Robust Adaptive Neural Network Tracking Control for Quadrotor Unmanned Aerial Vehicle via Reinforcement Learning Strategy

Baisong Wang[1,*], Kai Wu[2], Yang Wang[1], Zhugang Chu[1] and Kebiao Liu[1]

[1] State Grid Anqing Power Supply Company, Anqing, 246000, China

[2] Electric Power Research Institute of State Grid Anhui Electric Power Co., Ltd., Hefei, 230041, China

INFORMATION

Keywords:
Tracking control
reinforcement learning
nonlinear neural network
disturbance observer
event-triggered controller
quadrotor unmanned aerial vehicle

DOI: 10.23967/j.rimni.2026.10.72590

SCIPEDIA Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería

Revista Internacional Métodos numéricos para cálculo y diseño en ingeniería RIMNI

In cooperation with CIMNE

Robust Adaptive Neural Network Tracking Control for Quadrotor Unmanned Aerial Vehicle via Reinforcement Learning Strategy

Baisong Wang[1,*], **Kai Wu**[2], **Yang Wang**[1], **Zhugang Chu**[1] **and Kebiao Liu**[1]

[1]State Grid Anqing Power Supply Company, Anqing, 246000, China

[2]Electric Power Research Institute of State Grid Anhui Electric Power Co., Ltd., Hefei, 230041, China

ABSTRACT

This paper proposes an adaptive event-triggered trajectory and attitude tracking control framework for quadrotor unmanned aerial vehicle (QUAV) with external disturbances. To handle unknown uncertainties in QUAV control system, we propose a dual neural network (NN) architecture: combining reinforcement learning with disturbance estimation for real-time disturbance compensation. Specifically, the Actor-NN generates compensation signals to offset uncertainties, while the Critic-NN dynamically evaluates control performance to adjust the learning process. A nonlinear neural network disturbance observer (NNDO) is incorporated to estimate the lumped total disturbances in real time. By combining backstepping control approach with event-triggered mechanism, the proposed control strategy achieves rigorous closed-loop stability with guaranteed exclusion of Zeno behavior. Experimental validation on QUAV demonstrates the effectiveness of the proposed scheme in balancing computational efficiency and control performance.

OPEN ACCESS

Received: 30/08/2025

Accepted: 12/12/2025

DOI
10.23967/j.rimni.2026.10.72590

Keywords:
Tracking control
reinforcement learning
nonlinear neural network
disturbance observer
event-triggered controller
quadrotor unmanned aerial vehicle

1 Introduction

In recent years, quadrotor unmanned aerial vehicles (QUAVs) have witnessed rapid development and found extensive applications in various fields, including aerial photography, environmental monitoring, disaster relief, and delivery services. Their distinct advantages, including vertical takeoff and landing capability, high maneuverability, and facile deployability, render them exceptionally suitable for complex and dynamically evolving operational tasks [1–8]. However, the high-precision trajectory and attitude tracking control of QUAVs under external disturbances remains a challenging problem.

Intelligent control techniques have emerged as promising alternatives for addressing complex control challenges in various engineering systems [9–15]. Among these techniques, neural networks exhibit excellent approximation capabilities for complex nonlinear functions, making them a key focus in advanced control research. A neural network-based model predictive control (NNMPC)

*Correspondence: Baisong Wang (wangbs5335@ah.sgcc.com.cn).

optimization strategy is presented in [10], which centers on integrating rich historical operational data such as multi-flight attitude records and control responses under diverse wind conditions. In [12], an adaptive neural network control approach was proposed to approximate partially unknown system dynamics, but it relies on precise parameter tuning for QUAVs with strong coupling. The literature [13] developed a neuroadaptive fault-tolerant scheme, yet it does not address external disturbances like wind gusts. In [16], although neural networks were used as compensators to reduce model errors, but its computational complexity limits real-time application on QUAVs. To solve these issues, this paper integrates reinforcement learning with disturbance estimation to balance control precision and efficiency.

Due to the inherent nonlinearity, strong coupling, and parameter uncertainties of QUAV, these methods may face difficulties in achieving satisfactory performance. The time-consuming parameter tuning always be required. To learn the mapping relationship between system states and control inputs from data, a flexible and effective approach was designed to enhance the control performance in [17–19]. In [20], a learning-based robust tracking control scheme was developed for QUAV with time-varying and coupling uncertainties. By designing position and attitude tracking error subsystems and applying neural networks to approximate optimal control laws, the tracking errors of the system were convergent and a solid theoretical foundation for verifying the effectiveness and reliability of the proposed control scheme was provided. In parallel with neural network-based control methods, reinforcement learning (RL) has recently garnered growing attention. Owing to its unique advantage of enabling agents to learn optimal control policies through interaction with the environment (without relying on prior accurate system models). RL is a powerful data-driven learning paradigm that enables agents to learn optimal control policies through trial-and-error interactions with the environment.

However, in order to perform tasks with minimal communication and computational overhead, the event-triggered control strategies have been increasingly applied to uncertain nonlinear systems for enhanced performance and efficiency [21–29]. In [21], the robust optimal attitude control problem was studied for QUAV via dynamic event-triggered mechanisms. By constructing an augmented system, the tracking problem is transformed into a stabilization problem. A dynamic event-triggering mechanism is designed to reduce communication load, and an event-driven critic-only neural network reinforcement learning algorithm is developed to achieve optimal control. RL has been explored in the context of cellular Internet of UAVs for protocol design, trajectory control, and resource management. This research highlights the potential of reinforcement learning to enhance the quality of service in dynamic and uncertain environments in [22]. An adaptive attitude control method for QUAV that combines event-triggered control and optimized hybrid quantizers. By incorporating prescribed-time predefined performance functions, the convergence of attitude tracking errors is significantly enhanced. The event-triggering mechanism and optimized quantization reduce communication burdens in [23]. In [24], the event-triggered adaptive control problem for uncertain nonlinear systems was proposed. A co-design framework simultaneously develops the adaptive controller and triggering events, circumventing the traditional requirement for input-to-state stability of measurement errors. A novel switching threshold strategy balances system performance and network constraints. An event-triggered deep learning control strategy is proposed for real-time quadrotor trajectory tracking in [29].

External disturbances, such as wind gusts, air turbulence, and electromagnetic interference, can significantly affect the stability and control performance of QUAV [30–33]. The presence of external detrimental perturbations inevitably compromises the operational stability and precision of control systems, necessitating advanced countermeasures to enhance disturbance rejection capabilities. Contemporary research has progressively shifted toward observer-based compensation methodologies, where researchers have developed integrated control architectures incorporating disturbance observers

(DOs) to neutralize the detrimental impacts of perturbations. Notable contributions include comprehensive reviews analyzing hierarchical DO frameworks in [34], and neural network-enhanced DO implementations for multi-source uncertainty estimation [35–37]. In [38], a novel online self-learning disturbance observer is proposed for nonlinear systems by integrating feedback error learning, type-2 neuro-fuzzy structures, and sliding mode control theory. The observer leverages type-2 neuro-fuzzy systems to learn system uncertainties and employs sliding mode principles to design adaptive learning algorithms, enabling precise estimation of time-varying disturbances and robust system performance. In [39], an adaptive event-triggered constrained control strategy is proposed for uncertain nonlinear systems with input constraints, by utilizing self-learning disturbance observer and actor-critic neural network to compensate the unknown uncertainties and external disturbances.

Based on the above discussion and analysis, this paper mainly focuses on the trajectory and attitude tracking control problem of QUAV with external disturbances. The main contributions are listed as follows:

(I) Reinforcement learning (RL) is adopted to improve the approximation accuracy of the neural network (NN) for effective estimation of the system uncertainties, and a self-learning disturbance observer (SLDO) is designed to suppress external disturbances and NN approximation errors.

(II) The serial-parallel estimation methodology (SPEM) is integrated into the QUAV control framework to enhance the adaptive observation and approximation of complex unmodeled system dynamics and composite disturbances, laying a foundation for accurate uncertainty compensation.

(III) An improved event-triggered control (ETC) technique is presented to conserve communication and control resources of QUAV, while ensuring that all signals in the entire closed-loop system are ultimately uniformly bounded.

The remainder of this paper is organized as follows. Section 2 presents the mathematical model of the quadrotor UAV and formulates the trajectory and attitude tracking control problem. Section 3 details the proposed reinforcement learning-based control method. Section 4 reports the simulation and experimental results, followed by a discussion of the findings. Finally, Section 5 concludes the paper and outlines directions for future research.

2 Problem Statement

2.1 Control Objective

The simplified dynamic model of QUAV is described as [40]

$$\begin{cases} \ddot{x} = \dfrac{u_l}{m}(\cos\phi\sin\theta\cos\psi + \sin\phi\sin\psi) - \dfrac{a_x}{m}\dot{x} + d_x \\ \ddot{y} = \dfrac{u_l}{m}(\cos\phi\sin\theta\sin\psi - \sin\phi\cos\psi) - \dfrac{a_y}{m}\dot{y} + d_y \\ \ddot{z} = \dfrac{u_l}{m}\cos\phi\cos\theta - g - \dfrac{a_z}{m}\dot{z} + d_z \\ \ddot{\phi} = \dfrac{L}{I_x}u_\phi + \dot{\theta}\dot{\psi}\dfrac{I_y - I_z}{I_x} - \dfrac{a_\phi L}{I_x}\dot{\phi} + d_\phi \\ \ddot{\theta} = \dfrac{L}{I_y}u_\theta + \dot{\phi}\dot{\psi}\dfrac{I_z - I_x}{I_y} - \dfrac{a_\theta L}{I_y}\dot{\theta} + d_\theta \\ \ddot{\psi} = \dfrac{L}{I_z}u_\psi + \dot{\phi}\dot{\theta}\dfrac{I_x - I_y}{I_z} - \dfrac{a_\psi L}{I_z}\dot{\psi} + d_\psi \end{cases} \tag{1}$$

where x, y, z denote three position component, ϕ, θ, ψ represent the roll, pitch and yaw angles, respectively. m is the total mass, L stands for the length from the center of mass to the rotor center, g represents the gravitational acceleration, and I_x, I_y, I_z are the rotary inertia. For $i = x, y, z, \phi, \theta, \psi$, a_i denote the drag coefficients, d_i stand for the external disturbances which satisfy $|d_i| \le \bar{d}_i$.u_l is the total lift force, u_ϕ, u_θ, u_ψ are the control inputs of roll, pitch and yaw moments.

