsymbol{\mathcal{D}} - \boldsymbol{\mathcal{A}}. \tag{12}$$

By augmenting the graph $\mathcal{G}$ with one leader UAV, i.e., UAV #0, the leader adjacency matrix is defined as $\boldsymbol{\mathcal{B}} = \text{diag}\{b_1, b_2, \ldots, b_N\}$, and the element b_i is positive if the ith UAV has access to the leader UAV; otherwise, $b_i = 0$. The augmented graph is called a connected graph if a path can be found from the leader UAV to every follower UAV.

Lemma 1 [48]: If the augmented graph consisting of N follower UAVs and one leader UAV is undirected and connected, then the matrix $\boldsymbol{\mathcal{L}} + \boldsymbol{\mathcal{B}}$ is symmetric and positive definite.

E. Control Objective

The control objective is to design an FO FTFC scheme for each follower UAV with reinforcement learning mechanism, such that the distributed tracking errors of all follower UAVs can be strictly confined within the user-specified bounds in finite time, even when a portion of UAVs is subjected to unexpected faults.

III. Main Results

In this section, a reinforcement learning-based FO FTFC scheme is developed to achieve safe formation flight. To strictly confine the distributed tracking errors against faults, finite-time PPFs are utilized to construct new errors with the involvement of FO calculus. The actor-critic NNs and nonlinear DOs with auxiliary learning errors are innovatively constructed for approximating the unknown nonlinear term for each UAV.

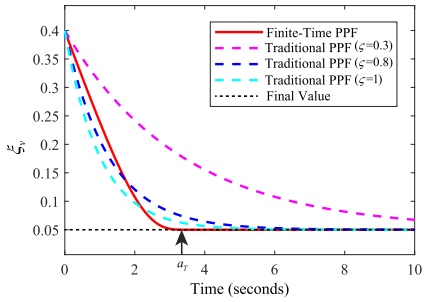


Fig. 2. Illustrative difference between finite-time PPF and traditional PPF.

A. Prescribed Performance Constraints With Finite-Time Requirements

Define the distributed tracking error as follows:

$$\boldsymbol{e}_i = \sum_{j \in N_i} a_{ij}\left(\boldsymbol{q}_i - \boldsymbol{q}_j - \boldsymbol{\delta}_{ij}\right) + b_i\left(\boldsymbol{q}_i - \boldsymbol{q}_0 - \boldsymbol{\delta}_i\right) \tag{13}$$

where $i \in \{1, \ldots, N\}$, $\boldsymbol{e}_i = [e_{i1}, e_{i2}, e_{i3}]^T$. $\boldsymbol{q}_0 = [q_{01}, q_{02}, q_{03}]^T$ is the position vector of the leader UAV. $\boldsymbol{\delta}_i = [\delta_{i1}, \delta_{i2}, \delta_{i3}]^T$ denotes the offset vector of the ith UAV with respect to the leader UAV. $\boldsymbol{\delta}_{ij} = [\delta_{ij1}, \delta_{ij2}, \delta_{ij3}]^T = \boldsymbol{\delta}_i - \boldsymbol{\delta}_j$ represents the relative offset between the ith and the jth UAVs.

Then, the following user-specified error constraint is imposed:

$$-\underline{k}_v \xi_v < e_{iv} < \overline{k}_v \xi_v \tag{14}$$

where $v = 1$–3. $\underline{k}_v$ and $\overline{k}_v$ are positive design parameters. ξ_v is the PPF. The initial error $e_{iv}(0)$ satisfies $-\underline{k}_v(0)\xi_v < e_{iv}(0) < \overline{k}_v\xi_v(0)$.

In this article, a finite-time PPF is chosen as follows:

$$\xi_v = \begin{cases} \text{csch}\left(\xi_{v0}\dfrac{a_T}{a_T - t}\right) + \xi_{vT}, & 0 \le t < a_T \\ \xi_{vT}, & t \ge a_T \end{cases} \tag{15}$$

where $\text{csch}(\cdot)$ is the hyperbolic cosecant function. ξ_{v0} and ξ_{vT} are positive design parameters, satisfying $\text{csch}(\xi_{v0}) > \xi_{vT}$. a_T represents the user-specified time constant. The PPF is strictly decreasing during $t \in [0, a_T)$ and reaches to the user-specified minimal value ξ_{vT} during $t \in [a_T, \infty)$, thus ensuring the finite-time convergence of the error $\boldsymbol{e}_i$. Moreover, the PPF is continuous on the time span $[0, \infty)$ [29].

Remark 3: As illustrated in Fig. 2, compared with the commonly investigated PPC with traditional PPF $\xi_v = (\xi_{v0} - \xi_{v\infty})e^{-\zeta t} + \xi_{v\infty}$ [49], which can only ensure the exponential convergence and ξ_v reaches the expected value $\xi_{v\infty}$ at $t = \infty$, the finite-time PPF (15) is adopted in this article, such that the user-specified time a_T can be involved. With this strategy, the distributed tracking errors are strictly confined within the user-specified bounds $(-\underline{k}_v\xi_v, \overline{k}_v\xi_v)$ in finite time and eventually stay within the bounds $(-\underline{k}_v\xi_{vT}, \overline{k}_v\xi_{vT})$ against unexpected faults, thus significantly enhancing the finite-time FTFC capabilities. Moreover, different FTFC requirements can be imposed on the formation team by adjusting the values of ξ_{vT} and a_T.

Remark 4: To date, the finite-time PPFs have been gradually used to facilitate the constraint controller design for

stochastic nonlinear system [29], switched nonlinear system [50], and quadrotor UAV [28]. Especially, only several results have been reported on the FTC investigations, such as [50], [51], and [52]. However, different from the existing results on finite-time PPFs, the research topic of this article is the FTFC problem of networked high-speed fixed-wing UAVs against faults. The high-speed characteristics of fixed-wing UAVs increase the finite-time PPC design challenge. Moreover, the existing results are mainly focused on the finite-time PPC design for a single system. For the networked systems, very rare results have been reported, let alone the FTFC design for networked fixed-wing UAVs. In such a situation, the fault effects will rapidly spread through the communication links due to the high flight speeds, causing the performance degradation of neighboring UAVs. Therefore, the FTFC design for networked fixed-wing UAVs with finite-time PPFs is still an open topic.

To facilitate the control design, the error constraint inequality (14) can be transformed into the following equality:

$$e_{iv} = \xi_v \phi_{iv}(\eta_{iv}) \tag{16}$$

where $\phi_{iv}(\eta_{iv})$ can be expressed as follows:

$$\phi_{iv}(\eta_{iv}) = \frac{\bar{k}_v e^{\eta_{iv}} - \underline{k}_v e^{-\eta_{iv}}}{e^{\eta_{iv}} + e^{-\eta_{iv}}}. \tag{17}$$

The inverse function of (17) is derived as follows:

$$\eta_{iv} = \phi_{iv}^{-1}\left(\frac{e_{iv}}{\xi_v}\right) = \frac{1}{2}\ln\frac{\phi_{iv}(\eta_{iv}) + \underline{k}_v}{\bar{k}_v - \phi_{iv}(\eta_{iv})} \tag{18}$$

where $\phi_{iv}(\eta_{iv})$ can be derived as $\phi_{iv}(\eta_{iv}) = e_{iv}/\xi_v$ from (16).

By defining the transformed error as $E_{iv} = \eta_{iv} - (1/2)\ln(\underline{k}_v/\bar{k}_v)$, the control problem of finite-time error constraints in the presence of faults is converted to the control problem of uniformly ultimate boundedness of E_{iv} against faults.

Differentiating the transformed error E_{iv} yields

$$\begin{aligned}\dot{E}_{iv} &= \left[\frac{1}{2\xi_v}\left(\frac{1}{\phi_{iv} + \underline{k}_v} - \frac{1}{\phi_{iv} - \bar{k}_v}\right)\right]\cdot\left[\dot{e}_{iv} - \frac{e_{iv}\dot{\xi}_{iv}}{\xi_{iv}}\right] \\ &= \Xi_{iv}\left(\dot{e}_{iv} - \frac{e_{iv}\dot{\xi}_{iv}}{\xi_{iv}}\right)\end{aligned} \tag{19}$$

where $\Xi_{iv} = [(1/2\xi_v)((1/(\phi_{iv} + \underline{k}_v)) - (1/(\phi_{iv} - \bar{k}_v)))]$.

The compact form of (19) can be expressed as follows:

$$\dot{\boldsymbol{E}}_i = \boldsymbol{\Xi}_i\left(\dot{\boldsymbol{e}}_i - \boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i\boldsymbol{e}_i\right) \tag{20}$$

where $\boldsymbol{E}_i = [E_{i1}, E_{i2}, E_{i3}]^T$, $\dot{\boldsymbol{E}}_i = [\dot{E}_{i1}, \dot{E}_{i2}, \dot{E}_{i3}]^T$, $\boldsymbol{\Xi}_i = \text{diag}\{\Xi_{i1}, \Xi_{i2}, \Xi_{i3}\}$, $\dot{\boldsymbol{e}}_i = [\dot{e}_{i1}, \dot{e}_{i2}, \dot{e}_{i3}]^T$, $\boldsymbol{\xi}_i = \text{diag}\{\xi_{i1}, \xi_{i2}, \xi_{i3}\}$, and $\dot{\boldsymbol{\xi}}_i = \text{diag}\{\dot{\xi}_{i1}, \dot{\xi}_{i2}, \dot{\xi}_{i3}\}$.

By differentiating (20) again, one can obtain

$$\begin{aligned}\ddot{\boldsymbol{E}}_i = {} & \dot{\boldsymbol{\Xi}}_i\dot{\boldsymbol{e}}_i - \dot{\boldsymbol{\Xi}}_i\boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i\boldsymbol{e}_i + \boldsymbol{\Xi}_i\ddot{\boldsymbol{e}}_i + \boldsymbol{\Xi}_i\boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i\boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i\boldsymbol{e}_i \\ & - \boldsymbol{\Xi}_i\boldsymbol{\xi}_i^{-1}\ddot{\boldsymbol{\xi}}_i\boldsymbol{e}_i - \boldsymbol{\Xi}_i\boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i\dot{\boldsymbol{e}}_i.\end{aligned} \tag{21}$$

B. FO FTFC Design With Reinforcement Learning Mechanism

In this section, the FO FTFC scheme is developed based on (21). By recalling (13), one has

$$\begin{aligned}\ddot{\boldsymbol{e}}_i = {} & \left(\sum_{j\in N_i} a_{ij} + b_i\right)\ddot{\boldsymbol{q}}_i - \sum_{j\in N_i} a_{ij}\ddot{\boldsymbol{q}}_j - \sum_{j\in N_i} a_{ij}\ddot{\boldsymbol{\delta}}_{ij} \\ & - b_i\ddot{\boldsymbol{q}}_0 - b_i\ddot{\boldsymbol{\delta}}_i.\end{aligned} \tag{22}$$

Then, (21) is further obtained as follows:

$$\begin{aligned}\ddot{\boldsymbol{E}}_i = {} & \boldsymbol{\Xi}_i\left[\left(\sum_{j\in N_i} a_{ij} + b_i\right)\left(\boldsymbol{f}_i + \boldsymbol{g}_i\boldsymbol{\Lambda}_i\boldsymbol{u}_{i0} + \boldsymbol{g}_i\boldsymbol{b}_{if}\right)\right. \\ & \left. - \sum_{j\in N_i} a_{ij}\ddot{\boldsymbol{q}}_j - \sum_{j\in N_i} a_{ij}\ddot{\boldsymbol{\delta}}_{ij} - b_i\ddot{\boldsymbol{q}}_0 - b_i\ddot{\boldsymbol{\delta}}_i\right] \\ & + \dot{\boldsymbol{\Xi}}_i\dot{\boldsymbol{e}}_i - \dot{\boldsymbol{\Xi}}_i\boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i\boldsymbol{e}_i + \boldsymbol{\Xi}_i\boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i\boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i\boldsymbol{e}_i \\ & - \boldsymbol{\Xi}_i\boldsymbol{\xi}_i^{-1}\ddot{\boldsymbol{\xi}}_i\boldsymbol{e}_i - \boldsymbol{\Xi}_i\boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i\dot{\boldsymbol{e}}_i.\end{aligned} \tag{23}$$

To achieve refined adjustment performance against faults, the following FO sliding-mode error is proposed:

$$\boldsymbol{S}_i = \dot{\boldsymbol{E}}_i + \lambda_{11}D^{a-1}\left(\text{sig}^{\lambda_{12}}(\boldsymbol{E}_i)\right) \tag{24}$$

where $\lambda_{11} > 0$ and $\lambda_{12} > 0$. $a \in (0, 1)$ is the FO operator. $\boldsymbol{S}_i = [s_{i1}, s_{i2}, s_{i3}]^T$.

