

Article

Deep-Reinforcement-Learning-Based Active Disturbance Rejection Control for Lateral Path Following of Parafoil System

Yuemin Zheng ¹, Jin Tao ^{1,2,*}, Qinglin Sun ¹, Hao Sun ¹, Zengqiang Chen ^{1,3}, Mingwei Sun ¹ and Feng Duan ^{1,*}

¹ College of Artificial Intelligence, Nankai University, Tianjin 300350, China

² Silo AI, 00100 Helsinki, Finland

³ Key Laboratory of Intelligent Robotics of Tianjin, Nankai University, Tianjin 300350, China

* Correspondence: taoj@nankai.edu.cn (J.T.); duanf@nankai.edu.cn (F.D.)

Abstract: The path-following control of the parafoil system is essential for executing missions, such as accurate homing and delivery. In this paper, the lateral path-following control of the parafoil system is studied. First, considering the relative motion between the parafoil canopy and the payload, an eight-degree-of-freedom (DOF) model of the parafoil system is constructed. Then, a guidance law containing the position deviation and heading angle deviation is proposed. Moreover, a linear active disturbance rejection controller (LADRC) is designed based on the guidance law to allow the parafoil system to track the desired path under internal unmodeled dynamics or external environmental disturbances. For the adaptive tuning of the controller parameters, a deep Q-network (DQN) is applied to the LADRC-based path-following control system, and the controller parameters can be adjusted in real time according to the system's states. Finally, the effectiveness of the proposed method is applied to a parafoil system following circular and straight paths in an environment with wind disturbances. The simulation results show that the proposed method is an effective means to realize the lateral path-following control of the parafoil system, and it can also promote the development of intelligent controllers.

Keywords: path-following control; parafoil system; linear active disturbance rejection control; deep Q-network; parameter optimization



Citation: Zheng, Y.; Tao, J.; Sun, Q.; Sun, H.; Chen, Z.; Sun, M.; Duan, F. Deep-Reinforcement-Learning-Based Active Disturbance Rejection Control for Lateral Path Following of Parafoil System. *Sustainability* **2023**, *15*, 435. <https://doi.org/10.3390/su15010435>

Academic Editor: Shuai Su

Received: 9 November 2022

Revised: 19 December 2022

Accepted: 22 December 2022

Published: 27 December 2022



Copyright: © 2022 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

A parafoil system is a unique aircraft consisting of a parafoil canopy, connecting rope, and payload. With its excellent gliding ability and payload characteristics, the parafoil system is currently widely used in aerospace, military, and civil fields. However, the parafoil system applies a flexible canopy to provide the lift force, so it has complicated dynamic characteristics and strong nonlinearity [1]. During flight, the parafoil can be affected by unpredictable wind disturbances. Therefore, determining how to overcome these disturbances and accurately control the parafoil system to follow the desired path is the key to completing missions.

Limited by actual flight tests, which require considerable preparation work and are time-consuming and expensive, a mathematical model is a prerequisite for analyzing the motion characteristics of the parafoil system. To the best of our knowledge, according to the division of the DOFs, the current existing modeling methods include longitudinal four-DOF [2], six-DOF [3,4], eight-DOF [5,6], and nine-DOF [7] models. These models were all obtained through force analysis of the parafoil system. In this paper, considering the relative pitch and yaw motions between the parafoil canopy and the payload, we constructed an eight-DOF model based on a six-DOF model [5].

There have been some research results for the path-following control of the parafoil system. For example, Tao et al. [8] applied a generalized predictive control (GPC)-based

method for a parafoil system to follow the designed path for a better control effect, where the guidance law was based on a combination of tracking errors. Zhao et al. [9] introduced a model-free adaptive control method based on iterative feedback tuning (IFT-MFAC) method for parafoil systems, where only the input/output data are needed during the construction. Guo et al. [10] proposed an improved adaptive path-following guidance law to reduce the effect of time-varying wind disturbances, where the convergence rate was improved by replacing the function. Li et al. [11] designed three proportional–integral–derivative (PID) controllers for the lateral motion, longitudinal motion, and velocity on the basis of the motion characteristics of the parafoil system, which overcame the limitations of the traditional guidance-based tracking strategy. As for actual airdrop scenarios, the PID controller still occupies a dominant position, but the PID controller cannot achieve high tracking accuracy, especially under the disturbances of a complex environment. Tao et al. [12,13] used linear active disturbance rejection control (LADRC) to realize the accurate trajectory tracking of the parafoil system. LADRC is currently the most widely used control strategy in practice besides PID; however, the adjustment of the LADRC parameters remains a challenging problem to be studied.

Active disturbance rejection control (ADRC) [14] was first proposed by Han [15], which combines the state observer in modern theory with the error-based ideas in PID. Specifically, ADRC uses an extended state observer (ESO) to observe the unknown disturbances in the system and uses a state error feedback (SEF) control law to eliminate the disturbance. With model-free characteristics and good control effects, ADRC has attracted the attention of many scholars. Gao [16] developed LADRC through the linearization of the ESO and SEF, which significantly promoted the theoretical and engineering application research of ADRC. In terms of theory, Chen [17] and Wang [18] provided proof of the stability of LADRC. LADRC has demonstrated its control advantages in applications such as power system load frequency control (LFC) [19], heading angle control [20], path-following control [21,22], and an electromechanical servo system [23]. For example, Li et al. [24] proposed a guidance law based on a ship's nonlinear combination of lateral error and heading angle error, and LADRC was used to estimate and eliminate the disturbances. However, this guidance law could only realize path tracking in the y-direction. Inspired by this result, this paper proposes a new guidance law.

Although the number of parameters that need to be adjusted in the LADRC controller has significantly been reduced compared to that in the ADRC, parameter optimization is still a non-negligible part of the controller design process. In most cases, researchers will manually adjust the controller parameters, which makes it challenging to achieve the system's optimal performance. Therefore, various optimization algorithms have emerged continually. The robustness of heuristic algorithms that can optimize a set of fixed parameters, such as the particle swarm optimization (PSO) algorithm [25] and genetic algorithm (GA) [26], is somewhat limited. For algorithms that can achieve adaptive optimization, such as neural networks [27] and fuzzy control [28], these algorithms have difficulty achieving better results when encountering unknown emergencies.

Deep reinforcement learning (DRL) is a combination of reinforcement learning (RL) and neural networks that uses the computing power of neural networks and encompasses the decision-making ability of RL. With the intelligent characteristics of not relying on models and being able to make decisions autonomously, similar to the human brain, DRL is favored by many scholars. The deep Q-network (DQN) [29] is one of the most classical algorithms in DRL, and it overcomes the shortcomings of the Q table in Q-learning of RL whereby it is difficult to express all states. In other words, the DQN can handle systems with continuous states. The application of DQN in path-following control is mainly to make direct decisions in terms of the control variable. For example, Zhao et al. [30] used DQN to determine the rudder angle and propeller speed during the path tracking of an unmanned surface vessel (USV), which completely separated decision-making and control theory. At present, there are few studies on the application of RL in parafoil systems. Therefore, to

promote the development of intelligent controllers in parafoil systems, this paper uses the DQN algorithm to optimize the parameters of the LADRC controller.

Inspired by the previous research, we proposed a DQN-optimized LADRC method for lateral path-following control of the parafoil system in this paper. The main contributions of this paper are summarized as follows:

- A new guidance law for the lateral path-following control of the parafoil system is proposed.
- Based on the guidance law, an LADRC controller is designed to overcome the influence of unknown disturbances during parafoil flight.
- The DQN algorithm of RL is applied to obtain the real-time parameters of the controller based on the parafoil flight states.

The rest of the paper is organized as follows. The eight-DOF parafoil system model is established in Section 2. Section 3 introduces the guidance law and provides the corresponding design process of the LADRC-based path-following controller. Section 4 presents the design process of the DQN-optimized LADRC. The simulation results are given in Section 5, and Section 6 concludes this paper.