For simplicity's sake, the dynamic model (1) can be rewritten as

$$\ddot{\mathcal{X}}_i = \mathcal{G}_i\mathcal{U}_i + \mathcal{F}_i + \mathcal{D}_i, \quad i = 1, 2, \ldots, 6 \tag{2}$$

where $(\mathcal{X}_1, \mathcal{X}_2, \mathcal{X}_3, \mathcal{X}_4, \mathcal{X}_5, \mathcal{X}_6) = (\phi, \theta, \psi, z, x, y)$, $(\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3, \mathcal{G}_4, \mathcal{G}_5, \mathcal{G}_6) = (L/I_x, L/I_y, L/I_z, 1/m, 1/m, 1/m)$, $(u_1, u_2, u_3, u_4, u_5, u_6) = (u_\phi, u_\theta, u_\psi, u_l\cos\phi\cos\theta - Mg, u_l(\cos\phi\sin\theta\cos\psi + \sin\phi\sin\psi), u_l(\cos\phi\sin\theta\sin\psi - \sin\phi\cos\psi))$, $(\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3, \mathcal{F}_4, \mathcal{F}_5, \mathcal{F}_6) = (\dot{\theta}\dot{\psi}(I_y - I_z)/I_x - a_\phi L\dot{\phi}/I_x, \dot{\phi}\dot{\psi}(I_z - I_x)/I_y - a_\theta L\dot{\theta}/I_y, \dot{\phi}\dot{\theta}(I_x - I_y)/I_z - a_\psi L\dot{\psi}/I_z, -a_z\dot{z}/m, -a_x\dot{x}/m, -a_y\dot{y}/m)$, $(\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4, \mathcal{D}_5, \mathcal{D}_6) = (d_\phi, d_\theta, d_\psi, d_z, d_x, d_y)$.

Denoting $\mathcal{X}_{i1} = \mathcal{X}_i$ and $\mathcal{X}_{i2} = \dot{\mathcal{X}}_i$, the system (2) can be rewritten as:

$$\begin{cases} \dot{\mathcal{X}}_{i1} = \mathcal{X}_{i2} \\ \dot{\mathcal{X}}_{i2} = \mathcal{G}_i\mathcal{U}_i + \mathcal{F}_i + \mathcal{D}_i, \end{cases} \quad i = 1, 2, \ldots, 6 \tag{3}$$

The main control objective of this paper is to develop an event-triggered adaptive trajectory and attitude tracking controller based on reinforcement learning strategy and nonlinear disturbance observer for the QUAV, which enables all the closed-loop signals to be uniformly ultimately bounded. The tracking control diagram is shown in Fig. 1. Meanwhile, the trajectory and attitude signals of the QUAV can track the ideal trajectories r_x, r_y, r_z and the desired reference attitude signal r_ψ, r_ϕ, r_θ. Note that in many cases, according to the actual physical characteristics of the QUAV, it is usually known that the expected signals of roll and pitch angles are calculated by $\mathcal{U}_4$, $\mathcal{U}_5$, $\mathcal{U}_6$, and ψ, which are expressed in the following form:

$$\begin{cases} r_\phi = \arcsin\left(\dfrac{\mathcal{U}_5\sin\psi - \mathcal{U}_6\cos\psi}{\sqrt{(\mathcal{U}_4 + mg)^2 + \mathcal{U}_5^2 + \mathcal{U}_6^2}}\right), \\ r_\theta = \arctan\left(\dfrac{\mathcal{U}_5\cos\psi + \mathcal{U}_6\sin\psi}{\mathcal{U}_4 + mg}\right) \end{cases} \tag{4}$$

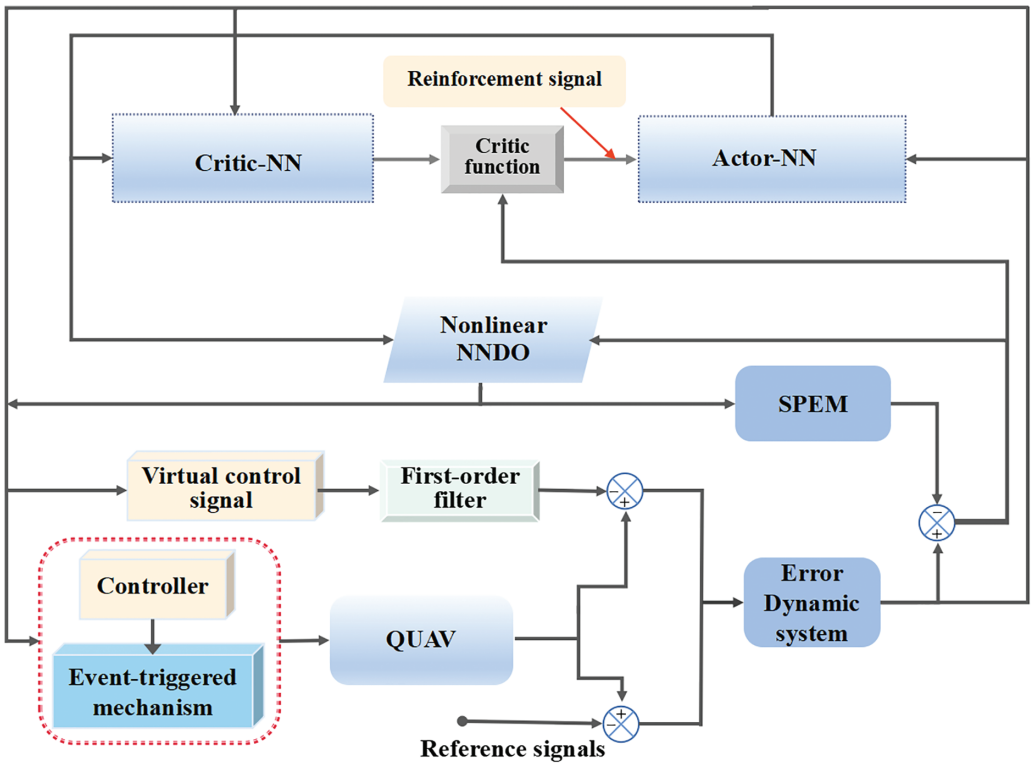


Figure 1: The tracking control diagram

Assumption 1: *The desired reference signals* r_j, $(j = x, y, z, \phi, \theta, \psi)$ *are bounded and differentiable, and their respective derivatives are also bounded.*

Assumption 2: *In view of the basic physical characteristics and feasible flight conditions, it is generally assumed that both the external disturbances* $\mathcal{D}_i, (i = 1, 2, \ldots, 6)$ *exist in each channel and their first-order derivatives* $\dot{\mathcal{D}}_i$ *are bounded.*

2.2 Preliminary

Some corresponding basic theoretical basis and assumptions need to be given in advance, and will be listed as follows:

Lemma 1 [39]: *For* $\forall x \in R^n$ *and* $\varsigma \in R$, *the following inequality holds:*

$$\varsigma\|x\| - x^{\mathrm{T}}\varsigma \tanh\left(\frac{\sigma x}{\varkappa}\right) \leq n\bar{\varkappa} \tag{5}$$

where $\bar{\varkappa} = 0.2785\varkappa$ *with* $\varkappa > 0$, *and* $\tanh\left(\frac{\sigma x}{\varkappa}\right)$ *denotes a vector.*

Lemma 2 [39]: *For* $\forall x, y \in R$, *there must exist positive constants* $\boldsymbol{a}, \boldsymbol{b}$ *and* $\boldsymbol{p}$ *satisfying* $\boldsymbol{a} > \mathbf{1}$, $\boldsymbol{b} > \mathbf{1}$, $\boldsymbol{p} > \mathbf{0}$ *and* $(\boldsymbol{a} - \mathbf{1})(\boldsymbol{b} - \mathbf{1}) = 1$, *such that*

$$xy \leq \frac{\boldsymbol{p}^{\boldsymbol{a}}}{\boldsymbol{a}}|x|^{\boldsymbol{a}} + \frac{1}{\boldsymbol{b}\boldsymbol{p}^{\boldsymbol{b}}}|y|^{\boldsymbol{b}} \tag{6}$$

A relative-threshold-based control strategy is defined as follows [41]:

$$\begin{cases} \bar{v}_i(t) = -(1+\rho_i)\left(v_{i2}\tanh\left(\dfrac{\mathcal{E}_{i2}v_{i2}}{\tau_i}\right) + \gamma_{i1}\tanh\left(\frac{\varepsilon_{i2}\gamma_{i1}}{\tau_i}\right)\right) \\ \mathcal{U}_i(t) = \bar{v}_i(t_k), \quad \forall t \in [t_k, t_{k+1}) \\ t_{k+1} = \inf\{t \in R || \varepsilon_i(t)| \geq \rho_i|\mathcal{U}_i(t)| + \gamma_{i2}\} \end{cases} \tag{7}$$

where v_{i2} denote the virtual control signals which will be designed later. $\varepsilon_i(t) = \bar{v}_i(t) - \mathcal{U}_i(t).\tau_i, \gamma_{i1}, \gamma_{i2}$, are positive design parameters, and $\gamma_{i2} > 0, 0 < \rho_i < 1, \gamma_{i1} > \dfrac{\gamma_{i2}}{1-\rho_i}, k \in Z^+$, are controller updating time instants, i.e., whenever (7) is triggered, the instants will be marked as t_{k+1} and control values $\mathcal{U}_i(t_{k+1})$ will be applied to the system (1). During the time interval $[t_k, t_{k+1})$, the control signal is a constant, i.e., $\bar{v}_i(t_k)$. In addition, $\forall \tau_i > 0$, the hyperbolic tangent function satisfies

$$0 \leq |\mathcal{E}_{i2}v_{i2}| - \mathcal{E}_{i2}v_{i2}\tanh\left(\frac{\mathcal{E}_{i2}v_{i2}}{\tau_i}\right) \leq \tau_i\varrho$$

$$0 \leq |\mathcal{E}_{i2}\gamma_{i1}| - \mathcal{E}_{i2}\gamma_{i1}\tanh\left(\frac{\mathcal{E}_{i2}\gamma_{i1}}{\tau_i}\right) \leq \tau_i\varrho$$

where ϱ is a constant, and $\varrho = \mathrm{e}^{-(\varrho+1)}$ (i.e., $\varrho \approx 0.2785$).

3 Tracking Controller Design via Reinforcement Learning

3.1 Error System Modeling Based on Reinforcement Learning

On the basis of the basic process of the classic backstepping control method, the tracking errors are defined as:

$$\mathcal{E}_{i1} = \mathcal{X}_{i1} - r_i,\ \mathcal{E}_{i2} = \mathcal{X}_{i2} - v_{ci}, i = 1, 2, \ldots, 6 \tag{8}$$

where v_{ci} denote the output of the following first-order filter:

$$\sigma_i\dot{v}_{ci} = -v_{ci} + v_i,\ v_{ci}(0) = v_i(0) \tag{9}$$

where $\sigma_i > 0$ is a positive constant, v_i is considered as the input of the filter, which is the virtual control signal to be designed. Moreover, defining the error variable $\varsigma_i = v_{ci} - v_i$.