Based on (23), the FO sliding-mode error dynamics can be obtained as follows:

$$\begin{aligned}\dot{\boldsymbol{S}}_i = {} & \boldsymbol{\Xi}_i\left[\left(\sum_{j\in N_i} a_{ij} + b_i\right)\left(\boldsymbol{f}_i + \boldsymbol{g}_i\boldsymbol{\Lambda}_i\boldsymbol{u}_{i0} + \boldsymbol{g}_i\boldsymbol{b}_{if}\right)\right. \\ & \left. - \sum_{j\in N_i} a_{ij}\ddot{\boldsymbol{q}}_j - \sum_{j\in N_i} a_{ij}\ddot{\boldsymbol{\delta}}_{ij} - b_i\ddot{\boldsymbol{q}}_0 - b_i\ddot{\boldsymbol{\delta}}_i\right] \\ & + \dot{\boldsymbol{\Xi}}_i\dot{\boldsymbol{e}}_i - \dot{\boldsymbol{\Xi}}_i\boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i\boldsymbol{e}_i + \boldsymbol{\Xi}_i\boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i\boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i\boldsymbol{e}_i \\ & - \boldsymbol{\Xi}_i\boldsymbol{\xi}_i^{-1}\ddot{\boldsymbol{\xi}}_i\boldsymbol{e}_i - \boldsymbol{\Xi}_i\boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i\dot{\boldsymbol{e}}_i \\ & + \lambda_{11}D^a\left(\text{sig}^{\lambda_{12}}(\boldsymbol{E}_i)\right) \\ = {} & \left(\sum_{j\in N_i} a_{ij} + b_i\right)\boldsymbol{\Xi}_i\boldsymbol{f}_i + \left(\sum_{j\in N_i} a_{ij} + b_i\right)\boldsymbol{\Xi}_i\boldsymbol{g}_i\boldsymbol{\Lambda}_i\boldsymbol{u}_{i0} \\ & + \left(\sum_{j\in N_i} a_{ij} + b_i\right)\left(\boldsymbol{\Xi}_i\boldsymbol{g}_i\boldsymbol{b}_{if} - \underline{\lambda}_i\boldsymbol{\Xi}_i\boldsymbol{g}_i\boldsymbol{u}_{i0}\right) \\ & + \left(\sum_{j\in N_i} a_{ij} + b_i\right)\underline{\lambda}_i\boldsymbol{\Xi}_i\boldsymbol{g}_i\boldsymbol{u}_{i0} \\ & - \boldsymbol{\Xi}_i\left(\sum_{j\in N_i} a_{ij}\ddot{\boldsymbol{q}}_j + \sum_{j\in N_i} a_{ij}\ddot{\boldsymbol{\delta}}_{ij} + b_i\ddot{\boldsymbol{q}}_0 + b_i\ddot{\boldsymbol{\delta}}_i\right) \\ & + \dot{\boldsymbol{\Xi}}_i\dot{\boldsymbol{e}}_i - \dot{\boldsymbol{\Xi}}_i\boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i\boldsymbol{e}_i + \boldsymbol{\Xi}_i\boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i\boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i\boldsymbol{e}_i \\ & - \boldsymbol{\Xi}_i\boldsymbol{\xi}_i^{-1}\ddot{\boldsymbol{\xi}}_i\boldsymbol{e}_i - \boldsymbol{\Xi}_i\boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i\dot{\boldsymbol{e}}_i \\ & + \lambda_{11}D^a\left(\text{sig}^{\lambda_{12}}(\boldsymbol{E}_i)\right).\end{aligned} \tag{25}$$

One can further obtain

$$\begin{aligned}\dot{\boldsymbol{S}}_i = {} & \boldsymbol{F}_{if} + \left(\sum_{j\in N_i} a_{ij} + b_i\right)\underline{\lambda}_i \boldsymbol{\Xi}_i \boldsymbol{g}_i \boldsymbol{u}_{i0} \\ & - \boldsymbol{\Xi}_i \left(\sum_{j\in N_i} a_{ij}\ddot{\boldsymbol{\delta}}_{ij} + b_i \ddot{\boldsymbol{q}}_0 + b_i \ddot{\boldsymbol{\delta}}_i\right) \\ & + \lambda_{11} D^a \left(\mathrm{sig}^{\lambda_{12}}(\boldsymbol{E}_i)\right) + \boldsymbol{\varepsilon}_{i0} \end{aligned} \tag{26}$$

where $\boldsymbol{F}_{if} = (\sum_{j\in N_i} a_{ij} + b_i)\boldsymbol{\Xi}_i \boldsymbol{f}_i + (\sum_{j\in N_i} a_{ij} + b_i)\boldsymbol{\Xi}_i \boldsymbol{g}_i \boldsymbol{\Lambda}_i \boldsymbol{u}_{i0f} - (\sum_{j\in N_i} a_{ij} + b_i)\underline{\lambda}_i \boldsymbol{\Xi}_i \boldsymbol{g}_i \boldsymbol{u}_{i0f} + (\sum_{j\in N_i} a_{ij} + b_i)\boldsymbol{\Xi}_i \boldsymbol{g}_i \boldsymbol{b}_{if} - \boldsymbol{\Xi}_i \sum_{j\in N_i} a_{ij}\ddot{\boldsymbol{q}}_j + \dot{\boldsymbol{\Xi}}_i \dot{\boldsymbol{e}}_i - \dot{\boldsymbol{\Xi}}_i \boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i \boldsymbol{e}_i + \boldsymbol{\Xi}_i \boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i \boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i \boldsymbol{e}_i - \boldsymbol{\Xi}_i \boldsymbol{\xi}_i^{-1}\ddot{\boldsymbol{\xi}}_i \boldsymbol{e}_i - \boldsymbol{\Xi}_i \boldsymbol{\xi}_i^{-1}\dot{\boldsymbol{\xi}}_i \dot{\boldsymbol{e}}_i$. $\boldsymbol{\varepsilon}_{i0}$ is the error caused by breaking the algebraic loop due to the involvement of the signal $\boldsymbol{u}_{i0}$ in $\boldsymbol{F}_i$, i.e., $\boldsymbol{F}_i = \boldsymbol{F}_{if} + \boldsymbol{\varepsilon}_{i0}$. $\boldsymbol{u}_{i0f}$ is the filtered signal. For more details about the algebraic loop-breaking methods, please refer to [53] and [54].

To evaluate the FO FTFC performance of each UAV, the long-term performance index is defined as follows:

$$\boldsymbol{\Upsilon}_i(t) = \int_t^\infty \sigma_i^{\frac{-\tau+t}{T}} \boldsymbol{\zeta}_i(\boldsymbol{S}_i(\tau))\, d\tau \tag{27}$$

where T is a positive integral reinforcement interval and $0 < \sigma_i < 1$ is the discount factor. $\boldsymbol{\Upsilon}_i(t) = [\Upsilon_{i1}(t), \Upsilon_{i2}(t), \Upsilon_{i3}(t)]^T$, $\boldsymbol{\zeta}_i(\boldsymbol{S}_i(\tau)) = [\zeta_{i1}(S_{i1}(\tau)), \zeta_{i2}(S_{i2}(\tau)), \zeta_{i3}(S_{i3}(\tau))]^T$, and the corresponding element can be defined as follows:

$$\begin{aligned}\zeta_{iv}(S_{iv}(\tau)) &= \begin{cases} 0, & S_{iv}(\tau)^2 \le C_{\zeta_v} \\ 1, & S_{iv}(\tau)^2 > C_{\zeta_v} \end{cases} \\ \tau &\in [t-T, t), \quad v = 1\text{–}3 \end{aligned} \tag{28}$$

and $C_{\zeta_v} > 0$.

The relationship between $\boldsymbol{\Upsilon}_i(t-T)$ and $\boldsymbol{\Upsilon}_i(t)$ can be obtained as follows:

$$\begin{aligned}\boldsymbol{\Upsilon}_i(t-T) &= \int_{t-T}^\infty \sigma_i^{\frac{-\tau+t-T}{T}} \boldsymbol{\zeta}_i(\boldsymbol{S}_i(\tau))\, d\tau \\ &= \sigma_i^{-1}\boldsymbol{\Upsilon}_i(t) + \int_{t-T}^t \sigma_i^{\frac{-\tau+t-T}{T}} \boldsymbol{\zeta}_i(\boldsymbol{S}_i(\tau))\, d\tau \\ &= \sigma_i^{-1}(\boldsymbol{\Upsilon}_i(t) + \boldsymbol{P}_{ic}(t)) \end{aligned} \tag{29}$$

where $\boldsymbol{P}_{ic}(t) = [P_{ic1}(t), P_{ic2}(t), P_{ic3}(t)]^T = \int_{t-T}^t \sigma_i^{((-\tau+t)/T)} \boldsymbol{\zeta}_i(\boldsymbol{S}_i(\tau))d\tau$.

By substituting (28) into the expression of the element $P_{icv}(t)$, $v = 1$–3, one has

$$P_{icv}(t) = \begin{cases} 0, & S_{iv}(\tau)^2 \le C_{\zeta_v} \\ \dfrac{T}{\ln\sigma_i}(\sigma_i - 1), & S_{iv}(\tau)^2 > C_{\zeta_v}. \end{cases} \tag{30}$$

Consider the fact that the long-term performance index (27) contains future information, it is very difficult to obtain the solution at the current time instant. To solve such a difficulty, the following critic RBF NN is used:

$$\boldsymbol{\Upsilon}_i(t) = \boldsymbol{W}_{ic}^{*T}\boldsymbol{\varphi}_{ic}(\boldsymbol{r}_{ic}) + \boldsymbol{\varepsilon}_{ic}^*(\boldsymbol{r}_{ic}) \tag{31}$$

where $\boldsymbol{W}_{ic}^*$ and $\boldsymbol{\varepsilon}_{ic}^*(\boldsymbol{r}_{ic})$ are the bounded ideal weighting matrix and minimum approximation error vector of the critic RBF NN for the ith UAV, respectively. $\boldsymbol{\varphi}_{ic}(\boldsymbol{r}_{ic})$ is the Gaussian basis function vector, and it is bounded according to the expression of (11), i.e., $\|\boldsymbol{\varphi}_{ic}\| \le \bar{\varphi}_{ic}$.

Then, the learned results can be described as follows:

$$\hat{\boldsymbol{\Upsilon}}_i(t) = \hat{\boldsymbol{W}}_{ic}^T \boldsymbol{\varphi}_{ic}(\boldsymbol{r}_{ic}(t)) \tag{32}$$

$$\hat{\boldsymbol{\Upsilon}}_i(t-T) = \hat{\boldsymbol{W}}_{ic}^T \boldsymbol{\varphi}_{ic}(\boldsymbol{r}_{ic}(t-T)). \tag{33}$$

The temporal difference error is defined as follows:

$$\begin{aligned}\boldsymbol{e}_{i\Upsilon c} &= \hat{\boldsymbol{\Upsilon}}_i(t) - \sigma_i \hat{\boldsymbol{\Upsilon}}_i(t-T) + \boldsymbol{P}_{ic}(t) \\ &= \hat{\boldsymbol{W}}_{ic}^T \Delta\boldsymbol{\varphi}_{ic}(t) + \boldsymbol{P}_{ic}(t) \\ &= \tilde{\boldsymbol{W}}_{ic}^T \Delta\boldsymbol{\varphi}_{ic}(t) + \boldsymbol{P}_{ic}(t) + \boldsymbol{W}_{ic}^{*T}\Delta\boldsymbol{\varphi}_{ic}(t) \end{aligned} \tag{34}$$

where $\Delta\boldsymbol{\varphi}_{ic}(t) = \boldsymbol{\varphi}_{ic}(t) - \sigma_i\boldsymbol{\varphi}_{ic}(t-T)$ and is bounded by $\|\Delta\boldsymbol{\varphi}_{ic}(t)\| \le (1+\sigma_i)\bar{\varphi}_{ic}$.

To handle the unknown nonlinear term $\boldsymbol{F}_i$ in (26), the following actor RBF NN is developed:

$$\boldsymbol{F}_{if} = \Gamma_1^{-1}\boldsymbol{W}_{ia}^{*T}\boldsymbol{\varphi}_{ia}(\boldsymbol{r}_{ia}) + \Gamma_1^{-1}\boldsymbol{\varepsilon}_{ia} \tag{35}$$

where Γ_1 is a design parameter, and $\boldsymbol{W}_{ia}^*$ and $\boldsymbol{\varepsilon}_{ia}$ are the bounded ideal weighting matrix and minimum approximation error vector of the actor RBF NN, respectively. $\boldsymbol{\varphi}_{ia}(\boldsymbol{r}_{ia})$ denotes the Gaussian basis function vector and is bounded as $\|\boldsymbol{\varphi}_{ia}\| \le \bar{\varphi}_{ia}$.

Then, the learned result of $\boldsymbol{F}_{if}$ is

$$\hat{\boldsymbol{F}}_{if} = \Gamma^{-1}\hat{\boldsymbol{W}}_{ia}^T\boldsymbol{\varphi}_{ia}(\boldsymbol{r}_{ia}) \tag{36}$$

where $\hat{\boldsymbol{W}}_{ia}$ is the estimated weighting matrix of $\boldsymbol{W}_{ia}^*$.

Moreover, (26) can be derived as follows:

$$\begin{aligned}\dot{\boldsymbol{S}}_i = {} & \Gamma_1^{-1}\boldsymbol{W}_{ia}^{*T}\boldsymbol{\varphi}_{ia}(\boldsymbol{r}_{ia}) + \boldsymbol{D}_i + \left(\sum_{j\in N_i} a_{ij} + b_i\right)\underline{\lambda}_i\boldsymbol{\Xi}_i\boldsymbol{g}_i\boldsymbol{u}_{i0} \\ & - \boldsymbol{\Xi}_i\left(\sum_{j\in N_i} a_{ij}\ddot{\boldsymbol{\delta}}_{ij} + b_i\ddot{\boldsymbol{q}}_0 + b_i\ddot{\boldsymbol{\delta}}_i\right) + \lambda_{11}D^a\left(\mathrm{sig}^{\lambda_{12}}(\boldsymbol{E}_i)\right) \end{aligned} \tag{37}$$

where $\boldsymbol{D}_i = \boldsymbol{\varepsilon}_{i0} + \Gamma_1^{-1}\boldsymbol{\varepsilon}_{ia}$ and $\dot{\boldsymbol{D}}_i^T\dot{\boldsymbol{D}}_i \le \bar{D}_i$.

