

Article

Approximate Optimal Tracking Control for Partially Unknown Nonlinear Systems via an Adaptive Fixed-Time Observer

Yanping Gao * and Zuojun Liu 

School of Artificial Intelligence, Hebei University of Technology, Tianjin 300401, China; liuzuojun@hebut.edu.cn

* Correspondence: zbkz_gao@163.com

Abstract: This paper investigates a novel adaptive fixed-time disturbance observer (AFXDO)-based approximate optimal tracking control architecture for nonlinear systems with partially unknown dynamic drift and perturbation under an adaptive dynamic programming (ADP) scheme. To attenuate the impact of disturbance, a novel AFXDO was designed based on the principle of a fixed-time stable system without prior information of disturbance, making disturbance observer errors converge to zero in a fixed time independent of initial estimation error. Additionally, approximate optimal control is conducted by incorporating the real-time estimation of AFXDO into a critic-only ADP framework to stabilize the dynamics of tracking errors and strike a balance between consumption and performance. In particular, to address the heavy calculation burden and oscillation phenomenon in the traditional actor-critic structure, an improved adaptive update law with a variable learning rate was developed to update the weight for adjusting the optimal cost function and optimal control policy simultaneously, avoiding the initial chattering phenomenon and achieving a prescribed convergence without resorting to dual networks. With the efforts of AFXDO and a weight law with a variable learning rate, the track errors were achieved with fast transient performance and low control consumptions in a fixed time. By revisiting Lyapunov stability, the tracking error and weight estimation error were proven to be uniformly ultimately bounded, and the designed control tended to optimal control. The simulations were carried out on quadrotor tracking to demonstrate the effectiveness of the developed control scheme, which achieves rapid convergence by lower control consumption in 4 s, where the cost function is reduced by 19.13%.

Keywords: disturbance attenuation; fixed-time disturbance observer; partially unknown nonlinear systems; critic-only; adaptive dynamic programming



Citation: Gao, Y.; Liu, Z.

Approximate Optimal Tracking Control for Partially Unknown Nonlinear Systems via an Adaptive Fixed-Time Observer. *Symmetry* **2023**, *15*, 1136. <https://doi.org/10.3390/sym15061136>

Academic Editor: Alexander Zaslavski

Received: 27 April 2023

Revised: 17 May 2023

Accepted: 20 May 2023

Published: 23 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In engineering applications, tracking control for nonlinear systems has been treated as an ever-increasing hot research topic over the last few years by virtue of its extensive application scenarios including the robot [1], permanent magnet synchronous motor [2], unmanned surface vehicle [3], and quadrotor [4]. Due to the inevitable presence of unknown system dynamics or uncertain parameters with the property of symmetry or asymmetry, as well as the external disturbances induced by working scenarios, the investigated model with uncertain system dynamic drift and extraneous disturbances widely exists. Conventional proportional integral control can hardly suppress the disturbance and dynamic drift to guarantee a robust tracking performance. To date, existing control methods to enhance the tracking performance for nonlinear systems can be roughly classified as sliding mode control [5], adaptive control [6,7], etc. The neural network and fuzzy logic system are two tools to make decisions for partially unknown systems, and their effectiveness can be verified in many areas, such as [6,8–10]. There have been many studies devoted to improving the estimation performance of neural networks and fuzzy logic systems [11–19]. However, the heavy computation burden caused by FLS and NN with adjusting parameters is impractical in real-world applications.

Aside from the classical robust control to handle partially unknown nonlinear systems, many scholars have shed light on disturbance observer (DO) for nonlinear uncertain systems on account of its concise form and validity. Its principal advantage is that DO-based controllers not only significantly improve the anti-interference ability but also retain the desired control performance among theories and practical engineering applications [20–23]. For example, approximate tracking control is put forward with neural network (NN) and nonlinear DO to approximate the system uncertainties [20]. In [22], a sliding mode control was designed for spacecraft to regulate attitude, where a novel integral DO with an extra integral term was inserted to estimate the lumped disturbances with high precision. In [23], an improvement of the reduced-order proportional integral observer was considered for feedforward compensation to avoid the effect caused by disturbance. Although there emerge various DOs focusing on disturbance attenuation, the convergence results of observation errors can only be assured when the time tends to infinity, which is uneconomic for real-world plants in engineering applications.

Therefore, a novel disturbance observer is developed, while an optimal tracking controller is proposed with a variable learning rate weight law. The investigated framework can achieve high precision and the rapid convergence of tracking performance with lower control consumption by designing the following goals:

- An adaptive fixed-time disturbance observer (AFXDO) is put forward for the estimation of lumped disturbance;
- The controller is set up based on the proposed AFXDO;
- The weight law with a variable learning rate is proposed to avoid the oscillation phenomenon during the learning process.

2. Related Works

Thus, aiming to achieve zero-disturbance within an effective short time and obtain fast tracking performance, finite-time stability with adjustment settling time has received considerable attention [24–28], which means the time from the initial state to the stable state is adjustable. In [24], a new super-twisting algorithm is provided for high-order systems with matched and mismatched disturbances, where a finite-time DO was designed to enable disturbance estimation errors to converge to zero in a finite time. For a second-order multiagent system, a centralized controller was developed in [26] based on the estimation of disturbance, where a finite-time observer was proposed for each agent to estimate disturbance. In [28], adaptive control was developed for rigid spacecraft to achieve attitude-tracking performance, where actuator faults are estimated by designing finite-time DO in the spirit of an improved finite-time stable system. Note that finite-time DOs are preferable to ensure higher accuracy and stronger robustness than conventional DO with asymptotic convergence. However, the settling time of finite-time DO is closely related to initial estimation error, which cannot be estimated without the known initial circumstance. Moreover, the settling time will tend to infinity when the initial estimation is far from the actual perturbation, which promotes further extension of finite-time stability, named the fixed-time technique, explored by Polykov [29]. In terms of disturbance attenuation, fixed-time DO is featured with fixed-time stability, i.e., the settling time of estimation error is dominated chiefly by control parameters rather than the initial estimation error [30–32]. By importing a fixed-time disturbance observer, a novel composite control consisting of a fixed-time controller and an anti-saturation controller was developed in [30] for an uncertainties model with actuator saturation to guarantee trajectory tracking and attitude regulation. To enhance the robustness of disturbance, tracking control for velocity and attitude was conducted by non-smooth backstepping control and fixed-time DO to compensate for the lumped disturbances [31]. In [32], a fixed-time DO was developed to approximate lumped perturbations by utilizing the proposed new fast fixed-time stable system. Note that the published outcomes related to fixed-time stability are subject to the following defects: One is that the observer gains are always selected and associated with the upper bound of disturbance and its derivative to assure stability. Another is that it is inferior in handling

performance optimization. Thus, in the absence of an upper bound of disturbance and its derivative, how to develop a novel control policy capable of performance optimization and fixed-time disturbance attenuation deserves deep investigation.

Reinforcement learning (RL) provides an efficient way to derive performance optimization by pursuing the maximum cost function, which offers a numerical method for solving optimal control. In the control field, it is equivalent to approximate dynamic programming (ADP). Its core is to solve the Hamilton–Jacobi–Bellman (HJB) equation for approximating optimal policy and corresponding optimal cost function via NN approximators, avoiding the curse of dimensionality existing in solving the optimal control of nonlinear systems. In particular, the optimal cost function and optimal control policy are estimated with an actor–critic dual networks framework. For example, the nonlinear control is designed for a nonlinear perturbation system, where the worst-case disturbance is attenuated by incorporating zero-sum differential game theory into the ADP scheme [33]. Note that the worst case is considered in control with over-conservative properties. To achieve disturbance attenuation, there emerges a wide range of results related to the ADP technique. For instance, the identifier–critic–actor architecture is developed in [34] with a fuzzy logic system to derive optimal control for a nonlinear system. The fuzzy logic model is well-identified, accompanied by trial and error to tune parameters heuristically. In [35], the optimal control is designed within an actor–critic framework, where the extra neural network is employed to identify the unmodeled dynamic, incurring a heavy computation burden. Thus, some scholars have focused on critic-only ADP with less computation complexity. The critic NN is adopted to update optimal cost function and optimal policy for nonlinear known systems simultaneously [36]. Subsequently, a series of ADP results are reported for partially known nonlinear systems, where the robustness of a closed-loop system can be ensured by game theory [37,38], adding robust terms [39], sliding mode control [40], or DO [41]. Specifically, DO-based ADP is a mainstream solution to achieve robustness and optimal performance [42–45]. On the condition that the derivative of disturbance is close to zero, the optimal tracking control is considered in [42] by integrating improved DO with dual network ADP. In [43], by introducing finite-time DO with known upper bound information, an optimal sliding mode controller is designed based on actor–critic ADP. In [44], by combining with DO, the actor–critic ADP is employed to derive the optimal control of sliding-mode dynamics. The steady control input is put forward based on novel DO, and optimal control is derived for optimal regulation problems in a single-network ADP in [45].

It would be worth mentioning that the above-involved DOs were developed under the assumption that the upper bound of disturbance is available or its derivative tends to zero, which are troublesome and unsuitable for dealing with various kinds of disturbances in practical implementations. In addition, only the asymptotic or finite-time convergence of disturbance estimation errors can be achieved in the existing whole ADP control scheme, and hysteretic disturbance estimation may induce poor tracking performance. Note that fixed-time DO can enable disturbance estimation errors to converge to zero in a preassigned time in spite of initial estimation errors, but previously reported fixed-time DOs are established on the condition that the upper bound of disturbance is available. It is a challenge to estimate lumped disturbances for identifying a system without the assumption of bounded information on lumped disturbance. Moreover, the critic NN is conducted to approximate the optimal cost function and optimal policy to derive optimal control without actor NN.

Motivated by the previous discussion, a novel adaptive fixed-time disturbance observer (AFXDO)-based approximate optimal tracking control architecture for nonlinear partially unknown systems is considered, where AFXDO is designed such that the estimation of perturbation can be accomplished within a fixed time. Then, approximate optimal control is conducted by incorporating the real-time estimation of AFXDO into a critic-only ADP framework, where weight law with variable learning rate is embedded in a single network. The outstanding features of the devised scheme are:

(1) Compared with the presenting DO, the stability of perturbation estimation error can be guaranteed with finite-time convergence [24–28] or asymptotic convergence [21–23],

and the investigated AFXDO can achieve the fixed-time stability of the estimation dynamic with the core of fixed-time stability system, ensuring that estimation error with fixed-time convergence with one power term. In addition, in view of the mentioned disturbance observer with a strong assumption about perturbation, the adaptive rule is employed to adjust observer gain, removing the rigid hypothesis on the known boundedness of perturbation [21–24].

(2) In contrast to the existing ADP framework with the actor–critic framework [42–44], the proposed ADP is a much simpler structure without using actor NN to estimate optimal policy. Different from the critic-only ADP framework with known upper-bounded information where the Bellman error is minimized by gradient descent [39,40], herein, by adaptive adjusting observer gain to estimate the perturbation, a controller is set up to eliminate the impact of perturbation, while the optimal controller is proposed with critic-only ADP with the weight law driven by extracted weight error conducted via filter operation, where the variable learning rate is utilized to fulfill faster weight convergence without initial chattering phenomenon. A brief comparison between current controllers and the proposed scheme is summarized in Table 1.

Table 1. Comparison between the existing controllers and proposed scheme.

Literature	Controller Design				
	Anti-Interference Manner	Assumptions	Convergence Results	Optimal Manner	Weight Update
[22]	Integral DO	Known	Asymptotic convergence	Unconsidered	Unconsidered
[24–28]	Finite-time DO	Known	Finite-time convergence	Unconsidered	Unconsidered
[30–32]	Fixed-time DO	Known	Fixed-time convergence	Unconsidered	Unconsidered
[39]	Value function	Known	Unconsidered	Critic-only	Gradient Descent
[40]	Sliding mode control	Known	Unconsidered	Critic-only	Gradient Descent
[42,44]	DO	Known	Asymptotic convergence	Actor–critic	Gradient Descent
[43]	Finite-time DO	Known	Finite-time Convergence	Actor–critic	Gradient Descent
Our method	Fixed-time DO	Unknown	Fixed-time convergence	Critic-only	Improved gradient

The remainder of this research is organized as follows: The problem formulation and its preliminaries are described in Section 3. In Section 4, the novel fixed-time observer is designed, while the optimal control policy is put forward with improved adaptive law and the stability analysis of improved adaptive law and error system are discussed. The simulation cases are carried out in Section 5. In Section 6, we present some conclusions.

3. Problem Formulation and Preliminaries

3.1. Problem Formulation

The nonlinear multi-input multi-output (MIMO) system is described with uncertain dynamics and perturbation as follows:

$$\dot{x} = f(x) + g(x)u + d(t) \quad (1)$$

where $x \in \mathbb{R}^n$ denotes the state vector and $u \in \mathbb{R}^m$ is the control input. $f(x) \in \mathbb{R}^n$ and $g(x) \in \mathbb{R}^{n \times m}$ indicate the dynamic drift and control gain matrix. $f(x)$ is assumed to be an uncertain smooth and nonlinear function, which can be decomposed as $f(x) = f_0(x) + \Delta f(x)$ with known $f_0(x)$ and bounded uncertainty $\Delta f(x)$. The time-varying disturbance $d(t) \in \mathbb{R}^n$ represents bounded perturbation. The desired tracking trajectory x_d and its first derivate $\dot{x}_d$ are supposed to be bounded, which can conduce the system being controllable on a compact set [46].