2. Dynamic Model of Parafoil System

In the six-DOF parafoil model, the parafoil canopy and payload are regarded as rigid connections, and the relative motion between the two is ignored. In order to describe the motion characteristics of the parafoil more accurately, an eight-DOF dynamic model is established in this paper. The slew, sway, yaw, heave, pitch, and roll motions of the parafoil canopy, and the relative pitch and relative yaw motion between the parafoil canopy and the payload are considered. In this way, we can better observe the attitude of the parafoil system during the movement. Three coordinate systems are established to facilitate parafoil system modeling: the ground coordinate system $O_d x_d y_d z_d$, the parafoil coordinate system $O_s x_s y_s z_s$, and the payload coordinate system $O_w x_w y_w z_w$, as shown in Figure 1.

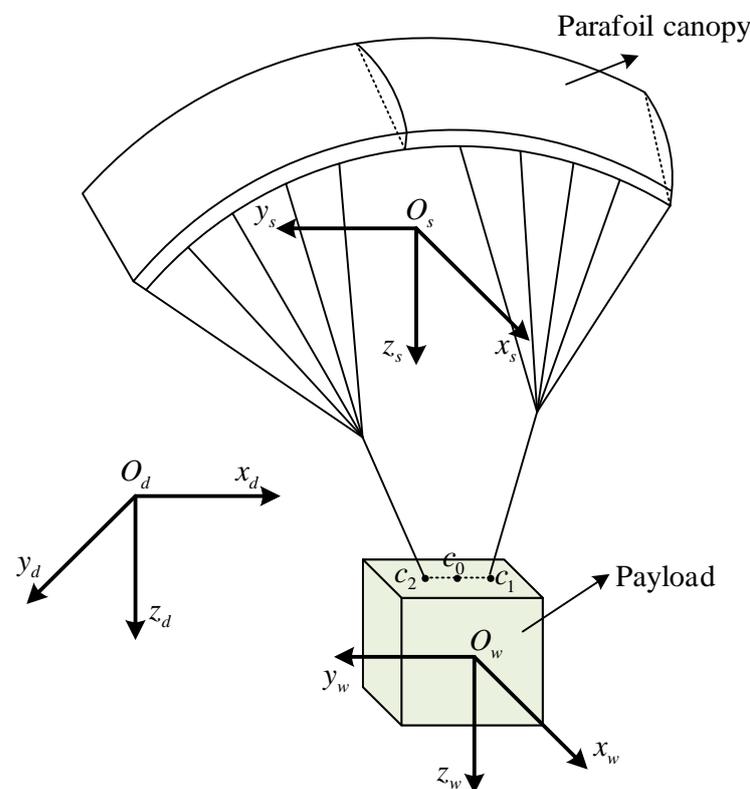


Figure 1. Schematic diagram of parafoil system coordinates.

Furthermore, since the actual model is very complex, in the process of building the model in this paper, the following three assumptions are adopted, and certain simplifications are made:

- The mass center of the parafoil canopy coincides with the aerodynamic pressure center but does not coincide with the gravity center.
- The lift of the payload is ignored and only the aerodynamic drag is considered.
- The elastic deformation of the connecting ropes is ignored.

2.1. Motion Equation of Payload

In this paper, the payload is regarded as a rigid body subjected to aerodynamic forces, gravity, and the pulling force of the connecting rope. The force and moment balance equations of the payload are expressed as

$$\begin{cases} \frac{\partial P_w}{\partial t} + W_w \times P_w = F_w^{aero} + F_w^G + F_w^t \\ \frac{\partial H_w}{\partial t} + W_w \times H_w = M_w^{aero} + M_w^f + M_w^t \end{cases} \quad (1)$$

where P and H represent the momentum and angular momentum, defined below. The superscripts *aero*, *G*, *t*, and *f* represent the aerodynamic force, gravity, connecting rope tension, and friction, respectively. $V_w = [u_w, v_w, w_w]^T$ denotes the velocity vector, and $W_w = [p_w, q_w, r_w]^T$ denotes the angular velocity vector. The momentum and angular momentum are defined, respectively, as follows:

$$\begin{cases} P_w = m_w V_w \\ H_w = J_w W_w \end{cases} \quad (2)$$

The tension and tension moment of the connecting rope are expressed as

$$\begin{cases} F_w^t = F_1 + F_2 \\ M_w^t = [L_{w1} \quad L_{w2}] [F_{w1}^T \quad F_{w2}^T]^T \end{cases} \quad (3)$$

where F_1 and F_2 represent the tension at points c_1 and c_2 , respectively. $L_{w1} = l_2 \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$,

$L_{w2} = \frac{l_1}{2} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$, and l_1 refers to the distance between the connection points c_1 and c_2 . l_2 is the distance between the O_w of the payload coordinate system and the midpoint c_0 of the two connecting points.

2.2. Motion Equation of Parafoil Canopy

When the canopy is fully inflated and unfolded, the force on the canopy includes aerodynamic forces, gravity, and the pulling force of the connecting rope. The equations of the conservation of linear and angular momentum of the parafoil canopy are expressed, respectively, as

$$\begin{cases} \frac{\partial P_s}{\partial t} + W_s \times P_s = F_s^{aero} + F_s^G + F_s^t \\ \frac{\partial H_s}{\partial t} + W_s \times H_s + V_s \times P_s = M_s^{aero} + M_s^G + M_s^f + M_s^t \end{cases} \quad (4)$$

Similarly, the velocity vector $V_s = [u_s, v_s, w_s]^T$ and angular velocity vector $W_s = [p_s, q_s, r_s]^T$ are defined. It should be pointed out that the movement of the parafoil canopy in the air can be approximately regarded as the movement in an ideal fluid. Hence, the influence of the

additional mass on the system needs to be considered. Therefore, the linear momentum and angular momentum of the parafoil canopy considering the additional mass are described as

$$\begin{bmatrix} P_s \\ H_s \end{bmatrix} = [A_a + A_r] \begin{bmatrix} V_s \\ W_s \end{bmatrix}, \quad (5)$$

where A_a denotes the inertia matrix of the additional mass, and A_r represents the inertia matrix of the canopy's real mass.

Similarly, in the parafoil coordinate system, there is an expression of the tension of the connecting rope relative to the moment of the parasol center of mass:

$$M_s^t = [L_{s1}T_{w-s} \quad L_{s2}T_{w-s}] [F_{w1}^T \quad F_{w2}^T]^T, \quad (6)$$

where T_{w-s} is defined below. $L_{s1} = l_3 \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $L_{s2} = \frac{l_1}{2} \begin{bmatrix} 0 & 0 & -\cos \psi_r \\ 0 & 0 & -\sin \psi_r \\ \cos \psi_r & \sin \psi_r & 0 \end{bmatrix}$, and l_3 represents the distance between the O_s of the payload coordinate system and c_0 .

2.3. Constraints of Velocity and Angular Velocity

As shown in Figure 1, the canopy and the payload are connected by ropes at two connection points, c_1 and c_2 , and then the midpoint c_0 of the two points can be regarded as the total connection point. At this point, there is the following velocity constraint:

$$V_w + W_w \times L_{w-c} = V_s + W_s \times L_{s-c}, \quad (7)$$

where L_{w-c} and L_{s-c} represent the distance from the payload centroid to c_0 and the distance from the parafoil canopy centroid to c_0 , respectively. Differentiation of Equation (7) can yield the following formula:

$$\begin{aligned} \dot{V}_s - L_{s-c} \times \dot{W}_s - T_{w-s} \dot{V}_w + T_{w-s} L_{w-c} \times \dot{W}_w \\ = T_{w-s} W_w \times (V_w + W_w \times L_{w-c}) - W_s \times (V_s + W_s \times L_{s-c}), \end{aligned} \quad (8)$$

where T_{w-s} is the conversion matrix from the payload coordinate system to the canopy coordinate system. T_{w-s} is expressed as follows:

$$T_{w-s} = \begin{bmatrix} \cos \theta_r \cos \psi_r & -\sin \psi_r & \sin \theta_r \cos \psi_r \\ \cos \theta_r \sin \psi_r & \cos \psi_r & \sin \theta_r \sin \psi_r \\ -\sin \theta_r & 0 & \cos \theta_r \end{bmatrix}, \quad (9)$$

where θ_r and ψ_r denote the relative pitch angle and relative yaw angle, respectively.