Define the following auxiliary learning error:

$$\begin{cases} \boldsymbol{\Psi}_i = \boldsymbol{S}_i - \hat{\boldsymbol{S}}_i \\ \dot{\hat{\boldsymbol{S}}}_i = \Gamma_1^{-1}\hat{\boldsymbol{W}}_{ia}^T\boldsymbol{\varphi}_{ia} + \hat{\boldsymbol{D}}_i + \left(\sum_{j\in N_i} a_{ij} + b_i\right)\underline{\lambda}_i\boldsymbol{\Xi}_i\boldsymbol{g}_i\boldsymbol{u}_{i0} \\ \qquad - \boldsymbol{\Xi}_i\left(\sum_{j\in N_i} a_{ij}\ddot{\boldsymbol{\delta}}_{ij} + b_i\ddot{\boldsymbol{q}}_0 + b_i\ddot{\boldsymbol{\delta}}_i\right) \\ \qquad + \lambda_{11}D^a\left(\mathrm{sig}^{\lambda_{12}}(\boldsymbol{E}_i)\right) + \kappa_{01}\boldsymbol{\Psi}_i \end{cases} \tag{38}$$

where $\kappa_{01} > 0$ is a design parameter and $\hat{\boldsymbol{D}}_i$ is the estimation of $\boldsymbol{D}_i$.

Taking the time derivative of (38) yields

$$\begin{aligned}\dot{\boldsymbol{\Psi}} &= \dot{\boldsymbol{S}}_i - \dot{\hat{\boldsymbol{S}}}_i \\ &= \Gamma_1^{-1}\boldsymbol{W}_{ia}^{*T}\boldsymbol{\varphi}_{ia}(\boldsymbol{r}_{ia}) + \boldsymbol{D}_i + \left(\sum_{j\in N_i} a_{ij} + b_i\right)\underline{\lambda}_i\boldsymbol{\Xi}_i\boldsymbol{g}_i\boldsymbol{u}_{i0} \end{aligned}$$

$$
\begin{aligned}
&- \Xi_i \left(\sum_{j\in N_i} a_{ij}\ddot{\boldsymbol{\delta}}_{ij} + b_i\ddot{\boldsymbol{q}}_0 + b_i\ddot{\boldsymbol{\delta}}_i \right) \\
&+ \lambda_{11} D^a \left(\mathrm{sig}^{\lambda_{12}}(\boldsymbol{E}_i) \right) \\
&- \Bigg[\Gamma_1^{-1}\hat{\boldsymbol{W}}_{ia}^T\boldsymbol{\varphi}_{ia} + \hat{\boldsymbol{D}}_i + \left(\sum_{j\in N_i} a_{ij} + b_i \right) \underline{\lambda}_i \Xi_i \boldsymbol{g}_i \boldsymbol{u}_{i0} \\
&\qquad - \Xi_i \left(\sum_{j\in N_i} a_{ij}\ddot{\boldsymbol{\delta}}_{ij} + b_i\ddot{\boldsymbol{q}}_0 + b_i\ddot{\boldsymbol{\delta}}_i \right) \\
&\qquad + \lambda_{11} D^a \left(\mathrm{sig}^{\lambda_{12}}(\boldsymbol{E}_i) \right) + \kappa_{01}\boldsymbol{\Psi}_i \Bigg] \\
&= -\Gamma_1^{-1}\tilde{\boldsymbol{W}}_{ia}^T\boldsymbol{\varphi}_{ia} + \tilde{\boldsymbol{D}}_i - \kappa_{01}\boldsymbol{\Psi}_i
\end{aligned} \tag{39}
$$

where $\tilde{\boldsymbol{W}}_{ia} = \hat{\boldsymbol{W}}_{ia} - \boldsymbol{W}_{ia}^*$ and $\tilde{\boldsymbol{D}}_i = \boldsymbol{D}_i - \hat{\boldsymbol{D}}_i$.

Based on the auxiliary learning error $\boldsymbol{\Psi}_i$, the following nonlinear DO is developed:

$$
\begin{cases}
\hat{\boldsymbol{D}}_i = \bar{\boldsymbol{\omega}}_i + \kappa_{02}\boldsymbol{S}_i \\
\dot{\bar{\boldsymbol{\omega}}}_i = -\kappa_{02}\boldsymbol{\omega}_i \\
\qquad - \kappa_{02} \Bigg[\Gamma_1^{-1}\hat{\boldsymbol{W}}_{ia}^T\boldsymbol{\varphi}_{ia} + \left(\sum_{j\in N_i} a_{ij} + b_i \right) \\
\qquad\qquad \cdot \underline{\lambda}_i \Xi_i \boldsymbol{g}_i \boldsymbol{u}_{i0} - \Xi_i \left(\sum_{j\in N_i} a_{ij}\ddot{\boldsymbol{\delta}}_{ij} + b_i\ddot{\boldsymbol{q}}_0 + b_i\ddot{\boldsymbol{\delta}}_i \right) \\
\qquad\qquad + \lambda_{11} D^a \left(\mathrm{sig}^{\lambda_{12}}(\boldsymbol{E}_i) \right) + \kappa_{02}^{-1}(\kappa_{03}\boldsymbol{\Psi}_i + \boldsymbol{S}_i) \Bigg]
\end{cases} \tag{40}
$$

where $\kappa_{02} > 0$ and $\kappa_{03} > 0$ are design parameters.

By taking the time derivative of (40), one has

$$
\begin{aligned}
\dot{\hat{\boldsymbol{D}}}_i &= \dot{\bar{\boldsymbol{\omega}}}_i + \kappa_{02}\dot{\boldsymbol{S}}_i \\
&= -\kappa_{02}\left(\hat{\boldsymbol{D}}_i - \kappa_{02}\boldsymbol{S}_i \right) - \kappa_{02}\Gamma_1^{-1}\tilde{\boldsymbol{W}}_{ia}^T\boldsymbol{\varphi}_{ia} \\
&\quad + \kappa_{02}\boldsymbol{D}_i - \kappa_{02}^2\boldsymbol{S}_i + \kappa_{03}\boldsymbol{\Psi}_i + \boldsymbol{S}_i \\
&= \kappa_{02}\tilde{\boldsymbol{D}}_i - \kappa_{02}\Gamma_1^{-1}\tilde{\boldsymbol{W}}_{ia}^T\boldsymbol{\varphi}_{ia} + \kappa_{03}\boldsymbol{\Psi}_i + \boldsymbol{S}_i.
\end{aligned} \tag{41}
$$

The learning law of the critic NN in the reinforcement learning architecture is designed as follows:

$$
\dot{\hat{\boldsymbol{W}}}_{ic} = -\kappa_{11}\left[\Delta\boldsymbol{\varphi}_{ic}(t) \left(\hat{\boldsymbol{W}}_{ic}^T\Delta\boldsymbol{\varphi}_{ic}(t) + \boldsymbol{P}_{ic} \right)^T + \kappa_{12}\hat{\boldsymbol{W}}_{ic} \right] \tag{42}
$$

where $\kappa_{11} > 0$ and $\kappa_{12} > 0$.

Based on the critic RBF NN and the auxiliary learning error $\boldsymbol{\Psi}_i$, the following learning law of the actor NN is developed:

$$
\dot{\hat{\boldsymbol{W}}}_{ia} = \kappa_{21}\left[\Gamma_1^{-1}\boldsymbol{\varphi}_{ia}(\boldsymbol{r}_{ia}) \left(\boldsymbol{S}_i + \hat{\boldsymbol{W}}_{ic}^T\boldsymbol{\varphi}_{ic}(\boldsymbol{r}_{ic}) + \kappa_{03}\boldsymbol{\Psi}_i \right)^T - \kappa_{22}\hat{\boldsymbol{W}}_{ia} \right] \tag{43}
$$

where $\kappa_{21} > 0$ and $\kappa_{22} > 0$.

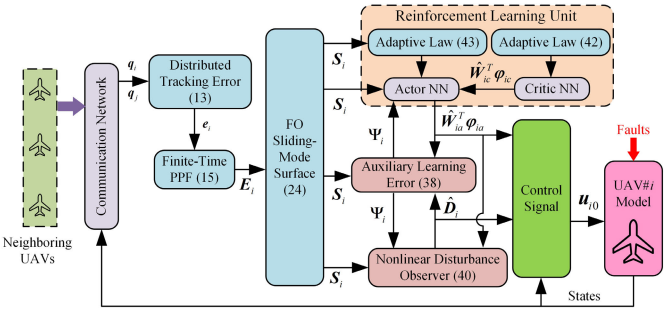


Fig. 3. Overall control architecture for the ith follower UAV.

Then, the control signal is developed as follows:

$$
\begin{aligned}
\boldsymbol{u}_{i0} = &\left[\left(\sum_{j\in N_i} a_{ij} + b_i \right) \underline{\lambda}_i \Xi_i \boldsymbol{g}_i \right]^{-1} \\
&\times \Bigg[-\boldsymbol{K}_2\boldsymbol{S}_i - \Gamma_1^{-1}\hat{\boldsymbol{W}}_{ia}^T\boldsymbol{\varphi}_{ia} \\
&\quad - \hat{\boldsymbol{D}}_i + \Xi_i \left(\sum_{j\in N_i} a_{ij}\ddot{\boldsymbol{\delta}}_{ij} + b_i\ddot{\boldsymbol{q}}_0 + b_i\ddot{\boldsymbol{\delta}}_i \right) \\
&\quad - \lambda_{11} D^a \left(\mathrm{sig}^{\lambda_{12}}(\boldsymbol{E}_i) \right) \Bigg]
\end{aligned} \tag{44}
$$

where $\boldsymbol{K}_2$ is a positive diagonal matrix.

To this end, the overall control architecture for the ith fixed-wing UAV is illustrated in Fig. 3. Moreover, Algorithm 1 is presented to clarify the structure of the developed reinforcement learning-based FO FTFC scheme for the ith fixed-wing UAV.

Algorithm 1 Reinforcement Learning-Based FO FTFC Method

Initialization:

Set $\hat{\boldsymbol{S}}_i = \boldsymbol{0}$, $\bar{\boldsymbol{\omega}}_i = \boldsymbol{0}$, $\hat{\boldsymbol{W}}_{ic} = \boldsymbol{0}$, $\hat{\boldsymbol{W}}_{ia} = \boldsymbol{0}$ for $i = 1, 2, \ldots, N$.

for follower fixed-wing UAV#i **do**

Compute $\boldsymbol{e}_i$ from (13)
Compute $\boldsymbol{E}_i$ from (18)
Compute $\boldsymbol{S}_i$ from (24)
Compute $\boldsymbol{\Psi}_i$ from the first equation in (38)
Compute $\hat{\boldsymbol{D}}_i$ from the first equation in (40)
Compute $\boldsymbol{u}_{i0}$ from (44)
Update $\hat{\boldsymbol{S}}_i$ from the second equation in (38)
Update $\bar{\boldsymbol{\omega}}_i$ from the second equation in (40)
Update $\hat{\boldsymbol{W}}_{ic}$ from (42)
Update $\hat{\boldsymbol{W}}_{ia}$ from (43)

end for

C. Stability Analysis

Theorem 1: Considering N follower fixed-wing UAVs and one leader fixed-wing UAV with undirected communications in a connected graph, if constraint (14) and the finite-time PPF (15) are combined for transforming the distributed tracking

error (13), the FO error is constructed as (24) to filter the transformed error $\boldsymbol{E}_i$, the critic and actor NNs are developed as (32) and (36) to, respectively, approximate the long-term performance index and unknown term with the adaptive learning laws (42) and (43), the auxiliary learning error is constructed as (38), the nonlinear DO is designed as (40), and the control signal is developed as (44), then all follower UAVs can track the leader UAV with predesigned offsets, and the distributed errors are strictly confined within the user-specified bounds.

Proof: Choose the following Lyapunov function:

$$L_i = \frac{1}{2}\boldsymbol{S}_i^T\boldsymbol{S}_i + \frac{1}{2\kappa_{21}}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ia}^T\tilde{\boldsymbol{W}}_{ia}\right] + \frac{1}{2\kappa_{11}}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^T\tilde{\boldsymbol{W}}_{ic}\right] + \frac{1}{2}\kappa_{03}\boldsymbol{\Psi}_i^T\boldsymbol{\Psi}_i + \frac{1}{2}\tilde{\boldsymbol{D}}_i^T\tilde{\boldsymbol{D}}_i \tag{45}$$

where $\tilde{\boldsymbol{W}}_{ic} = \hat{\boldsymbol{W}}_{ic} - \boldsymbol{W}_{ic}^*$.

