

Article

# Robust Approximate Optimal Trajectory Tracking Control for Quadrotors

Rong Li <sup>1,2,\*</sup> , Zhengliang Yang <sup>1</sup>, Gaowei Yan <sup>1</sup>, Long Jian <sup>1</sup> , Guoqiang Li <sup>2</sup> and Zhiqiang Li <sup>3</sup><sup>1</sup> College of Electrical and Power Engineering, Taiyuan University of Technology, Taiyuan 030024, China<sup>2</sup> Nuclear Emergency and Nuclear Safety Department, China Institute for Radiation Protection, Taiyuan 030006, China<sup>3</sup> College of Mechanical and Vehicle Engineering, Taiyuan University of Technology, Taiyuan 030024, China

\* Correspondence: lirong@tyut.edu.cn

**Abstract:** This paper uses the adaptive dynamic programming (ADP) method to achieve optimal trajectory tracking control for quadrotors. Relying on an established mathematical model of a quadrotor, the approximate optimal trajectory tracking control, which consists of the steady-state control input and the approximate optimal feedback control input, is designed for a nominal system. Considering the compound disturbances in position and attitude dynamic models, disturbance observers are introduced. The estimated values are used to design robust compensation inputs to suppress the effect of the compound disturbances for good trajectory tracking performance. Theoretically, the Lyapunov theorem demonstrates the stability of a closed-loop system. The robustness and effectiveness of the proposed controller are confirmed by the simulation results.

**Keywords:** ADP; quadrotor; trajectory tracking control; disturbance observer



**Citation:** Li, R.; Yang, Z.; Yan, G.; Jian, L.; Li, G.; Li, Z. Robust Approximate Optimal Trajectory Tracking Control for Quadrotors. *Aerospace* **2024**, *11*, 149. <https://doi.org/10.3390/aerospace11020149>

Academic Editor: Karim Abu Salem

Received: 21 December 2023

Revised: 5 February 2024

Accepted: 7 February 2024

Published: 13 February 2024



**Copyright:** © 2024 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

The miniaturization and reduction in cost of the relevant control components in aircraft, as well as the development and progress of computer and sensing and measurement technologies, have improved the stability of flight control systems and greatly facilitated the development of quadrotors [1]. The high operability, strong mobility and flexibility of quadrotors allow them to meet the specific needs of many projects, generally used in military, industrial and other fields [2,3]. A quadrotor system is multivariable, nonlinear and strongly coupled, and quadrotors will also be disturbed by the surrounding environment during flight [4]. These factors can affect the accuracy of quadrotor control systems. The requirements for high-accuracy and robust flight control in the design of controllers for quadrotors are stringent, and the design of a core control algorithm is a prerequisite for quadrotors to achieve a stable and high-precision flight performance. Therefore, the research and development of controllers for quadrotor systems is of great significance.

At present, it is no longer a problem to ensure the uniformity of quadrotors through control algorithms. Many controllers for quadrotors have been designed and are already in application [5]. Since the dynamics of quadrotors can be linearized around the equilibrium point, traditional linear control methods are used for a designed controller [6]. On this basis, linear techniques are employed in the flight control of quadrotors, such as linear quadratic regulator (LQR) control [7]. However, quadrotors need to be controlled away from the equilibrium point to accomplish complex control tasks and withstand external disturbances. As a result, a technique has been devised that is regarded as a robust feedback linearization method that uses extended state observers to estimate the nonlinear state feedback term online, containing aerodynamic forces, moments and unknown disturbances, and obtains the desired closed-loop dynamics via pole assignment [8]. Moreover, several robust controllers relying on nonlinear techniques have been proposed, such as sliding mode control [9], adaptive control [10], backstepping-based control [11] and robust control [12].

These control methods ensure the stability and robustness of nonlinear systems and have generally been used for the tracking control of these systems, but their optimal properties have not been considered. Therefore, the concept of optimization has been introduced into control design.

To derive the optimal control policy for the infinite horizon optimal control problem, solving the Hamilton–Jacobi–Bellman (HJB) equation or the Hamilton–Jacobi–Isaacs (HJI) equation for the  $H_\infty$  optimal control problem considering uncertainties is essential. Nevertheless, it is difficult to mathematically derive the corresponding analytical solutions in most cases. Neural networks are an optional method to overcome this problem [13–15]. The approximation property of neural networks makes it possible to find approximate solutions to partial differential equations. The convergence of neural networks can be ensured by penalizing them to ensure they satisfy the given partial differential equations. The ADP method is the combination of reinforcement learning, dynamic programming and neural network adaptive methods to derive approximate solutions of the HJB/HJI equations using function approximate structures to address nonlinear optimal control problems [16,17]. The ADP method is used for control design with suitable performance index functions to derive the desired dynamic performance and stabilize a nominal system with uncertainties. However, most nonlinear optimal control methods using the ADP method are aimed at nominal systems or uncertain systems satisfying specific conditions [18–20], while the immunity to disturbances is still weak for such systems with external time-varying disturbances independent of the state, and the control effect under stronger disturbances is not ideal. The ADP method has been used in the design of controllers for quadrotors and efforts have been made to improve the robustness, but the designed controllers are more geared towards linear systems and design uncertainty is a unique problem [21,22]. Quadrotors will often experience various external effects in flight, requiring strong adaptive and anti-disturbance capabilities in flight control. The disturbance observer technique achieves disturbance suppression of the target utilizing feedback regulation [23], which can attenuate compound disturbances containing external disturbances and model uncertainties, thus improving the system robustness. A disturbance observer can accurately estimate compound disturbances in a system, which greatly reduces the conservatism of the control. In addition, since a disturbance observer can usually be designed independently of the controller, this ensures that the method can be easily combined with other advanced control methods and more flexible in its application. There are experiments suggesting that the introduction of a disturbance observer significantly improves performance, which is a good reference for methods for quadrotors to overcome disturbances [24].

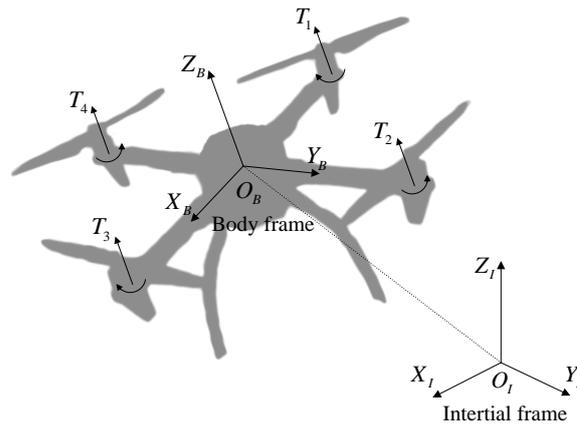
Considering the above analysis, a robust approximate optimal trajectory tracking control method is proposed for quadrotors to solve the optimal control problem under the conditions of compound disturbances. The main contributions are summarized as follows:

- (1) The combination of modeling uncertainties and external time-varying disturbances is considered as compound disturbances. Disturbance observers are introduced to estimate the compound disturbances in the position and attitude subsystems, and the estimated values are used to design robust compensation inputs to suppress the effects of the compound disturbances and ensure the stability of a quadrotor system under the ADP method.
- (2) To obtain optimal trajectory tracking control for a quadrotor without composite disturbances, the ADP method is used to design approximate optimal control inputs for the nominal system of a quadrotor.

The rest of the paper is organized as follows. In Section 2, the quadrotor mathematical model is developed, and the quadrotor system is divided into two subsystems. Section 3 describes the design of the robust approximate optimal trajectory tracking control and the stability analysis of the closed-loop system. Section 4 describes the robust approximate optimal trajectory tracking control for the quadrotor. The results of the corresponding simulation and the results of the comparative simulation without disturbance observer are presented in Section 5. Section 6 gives the conclusion of the paper.

## 2. Mathematical Modeling of a Quadrotor

The quadrotor has four evenly spaced, cross-symmetrical brushless motors in the plane. The rotors of motor 1 and motor 3 rotate clockwise, while the rotors of motor 2 and motor 4 rotate counterclockwise. By changing the rotational speed of the four rotors, the quadrotor generates different magnitudes of lift forces and torques, which can control the takeoff, landing and attitude motions of the quadrotor. As a result, the location of the quadrotor can be altered in the three-dimensional space. Figure 1 depicts the basic structure of the quadrotor.



**Figure 1.** Basic structure of the quadrotor.

To clarify the mathematical model of the quadrotor system and satisfy the implementation of the control method, the earth-fixed inertial frame  $O_I X_I Y_I Z_I$  and the body-fixed body frame  $O_B X_B Y_B Z_B$  are established. To ensure that the constructed mathematical model does not lose the generality, it is assumed that the deformation and elastic vibration properties of the rotors and body are neglected, and the quadrotor is considered as an ideal rigid body; the quadrotor's structure is symmetrical, its mass is uniformly distributed, and its center of mass is located at the geometric center. The translational and rotational motions of the quadrotor are satisfied by [25]

$$\dot{P} = v, \quad (1)$$

$$\dot{\Theta} = W_B^I \omega, \quad (2)$$

where  $P = [x, y, z]^T \in \mathbb{R}^3$  represents the position of the quadrotor in the inertial frame and  $v = [v_x, v_y, v_z]^T \in \mathbb{R}^3$  represents the corresponding velocity.  $\Theta = [\phi, \theta, \psi]^T \in \mathbb{R}^3$  denotes the vector of Euler angles.  $\omega = [p, q, r]^T \in \mathbb{R}^3$  denotes the angular velocity of the quadrotor in the body frame.  $W_B^I \in \mathbb{R}^{3 \times 3}$  is the rotation matrix for the angular velocity in the form of

$$W_B^I = \begin{bmatrix} 1 & s_\phi t_\theta & c_\phi t_\theta \\ 0 & c_\phi & -s_\phi \\ 0 & s_\phi / c_\theta & c_\phi / c_\theta \end{bmatrix}, \quad (3)$$

in which  $s_* = \sin(*)$ ,  $c_* = \cos(*)$  and  $t_* = \tan(*)$ .

Relying on the Newton–Euler method, the dynamical equation of quadrotor with compound disturbances is represented by [26]

$$m\dot{v} = F - kv + d_p, \quad (4)$$

$$I\dot{\omega} + \omega \times I\omega = \tau + d_a, \quad (5)$$

where  $m \in \mathbb{R}$  represents the mass of the quadrotor.  $I \in \mathbb{R}^{3 \times 3}$  represents the inertia matrix of the quadrotor. As the assumptions of the quadrotor structure, its inertia matrix can be defined as the diagonal array  $I \triangleq \text{diag}\{I_{xx}, I_{yy}, I_{zz}\}$ .  $k = \text{diag}\{k_x, k_y, k_z\} \in \mathbb{R}^{3 \times 3}$  is the

drag coefficient matrix.  $F = [F_x, F_y, F_z]^T \in \mathbb{R}^3$  represents the resultant force consisting of the gravity and the total lift in the inertial frame.  $\tau = [\tau_x, \tau_y, \tau_z]^T \in \mathbb{R}^3$  represents the torque in the body frame.  $d_p \in \mathbb{R}^3$  and  $d_a \in \mathbb{R}^3$  are the compound disturbances in position and attitude dynamic models, which contain modeling uncertainties and external time-varying disturbances.

According to the mechanical analysis, the quadrotor is affected by gravity and lift forces. Since the special structure of the quadrotor, the lift forces are along the z-axis direction of the body frame. Then, the resultant force expressed in the inertial frame is [27]

$$F = R_B^I z_u T - z_u m g_I, \quad (6)$$

where  $T \in \mathbb{R}$  represents the total lift force and  $g_I \in \mathbb{R}$  represents the gravity acceleration.  $z_u = [0, 0, 1]^T$ .  $R_B^I \in \mathbb{R}^{3 \times 3}$  is the rotation matrix of the body frame transformed into the inertial frame in the form of

$$R_B^I = \begin{bmatrix} c_\theta c_\psi & s_\phi s_\theta c_\psi - c_\phi s_\psi & c_\phi s_\theta c_\psi + s_\phi s_\psi \\ c_\theta s_\psi & s_\phi s_\theta s_\psi + c_\phi c_\psi & c_\phi s_\theta s_\psi - s_\phi c_\psi \\ -s_\theta & s_\phi c_\theta & c_\phi c_\theta \end{bmatrix}. \quad (7)$$

**Assumption 1** ([28]). The pitch and roll angles hold the conditions  $|\phi| < \pi/2$  and  $|\theta| < \pi/2$  to avoid the singularities of the matrices  $W_B^I$  and  $R_B^I$ .

**Assumption 2** ([29]). In the control process, the total compound disturbance  $d_t = [d_p^T, d_a^T]^T \in \mathbb{R}^6$  has finite energy. In addition,  $d_t$  is a continuous function and its norm is bounded such that  $\|d_t\| \leq d_{tM}$ , where  $d_{tM}$  is an unknown positive constant. Simultaneously, the compound disturbances are usually considered to be superimposed by the low-frequency period signals. Hence, it is assumed that the total compound disturbance has a low change rate and its rate of change is slow compared to the dynamic properties of the disturbance observer, which can be considered that  $\dot{d}_t \simeq 0$ .

**Assumption 3** ([30]). The desired trajectory of position  $P_d = [x_d, y_d, z_d]^T \in \mathbb{R}^3$  and the desired trajectory of yaw angle  $\psi_d \in \mathbb{R}$  and their higher order derivatives are known, continuous and bounded.

**Remark 1.** Assumption 2 is common in control studies using disturbance observers [31–33], while there are different considerations for compound disturbances in [34]. In the case of this paper, the considerations in Assumption 2 are used. Assumption 3 ensures that the ADP method can be utilized for the control design and the stability analysis.