As for the relative rotation between the payload and the canopy, the following angular velocity constraint is satisfied:

$$W_w = W_s + \tau_s + \kappa_w, \quad (10)$$

where $\tau_s = [0, 0, \dot{\psi}_r]^T$, and $\kappa_w = [0, \dot{\theta}_r, 0]^T$. The derivative of Equation (10) is as follows:

$$T_{w-s} \dot{W}_w - \dot{W}_s - T_{w-s} \dot{\kappa}_w - \dot{\tau}_s = (W_s - T_{w-s} W_w) \times W_s - (T_{w-s} \kappa_w) \times \tau_s. \quad (11)$$

Define the state variable as $x_s = [V_w^T, W_w^T, V_s^T, W_s^T, \dot{\psi}_r, \dot{\theta}_r]^T$. Then, by combining Equations (1)–(4), (8) and (11), the eight-DOF dynamic model of the parafoil can be expressed as follows:

$$\dot{x}_s = \left([D_1^T, D_2^T, D_3^T, D_4^T]^T \right)^{-1} [E_1^T, E_2^T, E_3^T, E_4^T]^T. \quad (12)$$

Refer to [4] for a more detailed modeling process.

By denoting $\sin\theta$ as $s\theta$ and $\cos\theta$ as $c\theta$, the following equation can be obtained:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} c\theta c\psi & c\theta s\psi & -s\theta \\ s\phi s\theta c\psi - c\phi s\psi & s\phi s\theta s\psi + c\phi c\psi & s\phi c\theta \\ c\phi s\theta c\psi + s\phi s\psi & c\phi s\theta s\psi - s\phi c\psi & c\phi c\theta \end{bmatrix}^T V_s, \quad (13)$$

where ϕ , θ , and ψ are three Euler angles, i.e., the roll angle, pitch angle, and yaw angle, respectively. Based on Equations (12) and (13), the position of the parafoil can be obtained.

In this paper, the parafoil system is described by the dynamic and kinematic parameters. The dynamic parameters of the parafoil system are the parameters that describe the forces and moments in Equations (1) and (4), which are mainly used to express the relationship between the force of the parafoil system and its motion. The kinematic parameters of the parafoil system are the parameters describing the speed, Euler angles, and acceleration of the parafoil system in Equation (13), which reflects the variation of the position of the parafoil system with time. It is worth mentioning that during the movement of the parafoil, unmodeled dynamics and disturbances in the external environment will cause unpredictable effects. Therefore, the mismatch between theory and reality challenges the controller design.

3. Linear Active Disturbance Rejection Controller (LADRC)-Based Path-Following Control

As mentioned above, the direction of the parafoil system is manipulated by control ropes connecting two sides of the canopy. Specifically, the left control rope is pulled downward to realize a left turn and the right control rope is pulled down to realize a right turn. That is, by controlling the control ropes on both sides, the position and heading angle of the parafoil system are changed simultaneously to track the desired path. Overall, the control command produced by the control system will result in the increase in the force of one of the ropes. Moreover, in the actual control process, the control signal is usually transmitted to the rope in the form of a differential or pulse (1/0), which also brings difficulty to the control system.

3.1. Guidance Law

To achieve convergence of the parafoil position deviation and heading angle deviation to zero simultaneously, according to a guidance law designed in Reference [24], a new guidance law is established as follows:

$$g = g_0 \tanh(g_1 \cdot \Delta d) + \psi - \psi_d, \quad (14)$$

where g_0 and g_1 are adjustable parameters ($g_0 > 0, g_1 > 0$). ψ and ψ_d represent the true value and the planned value of the parafoil's heading angle (yaw angle), respectively. Δd is the position tracking error, as shown in Figure 2, expressed as follows:

$$\begin{cases} \Delta d = \frac{\hat{x}\Delta y - \hat{y}\Delta x}{\sqrt{\Delta x^2 + \Delta y^2}} \\ \psi_d = \arctan(\Delta y / \Delta x) \end{cases}. \quad (15)$$

As shown in Figure 2, the red point $(x(t), y(t))$ represents the current position of the parafoil, and the blue points $(x_r(i-1), y_r(i-1))$ and $(x_r(i), y_r(i))$ are the two adjacent points in the desired path. Define

$$\begin{cases} \Delta x = x_r(i) - x_r(i-1) \\ \Delta y = y_r(i) - y_r(i-1) \\ \hat{x} = x_r(i) - x(t) \\ \hat{y} = y_r(i) - y(t) \end{cases}. \quad (16)$$

The idea of this article is to directly control the flight direction by controlling the tension of the connecting rope so that Δd can be 0. From Equation (14), it can be seen that

the path tracking error and heading angle tracking error are nonlinearly combined, and we hope that both Δd and $(\psi - \psi_d)$ can be stabilized to 0, that is, $g = 0$. There exists the following theorem.

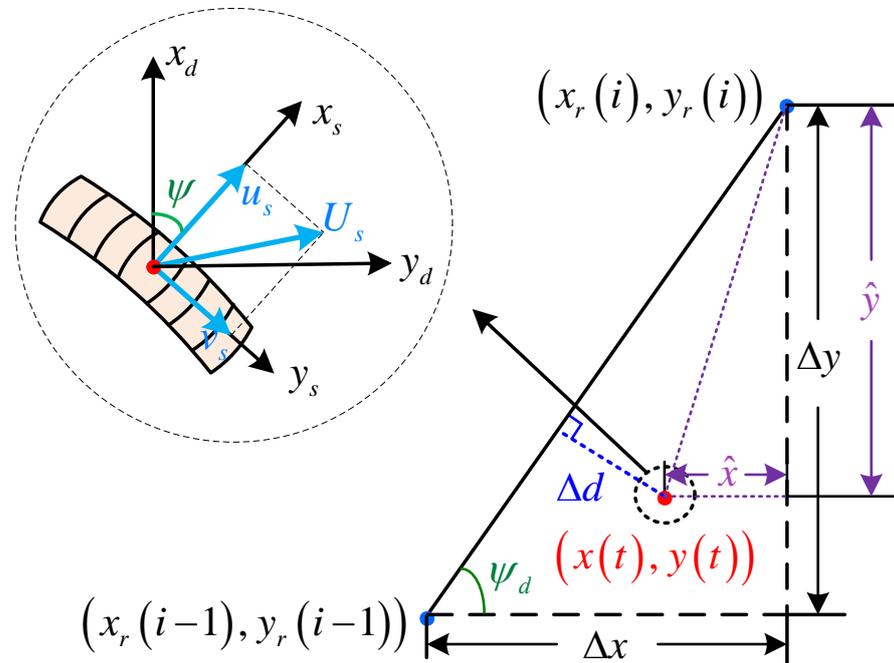


Figure 2. Schematic diagram of path following.