Differentiating (45) has

$$\begin{aligned}
\dot{L}_i =& \boldsymbol{S}_i^T\dot{\boldsymbol{S}}_i + \frac{1}{\kappa_{21}}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ia}^T\left(\dot{\hat{\boldsymbol{W}}}_{ia} - \dot{\boldsymbol{W}}_{ia}^*\right)\right] \\
&+ \frac{1}{\kappa_{11}}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^T\left(\dot{\hat{\boldsymbol{W}}}_{ic} - \dot{\boldsymbol{W}}_{ic}^*\right)\right] + \kappa_{03}\boldsymbol{\Psi}_i^T\dot{\boldsymbol{\Psi}}_i + \tilde{\boldsymbol{D}}_i^T\dot{\tilde{\boldsymbol{D}}}_i \\
=& \boldsymbol{S}_i^T\Bigg[\boldsymbol{F}_{if} + \boldsymbol{\varepsilon}_{i0} + \left(\sum_{j\in N_i} a_{ij} + b_i\right)\underline{\lambda}_i\boldsymbol{\Xi}_i\boldsymbol{g}_i\boldsymbol{u}_{i0} \\
&- \boldsymbol{\Xi}_i\left(\sum_{j\in N_i} a_{ij}\ddot{\boldsymbol{\delta}}_{ij} + b_i\ddot{\boldsymbol{q}}_0 + b_i\ddot{\boldsymbol{\delta}}_i\right) \\
&+ \lambda_{11}D^a\left(\mathrm{sig}^{\lambda_{12}}(\boldsymbol{E}_i)\right)\Bigg] \\
&+ \frac{1}{\kappa_{21}}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ia}^T\dot{\hat{\boldsymbol{W}}}_{ia}\right] + \frac{1}{\kappa_{11}}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^T\dot{\hat{\boldsymbol{W}}}_{ic}\right] \\
&+ \kappa_{03}\boldsymbol{\Psi}_i^T\dot{\boldsymbol{\Psi}}_i + \tilde{\boldsymbol{D}}_i^T\dot{\tilde{\boldsymbol{D}}}_i \\
=& \boldsymbol{S}_i^T\Bigg[\Gamma_1^{-1}\boldsymbol{W}_{ia}^{*T}\boldsymbol{\varphi}_{ia} + \boldsymbol{D}_i - \boldsymbol{K}_2\boldsymbol{S}_i - \Gamma_1^{-1}\hat{\boldsymbol{W}}_{ia}^T\boldsymbol{\varphi}_{ia} - \hat{\boldsymbol{D}}_i \\
&+ \boldsymbol{\Xi}_i\left(\sum_{j\in N_i} a_{ij}\ddot{\boldsymbol{\delta}}_{ij} + b_i\ddot{\boldsymbol{q}}_0 + b_i\ddot{\boldsymbol{\delta}}_i\right) \\
&- \lambda_{11}D^a\left(\mathrm{sig}^{\lambda_{12}}(\boldsymbol{E}_i)\right) \\
&- \boldsymbol{\Xi}_i\left(\sum_{j\in N_i} a_{ij}\ddot{\boldsymbol{\delta}}_{ij} + b_i\ddot{\boldsymbol{q}}_0 + b_i\ddot{\boldsymbol{\delta}}_i\right) \\
&+ \lambda_{11}D^a\left(\mathrm{sig}^{\lambda_{12}}(\boldsymbol{E}_i)\right)\Bigg] \\
&+ \frac{1}{\kappa_{21}}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ia}^T\dot{\hat{\boldsymbol{W}}}_{ia}\right] + \frac{1}{\kappa_{11}}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^T\dot{\hat{\boldsymbol{W}}}_{ic}\right] \\
&+ \kappa_{03}\boldsymbol{\Psi}_i^T\dot{\boldsymbol{\Psi}}_i + \tilde{\boldsymbol{D}}_i^T\dot{\tilde{\boldsymbol{D}}}_i.
\end{aligned} \tag{46}$$

Substituting (43) into (46) yields

$$\begin{aligned}
\dot{L}_i =& -\Gamma_1^{-1}\boldsymbol{S}_i^T\tilde{\boldsymbol{W}}_{ia}^T\boldsymbol{\varphi}_{ia} + \boldsymbol{S}_i^T\tilde{\boldsymbol{D}}_i - \boldsymbol{S}_i^T\boldsymbol{K}_2\boldsymbol{S}_i \\
&+ \frac{1}{\kappa_{21}}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ia}^T\dot{\hat{\boldsymbol{W}}}_{ia}\right] + \frac{1}{\kappa_{11}}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^T\dot{\hat{\boldsymbol{W}}}_{ic}\right] \\
&+ \kappa_{03}\boldsymbol{\Psi}_i^T\dot{\boldsymbol{\Psi}}_i + \tilde{\boldsymbol{D}}_i^T\dot{\tilde{\boldsymbol{D}}}_i \\
=& -\boldsymbol{S}_i^T\boldsymbol{K}_2\boldsymbol{S}_i - \Gamma_1^{-1}\boldsymbol{S}_i^T\tilde{\boldsymbol{W}}_{ia}^T\boldsymbol{\varphi}_{ia} + \boldsymbol{S}_i^T\tilde{\boldsymbol{D}}_i \\
&+ \mathrm{tr}\left[\Gamma_1^{-1}\tilde{\boldsymbol{W}}_{ia}^T\boldsymbol{\varphi}_{ia}(\boldsymbol{r}_{ia})\left[\boldsymbol{S}_i + \hat{\boldsymbol{W}}_{ic}\boldsymbol{\varphi}_{ic}(\boldsymbol{r}_{ic}) + \kappa_{03}\boldsymbol{\Psi}_i\right]^T\right. \\
&\left. - \kappa_{22}\tilde{\boldsymbol{W}}_{ia}\hat{\boldsymbol{W}}_{ia}\right] + \frac{1}{\kappa_{11}}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^T\dot{\hat{\boldsymbol{W}}}_{ic}\right] \\
&+ \kappa_{03}\boldsymbol{\Psi}_i^T\dot{\boldsymbol{\Psi}}_i + \tilde{\boldsymbol{D}}_i^T\dot{\tilde{\boldsymbol{D}}}_i \\
=& -\boldsymbol{S}_i^T\boldsymbol{K}_2\boldsymbol{S}_i - \Gamma_1^{-1}\boldsymbol{S}_i^T\tilde{\boldsymbol{W}}_{ia}^T\boldsymbol{\varphi}_{ia} + \underbrace{\boldsymbol{S}_i^T\tilde{\boldsymbol{D}}_i}_{\text{Term 1}} \\
&+ \mathrm{tr}\left[\Gamma_1^{-1}\tilde{\boldsymbol{W}}_{ia}^T\boldsymbol{\varphi}_{ia}(\boldsymbol{r}_{ia})\boldsymbol{S}_i^T\right] \\
&+ \underbrace{\mathrm{tr}\left[\Gamma_1^{-1}\tilde{\boldsymbol{W}}_{ia}^T\boldsymbol{\varphi}_{ia}(\boldsymbol{r}_{ia})\boldsymbol{\varphi}_{ic}^T\hat{\boldsymbol{W}}_{ic}\right]}_{\text{Term 2}} \\
&- \underbrace{\mathrm{tr}\left[\kappa_{22}\tilde{\boldsymbol{W}}_{ia}^T\hat{\boldsymbol{W}}_{ia}\right]}_{\text{Term 3}} + \underbrace{\frac{1}{\kappa_{11}}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^T\dot{\hat{\boldsymbol{W}}}_{ic}\right]}_{\text{Term 4}} \\
&+ \Gamma_1^{-1}\kappa_{03}\mathrm{tr}\left(\tilde{\boldsymbol{W}}_{ia}^T\boldsymbol{\varphi}_{ia}(\boldsymbol{r}_{ia})\boldsymbol{\Psi}_i^T\right) \\
&+ \kappa_{03}\boldsymbol{\Psi}_i^T\dot{\boldsymbol{\Psi}}_i + \tilde{\boldsymbol{D}}_i^T\dot{\tilde{\boldsymbol{D}}}_i.
\end{aligned} \tag{47}$$

With respect to Term 1 in (47), one has

$$\boldsymbol{S}_i^T\tilde{\boldsymbol{D}}_i \le \frac{1}{2\pi_0^2}\boldsymbol{S}_i^T\boldsymbol{S}_i + \frac{\pi_0^2}{2}\tilde{\boldsymbol{D}}_i^T\tilde{\boldsymbol{D}}_i \tag{48}$$

where π_0 is a positive constant.

For Term 2 in (47), the following inequality can be obtained:

$$\mathrm{tr}\left[\Gamma_1^{-1}\tilde{\boldsymbol{W}}_{ia}^T\boldsymbol{\varphi}_{ia}(\boldsymbol{r}_{ia})\boldsymbol{\varphi}_{ic}^T\hat{\boldsymbol{W}}_{ic}\right] \le \Gamma_1^{-1}\bar{\varphi}_{ia}\bar{\varphi}_{ic}\left\|\tilde{\boldsymbol{W}}_{ia}\right\|_F\cdot\left\|\tilde{\boldsymbol{W}}_{ic}\right\|_F \tag{49}$$

where $\|\boldsymbol{\varphi}_{ia}(\cdot)\| \le \bar{\varphi}_{ia}$ and $\|\boldsymbol{\varphi}_{ic}(\cdot)\| \le \bar{\varphi}_{ic}$ are used in (49).

For Term 3 in (47), one can further obtain

$$-\mathrm{tr}\left[\kappa_{22}\tilde{\boldsymbol{W}}_{ia}^T\hat{\boldsymbol{W}}_{ia}\right] \le -\frac{\kappa_{22}}{2}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ia}^T\tilde{\boldsymbol{W}}_{ia}\right] + \frac{\kappa_{22}\bar{w}_a^2}{2} \tag{50}$$

where $\|\boldsymbol{W}_{ia}^*\| \le \bar{w}_{ia}$ is used in (50).

Regarding Term 4, one has

$$\begin{aligned}
&\frac{1}{\kappa_{11}}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^T\dot{\hat{\boldsymbol{W}}}_{ic}\right] \\
&= \mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^T\left[-\Delta\boldsymbol{\varphi}_{ic}(t)\left[\hat{\boldsymbol{W}}_{ic}^T\Delta\boldsymbol{\varphi}_{ic}(t) + \boldsymbol{P}_{ic}\right]^T\right.\right. \\
&\quad\left.\left. - \kappa_{12}\hat{\boldsymbol{W}}_{ic}\right]\right] \\
&= -\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^T\Delta\boldsymbol{\varphi}_{ic}(t)\left[\hat{\boldsymbol{W}}_{ic}^T\Delta\boldsymbol{\varphi}_{ic}(t) + \boldsymbol{P}_{ic}\right]^T\right] \\
&\quad - \kappa_{12}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^T\hat{\boldsymbol{W}}_{ic}\right] \\
&= -\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^T\Delta\boldsymbol{\varphi}_{ic}(t)\tilde{\boldsymbol{W}}_{ic}^T\Delta\boldsymbol{\varphi}_{ic}(t)\right] \\
&\quad - \mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^T\Delta\boldsymbol{\varphi}_{ic}\left(\boldsymbol{W}_{ic}^{*T}\Delta\boldsymbol{\varphi}_{ic}(t) + \boldsymbol{P}_{ic}\right)\right]^T \\
&\quad - \kappa_{12}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^T\tilde{\boldsymbol{W}}_{ic}\right] - \kappa_{12}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^T\boldsymbol{W}_{ic}^*\right] \\
&\le -\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^T\left(\Delta\boldsymbol{\varphi}_{ic}(t)\Delta\boldsymbol{\varphi}_{ic}^T(t) + \kappa_{12}\boldsymbol{I}\right)\tilde{\boldsymbol{W}}_{ic}\right]
\end{aligned}$$

$$+\left\|\tilde{\boldsymbol{W}}_{ic}\right\|_F \cdot \underbrace{\left[\left\|\Delta\boldsymbol{\varphi}_{ic}(t)\right\| \cdot \left\|\boldsymbol{W}_{ic}^{*T}\Delta\boldsymbol{\varphi}_{ic}(t)+\boldsymbol{P}_{ic}\right\|+\kappa_{12}\left\|\boldsymbol{W}_{ic}^{*}\right\|_F\right]}_{\bar{\eta}_{i0}}$$

$$\leq -\kappa_{12}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^{T}\tilde{\boldsymbol{W}}_{ic}\right]+\bar{\eta}_{i1}\left\|\tilde{\boldsymbol{W}}_{ic}\right\|_F \tag{51}$$

where $\bar{\eta}_{i0} \leq \bar{\eta}_{i1}$.

Then, by incorporating (48)–(51) into (47), one has

$$\begin{aligned}\dot{L}_i \leq\ & -\boldsymbol{S}_i^T\boldsymbol{K}_2\boldsymbol{S}_i+\Gamma_1^{-1}\kappa_{03}\mathrm{tr}\left(\tilde{\boldsymbol{W}}_{ia}^{T}\boldsymbol{\varphi}_{ia}(\boldsymbol{r}_{ia})\boldsymbol{\Psi}_i^T\right)\\ &+\kappa_{03}\boldsymbol{\Psi}_i^T\dot{\boldsymbol{\Psi}}_i+\tilde{\boldsymbol{D}}_i^T\dot{\tilde{\boldsymbol{D}}}_i\\ &+\bar{\varphi}_{ia}\bar{\varphi}_{ic}\left\|\tilde{\boldsymbol{W}}_{ia}\right\|_F\cdot\left\|\tilde{\boldsymbol{W}}_{ic}\right\|_F\\ &-\frac{\kappa_{22}}{2}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ia}^{T}\tilde{\boldsymbol{W}}_{ia}\right]+\frac{\kappa_{22}\bar{w}_a^2}{2}-\kappa_{12}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^{T}\tilde{\boldsymbol{W}}_{ic}\right]\\ &+\bar{\eta}_i\left\|\tilde{\boldsymbol{W}}_{ic}\right\|_F+\frac{1}{2\pi_0^2}\boldsymbol{S}_i^T\boldsymbol{S}_i+\frac{\pi_0^2}{2}\tilde{\boldsymbol{D}}_i^T\tilde{\boldsymbol{D}}_i\\ \leq\ & -\lambda_{\min}(\boldsymbol{K}_2)\boldsymbol{S}_i^T\boldsymbol{S}_i+\Gamma_1^{-1}\kappa_{03}\mathrm{tr}\left(\tilde{\boldsymbol{W}}_{ia}^{T}\boldsymbol{\varphi}_{ia}(\boldsymbol{r}_{ia})\boldsymbol{\Psi}_i^T\right)\\ &+\kappa_{03}\boldsymbol{\Psi}_i^T\dot{\boldsymbol{\Psi}}_i+\tilde{\boldsymbol{D}}_i^T\dot{\tilde{\boldsymbol{D}}}_i\\ &+\frac{\bar{\varphi}_{ia}\bar{\varphi}_{ic}}{2}\mathrm{tr}\left(\tilde{\boldsymbol{W}}_{ia}^{T}\tilde{\boldsymbol{W}}_{ia}\right)+\frac{\bar{\varphi}_{ia}\bar{\varphi}_{ic}}{2}\mathrm{tr}\left(\tilde{\boldsymbol{W}}_{ic}^{T}\tilde{\boldsymbol{W}}_{ic}\right)\\ &-\frac{\kappa_{22}}{2}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ia}^{T}\tilde{\boldsymbol{W}}_{ia}\right]+\frac{\kappa_{22}\bar{w}_{ia}^2}{2}+\frac{\pi_0^2\tilde{\boldsymbol{D}}_i^T\tilde{\boldsymbol{D}}_i}{2}+\frac{\boldsymbol{S}_i^T\boldsymbol{S}_i}{2\pi_0^2}\\ &-\kappa_{12}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^{T}\tilde{\boldsymbol{W}}_{ic}\right]+\frac{\bar{\eta}_i^2}{2}+\frac{1}{2}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^{T}\tilde{\boldsymbol{W}}_{ic}\right].\end{aligned} \tag{52}$$