The total lift and torque of the quadrotor are related to the force and torque of the four rotors as follows [35]:

$$\begin{cases} T = T_1 + T_2 + T_3 + T_4 \\ \tau_x = l(T_2 - T_4) \\ \tau_y = l(T_1 - T_3) \\ \tau_z = \tau_1 - \tau_2 + \tau_3 - \tau_4 \end{cases}, \quad (8)$$

where  $T_i$  and  $\tau_i$  ( $i = 1, 2, 3, 4$ ) are the lift and torque generated by the four rotors of the quadrotor, respectively.  $l$  represents the length from each rotor to the center of the body.

The rotor speeds are related to pulse-width modulated (PWM) signals through the motors. The lift forces and torques generated by the four motors are related to the pulse width of the input signals as follows [36]:

$$\begin{cases} T_i = K_t \frac{B_w}{s+B_w} u_i \\ \tau_i = K_o \frac{B_w}{s+B_w} u_i \end{cases}, \quad (9)$$

where  $K_t$  and  $K_o$  are the positive gains of the lift coefficient and the inverse torque coefficient, respectively.  $B_w$  is the motor bandwidth and  $u_i$  represents the PWM signals of each corresponding motor, which should be limited between 0 and 1.

Assuming that the motors have a sufficiently fast response speed, then the motor model can be simplified as [37]

$$\begin{cases} T_i = K_t u_i \\ \tau_i = K_o u_i \end{cases} \quad (10)$$

Hence, (8) can be rewritten as [38]

$$\begin{bmatrix} T \\ \tau_x \\ \tau_y \\ \tau_z \end{bmatrix} = \begin{bmatrix} K_t & K_t & K_t & K_t \\ 0 & K_t l & 0 & -K_t l \\ K_t l & 0 & -K_t l & 0 \\ K_o & -K_o & K_o & -K_o \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} \quad (11)$$

Considering the trajectory tracking control for the quadrotor, the control objective is to design a controller that allows the position and attitude to track the desired trajectory asymptotically within a small error.

Combining (1), (2), (4) and (5), the overall model of the quadrotor can be decomposed into a position subsystem and an attitude subsystem. The position subsystem can be represented as

$$\dot{x}_1 = f_1(x_1) + g_1(x_1)(F + d_p), \quad (12)$$

with

$$\begin{aligned} x_1 &= [P^T, v^T]^T = [x, y, z, v_x, v_y, v_z]^T \in \mathbb{R}^6, \\ f_1(x_1) &= [v_x, v_y, v_z, -k_x v_x/m, -k_y v_y/m, -k_z v_z/m]^T \in \mathbb{R}^6, \\ g_1(x_1) &= \begin{bmatrix} 0 & 0 & 0 & 1/m & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/m & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/m \end{bmatrix}^T \in \mathbb{R}^{6 \times 3}. \end{aligned}$$

While the attitude subsystem is expressed in the form of

$$\dot{x}_2 = f_2(x_2) + g_2(x_2)(\tau + d_a), \quad (13)$$

with

$$\begin{aligned} x_2 &= [\Theta^T, \omega^T]^T = [\phi, \theta, \psi, p, q, r]^T \in \mathbb{R}^6, \\ f_2(x_2) &= [p + qs_\phi t_\theta + rc_\phi t_\theta, qc_\phi - rs_\phi, qs_\phi/c_\theta + rc_\phi/c_\theta, \\ &\quad qr(I_{yy} - I_{zz})/I_{xx}, pr(I_{zz} - I_{xx})/I_{yy}, pq(I_{xx} - I_{yy})/I_{zz}]^T \in \mathbb{R}^6, \\ g_2(x_2) &= \begin{bmatrix} 0 & 0 & 0 & 1/I_{xx} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/I_{yy} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/I_{zz} \end{bmatrix}^T \in \mathbb{R}^{6 \times 3}. \end{aligned}$$

In the next section, (12) and (13) will be the focus of our research.

### 3. Robust Approximate Optimal Trajectory Tracking Control Design

Considering the convenience of describing the control design process, (12) and (13) is represented in the uniform form

$$\dot{X} = f(X) + g(X)(U + D), \quad (14)$$

in which  $f(X) \in \mathbb{R}^6$  and  $g(X) \in \mathbb{R}^{6 \times 3}$  represent the drift dynamics and the input dynamics of the system, respectively.  $X \in \mathbb{R}^6$  denotes the observable state vector,  $U \in \mathbb{R}^3$  denotes the control input, and  $D \in \mathbb{R}^3$  denotes the compound disturbance.

**Definition 1** ([39]). A state vector  $\mathcal{X}$  is said to be uniformly ultimately bounded (UUB) if there exists a compact set  $\mathcal{O}_{\mathcal{X}}$ , a positive number  $b_{\mathcal{X}}$  and a time  $t_b(\mathcal{X}(t_0), b_{\mathcal{X}})$  such that  $\|\mathcal{X}\| \leq b_{\mathcal{X}}$  for all state variable initial value  $\mathcal{X}(t_0) \in \mathcal{O}_{\mathcal{X}}$  and all  $t \geq t_0 + t_b$ .

**Lemma 1** ([40]).  $\mathcal{X}$  is UUB if the time derivative of a positive definite function  $L_{\mathcal{X}}(\mathcal{X})$  is negative when  $\|\mathcal{X}\| > b_{\mathcal{X}}$  for a positive constant  $b_{\mathcal{X}}$ .

To realize the trajectory tracking control with robustness for the system, the designed controller consists of two parts, the form of which is as follows:

$$U = U^N + U^R, \quad (15)$$

where  $U^R$  is the robust compensation input designed through the disturbance observer for suppressing the effect of compound disturbances in the system.  $U^N$  is the control input designed based on the ADP method for the nominal system, which takes the form of

$$U^N = U_d + U_E, \quad (16)$$

where  $U_d$  represents the steady-state control input and  $U_E$  represents the feedback control input.

### 3.1. Disturbance Observer Design

The disturbance observer is applied to derive the estimate of the compound disturbance. The estimated value is then used for the design of the robust compensation input to improve robustness. The disturbance observer is designed as

$$\begin{cases} \dot{Z} = -l_D(X)(f(X) + g(X)(p_D(X) + U + Z)) \\ \dot{\hat{D}} = Z + p_D(X) \end{cases}, \quad (17)$$

in which  $\hat{D} \in \mathbb{R}^3$  represents the estimate of the unknown compound disturbance,  $p_D(X) \in \mathbb{R}^3$  represents the designed vector-valued function,  $l_D(X) = \partial p_D(X) / \partial X \in \mathbb{R}^{3 \times 6}$  is the observer gain and  $Z \in \mathbb{R}^3$  represents the auxiliary variable vector of the disturbance observer.

**Remark 2.** In the disturbance observer (17), the derivative of the state is required, which is unknown because the compound disturbance is unknown. Then, the auxiliary variable vector is given to avoid calculating the derivative of the state.

Define the estimation error of compound disturbance as  $\tilde{D} = D - \hat{D}$ . With regard to Assumption 2 and the disturbance observer (17), the time derivative of  $\tilde{D}$  is developed as

$$\begin{aligned} \dot{\tilde{D}} &= \dot{D} - \dot{\hat{D}} = -\dot{Z} - l_D(X)\dot{X} \\ &= l_D(X)(f(X) + g(X)(p_D(X) + U + Z)) - l_D(X)\dot{X} \\ &= l_D(X)g(X)(Z + p_D(X)) - l_D(X)(\dot{X} - f(X) - g(X)U). \end{aligned} \quad (18)$$

Combined with (14), we have

$$\begin{aligned} \dot{\tilde{D}} &= l_D(X)g(X)(Z + p_D(X)) - l_D(X)g(X)D \\ &= -l_D(X)g(X)(D - \hat{D}) \\ &= -l_D(X)g(X)\tilde{D}. \end{aligned} \quad (19)$$

Then,  $\tilde{D}$  is convergent by appropriately designing the vector-valued function  $p_D(X)$ .

**Theorem 1.** Considering System (14), the disturbance observer is designed as (17). If  $l_D(X)g(X)$  is ensured to be positive definite for the design of the vector-valued function  $p_D(X)$ , then the

estimated compound disturbance  $\hat{D}$  would follow the compound disturbance  $D$ , which means the estimation error  $\tilde{D}$  could converge to zero.

**Proof.** Select the candidate Lyapunov function as follows:

$$L_D = \frac{1}{2} \tilde{D}^T \tilde{D}. \quad (20)$$

Combined with (18), the time derivative of  $L_D$  is

$$\dot{L}_D = \tilde{D}^T \dot{\tilde{D}} = -\tilde{D}^T l_D(X) g(X) \tilde{D}. \quad (21)$$

In the case where  $l_D(X)g(X)$  is positive definite, then we derive

$$\dot{L}_D \leq -\kappa \|\tilde{D}\|^2, \quad (22)$$

where  $\kappa = \lambda_{\min}(l_D(X)g(X))$  and  $\lambda_{\min}(\ast)$  denotes the minimum eigenvalue. Obviously,  $\dot{L}_D < 0$  when  $\tilde{D} \neq 0$ . Hence, the disturbance observer (17) can estimate  $D$  and  $\tilde{D}$  will converge to zero. This completes the proof.  $\square$

Then, the robust compensation input  $U^R$  is designed as

$$U^R = -\hat{D}. \quad (23)$$

### 3.2. Optimal Trajectory Tracking Control Design and Analysis

The compound disturbance is estimated by the disturbance observer. The robust compensation input is designed by the estimated value to suppress the effect of the compound disturbances. As a result, converting the trajectory tracking control problem of the nonlinear system with the compound disturbance into the trajectory tracking control problem of the nominal system is possible. In order to derive the optimal control for the nominal system, deriving the solution of the associated HJB equation is essential. Unfortunately, deriving the analytical solution is difficult for the nonlinear system by the direct solution method. Then, the ADP method is utilized for achieving the approximate optimal control by constructing the critic network. The weight update law designed for the critic network ensures the convergence of the weight and the stability of the closed-loop system.

For System (14), the nominal system is represented by

$$\dot{X} = f(X) + g(X)U. \quad (24)$$

Given the desired trajectory  $X_d \in \mathbb{R}^6$ , the steady-state control input  $U_d$  is obtained from (24) as

$$U_d = g^+(X_d)(\dot{X}_d - f(X_d)), \quad (25)$$

in which  $g^+(X_d)$  denotes the pseudo-inverse of  $g(X_d)$ .

Define the tracking error as  $E = X - X_d \in \mathbb{R}^6$ . Combined with (14) and (15), the error system is developed as

$$\begin{aligned} \dot{E} &= f(X) + g(X)(U + D) - \dot{X}_d \\ &= f(E + X_d) + g(E + X_d)U_d - \dot{X}_d + g(E + X_d)U_E + g(E + X_d)\tilde{D}. \end{aligned} \quad (26)$$

Let  $f_E = f(E + X_d) + g(E + X_d)U_d - \dot{X}_d$  and  $g_E = g(E + X_d)$ , then we have

$$\dot{E} = f_E + g_E U_E + g_E \tilde{D}. \quad (27)$$

Noting that  $g_E = g(X)$ , the norm of  $g_E$  is bounded such that  $g_m \leq \|g_E\| \leq g_M$  for the positive constants  $g_m$  and  $g_M$ .

As a result of Theorem 1, the disturbance observer (20) can successfully estimate the compound disturbance  $D$  and the estimation error of compound disturbance  $\tilde{D}$  can

converge to zero. Therefore, it is possible to neglect  $\tilde{D}$  in the error system (27) for the optimal control design [41,42]. However,  $\tilde{D}$  would still be considered in the stability analysis. Then, the nominal error system is represented by

$$\dot{E} = f_E + g_E U_E. \quad (28)$$

Define the cost function as

$$V(E) = \int_{t_0}^{\infty} (E^T Q E + U_E^T R U_E) dt, \quad (29)$$

where  $Q \in \mathbb{R}^{6 \times 6}$  and  $R \in \mathbb{R}^{3 \times 3}$  are the designed symmetric positive definite matrices.

The nonlinear Lyapunov equation for (29) is achieved as

$$\nabla V^T (f_E + g_E U_E) + E^T Q E + U_E^T R U_E = 0, \quad (30)$$

where  $\nabla V = \partial V(E)/\partial E$  and  $V(0) = 0$ .

**Definition 2** ([43]). A control policy  $\mu(E)$  is said to be admissible on the compact set  $\mathcal{O}$  for (29) if  $\mu(E)$  is continuous on  $\mathcal{O}$ ,  $\mu(0) = 0$ ,  $\mu(E)$  stabilizes (28) on  $\mathcal{O}$  and  $V(E)$  is finite  $\forall E \in \mathcal{O}$ . This is represented by  $\mu(E) \in \Psi(\mathcal{O})$ , where  $\Psi(\mathcal{O})$  denotes the set of admissible control policies.

The Hamiltonian function takes the following form

$$H(E, U_E, \nabla V) = \nabla V^T (f_E + g_E U_E) + E^T Q E + U_E^T R U_E. \quad (31)$$

The optimal cost function is represented by

$$V^*(E) = \min_{U_E \in \Psi(\mathcal{O})} \int_{t_0}^{\infty} (E^T Q E + U_E^T R U_E) dt, \quad (32)$$

and the following relation is satisfied

$$\min_{U_E \in \Psi(\mathcal{O})} H(E, U_E, \nabla V^*) = 0, \quad (33)$$

where  $\nabla V^* = \partial V^*(E)/\partial E$ .