To fulfill optimal tracking control for system (1) with reference command, the tracking error is defined as $e = x - x_d = [e_1, e_2, \dots, e_n]^T$. Thus, the tracking error of the system can be obtained by

$$\dot{e} = \dot{x} - \dot{x}_d = f_0(x) + g(x)u + \Delta f(x) + d(t) - \dot{x}_d \quad (2)$$

For the purpose of deriving the tracking error accurately, the controller is decomposed into two parts: the feedforward controller and the optimal tracking controller $u = u_f + u_q$. The feedforward controller $u_f = [g^T g]^{-1} g^T (\dot{x}_d - f_0 - Ke)$ with a positive-definite diagonal matrix $K \in \mathbb{R}^{n \times n}$ is developed by compensating the reference command and known dynamic drift $f_0(x)$ to reduce the region of tracking error. Thus, by substituting u_f into (2), the dynamic of the tracking error can be converted into

$$\dot{e} = -Ke + g(x)u_q + \Delta f(x) + d(t) \quad (3)$$

The control objective of this paper is to develop the optimal control policy u_q for a nonlinear system in despite of perturbation and uncertain mode uniformly, such that the following goals are achieved:

- (1) The disturbance estimation error can converge to zero in a fixed time.
- (2) The trajectory tracking error and weight estimation error are uniformly ultimately bounded (UUB).
- (3) The designed control input approximates the optimal control policy.

Remark 1. In essence, there exist numerous nonlinear systems with disturbance and uncertainty, which can be expressed in the form of nonlinear MIMO systems (1) or high-order versions, such as robots [1], quadrotors [4], and MEMS [47], where assumptions of the boundedness of the uncertainty and disturbance are acceptable. For the sake of analysis, we mainly focus on (1) as a first-order system, although the high-order system can be investigated via the proposed controller scheme incorporated with the backstepping technique. Different from the vast controllers explored in former works [22–32], where controllers lean on Lyapunov stability in the traditional framework, the investigated controller is based on the ADP scheme. Noting the uncertain term and disturbance in the system (3), the optimal tracking control will be clarified with AFXDO.

3.2. Related Lemmas

Lemma 1 ([32]). Suppose that there exists a positive function V satisfying

$$\dot{V}(t) \leq -\ell V^\chi(t) - \iota V^\alpha(t) \quad (4)$$

where $\chi = \frac{\alpha}{2} + \frac{\beta}{2} + \left(\frac{\alpha}{2} - \frac{\beta}{2}\right) \text{sign}(V(t) - 1)$ with $0 < \beta < 1$ and $\alpha > 1$ and positive constants ℓ and ι , then $V(t)$ can converge to the origin before the settling time:

$$T = \frac{1}{\ell(1-\beta)} + \frac{1}{(\ell + \iota)(\alpha - 1)} \quad (5)$$

Lemma 2 ([48]). For positive number ξ and positive integer K , there holds

$$\sum_{k=1}^K \xi^m \geq \left(\sum_{k=1}^K \xi \right)^m, \text{ for } 0 < m \leq 1 \quad (6)$$

$$\sum_{k=1}^K \xi^m \geq K^{1-m} \left(\sum_{k=1}^K \xi \right)^m, \text{ for } m > 1 \quad (7)$$

Lemma 3 ([49]). For positive constants a and b with $p > 1$ satisfying $\frac{1}{p} + \frac{1}{q} = 1$, there holds

$$ab \leq \frac{\tau a^p}{p} + \frac{b^q}{\tau q} \quad (8)$$

with positive τ .

4. Main Results

In this section, the optimal tracking control based on AFXDO is developed to fulfill the specified goals in the above section. For the controller u_q , the novel fixed-time observer is ingeniously investigated to estimate the synthesized unknown term with adaptive observer gains. By virtue of the designed observer, the optimal control is derived with a critic-only ADP scheme, where an improved adaptive law with a variable learning rate is established for weight update. The flowchart of the proposed controller can be described in Figure 1.

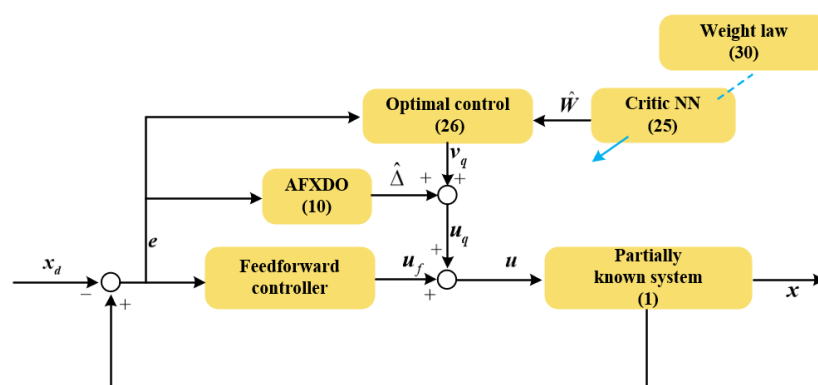


Figure 1. The flowchart of the proposed controller.

4.1. Adaptive Fixed-Time Disturbance Observer

Inspired by [20], a novel AFXDO is developed without prior information of disturbance, ensuring that the estimation error converges to zero in a fixed time by virtue of adaptive observer gains. Noting that the uncertain dynamic drift and external disturbance are unknown, the AFXDO is conducted to estimate the synthesized unknown term as lumped disturbances:

$$\Delta = \Delta f(x) + d(t) \quad (9)$$

Thus, the system (3) is transformed into $\dot{e} = g(x)u_q + \Delta = G(x, u_q) = [G_1, G_2, \dots, G_n]^T$, and the AFXDO is devised for each component of tracking error and lumped disturbances as

$$\begin{cases} \dot{\hat{e}}_i = k_1 \tilde{e}_i^\chi \text{sign}(\tilde{e}_i) + G_i + \hat{\Delta}_i + \lambda_1 \tilde{h}_i \\ \dot{\hat{\Delta}}_i = k_2 \tilde{h}_i^\chi \text{sign}(\tilde{h}_i) + \lambda_2 \text{sign}(\tilde{h}_i) \end{cases} \quad (10)$$

where $\hat{e} = [\hat{e}_1, \hat{e}_2, \dots, \hat{e}_n]^T$ and $\hat{\Delta} = [\hat{\Delta}_1, \hat{\Delta}_2, \dots, \hat{\Delta}_n]^T$ represent the estimation of tracking error e and synthesized uncertainties Δ , $\tilde{e}_i = e_i - \hat{e}_i$, $\tilde{h}_i = r \text{sign}(\tilde{e}_i)$ for $i = 1, 2, \dots, n$ with positive constant r . The parameters k_1 and k_2 are positive parameters, $\chi = \frac{\alpha}{2} + \frac{\beta}{2} + \left(\frac{\alpha}{2} - \frac{\beta}{2}\right) \text{sign}(|\tilde{e}_i| - 1)$ with $\alpha > 1$ and $0 < \beta < \frac{1}{2}$. The observer gains λ_1 and λ_2 are adapted from following rules:

$$\begin{cases} \dot{\lambda}_1 = l_1 |\tilde{e}_i|^{1-2\beta} + \mu_1 \\ \dot{\lambda}_2 = l_2 |(\tilde{h}_i)_{eq}|^{1-2\beta} + \mu_2 \end{cases} \quad (11)$$

with positive parameters l_1, l_2, μ_1 , and μ_2 to be chosen and $(\tilde{h}_i)_{eq} = r \text{sign}(\tilde{e}_i)$. Theorem 1 is given for disturbance estimation error, which is proved in Appendix A.

Theorem 1. *Considering the system (1) and novel AFXDO (10) with adaptive gain update (11), the disturbance estimation error can converge to zero in a fixed time.*

In view of the tracking error system (3), the feedback controller is set as follows:

$$u_q = -[g^T g]^{-1} g^T \hat{\Delta} + v_q \quad (12)$$

where $-[g^T g]^{-1} g^T \hat{\Delta}$ is employed to attenuate the impression of uncertainty and the extra term is the optimal control policy of the following system:

$$\dot{e} = -Ke + g(e)v_q + \tilde{\Delta} \quad (13)$$

Remark 2. *The settling time is only related to the control parameters and does not vary along with the initial condition, which dominates superiority among other DOs with finite-time convergence [24–28] and asymptotic convergence [20–23]. In addition, there is one power term in the observer (10) rather than existing fixed-time DOs with two power terms [18–20] to achieve fixed-time convergence. Note that the fixed-time stable system in [32] can be shown in the following equation as*

$$\dot{z} = -(\gamma_1 |z|^\alpha + \gamma_2 |z|^\beta) \text{sign}(z(t)) \quad (14)$$

with positive constants γ_1 and γ_2 . Based on Lemma 1, we can derive settling time as

$$T = \frac{1}{\gamma_1(\alpha - 1)} + \frac{1}{(\gamma_1 + \gamma_2)(1 - \beta)} \quad (15)$$

The improved version of (14) consisting of only the power term can be represented with positive constant γ_1' as

$$\dot{z} = -\gamma_1' |z|^\chi \text{sign}(z(t)) \quad (16)$$

which is the core of the designed observer. Moreover, the adaptive rule (11), which is simple and easily operated, can enable sufficiently large increases in λ_1 and λ_2 , removing the strict assumptions on perturbations, such as the known bounded information of the perturbation or its derivative and invariant perturbation [42,43]. Noting that there exist many parameters involved in the AFXDO, we summarize the fundamental guidelines for adjusting the observer performance as follows: The parameters k_1 , k_2 , λ_1 , λ_2 , and exponential χ related to α, β are key factors to guarantee the settling time and accuracy. Noting that the k_1 is the identical gain for $\tilde{e}_i > 1$ and $\tilde{e}_i < 1$, we should make a compromise on the convergence rate and accuracy. In addition, the adaptive rules (11) are adopted for increasing gains λ_1 and λ_2 with adequate parameters μ_1 , μ_2 , l_1 , l_2 for observer stability.

Remark 3. *Existing optimal controls based on the ADP scheme mainly concentrate on the optimal regulation problem, which is challenging for designing optimal tracking control, especially for uncertainty systems as well as nonvanishing tracking trajectories. The feedforward controller is conducted to compensate for the nonvanishing tracking trajectory and known drift dynamic. Thus, the optimal trajectory controller for system (1) is converted to an optimal regulation problem (3). For addressing the uncertainty, a novel AFXDO is ingeniously conducted for the estimation of lumped disturbances, where the disturbance estimation errors can converge to zero in a fixed time and observer gains are tuned adaptively without an assumption of the known upper bound of Δ .*

4.2. Optimal Tracking Control Scheme

The optimal control is considered to stabilize the system (13). Unlike the traditional controller in [29–32], there is only taking tracking performance into consideration without further exploring optimal control. A cost function is denoted related to tracking error and control input v_q as

$$V(e) = \int_t^\infty c(e, v_q) d\varsigma \quad (17)$$

where $c(e, v_q) = e^T Q e + v_q^T R v_q$ with positive-definite diagonal matrices $Q \in \mathbb{R}^{n \times n}$ and $R \in \mathbb{R}^{m \times m}$. One can denote the optimal cost function:

$$V^*(e) = \min_{v_q \in \Psi(\Omega)} \int_t^\infty c(e, v_q^*) d\zeta \quad (18)$$

Taking the time derivative of (18) along with (13), one can obtain the HJB equation as

$$H(e, v_q^*, V^*) = V_e^{*T} [-K e + g v_q^* + \tilde{\Delta}] + e^T Q e + v_q^{*T} R v_q^* = 0 \quad (19)$$

where $V_e^* = \frac{\partial V^*(e)}{\partial e}$. According to (19), optimal control v_q^* can be conducted by $\frac{\partial H(e, v_q^*, V^*)}{\partial v_q^*} = 0$ as

$$v_q^* = -\frac{1}{2} R^{-1} g^T V_e^* \quad (20)$$

Noting that the optimal control v_q^* (20) depends on the gradient of the cost function, i.e., V_e^* , by substituting (20) into (19), the HJB Equation (19) can be represented as

$$V_e^{*T} [-K e + \tilde{\Delta}] + e^T Q e - \frac{1}{4} V_e^{*T} g R g^T V_e^* = 0 \quad (21)$$

However, it is difficult and even impossible to achieve the analytical solution due to the nonlinearity of the HJB equation. For addressing this issue, an NN approximator named critic-only NN is introduced to approximate the optimal cost function $V^*(e)$, and the optimal control policy can be achieved immediately rather than depending on another NN.