Taking time derivative of (4), and combining with (1) and (5), the dynamic system error can be described as follows:

$$\begin{cases} \dot{\mathcal{E}}_{i1} = \mathcal{E}_{i2} + \varsigma_i + v_i - \dot{r}_i \\ \dot{\mathcal{E}}_{i2} = \mathcal{G}_i\mathcal{U}_i + \bar{\mathcal{F}}_i + \mathcal{D}_i, \end{cases} \tag{10}$$

where $\bar{\mathcal{F}}_i = \mathcal{F}_i + \dfrac{1}{\sigma_i}(v_{ci} - v_i)$ is the lumped uncertain term, which can be approximated by the NNs via reinforcement learning strategy.

As a consequence, the term $\bar{\mathcal{F}}_i$ is unknown but continuous nonlinear function which can be approximated by the NN, that is

$$\bar{\mathcal{F}}_i(Z_i) = \mathcal{W}_i^{*T}\Xi_i(Z_i) + \varepsilon_i^*(Z_i) \tag{11}$$

where $Z_i = \left[\bar{\chi}_i^T, \zeta_i^T, \dot{\zeta}_i^T\right]^T$. $\mathcal{W}_i^*$ and $\Xi_i(Z_i)$ denote the optimal weight value vector and the basis function vector, respectively. $\varepsilon_i^*(Z_i)$ is the approximation error. A brief description of neural network approximation method can be found in Remark 1.

Substituting (11) into (10), we have

$$\begin{cases} \dot{\mathcal{E}}_{i1} = \mathcal{E}_{i2} + \varsigma_i + \nu_i - \dot{r}_i \\ \dot{\mathcal{E}}_{i2} = \mathcal{G}_i \mathcal{U}_i + \mathcal{W}_i^{*T} \Xi_i(Z_i) + \bar{\mathcal{D}}_i, \end{cases} \tag{12}$$

where $\bar{\mathcal{D}}_i = \mathcal{D}_i + \varepsilon_i^*(Z_i)$ denotes the total disturbance, and it is obviously assumed that $\bar{\mathcal{D}}_i$ is differentiable and bounded. Remark 2 provides specific explanations about the composite disturbance.

Remark 1: *Over recent decades, intelligent control methodologies utilizing neural networks (NNs) and fuzzy logic systems have emerged as prominent approaches in the field. Notably, NNs are distinguished by their exceptional function approximation capabilities, a characteristic that has led to their widespread adoption in managing undefined dynamics within complex systems. As theoretically established, any arbitrary continuous function $f(x)$, $x \in S$ can be systematically approximated through neural architectures, with a representative mathematical formulation expressed as [29]:*

$$f_{continuous}(x) = \mathcal{W}^T \Xi(x) + \varepsilon(x) \tag{13}$$

where $\mathcal{W}$ is the weight vector, $\Xi(x)$ is the basis vector function, and $\varepsilon(x)$ is the error. According to the theoretical properties of the NN approximation, there must exist an optimal weight vector $\mathcal{W}^*$ with the minimum approximation error $\varepsilon^*(x)$, that is,

$$\begin{cases} f_{continuous}(x) = \mathcal{W}^{*T} \Xi(x) + \varepsilon^*(x) \\ \mathcal{W}^* := \arg\min\limits_{\mathcal{W}} \left\{ \sup\limits_{x \in S} \left\| f_{continuous}(x) - \mathcal{W}^T \Xi(x) \right\| \right\} \end{cases} \tag{14}$$

Remark 2: *In accordance with Assumption 2, obviously, it is known that both the composite disturbance $\bar{\mathcal{D}}_i$ and its change rate are all bounded, that is,*

$$|\dot{\bar{\mathcal{D}}}_i| \leq \bar{d}_i \tag{15}$$

where $\bar{d}_i > 0$ is the unknown boundary constant.

In this paper, the reinforcement learning strategy is introduced to enhance the precision of approximating the unknown uncertain terms. As a consequence, two types of neural networks (called Actor neural network and Critic neural network, respectively) need to be designed. First, the Actor neural network architecture is mathematically represented by the following formulation:

$$\hat{\bar{\mathcal{F}}}_i(Z_i) = \hat{\mathcal{W}}_{ai}^T \Xi_{ai}(Z_i) \tag{16}$$

where $\hat{\mathcal{W}}_{ai}$ denotes the estimation of the optimal weight value vector $\mathcal{W}_{ai} = \mathcal{W}_i^*$. Additionally, defining the approximation error as $\tilde{\mathcal{W}}_{ai} = \mathcal{W}_{ai} - \hat{\mathcal{W}}_{ai}$.

Remark 3: *As an effective approach in machine learning, reinforcement learning (RL) has demonstrated significant potential for real-time learning applications in intricate decision-making architectures. Recent advancements have facilitated the integration of RL techniques into adaptive and optimal control domains [38,39], enabling control systems to autonomously update strategies during operation while addressing optimization challenges defined by operational requirements. The fundamental mechanism of RL involves intelligent agents continuously engaging with their operating environments to refine decision-making policies. Typical RL frameworks incorporate dual neural network components: the Actor network generates executable control actions (functioning as compensatory signals), while the Critic network performs action quality evaluation (serving as a supervisory mechanism to assess both control effectiveness and Actor network performance metrics).*

To address unmodeled system dynamics and exogenous perturbations, the serial-parallel estimation methodology (SPEM) has been developed to approximate these complex disturbances through adaptive observation techniques [39]. Building upon this framework, the subsequent development of a nonlinear neural network disturbance observer (NNDO) enables precise reconstruction of composite disturbances. Specifically, the proposed NNDO architecture facilitates accurate acquisition of the composite disturbance estimation $\hat{\bar{D}}_i$ corresponding to the actual disturbance $\bar{D}_i$ in dynamical systems.

$$\begin{cases} \dot{\hat{\mathcal{E}}}_{i1} = p_{i1}\mathcal{Z}_{i1} + \mathcal{E}_{i2} + \varsigma_i + \nu_i - \dot{r}_i \\ \dot{\hat{\mathcal{E}}}_{i2} = p_{i2}\mathcal{Z}_{i2} + \mathcal{G}_i\mathcal{U}_i + \hat{\mathcal{W}}_{ai}^T \Xi_{ai}(Z_i) + \hat{\bar{\mathcal{D}}}_i, \end{cases} \tag{17}$$

where $p_{ij} > 0\,(j = 1, 2)$ are the positive constants, and $\mathcal{Z}_{ij} = \mathcal{E}_{ij} - \hat{\mathcal{E}}_{ij}, (j = 1, 2)$ denote the prediction errors. Obviously, the dynamic of $\mathcal{Z}_{ij}$ can be expressed as:

$$\begin{cases} \dot{\mathcal{Z}}_{i1} = -p_{i1}\mathcal{Z}_{i1}, \\ \dot{\mathcal{Z}}_{i2} = -p_{i2}\mathcal{Z}_{i2} + \tilde{\mathcal{W}}_{ai}^T \Xi_{ai}(Z_i) + \tilde{\bar{\mathcal{D}}}_i, \end{cases} \tag{18}$$

As noted previously, to estimate the composite disturbance $\bar{\mathcal{D}}_i$, a nonlinear NNDO is constructed in the following form:

$$\begin{cases} \dot{\varpi}_i = \mathcal{Z}_{i2} + \mathcal{E}_{i2} - c_i[\varpi_i + c_i\mathcal{E}_{i2} + \mathcal{G}_i\mathcal{U}_i + \hat{\mathcal{W}}_{ai}^T \Xi_{ai}(Z_i)] \\ \hat{\bar{\mathcal{D}}}_i = \varpi_i + c_i\mathcal{E}_{i2} \end{cases} \tag{19}$$

where $c_i > 0$ denotes the oberver gain parameter, and ϖ_i is an intermediate variable.

Denoting $\tilde{\bar{\mathcal{D}}}_i = \bar{\mathcal{D}}_i - \hat{\bar{\mathcal{D}}}_i$, it is naturally to obtain the estimation error dynamic, that is,

$$\dot{\tilde{\bar{\mathcal{D}}}}_i = -c_i\tilde{\bar{\mathcal{D}}}_i - \mathcal{Z}_{i2} - \mathcal{E}_{i2} - c_i\tilde{\mathcal{W}}_{ai}^T \Xi_{ai}(Z_i) + \dot{\bar{\mathcal{D}}}_i \tag{20}$$

Furthermore, to enable real-time implementation of the Actor-NN through reinforcement learning, an ideal value function (critic function) is formally defined as:

$$\zeta_i = \mathcal{Z}_{i2} + \mathcal{E}_{i2}^2 \mathcal{W}_{ci}^T \Xi_{ci}(Z_i) \tag{21}$$

where $\mathcal{W}_{ci}^T \Xi_{ci}(Z_i)$ denotes the theoretically optimal output of the Critic neural network. This configuration establishes a dual-component evaluative mechanism in (21): the primary term $\mathcal{Z}_{i2}$ quantifies subsystem approximation errors, while the secondary term embodies the stability criterion through reinforcement signal integration.

In practical implementation process, the ideal weight vector $\mathcal{W}_{ci}$ remains non-measurable. Through deployment of an online learning framework, based on the actual Critic-NN output $\hat{\mathcal{W}}_{ci}^T \Xi_{ci}(Z_i)$, the critic signal is generated via the following formula:

$$\hat{\zeta}_i = \mathcal{Z}_{i2} + \mathcal{E}_{i2}^2 \hat{\mathcal{W}}_{ci}^T \Xi_{ci}(Z_i) \tag{22}$$

where $\hat{\mathcal{W}}_{ci}$ denotes the estimation of the optimal weight value vector $\mathcal{W}_{ci}$. Additionally, defining the approximation error as $\tilde{\mathcal{W}}_{ci} = \mathcal{W}_{ci} - \hat{\mathcal{W}}_{ci}$.

The subsequent update laws are presented to generate $\hat{\mathcal{W}}_{ai}$ and $\hat{\mathcal{W}}_{ci}$:

$$\dot{\hat{\mathcal{W}}}_{ai} = \left(\hat{\zeta}_i + \mathcal{E}_{i2}\right) \Gamma_{ai} \Xi_{ai}(Z_i) - \Gamma_{ai}\eta_{i1}\hat{\mathcal{W}}_{ai} \tag{23}$$

$$\dot{\hat{\mathcal{W}}}_{ci} = -\mathcal{E}_{i2}^2 \Gamma_{ci} \Xi_{ci}(Z_i)\, \hat{\mathcal{W}}_{ai}^{\mathrm{T}} \Xi_{ai}(Z_i) - \Gamma_{ci}\eta_{i2}\hat{\mathcal{W}}_{ci} \tag{24}$$

where Γ_{ai} and Γ_{ci} are the diagonal positive definite gain matrices, and $\eta_{i1} > 0, \eta_{i2} > 0$ are the constant parameters.