Under the existence condition of the optimal solution  $\partial H(E, U_E^*, \nabla V^*)/\partial U_E^* = 2R U_E^* + g_E^T \nabla V^* = 0$ , the optimal feedback control input is derived by

$$U_E^* = -\frac{1}{2} R^{-1} g_E^T \nabla V^*. \quad (34)$$

Substituting (34) and (31) into (33), the HJB equation is developed as

$$\nabla V^{*T} f_E + E^T Q E - \frac{1}{4} \nabla V^{*T} g_E R^{-1} g_E^T \nabla V^* = 0. \quad (35)$$

### 3.3. Approximate Optimal Control Design

Clearly, it is necessary to derive  $\nabla V^*$  by solving the HJB Equation (35) for deriving the optimal feedback control input (34). However, (35) is a typical nonlinear partial differential equation and its solution is difficult to derive in the analytic form [44,45]. To overcome the difficulty, the ADP method relying on the policy iteration technique is utilized to derive the approximate solution.

**Assumption 4** ([46]). The continuously differentiable Lyapunov function candidate  $J(E)$  for the nominal error system (28) satisfies  $\nabla J^T (f_E + g_E U_E^*) < 0$ , where  $\nabla J = \partial J(E)/\partial E$ . Meanwhile, there exists a symmetric positive definite matrix  $\Lambda(E)$  such that  $\nabla J^T (f_E + g_E U_E^*) =$

$-\nabla J^T \Lambda(E) \nabla J$ . Moreover, the relation  $\Lambda_m \leq \|\Lambda(E)\| \leq \Lambda_M$  holds for positive constants  $\Lambda_m, \Lambda_M$ .

**Remark 3.** Assumption 4 is a common assumption that has been used for the ADP method. Generally, it is assumed that the closed-loop dynamics with the optimal feedback control is bounded by a function of the system state on the compact set. In such a situation, there exists a positive constant  $\eta$  such that  $\|f_E + g_E U_E^*\| \leq \eta \|\nabla J\|$ . Hence, we can further derive  $\|\nabla J^T (f_E + g_E U_E^*)\| \leq \eta \|\nabla J\|^2$ . Furthermore, the function  $J(E)$  can be correctly selected as a quadratic polynomial [47], such as  $J(E) = \frac{1}{2} E^T E$ .

Considering the uniform estimation property of neural networks, the optimal cost function is approximated by

$$V^*(E) = W_c^T \varphi_c(E) + \varepsilon_c(E), \quad (36)$$

where  $W_c \in \mathbb{R}^N$  represents the unknown ideal constant weight,  $\varphi_c(E) \in \mathbb{R}^N$  represents the activation function,  $\varepsilon_c(E)$  represents the approximate error, and  $N$  represents the number of neurons. This neural network is called the critic network in the ADP method.

**Lemma 2** ([48]). The estimation error  $\varepsilon_c(E)$  is expected to be bounded when the approximated function  $V^*(E)$  is bounded.

Then, by the definition of  $\nabla V^*$ , it is developed as follows

$$\nabla V^* = \nabla \varphi_c^T W_c + \nabla \varepsilon_c, \quad (37)$$

where  $\nabla \varphi_c = \partial \varphi_c(E) / \partial E$  and  $\nabla \varepsilon_c = \partial \varepsilon_c(E) / \partial E$ .

Invoking (37), the optimal feedback control input (34) is developed as

$$U_E^* = -\frac{1}{2} R^{-1} g_E^T (\nabla \varphi_c^T W_c + \nabla \varepsilon_c). \quad (38)$$

Substituting (37) into (35), the HJB equation is developed as

$$W_c^T \nabla \varphi_c f_E + E^T Q E - \frac{1}{4} W_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T W_c + \varepsilon_H = 0, \quad (39)$$

where  $\Xi = g_E R^{-1} g_E^T$ .  $\varepsilon_H$  represents the residual error, which takes the form of

$$\begin{aligned} \varepsilon_H &= \nabla \varepsilon_c^T f_E - \frac{1}{2} \nabla \varepsilon_c^T \Xi \nabla \varphi_c^T W_c - \frac{1}{4} \nabla \varepsilon_c^T \Xi \nabla \varepsilon_c \\ &= \nabla \varepsilon_c^T (f_E + g_E U_E^*) + \frac{1}{4} \nabla \varepsilon_c^T \Xi \nabla \varepsilon_c. \end{aligned} \quad (40)$$

Since  $\|g_E\|$  is bounded, there exists the positive constants  $\Xi_m$  and  $\Xi_M$  such that  $\Xi_m \leq \|\Xi\| \leq \Xi_M$ .

Define the estimate of  $W_c$  as  $\hat{W}_c$ , then the estimate of  $V^*(E)$  is derived as follows:

$$\hat{V}(E) = \hat{W}_c^T \varphi_c(E). \quad (41)$$

Moreover, the approximate optimal feedback control input is derived as

$$U_E = -\frac{1}{2} R^{-1} g_E^T \nabla \varphi_c^T \hat{W}_c. \quad (42)$$

**Remark 4.** The classical ADP method utilizes the critic network and the actor network to approximate the optimal cost function and the optimal feedback control, respectively [43,49,50]. Considering the association between the optimal cost function and the optimal feedback control for

the continuous affine nonlinear system, it is possible to omit the actor network and only use the critic network [51,52]. This framework provides smaller computational effort, faster convergence and compared to the actor–critic network framework, which has a better practical value.

Combining (31), (41) and (42), the approximate Hamiltonian function is developed as

$$H(E, \hat{W}_c) = \hat{W}_c^T \nabla \varphi_c f_E + E^T Q E - \frac{1}{4} \hat{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \hat{W}_c \triangleq e_c. \tag{43}$$

Define the objective function as

$$E_c = \frac{1}{2} e_c^2. \tag{44}$$

Moreover, the weight update law is designed as

$$\dot{\hat{W}}_c = -\frac{\alpha_1 \sigma}{\sigma_c^2} (\hat{W}_c^T \nabla \varphi_c f_E + E^T Q E - \frac{1}{4} \hat{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \hat{W}_c) + \frac{\alpha_2}{2} \Pi(E, U_E) \nabla \varphi_c \Xi \nabla J, \tag{45}$$

where  $\alpha_1 > 0$ ,  $\alpha_2 > 0$  are the learning rates to be designed.  $\sigma = \nabla \varphi_c(f_E + g_E U_E)$  and  $\sigma_c = \sigma^T \sigma + 1$ .  $\nabla J$  is given in Assumption 4.  $\Pi(E, U_E)$  in the last term is defined as

$$\Pi(E, U_E) = \begin{cases} 0, & \text{if } \nabla J^T (f_E + g_E U_E) + \alpha_3 \nabla J^T g_E g_E^T \nabla J < 0, \\ 1, & \text{else} \end{cases}, \tag{46}$$

where  $\alpha_3$  is a designed positive constant.

**Remark 5.** The first term in (45) is employed for minimizing the objective function (44). To ensure that  $\hat{W}_c$  will converge to  $W_c$ , the existence of the persistence of excitation (PE) condition is essential during the learning process is necessary [49]. In addition, the probing noise is typically introduced to the control input for satisfying this condition, which may enable the closed-loop system to become unstable during the learning process [53,54]. The second term in (45) is employed for the stability of the closed-loop system.

Define the weight estimation error as  $\tilde{W}_c = W_c - \hat{W}_c$ . Observing that  $\dot{\tilde{W}}_c = -\dot{\hat{W}}_c$ ,  $\sigma = \nabla \varphi_c(f_E + g_E U_E) = \nabla \varphi_c \dot{E}^* + \frac{1}{2} \nabla \varphi_c \Xi \nabla \varepsilon_c + \frac{1}{2} \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c$  where  $\dot{E}^* = f_E + g_E U_E^*$ , and using (39) and (45), we have

$$\begin{aligned} \dot{\tilde{W}}_c = & -\frac{\alpha_1}{\sigma_c^2} (\nabla \varphi_c \dot{E}^* + \frac{1}{2} \nabla \varphi_c \Xi \nabla \varepsilon_c + \frac{1}{2} \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c) (\tilde{W}_c^T \nabla \varphi_c \dot{E}^* \\ & + \frac{1}{2} \tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varepsilon_c + \frac{1}{4} \tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c + \varepsilon_H) - \frac{\alpha_2}{2} \Pi(E, U_E) \nabla \varphi_c \Xi \nabla J. \end{aligned} \tag{47}$$

### 3.4. Stability Analysis

**Assumption 5** ([50]). The ideal weight  $W_c$  have bound over the compact set  $\mathcal{O}$  such that  $\|W_c\| \leq W_{cM}$  for a positive constant  $W_{cM}$ . Meanwhile, the activation function  $\varphi_c$  and the approximate error  $\varepsilon_c$  are bounded such that  $\|\varphi_c\| \leq \varphi_{cM}$ ,  $\|\varepsilon_c\| \leq \varepsilon_{cM}$  for positive constants  $\varphi_{cM}$  and  $\varepsilon_{cM}$ , and their derivatives are also bounded such that  $\|\nabla \varphi_c\| \leq \bar{\varphi}_{cM}$  and  $\|\nabla \varepsilon_c\| \leq \bar{\varepsilon}_{cM}$  for positive constants  $\bar{\varphi}_{cM}$  and  $\bar{\varepsilon}_{cM}$ . Moreover, the residual error  $\varepsilon_H$  will converge to zero when the number of neurons  $N$  is sufficiently large, as suggested by Remark 3 and the bound of  $\|\Xi\|$ . That is, the relation  $\|\varepsilon_H\| \leq \varepsilon_{HM}$  exists for the positive constant  $\varepsilon_{HM}$ .

**Theorem 2.** Considering System (14), the robust approximate optimal controller for the trajectory tracking control is designed as (15), which consists of the robust compensation input (23) and the nominal system control input (16), and the weight update law is designed as (45) for the critic network, then it is ensured that the tracking error  $E$  of the closed-loop system and the weight estimation error  $\tilde{W}_c$  are UUB.

**Proof.** Select the candidate Lyapunov function as follows

$$L = L_D + L_J + L_W, \tag{48}$$

where  $L_D$  is designed as (20),  $L_J = \alpha_2 J(E)$  and  $L_W = \frac{1}{2} \tilde{W}_c^T \tilde{W}_c$ .

Considering the second term in (48) and using (27), the time derivative is developed as

$$\dot{L}_J = \alpha_2 \nabla J^T (f_E + g_E U_E) + \alpha_2 \nabla J^T g_E \tilde{D}. \tag{49}$$

Considering the third term in (48) and according to (47), the time derivative is developed as

$$\begin{aligned} \dot{L}_W &= \tilde{W}_c^T \dot{\tilde{W}}_c \\ &= -\frac{\alpha_1}{\sigma_c^2} (\tilde{W}_c^T \nabla \varphi_c \dot{E}^* + \frac{1}{2} \tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varepsilon_c + \frac{1}{2} \tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c) (\tilde{W}_c^T \nabla \varphi_c \dot{E}^* \\ &\quad + \frac{1}{2} \tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varepsilon_c + \frac{1}{4} \tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c + \varepsilon_H) - \frac{\alpha_2}{2} \Pi(E, U_E) \tilde{W}_c^T \nabla \varphi_c \Xi \nabla J \\ &= -\frac{\alpha_1}{\sigma_c^2} (\tilde{W}_c^T \nabla \varphi_c \dot{E}^*)^2 - \frac{\alpha_1}{4\sigma_c^2} (\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varepsilon_c)^2 - \frac{\alpha_1}{8\sigma_c^2} (\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c)^2 \\ &\quad - \frac{\alpha_1}{\sigma_c^2} (\tilde{W}_c^T \nabla \varphi_c \dot{E}^*) (\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varepsilon_c) - \frac{3\alpha_1}{4\sigma_c^2} (\tilde{W}_c^T \nabla \varphi_c \dot{E}^*) (\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c) \\ &\quad - \frac{3\alpha_1}{8\sigma_c^2} (\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varepsilon_c) (\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c) - \frac{\alpha_1}{\sigma_c^2} \tilde{W}_c^T \nabla \varphi_c \dot{E}^* \varepsilon_H \\ &\quad - \frac{\alpha_1}{2\sigma_c^2} \tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varepsilon_c \varepsilon_H - \frac{\alpha_1}{2\sigma_c^2} \tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c \varepsilon_H - \frac{\alpha_2}{2} \Pi(E, U_E) \tilde{W}_c^T \nabla \varphi_c \Xi \nabla J. \end{aligned} \tag{50}$$

Since the first two terms in the final form of (50) are negative semi-definite, we then derive

$$\begin{aligned} \dot{L}_W &\leq -\frac{\alpha_1}{8\sigma_c^2} (\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c)^2 - \frac{\alpha_1}{\sigma_c^2} (\tilde{W}_c^T \nabla \varphi_c \dot{E}^*) (\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varepsilon_c) \\ &\quad - \frac{3\alpha_1}{4\sigma_c^2} (\tilde{W}_c^T \nabla \varphi_c \dot{E}^*) (\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c) \\ &\quad - \frac{3\alpha_1}{8\sigma_c^2} (\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varepsilon_c) (\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c) - \frac{\alpha_1}{\sigma_c^2} \tilde{W}_c^T \nabla \varphi_c \dot{E}^* \varepsilon_H \\ &\quad - \frac{\alpha_1}{2\sigma_c^2} \tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varepsilon_c \varepsilon_H - \frac{\alpha_1}{2\sigma_c^2} \tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c \varepsilon_H - \frac{\alpha_2}{2} \Pi(E, U_E) \tilde{W}_c^T \nabla \varphi_c \Xi \nabla J. \end{aligned} \tag{51}$$