Specifically, the optimal cost function $V^*(e)$ can be approximated by a single-layer NN, which can be expressed as

$$V^*(e) = W^T \sigma(e) + \varepsilon \quad (22)$$

where $W \in \mathbb{R}^p$ represents the ideal weight of the NN, with p indicating the number of neurons, $\sigma(e) \in \mathbb{R}^p$ being the corresponding activation function, and ε being the approximate error induced by the critic NN. From (22), the first derivative of $V^*(e)$ with respect to tracking error can be obtained as

$$\frac{\partial V^*(e)}{\partial e} = \nabla \sigma(e)^T W + \nabla \varepsilon \quad (23)$$

where $\nabla \sigma(e) = \frac{\partial \sigma(e)}{\partial e}$ and $\nabla \varepsilon = \frac{\partial \varepsilon}{\partial e}$ are the gradient of the activation function and approximation error with upper boundness $\bar{\sigma}$ and $\bar{\varepsilon}$ satisfying $\|\nabla \sigma\| \leq \bar{\sigma}$ and $\|\nabla \varepsilon\| \leq \bar{\varepsilon}$. Thus, the optimal control v_q^* can be given as:

$$v_q^* = -\frac{1}{2} R^{-1} g^T (\nabla \sigma(e)^T W + \nabla \varepsilon) \quad (24)$$

Since the ideal weights that can accurately estimate the cost function are unavailable, its approximation can be estimated by estimation weight and corresponding regressors:

$$\hat{V}(e) = \hat{W}^T \sigma(e) \quad (25)$$

with $\hat{W}$ representing an estimation of the ideal weight. The actual approximated optimal policy is induced by

$$v_q = -\frac{1}{2} R^{-1} g^T \nabla \sigma(e)^T \hat{W} \quad (26)$$

Then approximation HJB Equation (19) with (26) can be expressed as

$$\mathbf{W}^T \nabla \sigma(-\mathbf{K}\mathbf{e} + \mathbf{g}\mathbf{v}_q) + \mathbf{e}^T \mathbf{Q}\mathbf{e} + \mathbf{v}_q^T \mathbf{R}\mathbf{v}_q + \varepsilon_{HJB} = 0 \quad (27)$$

with ε_{HJB} being the HJB error associated with NN approximation error and disturbance error and denoting $\varepsilon_{HJB} = \nabla \varepsilon^T(-\mathbf{K}\mathbf{e} + \mathbf{g}\mathbf{v}_q + \tilde{\Delta}) + (\mathbf{W}^T \nabla \sigma)\tilde{\Delta}$. With sufficiently large p and the fixed-time convergence of AFXDO, it can be guaranteed that $\nabla \varepsilon$ and $\tilde{\Delta}$ converge to zero.

Then, (27) can be simplified as a linearization form:

$$\Psi = -\Phi^T \mathbf{W} - \varepsilon_{HJB} \quad (28)$$

with $\Psi = \mathbf{e}^T \mathbf{Q}\mathbf{e} + \mathbf{v}_q^T \mathbf{R}\mathbf{v}_q$ and $\Phi = \nabla \sigma(-\mathbf{K}\mathbf{e} + \mathbf{g}\mathbf{v}_q)$. There have been many update laws for online updating weight, such as gradient descent and least squares methods [39,40] with fixed learning rates. Different from the Bellman error driving weight update in the inherent ADP schemes, the weight errors can be extracted by auxiliary operation, leading to the fact that weight errors can converge to the neighborhood of ideal weight and further enhance weight estimation performance—specifically, by introducing auxiliary matrix $\mathbf{M} \in \mathbb{R}^{p \times p}$ and vector $\mathbf{N} \in \mathbb{R}^p$, which are defined as:

$$\begin{cases} \dot{\mathbf{M}} = -c\mathbf{M} + \Phi\Phi^T, & \mathbf{M}(0) = 0 \\ \dot{\mathbf{N}} = -c\mathbf{N} + \Phi\Psi, & \mathbf{N}(0) = 0 \end{cases} \quad (29)$$

with designed constant $c > 0$. Denoting $\mathfrak{S} = \mathbf{M}\hat{\mathbf{W}} + \mathbf{N}$, the estimation rule of weight in [50] can be employed for weight update as

$$\dot{\hat{\mathbf{W}}} = -\mathbf{Y}\mathfrak{S}, \quad \mathbf{Y} = \mathbf{Y}_0 e^{-\frac{\|\mathbf{e}\|^2}{\delta^2}} \quad (30)$$

with positive-definite diagonal matrices $\mathbf{Y} \in \mathbb{R}^{p \times p}$ and $\mathbf{Y}_0 \in \mathbb{R}^{p \times p}$ and constant $\delta > 0$. The weight convergence can be guaranteed by following the theorem with ingenious derivation, in which there contains the weight error $\tilde{\mathbf{W}} = \mathbf{W} - \hat{\mathbf{W}}$ in (30).

Theorem 2. For the approximation of the cost function with critic NN (22), if the estimation cost function can be evolved by the estimation of weight with update rule (30), the weight errors can be driven into the neighborhood of zero.

Proof of Theorem 2. By integrating both sides of (29), the solution can be formulated as

$$\begin{cases} \mathbf{M}(t) = \int_0^t e^{-c(t-r)} \Phi\Phi^T dr, & \mathbf{M}(0) = 0 \\ \mathbf{N}(t) = \int_0^t e^{-c(t-r)} \Phi\Psi^T dr, & \mathbf{N}(0) = 0 \end{cases} \quad (31)$$

Combined with (45), one can derive that there exists an auxiliary variable $\mathbf{o}$ being associated to HJB error, i.e., $\mathbf{o} = -\int_0^t e^{-c(t-r)} \Phi \varepsilon_{HJB} dr$, which satisfies $\dot{\mathbf{N}} = -\mathbf{M}\mathbf{W} + \mathbf{o}$. Following (30), this can be a further representation of the update rule:

$$\dot{\mathfrak{S}} = \mathbf{M}\dot{\hat{\mathbf{W}}} + \dot{\mathbf{N}} = -\mathbf{M}\tilde{\mathbf{W}} + \mathbf{o} \quad (32)$$

where there exists a positive constant ε_o , such that $\|\mathbf{o}\| \leq \varepsilon_o$ holds. Obviously, the weight errors can be extracted via the update rule, leading to weight convergence. Under the condition that Φ satisfies the persistent excitation (PE) condition, the positive-definite matrix $\mathbf{M}$ can be derived from Lemma 1 in [50]. Consider the following Lyapunov function:

$$V_W = \frac{1}{2} \tilde{\mathbf{W}}^T \mathbf{Y}_0^{-1} \tilde{\mathbf{W}} \quad (33)$$

According to (30), one can derive the time derivative of V_W , such that:

$$\begin{aligned}\dot{V}_W &= \tilde{W}^T \mathbf{Y}_0^{-1} \dot{\tilde{W}} = e^{-\frac{\|e\|^2}{\delta^2}} \tilde{W}^T (-M\tilde{W} + o) \\ &\leq e^{-\frac{\|e\|^2}{\delta^2}} \left(-\eta \|\tilde{W}\|^2 + \|\tilde{W}\| \|o\| \right) \\ &\leq -e^{-\frac{\|e\|^2}{\delta^2}} \|\tilde{W}\| \left(\eta \|\tilde{W}\| - \varepsilon_o \right)\end{aligned}\quad (34)$$

with $\eta > 0$ denoting the minimized eigenvalue of the matrix, i.e., $\eta = \lambda_{\min}(M)$. By revisiting the Lyapunov theory, the weight errors can converge to the neighborhood of origin. $\square$

Remark 4. Contrary to most ADP schemes, where the Hamiltonian function is structured for the nominal system, the HJB equation is conducted related to (13), consisting of the estimation error, which leads to challenges for adjusting the optimal control policy and weight update. As opposed to the existing control within the ADP framework, where the weight update is carried out by minimizing the Bellman error with gradient descent [39,40], by imposing an auxiliary operation on intermediate parameters and employing the Gaussian function to tune the learning rate, the adaptive rule (30), which contains the weight errors, can compel the approximation of weight converge to an ideal weight. In particular, when the tracking error is far from the origin, $\mathbf{E}$ is decreased to avoid the initial oscillation of weight, which further affects the control input. If the tracking error is close to zero, larger $\mathbf{Y}$ can enable online learning ability with an acceleration convergence of weight, and the PE condition can be easily verified by transferring it into the positive-definite property of matrix $\mathbf{N}$, avoiding introducing the concurrent learning or the experience replay employed in most ADP frameworks.

4.3. Stability Analysis

Theorem 3. For the partially known nonlinear system (3) with uncertain parameters and unknown disturbance, by designing control (26) with weight updating laws (30), the following goals can be achieved.

- (1) The tracking error and weight error are UUB.
- (2) The approximate actual control input approximates the optimal control policy.

Proof of Theorem 3. According to (13) and (26), there exist dynamics of tracking errors such as:

$$\begin{aligned}\dot{e} &= -Ke + gv_q + \tilde{\Delta} \\ &= -Ke - \frac{1}{2}gR^{-1}g^T \nabla \sigma^T \tilde{W} + \frac{1}{2}gR^{-1}g^T (\nabla \sigma^T W + \nabla \varepsilon) + gv_q^* + \tilde{\Delta} \\ &= -Ke + \frac{1}{2}gR^{-1}g^T \nabla \sigma^T \tilde{W} + \frac{1}{2}gR^{-1}g^T \nabla \varepsilon + gv_q^* + \tilde{\Delta}\end{aligned}\quad (35)$$

Thus, the following Lyapunov function is constructed for proving the stability of tracking errors:

$$V = \frac{1}{2} \tilde{W}^T \mathbf{Y}_0^{-1} \tilde{W} + \Gamma_1 e^T e + \Gamma_2 V^* + \Gamma_3 o^T o \quad (36)$$

with optimal function V^* denoted by (18) and positive constants $\Gamma_1, \Gamma_2, \Gamma_3$. In the following proof, each individual part of (36) is concretely analyzed based on Lemma 3.

First, we restate the convergence of $\tilde{W}$ as follows:

$$\begin{aligned}\dot{V}_1 &= e^{-\frac{\|e\|^2}{\delta^2}} \tilde{W}^T (-M\tilde{W} + o) \\ &\leq -e^{-\frac{\|e\|^2}{\delta^2}} \left(\eta \|\tilde{W}\|^2 + \|\tilde{W}^T o\| \right) \\ &= -e^{-\frac{\|e\|^2}{\delta^2}} \left[\left(\eta - \frac{1}{2\tau\Gamma_3} \right) \|\tilde{W}\|^2 - \frac{\tau\Gamma_3 \|o\|^2}{2} \right]\end{aligned}\quad (37)$$

Thus, by substituting (35) into (36), it can be further reformulated in a simple form:

$$\begin{aligned}
 \dot{V}_2 &= 2\Gamma_1 e^T \dot{e} + \Gamma_2 \dot{V}^* \\
 &= 2\Gamma_1 e^T \dot{e} + \Gamma_2 (-e^T Q e - v_q^{*T} R v_q^*) \\
 &= 2\Gamma_1 e^T \left[-\lambda_{\min}(K) e + \frac{1}{2} g R^{-1} g^T \nabla \sigma^T \tilde{W} + \frac{1}{2} g R^{-1} g^T \nabla \varepsilon + g v_q^* + \tilde{\Delta} \right] \\
 &\quad + \Gamma_2 (-e^T Q e - v_q^{*T} R v_q^*) \\
 &\leq -[2\lambda_{\min}(K)\Gamma_1 + \Gamma_2 \lambda_{\min}(Q) - (\|g R^{-1} g^T \nabla \sigma^T\| + \|g R^{-1} g^T\| + 2)\Gamma_1] \|e\|^2 \\
 &\quad + \frac{1}{4}\Gamma_1 \|g R^{-1} g^T \nabla \sigma^T\| \|\tilde{W}\|^2 + \frac{1}{4}\Gamma_1 \|g R^{-1} g^T\| \|\nabla \varepsilon\|^2 + \Gamma_1 \|\tilde{\Delta}\|^2 \\
 &\quad - (\Gamma_2 \lambda_{\min}(R) - \Gamma_1 \|g\|^2) \|v_q^*\|^2
 \end{aligned} \tag{38}$$

It can be derived from (31) that $\dot{o} = -co + \Phi \varepsilon_{HJB}$, so that

$$\begin{aligned}
 \dot{V}_3 &= 2\Gamma_3 o^T \dot{o} \\
 &= 2\Gamma_3 o^T (-co + \Phi \varepsilon_{HJB}) \\
 &= 2\Gamma_3 o^T \left\{ -co + \Phi \left[(W^T \nabla \sigma + \nabla \varepsilon) \tilde{\Delta} + \nabla \varepsilon \left(-\frac{1}{2} g R^{-1} g^T \nabla \sigma^T \tilde{W} \right) - \nabla \varepsilon K e \right] \right\} \\
 &= -\Gamma_3 (2c - 3\tau) \|o\|^2 + \frac{\Gamma_3}{\tau} \|\Phi (W^T \nabla \sigma + \nabla \varepsilon)\|^2 \|\tilde{\Delta}\|^2 + \frac{\Gamma_3}{4\tau} \|\Phi \nabla \varepsilon g R^{-1} g^T \nabla \sigma^T \tilde{W}\|^2 \\
 &\quad + \frac{\Gamma_3}{\tau} \|\Phi \nabla \varepsilon K\|^2 \|e\|^2
 \end{aligned} \tag{39}$$