Remark 4: *The Actor-Critic neural network framework operates through a temporal-forward computational paradigm, comprising two interactive components: the Actor module synthesizes adaptive correction outputs, while the Critic module executes real-time performance assessment via closed-loop state monitoring. Within this architecture, an enhanced value metric ζ_i is defined as (21), which serves as an enriched information carrier for the online adaptation process [38]. Given the inherent inaccessibility of optimal weight value parameter W_{ci}^*, in practical implementation process, an estimable counterpart $\hat{W}_{ci}^T \Xi_{ci}(Z_i)$ is exployed via the synthesized learning signal $\hat{\zeta}_i$ in (22). This discrepancy drives the parameter adaptation mechanism for both networks, see details as (23) and (24).*

3.2 Event-Triggered Tracking Controller Design

Step 1: According to the new error dynamic system (17), the virtual control law v_i is designed as:

$$v_{i1} = -\kappa_{i1}\mathcal{E}_{i1} + \dot{r}_i \tag{25}$$

where $\kappa_{i1} > 0$ is the designed constant parameter.

For the stability analysis, the Lyapunov function candidate can be selected as

$$\mathcal{V}_{i1} = \frac{1}{2}\mathcal{E}_{i1}^2 + \frac{1}{2}\mathcal{Z}_{i1}^2 + \frac{1}{2}\varsigma_i^2 \tag{26}$$

Taking the time derivative of (26), and combining with (9), (17), we have

$$\begin{aligned}\dot{\mathcal{V}}_{i1} &= \mathcal{E}_{i1}\dot{\mathcal{E}}_{i1} + \mathcal{Z}_{i1}\dot{\mathcal{Z}}_{i1} + \varsigma_i\dot{\varsigma}_i \\ &= \mathcal{E}_{i1}(\mathcal{E}_{i2} + \varsigma_i + v_{i1} - \dot{r}_i) - p_{i1}\mathcal{Z}_{i1}^2 + \varsigma_i\dot{\varsigma}_i\end{aligned} \tag{27}$$

According to (9), it yields $\dot{\varsigma}_i = -\frac{1}{\sigma_i}\varsigma_i - \dot{v}_{i1}$. Moreover, we can further assume that there exists an unknown positive constant $\bar{b}_i > 0$, such that $|\dot{\varsigma}_i + \frac{1}{\sigma_i}\varsigma_i| < \bar{b}_i$. Furthermore, one obtains

$$\varsigma_i\dot{\varsigma}_i \le -(\frac{1}{\sigma_i} - 4)\varsigma_i^2 + \frac{1}{16}\bar{b}_i^2 \tag{28}$$

Substituting (25) and (28) into (27), and based on fundamental inequalities, one has

$$\begin{aligned}\dot{\mathcal{V}}_{i1} &\le -\kappa_i\mathcal{E}_{i1}^2 + \mathcal{E}_{i1}\mathcal{E}_{i2} + \mathcal{E}_{i1}\varsigma_i - p_{i1}\mathcal{Z}_{i1}^2 - (\frac{1}{\varsigma_i} - 4)\varsigma_i^2 + \frac{1}{16}\bar{b}_i^2 \\ &\le -(\kappa_i - 1)\mathcal{E}_{i1}^2 + \frac{1}{2}\mathcal{E}_{i2}^2 - p_{i1}\mathcal{Z}_{i1}^2 - (\frac{1}{\varsigma_i} - \frac{9}{2})\varsigma_i^2 + \frac{1}{16}\bar{b}_i^2\end{aligned} \tag{29}$$

Step 2: According to (17), it yields

$$\dot{\mathcal{E}}_{i2} = \mathcal{G}_i\mathcal{U}_i + \hat{\mathcal{W}}_{ai}^T\Xi_i(Z_i) + \tilde{\mathcal{W}}_{ai}^T\Xi_i(Z_i) + \hat{\bar{\mathcal{D}}}_i + \tilde{\bar{\mathcal{D}}}_i, \tag{30}$$

To conduct stability analysis, the following Lyapunov function is selected.

$$\mathcal{V}_{i2} = \frac{1}{2}\mathcal{E}_{i2}^2 + \frac{1}{2}\tilde{\mathcal{W}}_{ai}^T\Gamma_{ai}^{-1}\tilde{\mathcal{W}}_{ai} + \frac{1}{2}\tilde{\mathcal{W}}_{ci}^T\Gamma_{ci}^{-1}\tilde{\mathcal{W}}_{ci} + \frac{1}{2}\mathcal{Z}_{i2}^2 + \frac{1}{2}\tilde{\bar{\mathcal{D}}}_i^2 + \frac{1}{2}\tilde{\Lambda}_i^T\Upsilon_i\tilde{\Lambda}_i \tag{31}$$

where $\tilde{\Lambda}_i = \Lambda_i - \hat{\Lambda}_i$ with $\lambda_{ij}, (j = 1, 2)$ being the unknown constants and $\Lambda_i = [\lambda_{i1}, \lambda_{i2}]^T$, $\hat{\Lambda}_i$ denotes the estimation of Λ_i.

Further, according to (18), (20), (23), and (24), the time derivative of $\mathcal{V}_{i2}$ can be obtained:

$$\begin{aligned}
\dot{\mathcal{V}}_{i2} &= \mathcal{E}_{i2}\dot{\mathcal{E}}_{i2} + \tilde{\mathcal{W}}_{ai}^T\Gamma_{ai}^{-1}\dot{\hat{\mathcal{W}}}_{ai} + \tilde{\mathcal{W}}_{ci}^T\Gamma_{ci}^{-1}\dot{\hat{\mathcal{W}}}_{ci} + \mathcal{Z}_{i2}\dot{\mathcal{Z}}_{i2} + \tilde{\bar{\mathcal{D}}}_i\dot{\bar{\mathcal{D}}}_i - \tilde{\Lambda}_i^T\Upsilon_i\dot{\hat{\Lambda}}_i \\
&= \mathcal{E}_{i2}(\mathcal{G}_i\mathcal{U}_i + \hat{\mathcal{W}}_{ai}^T\Xi_i(Z_i) + \tilde{\mathcal{W}}_{ai}^T\Xi_i(Z_i) + \hat{\bar{\mathcal{D}}}_i + \tilde{\bar{\mathcal{D}}}_i) \\
&\quad - \tilde{\mathcal{W}}_{ai}^T\Gamma_{ai}^{-1}[\left(\hat{\xi}_i + \mathcal{E}_{i2}\right)\Gamma_{ai}\Xi_{ai}(Z_i) - \Gamma_{ai}\eta_{i1}\hat{\mathcal{W}}_{ai}] \\
&\quad - \tilde{\mathcal{W}}_{ci}^T\Gamma_{ci}^{-1}[-\mathcal{E}_{i2}^2\Gamma_{ci}\Xi_{ci}(Z_i)\hat{\mathcal{W}}_{ai}^{\mathrm{T}}\Xi_{ai}(Z_i) - \Gamma_{ci}\eta_{i2}\hat{\mathcal{W}}_{ci}] \\
&\quad + \mathcal{Z}_{i2}(-p_{i2}\mathcal{Z}_{i2} + \mathcal{G}_i\mathcal{U}_i + \tilde{\mathcal{W}}_{ai}^T\Xi_{ai}(Z_i) + \tilde{\bar{\mathcal{D}}}_i) \\
&\quad + \tilde{\bar{\mathcal{D}}}_i[-c_i\tilde{\bar{\mathcal{D}}}_i - \mathcal{Z}_{i2} - \mathcal{E}_{i2} - c_i\tilde{\mathcal{W}}_{ai}^T\Xi_{ai}(Z_i) + \dot{\bar{\mathcal{D}}}_i] - \tilde{\Lambda}_i^T\Upsilon_i\dot{\hat{\Lambda}}_i \\
&= \mathcal{E}_{i2}\mathcal{G}_i\mathcal{U}_i + \mathcal{E}_{i2}\hat{\mathcal{W}}_{ai}^T\Xi_i(Z_i) + \mathcal{E}_{i2}\hat{\bar{\mathcal{D}}}_i + \eta_{i1}\tilde{\mathcal{W}}_{ai}^T\hat{\mathcal{W}}_{ai} + \eta_{i2}\tilde{\mathcal{W}}_{ci}^T\hat{\mathcal{W}}_{ci} \\
&\quad - \mathcal{E}_{i2}^2\hat{\mathcal{W}}_{ci}^T\Xi_{ci}(Z_i)\tilde{\mathcal{W}}_{ai}^T\Xi_{ai}(Z_i) + \mathcal{E}_{i2}^2\tilde{\mathcal{W}}_{ci}^T\Xi_{ci}(Z_i)\hat{\mathcal{W}}_{ai}^{\mathrm{T}}\Xi_{ai}(Z_i) \\
&\quad - p_{i2}\mathcal{Z}_{i2}^2 - c_i\tilde{\bar{\mathcal{D}}}_i^{\,2} - c_i\tilde{\mathcal{W}}_{ai}^T\Xi_{ai}(Z_i)\tilde{\bar{\mathcal{D}}}_i + \tilde{\bar{\mathcal{D}}}_i\dot{\bar{\mathcal{D}}}_i - \tilde{\Lambda}_i^T\Upsilon_i\dot{\hat{\Lambda}}_i
\end{aligned} \tag{32}$$

As a matter of convenience, denoting $\bar{\mathcal{W}}_i = [\|\mathcal{W}_{ai}^T\|, \|\hat{\mathcal{W}}_{ci}^T\|]^T$, we have

$$\begin{aligned}
&- \mathcal{E}_{i2}\tilde{\mathcal{W}}_{ai}^T\Xi_{ai}(Z_i)\hat{\mathcal{W}}_{ci}^T\Xi_{ci}(Z_i) + \mathcal{E}_{i2}\tilde{\mathcal{W}}_{ci}^T\Xi_{ci}(Z_i)\hat{\mathcal{W}}_{ai}^{\mathrm{T}}\Xi_{ai}(Z_i) \\
&= -\mathcal{E}_{i2}\mathcal{W}_{ai}^T\Xi_{ai}(Z_i)\hat{\mathcal{W}}_{ci}^T\Xi_{ci}(Z_i) + \mathcal{E}_{i2}\hat{\mathcal{W}}_{ai}^T\Xi_{ai}(Z_i)\hat{\mathcal{W}}_{ci}^T\Xi_{ci}(Z_i) \\
&+ \mathcal{E}_{i2}\mathcal{W}_{ci}^T\Xi_{ci}(Z_i)\hat{\mathcal{W}}_{ai}^{\mathrm{T}}\Xi_{ai}(Z_i) - \mathcal{E}_{i2}\hat{\mathcal{W}}_{ci}^T\Xi_{ci}(Z_i)\hat{\mathcal{W}}_{ai}^{\mathrm{T}}\Xi_{ai}(Z_i) \\
&= -\mathcal{E}_{i2}\mathcal{W}_{ai}^T\Xi_{ai}(Z_i)\hat{\mathcal{W}}_{ci}^T\Xi_{ci}(Z_i) + \mathcal{E}_{i2}\mathcal{W}_{ci}^T\Xi_{ci}(Z_i)\hat{\mathcal{W}}_{ai}^{\mathrm{T}}\Xi_{ai}(Z_i) \\
&\leq \mathcal{E}_{i2}^2\bar{\Xi}_{ai}\bar{\Xi}_{ci}\|\mathcal{W}_{ai}\|\|\hat{\mathcal{W}}_{ci}\| + \mathcal{E}_{i2}^2\bar{\Xi}_{ci}\bar{\Xi}_{ai}\|\mathcal{W}_{ci}\|\|\hat{\mathcal{W}}_{ai}\| \\
&=: \mathcal{E}_{i2}^2\Lambda_i^T\bar{\mathcal{W}}_i
\end{aligned} \tag{33}$$

where $\bar{\Xi}_{ai}$ and $\bar{\Xi}_{ci}$ denote the upper bounds of $\Xi_{ai}(Z_i)$ and $\Xi_{ai}(Z_i)$, repectively.