According to Remark 3 and Assumption 5, and considering the bound of  $\|\Xi\|$ , we assume that  $\lambda_{1m} \leq \|\nabla \varphi_c \Xi \nabla \varphi_c^T\| \leq \lambda_{1M}$ ,  $\|\Xi\| \leq \lambda_2$ ,  $\|\nabla \varphi_c \dot{E}^*\| \leq \lambda_3$ ,  $\|\nabla \varepsilon_c\| \leq \lambda_4$ ,  $\|\nabla \varphi_c \Xi \nabla \varepsilon_c\| \leq \lambda_5$  and  $\|\varepsilon_H\| \leq \lambda_6$ . Noticing that the PE condition guarantees  $\sigma_c$  to be bounded, there exists a positive constant  $\lambda_7$  such that  $\lambda_7 \leq 1/\sigma_c^2 \leq 1$ . In addition, based on Young's inequality, there exists the relation  $-ab \leq \frac{1}{2}(c^2 a^2 + \frac{b^2}{c^2})$ , where  $c$  is a nonzero constant. Then, we have

$$-\frac{\alpha_1}{8\sigma_c^2} (\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c)^2 \leq -\frac{\alpha_1}{8} \lambda_7 \lambda_{1m}^2 \|\tilde{W}_c\|^4, \tag{52}$$

$$\begin{aligned} -\frac{\alpha_1}{\sigma_c^2} (\tilde{W}_c^T \nabla \varphi_c \dot{E}^*) (\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varepsilon_c) &\leq \frac{\alpha_1}{2\sigma_c^2} (c_1^2 (\tilde{W}_c^T \nabla \varphi_c \dot{E}^*)^2 + \frac{(\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varepsilon_c)^2}{c_1^2}) \\ &\leq \frac{\alpha_1 c_1^2}{2} \lambda_3^2 \|\tilde{W}_c\|^2 + \frac{\alpha_1}{2c_1^2} \lambda_5^2 \|\tilde{W}_c\|^2, \end{aligned} \tag{53}$$

$$\begin{aligned}
 -\frac{3\alpha_1}{4\sigma_c^2}(\tilde{W}_c^T \nabla \varphi_c \dot{E}^*)(\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c) &\leq \frac{3\alpha_1}{8\sigma_c^2}(c_2^2(\tilde{W}_c^T \nabla \varphi_c \dot{E}^*)^2 + \frac{(\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c)^2}{c_2^2}) \\
 &\leq \frac{3\alpha_1 c_2^2}{8} \lambda_3^2 \|\tilde{W}_c\|^2 + \frac{3\alpha_1}{8c_2^2} \lambda_{1M}^2 \|\tilde{W}_c\|^4,
 \end{aligned} \tag{54}$$

$$\begin{aligned}
 -\frac{3\alpha_1}{8\sigma_c^2}(\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varepsilon_c)(\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c) &\leq \frac{3\alpha_1}{16\sigma_c^2}(c_3^2(\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varepsilon_c)^2 + \frac{(\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c)^2}{c_3^2}) \\
 &\leq \frac{3\alpha_1 c_3^2}{16} \lambda_5^2 \|\tilde{W}_c\|^2 + \frac{3\alpha_1}{16c_3^2} \lambda_{1M}^2 \|\tilde{W}_c\|^4,
 \end{aligned} \tag{55}$$

$$\begin{aligned}
 -\frac{\alpha_1}{\sigma_c^2} \tilde{W}_c^T \nabla \varphi_c \dot{E}^* \varepsilon_H &\leq \frac{\alpha_1}{2\sigma_c^2}(c_4^2(\tilde{W}_c^T \nabla \varphi_c \dot{E}^*)^2 + \frac{\varepsilon_H^2}{c_4^2}) \\
 &\leq \frac{\alpha_1 c_4^2}{2} \lambda_3^2 \|\tilde{W}_c\|^2 + \frac{\alpha_1}{2c_4^2} \lambda_6^2,
 \end{aligned} \tag{56}$$

$$\begin{aligned}
 -\frac{\alpha_1}{2\sigma_c^2} \tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varepsilon_c \varepsilon_H &\leq \frac{\alpha_1}{4\sigma_c^2}(c_5^2(\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varepsilon_c)^2 + \frac{\varepsilon_H^2}{c_5^2}) \\
 &\leq \frac{\alpha_1 c_5^2}{4} \lambda_5^2 \|\tilde{W}_c\|^2 + \frac{\alpha_1}{4c_5^2} \lambda_6^2,
 \end{aligned} \tag{57}$$

$$\begin{aligned}
 -\frac{\alpha_1}{2\sigma_c^2} \tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c \varepsilon_H &\leq \frac{\alpha_1}{4\sigma_c^2}(c_6^2(\tilde{W}_c^T \nabla \varphi_c \Xi \nabla \varphi_c^T \tilde{W}_c)^2 + \frac{\varepsilon_H^2}{c_6^2}) \\
 &\leq \frac{\alpha_1 c_6^2}{4} \lambda_{1M}^2 \|\tilde{W}_c\|^4 + \frac{\alpha_1}{4c_6^2} \lambda_6^2.
 \end{aligned} \tag{58}$$

Then, (51) is developed as

$$\dot{L}_W \leq -\alpha_1 \lambda_8 \|\tilde{W}_c\|^4 + \alpha_1 \lambda_9 \|\tilde{W}_c\|^2 + \alpha_1 \lambda_{10} - \frac{\alpha_2}{2} \Pi(E, U_E) \tilde{W}_c^T \nabla \varphi_c \Xi \nabla J, \tag{59}$$

where

$$\begin{aligned}
 \lambda_8 &= \frac{1}{8} \lambda_7 \lambda_{1m}^2 - \frac{3}{8c_2^2} \lambda_{1M}^2 - \frac{3}{16c_3^2} \lambda_{1M}^2 - \frac{c_6^2}{4} \lambda_{1M}^2, \\
 \lambda_9 &= \frac{c_1^2}{2} \lambda_3^2 + \frac{1}{2c_1^2} \lambda_5^2 + \frac{3c_2^2}{8} \lambda_3^2 + \frac{3c_3^2}{16} \lambda_5^2 + \frac{c_4^2}{2} \lambda_3^2 + \frac{c_5^2}{4} \lambda_5^2, \\
 \lambda_{10} &= \frac{1}{2c_4^2} \lambda_6^2 + \frac{1}{4c_5^2} \lambda_6^2 + \frac{1}{4c_6^2} \lambda_6^2,
 \end{aligned} \tag{60}$$

and  $c_j$  ( $j = 1, 2, \dots, 6$ ) are all non-zero constants whose selection guarantees  $\lambda_8 > 0$ . Combining the results of (22), (49) and (51), we have

$$\begin{aligned}
 \dot{L} &= \dot{L}_D + \dot{L}_J + \dot{L}_W \\
 &\leq -\kappa \|\tilde{D}\|^2 + \alpha_2 \nabla J^T (f_E + g_E U_E) + \alpha_2 \nabla J^T g_E \tilde{D} - \alpha_1 \lambda_8 \|\tilde{W}_c\|^4 \\
 &\quad + \alpha_1 \lambda_9 \|\tilde{W}_c\|^2 + \alpha_1 \lambda_{10} - \frac{\alpha_2}{2} \Pi(E, U_E) \tilde{W}_c^T \nabla \varphi_c \Xi \nabla J.
 \end{aligned} \tag{61}$$

By using Young’s inequality, the relation  $\alpha_2 \nabla J^T g_E \tilde{D} \leq \frac{\alpha_2 \alpha_3}{2} \nabla J^T g_E g_E^T \nabla J + \frac{\alpha_2}{2\alpha_3} \|\tilde{D}\|^2$  exists. Then, (61) is developed as

$$\begin{aligned}
 \dot{L} &\leq -(\kappa - \frac{\alpha_2}{2\alpha_3}) \|\tilde{D}\|^2 + \alpha_2 \nabla J^T (f_E + g_E U_E) + \frac{\alpha_2 \alpha_3}{2} \nabla J^T g_E g_E^T \nabla J \\
 &\quad - \alpha_1 \lambda_8 \|\tilde{W}_c\|^4 + \alpha_1 \lambda_9 \|\tilde{W}_c\|^2 + \alpha_1 \lambda_{10} - \frac{\alpha_2}{2} \Pi(E, U_E) \tilde{W}_c^T \nabla \varphi_c \Xi \nabla J.
 \end{aligned} \tag{62}$$

The following discussion is divided into two cases.

**Case 1.** In this case,  $\Pi(E, U_E) = 0$ . Since  $\nabla J^T(f_E + g_E U_E) + \alpha_3 \nabla J^T g_E g_E^T \nabla J < 0$ , we can derive that  $\nabla J^T(f_E + g_E U_E) < 0$ . According to the dense property of  $\mathbb{R}$ , there exists a positive constant  $\lambda_{11}$  such that  $0 < \lambda_{11} \|\nabla J\| \leq -\nabla J^T(f_E + g_E U_E)$  for all  $E \in \mathcal{O}$ . Then, (62) becomes

$$\dot{L} \leq -\left(\kappa - \frac{\alpha_2}{2\alpha_3}\right) \|\tilde{D}\|^2 - \frac{\alpha_2}{2} \lambda_{11} \|\nabla J\| - \alpha_1 \lambda_8 \|\tilde{W}_c\|^4 + \alpha_1 \lambda_9 \|\tilde{W}_c\|^2 + \alpha_1 \lambda_{10}. \tag{63}$$

By selecting  $\alpha_2$  and  $\alpha_3$ , such that  $\kappa - \frac{\alpha_2}{2\alpha_3} > 0$ , then  $\dot{L} < 0$  is satisfied provided that one of the following conditions holds:

$$\|\nabla J\| > \frac{\alpha_1(4\lambda_8\lambda_{10} + \lambda_9^2)}{2\alpha_2\lambda_8\lambda_{11}} \triangleq \ell_1, \tag{64}$$

or

$$\|\tilde{W}_c\| > \sqrt{\frac{\lambda_9 + \sqrt{4\lambda_8\lambda_{10} + \lambda_9^2}}{2\lambda_8}} \triangleq \tilde{h}_1. \tag{65}$$

**Case 2.** Considering the case  $\Pi(E, U_E) = 1$ , (62) is developed as

$$\begin{aligned} \dot{L} &\leq \alpha_2(\nabla J^T(f_E + g_E U_E) + \frac{\alpha_3}{2} \nabla J^T g_E g_E^T \nabla J) - \left(\kappa - \frac{\alpha_2}{2\alpha_3}\right) \|\tilde{D}\|^2 \\ &\quad - \alpha_1 \lambda_8 \|\tilde{W}_c\|^4 + \alpha_1 \lambda_9 \|\tilde{W}_c\|^2 + \alpha_1 \lambda_{10} - \frac{\alpha_2}{2} \tilde{W}_c^T \nabla \varphi_c \Xi \nabla J \\ &= \alpha_2(\nabla J^T(f_E + g_E U_E^*) + \frac{\alpha_3}{2} \nabla J^T g_E g_E^T \nabla J) + \frac{\alpha_2}{2} \nabla J^T \Xi \nabla \varepsilon_c \\ &\quad - \left(\kappa - \frac{\alpha_2}{2\alpha_3}\right) \|\tilde{D}\|^2 - \alpha_1 \lambda_8 \|\tilde{W}_c\|^4 + \alpha_1 \lambda_9 \|\tilde{W}_c\|^2 + \alpha_1 \lambda_{10}. \end{aligned} \tag{66}$$

Based on Assumption 4, and considering  $\|g_E\| \leq g_M$ , we have

$$\begin{aligned} \dot{L} &\leq -\alpha_2 \left(\Lambda_m - \frac{\alpha_3}{2} g_M^2\right) \|\nabla J\|^2 + \frac{\alpha_2}{2} \lambda_2 \lambda_4 \|\nabla J\| - \left(\kappa - \frac{\alpha_2}{2\alpha_3}\right) \|\tilde{D}\|^2 \\ &\quad - \alpha_1 \lambda_8 \|\tilde{W}_c\|^4 + \alpha_1 \lambda_9 \|\tilde{W}_c\|^2 + \alpha_1 \lambda_{10}. \end{aligned} \tag{67}$$

Similarly, by selecting  $\alpha_2$  and  $\alpha_3$  such that  $\lambda_{12} = \Lambda_m - \frac{\alpha_3}{2} g_M^2 > 0$  and  $\kappa - \frac{\alpha_2}{2\alpha_3} > 0$ , then it means that  $\dot{L} < 0$  holds as long as

$$\|\nabla J\| > \frac{\lambda_2 \lambda_4}{4\lambda_{12}} + \sqrt{\frac{\alpha_1(4\lambda_8\lambda_{10} + \lambda_9^2)}{4\alpha_2\lambda_8\lambda_{12}} + \frac{\lambda_2^2 \lambda_4^2}{16\lambda_{12}^2}} \triangleq \ell_2, \tag{68}$$

or

$$\|\tilde{W}_c\| > \sqrt{\frac{\lambda_9}{2\lambda_8} + \sqrt{\frac{\lambda_{10}}{\lambda_8} + \frac{\lambda_9^2}{4\lambda_8^2} + \frac{\alpha_2 \lambda_2^2 \lambda_4^2}{16\alpha_1 \lambda_8 \lambda_{12}}}} \triangleq \tilde{h}_2. \tag{69}$$

In conclusion,  $\dot{L} < 0$  when  $\|\nabla J\| > \max\{\ell_1, \ell_2\}$  or  $\|\tilde{W}_c\| > \max\{\tilde{h}_1, \tilde{h}_2\}$ . Relying on Lemma 1 and the standard Lyapunov extension theorem [55], it is further concluded that the tracking error  $E$  of the closed-loop system and the weight estimation error  $\tilde{W}_c$  are UUB. This completes the proof.  $\square$

**Remark 6.** As a result of Theorem 2, the approximate optimal cost function  $\hat{V}(E)$  in (41) and the approximate optimal feedback control input  $U_E$  in (42) can, respectively, converge to the neighborhoods of the optimal cost function  $V^*(E)$  and the optimal feedback control input  $U_E^*$  within finite bounds when the PE condition holds [41].