Therefore,

$$\begin{aligned}
 \dot{V} &= \dot{V}_1 + \dot{V}_2 + \dot{V}_3 \\
 &\leq - \left[e^{-\frac{\|e\|^2}{\delta^2}} \left(\eta - \frac{1}{2\tau\Gamma_3} \right) - \frac{1}{4}\Gamma_1 \|g R^{-1} g^T \nabla \sigma^T\| \right] \|\tilde{W}\|^2 - (K' \lambda_{\min}(R) - \Gamma' \|g\|^2) \|v_q^*\|^2 \\
 &\quad - \left[2\lambda_{\min}(K)\Gamma_1 + \Gamma_2 \lambda_{\min}(Q) - (\|g R^{-1} g^T \nabla \sigma^T\| + \|g R^{-1} g^T\| + 2)\Gamma_1 - \frac{\Gamma_3}{\tau} \|\Phi \nabla \varepsilon K\|^2 \right] \|e\|^2 \\
 &\quad - \Gamma_3 \left(2c + \left(\frac{1}{2} e^{-\frac{\|e\|^2}{\delta^2}} - 3 \right) \tau \right) \|o\|^2 + \frac{1}{4}\Gamma_1 \|g R^{-1} g^T\| \|\nabla \varepsilon\|^2 + \frac{\Gamma_3}{4\tau} \|\Phi \nabla \varepsilon g R^{-1} g^T \nabla \sigma^T \tilde{W}\|^2 \\
 &\quad + \left(\Gamma_1 + \frac{\Gamma_3}{\tau} \|\Phi (W^T \nabla \sigma + \nabla \varepsilon)\|^2 \right) \|\tilde{\Delta}\|^2
 \end{aligned} \tag{40}$$

As long as $\Gamma_2 \lambda_{\min}(R) - \Gamma_1 \|g\|^2 > 0$, (40) can be rewritten as

$$\dot{V} \leq -c_1 \|\tilde{W}\|^2 - c_2 \|e\|^2 - c_3 \|o\|^2 + \rho \tag{41}$$

with positive coefficients

$$c_1 = e^{-\frac{\|e\|^2}{\delta^2}} \left(\eta - \frac{1}{2\tau\Gamma_3} \right) - \frac{1}{4}\Gamma_1 \|g R^{-1} g^T \nabla \sigma^T\|$$

$$c_2 = 2\lambda_{\min}(K)\Gamma_1 + \Gamma_2 \lambda_{\min}(Q) - (\|g R^{-1} g^T \nabla \sigma^T\| + \|g R^{-1} g^T\| + 2)\Gamma_1 - \frac{\Gamma_3}{\tau} \|\Phi \nabla \varepsilon K\|^2$$

$$c_3 = \Gamma_3 \left(2c + \left(\frac{1}{2} e^{-\frac{\|e\|^2}{\delta^2}} - 3 \right) \tau \right)$$

$$\rho = \frac{1}{4}\Gamma_1 \|g R^{-1} g^T\| \|\nabla \varepsilon\|^2 + \frac{\Gamma_3}{4\tau} \|\Phi \nabla \varepsilon g R^{-1} g^T \nabla \sigma^T \tilde{W}\|^2 + \left(\Gamma_1 + \frac{\Gamma_3}{\tau} \|\Phi (W^T \nabla \sigma + \nabla \varepsilon)\|^2 \right) \|\tilde{\Delta}\|^2$$

Obviously, the upper bound of ρ can be easily guaranteed from the NN approximation errors and HJB errors, while the observer error $\tilde{\Delta}$ will converge to zero in a fixed time. If suitable parameters are selected to ensure the positiveness of c_1 , c_2 , and c_3 , it must satisfy the following conditions

$$\Gamma_1 < \frac{4\eta}{\|g R^{-1} g^T \nabla \sigma^T\|}$$

$$\begin{aligned}
\tau &> \max \left(\frac{1}{\Gamma_3 \left(2\eta - \frac{1}{2} \Gamma_2 \|gR^{-1}g^T \nabla \sigma^T\| \right)}, \frac{\Gamma_3 \|\Phi \nabla \varepsilon K\|^2}{2\lambda_{\min}(K)\Gamma_2} \right) \\
\Gamma_2 &> \max \left(\frac{-2\lambda_{\min}(K)\Gamma_1 + \frac{\Gamma_3}{\tau} \|\Phi \nabla \varepsilon K\|^2 - (\|gR^{-1}g^T \nabla \sigma^T\| + \|gR^{-1}g^T\| + 2)\Gamma_1}{\lambda_{\min}(Q)}, \frac{\Gamma_1 \|g\|^2}{\lambda_{\min}(R)} \right) \\
\lambda_{\min}(K) &> \frac{\|gR^{-1}g^T \nabla \sigma^T\| + \|gR^{-1}g^T\|}{2} + 1 \\
c &> \left(-\frac{1}{4} e^{-\frac{\|e\|^2}{\delta^2}} + \frac{3}{2} \right) \tau > \frac{5}{4} \tau, \Gamma_3 > 0.
\end{aligned}$$

Noting that critic NN errors can be bounded with finite constants, i.e., $\rho > 0$, it is obviously derived that $\dot{V} < 0$ as long as one of the following inequalities holds:

$$\|\tilde{W}\| > \sqrt{\frac{\rho}{c_1}}, \|e\| > \sqrt{\frac{\rho}{c_2}}, \|o\| > \sqrt{\frac{\rho}{c_3}}, \quad (42)$$

Furthermore, tracking error e and NN weights error $\tilde{W}$ are UUB. From (24) and (26), one can infer that

$$\begin{aligned}
\lim_{t \rightarrow \infty} \|v_q - v_q^*\| &= \lim_{t \rightarrow \infty} \left\| \frac{1}{2} R^{-1} g^T \nabla \sigma^T \tilde{W} + \frac{1}{2} R^{-1} g^T \nabla \varepsilon \right\| \\
&\leq \frac{1}{2} \|R^{-1} g^T\| (\bar{\sigma} \|\tilde{W}\| + \bar{\varepsilon}) \\
&\triangleq \varepsilon_u.
\end{aligned} \quad (43)$$

Especially, if the critic NN error is zero, i.e., $\nabla \varepsilon = 0$, we can obtain $\rho = 0$, resulting in

$$\dot{V} = -c_1 \|\tilde{W}\|^2 - c_2 \|e\|^2 \leq 0 \quad (44)$$

We can obtain $V \rightarrow 0$ from the Lyapunov Theorem, leading to $\tilde{W}$ and e converging to zero. Thus, the actual control input approach to the optimal control policy is as follows:

$$\begin{aligned}
\lim_{t \rightarrow \infty} \|v_q - v_q^*\| &= \lim_{t \rightarrow \infty} \left\| \frac{1}{2} R^{-1} g^T \nabla \sigma^T \tilde{W} + \frac{1}{2} R^{-1} g^T \nabla \varepsilon \right\| \\
&\leq \frac{\bar{\sigma}}{2} \|R^{-1} g^T\| \|\tilde{W}\| \\
&= 0.
\end{aligned} \quad (45)$$

□

5. Simulation and Discussion

We have simulated a typical disturbance nonlinear system and confirmed the availability and superiority of the optimal tracking control based on disturbance compensation. Whether it is a quadrotor or a UAV operating with a specific task, there is a complex nonlinear partially known system, which induces challenges for control design to realize tracking control. Many scholars have paid attention to this problem and achieved productive research [4,6]. The motion of the quadrotor can be described by the rigid body model, where the Euclidean position in an inertia frame and Euler angle in a body-fixed frame are expressed by $p = [p_x, p_y, p_z]^T$ and $v = \dot{p} = [v_x, v_y, v_z]^T$. Noting the rotation matrix

$$\mathbf{H} = \begin{bmatrix} \cos \theta \cos \psi & \cos \theta \sin \phi \cos \psi - \cos \phi \sin \psi & \sin \psi \sin \theta + \sin \theta \cos \psi \cos \phi \\ \cos \theta \sin \psi & \cos \phi \cos \psi + \sin \phi \sin \theta \cos \psi & \sin \theta \cos \phi \cos \psi - \cos \psi \sin \phi \\ -\sin \theta & \cos \theta \sin \phi & \cos \theta \cos \phi \end{bmatrix} \in \mathbb{R}^{3 \times 3} \quad (46)$$

with ϕ, θ, ψ representing roll, pitch, and yaw angles, and leaving out vane fluttering and air-operated friction, the dynamics of a quadrotor can be modeled as in the following formulae,

$$\begin{aligned}
\dot{p} &= v \\
m\dot{v} &= -mgh_3 + Hu - \Pi v + \wp
\end{aligned} \quad (47)$$

with m and g being mass and gravitational acceleration, $\mathbf{h}_3 = [0, 0, 1]^T$. $\mathbf{\Pi}$ and $\wp$ are unknown bounded uncertainties. To employ the designed controller, (47) is transferred into the following compact form as

$$\dot{\mathbf{X}} = \mathbf{f}_0(\mathbf{X}) + \mathbf{g}(\mathbf{X})\mathbf{u} + \Delta\mathbf{f}(\mathbf{X}) + \mathbf{d}(t) \quad (48)$$

where

$$\begin{aligned} \mathbf{X} &= \begin{bmatrix} \mathbf{p} \\ \mathbf{v} \end{bmatrix} \in \mathbb{R}^{6 \times 1}, \mathbf{f}_0(\mathbf{X}) = \begin{bmatrix} \mathbf{v} \\ -g\mathbf{h}_3 \end{bmatrix} \in \mathbb{R}^{6 \times 1}, \mathbf{g}(\mathbf{X}) = \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \in \mathbb{R}^{6 \times 3} \\ \mathbf{u} &= \begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix} \in \mathbb{R}^{3 \times 1}, \Delta\mathbf{f}(\mathbf{X}) = \begin{bmatrix} \mathbf{0} \\ -\mathbf{\Pi}\mathbf{v} \end{bmatrix} \in \mathbb{R}^{6 \times 1}, \mathbf{d}(t) = \begin{bmatrix} \mathbf{0} \\ \wp/m \end{bmatrix} \in \mathbb{R}^{6 \times 1} \end{aligned} \quad (49)$$

Denoting the synthesized unknown term as stated in (9), i.e., $\Delta = \Delta\mathbf{f}(\mathbf{x}) + \mathbf{d}(t)$, the system parameters and reference command of position and velocity are listed in Table 2. The initial position and velocity are $\mathbf{p}(0) = [-5, 3, 2]^T$, $\mathbf{v}(0) = [0, 0, 0]^T$.

Table 2. Model parameters and reference commands.

Variable	Value
m	2kg
g	9.8 m/s ²
$\mathbf{\Pi}$	diag{0.01, 0.01, 0.01}Nms ²
$\mathbf{d}$	$[\sin(4t) + \cos(2t) - \sin(t); \cos(4t) + \sin(2t) - \cos(t); \sin(3t) \cos(2t) - \cos(t)]^T$
$\mathbf{p}_d$	$[10(1 - \cos(0.1\pi t)), 5 \sin(0.2\pi t), 9(1 - e^{-0.3t})]^T$
$\mathbf{v}_d$	$[\pi \sin(0.1\pi t), \pi \cos(0.2\pi t), 2.7e^{-0.3t}]^T$

On account of the mentioned controller design, the devised controller is based on AFXDO and optimal control policy, where the unknown term is estimated by AFXDO. The optimal control policy is designed based on the ADP framework with the novel adaptive law (30). The optimal cost function is approximated by critic NN, in which Gaussian basis functions $\sigma = [\sigma_1, \sigma_2, \dots, \sigma_{15}]^T$ are considered regressors

$$\sigma_i = \exp\left[-\frac{\|\mathbf{e} - \mathbf{a}_i\|^2}{q_i^2}\right], \quad i = 1, 2, \dots, 15 \quad (50)$$

with $\mathbf{e} = [\mathbf{p}^T - \mathbf{p}_d^T, \mathbf{v}^T - \mathbf{v}_d^T]^T$, the width of the Gaussian function is set as $q_i = 5$, and the centers $\mathbf{a}_i = [a_{i1}, a_{i2}, \dots, a_{i6}]^T$ are evenly distributed over the state space $[-5, 5]$. Correspondingly, we choose the initial weight $\mathbf{W}(0) = [0, 0, \dots, 0]^T \in \mathbb{R}^{15}$ for updating critic NN weights $\mathbf{W} = [W_1, W_2, \dots, W_{15}]^T \in \mathbb{R}^{15}$. The computer programming is clearly clarified in Table 3 with specific steps. We will verify the validity of the proposed scheme from the observer and weight law.

Table 3. The computer programming of the proposed scheme.

Step 1	Initialize $\mathbf{W} = 0, \mathbf{M} = 0, \mathbf{N} = 0$.
Step 2	Compute the tracking errors, feedforward controller $\mathbf{u}_f$ and lumped disturbance approximation (10).
Step 3	Compute approximation of cost function (25) with critic NN.
Step 4	Update the auxiliary variables $\mathbf{M}$ and $\mathbf{N}$ (29).
Step 5	Update the weight of critic NN (30).
Step 6	Calculate the optimal control $\mathbf{v}_q$ (26) with the approximate weight of critic NN.
Step 7	Update the controller $\mathbf{u}_q$ according to lumped disturbance approximation and optimal control.
Step 8	If $t < t_f$, return to Step 2; else stop.

Case 1: Effect of disturbance observer on closed-loop performance

To verify the superiority of the proposed AFXDO, the following controllers are carried out, where different observers are inserted into the proposed ADP controller. For making a fair comparison, the controller parameters are set as: $K = \text{diag}[1, 1, 1, 1, 1, 1]$, $Q = I$, $R = I$, $Y_0 = 100I$, and $\delta = 2.1$.