Substituting (33) into (34), one has

$$\begin{aligned}
\dot{\mathcal{V}}_{i2} &\leq \mathcal{E}_{i2}\mathcal{G}_i\mathcal{U}_i + \mathcal{E}_{i2}\hat{\mathcal{W}}_{ai}^T\Xi_i(Z_i) + \mathcal{E}_{i2}\hat{\bar{\mathcal{D}}}_i + \eta_{i1}\tilde{\mathcal{W}}_{ai}^T\hat{\mathcal{W}}_{ai} + \eta_{i2}\tilde{\mathcal{W}}_{ci}^T\hat{\mathcal{W}}_{ci} \\
&\quad + \mathcal{E}_{i2}^2\Lambda_i^T\bar{\mathcal{W}}_i - p_{i2}\mathcal{Z}_{i2}^2 - c_i\tilde{\bar{\mathcal{D}}}_i^{\,2} - c_i\tilde{\mathcal{W}}_{ai}^T\Xi_{ai}(Z_i)\tilde{\bar{\mathcal{D}}}_i + \tilde{\bar{\mathcal{D}}}_i\dot{\bar{\mathcal{D}}}_i - \tilde{\Lambda}_i^T\Upsilon_i\dot{\hat{\Lambda}}_i
\end{aligned} \tag{34}$$

In this step, the actual event-triggered tracking controller will be proposed. Until then, the virtual control signal ν_{i2} and parameter adptive updating law will be first designed as follows:

$$\nu_{i2} = -\mathcal{G}_i^{-1}\left(\kappa_{i2}\mathcal{E}_{i2} + \frac{1}{2}\mathcal{E}_{i2} + \hat{\mathcal{W}}_{ai}^T\Xi_{ai}(Z_i) + \hat{\bar{\mathcal{D}}}_i + \mathcal{E}_{i2}\hat{\Lambda}_i^T\bar{\mathcal{W}}_i\right) \tag{35}$$

$$\dot{\hat{\Lambda}}_i = \Upsilon_i^{-1}(\mathcal{E}_{i2}^2\bar{\mathcal{W}}_i - \eta_{i3}\hat{\Lambda}_i) \tag{36}$$

Based on the relative-threshold-based control mechanism, in view of (7), an event-triggered robust tracking controller is developed in the following form:

$$\bar{\mathcal{U}}_i(t) = -(1+\rho_i)\left(\nu_{i2}\tanh\left(\frac{\mathcal{E}_{i2}\mathcal{G}_i\nu_{i2}}{\tau_i}\right) + \gamma_{i1}\tanh\left(\frac{\mathcal{E}_{i2}\mathcal{G}_i\gamma_{i1}}{\tau_i}\right)\right) \tag{37}$$

$$\mathcal{U}_i(t) = \bar{\mathcal{U}}_i(t_k), \quad \forall t \in [t_k, t_{k+1}), k \in \mathbb{Z}^+ \tag{38}$$

$$t_{k+1} = \inf\{t \in R || \epsilon_i(t)| \geq \rho_i |\mathcal{U}_i(t)| + \gamma_{i2}\} \tag{39}$$

where $\epsilon_i(t) = \bar{\mathcal{U}}_i(t) - \mathcal{U}_i(t)$, and the design parameters satisfy $\tau_i > 0$, $\gamma_{i1} > \dfrac{\gamma_{i2}}{1-\rho_i}$, $\gamma_{i2} > 0$, $0 < \rho_i < 1$. Obviously, the control signal is a constant in $[t_k, t_{k+1})$, that is $\bar{\mathcal{U}}_i(t_k)$.

From (7), we have $|\bar{\mathcal{U}}_i(t) - \mathcal{U}_i(t)| < \rho_i|\mathcal{U}_i(t)| + \gamma_{i2}$ in $[t_k, t_{k+1})$. Thereby, $\forall t \in [t_k, t_{k+1})$, there exist two continuous functions $\ell_{i1}(t)$ and $\ell_{i2}(t)$ satisfying $|\ell_{i1}(t)| \leq 1$ and $|\ell_{i2}(t)| \leq 1$, so that

$$\mathcal{U}_i(t) = \frac{\bar{\mathcal{U}}_i(t) - \gamma_{i2}\ell_{i2}(t)}{1+\rho_i\ell_{i1}(t)}. \tag{40}$$

By combining with (35)–(40), the comprehensive calculation is performed, one obtain

$$\begin{aligned}
\dot{\mathcal{V}}_{i2} \leq & -\kappa_{i2}\mathcal{E}_{i2}^2 - \frac{1}{2}\mathcal{E}_{i2}^2 - p_{i2}\mathcal{Z}_{i2}^2 - c_i\tilde{\bar{\mathcal{D}}}_i^2 - \mathcal{E}_{i2}\mathcal{G}_i\nu_{i2} \\
& - \frac{1+\rho_i}{1+\rho_i\ell_{i1}(t)}\mathcal{E}_{i2}\mathcal{G}_i\nu_{i2}\tanh\left(\frac{\mathcal{E}_{i2}\mathcal{G}_i\nu_{i2}}{\tau_i}\right) \\
& - \frac{1+\rho_i}{1+\rho_i\ell_{i1}(t)}\mathcal{E}_{i2}\mathcal{G}_i\gamma_{i1}\tanh\left(\frac{\mathcal{E}_{i2}\mathcal{G}_i\gamma_{i1}}{\tau_i}\right) \\
& - \frac{\mathcal{E}_{i2}\mathcal{G}_i\gamma_{i2}\ell_{i2}(t)}{1+\rho_i\ell_{i1}(t)} + \eta_{i1}\tilde{\mathcal{W}}_{ai}^T\hat{\mathcal{W}}_{ai} + \eta_{i2}\tilde{\mathcal{W}}_{ci}^T\hat{\mathcal{W}}_{ci} \\
& - c_i\tilde{\mathcal{W}}_{ai}^T\Xi_{ai}(Z_i)\tilde{\bar{\mathcal{D}}}_i + \tilde{\bar{\mathcal{D}}}_i\dot{\bar{\mathcal{D}}}_i + \eta_{i3}\tilde{\Lambda}_i^T\hat{\Lambda}_i
\end{aligned} \tag{41}$$

In fact, these two inequalities hold:

$$0 < 1 + \rho_i\ell_{i1}(t) \leq 1 + \rho_i \tag{42}$$

$$-\frac{\mathcal{E}_{i2}\mathcal{G}_i\gamma_{i2}\ell_{i2}(t)}{1+\rho_i\ell_{i1}(t)} \leq \left|\frac{\mathcal{E}_{i2}\mathcal{G}_i\gamma_{i2}}{1-\rho_i}\right| \tag{43}$$

According to (42) and (43), we have

$$\begin{aligned}
\dot{\mathcal{V}}_{i2} \leq & -\kappa_{i2}\mathcal{E}_{i2}^2 - \frac{1}{2}\mathcal{E}_{i2}^2 - p_{i2}\mathcal{Z}_{i2}^2 - c_i\tilde{\bar{\mathcal{D}}}_i^2 + |\mathcal{E}_{i2}\mathcal{G}_i\nu_{i2}| \\
& - \mathcal{E}_{i2}\mathcal{G}_i\nu_{i2}\tanh\left(\frac{\mathcal{E}_{i2}\mathcal{G}_i\nu_{i2}}{\tau_i}\right) + |\mathcal{E}_{i2}\mathcal{G}_i\gamma_{i1}|
\end{aligned}$$

$$
\begin{aligned}
&- \mathcal{E}_{i2}\mathcal{G}_i\gamma_{i1}\tanh\left(\frac{\mathcal{E}_{i2}\mathcal{G}_i\gamma_{i1}}{\tau_i}\right) - |\mathcal{E}_{i2}\mathcal{G}_i\gamma_{i1}| \\
&+ \left|\frac{\mathcal{E}_{i2}\mathcal{G}_i\gamma_{i2}}{1-\rho_i}\right| + \eta_{i1}\tilde{\mathcal{W}}_{ai}^T\hat{\mathcal{W}}_{ai} + \eta_{i2}\tilde{\mathcal{W}}_{ci}^T\hat{\mathcal{W}}_{ci} \\
&- c_i\tilde{\mathcal{W}}_{ai}^T\Xi_{ai}(Z_i)\tilde{\bar{\mathcal{D}}}_i + \tilde{\bar{\mathcal{D}}}_i\dot{\bar{\mathcal{D}}}_i + \eta_{i3}\tilde{\Lambda}_i^T\hat{\Lambda}_i
\end{aligned} \tag{44}
$$

On the basis of Young's inequality, it yields

$$\eta_{i1}\tilde{\mathcal{W}}_{ai}^T\hat{\mathcal{W}}_{ai} \le \frac{\eta_{i1}}{2}\|\mathcal{W}_{ai}\|^2 - \frac{\eta_{i1}}{2}\|\tilde{\mathcal{W}}_{ai}\|^2 \tag{45}$$

$$\eta_{i2}\tilde{\mathcal{W}}_{ci}^T\hat{\mathcal{W}}_{ci} \le \frac{\eta_{i2}}{2}\|\mathcal{W}_{ci}\|^2 - \frac{\eta_{i2}}{2}\|\tilde{\mathcal{W}}_{ci}\|^2 \tag{46}$$

$$-c_i\tilde{\mathcal{W}}_{ai}^T\Xi_{ai}(Z_i)\tilde{\bar{\mathcal{D}}}_i \le \frac{c_i}{2}\|\tilde{\mathcal{W}}_{ai}\|^2 + \frac{c_i}{2}\tilde{\bar{\mathcal{D}}}_i^2 \tag{47}$$

$$\tilde{\bar{\mathcal{D}}}_i\dot{\bar{\mathcal{D}}}_i \le \frac{1}{2}\tilde{\bar{\mathcal{D}}}_i^2 + \frac{\bar{d}_i^2}{2} \tag{48}$$

$$\eta_{i3}\tilde{\Lambda}_i^T\hat{\Lambda}_i \le \frac{\eta_{i3}}{2}\|\Lambda_i\|^2 - \frac{\eta_{i3}}{2}\|\tilde{\Lambda}_i\|^2 \tag{49}$$