#### 4. Robust Approximate Optimal Trajectory Tracking Control for a Quadrotor

Position and yaw angle are the system outputs for the quadrotor that tracks the desired trajectory of position and the desired trajectory of yaw angle. The desired trajectories of roll and pitch angles required by the attitude subsystem are generated according to the position subsystem control inputs. The tracking errors in lateral and longitudinal positions are eliminated by the attitude subsystem tracking the desired trajectories of roll and pitch angles. According to the description of the control design in the previous section, the control design for the quadrotor is shown in Figure 2, which can guarantee that the tracking error of the quadrotor remains within a small range.

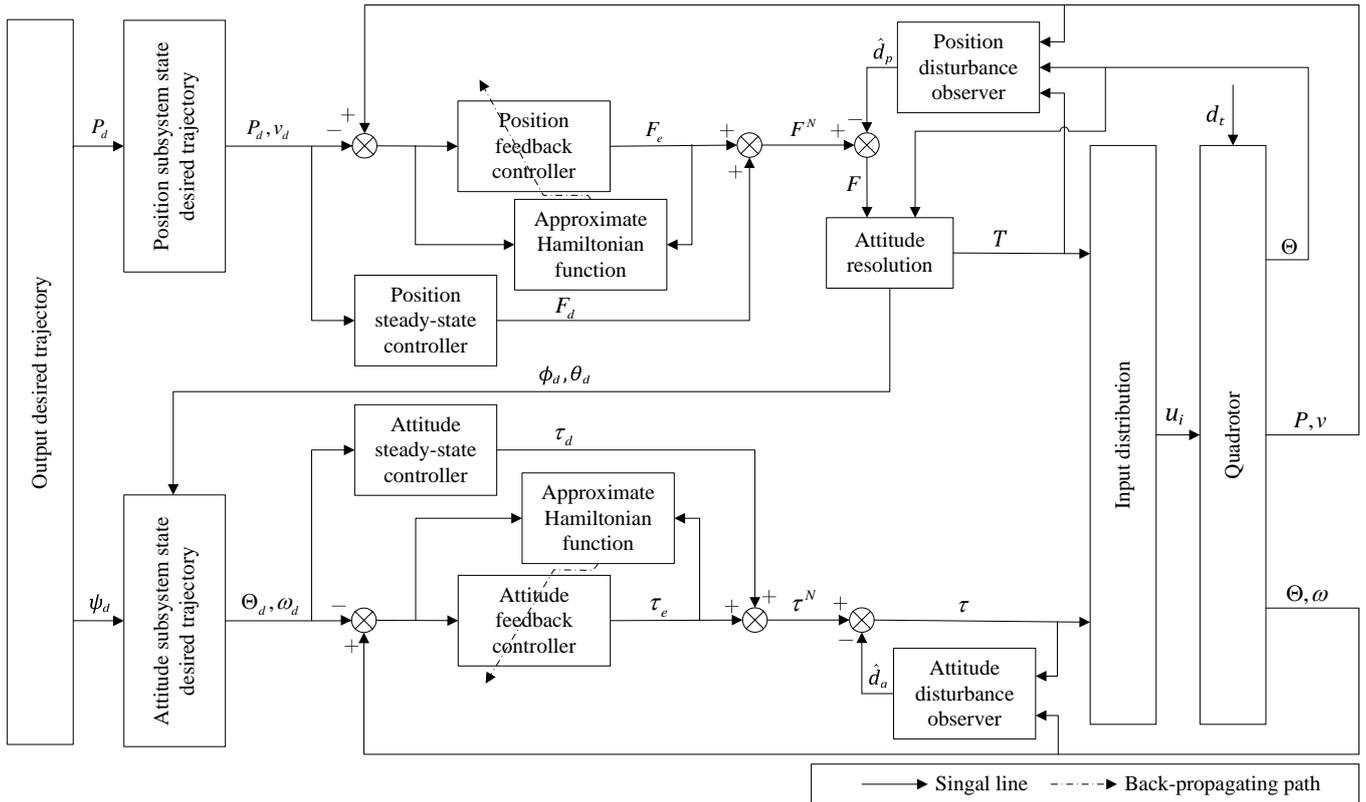


Figure 2. Control design of the quadrotor.

##### 4.1. Position Control Design

The estimated value of unknown compound disturbance  $\hat{d}_p$  in the position subsystem is derived by the following disturbance observer

$$\begin{cases} \dot{z}_1 = -l_1(x_1)(f_1(x_1) + g_1(x_1)(p_1(x_1) + F + z_1)) \\ \hat{d}_p = z_1 + p_1(x_1) \end{cases}, \tag{70}$$

where  $l_1(x_1) = \partial p_1(x_1)/\partial x_1$  denotes the observer gain of the disturbance observer in the position subsystem and  $F$  is derived by (6). Then, the position subsystem robust compensation input is designed as

$$F^R = -\hat{d}_p. \tag{71}$$

The steady-state control input for the position nominal system is designed as

$$F_d = g_1^+(x_{1d})(\dot{x}_{1d} - f_1(x_{1d})), \tag{72}$$

where  $x_{1d} = [P_d^T, v_d^T]^T \in \mathbb{R}^6$  and  $v_d = [v_{xd}, v_{yd}, v_{zd}]^T = \dot{P}_d \in \mathbb{R}^3$ .  $g_1^+(x_{1d})$  denotes the pseudo-inverse of  $g_1(x_{1d})$ . Then, define the position subsystem tracking error as

$$e_1 = x_1 - x_{1d} \triangleq [e_x, e_y, e_z, e_{v_x}, e_{v_y}, e_{v_z}]^T \in \mathbb{R}^6. \quad (73)$$

The cost function of the position subsystem is represented as

$$V_1(e_1) = \int_{t_0}^{\infty} (e_1^T Q_1 e_1 + F_e^T R_1 F_e) dt, \quad (74)$$

where  $Q_1 \in \mathbb{R}^{6 \times 6}$  and  $R_1 \in \mathbb{R}^{3 \times 3}$  are the designed symmetric definite matrices. The approximate optimal feedback control input in the position subsystem is

$$F_e = -\frac{1}{2} R_1^{-1} g_{e1}^T \nabla \varphi_{c1}^T \hat{W}_{c1}, \quad (75)$$

where  $g_{e1} = g_1(x_1)$  and  $\nabla \varphi_{c1} = \partial \varphi_{c1}(e_1) / \partial e_1$ .  $\varphi_{c1}(e_1)$  is the activation function and  $\hat{W}_{c1}$  represents the estimate of the ideal weight for the critic network of the position subsystem. The corresponding weight update law is designed as

$$\begin{aligned} \dot{\hat{W}}_{c1} = & -\frac{\alpha_{11} \sigma_1}{\sigma_{c1}^2} (e_1^T Q_1 e_1 + \hat{W}_{c1}^T \nabla \varphi_{c1} f_{e1} - \frac{1}{4} \hat{W}_{c1}^T \nabla \varphi_{c1} \Xi_1 \nabla \varphi_{c1}^T \hat{W}_{c1}) \\ & + \frac{\alpha_{12}}{2} \Pi(e_1, F_e) \nabla \varphi_{c1} \Xi_1 \nabla J_1, \end{aligned} \quad (76)$$

where  $\alpha_{11} > 0$ ,  $\alpha_{12} > 0$  are the designed learning rates.  $\sigma_1 = \nabla \varphi_{c1}(f_{e1} + g_{e1} F_e)$ ,  $\sigma_{c1} = \sigma_1^T \sigma_1 + 1$  and  $\Xi_1 = g_{e1} R_1^{-1} g_{e1}^T$ .  $\nabla J_1 = \partial J_1(e_1) / \partial e_1$ , where  $J_1(e_1)$  is the Lyapunov function candidate that satisfies Assumption 4.

Then, the robust approximate optimal trajectory tracking control in the position subsystem is designed as

$$F = F^N + F^R = F_d + F_e + F^R. \quad (77)$$

#### 4.2. Attitude Resolution

Since the system of the quadrotor is underactuated and strongly coupled, the information of the position subsystem is used to calculate the total lift force. The desired trajectories of roll and pitch angles are determined by the position subsystem through the relation between the kinematic equation and the Euler equation and passed to the attitude subsystem. For the position subsystem, the generated tracking error and the received compound disturbance can be eliminated by the attitude subsystem. By a matrix operation on (6), the following equations are derived:

$$\begin{aligned} F_x &= T(c_\phi s_\theta c_\psi + s_\phi s_\psi), \\ F_y &= T(c_\phi s_\theta s_\psi - s_\phi c_\psi), \\ F_z &= T c_\phi c_\theta - m g_I. \end{aligned} \quad (78)$$

The actual total lift force for the quadrotor system is designed as

$$T = (F_z + m g_I) / c_\phi c_\theta. \quad (79)$$

Substituting (79) into (78), the form is transformed as

$$\begin{bmatrix} F_x \\ F_y \end{bmatrix} = (F_z + m g_I) \begin{bmatrix} c_\psi & s_\psi \\ s_\psi & -c_\psi \end{bmatrix} \begin{bmatrix} t_\theta \\ t_\phi / c_\theta \end{bmatrix}. \quad (80)$$

The desired trajectories of the pitch and roll angles are derived by the following equations:

$$\begin{aligned} F_x c_\psi + F_y s_\psi &= (F_z + mg_I) t_{\theta_d}, \\ F_x s_\psi - F_y c_\psi &= (F_z + mg_I) \frac{t_{\phi_d}}{c_\theta}. \end{aligned} \quad (81)$$

Then, we have

$$\begin{aligned} \theta_d &= \arctan\left(\frac{F_x c_\psi + F_y s_\psi}{F_z + mg_I}\right), \\ \phi_d &= \arctan\left(c_\theta \frac{F_x s_\psi - F_y c_\psi}{F_z + mg_I}\right). \end{aligned} \quad (82)$$

#### 4.3. Attitude Control Design

Similarly, the estimated value of unknown compound disturbance  $\hat{d}_a$  in the attitude subsystem is derived by the following disturbance observer

$$\begin{cases} \dot{z}_2 = -l_2(x_2)(f_2(x_2) + g_2(x_2)(p_2(x_2) + \tau + z_2)) \\ \hat{d}_a = z_2 + p_2(x_2) \end{cases}, \quad (83)$$

where  $l_2(x_2) = \partial p_2(x_2)/\partial x_2$  denotes the observer gain of the disturbance observer in the attitude subsystem. Then, the attitude subsystem robust compensation input is

$$\tau^R = -\hat{d}_a. \quad (84)$$

The desired trajectory for the angular velocity is given by [56]

$$\omega_d = \begin{bmatrix} 1 & 0 & -s_{\theta_d} \\ 0 & c_{\phi_d} & s_{\phi_d} c_{\theta_d} \\ 0 & -s_{\phi_d} & c_{\phi_d} c_{\theta_d} \end{bmatrix} \dot{\Theta}_d, \quad (85)$$

in which  $\Theta_d = [\phi_d, \theta_d, \psi_d]^T \in \mathbb{R}^3$  is the desired trajectory of Euler angles and  $\omega_d = [p_d, q_d, r_d]^T \in \mathbb{R}^3$  is the desired trajectory of the angular velocity. The steady-state control input for the attitude nominal system is designed as

$$\tau_d = g_2^+(x_{2d})(\dot{x}_{2d} - f_2(x_{2d})), \quad (86)$$

where  $x_{2d} = [\Theta_d^T, \omega_d^T]^T \in \mathbb{R}^6$  and  $g_2^+(x_{2d})$  denotes the pseudo-inverse of  $g_2(x_{2d})$ . Then, define the attitude subsystem tracking error as

$$e_2 = x_2 - x_{2d} \triangleq [e_\phi, e_\theta, e_\psi, e_p, e_q, e_r]^T \in \mathbb{R}^6. \quad (87)$$

While the cost function of the attitude subsystem is represented as

$$V_2(e_2) = \int_{t_0}^{\infty} (e_2^T Q_2 e_2 + \tau_e^T R_2 \tau_e) dt, \quad (88)$$

where  $Q_2 \in \mathbb{R}^{6 \times 6}$  and  $R_2 \in \mathbb{R}^{3 \times 3}$  are the designed symmetric definite matrices. The approximate optimal feedback control input in the attitude subsystem is

$$\tau_e = -\frac{1}{2} R_2^{-1} g_{e2}^T \nabla \varphi_{c2}^T \hat{W}_{c2}, \quad (89)$$

where  $g_{e2} = g_2(x_2)$  and  $\nabla \varphi_{c2} = \partial \varphi_{c2}(e_2)/\partial e_2$ .  $\varphi_{c2}(e_2)$  is the activation function and  $\hat{W}_{c2}$  represents the estimate of the ideal weight for the critic network of the attitude subsystem. The corresponding weight update law is designed as

$$\begin{aligned} \dot{\hat{W}}_{c2} = & -\frac{\alpha_{21}\sigma_2}{\sigma_{c2}^2}(e_2^T Q_2 e_2 + \hat{W}_{c2}^T \nabla \varphi_{c2} f_{e2} - \frac{1}{4} \hat{W}_{c2}^T \nabla \varphi_{c2} \Xi_2 \nabla \varphi_{c2}^T \hat{W}_{c2}) \\ & + \frac{\alpha_{22}}{2} \Pi(e_2, \tau_e) \nabla \varphi_{c2} \Xi_2 \nabla J_2, \end{aligned} \tag{90}$$

where  $\alpha_{21} > 0, \alpha_{22} > 0$  are the learning rates,  $\sigma_2 = \nabla \varphi_{c2}(f_{e2} + g_{e2} \tau_e), \sigma_{c2} = \sigma_2^T \sigma_2 + 1$  and  $\Xi_2 = g_{e2} R_2^{-1} g_{e2}^T, \nabla J_2 = \partial J_2(e_2) / \partial e_2$ , where  $J_2(e_2)$  is the Lyapunov function candidate that satisfies Assumption 4.