- (1) The designed ADP controller with extended state observer (ESO-ADP): The observer in [4] is utilized for the estimation of disturbance with the ADP framework. The bandwidth is set as $w_0 = 8$.
- (2) The designed ADP controller with finite-time DO (FIDO-ADP): The finite-time DO in [28] is employed with adaptive gains to eliminate the effect of disturbance in a finite time. The parameters of the observer are set as $\beta = 5/11$, $k_1 = k_2 = 10$, $l_1 = l_2 = 0.01$, and $\mu_1 = \mu_2 = 0.001$.
- (3) The designed ADP controller with the proposed fixed-time DO (FXDO-ADP): During the operation of this method, we select the following parameters in AFXDO to ensure stability: $\alpha = 2$, $\beta = 5/11$, $k_1 = k_2 = 10$, $l_1 = l_2 = 0.01$, and $\mu_1 = \mu_2 = 0.001$.

It is worth noting that the different observers are employed in the critic-only ADP scheme in the above algorithm. Based on the structure of the observer and the ADP framework, the time and space complexity are consistent. The simulation results are illustrated in Figures 1–6. The disturbance estimation via three observers and real disturbance is shown in Figure 1, where the disturbance estimation of different observers can accurately estimate the real disturbance within 2 s. In addition, the designed AFXDO with faster transient performance can achieve accurate disturbance estimation in 0.5 s. In order to verify that AFXDO can make the disturbance estimation error converge within a fixed time, the disturbance estimation errors via different observers are shown in Figure 3. Noting that the convergence time of the disturbance estimation error with FIDO is affected by the initial disturbance estimation error, when the initial estimation is poor, the convergence time will be larger. Compared with the optimal control based on ESO and FIDO, the optimal control based on AFXDO achieves accurate disturbance estimation in 0.5 s. The settling time of disturbance estimation error is irrelevant to the initial perturbation estimation error, which is less than the theoretical settling time of 17/3 s from Theorem 1. In Figure 4, the observer gains of AFXDO are depicted. Due to the design of the observer without the upper bound of the disturbance, the observer gain is updated adaptively online by (10), avoiding the assumption of a known upper bound of the disturbance.

As shown in Figures 3 and 4, AFXDO achieves the rapid convergence of disturbance estimation. Due to the rapid convergence of disturbance, which will further affect the tracking performance, the position tracking error and velocity tracking error are shown under three controllers in Figures 5 and 6. The accuracy of the position tracking error under the three controllers reaches 10^{-1} within 6 s. In addition, the trajectory tracking error of FXDO-ADP can enter steady-state performance within 5 s, which is smaller than that of the other two controllers, indicating that the rapid estimation of disturbance accelerates the convergence of the trajectory tracking error. As shown in Figures 5 and 6, the velocity tracking errors with three controllers can enter the steady-state performance within 2 s. For ESO-ADP and FIDO-ADP, the tracking errors of FIDO-ADP can achieve steady-state performance within 1.3 s, which is faster than that of ESO-ADP. Further, the tracking errors of FXDO-ADP can enter the steady state within 1 s. Therefore, the transient performance of AFXDO-ADP is better than that of the ESO-ADP and FIDO-ADP controllers in terms of tracking errors of position and velocity. Due to the fact that AFXDO can make disturbance estimation follow real disturbance in a fixed time, the position and velocity tracking errors can enter a steady state within a short time.

Taking channel x as an example, Figure 7 compares the control inputs of the three controllers, where the control inputs of FXDO-ADP and ESO-ADP are less. Combined with tracking errors in Figures 5 and 6, the FXDO-ADP can achieve the rapid convergence of tracking errors with equivalent control consumption. Therefore, the controller makes the predefined cost function smaller among the three controllers.

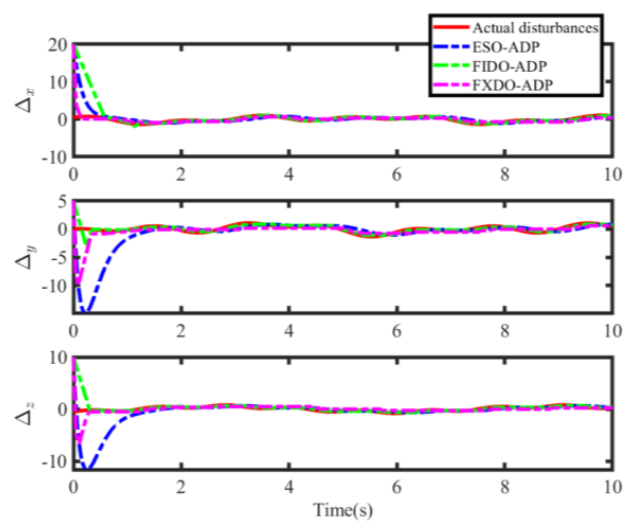


Figure 2. The disturbance estimation via different observers.

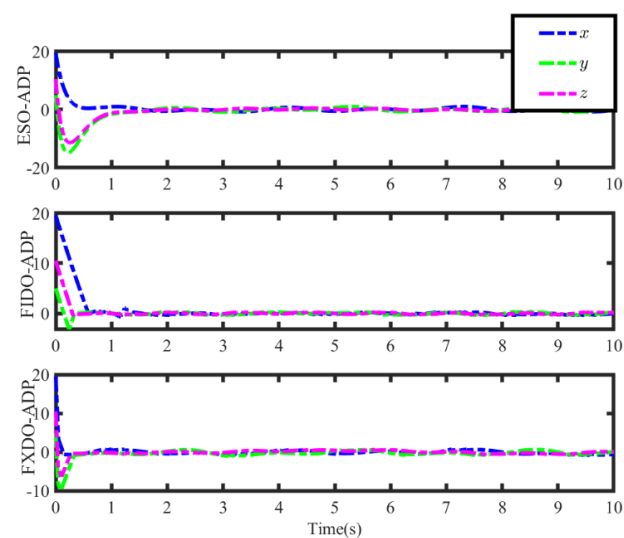


Figure 3. The disturbance estimation error via different observers.

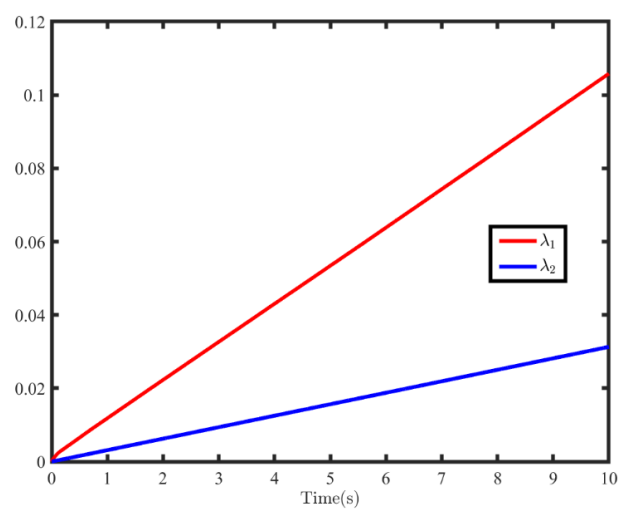


Figure 4. The adaptive observer gains of AFXDO.

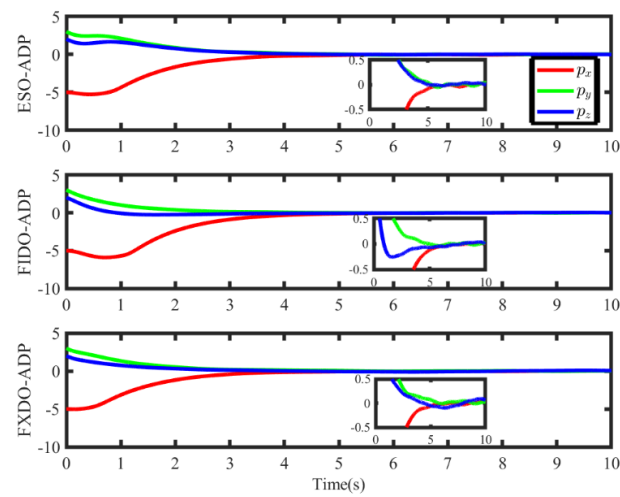


Figure 5. The position tracking errors under three controllers.

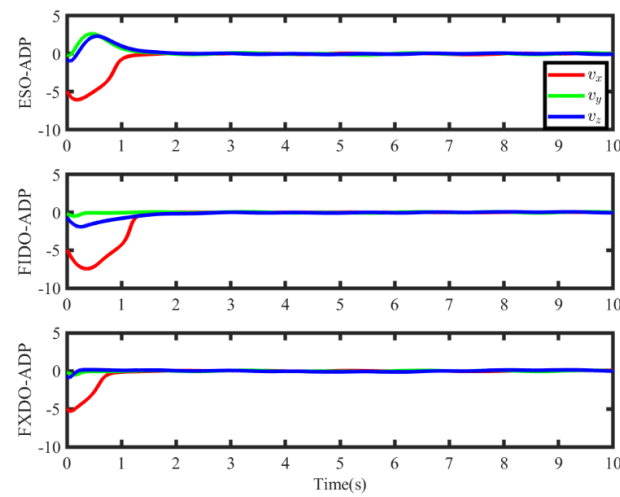


Figure 6. The velocity tracking errors under three controllers.

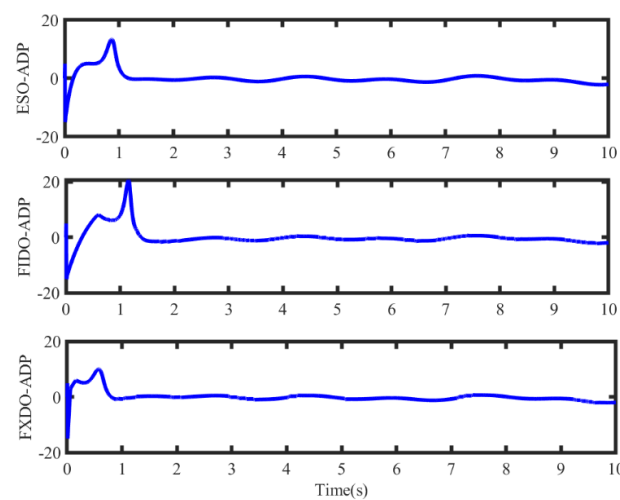


Figure 7. The control inputs under three controllers.

Case 2: Effect of weight update on learning performance

To explore the effectiveness of weight update for deriving the optimal policy, we compare the tracking performance and cost consumptions among the following schemes:

- (1) The designed ADP controller with a fixed learning rate (ADPFRL) [50]: The fixed learning rate is set as $\mathbf{Y} = 100\mathbf{I}_{15}$ to accelerate the convergence of tracking errors.
- (2) The designed ADP controller with a proposed variable learning rate (ADPVLR): The weight is updated by (30) rather than the fixed learning rate in ADPFRL. The parameters are selected as $\mathbf{Y} = 100\mathbf{I}_{15}$, $\delta = 2.1$.
- (3) The actor–critic ADP controller with the gradient method [35] (ADP2): The learning rate in actor and critic NNs are set as $\mathbf{Y}_a = 100\mathbf{I}_{15}$, $\mathbf{Y}_c = 10\mathbf{I}_{15}$.

Different from critic-only ADP, there exist the actor NN and critic NN to conduct the actor–critic ADP, which leads to the space and time complexity increasing due to the extra actor NN. In addition, on account of the variable learning rate designed with tracking errors, the space and time complexity are nonincreasing, where the tracking error is essential for other weight laws. The simulation results are shown in Figures 8–13. In Figures 8 and 9, the position tracking error and velocity tracking error under three controllers are shown, where different weight laws are employed under the ADP framework. The position tracking error enters the steady-state performance within 5 s, and the velocity tracking error converges to the neighborhood of origin within 2 s. In specific, the ADP based on critic NN has better transient performance than the ADP2 control based on the actor–critic NN in terms of velocity tracking errors and can achieve accurate tracking within 1 s, which is due to the hysteresis caused by weight convergence of the actor–critic NN.

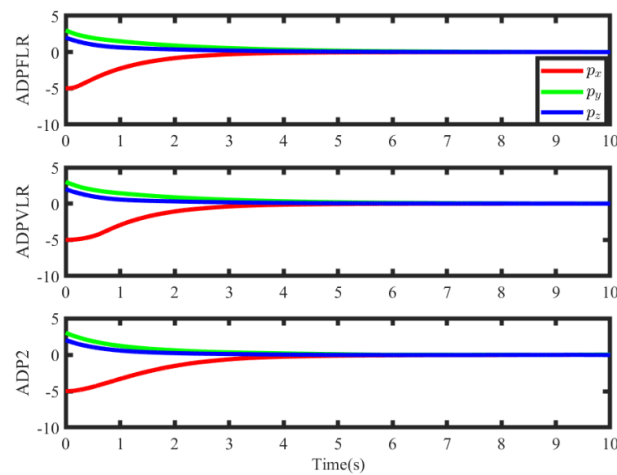


Figure 8. The position tracking errors under different controllers.