Theorem 1. *There exists a g_0 such that Δd converges to 0 in the case of $g = 0$.*

Proof of Theorem 1. We construct the Lyapunov function as $V = \frac{1}{2} \Delta d^2$. Furthermore, we can obtain

$$\dot{V} = \Delta d \cdot \Delta \dot{d}. \tag{17}$$

By combining Equations (15) and (16), we can obtain Δd in the following form:

$$\Delta d = (x_r(i) - x(t)) \sin \psi_d - (y_r(i) - y(t)) \cos \psi_d, \tag{18}$$

with the derivative

$$\Delta \dot{d} = -\dot{x}(t) \sin \psi_d + \dot{y}(t) \cos \psi_d. \tag{19}$$

By ignoring the influence of the pitch angle and the roll angle, Equation (13) can be simplified to obtain the following expressions:

$$\begin{cases} \dot{x} = u_s \cos \psi - v_s \sin \psi \\ \dot{y} = u_s \sin \psi + v_s \cos \psi \end{cases}. \tag{20}$$

Then, Equation (19) can be rearranged as

$$\begin{aligned} \Delta \dot{d} &= -(u_s \cos \psi - v_s \sin \psi) \sin \psi_d + (u_s \sin \psi + v_s \cos \psi) \cos \psi_d \\ &= \sqrt{u_s^2 + v_s^2} \sin(\beta + \psi - \psi_d), \end{aligned} \tag{21}$$

where $\beta = \arctan(u_s/v_s)$. During the forward flight of the parafoil, it can be approximated that $u_s \gg v_s$, that is, $\beta \approx 0$. Then, Equation (17) has the following expression:

$$\dot{V} = \Delta d \cdot \sqrt{u_s^2 + v_s^2} \sin(\psi - \psi_d). \tag{22}$$

In the case of $g = 0$, by substituting Equation (14) into Equation (22), we have

$$\dot{V} = \Delta d \cdot \sqrt{u_s^2 + v_s^2} \sin(-g_0 \tanh(g_1 \cdot \Delta d)). \quad (23)$$

There are the following three situations:

- (a) $\Delta d = 0$. In this case, $\dot{V} = 0$.
 (b) $\Delta d > 0$. In this case, $-g_0 < -g_0 \tanh(g_1 \cdot \Delta d) < 0$. Then, $0 < g_0 < \pi$, and thus, $\dot{V} < 0$.
 (c) $\Delta d < 0$. In this case, $0 < -g_0 \tanh(g_1 \cdot \Delta d) < g_0$. Then, $0 < g_0 < \pi$, and thus, $\dot{V} < 0$.

□

3.2. Design of LADRC

The design of the LADRC does not require the model of the system to be known, but it requires the order information of the controlled plant. The guidance law in Equation (14) can essentially be understood as controlling the position deviation through the heading angle. Generally, there is the following relationship between the Euler angles and the angular velocities:

$$\begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} 1 & \sin \phi \tan \theta & \cos \phi \tan \theta \\ 0 & \cos \phi & -\sin \phi \\ 0 & \sin \phi / \cos \theta & \cos \phi / \cos \theta \end{bmatrix} \begin{bmatrix} p_s \\ q_s \\ r_s \end{bmatrix}. \quad (24)$$

Further derivation yields the following:

$$\begin{aligned} \ddot{\psi} = & \frac{\sin \phi}{\cos \theta} \dot{q} + \frac{\cos \phi}{\cos \theta} \dot{r} + \frac{\sin \theta \sin 2\phi}{\cos^2 \theta} q^2 - \frac{\sin \theta \sin 2\phi}{\cos^2 \theta} r^2 \\ & + \frac{\cos \phi}{\cos \theta} p q - \frac{\sin \phi}{\cos \theta} p r + \frac{2 \sin \theta \cos 2\phi}{\cos^2 \theta} q r. \end{aligned} \quad (25)$$

By combining Equations (12) and (25), we can obtain the second-order relationship between ψ and the control variable u :

$$\ddot{\psi} = f_1(\cdot) + f_2(u), \quad (26)$$

where $f_1(\cdot)$ represents the disturbance, including the internal state and external disturbance information, and $f_2(u)$ is an expression of the control variable. u represents the deflection of the left and right trailing edges of the canopy.

Since g and ψ have the same order, we can rearrange Equation (26) as

$$\ddot{g} = f(\cdot) + f_2(u) - b_0 u + b_0 u = f + b_0 u, \quad (27)$$

where f can be regarded as the total disturbances of the parafoil system containing unmodeled dynamics inside the system and disturbances in the external environment, and b_0 is an adjustable parameter.

Then, the states can be defined as $x_1 = g$, $x_2 = \dot{g}$, and $x_3 = f$. The linear ESO (LESO) can be expressed as

$$\begin{cases} \dot{\hat{x}} = A\hat{x} + Bu + L(\hat{y} - \hat{y}) \\ \hat{y} = C\hat{x} \end{cases}, \quad (28)$$

where $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 \\ b_0 \\ 0 \end{bmatrix}$, $L = \begin{bmatrix} \beta_{01} \\ \beta_{02} \\ \beta_{03} \end{bmatrix}$, $C = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}^T$, and $\hat{x} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \end{bmatrix}$. In

addition, β_{01} , β_{02} , and β_{03} are observer gains, which are the prerequisite for realizing the observed state \hat{x} to estimate the true value of the state x . Usually, the pole configuration method is used to configure the observer gain at the pole $-\omega_o$, which is arranged as

$$\begin{aligned} |sI - (A - LC)| &= s^3 + \beta_{01}s^2 + \beta_{02}s + \beta_{03} \\ &= (s + \omega_o)^3. \end{aligned} \quad (29)$$

Thus, $\beta_{01} = 3\omega_o$, $\beta_{02} = 3\omega_o^2$, and $\beta_{03} = \omega_o^3$. In this way, by adjusting the parameter ω_o , the total system disturbance f can be estimated.

Under the premise of $\hat{x}_3 \approx f$, define u in Equation (27) as

$$u = \frac{u_0 - \hat{z}_3}{b_0}. \quad (30)$$

Then, Equation (31) can be obtained:

$$\ddot{g} \approx u_0, \quad (31)$$

where u_0 follows the proportional–derivative (PD) controller:

$$u_0 = k_p(g_d - \hat{x}_1) - k_d\hat{x}_2. \quad (32)$$

Since the target value of g is 0, g_d is 0. By using the pole placement method in the same way, we obtain $k_p = 2\omega_c$ and $k_d = \omega_c^2$.

As seen from the above process, the parameters that need to be adjusted are g_0 , g_1 , ω_o , ω_c , and b_0 . Here, g_1 keeps the heading angle and Δd at the same order of magnitude, and g_0 can limit the maximum heading angle of the system. Therefore, these two parameters can be adjusted manually. In addition, Ref. [31] shows that b_0 in the LADRC is not an essential factor for suppressing interference. Therefore, the parameters to be adjusted are ω_o and ω_c .

4. Deep Q-Network (DQN)-Optimized LADRC

We know that RL has the advantages of not relying on models and learning independently, which has shown excellent results in solving sequential decision-making problems. This paper uses the DQN in RL to optimize ω_o and ω_c for the LADRC.

4.1. Basics of DQN

If the entire parafoil system containing the LADRC-based path-following controller is regarded as the environment, then an agent can be designed in RL, similar to the human brain. A series of decisions can be completed through the continuous interaction between the environment and the agent. Usually, the process of RL is described by the Markov decision process (MDP): (S, A, P, R) , where S and A are the environment state space and the agent's action space, respectively. P represents the probability of the state transition. During the agent training process, the training target comes from the reward function R . On this basis, the cumulative reward R_c can be obtained for each episode:

$$R_c = \sum_{t=0}^{\infty} \gamma^t R_{t+1}, \quad (33)$$

where γ is the discount factor. It can reflect the importance of the reward value at a future moment. In Q-learning, the training goal of the agent is the evaluated value Q corresponding to the state–action pair:

$$Q(s, a) = E[R_c | S_t = s, A_t = a]. \quad (34)$$

Equation (34) shows the expected value of the cumulative reward when the state is s , and the action is selected as a at time t . In other words, when the agent is sufficiently trained, and when the current system state is s , the action value a corresponding to the maximum Q value can be selected according to the Q table.

As for the action space, the parameters that need to be adjusted are ω_o and ω_c , so the action space is obtained by permutation and combination based on these two parameters. The action space of ω_o and ω_c is expressed as

$$\begin{cases} \omega_o \in [\omega_o \min, \omega_o \max] \\ \omega_c \in [\omega_c \min, \omega_c \max] \end{cases}, \quad (38)$$

where the sampling interval is h for the two parameters. This means that ω_o and ω_c are $N_{\omega_o} = \frac{\omega_o \max - \omega_o \min}{h} + 1$ and $N_{\omega_c} = \frac{\omega_c \max - \omega_c \min}{h} + 1$, respectively. Therefore, the dimension of the action space is $N_{\omega_o} * N_{\omega_c}$.