Consider the fact that

$$\begin{aligned}&\Gamma_1^{-1}\kappa_{03}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ia}^{T}\boldsymbol{\varphi}_{ia}(\boldsymbol{r}_{ia})\boldsymbol{\Psi}_i^T\right]+\kappa_{03}\boldsymbol{\Psi}_i^T\dot{\boldsymbol{\Psi}}_i+\tilde{\boldsymbol{D}}_i^T\dot{\tilde{\boldsymbol{D}}}_i\\ &=\Gamma_1^{-1}\kappa_{03}\boldsymbol{\Psi}_i^T\tilde{\boldsymbol{W}}_{ia}^{T}\boldsymbol{\varphi}_{ia}(\boldsymbol{r}_{ia})\\ &\quad+\kappa_{03}\boldsymbol{\Psi}_i^T\left[-\Gamma_1^{-1}\tilde{\boldsymbol{W}}_{ia}^{T}\boldsymbol{\varphi}_{ia}+\tilde{\boldsymbol{D}}_i-\kappa_{01}\boldsymbol{\Psi}_i\right]\\ &\quad+\tilde{\boldsymbol{D}}_i^T\left[\dot{\boldsymbol{D}}_i-\kappa_{02}\tilde{\boldsymbol{D}}_i-\kappa_{02}\Gamma_1\tilde{\boldsymbol{W}}_{ia}^{T}\boldsymbol{\varphi}_{ia}-\kappa_{03}\boldsymbol{\Psi}_i-\boldsymbol{S}_i\right]\\ &=\kappa_{03}\boldsymbol{\Psi}_i^T\tilde{\boldsymbol{D}}_i-\kappa_{01}\kappa_{03}\boldsymbol{\Psi}_i^T\boldsymbol{\Psi}_i+\tilde{\boldsymbol{D}}_i^T\dot{\boldsymbol{D}}_i-\kappa_{02}\tilde{\boldsymbol{D}}_i^T\tilde{\boldsymbol{D}}_i\\ &\quad-\kappa_{02}\tilde{\boldsymbol{D}}_i^T\Gamma_1\tilde{\boldsymbol{W}}_{ia}^{T}\boldsymbol{\varphi}_{ia}-\kappa_{03}\tilde{\boldsymbol{D}}_i\boldsymbol{\Psi}_i-\tilde{\boldsymbol{D}}_i^T\boldsymbol{S}_i.\end{aligned} \tag{53}$$

Then, (52) can be further obtained as follows:

$$\begin{aligned}\dot{L}_i \leq\ & -\lambda_{\min}(\boldsymbol{K}_2)\boldsymbol{S}_i^T\boldsymbol{S}_i-\kappa_{01}\kappa_{03}\boldsymbol{\Psi}_i^T\boldsymbol{\Psi}_i+\frac{\tilde{\boldsymbol{D}}_i^T\tilde{\boldsymbol{D}}_i}{2\pi_1^2}\\ &+\frac{\pi_1^2\dot{\boldsymbol{D}}_i^T\dot{\boldsymbol{D}}_i}{2}-\kappa_{02}\tilde{\boldsymbol{D}}_i^T\tilde{\boldsymbol{D}}_i+\frac{\kappa_{02}\Gamma_1\tilde{\boldsymbol{D}}_i^T\tilde{\boldsymbol{D}}_i}{2\pi_2^2}\\ &+\frac{\kappa_{02}\Gamma_1\pi_2^2\bar{\varphi}_{ia}^2\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ia}^{T}\tilde{\boldsymbol{W}}_{ia}\right]}{2}+\frac{\tilde{\boldsymbol{D}}_i^T\tilde{\boldsymbol{D}}_i}{2\pi_3^2}+\frac{\pi_3^2\boldsymbol{S}_i^T\boldsymbol{S}_i}{2}\\ &+\frac{\bar{\varphi}_{ia}\bar{\varphi}_{ic}}{2}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ia}^{T}\tilde{\boldsymbol{W}}_{ia}\right]+\frac{\bar{\varphi}_{ia}\bar{\varphi}_{ic}}{2}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^{T}\tilde{\boldsymbol{W}}_{ic}\right]\\ &-\frac{\kappa_{22}}{2}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ia}^{T}\tilde{\boldsymbol{W}}_{ia}\right]+\frac{\kappa_{22}\bar{w}_{ia}^2}{2}+\frac{1}{2\pi_0^2}\boldsymbol{S}_i^T\boldsymbol{S}_i+\frac{\pi_0^2}{2}\tilde{\boldsymbol{D}}_i^T\tilde{\boldsymbol{D}}_i\\ &-\kappa_{12}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^{T}\tilde{\boldsymbol{W}}_{ic}\right]+\frac{\bar{\eta}_i^2}{2}+\frac{1}{2}\mathrm{tr}\left[\tilde{\boldsymbol{W}}_{ic}^{T}\tilde{\boldsymbol{W}}_{ic}\right]\\ \leq\ & -m_{i1}L_i+m_{i2}\end{aligned} \tag{54}$$

where $m_{i1} = \min\{2\lambda_{\min}(\boldsymbol{K}_2)-1/\pi_0^2-\pi_3^2, 2\kappa_{01}, 2\kappa_{02}-1/\pi_1^2-\kappa_{02}\Gamma_1/\pi_2^2-1/\pi_3^2-\pi_0^2, \kappa_{21}(\kappa_{22}-\kappa_{02}\Gamma_1\pi_2^2\bar{\varphi}_{ia}^2-\bar{\varphi}_{ia}\bar{\varphi}_{ic}), \kappa_{11}(2\kappa_{12}-\bar{\varphi}_{ia}\bar{\varphi}_{ic}-1)\}$ and $m_{i2} = \pi_1^2\dot{\boldsymbol{D}}_i^T\dot{\boldsymbol{D}}_i/2+\kappa_{22}\bar{w}_{ia}^2/2+\bar{\eta}_i^2/2$. π_1, π_2, and π_3 are positive constants and only involved in the stability analysis.

According to the stability theorem, one can conclude that the errors $\boldsymbol{S}_i$, $\tilde{\boldsymbol{W}}_{ia}$, $\tilde{\boldsymbol{W}}_{ic}$, $\boldsymbol{\Psi}_i$, and $\tilde{\boldsymbol{D}}_i$ are uniformly ultimately bounded. Then, the transformed error $\boldsymbol{E}_i$ is bounded by following the similar analysis procedure as [55]. Considering the problem transformation from the user-specified finite-time convergence requirement of the distributed tracking error $\boldsymbol{e}_i$ to the uniformly ultimate boundedness of the error $\boldsymbol{E}_i$ via the PPF (15), one can finally conclude that the distributed tracking error e_{iv} is finite-time convergent and strictly confined within the user-specified bound $(-\underline{k}_v\xi_v, \bar{k}_v\xi_v)$. This ends the proof. ■

Remark 5: Different from the several existing works on reinforcement learning-based control design without explicit consideration of learning errors [35], [41], [42], this article provides a creative solution to handle the reinforcement learning errors by developing nonlinear DOs with auxiliary learning errors. Therefore, the unknown nonlinear functions associated with the faults and highly complex terms are effectively learned by the composite learning algorithm consisting of reinforcement learning mechanisms and nonlinear DOs.

Remark 6: With respect to the proposed FTFC method, finite-time PPF (15) is first involved to transform the distributed tracking errors. By using such a strategy, the distributed tracking errors e_{i1}, e_{i2}, and e_{i3} are strictly confined even when a portion of UAVs is encountered by the faults. Moreover, FO calculus is incorporated into the FTFC architecture for achieving refined adjustments of transient and steady-state performances by involving the FO sliding-mode error (24). Furthermore, actor-critic NN-based reinforcement learning strategy is developed to handle the fault-induced term $\boldsymbol{F}_i$ with auxiliary learning error (38) and nonlinear DO (40), thus significantly enhancing the reinforcement learning capability. The adaptive laws (42) and (43) are developed to update the critic and actor NNs, respectively. Actually, the reinforcement learning functionality can be well achieved if enough neural nodes are adopted. However, too many nodes may induce an enormous computational burden, which may limit the practical implementation of the reinforcement learning method. Fortunately, the involvement of the nonlinear DO with auxiliary learning error can increase the learning capability with less nodes.

Remark 7: In the developed method, the control parameters $\boldsymbol{K}_2$, κ_{11}, κ_{12}, κ_{21}, κ_{22}, κ_{01}, κ_{02}, and κ_{03} and FO operator a are mainly involved to adjust the FTFC performance. Especially, $\boldsymbol{K}_2$ is the feedback gain matrix and can be slightly increased to reduce the distributed tracking errors. κ_{11} and κ_{21} are used to regulate the learning speeds of adaptive laws (42) and (43), respectively. κ_{12} and κ_{22} are mainly used to act as the σ-modification technique for avoiding the parameter drift. κ_{01}, κ_{02}, and κ_{03} are designed to regulate the learning error compensation capability. The FO operator a can be adjusted to change the transient and steady-state performances. All these

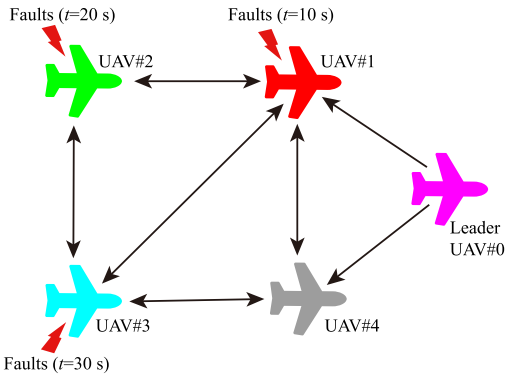


Fig. 4. Communication network.

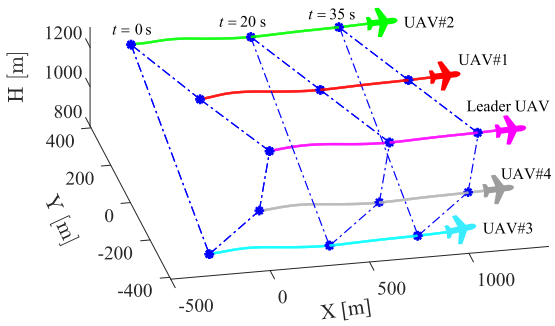


Fig. 5. Flight trajectories of all fixed-wing UAVs.

parameters should be chosen by trial and error until a good FTFC performance is obtained.

IV. Simulation Results

A. Description of the Simulation Scenarios

The proposed reinforcement learning-based FO FTFC scheme is demonstrated on a group of four follower fixed-wing UAVs and one leader UAV. The communication network is illustrated in Fig. 4, and the corresponding adjacency matrices are set as (55) and (56). The trajectory of leader UAV 0 is set as $\boldsymbol{q}_0(0) = [30t + 200,\ 10\sin(0.2t),\ 5\sin(0.3t) + 1000]^T$ m, and the initial positions of all follower UAVs are set as $\boldsymbol{q}_1(0) = [-2, 298, 998]^T$ m, $\boldsymbol{q}_2(0) = [-298, 402, 1203]^T$ m, $\boldsymbol{q}_3(0) = [-302, -402, 902]^T$ m, and $\boldsymbol{q}_4(0) = [0, -300, 1000]^T$ m. The initial states are set as $V_i(0) = 30$ m/s, $\chi_i(0) = 0°$, and $\gamma_i(0) = 0°$. The offset vectors of UAVs in Fig. 4 are given by $\boldsymbol{\delta}_1 = [-200, 300, 0]^T$ m, $\boldsymbol{\delta}_4 = [-200, -300, 0]^T$ m, $\boldsymbol{\delta}_{12} = [300, -100, -200]^T$ m, $\boldsymbol{\delta}_{13} = [300, 700, 100]^T$ m, $\boldsymbol{\delta}_{14} = [0, 600, 0]^T$ m, $\boldsymbol{\delta}_{21} = [-300, 100, 200]^T$ m, $\boldsymbol{\delta}_{23} = [0, 800, 300]^T$ m, $\boldsymbol{\delta}_{31} = [-300, -700, -100]^T$ m, $\boldsymbol{\delta}_{32} = [0, -800, -300]^T$ m, $\boldsymbol{\delta}_{34} = [-300, -100, -100]^T$ m, $\boldsymbol{\delta}_{41} = [0, -600, 0]^T$ m, and $\boldsymbol{\delta}_{43} = [300, 100, 100]^T$ m. In the simulation, the external disturbances $d_{i1} = 1.5$, $d_{i2} = 1.2$, and $d_{i3} = 2$ are suddenly suffered by all follower UAVs at 5 s. Moreover, to demonstrate the FTFC capability and ability for satisfying the user-specified transient and steady-state requirements, it is assumed that UAVs 1–3 are encountered by the velocity, heading angle, and flight path angle channel faults at 10, 20, and 30 s, respectively. During the simulation, UAV 4 is healthy, such that the developed reinforcement learning-based FO FTFC scheme can be demonstrated to be applicable to both faulty and healthy UAVs in the formation team. The fault signals are chosen as $\Lambda_{i1} = 0.3e^{-2(t-t_{if})} + 0.7$, $\Lambda_{i2} = 0.2e^{-2(t-t_{if})} + 0.8$, $\Lambda_{i3} = 0.25e^{-2(t-t_{if})} + 0.75$, $b_{if1} = 6e^{-2(t-t_{if})} - 6$, $b_{if2} = 2e^{-2(t-t_{if})} - 2$, $b_{if3} = 5e^{-2(t-t_{if})} - 5$, $i = 1$–3, $t_{1f} = 10$, $t_{2f} = 20$, and $t_{3f} = 30$