According to Lemma 2, and combining with (45)–(49), we obtain

$$
\begin{aligned}
\dot{\mathcal{V}}_{i2} \le\ & -\kappa_{i2}\mathcal{E}_{i2}^2 - \frac{1}{2}\mathcal{E}_{i2}^2 - \left(\frac{\eta_{i1}}{2} - \frac{c_i}{2}\right)\|\tilde{\mathcal{W}}_{ai}\|^2 - \frac{\eta_{i2}}{2}\|\tilde{\mathcal{W}}_{ci}\|^2 \\
&- p_{i2}\mathcal{Z}_{i2}^2 - \left(\frac{c_i}{2} - \frac{1}{2}\right)\tilde{\bar{\mathcal{D}}}_i^2 - \frac{\eta_{i3}}{2}\|\tilde{\Lambda}_i\|^2 + 2\varrho\tau_i \\
&+ \frac{\eta_{i1}}{2}\|\mathcal{W}_{ai}\|^2 + \frac{\eta_{i2}}{2}\|\mathcal{W}_{ci}\|^2 + \frac{\bar{d}_i^2}{2} + \frac{\eta_{i3}}{2}\|\Lambda_i\|^2
\end{aligned} \tag{50}
$$

To sum up, through systematic integration of the multi-stage synthesis framework encompassing disturbance estimation, neural approximation, and event-triggered compensation mechanisms, the core theoretical contribution of this work is formally encapsulated in the following theorem. Subsequently, a rigorous stability verification framework is established for the resulting closed-loop architecture.

Theorem 1: *Consider the QUAV system with six degree of freedom (1), external disturbances and the uncertain system terms, by introducing the reinforcement learning strategy, and developing the filter (9), the NNDO (19), two types of newral networks with the adaptive updating laws (23) and (24), the virtual control laws (25) and (35), the event-triggered robust tracking controller is designed as (40), and then the designed control can ensure that all the system signals are uniformly ultimately bounded. Meanwhile, Zeno behavior will not occur.*

Proof: To conduct the entire system stability analysis, we propose the following Lyapunov function candidate:

$$\mathcal{V}_i^* = \mathcal{V}_{i1} + \mathcal{V}_{i2} \tag{51}$$

Taking the time derivative of (51), and by utilizing (29) and (50), one has

$$\begin{aligned}\dot{\mathcal{V}}_i^* \le & -(\kappa_i-1)\mathcal{E}_{i1}^2-\kappa_{i2}\mathcal{E}_{i2}^2-p_{i1}\mathcal{Z}_{i1}^2-\left(\frac{1}{\varsigma_i}-\frac{9}{2}\right)\varsigma_i^2-\left(\frac{\eta_{i1}}{2}-\frac{c_i}{2}\right)\|\tilde{\mathcal{W}}_{ai}\|^2\\ & -\frac{\eta_{i2}}{2}\|\tilde{\mathcal{W}}_{ci}\|^2-p_{i2}\mathcal{Z}_{i2}^2-\left(\frac{c_i}{2}-\frac{1}{2}\right)\tilde{\bar{\mathcal{D}}}_i^2-\frac{\eta_{i3}}{2}\|\tilde{\Lambda}_i\|^2+\frac{1}{16}\bar{b}_i^2+2\varrho\tau_i\\ & +\frac{\eta_{i1}}{2}\|\mathcal{W}_{ai}\|^2+\frac{\eta_{i2}}{2}\|\mathcal{W}_{ci}\|^2+\frac{\bar{d}_i^2}{2}+\frac{\eta_{i3}}{2}\|\Lambda_i\|^2\\ & \le -\mathcal{A}\mathcal{V}_i^*+\mathcal{B}\end{aligned} \tag{52}$$

where $\mathcal{A}=\min\left\{2(\kappa_i-1),\ 2\kappa_{i2},\ 2p_{i1},\ 2p_{i2},\ \frac{2}{\varsigma_i}-9,\ \left(\frac{\eta_{i1}}{2}-\frac{c_i}{2}\right)/\lambda_{max}(\Gamma_{ai}^{-1}),\eta_{i2}/\lambda_{max}(\Gamma_{ci}^{-1}),\ c_i-1,\eta_{i3}/\lambda_{max}(\Upsilon_i^{-1})\right\}$, $\mathcal{B}=\frac{1}{16}\bar{b}_i^2+2\varrho\tau_i+\frac{\eta_{i1}}{2}\|\mathcal{W}_{ai}\|^2+\frac{\eta_{i2}}{2}\|\mathcal{W}_{ci}\|^2+\frac{\bar{d}_i^2}{2}+\frac{\eta_{i3}}{2}\|\Lambda_i\|^2$.

Furthermore, by using $e^{\mathcal{A}t}$ to multiply by (52), and calculating definite integral on $[0,t]$, one has

$$\mathcal{V}_i^*\le\left(\mathcal{V}_i^*(0)-\frac{\mathcal{B}}{\mathcal{A}}\right)e^{\mathcal{A}t}+\frac{\mathcal{B}}{\mathcal{A}}\le\mathcal{V}_i^*(0)+\frac{\mathcal{B}}{\mathcal{A}} \tag{53}$$

Obviously, in view of the definition of $\mathcal{V}_i^*$, it is implied that all the signals of the closed-loop system are bounded.

On the other hand, if it is proven that there is no Zeno behavior, it is actually to prove that there exist s a time constant t_i^* such that $t_{k+1}-t_k\ge t^*,\forall k\in Z^+$. In fact, note that $\lim\limits_{t\in t_{k+1}}\epsilon_i(t)=\rho_i|\mathcal{U}_i(t)\mid+\gamma_{i2}$ and $|\dot{\bar{\mathcal{U}}}_i|\le\iota_i$ with ι_i being a positive constant, then we have $t_{k+1}-t_k\ge\frac{\rho_i|\mathcal{U}_i(t)|+\gamma_{i2}}{\iota_i}$, that is, there is a lower bounds satisfying $t_i^*=\frac{\gamma_{i2}}{\iota_i}$. □

4 Simulation Experiment Results

In this section, to verify the effectiveness of the robust adaptive trajectory and attitude tracking control scheme for QUAV, a set of simulation experiments has been carried out and the satisfactory results have been obtained. The relevant model parameters are listed in Table 1.

Table 1: The relevant model parameters

Variable	Value	Unit
m	2	kg
g	9.8	m/s^2
L	0.325	m
a_x, a_y, a_z, a_ϕ, a_θ, a_ψ	0.6	N·s/rad
I_x	0.082	kg · m^2
I_y	0.082	kg · m^2
I_z	0.149	kg · m^2

The simulation configuration adopts the following parametric settings:

- **Reference signals:** $r_x = 2\sin\left(\frac{\pi}{10}t\right)$, $r_y = 2\cos\left(\frac{\pi}{10}t\right)$, $r_z = \frac{1}{6}t$, $r_\psi = \frac{\pi}{4}$, r_ϕ and r_θ are generated by (4).
- **External disturbances:** $d_x = 0.1\cos(30t)$, $d_y = 0.1\sin(20t)$, $d_z = 0.1\cos(30t)$, $d_\phi = 0.1\sin(30t)$, $d_\theta = \cos(30t)$, and $d_\psi = 0.1\sin(20t)$.
- **System control parameters:** unless otherwise specified, note that $i = 1, 2, \ldots, 6$.
- Filter parameters are set as: $\sigma_i = 0.5$. The gain parameters of SPEM are set as: $p_{11} = p_{21} = p_{31} = 6, p_{41} = p_{51} = p_{61} = 4$, $p_{12} = p_{22} = p_{32} = 10, p_{42} = p_{52} = p_{62} = 5$. The configurations of NNDO are set as: $c_1 = c_2 = c_3 = 10, c_4 = c_5 = c_6 = 5$. The parameters of actor-NN, critic-NN and adaptive estimation are selected as: $\Gamma_{ai} = diag\{20, 20, 20, 20, 20, 20\}$, $\Gamma_{ci} = diag\{10, 10, 10, 10, 10, 10\}$, $\Upsilon_i = diag\{10, 10, 10, 10, 10, 10\}$, $\eta_{i1} = \eta_{i2} = 2, \eta_{i3} = 1$. The event-triggered control parameters are selected as: $\kappa_{11} = \kappa_{21} = \kappa_{31} = 2.5, \kappa_{41} = \kappa_{51} = \kappa_{61} = 2$, $\kappa_{12} = \kappa_{22} = \kappa_{32} = 30, \kappa_{42} = \kappa_{52} = \kappa_{62} = 5$, $\rho_i = 0.01, \tau_i = 0.1, \gamma_{i1} = 0.11, \gamma_{i2} = 0.01$.
- **Initial conditions:** $[x(0), y(0), z(0)]^T = [1, 2, 0.5]^T$ and $[\phi(0), \theta(0), \psi(0)]^T = [0.6, 0.1, 0]^T$.

The proposed event-triggered robust constrained control scheme demonstrates superior performance through numerical simulations, as evidenced by the following quantitative analyses across Figs. 2–6. As visualized in Fig. 2, the QUAV exhibits precise trajectory tracking performance. Figs. 3 and 4 reveal that the actual position and attitude angular can effectively track the reference signals with a relatively small errors.