The reward function is a direct factor that affects the agent's training. In order to enable the parafoil system to track the desired path better, we hope that when Δd is relatively small, the reward function can be significant. Therefore, the reward function in this paper is designed as

$$R = -10 \text{sign}(|s_1| - 2) - |s_2|. \quad (39)$$

The reward function means that when Δd is less than 2, the agent can obtain a reward of 10. Otherwise, it can obtain a reward of -10 . The smaller Δd is, the greater the reward is. The training process of the DQN is described in Algorithm 1. The schematic diagram of the parameter optimization based on the DQN is shown in Figure 4. It should be pointed out that the process of optimizing the LADRC controller parameters by the DQN algorithm includes two parts: offline training, shown in Algorithm 1, and online parameter acquisition. During offline training, a Q -network is obtained, and the training time is related to the simulation sample step and the number of iterations. The smaller the simulation sample step is or the larger the number of iterations is, the longer the training time will be. Moreover, by inputting the system state at the current moment into the trained Q -network, the optimal action values can be obtained to realize adaptive controller parameters, which is the online parameter selection process.

Algorithm 1 DQN algorithm.

- 1: Initialize the replay buffer; Initialize the Q -network with random weights μ ; Initialize the target \hat{Q} -network with weights $\mu' = \mu$.
 - 2: **for** episode = 1: M **do**
 - 3: Initialize the states $s_1 = s_{01}, s_{02}$;
 - 4: **for** t = 1: T **do**
 - 5: Select a random action $a_t = \omega_o, \omega_c$ with the probability ε ; otherwise, select $a_t = \arg \max_a f_{\mu}(s_t, a)$;
 - 6: Execute action a_t in the LADRC and observe the reward R_t and s_{t+1} ;
 - 7: Set $s_{t+1} = s_t$;
 - 8: Store (s_t, a_t, R_t, s_{t+1}) in the replay buffer \mathcal{D} ;
 - 9: Randomly extract m sets of data (s_j, a_j, R_j, s_{j+1}) from the replay buffer \mathcal{D} ;
 - 10: Set $y_j = \begin{cases} R_j, & \text{if episode terminates at step } j + 1 \\ R_j + \gamma \max_{a'} f_{\mu'}(s_{j+1}, a_{j+1}), & \text{otherwise} \end{cases}$
 - 11: Perform a gradient descent step according to Equation (36);
 - 12: Every T_n steps, reset $\mu' = \mu$.
 - 13: **end for**
 - 14: **end for**
-

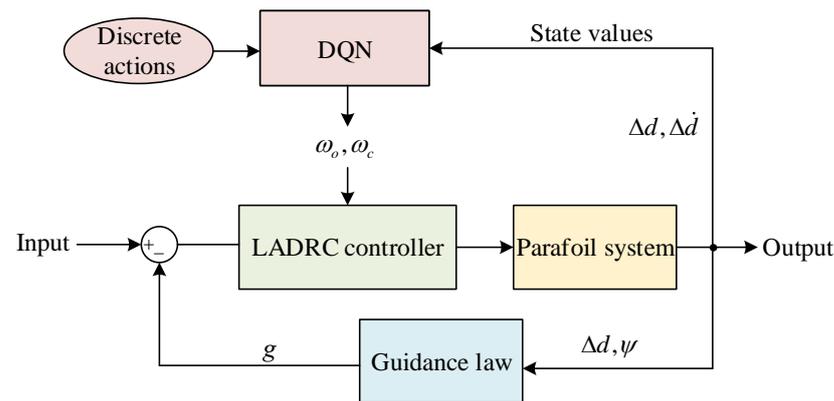


Figure 4. DQN-based linear active disturbance rejection controller (LADRC) control structure diagram.

5. Simulation and Analysis

5.1. Environment Setting

In this section, the effectiveness of the proposed DQN-LADRC method is verified by simulations. The parafoil system's parameters are shown in Table 1. In addition, the parameters involved in the DQN training process are shown in Table 2, where the Q-network is a fully connected neural network with two hidden layers. The network's inputs are the state values shown in Equation (37), and the outputs are the parameters to be optimized. Both hidden layers have 500 neurons. In addition, the learning rate value was fixed during training, and its value was small enough to ensure convergence of the networks.

Table 1. Physical parameters of the parafoil system.

Parameter Description	Value
Wing span	4.5 m
Mean aerodynamic chord	1.3 m
Mass of parafoil	1.7 kg
Mass of payload	20 kg
Wing area	6.5 m ²
Rope length	3 m

Table 2. Parameter settings of DQN.

Parameter Description	Value
Simulation sample step	0.02 s
Number of iterations	300
Learning rate of Q-network	10 ⁻⁴
Discount factor	0.99

Furthermore, the range limits of the parameter tuning in this paper were artificially selected to ensure that the simulation could proceed smoothly within the selected range. The change of the parameter range size, on the one hand, could cause the system to directly stop the training when the "bad" parameters were explored; on the other hand, if the parameter range was too large, it would greatly increase the training time unnecessarily.

It should also be noted that, since the state observer will have extreme values in the transient response, to avoid adverse effects on equipment caused by extreme values, the control quantity generated by the motor connecting the control rope of the dynamic parafoil is limited: $-1 \leq u \leq 1$ [32].

5.2. Performance Study of DQN-LADRC under Wind Disturbances

5.2.1. Circular Path Following

In order to verify the effectiveness of the proposed method, the DQN-LADRC-based lateral path-following control of the parafoil system was conducted under wind disturbances. The desired path was set as a circle with a center of (0 m, 0 m) and a radius of 200 m. In the initial state, the three coordinates coincided, and the initial velocity and angular velocity of the parafoil were $V_s = [14.9, 0, 2.1]^T$ m/s and $W_s = [0, 0, 0]^T$ rad/s, respectively. Furthermore, the parafoil's initial plane position and initial heading angle were (0 m, 150 m) and 0° , respectively. In the simulation environment, the average wind speed of 2 m/s along the y -axis square of the ground coordinate system was added at 50 s, and the duration was 20 s. $g_0, g_1,$ and b_0 were selected as 0.08, 0.2, and 0.2, respectively. This paper compares the control effect of the proposed method with the traditional LADRC controller under a wind disturbance of the same size and direction, in which the parameters of the LADRC were selected from two sets of boundary values of the action space, as follows:

$$\begin{cases} \omega_o \in [1.5, 2.7], \omega_c \in [0.4, 1], & \text{For DQN-LADRC, } h = 0.01 \\ \omega_o = 2.7, \omega_c = 1, & \text{For LADRC1} \\ \omega_o = 1.5, \omega_c = 0.4, & \text{For LADRC2} \end{cases} \quad (40)$$

The abovementioned trained agent was used to obtain the parameters of ω_o and ω_c , and the simulation results are shown in Figures 5–8.

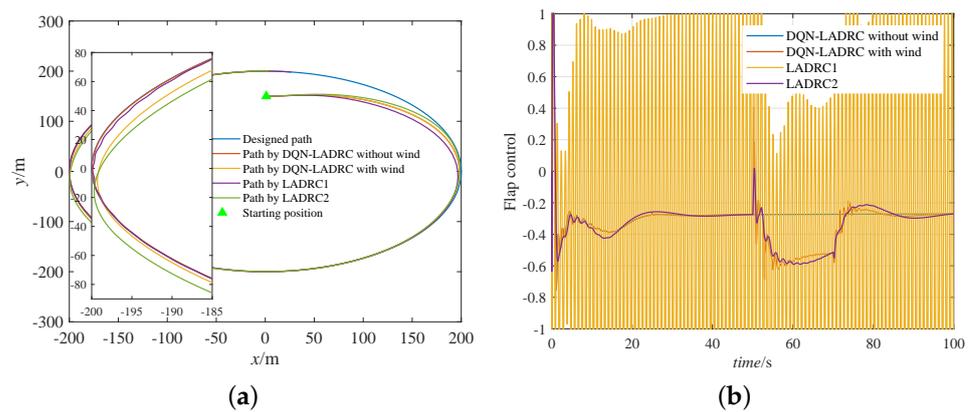


Figure 5. Circular-path-following control results. (a) Path-following trajectories. (b) Control variables.