$$\mathcal{A} = \begin{bmatrix} 0 & 0.9 & 0.2 & 0.4 \\ 0.9 & 0 & 0.3 & 0 \\ 0.2 & 0.3 & 0 & 0.8 \\ 0.4 & 0 & 0.8 & 0 \end{bmatrix} \tag{55}$$

$$\mathcal{B} = \text{diag}\{0.8, 0, 0, 0.9\}. \tag{56}$$

In the simulation, the effectiveness of the developed FO FTFC method is first demonstrated in Scenario 1. Then, in Scenario 2, to verify the refined adjustment performance of FO calculus, the comparisons between the developed FO FTFC method with different FO operators and the IO FTFC method are conducted, in which the widely used IO sliding-mode surface $\boldsymbol{S}_i = \dot{\boldsymbol{E}}_i + \lambda_{11}\boldsymbol{E}_i$ is used to facilitate the comparative IO FTFC design. Moreover, consider the fact that finite-time PPFs are incorporated into the FTFC architecture to constrain the distributed tracking errors and the reinforcement learning units are developed to compensate for the faults, the developed FO FTFC method, the comparative FTFC method obtained by removing the PPFs from the developed control scheme (labeled as "No_PPF" in the simulation), and the comparative strategy constructed by removing the PPFs and reinforcement learning from the developed method (labeled as "No_PPF_No_Learning" in the simulation) are also compared in Scenario 2 to show the superiority of the developed FO FTFC method. Furthermore, different initial distributed tracking errors are adopted in Scenario 2 to test the capability of the finite-time PPF-based FO FTFC method for pulling large initial errors into the small region containing zero. To ensure the fairness of comparisons in Scenario 2, the same parameter values are adopted in the developed and comparative methods. The parameter values are set as $a = 0.3$, $\boldsymbol{K}_2 = \text{diag}\{17, 8.6, 15.8\}$, $\kappa_{01} = 2.8$, $\kappa_{02} = 1$, $\kappa_{03} = 15.25$, $\Gamma_1 = 16.36$, $\kappa_{11} = 33.3$, $\kappa_{12} = 0.8$, $\kappa_{21} = 81.4$, $\kappa_{22} = 1.4$, $\lambda_{11} = 0.23$, $\lambda_{12} = 0.6$, $\xi_{10} = 1.1$, $\xi_{20} = 1.1$, $\xi_{30} = 1.1$, $\xi_{1T} = 0.2$, $\xi_{2T} = 0.2$, $\xi_{3T} = 0.2$, $\bar{k}_v = \underline{k}_v = 10$, $\Lambda_{iv} = 0.4$, $v = 1$–3, $T = 0.6$, $C_{\zeta 1} = 0.3$, $C_{\zeta 2} = 0.3$, $C_{\zeta 3} = 0.3$, and $a_T = 10$.

B. Scenario 1

In this scenario, the effectiveness of the developed FO FTFC scheme is mainly verified. Fig. 5 shows the flight trajectories of all UAVs, and it can be seen that follower UAVs 1–4 can track the leader UAV with predesigned offsets even when the disturbances are injected at 5 s, and UAVs 1–3 are encountered by the faults at 10, 20, and 30 s, respectively. Fig. 6 illustrates the distributed tracking errors e_{i1}, e_{i2}, and e_{i3} of four follower UAVs. From the initial simulation phases, it can be easily seen that nonzero initial distributed tracking errors $e_{11}(0) = -6$ m, $e_{21}(0) = 4.8$ m, $e_{31}(0) = -2.8$ m, $e_{41}(0) = 2.4$ m, $e_{12}(0) = -6$ m, $e_{22}(0) = 4.8$ m, $e_{32}(0) = -2.8$ m, $e_{42}(0) = 2.4$ m,

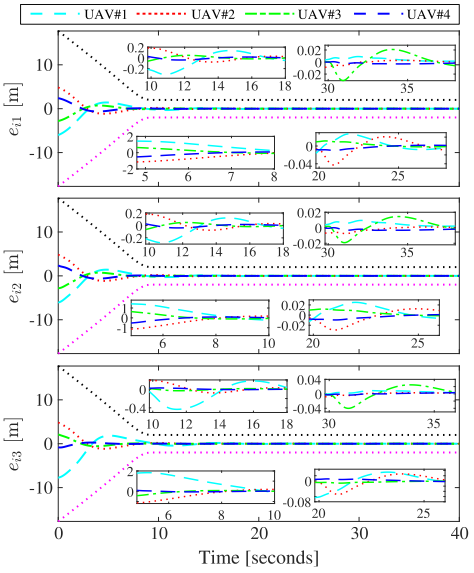


Fig. 6. Distributed tracking errors e_{ij}, $i = 1$–4 and $j = 1$–3.

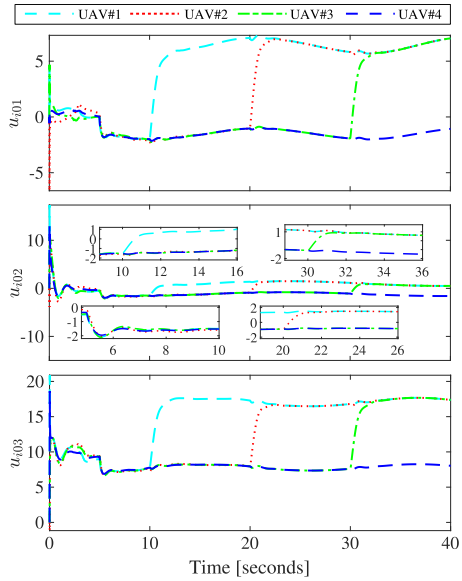


Fig. 8. Control inputs u_{i0j}, $i = 1$–4 and $j = 1$–3.

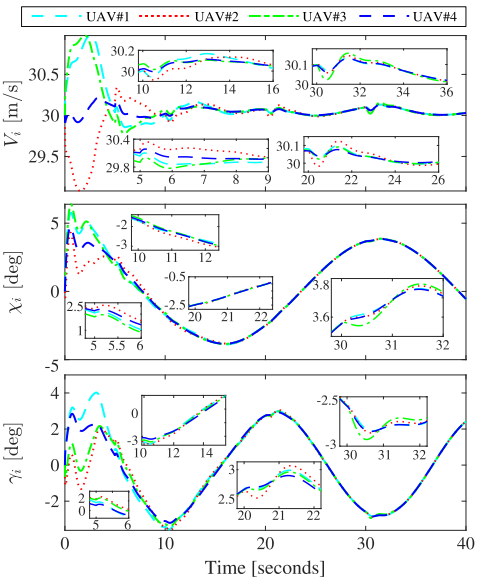


Fig. 7. Velocities, heading angles, and flight path angles of four follower UAVs.

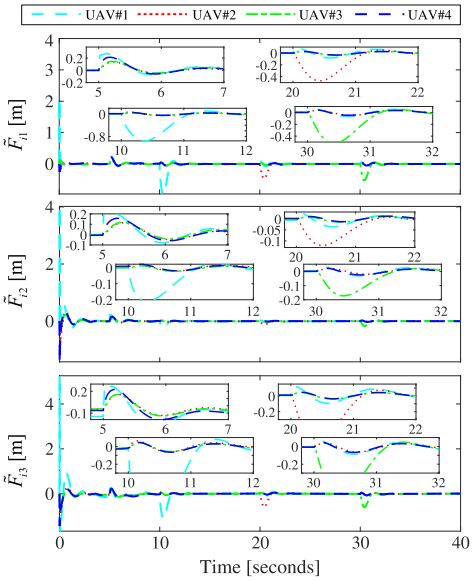


Fig. 9. Learning errors $\tilde{F}_{ij}$, $i = 1$–4 and $j = 1$–3.

$e_{13}(0) = -7.7$ m, $e_{23}(0) = 4.8$ m, $e_{33}(0) = 2.1$ m, and $e_{43}(0) = -0.8$ m are involved to demonstrate the feasibility of the finite-time PPF-based FO FTFC method. It can be further observed from Fig. 6 that slight performance degradation and variations are caused if the disturbances are injected into all follower UAVs at 5 s and follower UAVs 1–3 become faulty at 10, 20, and 30 s, respectively. However, under the developed reinforcement learning-based FO FTFC method, the variations are stabilized, and the errors are strictly confined within the user-specified ranges $(-\underline{k}_v\xi_v, \overline{k}_v\xi_v)$ after finite time 10 s.

Fig. 7 presents the velocities, heading angles, and flight path angles. Due to the nonzero initial distributed tracking errors shown in Fig. 6, relatively large variations of V_i, χ_i, and γ_i, $i = 1$–4, are activated at the beginning to attempt to reduce the distributed tracking errors. All states are bounded even if the follower UAVs are confronted with disturbances and faults.

The control inputs are shown in Fig. 8. To react to the disturbances at 5 s, control input signals u_{i01}, u_{i02}, and u_{i03} are expeditiously adjusted, ensuring the formation stability of networked UAVs. When follower UAV 1 becomes faulty at 10 s, u_{101}, u_{102}, and u_{103} are rapidly updated to stabilize faulty UAV 1 and the neighboring UAVs. Similar control actions are activated when UAV 2 and UAV 3 are subjected to faults at 20 and 30 s, respectively. Moreover, from the time responses of u_{i02} and u_{i03}, $i = 1$–4, it is observed that u_{402} and u_{403} are also slightly updated at 10 s, which is due to the fact that there exist information exchanges between UAV 1 and UAV 4. By using such a control response mode, the distributed tracking errors are convergent against disturbances and faults, and all follower UAVs can track the leader UAV with predesigned offsets.

Fig. 9 illustrates the learning errors $\tilde{F}_{ij}$ of reinforcement learning units and nonlinear DOs, $i = 1$–4 and $j = 1$–3.

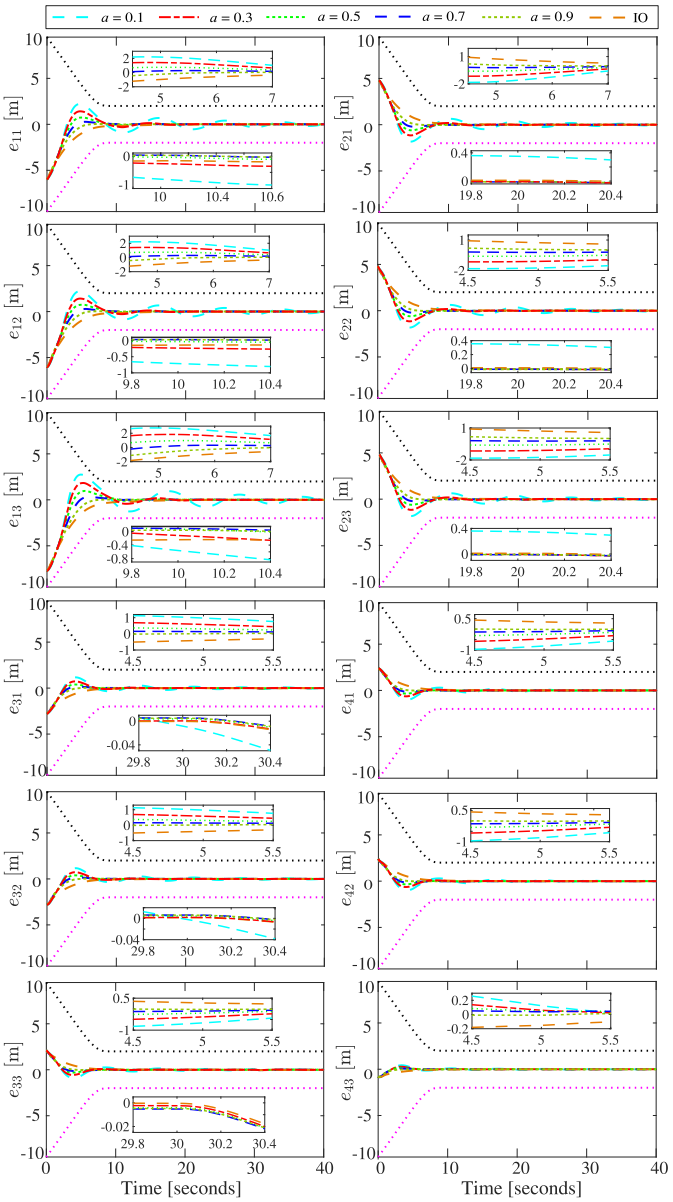


Fig. 10. Distributed tracking errors e_{ij} under the developed FO FTFC method with different FO operators and the IO FTFC method, $i = 1$–4 and $j = 1$–3.