Then, the robust approximate optimal trajectory tracking control in the attitude sub-system is designed as

$$\tau = \tau^N + \tau^R = \tau_d + \tau_e + \tau^R. \tag{91}$$

### 5. Simulation Results

In this section, the robustness and effectiveness of the designed controller are evaluated through numerical simulations. The quadrotor is considered to be in a flight environment with slow-changing disturbances. The parameters of the quadrotor model are presented in Table 1 [24].

**Table 1.** Parameters of quadrotor model.

Symbol	Value	Units
$m$	1.79	kg
$gI$	9.81	m/s <sup>2</sup>
$l$	0.20	m
$K_t$	12.0	N
$K_o$	0.40	N·m
$I_{xx} = I_{yy}$	0.03	kg·m <sup>2</sup>
$I_{zz}$	0.04	kg·m <sup>2</sup>
$k_x = k_y = k_z$	0.012	N·s/m

A representative desired trajectory is selected to emulate the trajectory tracking performance of the quadrotor. The desired trajectory is designed as  $P_d = [0.5 \cos(0.5t), 0.5 \sin(0.5t), 0.05t + 0.5]^T$  and  $\psi_d = \pi/12$ . In addition, referring to [57,58], the unknown compound disturbances considered are described as  $d_p = [0.3 + 0.5(\sin(t) + \sin(0.5t) - \cos(0.8t)); 0.3 + 0.5(\cos(t) + \sin(0.5t) - \cos(0.8t)); 0.2 + 0.5 \sin(1.5t)]^T$  and  $d_a = [0.1 + 0.2(\sin(t) + \sin(0.5t)); 0.1 + 0.2(\cos(0.5t) - \cos(0.8t)); 0.05 + 0.2 \sin(t) \sin(0.5t)]^T$ . In this way, the performance of the disturbance observers is reflected by comparing them with the estimates. The initial states of the quadrotor are all set to zero.

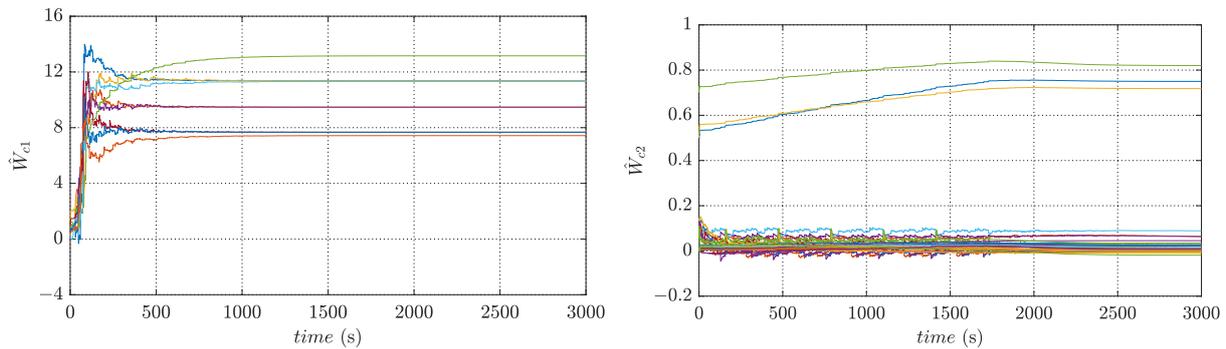
The vector-valued functions of the disturbance observers are designed as  $p_1(x_1) = l_1(x_1)x_1, p_2(x_2) = l_2(x_2)x_2$ , while the observer gains are selected as

$$l_1(x_1) = \begin{bmatrix} 0 & 0 & 0 & 60 & 0 & 0 \\ 0 & 0 & 0 & 0 & 60 & 0 \\ 0 & 0 & 0 & 0 & 0 & 60 \end{bmatrix}, l_2(x_2) = \begin{bmatrix} 0 & 0 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix}.$$

Clearly,  $l_1(x_1)g_1(x_1)$  and  $l_2(x_2)g_2(x_2)$  are positive definite and satisfy the design requirements of Theorem 1. To derive the appropriate dynamic performance, the parameters of the performance index functions are designed as  $Q_1 = \text{diag}\{7, 7, 10, 9, 9, 6\}, Q_2 = \text{diag}\{1.5, 1.5, 1.2, 0.3, 0.3, 0.4\}, R_1 = R_2 = I_3$ . The activation functions are designed as  $\varphi_{c1}(e_1) = [e_x^2, e_x e_{v_x}, e_y^2, e_y e_{v_y}, e_z^2, e_z e_{v_z}, e_{v_x}^2, e_{v_y}^2, e_{v_z}^2]^T, \varphi_{c2}(e_2) = [e_\phi^2, e_\phi e_p, e_\theta^2, e_\theta e_q, e_\psi^2, e_\psi e_r, e_p^2, e_p e_q, e_q^2, e_q e_r, e_r^2, e_\phi^2 e_q e_r, e_\phi e_p e_q e_r, e_\theta^2 e_p e_r, e_\theta e_p e_q e_r, e_\psi^2 e_p e_q, e_\psi e_p e_q e_r, e_p^4, e_p^3 e_q, e_p^3 e_r, e_p^2 e_q^2, e_p^2 e_q e_r, e_p^2 e_r^2, e_p e_q^3 e_r, e_p e_q e_r^2, e_p e_r^3, e_q^4, e_q^3 e_r, e_q^2 e_r^2, e_q e_r^3, e_r^4]^T$ . The relevant constants of the weight update laws are selected as  $\alpha_{11} = 10, \alpha_{12} = 0.01, \alpha_{13} = 0.1, \alpha_{21} = 20, \alpha_{22} = 0.001, \alpha_{23} = 0.1$ . The Lyapunov function candidates are selected as  $J_1(e_1) = \frac{1}{2} e_1^T e_1$  and  $J_2(e_2) = \frac{1}{2} e_2^T e_2$ . The initial weights are assigned values within the interval  $[0, 1]$ .

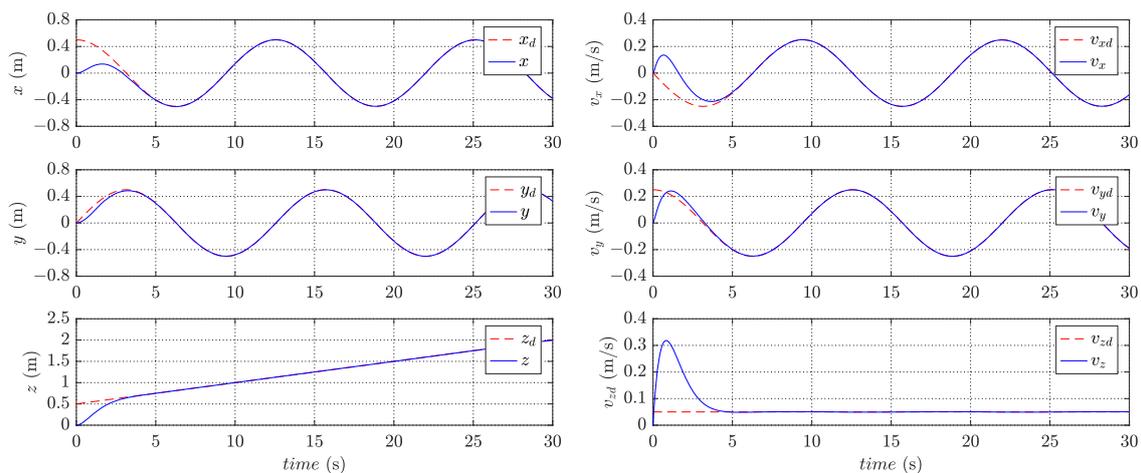
The PE condition is ensured by the method mentioned in Remark 5 to excite the system states. The weights gradually vary to become slower and stabilize during the learning process. The converged weights are already very close to the ideal weights after sufficient learning. The convergence of the whole critic network weights  $\hat{W}_{c1}$ ,  $\hat{W}_{c2}$  in the learning processes are depicted in Figure 3. The final converged values of  $\hat{W}_{c1}$ ,  $\hat{W}_{c2}$  are as follows

$$\begin{aligned}\hat{W}_{c1} &= [11.3714, 9.4718, 11.3713, 9.4718, 13.1589, 11.3203, 7.6718, 7.6718, 7.4279]^T, \\ \hat{W}_{c2} &= [0.7500, 0.0646, 0.7182, 0.0630, 0.8202, 0.0888, 0.0214, -0.0006, -0.0004, \\ &\quad 0.0221, -0.0024, 0.0287, 0.0092, 0.0045, 0.0346, -0.0059, -0.0096, 0.0453, \\ &\quad 0.0340, 0.0211, 0.0106, -0.0025, 0.0039, -0.0045, 0.0227, 0.0126, 0.0155, \\ &\quad 0.0076, 0.0256, 0.0020, 0.0016, 0.0090, -0.0185]^T.\end{aligned}$$



**Figure 3.** Convergence of critic network weights.

The converged weights are used to design the approximate feedback optimal control inputs. Figure 4 and Figure 5 present the variation of states in trajectory tracking control, revealing the corresponding tracking errors in Figure 6 and Figure 7. In addition, Figure 8 visualizes the path in three-dimensional space, whereas Figure 9 illustrates the PWM signals for the motors. The figures clearly demonstrate that the quadrotor system effectively tracks the desired trajectory and achieves a small convergence bound for the tracking error. These results highlight the rapidity and accuracy of the designed controller in the control process.



**Figure 4.** Variation of states in the position subsystem.

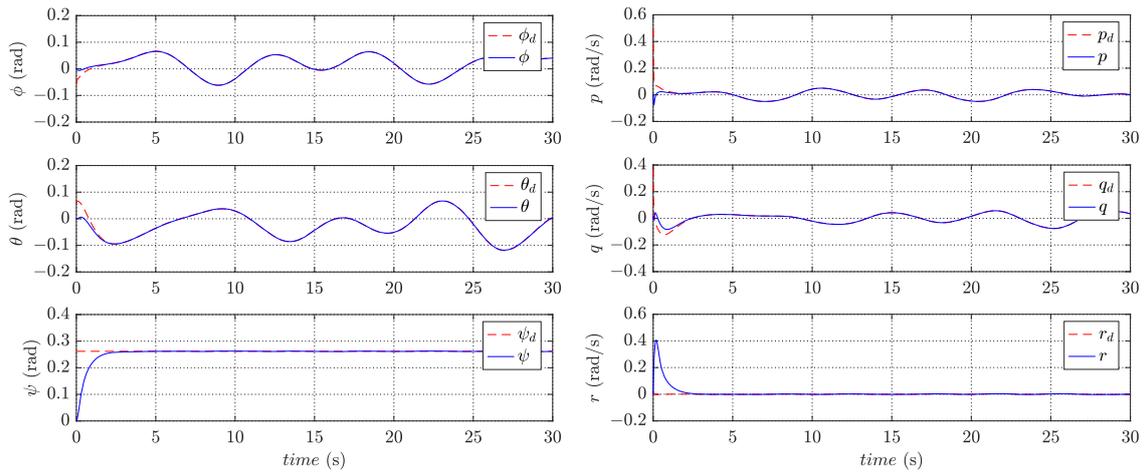


Figure 5. Variation of states in the attitude subsystem.

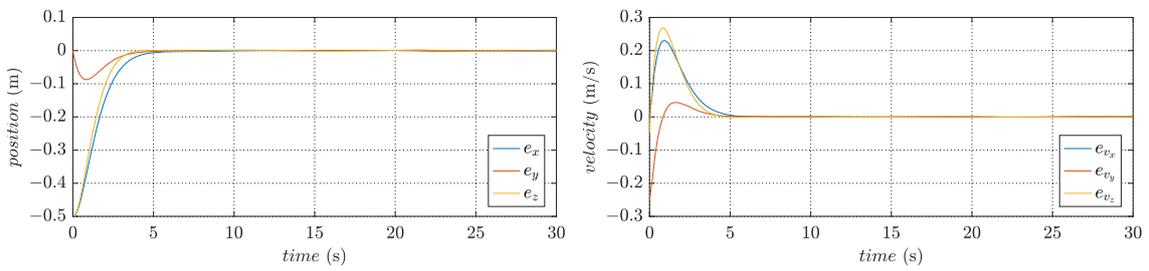


Figure 6. Tracking errors in the position subsystem.