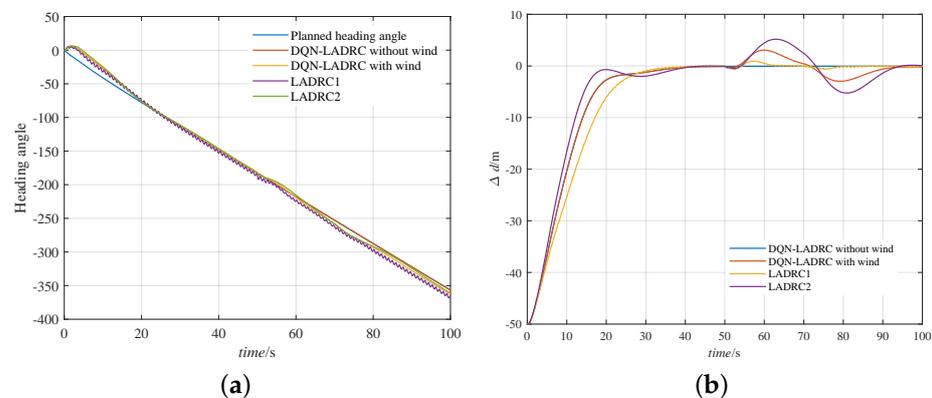


Figure 6. Circular-path-following control results of attitude angles. (a) Heading angle. (b) Position deviation.

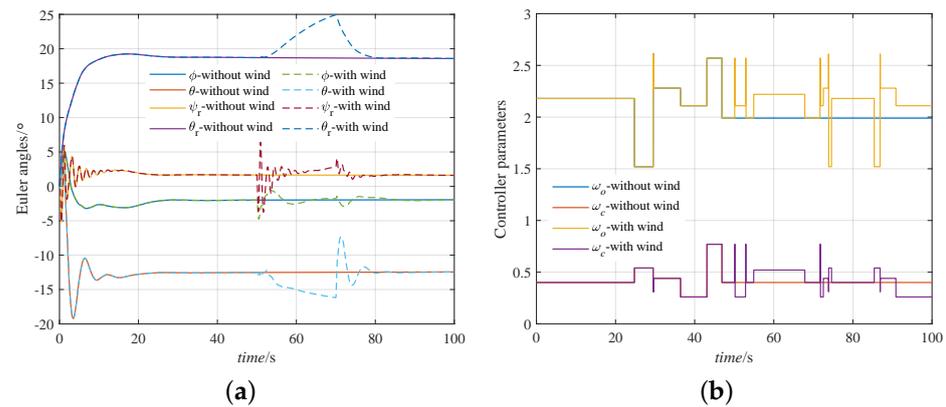


Figure 7. Variation process of Euler angles and controller parameters for circular-path-following control. (a) Euler angles. (b) Adaptive parameters.

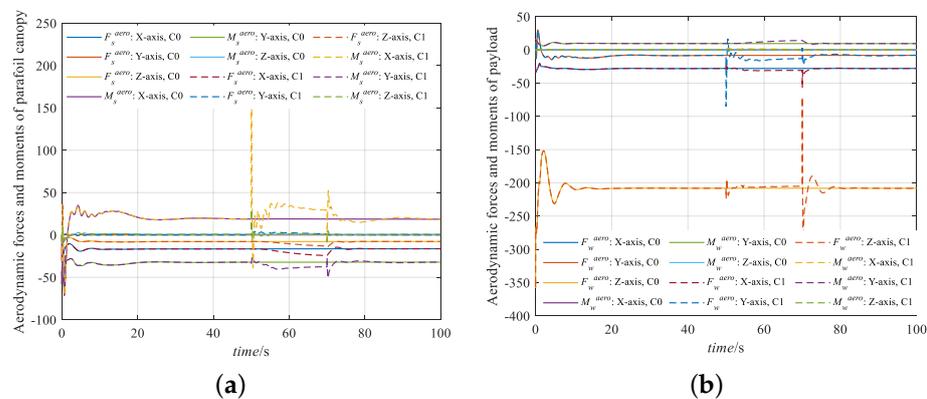


Figure 8. Changing process of aerodynamic forces and moments of parafoil canopy and payload for circular path following. (a) F_s^{aero} and M_s^{aero} in two conditions. C0: the system is not disturbed by wind; C1: the system is disturbed by wind. (b) F_w^{aero} and M_w^{aero} in two conditions. C0: the system is not disturbed by wind; C1: the system is disturbed by wind.

Figure 5 shows the intuitive effect of the parafoil tracking of the desired path with and without wind disturbances, respectively. We can see that both the DQN-LADRC and the traditional LADRC could overcome disturbances and achieve circular path tracking. Specifically, the proposed method achieved better results than the LADRC with the fixed parameters in terms of the control variables, as shown in Figure 5b, and the tracking errors, as shown in Figure 6. It can be found that when the smaller controller parameters were taken, the overshoot of the tracking error obtained by the LADRC was more significant. For more extensive controller parameters, although the overshoot became smaller, the system shock could not be ignored, and such drastic changes in the flap deflection could not be realized in practice. The Euler angles in Figure a show the flight attitude of the parafoil during path tracking. From Figure b, we can observe that the agent had significantly different action values under the influence of wind disturbances. This is also the core idea of the proposed method: to adjust the controller parameters in real time according to the system state. Figure 8 shows the aerodynamic forces and moments of parafoil canopy and payload in three coordinate axes.

5.2.2. Straight Path Following

To prove that the above results were not accidental, the straight-path-following control for the parafoil system was also studied. The desired path consisted of three line segments, and the initial position of the parafoil was (0, 80 m). With the initial heading angle of 0° and b_0 of 0.2, the simulation results by the DQN-LADRC are shown in Figures 9–12. In

order to reflect the anti-interference performance of the proposed method, a south wind of 5 m/s was added at 150 s and continued until the end of the simulation.

Figure 9 shows the tracking effect of the straight path. It can be seen that the system continuously adjusted the flight direction by controlling the rope after being affected by the wind. It can be observed from Figure 10 that the parafoil could fly according to the planned heading angle, and its attitude angles could finally be stabilized. That is, the parafoil could smoothly follow the designed path. The path tracking error is shown in Figure 11a, where the path tracking deviation was 0 at steady state. In addition, the parafoil had large tracking deviations during sharp turns. The adaptive controller parameters are given in Figure 11b, where $\omega_o \in [0.5, 1.5]$ and $\omega_c \in [0.2, 0.5]$, with sampling intervals of 0.01. To demonstrate the proposed method's effectiveness, the LADRC controller's control effect with fixed parameters of $\omega_o = 1.5$ and $\omega_c = 0.5$ was added to the results. The aerodynamic forces and moments on the canopy and payload are shown in Figure 12 in three coordinate axes. It can be seen that in the z-axis direction there are aerodynamic effects. This is because in the eight-DOF model in this paper, we default that the parafoil is free-falling in the longitudinal direction. Even so, the proposed method can enable the parafoil to track a given path on the plane.