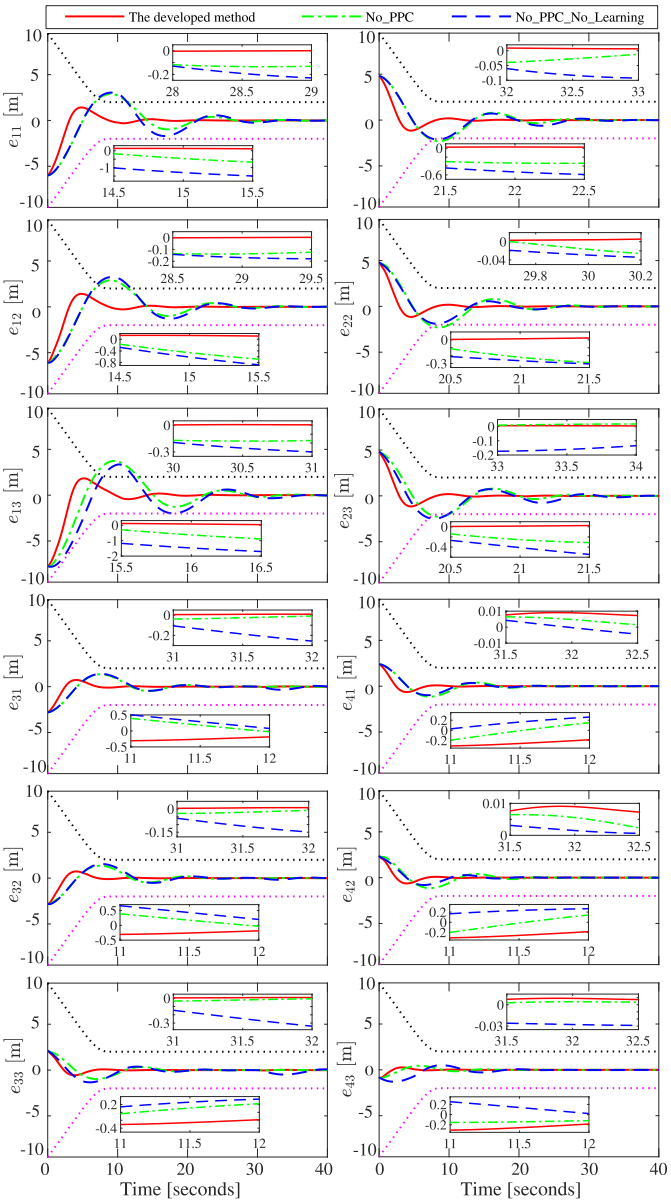


Fig. 11. Distributed tracking errors e_{ij} under the developed method and the comparative "No_PPF" and "No_PPF_No_Learning" methods, $i = 1$–4 and $j = 1$–3.

It can be easily seen that the learning errors are increased when the follower UAVs are encountered by the disturbances at 5 s and faults at 10, 20, and 30 s. Then, under the supervision of the developed modified reinforcement learning method with auxiliary learning error compensation technique, these perturbed errors are pulled into the very small region containing zero, thus ensuring successful fault compensations.

C. Scenario 2

In this scenario, comparative simulations are mainly conducted to show the superiority of the developed FO FTFC scheme with reinforcement learning mechanism and learning error compensation. Fig. 10 shows the distributed tracking errors under the developed FO FTFC method with different FO operators ($a = 0.1$, 0.3, 0.5, 0.7, and 0.9) and IO FTFC scheme. It can be found that if the FO operator is incorporated into the strategy, the transient and steady-state performances can be finely adjusted by changing the FO operator a, which significantly increases the flexibility of the FTFC method. Moreover, the comparative results among the developed FO FTFC method, the "No_PPF" scheme, and the "No_PPF_No_Learning" strategy are given in Fig. 11. It can be seen that the distributed tracking errors violate the prescribed bounds if the finite-time PPFs are removed from the developed method. Moreover, if the reinforcement learning units are further removed from the "No_PPF" method, it is observed that the enlarged distributed tracking error curves occur larger deviations from the small convergence region containing zero than the developed FO FTFC method and the "No_PPF" scheme.

Furthermore, to verify the capability of the proposed finite-time PPF-based FO FTFC method to pull large initial

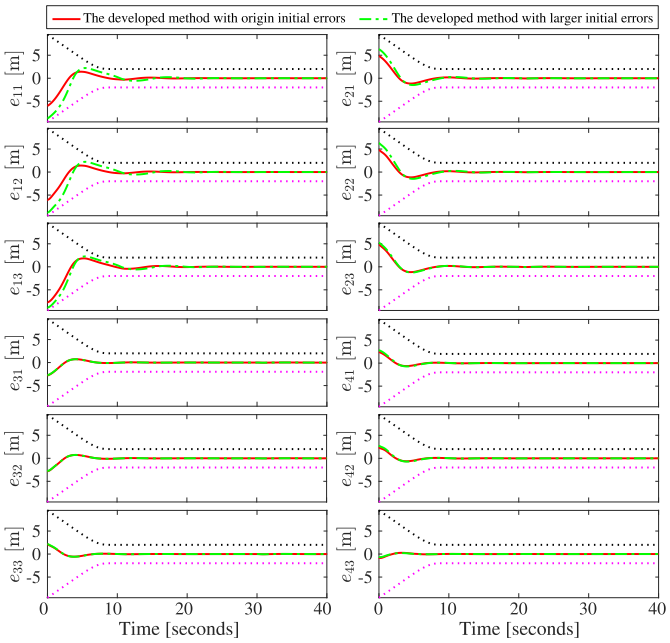


Fig. 12. Distributed tracking errors e_{ij} of the developed FO FTFC method with original and larger initial errors, $i = 1$–4 and $j = 1$–3.

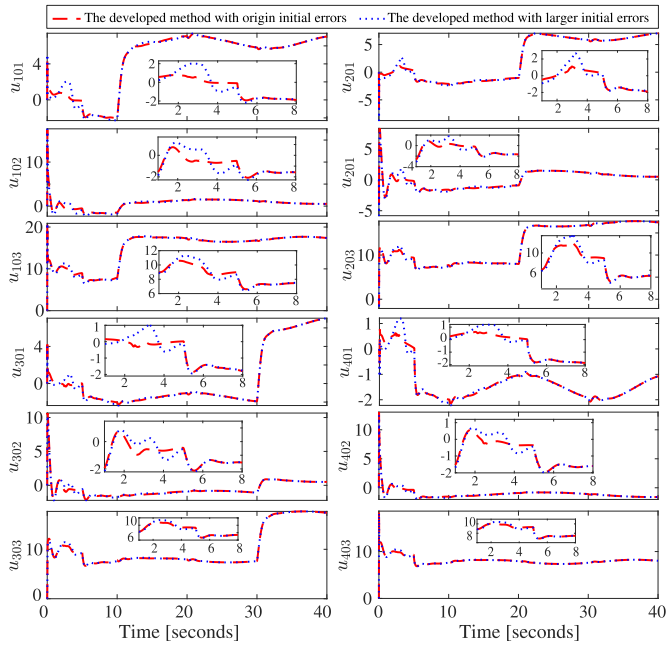


Fig. 13. Control inputs u_{i0j} under the developed FO FTFC method with original and larger initial errors, $i = 1$–4 and $j = 1$–3.

distributed tracking errors into the very small convergence region containing zero, the developed method with the original initial errors, i.e., the same initial errors as Scenario 1, and partial larger initial errors ($e_{11}(0) = -8.75$ m, $e_{21}(0) = 6.3$ m, $e_{31}(0) = -2.75$ m, $e_{41}(0) = 2.8$ m, $e_{12}(0) = -8.75$ m, $e_{22}(0) = 6.3$ m, $e_{32}(0) = -2.75$ m, $e_{42}(0) = 2.8$ m, $e_{13}(0) = -8.85$ m, $e_{23}(0) = 5.25$ m, $e_{33}(0) = 2.2$ m, and $e_{43}(0) = -0.6$ m) are considered. It can be seen from Fig. 12 that the developed finite-time PPF-based FO FTFC method can pull back the original and larger initial errors only if these initial errors are within the initial bounds. Moreover, the distributed errors are strictly confined within the overall prescribed bounds. However, it is clearly observed from Fig. 13 that if larger initial errors are involved, the control inputs u_{i0j} at the beginning phases will be perturbed and dynamically updated to stabilize the formation stability, $i = 1$–4 and $j = 1$–3, thus achieving reliable formation flight.

V. Conclusion and Future Works

This article has developed a reinforcement learning-based FO FTFC scheme for networked fixed-wing UAVs with user-specified performance constraints. The finite-time PPFs have been used to transform the distributed tracking errors of all follower UAVs, which are then filtered by FO sliding-mode surfaces. Based on the transformed FO error dynamics, the reinforcement learning with actor-critic NNs has been used to handle the approximation problems of unknown nonlinear terms and long-term performance indices. Nonlinear DOs and auxiliary learning errors have been skillfully constructed to compensate for the reinforcement learning errors. It has been shown by theoretical analysis and simulation results that all follower UAVs can track the leader UAV with predefined offsets, and the distributed tracking errors are strictly confined within the user-specified bounds.

With respect to the developed method, there is one disadvantage that the initial errors must fit within the specified initial bound $(-\underline{k}_v\xi_v(0), \bar{k}_v\xi_v(0))$, which may limit the implementation of the developed method. For very large initial errors, non-PPC strategy should be first adopted to steer the errors into the prescribed bounds, and the PPC method is still effective once these errors are pulled into the bounds. As one of future works, very large initial errors outside the user-specified initial bounds will be considered in the development of the reinforcement learning-based FO FTFC strategy for networked fixed-wing UAVs, leading to a piecewise control architecture.

References

[1] Y. Tian et al., "Search and rescue under the forest canopy using multiple UAVs," *Int. J. Robot. Res.*, vol. 39, nos. 10–11, pp. 1201–1221, Sep. 2020.

[2] C. Yuan, Y. M. Zhang, and Z. Liu, "A survey on technologies for automatic forest fire monitoring, detection, and fighting using unmanned aerial vehicles and remote sensing techniques," *Can. J. Forest Res.*, vol. 45, no. 7, pp. 783–792, Jul. 2015.

[3] P. Ladosz, H. Oh, and W.-H. Chen, "Trajectory planning for communication relay unmanned aerial vehicles in urban dynamic environments," *J. Intell. Robotic Syst.*, vol. 89, nos. 1–2, pp. 7–25, Jan. 2018.

[4] Z. Q. Yu, Y. M. Zhang, B. Jiang, and X. Yu, "Fault-tolerant time-varying elliptical formation control of multiple fixed-wing UAVs for cooperative forest fire monitoring," *J. Intell. Robotic Syst.*, vol. 101, no. 3, pp. 1–15, Mar. 2021.

[5] Z. Q. Yu, Y. M. Zhang, B. Jiang, J. Fu, and Y. Jin, "A review on fault-tolerant cooperative control of multiple unmanned aerial vehicles," *Chin. J. Aeronaut.*, vol. 35, no. 1, pp. 1–18, Jan. 2022.

[6] Z. Q. Yu, Y. H. Qu, and Y. M. Zhang, "Safe control of trailing UAV in close formation flight against actuator fault and wake vortex effect," *Aerosp. Sci. Technol.*, vol. 77, pp. 189–205, Jun. 2018.

[7] A. Chakravarty and C. Mahanta, "Actuator fault-tolerant control (FTC) design with post-fault transient improvement for application to aircraft control," *Int. J. Robust Nonlinear Control*, vol. 26, no. 10, pp. 2049–2074, Jul. 2016.

[8] Y. M. Zhang and J. Jiang, "Bibliographical review on reconfigurable fault-tolerant control systems," *Annu. Rev. Control*, vol. 32, no. 2, pp. 229–252, Dec. 2008.

[9] J. Jiang and X. Yu, "Fault-tolerant control systems: A comparative study between active and passive approaches," *Annu. Rev. Control*, vol. 36, no. 1, pp. 60–72, Apr. 2012.

[10] W. Zhao, H. Liu, and F. L. Lewis, "Data-driven fault-tolerant control for attitude synchronization of nonlinear quadrotors," *IEEE Trans. Autom. Control*, vol. 66, no. 11, pp. 5584–5591, Nov. 2021.

[11] T. Krüger, P. Schnetter, R. Placzek, and P. Vörsmann, "Fault-tolerant nonlinear adaptive flight control using sliding mode online learning," *Neural Netw.*, vol. 32, pp. 267–274, Aug. 2012.

[12] X. Zhao, Q. Zong, B. Tian, D. Wang, and M. You, "Finite-time fault-tolerant formation control for multiquadrotor systems with actuator fault," *Int. J. Robust Nonlinear Control*, vol. 28, no. 17, pp. 5386–5405, Nov. 2018.

[13] J. Shi, Y. Yang, J. Sun, X. He, D. Zhou, and Y. Zhong, "Fault-tolerant formation control of non-linear multi-vehicle systems with application to quadrotors," *IET Control Theory Appl.*, vol. 11, no. 17, pp. 3179–3190, Nov. 2017.

[14] Z. Q. Yu, Y. H. Qu, and Y. M. Zhang, "Distributed fault-tolerant cooperative control for multi-UAVs under actuator fault and input saturation," *IEEE Trans. Control Syst. Technol.*, vol. 27, no. 6, pp. 2417–2429, Nov. 2019.

[15] Z. Q. Yu, Z. X. Liu, Y. M. Zhang, Y. H. Qu, and C.-Y. Su, "Distributed finite-time fault-tolerant containment control for multiple unmanned aerial vehicles," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 6, pp. 2077–2091, Jun. 2020.

[16] Z. Q. Yu, Y. M. Zhang, B. Jiang, J. Fu, Y. Jin, and T. Y. Chai, "Composite adaptive disturbance observer-based decentralized fractional-order fault-tolerant control of networked UAVs," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 52, no. 2, pp. 799–813, Feb. 2022.