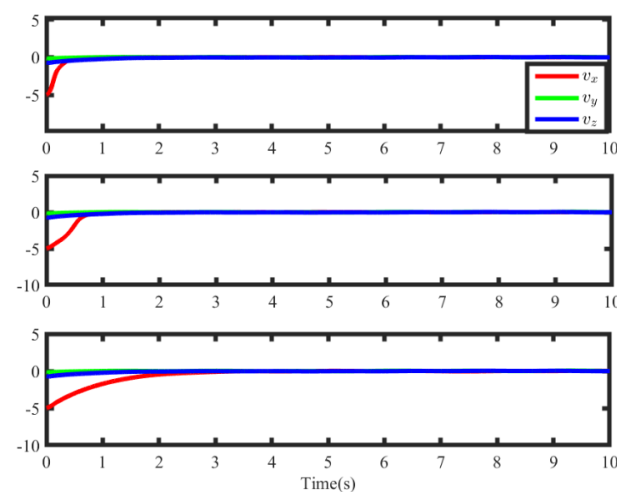


Figure 9. The velocity tracking errors under different controllers.

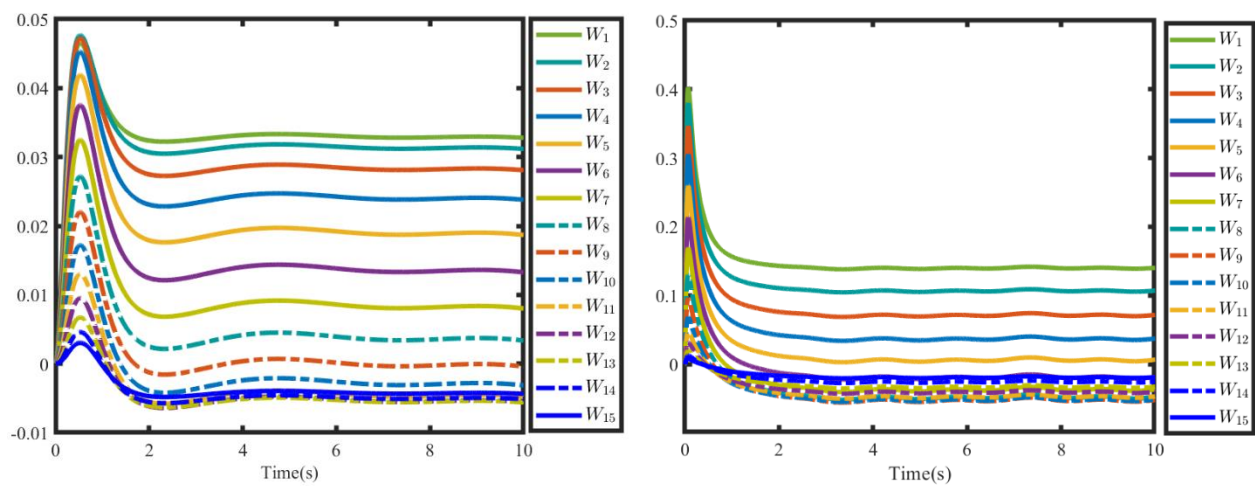


Figure 10. The weights in actor and critic NNs of ADP2.

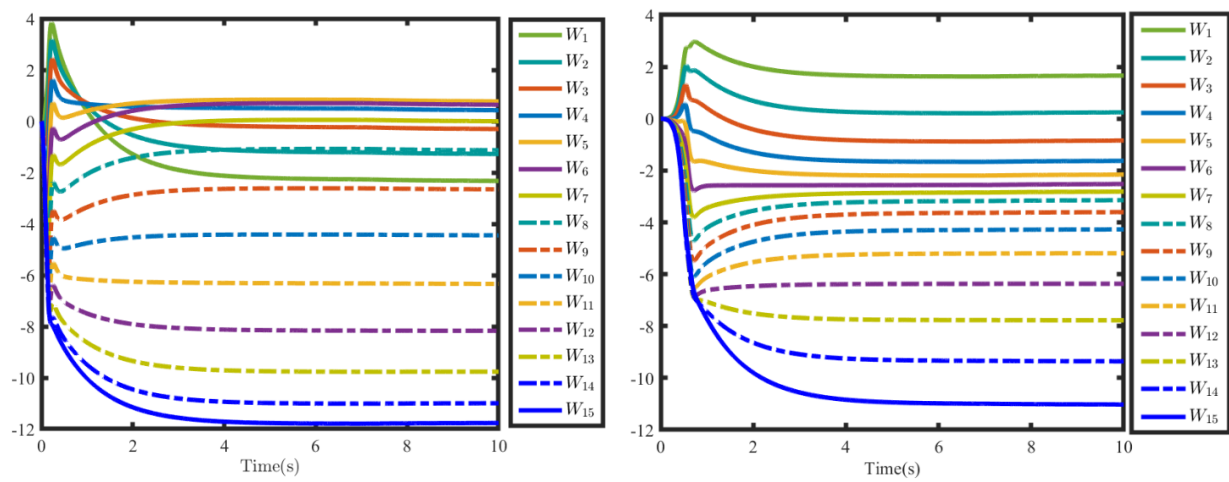


Figure 11. The weights of ADPFRL and ADPVLR.

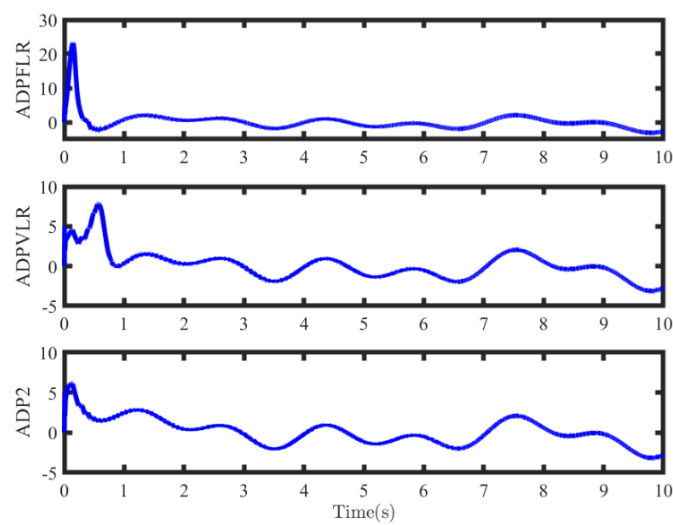


Figure 12. The control inputs under different controllers.

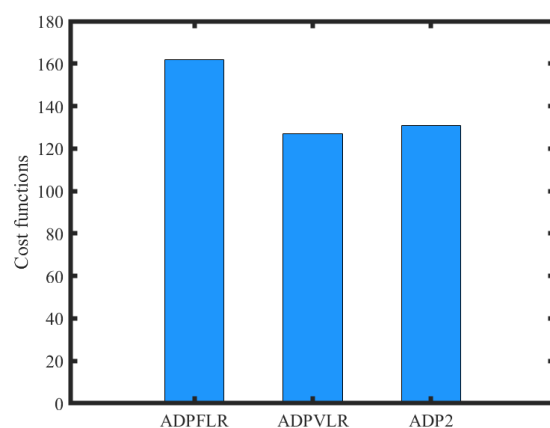


Figure 13. The comparison of cost functions.

Figures 10 and 11 show the weight of the NN under the three controllers. The weights of the actor NN and critic NN in ADP2 are described in Figure 10, where the weight estimation of the critic NN and actor NN have obvious chattering at the initial stage. Conversely, the weight of the critic NN converges to an ideal weight based on critic-only ADP, as shown in Figure 11. Under the critic-only ADP framework, the weight update law of critic NN extracts the weight error through the auxiliary filtering operation, such that the weight update is driven by the weight error implied in the updated law, resulting in the fast convergence of weight. Furthermore, Figure 11 compares the weight of ADPFLR and ADPVLR. As shown in Figure 11, the ADPVLR significantly reduces the chattering occurring at the initial stage of ADPFLR with a fixed learning rate, which indicates that the proposed variable learning rate is superior to the weight update with a fixed learning rate. Specifically, its key features can be described thus: when the tracking error is large, the learning rate is reduced to avoid chattering; when the tracking error is small, the learning rate is increased to accelerate the convergence rate.

Since the convergence of weight affects the optimal control, the control inputs of three controllers in the x channel are shown in Figure 12. Compared with ADPFLR and ADPVLR, the control input of ADPVLR is smaller, which can be due to the weight of the variable learning rate when avoiding the initial large control input. In addition, the control input of ADP2 is small, which leads to the slow convergence of tracking errors. Considering both tracking errors and control consumption, the cost functions of the three controllers are shown in Figure 13, where ADPFLR with a fixed learning rate requires high control consumption to achieve fast tracking performance. By the weight update law (30), the cost function of ADPVLR is decreased by 19.13% from that of ADPFLR. By introducing the weight updating law of variable learning rate, ADPVLR has a tradeoff ability in control consumption and tracking performance, which not only achieves better tracking performance but also ensures lower control consumption.

By analyzing the effectiveness of disturbance observer and new weight law, although AFXDO is carried with large control consumption to achieve faster convergence, by combining AFXDO and weight law with a variable learning rate, the optimal control consumption is reduced, so the total control consumption is decreased, which induces lower cost functions.

Different from the uncertainty of the above system, $\Delta = \Delta f(x) + d(t)$, we set the only $d(t)$ to unknown for a sensitive analysis of the system. The tracking performance of position and velocity are shown in Figures 14 and 15. The position and velocity can follow the reference position and velocity in 4 s. The weight of the critic NN is described in Figure 16, avoiding the initial chattering phenomenon. We will validate the effectiveness of the completely unknown system in actual application.

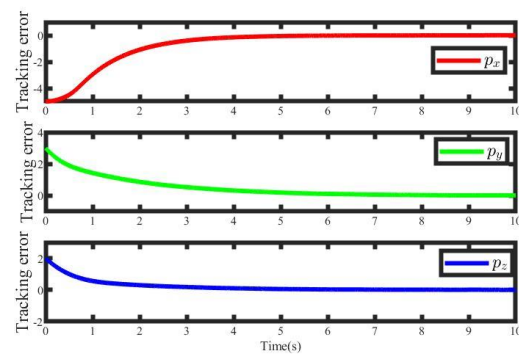


Figure 14. The position tracking errors.

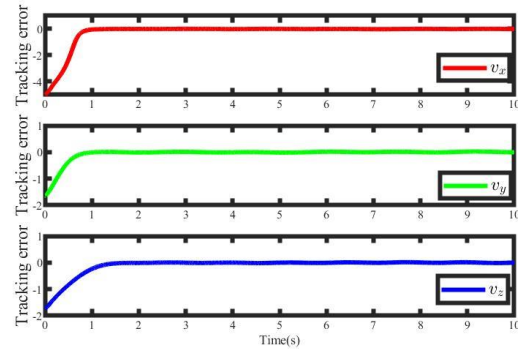


Figure 15. The velocity tracking errors.

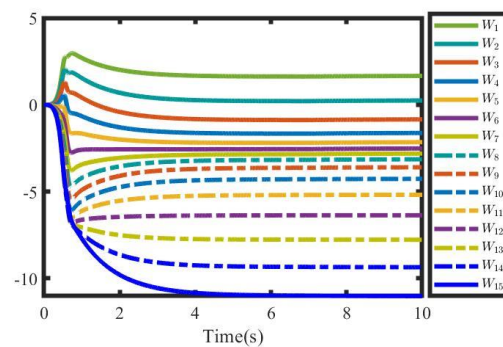


Figure 16. The weight of critic NN.

6. Conclusions

The approximate tracking control with adaptive fixed-time estimation is developed to address the tracking problem of partially unknown nonlinear systems. After the procession of feedforward control for the error system, the novel adaptive fixed-time disturbance observer is designed such that the estimation error for disturbance can converge to zero within a fixed time. Then, an optimal control policy is derived by the critic-only ADP framework, where a novel weight adaptive law with a variable learning rate is proposed for updating weight without initial oscillation phenomena. The tracking errors and weight estimation errors can be proven to be UUB. The designed approximate optimal policy approaches the optimal control. The superiority of the proposed controller can be verified in the simulation, which well balances tracking performance and control consumption. Noting that control input and tracking error may exceed the capacity of the actuator and predefined region, which is not considered in controller design, future work will concentrate on the extension of the proposed controller with a new cost function for a partially unknown nonlinear system with input saturation and output constraints to reduce control consumption and improve tracking performance.

Author Contributions: Conceptualization, Y.G. and Z.L.; methodology, Y.G. and Z.L.; software, Y.G.; validation, Y.G.; formal analysis, Y.G.; writing—original draft preparation, Y.G.; writing—review and editing, Y.G. and Z.L.; supervision, Z.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the first author.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

The following derivation process is given for the settling time of disturbance estimation errors.