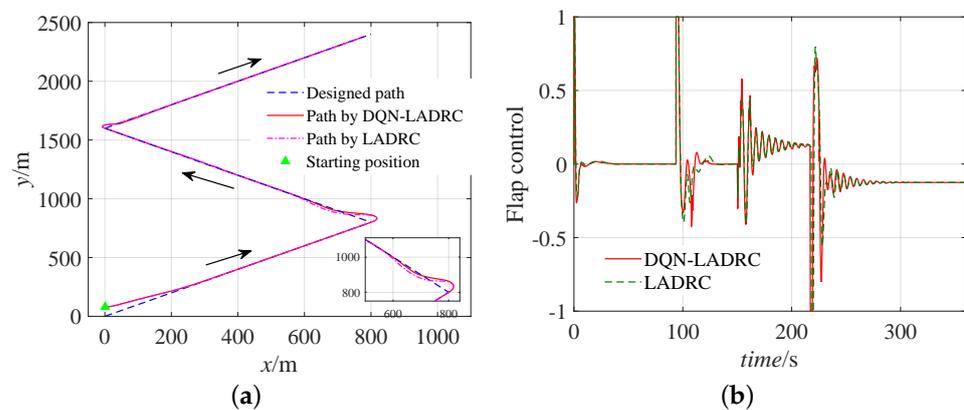


Figure 9. Straight-path-following control results under wind disturbances. (a) Path-following trajectories. (b) Control variables.

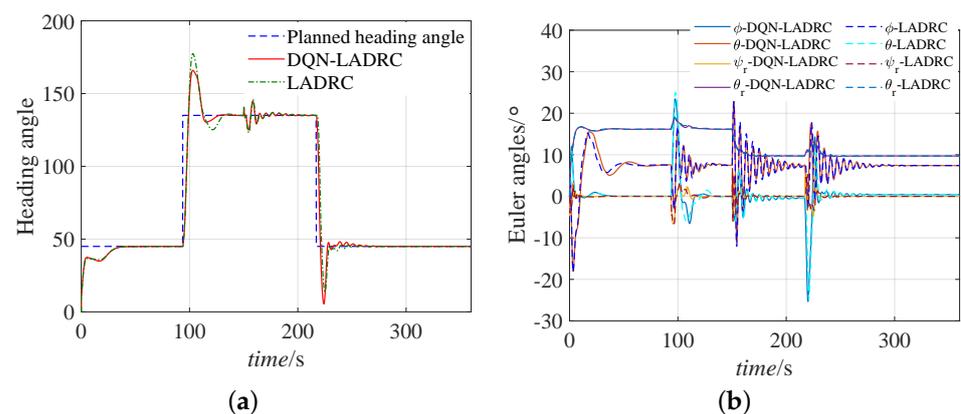


Figure 10. Straight-path-following control results of attitude angles under wind disturbances. (a) Heading angles. (b) Euler angles.

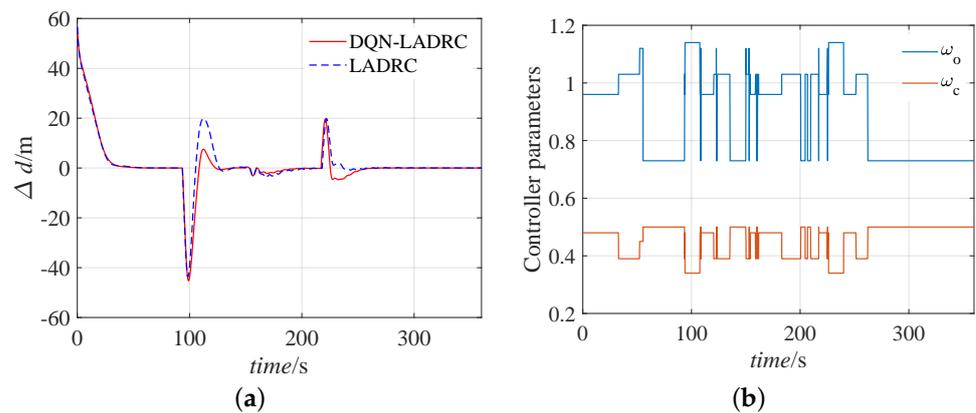


Figure 11. Tracking error and adaptive controller parameters. (a) Position derivation. (b) Controller parameters.

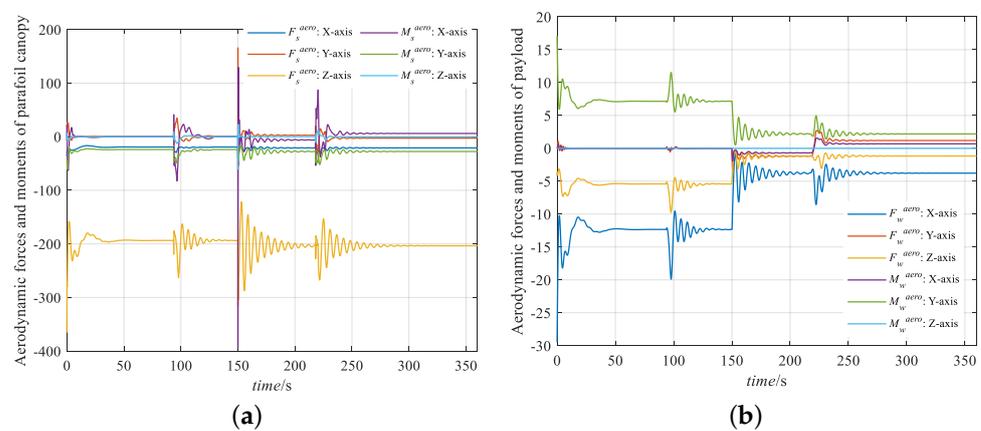


Figure 12. Changing process of aerodynamic forces and moments of parafoil canopy and payload for straight path following. (a) F_s^{aero} and M_s^{aero} . (b) F_w^{aero} and M_w^{aero} .

By comparing the tracking effect of the circular path and the straight path, it can be seen that the tracking effect of the straight path was better. This was because the guidance law is based on the idea of tracking a straight line. In circular path tracking, a circle is approximated as an infinite number of straight line segments.

6. Conclusions

This paper studied the lateral path-following control of the parafoil system. First, an eight-DOF dynamic model of the parafoil system was constructed, considering the relative motion between the parafoil canopy and the payload. Then, using the error conversion between the parafoil system's position information and the desired path, a guidance law was proposed, which essentially converts the path following into the control of the parafoil system's heading angle. Furthermore, a second-order LADRC was designed based on the guidance law, and the DQN was applied to optimize the controller parameters adaptively. Finally, the proposed DNQ-LADRC was applied to control the parafoil system to follow a circular path under environments with and without disturbances and a straight path under wind disturbances. The simulation results demonstrated that the influence of wind disturbances on the parafoil could not be ignored. The proposed method could overcome the effect of wind disturbances and realize the tracking control of straight or circular paths. Compared with the traditional LADRC controller, the proposed method has certain advantages in terms of the settling time and overshoot.

In future work, we will simultaneously consider implementing lateral and longitudinal path tracking control of the parafoil system and applying the proposed method to actual

experiments for verification. In addition, the DQN can realize the optimization of controller parameters, but it also requires human experience to divide the action space. The agent training of the DQN will be difficult when the action space is too large. Therefore, one of our goals is to explore a fully intelligent reinforcement learning algorithm. In this study, only the disturbances caused by uniform wind were considered, but in the actual flight process, the parafoil system will be affected by various complex wind fields. Thus, path-following control of the parafoil based on wind field identification is what we hope to achieve.

Author Contributions: Conceptualization, Y.Z. and J.T.; Methodology, Y.Z. and H.S.; Software, Y.Z.; Validation, Q.S. and Z.C.; Formal analysis, Y.Z.; Writing—original draft preparation, Y.Z.; Writing—review and editing, J.T.; Supervision, J.T., Q.S., H.S., M.S. and F.D.; Funding acquisition, J.T. 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 Nos. 61973172, 61973175, 62003175, and 62003177), the National Key Research and Development Project (Grant No. 2019YFC1510900), and the Key Technologies Research and Development Program of Tianjin (Grant No. 19JCZDJC32800). This project was also funded by China Postdoctoral Science Foundation (Grant No. 2020M670633) and Tianjin Postgraduate Research and Innovation Project 2021YJSB084.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable.