[17] Z. Yu et al., "Fractional-order adaptive fault-tolerant synchronization tracking control of networked fixed-wing UAVs against actuator-sensor faults via intelligent learning mechanism," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 12, pp. 5539–5553, Dec. 2021.

[18] Z. Shen, F. Li, X. Cao, and C. Guo, "Prescribed performance dynamic surface control for trajectory tracking of quadrotor UAV with uncertainties and input constraints," *Int. J. Control*, vol. 94, no. 11, pp. 2945–2955, Nov. 2021.

[19] Q. Guo, Y. Zhang, B. G. Celler, and S. W. Su, "Neural adaptive backstepping control of a robotic manipulator with prescribed performance constraint," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 12, pp. 3572–3583, Dec. 2019.

[20] N. Wang, Y. Gao, and X. Zhang, "Data-driven performance-prescribed reinforcement learning control of an unmanned surface vehicle," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 12, pp. 5456–5467, Dec. 2021.

[21] S. Dai, S. He, M. Wang, and C. Yuan, "Adaptive neural control of underactuated surface vessels with prescribed performance guarantees," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 12, pp. 3686–3698, Dec. 2019.

[22] O. Elhaki and K. Shojaei, "A robust neural network approximation-based prescribed performance output-feedback controller for autonomous underwater vehicles with actuators saturation," *Eng. Appl. Artif. Intell.*, vol. 88, Feb. 2020, Art. no. 103382.

[23] O. Elhaki and K. Shojaei, "Robust prescribed performance-based control of autonomous tractor-trailers convoy with limited communication range," *Int. J. Syst. Sci.*, vol. 52, no. 3, pp. 555–582, Feb. 2021.

[24] C. Wei, J. Luo, H. Dai, and G. Duan, "Learning-based adaptive attitude control of spacecraft formation with guaranteed prescribed performance," *IEEE Trans. Cybern.*, vol. 49, no. 11, pp. 4004–4016, Nov. 2019.

[25] Y. Nai, Q. Yang, and Z. Wu, "Prescribed performance adaptive neural compensation control for intermittent actuator faults by state and output feedback," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 11, pp. 4931–4945, Nov. 2021.

[26] M.-S. Qian, B. Jiang, and H. H.-T. Liu, "Dynamic surface active fault tolerant control design for the attitude control systems of UAV with actuator fault," *Int. J. Control, Autom. Syst.*, vol. 14, no. 3, pp. 723–732, Jun. 2016.

[27] M. Chen, H. Wang, and X. Liu, "Adaptive practical fixed-time tracking control with prescribed boundary constraints," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 68, no. 4, pp. 1716–1726, Apr. 2021.

[28] G. Cui, W. Yang, J. Yu, Z. Li, and C. Tao, "Fixed-time prescribed performance adaptive trajectory tracking control for a QUAV," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 69, no. 2, pp. 494–498, Feb. 2022.

[29] Y. Liu, H. Zhang, Y. Wang, H. Ren, and Q. Li, "Adaptive fuzzy prescribed finite-time tracking control for nonlinear system with unknown control directions," *IEEE Trans. Fuzzy Syst.*, vol. 30, no. 6, pp. 1993–2003, Jun. 2022.

[30] Z. Yu et al., "Enhanced recurrent fuzzy neural fault-tolerant synchronization tracking control of multiple unmanned airships via fractional calculus and fixed-time prescribed performance function," *IEEE Trans. Fuzzy Syst.*, vol. 30, no. 10, pp. 4515–4529, Oct. 2022.

[31] Y. Yu, J. Guo, C. K. Ahn, and Z. Xiang, "Neural adaptive distributed formation control of nonlinear multi-UAVs with unmodeled dynamics," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Mar. 16, 2022, doi: 10.1109/TNNLS.2022.3157079.

[32] F. Sedghi, M. M. Arefi, A. Abooee, and S. Yin, "Distributed adaptive-neural finite-time consensus control for stochastic nonlinear multiagent systems subject to saturated inputs," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Feb. 14, 2022, doi: 10.1109/TNNLS.2022.3145975.

[33] Y. Ouyang, C. Sun, and L. Dong, "Actor–critic learning based coordinated control for a dual-arm robot with prescribed performance and unknown backlash-like hysteresis," *ISA Trans.*, vol. 126, pp. 1–13, Jul. 2022.

[34] Y. Song, L. He, D. Zhang, J. Qian, and J. Fu, "Neuroadaptive fault-tolerant control of quadrotor UAVs: A more affordable solution," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 7, pp. 1975–1983, Jul. 2019.

[35] B. Xian, X. Zhang, H. Zhang, and X. Gu, "Robust adaptive control for a small unmanned helicopter using reinforcement learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 12, pp. 7589–7597, Dec. 2022.

[36] D. Nodland, H. Zargarzadeh, and S. Jagannathan, "Neural network-based optimal adaptive output feedback control of a helicopter UAV," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 7, pp. 1061–1073, Jul. 2013.

[37] X. Bu, "An improvement of single-network adaptive critic design for nonlinear systems with asymmetry constraints," *J. Franklin Inst.*, vol. 356, no. 16, pp. 9646–9664, Nov. 2019.

[38] Y. Ouyang, L. Dong, and C. Sun, "Critic learning-based control for robotic manipulators with prescribed constraints," *IEEE Trans. Cybern.*, vol. 52, no. 4, pp. 2274–2283, Apr. 2022.

[39] S. Bhasin, R. Kamalapurkar, M. Johnson, K. G. Vamvoudakis, F. L. Lewis, and W. E. Dixon, "A novel actor–critic–identifier architecture for approximate optimal control of uncertain nonlinear systems," *Automatica*, vol. 49, no. 1, pp. 82–92, Jan. 2013.

[40] O. Elhaki and K. Shojaei, "A novel model-free robust saturated reinforcement learning-based controller for quadrotors guaranteeing prescribed transient and steady state performance," *Aerosp. Sci. Technol.*, vol. 119, Dec. 2021, Art. no. 107128.

[41] L. Liu, Z. Wang, and H. Zhang, "Adaptive fault-tolerant tracking control for MIMO discrete-time systems via reinforcement learning algorithm with less learning parameters," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 1, pp. 299–313, Jan. 2017.

[42] T. Li, W. Bai, Q. Liu, Y. Long, and C. L. P. Chen, "Distributed fault-tolerant containment control protocols for the discrete-time multiagent systems via reinforcement learning method," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Nov. 1, 2021, doi: 10.1109/TNNLS.2021.3121403.

[43] X. Xiang, C. Liu, H. Su, and Q. Zhang, "On decentralized adaptive full-order sliding mode control of multiple UAVs," *ISA Trans.*, vol. 71, pp. 196–205, Nov. 2017.

[44] I. Podlubny, *Fractional Differential Equations*. New York, NY, USA: Academic, 1998.

[45] M. P. Aghababa, "A fractional sliding mode for finite-time control scheme with application to stabilization of electrostatic and electromechanical transducers," *Appl. Math. Model.*, vol. 39, no. 20, pp. 6103–6113, Oct. 2015.

[46] C. Li, D. Qian, and Y. Chen, "On Riemann–Liouville and Caputo derivatives," *Discrete Dyn. Nature Soc.*, vol. 2011, Mar. 2011, Art. no. 562494.

[47] R. M. Sanner and J.-J. E. Slotine, "Gaussian networks for direct adaptive control," *IEEE Trans. Neural Netw.*, vol. 3, no. 6, pp. 837–863, Nov. 1992.

[48] Z. H. Qu, "Matrix theory for cooperative systems," in *Cooperative Control of Dynamical Systems: Applications to Autonomous Vehicles*. Berlin, Germany: Springer, 2009, pp. 153–193.

[49] X. Bu, "Prescribed performance control approaches, applications and challenges: A comprehensive survey," *Asian J. Control*, vol. 25, no. 1, pp. 241–261, Jan. 2023.

[50] X. Wang, B. Niu, P. Zhao, and X. Song, "Neural networks-based adaptive finite-time prescribed performance fault-tolerant control of switched nonlinear systems," *Int. J. Adapt. Control Signal Process.*, vol. 35, no. 4, pp. 532–548, Apr. 2021.

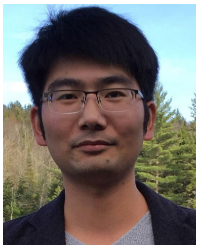
[51] H. Wang, W. Bai, X. Zhao, and P. X. Liu, "Finite-time-prescribed performance-based adaptive fuzzy control for strict-feedback nonlinear systems with dynamic uncertainty and actuator faults," *IEEE Trans. Cybern.*, vol. 52, no. 7, pp. 6959–6971, Jul. 2022.

[52] J. Tan, Y. Dong, P. Shao, and G. Qu, "Anti-saturation adaptive fault-tolerant control with fixed-time prescribed performance for UAV under AOA asymmetric constraint," *Aerosp. Sci. Technol.*, vol. 120, Jan. 2022, Art. no. 107264.

[53] A.-M. Zou, Z.-G. Hou, and M. Tan, "Adaptive control of a class of nonlinear pure-feedback systems using fuzzy backstepping approach," *IEEE Trans. Fuzzy Syst.*, vol. 16, no. 4, pp. 886–897, Aug. 2008.

[54] Y. Li and S. Tong, "Adaptive fuzzy output-feedback control of pure-feedback uncertain nonlinear systems with unknown dead zone," *IEEE Trans. Fuzzy Syst.*, vol. 22, no. 5, pp. 1341–1347, Oct. 2014.

[55] Y. Wang, L. Gu, Y. Xu, and X. Cao, "Practical tracking control of robot manipulators with continuous fractional-order nonsingular terminal sliding mode," *IEEE Trans. Ind. Electron.*, vol. 63, no. 10, pp. 6194–6204, Oct. 2016.



Ziquan Yu (Member, IEEE) received the Ph.D. degree in control science and engineering from Northwestern Polytechnical University, Xi'an, China, in 2019.

From 2017 to 2019, he was a Joint Ph.D. Student supported by the China Scholarship Council with the Department of Mechanical, Industrial and Aerospace Engineering, Concordia University, Montreal, QC, Canada. He is currently with the College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China. His current research interests include fault-tolerant cooperative control of safety-critical systems, and guidance, navigation, and control of unmanned aerial vehicles.

Jiaxu Li received the B.E. degree in automation from Northeast Forestry University, Harbin, China, in 2021. He is currently pursuing the M.S. degree in control science and engineering with the College of Automation, Nanjing University of Aeronautics and Astronautics, Nanjing, China.

His research interests include fault diagnosis and fault-tolerant control of unmanned aerial vehicles.

Yiwei Xu received the M.S. degree in control engineering from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2023.

Her research interests include fault-tolerant control of unmanned aerial vehicle, and data-based fault diagnosis and fault prediction.

Youmin Zhang (Fellow, IEEE) received the B.S., M.S., and Ph.D. degrees from Northwestern Polytechnical University, Xi'an, China, in 1983, 1986, and 1995, respectively, all in automatic control.

He is currently a Professor with the Department of Mechanical, Industrial and Aerospace Engineering and the Concordia Institute of Aerospace Design and Innovation, Concordia University, Montreal, QC, Canada. He has authored eight books and more than 600 journal and conference papers. His current research interests include guidance, navigation, and control (GNC); fault detection and diagnosis (FDD); fault-tolerant control (FTC); and remote sensing with applications to unmanned aerial/space/ground/marine vehicles, smart grids, smart cities, and cyber-physical systems (CPS).

Dr. Zhang is a fellow of Canadian Society for Mechanical Engineering (CSME), a Senior Member of American Institute of Aeronautics and Astronautics (AIAA), and a member of the Technical Committee for several scientific societies. He has served as the honorary general chair, the general chair, the program chair, and an International Program Committee (IPC) member of several unmanned systems and renewable energies relevant international conferences. He was the President of the International Society of Intelligent Unmanned Systems (ISIUS) from 2019 to 2022. He is a registered Professional Engineer in ON, Canada. He is an Editorial Board Member, a Deputy Editor-in-Chief, and an Editor/Associate Editor of several international journals, including the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, the IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—II, the *Journal of Intelligent and Robotic Systems*, *Unmanned Systems*, *Guidance, Navigation and Control*, and *Security and Safety*.

Bin Jiang (Fellow, IEEE) received the Ph.D. degree in automatic control from Northeastern University, Shenyang, China, in 1995.

He has been a post-doctoral fellow or a research fellow in Singapore, France, and USA, and a visiting professor in Canada. He is currently the Chair Professor of the Cheung Kong Scholar Program, Ministry of Education, China, and the Vice President of the Nanjing University of Aeronautics and Astronautics, Nanjing, China. His current research interests include fault diagnosis and fault-tolerant control and their applications.

Dr. Jiang is a member of the International Federation of Automatic Control Technical Committee on Fault Detection, Supervision, and Safety of Technical Processes. He is the Chair of the Control Systems Chapter of the IEEE Nanjing Section. He serves as a Senior Editor, an Editor, an Associate Editor, or an Editorial Board Member for the IEEE TRANSACTIONS ON CYBERNETICS, the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, the *International Journal of Control, Automation, and Systems*, *Neurocomputing*, *Acta Automatica Sinica*, the *Journal of Astronautics*, and *Guidance, Navigation and Control*.

Chun-Yi Su (Senior Member, IEEE) received the Ph.D. degree in control engineering from the South China University of Technology, Guangzhou, China, in 1990.

He was with the University of Victoria, Victoria, BC, Canada. In 1998, he joined Concordia University, Montreal, QC, Canada. He has authored or coauthored more than 400 publications in journals, book chapters, and conference proceedings. His current research interests include the application of automatic control theory to mechanical systems, especially the control of systems involving hysteresis nonlinearities.