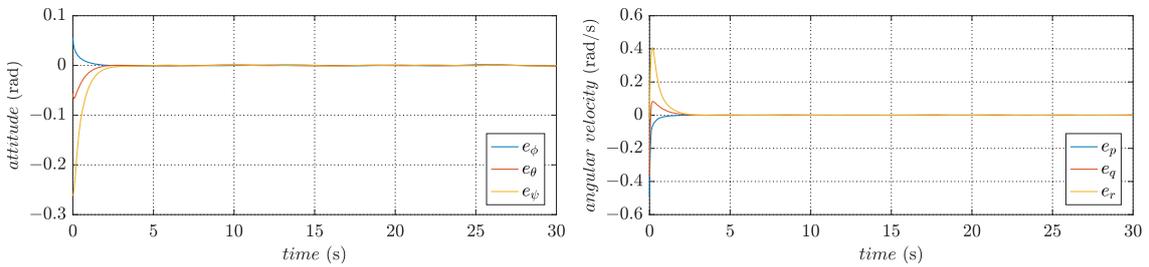


Figure 7. Tracking errors in the attitude subsystem.

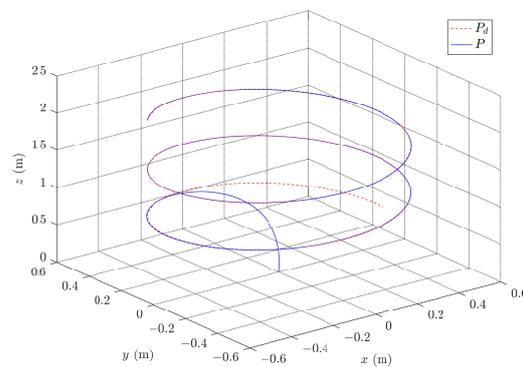


Figure 8. Results of three-dimensional path.

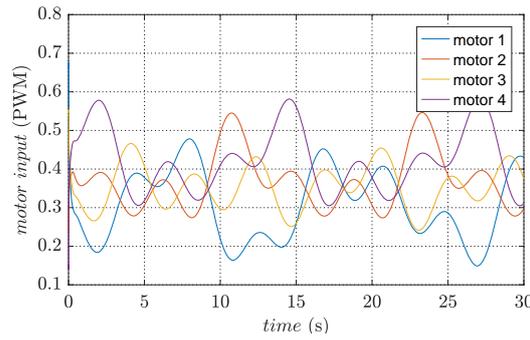


Figure 9. Pulse-width of input signals.

The estimates for the compound disturbances are depicted in Figure 10. It shows that the estimated values from the disturbance observers can quickly follow the actual compound disturbances. Moreover, the trajectory tracking control performs well in the presence of compound disturbances, which implies the robustness of the designed controller.

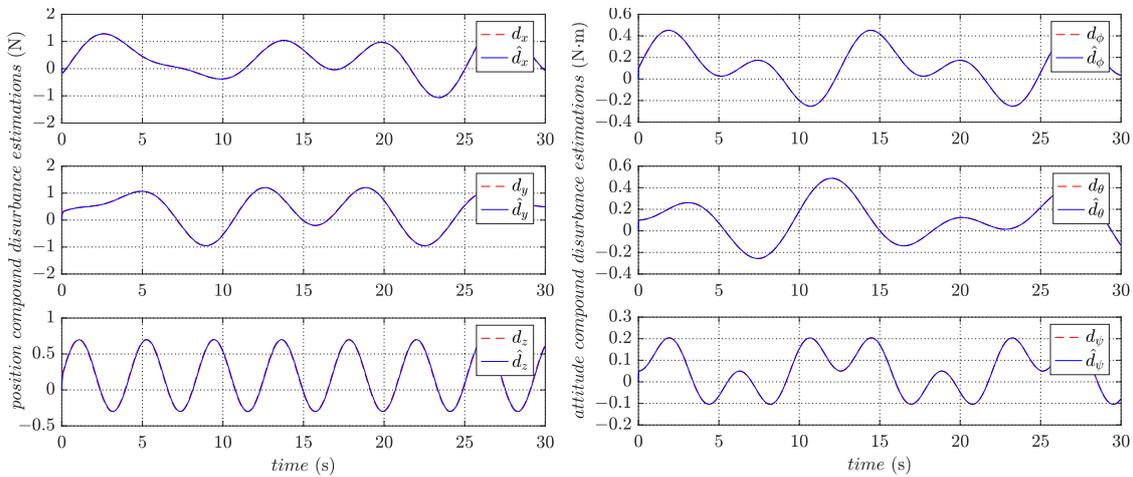


Figure 10. Estimates of compound disturbances.

In order to verify that the designed controller rejects the compound disturbances, a comparative simulation is performed without the disturbance observers in the position subsystem and the attitude subsystem. The control inputs use only the control inputs designed for the nominal system. Under such control, the variation of states is presented in Figures 11 and 12, while Figures 13 and 14 show the corresponding tracking errors.

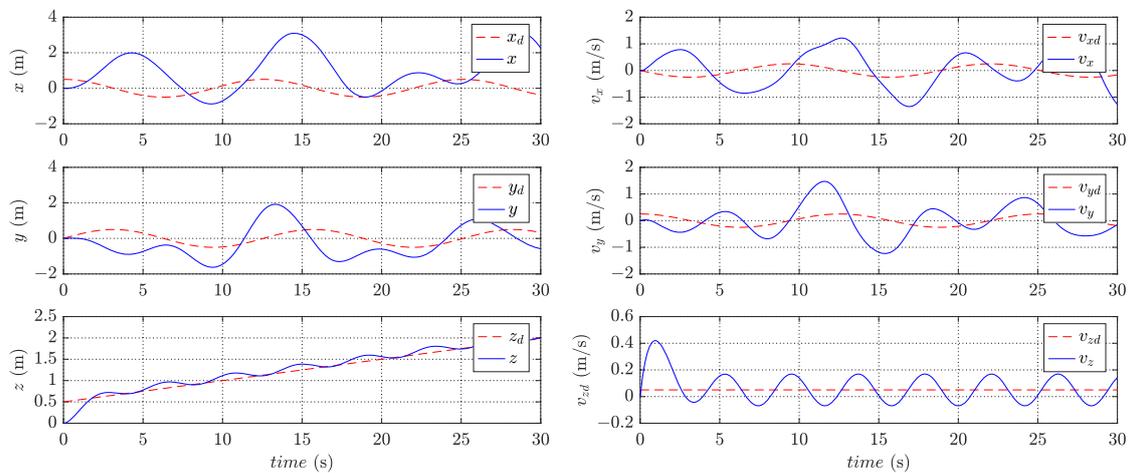


Figure 11. Variation of states in the position subsystem without disturbance observers.

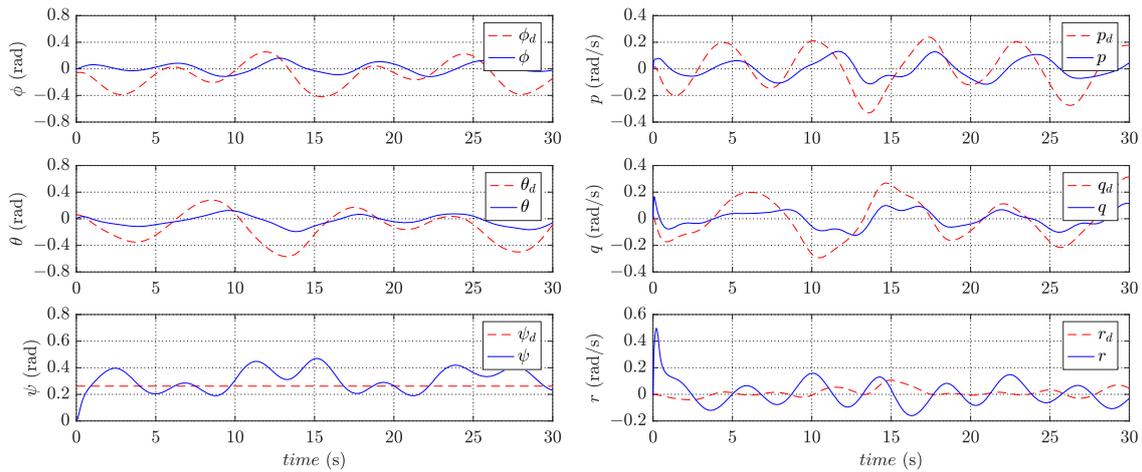


Figure 12. Variation of states in the attitude subsystem without disturbance observers.

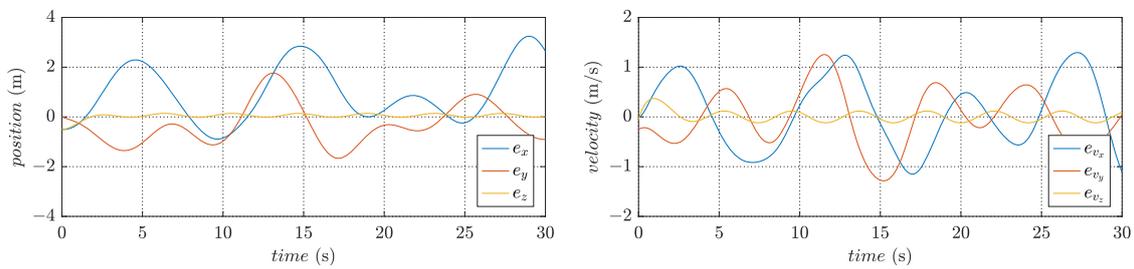


Figure 13. Tracking errors in the position subsystem without disturbance observers.

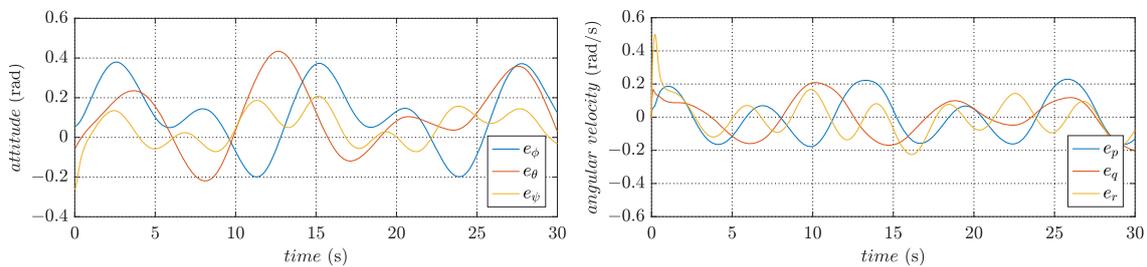


Figure 14. Tracking errors in the attitude subsystem without disturbance observers.

By comparing the simulation results, it is clear that the trajectory tracking control of the quadrotor cannot be realized without the robust compensation inputs. Thus, further demonstrating the robustness of the designed controller. Moreover, the corresponding path in three-dimensional space and the PWM signals of the motors are shown in Figure 15 and Figure 16, respectively.

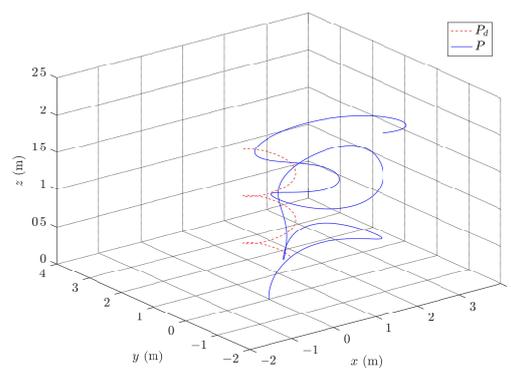
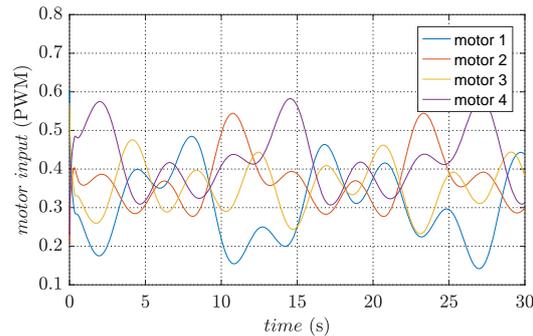


Figure 15. Results of three-dimensional path without disturbance observers.



**Figure 16.** Pulse-width of input signals without disturbance observers.

In summary, the controller designed for quadrotor trajectory tracking control has good dynamic performance, high tracking accuracy and strong robustness when the quadrotor is subjected to compound disturbances.

## 6. Conclusions

This paper proposes a robust approximate optimal controller for the trajectory tracking control of the quadrotor with unknown compound disturbances. By incorporating the estimated values of compound disturbances that are estimated by the disturbance observers into the control design, the effect of compound disturbances can be suppressed, resulting in ensured tracking accuracy and improved robustness. Moreover, the ADP method can then be utilized in the nominal system for ensuring the performance index of the control. The stability of the closed-loop system is analyzed by the Lyapunov theorem, which demonstrates that the tracking errors are UUB. Simulation results further confirm the robustness and effectiveness of the designed controller. In future work, experiments will be considered to validate the performance of the proposed controller.

**Author Contributions:** Conceptualization, R.L. and Z.Y.; methodology, R.L.; software, Z.Y.; validation, R.L., Z.Y. and G.Y.; formal analysis, R.L.; investigation, Z.Y.; resources, R.L.; data curation, L.J.; writing—original draft preparation, Z.Y.; writing—review and editing, R.L., Z.Y. and G.Y.; visualization, G.L.; supervision, Z.L.; project administration, R.L.; funding acquisition, R.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Natural Science Foundation of China (Grant No. 62003233), the Fundamental Research Program of Shanxi Province (Grant Nos. 201901D211083 and 20210302124552), and the Science and Technology Innovation Project of Higher Education Institutions in Shanxi Province (Grant No. 2019L0236).