Conflicts of Interest: The authors declare that they do not have any conflict of interest.

References

- Zhu, H.; Sun, Q.; Liu, X.; Liu, J. Fluid–structure interaction-based aerodynamic modeling for flight dynamics simulation of parafoil system. *Nonlinear Dyn.* **2021**, *104*, 3445–3466. [[CrossRef](#)]
- Guo, Y.; Yan, J.; Wu, C.; Xiao, B. Modeling and practical fixed-time attitude tracking control of a paraglider recovery system. *ISA Trans.* **2022**, *128*, 391–401. [[CrossRef](#)] [[PubMed](#)]
- Slegers, N.; Costello, M. Model predictive control of a parafoil and payload system. *J. Guid. Control Dyn.* **2005**, *28*, 816–821. [[CrossRef](#)]
- Zhang, Z.; Zhao, Z.; Fu, Y. Dynamics analysis and simulation of six DOF parafoil system. *Clust. Comput.* **2018**, *22*, 12669–12680. [[CrossRef](#)]
- Zhu, E.; Sun, Q.; Tan, P.; Chen, Z. Modeling of powered parafoil based on Kirchhoff motion equation. *Nonlinear Dyn.* **2014**, *79*, 617–629. [[CrossRef](#)]
- Tan, P.; Sun, M.; Sun, Q.; Chen, Z. Dynamic modeling and experimental verification of powered parafoil with two suspending points. *IEEE Access* **2020**, *8*, 12955–12966. [[CrossRef](#)]
- Prakash, O. NDI based generic heading tracking control law for parafoil/payload system. In Proceedings of the AIAA Aviation 2020 Forum, Virtual Event, 15–19 June 2020. [[CrossRef](#)]
- Tao, J.; Dehmer, M.; Xie, G.; Zhou, Q. A generalized predictive control-based path following method for parafoil systems in wind environment. *IEEE Access* **2019**, *7*, 42586–42595. [[CrossRef](#)]
- Zhao, L.; He, W.; Lv, F. Model-free adaptive control for parafoil systems based on the iterative feedback tuning method. *IEEE Access* **2021**, *9*, 35900–35914. [[CrossRef](#)]
- Guo, Y.; Yan, J.; Wu, C.; Wu, X.; Chen, M.; Xing, X. Autonomous homing design and following for parafoil/rocket system with high-altitude. *J. Intell. Robot. Syst.* **2021**, *101*, 73. [[CrossRef](#)]
- Li, Y.; Zhao, M.; Yao, M.; Chen, Q.; Guo, R.; Sun, T.; Jiang, T.; Zhao, Z. 6-DOF modeling and 3D trajectory tracking control of a powered parafoil system. *IEEE Access* **2020**, *8*, 151087–151105. [[CrossRef](#)]
- Tao, J.; Sun, Q.; Sun, H.; Chen, Z.; Dehmer, M.; Sun, M. Dynamic modeling and trajectory tracking control of parafoil system in wind environment. *IEEE ASME Trans. Mechatron.* **2017**, *22*, 2736–2745. [[CrossRef](#)]
- Tao, J.; Liang, W.; Sun, Q.; Tan, P.; Luo, S.; Chen, Z.; He, Y. Modeling and control of a powered parafoil in wind and rain environments. *IEEE Trans. Aerosp. Electron. Syst.* **2017**, *53*, 1642–1659. [[CrossRef](#)]
- Han, J. From PID to active disturbance rejection control. *IEEE Trans. Ind. Electron.* **2019**, *56*, 900–906. [[CrossRef](#)]
- Han, J. Auto-disturbance-rejection controller and its applications. *Contr. Decis.* **1998**, *13*, 19–23.
- Gao, Z. On the foundation of active disturbance rejection control. *Control. Theory Appl.* **2013**, *30*, 1498–1510.
- Chen, Z.; Wang, Y.; Sun, M.; Sun, Q. Convergence and stability analysis of active disturbance rejection control for first-order nonlinear dynamic systems. *Trans. Inst. Meas. Control* **2019**, *41*, 2064–2076. [[CrossRef](#)]

18. Wang, Y.; Chen, Z.; Sun, M.; Sun, Q. On the stability and convergence rate analysis for the nonlinear uncertain systems based upon active disturbance rejection control. *Int. J. Robust Nonlinear Control* **2020**, *30*, 5728–5750. [[CrossRef](#)]
19. Zheng, Y.; Chen, Z.; Huang, Z.; Sun, M. Active disturbance rejection controller for multi-area interconnected power system based on reinforcement learning. *Neurocomputing* **2021**, *425*, 149–159. [[CrossRef](#)]
20. Zheng, Y.; Tao, J.; Sun, H.; Sun, Q. An intelligent course keeping active disturbance rejection controller based on Double Deep Q-network for towing system of unpowered cylindrical drilling platform. *Int. J. Robust Nonlinear Control* **2021**, *31*, 8463–8480. [[CrossRef](#)]
21. Tao, J.; Du, L.; Dehmer, M.; Wen, Y. Path following control for towing system of cylindrical drilling platform in presence of disturbances and uncertainties. *ISA Trans.* **2019**, *95*, 185–193. [[CrossRef](#)]
22. Zeng, D.; Yu, Z.; Xiong, L.; Fu, Z.; Li, Z. HFO-LADRC lateral motion controller for autonomous road sweeper. *Sensors* **2020**, *20*, 2274. [[CrossRef](#)] [[PubMed](#)]
23. Liu, C.; Luo, G.; Duan, X. Adaptive LADRC-based disturbance rejection method for electromechanical servo system. *IEEE Trans. Ind. Appl.* **2020**, *56*, 876–889. [[CrossRef](#)]
24. Li, R.; Li, T.; Bu, R.; Zheng, Q.; Chen, C.L. Active disturbance rejection with sliding mode control-based course and path following for underactuated ships. *Math. Probl. Eng.* **2013**, *2013*, 743716-1–743716-9. [[CrossRef](#)]
25. Zhao, P.; Nagamune, R. Switching linear parameter-varying control with improved local performance and optimized switching surfaces. *Int. J. Robust Nonlinear Control* **2018**, *28*, 3403–3421. [[CrossRef](#)]
26. Li, Y.; Jia, M.; Han, X. Towards a comprehensive optimization of engine efficiency and emissions by coupling artificial neural network (ANN) with genetic algorithm (GA). *Energy* **2021**, *225*, 120331. [[CrossRef](#)]
27. Ho, S.; Shu, L. Optimizing fuzzy neural networks for tuning PID controllers using an orthogonal simulated annealing algorithm OSA. *IEEE Trans. Fuzzy Syst.* **2006**, *14*, 421–434. [[CrossRef](#)]
28. Sun, C.; Liu, M.; Liu, C.A.; Feng, X.; Wu, H. An industrial quadrotor UAV control method based on fuzzy adaptive linear active disturbance rejection control. *Electronics* **2021**, *10*, 376. [[CrossRef](#)]
29. Mnih, V.; Kavukcuoglu, K.; Silver, D. Human-level control through deep reinforcement learning. *Nature* **2015**, *518*, 529–533. [[CrossRef](#)]
30. Zhao, Y.; Qi, X.; Ma, Y.; Li, Z.; Malekian, R.; Sotelo, M.A. Path following optimization for an underactuated USV using smoothly-convergent deep reinforcement learning. *IEEE Trans. Intell. Transp. Syst.* **2020**, *22*, 1–13. [[CrossRef](#)]
31. Xue, W.; Huang, Y. Performance analysis of active disturbance rejection tracking control for a 513 class of uncertain LTI systems. *ISA Trans.* **2015**, *58*, 133–154. [[CrossRef](#)]
32. Sun, Q.; Yu, L.; Zheng, Y.; Tao, J.; Sun, H.; Sun, M.; Dehmer, M.; Chen, Z. Trajectory tracking control of powered parafoil system based on sliding mode control in a complex environment. *Aerosp. Sci. Technol.* **2022**, *122*, 107406. [[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.