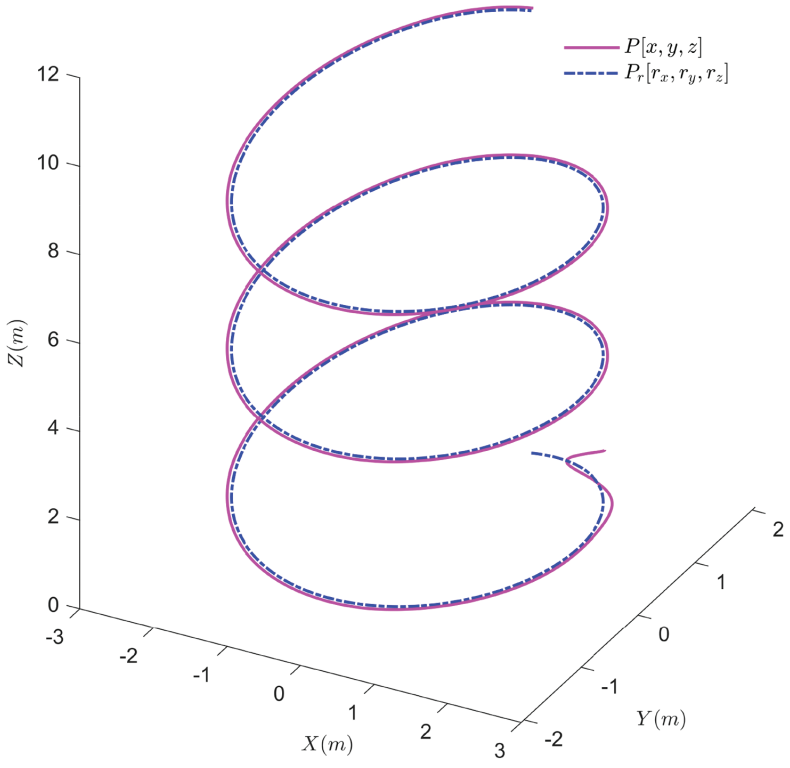
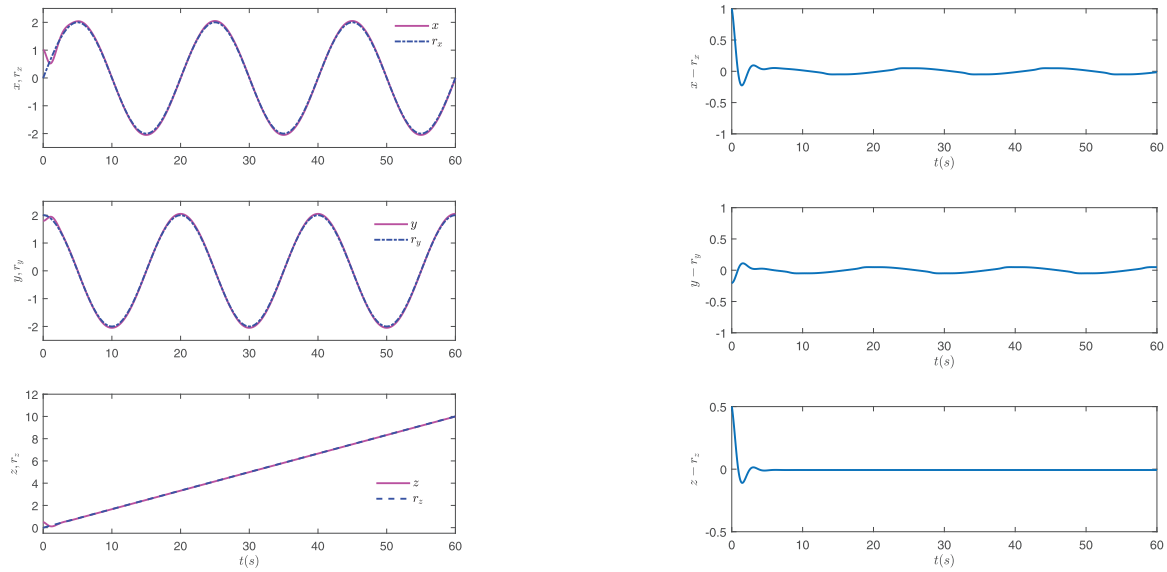
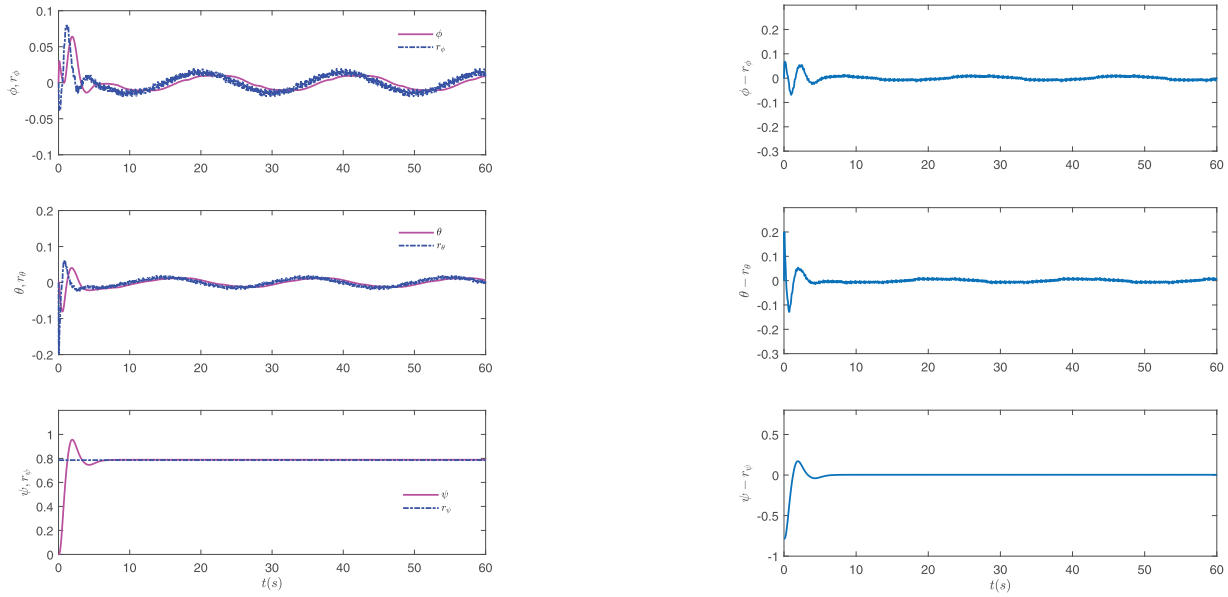


Figure 2: The 3D trajectory tracking results



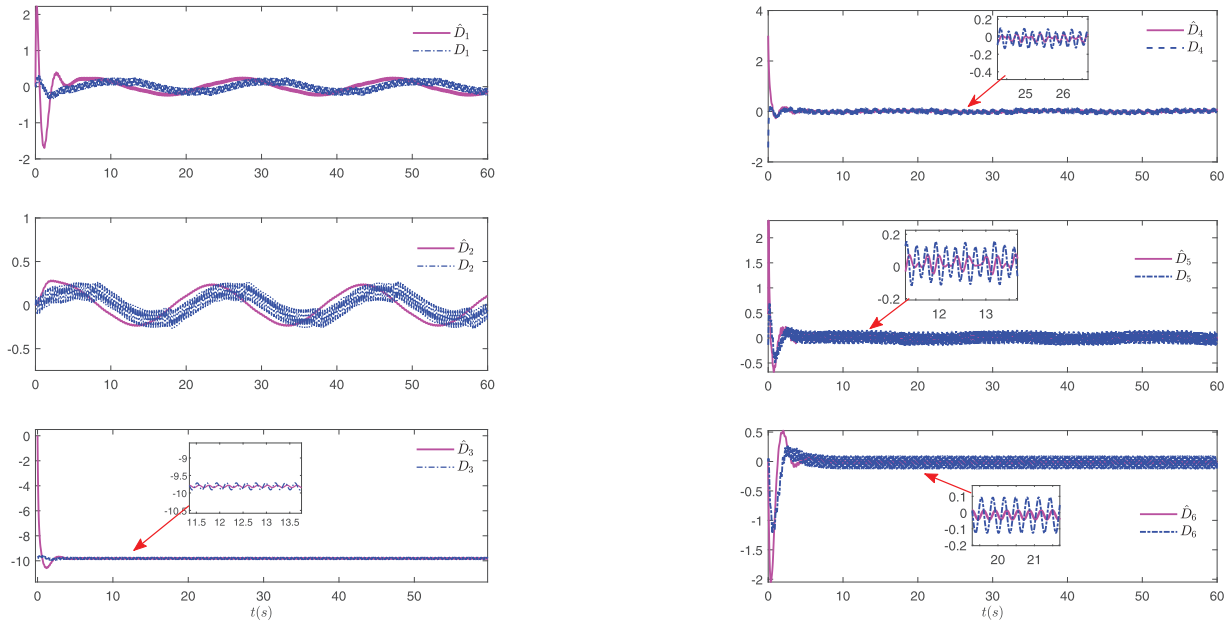
(a) (b)

Figure 3: The tracking results of position components: (**a**) The tracking curves of position components. (**b**) The curves of position tracking errors



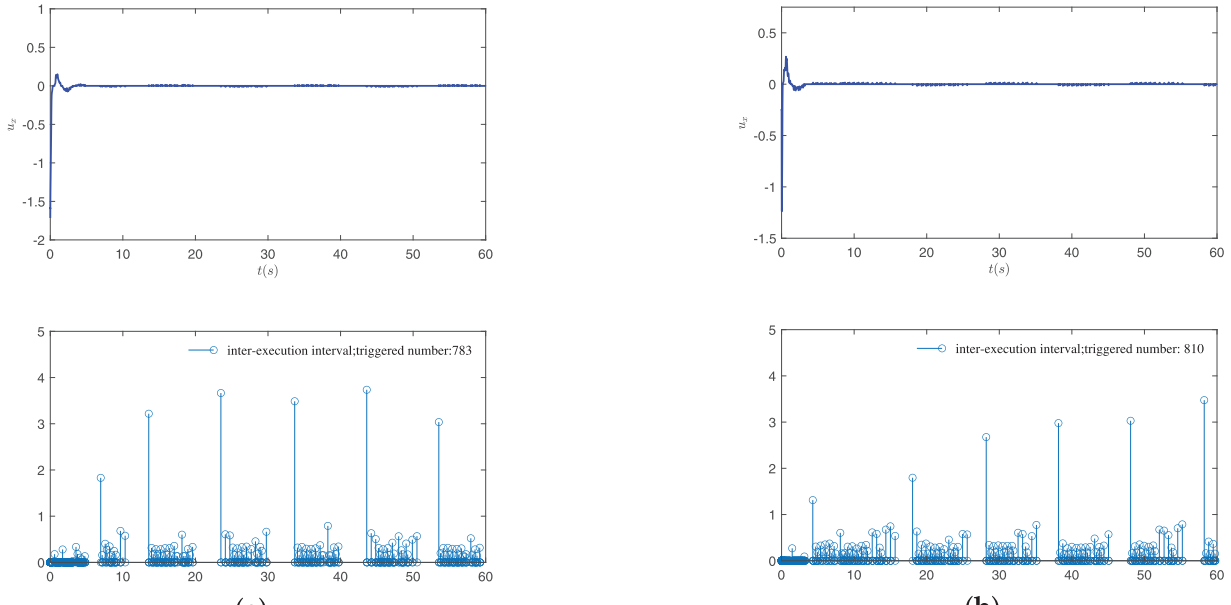
(a) (b)

Figure 4: The tracking results of attitude angular components: (**a**) The tracking curves of attitude angular components. (**b**) The curves of attitude angular tracking errors



(a) (b)

Figure 5: The disturbance estimation results of NNDOs: (**a**) Position loop. (**b**) Attitude loop



(a) (b)

Figure 6: The curves of event-triggered control inputs, taking (**a**) u_ϕ and (**b**) u_θ as examples

The developed NNDOs achieve the satisfactory effect for composite disturbance estimation, as quantified in Fig. 5. The observer's mean square errors (MSEs) are shown in Table 2. The fact verifies the dual neural network architecture's capability in uncertainty approximation and disturbance rejection.