Proof. Based on the observer (10), the observer error dynamics can be derived as

$$\begin{cases} \dot{\tilde{e}}_i = -k_1 \tilde{e}_i^\alpha \text{sign}(\tilde{e}_i) + \tilde{\Delta}_i - \lambda_1 \tilde{h}_i \\ \dot{\tilde{\Delta}}_i = -k_2 \tilde{h}_i^\chi \text{sign}(\tilde{h}_i) - \lambda_2 \text{sign}(\tilde{h}_i) + \dot{\Delta}_i \end{cases} \quad (\text{A1})$$

For simplicity and clarity, the following analyses are based on each component of estimation errors. Denote $\bar{\lambda}_1 = \max\{\max_{[0,T]} r\lambda_1 + \varepsilon_1, \max_{[0,T]} |\tilde{\Delta}_i|\}$ and the following candidate Lyapunov function is chosen:

$$V_e^i = \frac{1}{2-2\beta} \tilde{e}_i^{2-2\beta} + \frac{1}{2r\lambda_1} \kappa_1 (r\lambda_1 - \bar{\lambda}_1)^2 \quad (\text{A2})$$

with $\kappa_1 = \frac{2+\bar{\lambda}_1^\omega}{2} + \frac{\bar{\lambda}_1^\omega}{2} (\text{sign}(|\tilde{e}_i|) - 1)$ and $\omega = \frac{\alpha-\beta}{1-\beta}$. When $|\tilde{e}_i| > 1$, differentiating V_e^i along (11) and (12) gives

$$\begin{aligned} \dot{V}_{ei} &= \tilde{e}_i^{1-2\beta} \left(-k_1 \tilde{e}_i^\alpha \text{sign}(\tilde{e}_i) + \tilde{\Delta}_i - \lambda_1 r \text{sign}(\tilde{e}_i) \right) + (1 + \bar{\lambda}_1^\omega) (r\lambda_1 - \bar{\lambda}_1) \left(|\tilde{e}_i|^{1-2\beta} + \frac{\mu_1}{l_1} \right) \\ &= -k_1 \tilde{e}_i^{1-2\beta+\alpha} \text{sign}(\tilde{e}_i) + \tilde{e}_i^{1-2\beta} \tilde{\Delta}_i - \lambda_1 r \text{sign}(\tilde{e}_i) \tilde{e}_i^{1-2\beta} + \bar{\lambda}_1^\omega |\tilde{e}_i|^{1-2\beta} (r\lambda_1 - \bar{\lambda}_1) \\ &\quad + \frac{\mu_1}{l_1} \bar{\lambda}_1^\omega (r\lambda_1 - \bar{\lambda}_1) + r\lambda_1 |\tilde{e}_i|^{1-2\beta} - \bar{\lambda}_1 |\tilde{e}_i|^{1-2\beta} + \frac{\mu_1}{l_1} (r\lambda_1 - \bar{\lambda}_1) \end{aligned} \quad (\text{A3})$$

As $\tilde{e}_i^{1-2\beta} \tilde{\Delta}_i - \bar{\lambda}_1 |\tilde{e}_i|^{1-2\beta} \leq 0$, $\bar{\lambda}_1^\omega |\tilde{e}_i|^{1-2\beta} (r\lambda_1 - \bar{\lambda}_1) \leq 0$ and $\lambda_1 r \text{sign}(\tilde{e}_i) \tilde{e}_i^{1-2\beta} = r\lambda_1 |\tilde{e}_i|^{1-2\beta}$, it follows along (7) in Lemma 2 that

$$\begin{aligned} \dot{V}_{ei} &\leq -k_1 \tilde{e}_i^{1-2\beta+\alpha} \text{sign}(\tilde{e}_i) + \frac{\mu_1}{l_1} \bar{\lambda}_1^\omega (r\lambda_1 - \bar{\lambda}_1) + \frac{\mu_1}{l_1} (r\lambda_1 - \bar{\lambda}_1) \\ &\leq -k_1 \tilde{e}_i^{1-2\beta+\alpha} \text{sign}(\tilde{e}_i) - |r\lambda_1 - \bar{\lambda}_1|^{\omega+1} \frac{\mu_1}{l_1} - |r\lambda_1 - \bar{\lambda}_1| \frac{\mu_1}{l_1} \\ &\leq -k_1 \tilde{e}_i^{1-2\beta+\alpha} \text{sign}(\tilde{e}_i) - \frac{\mu_1}{l_1} |r\lambda_1 - \bar{\lambda}_1|^{\omega+1} \\ &= -k_1 \left(\frac{1}{2-2\beta} \right)^{-\frac{\omega+1}{2}} \left[\left(\frac{1}{2-2\beta} \right)^{\frac{\omega+1}{2}} \tilde{e}_i^{1-2\beta+\alpha} \text{sign}(\tilde{e}_i) \right] - \left(\frac{1}{2r\lambda_1} \right)^{-\frac{\omega+1}{2}} \left[\left(\frac{1}{2r\lambda_1} \right)^{\frac{\omega+1}{2}} |r\lambda_1 - \bar{\lambda}_1|^{\omega+1} \right] \frac{\mu_1}{l_1} \\ &\leq -\sigma_1 V_{ei}^{\frac{\omega+1}{2}} \end{aligned} \quad (\text{A4})$$

where

$$\sigma_1 = \min \left\{ k_1 \left(\frac{1}{2-2\beta} \right)^{-\frac{\omega+1}{2}}, \left(\frac{1}{2r\lambda_1} \right)^{-\frac{\omega+1}{2}} \frac{\mu_1}{l_1} \right\} \quad (\text{A5})$$

Furthermore, it follows for $|\tilde{e}_i| < 1$ that

$$\begin{aligned} \dot{V}_{ei} &= \tilde{e}_i^{1-2\beta} \left(-k_1 \tilde{e}_i^\beta \text{sign}(\tilde{e}_i) + \tilde{\Delta}_i - \lambda_1 r \text{sign}(\tilde{e}_i) \right) + (r\lambda_1 - \bar{\lambda}_1) \times \left(|\tilde{e}_i|^{1-2\beta} + \frac{\mu_1}{l_1} \right) \\ &= -k_1 \tilde{e}_i^{1-\beta} \text{sign}(\tilde{e}_i) + \tilde{e}_i^{1-2\beta} \tilde{\Delta}_i - \lambda_1 r \text{sign}(\tilde{e}_i) \tilde{e}_i^{1-2\beta} + r\lambda_1 |\tilde{e}_i|^{1-2\beta} - \bar{\lambda}_1 |\tilde{e}_i|^{1-2\beta} + \frac{\mu_1}{l_1} (r\lambda_1 - \bar{\lambda}_1) \\ &= -k_1 \tilde{e}_i^{1-\beta} \text{sign}(\tilde{e}_i) + \tilde{e}_i^{1-2\beta} \tilde{\Delta}_i - \bar{\lambda}_1 |\tilde{e}_i|^{1-2\beta} + \frac{\mu_1}{l_1} (r\lambda_1 - \bar{\lambda}_1) \end{aligned} \quad (\text{A6})$$

As $\tilde{e}_i^{1-2\beta} \tilde{\Delta}_i - \bar{\lambda}_1 |\tilde{e}_i|^{1-2\beta} < 0$, it follows from (6) in Lemma 2 that

$$\begin{aligned} \dot{V}_{ei} &\leq -k_1 \left(\frac{1}{2-2\beta} \right)^{-\frac{1}{2}} \left[\left(\frac{1}{2-2\beta} \right)^{\frac{1}{2}} \tilde{e}_i^{1-\beta} \text{sign}(\tilde{e}_i) \right] - \left(\frac{1}{2r\lambda_1} \right)^{-\frac{1}{2}} \left[\left(\frac{1}{2r\lambda_1} \right)^{\frac{1}{2}} (r\lambda_1 - \bar{\lambda}_1) \right] \frac{\mu_1}{l_1} \\ &\leq -\sigma_2 V_{ei}^{\frac{1}{2}} \end{aligned} \quad (\text{A7})$$

where

$$\sigma_2 = \min \left\{ k_1 \left(\frac{1}{2-2\beta} \right)^{-\frac{1}{2}}, \left(\frac{1}{2r\lambda_1} \right)^{-\frac{1}{2}} \frac{\mu_1}{l_1} \right\} \quad (\text{A8})$$

Thus, the settling time of state error can be obtained as

$$\begin{aligned} T_1 &= -\frac{1}{\sigma_1} \frac{1}{\frac{\omega+1}{2}-1} V_{ei}^{-\frac{\omega+1}{2}+1} \Big|_1^{V_{ei}(0)} + \frac{1}{\sigma_2} \frac{1}{1-\frac{1}{2}} V_{ei}^{-\frac{1}{2}+1} \Big|_0^1 \\ &\leq \frac{1}{\sigma_1} \frac{2}{\omega-1} \left(1 - \lim_{V_{ei}(0) \rightarrow \infty} V_{ei}(0)^{-\frac{\omega+1}{2}+1} \right) + \frac{2}{\sigma_2} \\ &= \frac{1}{\sigma_1} \frac{2}{\omega-1} + \frac{2}{\sigma_2} \end{aligned} \quad (\text{A9})$$

After $\tilde{e}_i$ enters the sliding mode surface $\tilde{e}_i \equiv 0$, $\lambda_1(\tilde{h}_i)_{eq} = \tilde{\Delta}$ holds according to the sliding-mode equivalent control theory, which yields the restatement of the disturbance dynamic as

$$\dot{\tilde{\Delta}}_i = -k_2 \tilde{e}_i \text{sign}(\tilde{e}_i) - \lambda_2 \text{sign}(\tilde{e}_i) + \dot{\Delta}_i \quad (\text{A10})$$

with

$$\dot{\lambda}_2 = l_2 \lambda_2^{2\beta-1} |\tilde{e}_i|^{1-2\beta} + \mu_2 \quad (\text{A11})$$

Another Lyapunov function is considered as

$$V_{\Delta i} = \frac{1}{2-2\beta} \tilde{e}_i^{2-2\beta} + \frac{\lambda_2^{2-2\beta}}{2l_2} \kappa_2 (r\lambda_2 - \bar{\lambda}_2)^2 \quad (\text{A12})$$

where $\kappa_2 = \frac{2+\bar{\lambda}_2^\omega}{2} + \frac{\bar{\lambda}_2^\omega}{2} (\text{sign}(|\tilde{e}_i| - 1))$. Considering the above-mentioned two cases and following a similar technique, one can derive from $|\tilde{h}_i| > 1$ that

$$\begin{aligned} \dot{V}_{\Delta i} &\leq -k_2 \left(\frac{1}{2-2\beta} \right)^{-\frac{\omega+1}{2}} \left[\left(\frac{1}{2-2\beta} \right)^{\frac{\omega+1}{2}} \tilde{e}_i^{1-2\beta+\alpha} \text{sign}(\tilde{e}_i) \right] - \left(\frac{1}{2rl_2} \right)^{-\frac{\omega+1}{2}} \left[\left(\frac{1}{2rl_2} \right)^{\frac{\omega+1}{2}} |r\lambda_2 - \bar{\lambda}_2|^{\omega+1} \right] \frac{\mu_2}{l_2} \\ &\leq -\sigma_3 V_{\Delta i}^{\frac{\omega+1}{2}} \end{aligned} \quad (\text{A13})$$

with

$$\sigma_3 = \min \left\{ k_2 \left(\frac{1}{2-2\beta} \right)^{-\frac{\omega+1}{2}}, \left(\frac{1}{2rl_2} \right)^{-\frac{\omega+1}{2}} \frac{\mu_2}{l_2} \right\} \quad (\text{A14})$$

Otherwise, it follows from $|\tilde{h}_i| < 1$ that

$$\begin{aligned} \dot{V}_{\Delta i} &\leq -k_2 \left(\frac{1}{2-2\beta} \right)^{-\frac{1}{2}} \left[\left(\frac{1}{2-2\beta} \right)^{\frac{1}{2}} \tilde{e}_i^{1-\beta} \text{sign}(\tilde{e}_i) \right] - \left(\frac{1}{2rl_2} \right)^{-\frac{1}{2}} \left[\left(\frac{1}{2rl_2} \right)^{\frac{1}{2}} (r\lambda_2 - \bar{\lambda}_2) \right] \frac{\mu_2}{l_2} \\ &\leq -\sigma_4 V_{\Delta i}^{\frac{1}{2}} \end{aligned} \quad (\text{A15})$$

with

$$\sigma_4 = \min \left\{ k_2 \left(\frac{1}{2-2\beta} \right)^{-\frac{1}{2}}, \left(\frac{1}{2rl_2} \right)^{-\frac{1}{2}} \frac{\mu_2}{l_2} \right\} \quad (\text{A16})$$

After the settling time T_1 , the settling time of the perturbation estimation error can be computed as

$$\begin{aligned} T_2 &= -\frac{1}{\sigma_3} \frac{1}{\frac{\omega+1}{2}-1} V_{\Delta i}^{-\frac{\omega+1}{2}+1} \Big|_1^{V_{\Delta i}(0)} + \frac{1}{\sigma_4} \frac{1}{1-\frac{1}{2}} V_{\Delta i}^{-\frac{1}{2}+1} \Big|_0^1 \\ &\leq \frac{1}{\sigma_3} \frac{2}{\omega-1} \left(1 - \lim_{V_{\Delta i}(0) \rightarrow \infty} V_{\Delta i}(0)^{-\frac{\omega+1}{2}+1} \right) + \frac{2}{\sigma_4} \\ &= \frac{1}{\sigma_3} \frac{2}{\omega-1} + \frac{2}{\sigma_4} \end{aligned} \quad (\text{A17})$$

Therefore, one can achieve $V_{\Delta i} = 0$ after the settling time $T = T_1 + T_2$, which indicates perturbation estimation errors equal to zero. Among the observers, the adaptive observer gains λ_1 and λ_2 are increasing sufficiently to compensate for the impression of perturbation and remain unchanged after entering sliding mode surfaces for the estimation of state and perturbation. $\square$