**Data Availability Statement:** Data are contained within the article.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

- Hassanalian, M.; Abdelkefi, A. Classifications, applications, and design challenges of drones: A review. *Prog. Aerosp. Sci.* **2017**, *91*, 99–131. [[CrossRef](#)]
- Salem, K.A.; Palaia, G.; Chiarelli, M.R.; Bianchi, M. A simulation framework for aircraft take-off considering ground effect aerodynamics in conceptual design. *Aerospace* **2023**, *10*, 459. [[CrossRef](#)]
- Salem, K.A.; Palaia, G.; Quarta, A.A. Review of hybrid-electric aircraft technologies and designs: Critical analysis and novel solutions. *Prog. Aerosp. Sci.* **2023**, *141*, 100924. [[CrossRef](#)]
- Shao, S.; Chen, M.; Hou, J.; Zhao, Q. Event-triggered-based discrete-time neural control for a quadrotor UAV using disturbance observer. *IEEE/ASME Trans. Mechatronics* **2021**, *26*, 689–699. [[CrossRef](#)]
- Idrissi, M.; Salami, M.; Annaz, F. A review of quadrotor unmanned aerial vehicles: Applications, architectural design and control algorithms. *J. Intell. Robot. Syst.* **2022**, *104*, 22. [[CrossRef](#)]
- Rinaldi, F.; Chiesa, S.; Quagliotti, F. Linear quadratic control for quadrotors UVAs dynamics and formation flight. *J. Intell. Robot. Syst.* **2013**, *70*, 203–220. [[CrossRef](#)]
- Dharmawan, A.; Priyambodo, T.K. Model of linear quadratic regulator (lqr) control method in hovering state of quadrotor. *J. Telecommun. Electron. Comput. Eng. (JTREC)* **2017**, *9*, 135–143.

8. Alonge, F.; D'Ippolito, F.; Fagiolini, A.; Garraffa, G.; Sferlazza, A. Trajectory robust control of autonomous quadcopters based on model decoupling and disturbance estimation. *Int. J. Adv. Robot. Syst.* **2021**, *18*, 1729881421996974. [[CrossRef](#)]
9. Yang, Y.; Yan, Y. Attitude regulation for unmanned quadrotors using adaptive fuzzy gain-scheduling sliding mode control. *Aerosp. Sci. Technol.* **2016**, *54*, 208–217. [[CrossRef](#)]
10. Avram, R.C.; Zhang, X.; Muse, J. Nonlinear adaptive fault-tolerant quadrotor altitude and attitude tracking with multiple actuator faults. *IEEE Trans. Control. Syst. Technol.* **2017**, *26*, 701–707. [[CrossRef](#)]
11. Chen, F.; Lei, W.; Zhang, K.; Tao, G.; Jiang, B. A novel nonlinear resilient control for a quadrotor UVA via backstepping control and nonlinear disturbance observer. *Nonlinear Dyn.* **2016**, *85*, 1281–1295. [[CrossRef](#)]
12. Liu, H.; Xi, J.; Zhong, Y. Robust attitude stabilization for nonlinear quadrotor systems with uncertainties and delays. *IEEE Trans. Ind. Electron.* **2017**, *64*, 5585–5594. [[CrossRef](#)]
13. Liu, E.; Yan, Y.; Yang, Y. Neural network approximation-based backstepping sliding mode control for spacecraft with input saturation and dynamics uncertainty. *Acta Astronaut.* **2022**, *191*, 1–10. [[CrossRef](#)]
14. Li, R.; Chen, M.; Wu, Q. Robust control for an unmanned helicopter with constrained flapping dynamics. *Chin. J. Aeronaut.* **2018**, *31*, 2136–2148. [[CrossRef](#)]
15. Li, R.; Chen, M.; Wu, Q. Adaptive neural tracking control for uncertain nonlinear systems with input and output constraints using disturbance observer. *Neurocomputing* **2017**, *235*, 27–37. [[CrossRef](#)]
16. Yang, Y.; Modares, H.; Vamvoudakis, K.G.; He, W.; Xu, C.Z.; Wunsch, D.C. Hamiltonian-driven adaptive dynamic programming with approximation errors. *IEEE Trans. Cybern.* **2021**, *52*, 13762–13773. [[CrossRef](#)]
17. 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](#)]
18. Du, Y.; Jiang, B.; Ma, Y.; Cheng, Y. Robust ADP-based sliding-mode fault-tolerant control for nonlinear systems with application to spacecraft. *Appl. Sci.* **2022**, *12*, 1673. [[CrossRef](#)]
19. Huang, Y.; Wang, D.; Liu, D. Bounded robust control design for uncertain nonlinear systems using single-network adaptive dynamic programming. *Neurocomputing* **2017**, *266*, 128–140. [[CrossRef](#)]
20. Wang, D.; Liu, D.; Li, H. Policy iteration algorithm for online design of robust control for a class of continuous-time nonlinear systems. *IEEE Trans. Autom. Sci. Eng.* **2014**, *11*, 627–632. [[CrossRef](#)]
21. Dou, L.; Su, X.; Zhao, X.; Zong, Q.; He, L. Robust tracking control of quadrotor via on-policy adaptive dynamic programming. *Int. J. Robust Nonlinear Control* **2021**, *31*, 2509–2525. [[CrossRef](#)]
22. Mu, C.; Zhang, Y. Learning-based robust tracking control of quadrotor with time-varying and coupling uncertainties. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *31*, 259–273. [[CrossRef](#)]
23. Chen, W.H.; Yang, J.; Guo, L.; Li, S. Disturbance-observer-based control and related methods—An overview. *IEEE Trans. Ind. Electron.* **2015**, *63*, 1083–1095. [[CrossRef](#)]
24. Chen, M.; Xiong, S.; Wu, Q. Tracking flight control of quadrotor based on disturbance observer. *IEEE Trans. Syst. Man Cybern. Syst.* **2019**, *51*, 1414–1423. [[CrossRef](#)]
25. Chen, F.; Jiang, R.; Zhang, K.; Jiang, B.; Tao, G. Robust backstepping sliding-mode control and observer-based fault estimation for a quadrotor UVA. *IEEE Trans. Ind. Electron.* **2016**, *63*, 5044–5056.
26. Shao, X.; Liu, J.; Cao, H.; Shen, C.; Wang, H. Robust dynamic surface trajectory tracking control for a quadrotor UVA via extended state observer. *Int. J. Robust Nonlinear Control* **2018**, *28*, 2700–2719. [[CrossRef](#)]
27. Mofid, O.; Mobayen, S. Adaptive sliding mode control for finite-time stability of quad-rotor UAVs with parametric uncertainties. *ISA Trans.* **2018**, *72*, 1–14. [[CrossRef](#)] [[PubMed](#)]
28. Lei, W.; Li, C.; Chen, M.Z. Robust adaptive tracking control for quadrotors by combining PI and self-tuning regulator. *IEEE Trans. Control Syst. Technol.* **2018**, *27*, 2663–2671. [[CrossRef](#)]
29. Maqsood, H.; Qu, Y. Nonlinear disturbance observer based sliding mode control of quadrotor helicopter. *J. Electr. Eng. Technol.* **2020**, *15*, 1453–1461. [[CrossRef](#)]
30. Hua, H.; Fang, Y.; Zhang, X.; Lu, B. A novel robust observer-based nonlinear trajectory tracking control strategy for quadrotors. *IEEE Trans. Control Syst. Technol.* **2020**, *29*, 1952–1963. [[CrossRef](#)]
31. 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](#)]
32. Lee, D. Nonlinear disturbance observer-based robust control for spacecraft formation flying. *Aerosp. Sci. Technol.* **2018**, *76*, 82–90. [[CrossRef](#)]
33. Yuan, W.; Gao, G. Sliding mode control of the automobile electro-coating conveying mechanism with a nonlinear disturbance observer. *Adv. Mech. Eng.* **2018**, *10*, 1687814018795748. [[CrossRef](#)]
34. Orozco Soto, S.M.; Cacace, J.; Ruggiero, F.; Lippiello, V. Active Disturbance Rejection Control for the Robust Flight of a Passively Tilted Hexarotor. *Drones* **2022**, *6*, 258. [[CrossRef](#)]
35. Wang, Y.; Sun, J.; He, H.; Sun, C. Deterministic policy gradient with integral compensator for robust quadrotor control. *IEEE Trans. Syst. Man Cybern. Syst.* **2019**, *50*, 3713–3725. [[CrossRef](#)]
36. Li, C.; Wang, Y.; Yang, X. Adaptive fuzzy control of a quadrotor using disturbance observer. *Aerosp. Sci. Technol.* **2022**, *128*, 107784. [[CrossRef](#)]

37. Fan, Y.; Guo, H.; Han, X.; Chen, X. Research and verification of trajectory tracking control of a quadrotor carrying a load. *Appl. Sci.* **2022**, *12*, 1036. [[CrossRef](#)]
38. Wang, B.; Yu, X.; Mu, L.; Zhang, Y. Disturbance observer-based adaptive fault-tolerant control for a quadrotor helicopter subject to parametric uncertainties and external disturbances. *Mech. Syst. Signal Process.* **2019**, *120*, 727–743. [[CrossRef](#)]
39. Fei, Y.; Shi, P.; Lim, C.C. Robust and collision-free formation control of multiagent systems with limited information. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *34*, 4286–4295. [[CrossRef](#)]
40. Fei, Y.; Shi, P.; Lim, C.C. Robust formation control for multi-agent systems: A reference correction based approach. *IEEE Trans. Circuits Syst. Regul. Pap.* **2021**, *68*, 2616–2625. [[CrossRef](#)]
41. 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](#)]
42. Sun, J.; Liu, C. Disturbance observer-based robust missile autopilot design with full-state constraints via adaptive dynamic programming. *J. Frankl. Inst.* **2018**, *355*, 2344–2368. [[CrossRef](#)]
43. Zhang, H.; Cui, L.; Zhang, X.; Luo, Y. Data-driven robust approximate optimal tracking control for unknown general nonlinear systems using adaptive dynamic programming method. *IEEE Trans. Neural Netw.* **2011**, *22*, 2226–2236. [[CrossRef](#)] [[PubMed](#)]
44. Xu, N.; Niu, B.; Wang, H.; Huo, X.; Zhao, X. Single-network ADP for solving optimal event-triggered tracking control problem of completely unknown nonlinear systems. *Int. J. Intell. Syst.* **2021**, *36*, 4795–4815. [[CrossRef](#)]
45. Xia, R.; Wu, Q.; Chen, M. Disturbance observer-based optimal longitudinal trajectory control of near space vehicle. *Sci. China Inf. Sci.* **2019**, *62*, 1–3. [[CrossRef](#)]
46. Sun, J.; Liu, C. Backstepping-based adaptive dynamic programming for missile-target guidance systems with state and input constraints. *J. Frankl. Inst.* **2018**, *355*, 8412–8440. [[CrossRef](#)]
47. Wang, D.; Liu, D.; Li, H.; Ma, H. Neural-network-based robust optimal control design for a class of uncertain nonlinear systems via adaptive dynamic programming. *Inf. Sci.* **2014**, *282*, 167–179. [[CrossRef](#)]
48. Zheng, S.; Shi, P.; Wang, S.; Shi, Y. Adaptive neural control for a class of nonlinear multiagent systems. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *32*, 763–776. [[CrossRef](#)] [[PubMed](#)]
49. Fan, Q.Y.; Yang, G.H. Adaptive actor–critic design-based integral sliding-mode control for partially unknown nonlinear systems with input disturbances. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *27*, 165–177. [[CrossRef](#)] [[PubMed](#)]
50. 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](#)]
51. Liu, D.; Xue, S.; Zhao, B.; Luo, B.; Wei, Q. Adaptive dynamic programming for control: A survey and recent advances. *IEEE Trans. Syst. Man Cybern. Syst.* **2020**, *51*, 142–160. [[CrossRef](#)]
52. Zhao, B.; Liu, D.; Luo, C. Reinforcement learning-based optimal stabilization for unknown nonlinear systems subject to inputs with uncertain constraints. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *31*, 4330–4340. [[CrossRef](#)] [[PubMed](#)]
53. Wang, D.; Liu, D.; Zhang, Y.; Li, H. Neural network robust tracking control with adaptive critic framework for uncertain nonlinear systems. *Neural Netw.* **2018**, *97*, 11–18. [[CrossRef](#)] [[PubMed](#)]
54. Liu, D.; Wei, Q.; Wang, D.; Yang, X.; Li, H. *Adaptive Dynamic Programming with Applications in Optimal Control*; Springer International Publishing: Berlin/Heidelberg, Germany, 2017.
55. Lewis, F.L.; Jagannathan, S.; Yesildirek, A. *Neural Network Control of Robot Manipulators and Nonlinear Systems*; Taylor & Francis: London, UK, 1999.
56. Castillo, A.; Sanz, R.; Garcia, P.; Qiu, W.; Wang, H.; Xu, C. Disturbance observer-based quadrotor attitude tracking control for aggressive maneuvers. *Control Eng. Pract.* **2019**, *82*, 14–23. [[CrossRef](#)]
57. Mobayen, S.; El-Sousy, F.F.; Alattas, K.A.; Mofid, O.; Fekih, A.; Rojsiraphisal, T. Adaptive fast-reaching nonsingular terminal sliding mode tracking control for quadrotor UAVs subject to model uncertainties and external disturbances. *Ain Shams Eng. J.* **2023**, *14*, 102059. [[CrossRef](#)]
58. Shao, X.; Yue, X.; Li, J. Event-triggered robust control for quadrotors with preassigned time performance constraints. *Appl. Math. Comput.* **2021**, *14*, 102059. [[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.