Table 2: The mean square errors of NNDOs

Disturbance estimation	$\hat{\bar{\mathcal{D}}}_1$	$\hat{\bar{\mathcal{D}}}_2$	$\hat{\bar{\mathcal{D}}}_3$	$\hat{\bar{\mathcal{D}}}_3$	$\hat{\bar{\mathcal{D}}}_5$	$\hat{\bar{\mathcal{D}}}_6$
MSEs	0.0120	0.0143	0.0962	0.0397	0.0052	0.0174

The event-triggering mechanism demonstrates superior resource efficiency, as evidenced by the inter-event intervals shown in Fig. 6, and it is shown that the developed event-triggered tracking controller can effectively reduce the consumption with ensuring control stability.

5 Conclusions

This study has established a novel adaptive event-triggered control framework for QUAV that systematically addresses trajectory and attitude tracking under external disturbances. By combining reinforcement learning-based neural networks with nonlinear disturbance estimation, this architecture achieves robust uncertainty compensation. The dual NN mechanism enables dynamic performance evaluation and adaptive policy refinement. This paper verifies the scheme's effectiveness under single QUAV, but it still has limitations: (1) It does not consider QUAV's payload changes (e.g., delivery tasks); (2) The event-triggered threshold is fixed, which may not adapt to dynamic environments. Future work will: (1) Integrate payload estimation into the NNDO to handle mass variations; (2) Extend the single QUAV control framework to swarm systems using federated learning, ensuring data privacy during collaborative control.

Acknowledgement: Not applicable.

Funding Statement: This research was funded by State Grid Anhui Electric Power Co., Ltd. Technology Project under grant 5212D0240002.

Author Contributions: The authors confirm contribution to the paper as follows: Methodology and writing—original draft preparation, Baisong Wang; methodology and review, Kai Wu; resources and writing—review, Yang Wang, Zhugang Chu; formal analysis and software, Kebiao Liu. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: Not applicable.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest to report regarding the present study.

References

1. Zhou X, Yu X, Guo K, Zhou S, Guo L, Zhang Y, et al. Safety flight control design of a quadrotor UAV with capability analysis. IEEE Trans Cybern. 2023;53(3):1738–51.
2. Wang B, Zhang Y, Zhang W. Integrated path planning and trajectory tracking control for QUAV with obstacle avoidance in the presence of environmental and systematic uncertainties: theory and experiment. Aerosp Sci Technol. 2022;120:107277.
3. Invernizzi D, Giurato M, Gattazzo P. Comparison of control methods for trajectory tracking in fully actuated unmanned aerial vehicles. IEEE Trans Control Syst Technol. 2020;29(3):1147–60.

4. Zahid F, Abro GEM, Alansari Z. Recent advancements in manipulator-equipped unmanned aerial vehicles (UAVs). In: Proceedings of the 1st International Conference on Innovative Engineering Sciences and Technological Research (ICIESTR); 2024 May 14–15; Muscat, Oman. p. 1–4.
5. Liu Z, Xiong L, Shi W. Trajectory tracking control for a QUAV with performance constraints. IEEE Access. 2019;7:142467–77.
6. Abro GEM, Abdallah AM, Zahid F, Ahmed S. A comprehensive review of next-gen UAV swarm robotics: optimisation techniques and control strategies for dynamic environments. Intell Autom Soft Comput. 2025;40(1):99–123.
7. Tian B, Cui J, Lu H, Zuo Z, Zong Q. Adaptive finite-time attitude tracking of quadrotors with experiments and comparisons. IEEE Trans Ind Electron. 2019;66(12):9428–38.
8. Wang H, Li Z, Xiong H, Nian X. Robust H_∞ attitude tracking control of a quadrotor UAV on SO (3) via variation-based linearization and interval matrix approach. ISA Trans. 2019;87:10–6.
9. Chen M, Tao G, Jiang B. Dynamic surface control using neural networks for a class of uncertain nonlinear systems with input saturation. IEEE Trans Neural Netw Learn Syst. 2015;26(9):2086–97.
10. Puga-Guzmá S, Moreno-Valenzuela J, Santibáñez V. Neural controller for the trajectory tracking control of an inertia wheel pendulum. Revista Internacional De Métodos Numéricos Para Cálculo Y Diseño En Ingeniería. 2016;32(4):204–11.
11. Li J, Sun H, Sun Q. Neural network model predictive control applied to parafoil under turbulence wind field. Revista Internacional De Métodos Numéricos Para Cálculo Y Diseño En Ingeniería. 2025;41(3):51.
12. Yu G, Reis J, Silvestre C. Quadrotor neural network adaptive control: design and experimental validation. IEEE Robot Autom Lett. 2023;8(5):2574–81.
13. Song Y, He L, Zhang D, Qian J, Fu J. Neuroadaptive fault-tolerant control of QUAV: a more affordable solution. IEEE Trans Neural Netw Learn Syst. 2018;30(7):1975–83.
14. Zhang Y, Su X, Liu Z, Chen CLP. Event-triggered adaptive fuzzy tracking control with guaranteed transient performance for MIMO nonlinear uncertain systems. IEEE Trans Cybern. 2021;51(2):736–49.
15. Zhang J. Adaptive fuzzy global sliding mode control for trajectory tracking of quadrotor UAVs. Nonlinear Dyn. 2019;97:609–27.
16. Jiang F, Pourpanah F, Hao Q. Design, implementation, and evaluation of a neural-network-based quad-copter UAV system. IEEE Trans Ind Electron. 2019;67(3):2076–85.
17. Guo Z, Ren H, Li H, Zhou Q. Adaptive-critic-based event-triggered intelligent cooperative control for a class of second-order constrained multiagent systems. IEEE Trans Artif Intell. 2023;4(6):1654–65.
18. Wu H, Xu D, Jayawardhana B. On self-learning mechanism for the output regulation of second-order affine nonlinear systems. IEEE Trans Automat Contr. 2022;67(11):5964–79.
19. Wu D, Zhang W, Du H. Robust adaptive finite-time trajectory tracking control of a quadrotor aircraft. Int J Robust Nonlinear Control. 2021;31:8030–54.
20. Mu C, Zhang Y. Learning-based robust tracking control of quadrotor with time-varying and coupling uncertainties. IEEE Trans Neural Netw Learn Syst. 2020;31(1):259–73.
21. Liu H. Reinforcement learning-based tracking control for a quadrotor unmanned aerial vehicle under external disturbances. Int J Robust Nonlinear Control. 2023;33(17):10360–77.
22. Hu J, Zhang H, Song L, Han Z, Poor HV. Reinforcement learning for a cellular internet of UAVs: protocol design, trajectory control, and resource management. IEEE Wirel Commun. 2020;27(1):116–23.
23. Jin P, Ma Q, Xu S. Dynamic event-triggered robust optimal attitude control of QUAV using reinforcement learning. IEEE Trans Aerosp Electron Syst. 2023;59(6):7798–807.
24. Postoyan R, Tabuada P, Nešić D, Anta A. A framework for the event-triggered stabilization of nonlinear systems. IEEE Trans Automat Contr. 2015;60(4):982–96.

25. Liu J, Wang Q-G, Yu J. Event-triggered adaptive neural network tracking control for uncertain systems with unknown input saturation based on command filters. IEEE Trans Neural Netw Learn Syst. 2024;35(6):8702–7.
26. Xing J, Wen C, Liu Z, Su H, Cai J. Event-triggered adaptive control for a class of uncertain nonlinear systems. IEEE Trans Automat Contr. 2017;62(4):2071–6.
27. Sun Y. Adaptive attitude control for QUAVs using event-triggered control and optimized quantized communication with prescribed performance. Fluctuation Noise Lett. 2024;23(4):2450036.
28. Tian B, Cui J, Lu H, Liu L, Zong Q. Attitude control of UAVs based on event-triggered supertwisting algorithm. IEEE Trans Ind Inform. 2021;17(2):1029–38.
29. Zhu C, Chen J, Iwasaki M, Zhang H. Event-triggered deep learning control of quadrotors for trajectory tracking. IEEE Trans Ind Electron. 2024;71(3):2726–36.
30. Yu D, Ma S, Wang Z, Liu Y, Chen CLP. Fixed-time attitude control for QUAV with stochastic disturbances: theory and experiment. IEEE Trans Ind Electron. 2024;71(12):16484–92.
31. Qi G, Hu J, Li L, Li K. Integral compensation function observer and its application to disturbance-rejection control of QUAV attitude. IEEE Trans Cybern. 2024;54(7):4088–99.
32. Bi H, Qi G, Li X. Characteristic analyzes, experimental testing and control for attitude system of QUAV under disturbance. Appl Math Model. 2021;100:77–91.
33. Yang Y, Gorbachev S, Zhao B, Liu Q, Shu Z, Yue D. Predictor-based neural attitude control of a quadrotor with disturbances. IEEE Trans Ind Inform. 2024;24(1):169–78.
34. Chen W-H, Yang J, Guo L, Li S. Disturbance-observer-based control and related methods: an overview. IEEE Trans Ind Electron. 2016;63(2):1083–95.
35. Sariyildiz E, Oboe R, Ohnishi K. Disturbance observer-based robust control and its applications: 35th anniversary overview. IEEE Trans Ind Electron. 2019;67(3):2042–53.
36. Yang J, Li S, Yu X. Sliding-mode control for systems with mismatched uncertainties via a disturbance observer. IEEE Trans Ind Electron. 2012;60(1):160–9.
37. Tao X, Jiang J. Robust adaptive fault tolerant control for nonlinear systems with actuator failure and mismatched disturbance. Revista Internacional De Métodos Numéricos Para Cálculo Y Diseño En Ingeniería. 2024;40(1):8.
38. Kayacana E, Peschelb J, Chowdharya G. A self-learning disturbance observer for nonlinear systems in feedback-error learning scheme. Eng Appl Artif Intell. 2017;62:276–85.
39. Shao S, Yan X, Chen M, An Z. Event-triggered robust constrained control of uncertain nonlinear systems with input saturation based on self-learning disturbance observer. IEEE Trans Instrum Meas. 2024;73:1–15.
40. Cui G, Yang W, Yu J, Li Z, Tao C. Fixed-time prescribed performance adaptive trajectory tracking control for a QUAV. IEEE Trans Circ Syst II Express Briefs. 2022;69(2):494–8.
41. Sun K, Qiu J, Karimi H, Fu Y. Event-triggered robust fuzzy adaptive finite-time control of nonlinear systems with prescribed performance. IEEE Trans Fuzzy Syst. 2021;29(6):1460–71.