References

- Castellini, A.; Marchesini, E.; Farinelli, A. Partially observable monte carlo planning with state variable constraints for mobile robot navigation. *Eng. Appl. Artif. Intell.* **2021**, *104*, 104382. [\[CrossRef\]](#)
- Dai, Y.; Ni, S.; Xu, D.; Zhang, L.; Yan, X.G. Disturbance-observer based prescribed-performance fuzzy sliding mode control for PMSM in electric vehicles. *Eng. Appl. Artif. Intell.* **2021**, *104*, 104361. [\[CrossRef\]](#)

3. Park, B.S.; Yoo, S.J. Quantized-communication-based neural network control for formation tracking of networked multiple unmanned surface vehicles without velocity information. *Eng. Appl. Artif. Intell.* **2022**, *114*, 105160. [\[CrossRef\]](#)
4. Shao, X.; Yue, X.; Li, J. Event-triggered robust control for quadrotors with preassigned time performance constraints. *Appl. Math. Comput.* **2021**, *392*, 125667. [\[CrossRef\]](#)
5. Wang, Y.; Luo, G.; Wang, D. Observer-based fixed-time adaptive fuzzy control for SbW systems with prescribed performance. *Eng. Appl. Artif. Intell.* **2022**, *114*, 105026. [\[CrossRef\]](#)
6. Wu, L.; Li, Z.; Liu, S.; Li, Z.; Sun, D. A novel multi-agent model-free adaptive control algorithm for a class of multivehicle systems with constraints. *Symmetry* **2023**, *15*, 168. [\[CrossRef\]](#)
7. Wu, J.; Sun, W.; Su, S.F.; Wu, Y. Adaptive asymptotic tracking control for input-quantized nonlinear systems with multiple unknown control directions. *IEEE Trans. Cybern.* **2022**, *in press*. [\[CrossRef\]](#)
8. Qiyas, M.; Abdullah, S.; Khan, F.; Naeem, M. Banzhaf-Choquet-Copula-based aggregation operators for managing fractional orthotriple fuzzy information. *Alex. Eng. J.* **2022**, *61*, 4659–4677. [\[CrossRef\]](#)
9. Khan, A.; Ashraf, S.; Abdullah, S.; Muhammad, A.; Thongchai, B. A novel decision aid approach based on spherical hesitant fuzzy Aczel-Alsina geometric aggregation information. *AIMS Math.* **2023**, *8*, 5148–5174. [\[CrossRef\]](#)
10. Qiyas, M.; Naeem, M.; Abdullah, S.; Khan, F.; Khan, N.; Garg, H. Fractional orthotriple fuzzy rough Hamacher aggregation operators and their application on service quality of wireless network selection. *Alex. Eng. J.* **2022**, *61*, 10433–10452. [\[CrossRef\]](#)
11. Yahya, M.; Abdullah, S.; Almagrabi, A.O.; Botmart, T. Analysis of S-box based on image encryption application using complex fuzzy credibility Frank aggregation operators. *IEEE Access* **2022**, *10*, 88858–88871. [\[CrossRef\]](#)
12. Mohammad, M.M.S.; Abdullah, S.; Al-Shomrani, M.M. Some linear Diophantine fuzzy similarity measures and their application in decision making problem. *IEEE Access* **2022**, *10*, 29859–29877. [\[CrossRef\]](#)
13. Qiyas, M.; Madrar, T.; Khan, S.; Abdullah, S.; Botmart, T.; Jirawattanapant, A. Decision support system based on fuzzy credibility Dombi aggregation operators and modified TOPSIS method. *AIMS Math.* **2022**, *7*, 19057–19082. [\[CrossRef\]](#)
14. Ahmad, S.; Basharat, P.; Abdullah, S.; Botmart, T.; Jirawattanapanit, A. MABAC under non-linear diophantine fuzzy numbers: A new approach for emergency decision support systems. *AIMS Math.* **2022**, *7*, 17699–17736. [\[CrossRef\]](#)
15. Midrar, T.; Khan, S.; Abdullah, S.; Botmart, T. Entropy based extended TOPOSIS method for MCDM problem with fuzzy credibility numbers. *AIMS Math.* **2022**, *7*, 17286–17312. [\[CrossRef\]](#)
16. Ashraf, S.; Rehman, N.; Abdullah, S.; Batool, B.; Lin, M.; Aslam, M. Decision support model for the patient admission scheduling problem based on picture fuzzy aggregation information and TOPSIS methodology. *Math. Biosci. Eng.* **2022**, *19*, 3147–3176. [\[CrossRef\]](#)
17. Batool, B.; Abdullah, S.; Ashraf, S.; Ahmad, M. Pythagorean probabilistic hesitant fuzzy aggregation operators and their application in decision-making. *Kybernetes* **2022**, *51*, 1626–1652. [\[CrossRef\]](#)
18. Abdullah, S.; Al-Shomrani, M.M.; Liu, P.; Ahmad, S. A new approach to three-way decisions making based on fractional fuzzy decision-theoretical rough set. *Int. J. Intell. Syst.* **2022**, *37*, 2428–2457. [\[CrossRef\]](#)
19. Ashraf, S.; Abdullah, S.; Chinram, R. Emergency decision support modeling under generalized spherical fuzzy Einstein aggregation information. *J. Ambient Intell. Humaniz. Comput.* **2022**, *13*, 2091–2117. [\[CrossRef\]](#)
20. Shao, S.; Chen, M.; Zheng, S.; Lu, S.; Zhao, Q. Event-triggered fractional-order tracking control for an uncertain nonlinear system with output saturation and disturbances. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *in press*. [\[CrossRef\]](#)
21. Wang, Y.; Wang, Y.; Tie, M. Hybrid adaptive learning neural network control for steer-by-wire systems via sigmoid tracking differentiator and disturbance observer. *Eng. Appl. Artif. Intell.* **2021**, *104*, 104393. [\[CrossRef\]](#)
22. Zhang, J.; Zhao, W.; Shen, G.; Xia, Y. Disturbance observer-based adaptive finite-time attitude tracking control for rigid spacecraft. *IEEE Trans. Syst. Man Cybern. Syst.* **2020**, *51*, 6606–6613. [\[CrossRef\]](#)
23. Nguyen, T.H.; Nguyen, T.T.; Nguyen, V.Q.; Le, K.M.; Tran, H.N.; Jeon, J.W. An adaptive sliding-mode controller with a modified reduced-order proportional integral observer for speed regulation of a permanent magnet synchronous motor. *IEEE Trans. Ind. Electron.* **2021**, *69*, 7181–7191. [\[CrossRef\]](#)
24. Nguyen, N.P.; Oh, H.; Kim, Y.; Moon, J.; Yang, J.; Chen, W.H. Finite-time disturbance observer-based modified super-twisting algorithm for systems with mismatched disturbances: Application to fixed-wing UAVs under wind disturbances. *Int. J. Robust Nonlin. Control* **2021**, *31*, 7317–7343. [\[CrossRef\]](#)
25. Mirzaei, M.J.; Mirzaei, M.; Aslmostafa, E.; Asadollahi, M. Robust observer-based stabilizer for perturbed nonlinear complex financial systems with market confidence and ethics risks by finite-time integral sliding mode control. *Nonlinear Dyn.* **2021**, *105*, 2283–2297. [\[CrossRef\]](#)
26. Wang, X.; Zheng, W.X.; Wang, G. Distributed finite-time optimization of second-order multiagent systems with unknown velocities and disturbances. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *in press*. [\[CrossRef\]](#)
27. Wang, H.; Zhang, Y.; Zhao, Z.; Tang, X.; Yang, J.; Chen, I. Finite-time disturbance observer-based trajectory tracking control for flexible-joint robots. *Nonlinear Dyn.* **2021**, *106*, 459–471. [\[CrossRef\]](#)
28. Huang, D.; Huang, T.; Qin, N.; Li, Y.; Yang, Y. Finite-time control for a UAV system based on finite-time disturbance observer. *Aerosp. Sci. Technol.* **2022**, *129*, 107825. [\[CrossRef\]](#)
29. Polyakov, A. Nonlinear feedback design for fixed-time stabilization of linear control systems. *IEEE Trans. Autom. Control.* **2011**, *57*, 2106–2110. [\[CrossRef\]](#)

30. Sun, L.; Sun, G.; Jiang, J. Disturbance observer-based saturated fixed-time pose tracking for feature points of two rigid bodies. *Automatica* **2022**, *144*, 110475. [\[CrossRef\]](#)
31. Sun, J.; Yi, J.; Pu, Z.; Tan, X. Fixed-time sliding mode disturbance observer-based nonsmooth backstepping control for hypersonic vehicles. *IEEE Trans. Syst. Man, Cybern. Syst.* **2018**, *50*, 4377–4386. [\[CrossRef\]](#)
32. Gao, J.; Fu, Z.; Zhang, S. Adaptive fixed-time attitude tracking control for rigid spacecraft with actuator faults. *IEEE Trans. Ind. Electron.* **2018**, *66*, 7141–7149. [\[CrossRef\]](#)
33. Hu, G.; Guo, J.; Guo, Z.; Cieslak, J.; Henry, D. ADP-based intelligent tracking algorithm for reentry vehicles subjected to model and state uncertainties. *IEEE Trans. Ind. Informat.* **2022**, *in press*. [\[CrossRef\]](#)
34. Yang, Y.; Gao, W.; Modares, H.; Xu, C.Z. Robust actor-critic learning for continuous-time nonlinear systems with unmodeled dynamics. *IEEE Trans. Fuzzy Syst.* **2021**, *30*, 2101–2112. [\[CrossRef\]](#)
35. El-Sousy, F.F.; Amin, M.M.; Al-Durra, A. Adaptive optimal tracking control via actor-critic-identifier based adaptive dynamic programming for permanent-magnet synchronous motor drive system. *IEEE Trans. Ind. Appl.* **2021**, *57*, 6577–6591. [\[CrossRef\]](#)
36. Dierks, T.; Jagannathan, S. Optimal Control of Affine Nonlinear Continuous-Time Systems. In Proceedings of the 2010 American Control Conference, Baltimore, MD, USA, 30 June–2 July 2010; pp. 1568–1573.
37. Xue, S.; Luo, B.; Liu, D. Event-triggered adaptive dynamic programming for zero-sum game of partially unknown continuous-time nonlinear systems. *IEEE Trans. Syst. Man Cybern. Syst.* **2018**, *50*, 3189–3199. [\[CrossRef\]](#)
38. Wang, D.; Mu, C.; Liu, D.; Ma, H. On mixed data and event driven design for adaptive-critic-based nonlinear H_∞ control. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *29*, 993–1005. [\[CrossRef\]](#)
39. Yang, H.; Hu, Q.; Dong, H.; Zhao, X. ADP-based spacecraft attitude control under actuator misalignment and pointing constraints. *IEEE Trans. Ind. Electron.* **2017**, *69*, 9342–9352. [\[CrossRef\]](#)
40. Wang, J.; Zhang, Z.T.; Tian, B.L.; Zong, Q. Event-based robust optimal consensus control for nonlinear multiagent system with local adaptive dynamic programming. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *in press*. [\[CrossRef\]](#)
41. Song, R.; Lewis, F.L. Robust optimal control for a class of nonlinear systems with unknown disturbances based on disturbance observer and policy iteration. *Neurocomputing* **2020**, *390*, 185–195. [\[CrossRef\]](#)
42. Pham, T.L.; Dao, P.N. Disturbance observer-based adaptive reinforcement learning for perturbed uncertain surface vessels. *ISA Trans.* **2022**, *130*, 277–292.
43. Dong, H.; Yang, X. Learning-based online optimal sliding-mode control for space circumnavigation missions with input constraints and mismatched uncertainties. *Neurocomputing* **2022**, *484*, 13–25. [\[CrossRef\]](#)
44. Zhang, H.; Park, J.H.; Yue, D.; Zhao, W. Nearly optimal integral sliding-mode consensus control for multiagent systems with disturbances. *IEEE Trans. Syst. Man Cybern. Syst.* **2019**, *51*, 4741–4750. [\[CrossRef\]](#)
45. Xia, R.; Wu, Q.; Shao, S. Disturbance observer-based optimal flight control of near space vehicle with external disturbance. *Trans. Inst. Meas. Control.* **2020**, *42*, 272–284. [\[CrossRef\]](#)
46. Vamvoudakis, K.G.; Lewis, F.L. Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem. *Automatica* **2010**, *46*, 878–888. [\[CrossRef\]](#)
47. Shao, X.; Shi, Y.; Zhang, W. Input-and-measurement event-triggered output-feedback chattering reduction control for MEMS gyroscopes. *IEEE Trans. Syst. Man Cybern. Syst.* **2021**, *52*, 5579–5590. [\[CrossRef\]](#)
48. Zuo, Z. Non-singular fixed-time terminal sliding mode control of non-linear systems. *IET Control. Theory Appl.* **2015**, *9*, 545–552. [\[CrossRef\]](#)
49. Cruz-Zavala, E.; Moreno, J.A.; Fridman, L.M. Uniform robust exact differentiator. *IEEE Trans. Autom. Control.* **2011**, *56*, 2727–2733. [\[CrossRef\]](#)
50. Na, J.; Lv, Y.; Zhang, K.; Zhao, J. Adaptive identifier-critic-based optimal tracking control for nonlinear systems with experimental validation. *IEEE Trans. Syst. Man Cybern. Syst.* **2020**, *52*, 459–472. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.