




Article

Intelligent Trajectory Tracking Linear Active Disturbance Rejection Control of a Powered Parafoil Based on Twin Delayed Deep Deterministic Policy Gradient Algorithm Optimization

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

¹ College of Artificial Intelligence, Nankai University, Tianjin 300350, China; zhengyueminyln@163.com (Y.Z.); sunh@nankai.edu.cn (H.S.); chenqz@nankai.edu.cn (Z.C.); smw_sunmingwei@163.com (M.S.)

² Beijing Institute of Spacecraft Environment Engineering, Beijing 100081, China; feizelin0901@163.com

³ Silo AI, 00100 Helsinki, Finland

* Correspondence: taoj@nankai.edu.cn (J.T.); sunql@nankai.edu.cn (Q.S.)

Abstract: Powered parafoils, known for their impressive load-bearing capacity and extended endurance, have garnered significant interest. However, the parafoil system is a highly complex nonlinear system. It primarily relies on the steering gear to change flight direction and utilizes a thrust motor for climbing. However, achieving precise trajectory tracking control presents a challenge due to the interdependence of direction and altitude control. Furthermore, underactuation and wind disturbances bring additional difficulties for trajectory tracking control. Consequently, realizing trajectory tracking control for powered parafoils holds immense significance. In this paper, we propose a trajectory tracking method based on Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm-optimized Linear Active Disturbance Rejection Control (LADRC). Our method addresses the underactuation issue by incorporating a guiding law while utilizing two LADRC methods to achieve decoupling and compensate for disturbances. Moreover, we employ the TD3 algorithm to dynamically adjust controller parameters, thus enhancing the controller performance. The simulation results demonstrate the effectiveness of our proposed method as a trajectory tracking control approach. Additionally, since the control process is not reliant on system-specific models, our method can also provide guidance for trajectory tracking control in other aircraft.

Keywords: trajectory tracking control; powered parafoil system; linear active disturbance rejection control; twin delayed deep deterministic policy gradient



Citation: Zheng, Y.; Fei, Z.; Tao, J.; Sun, Q.; Sun, H.; Chen, Z.; Sun, M. Intelligent Trajectory Tracking Linear Active Disturbance Rejection Control of a Powered Parafoil Based on Twin Delayed Deep Deterministic Policy Gradient Algorithm Optimization. *Appl. Sci.* **2023**, *13*, 12555. <https://doi.org/10.3390/app132312555>

Academic Editors: Seppo Sierla, David Hästbacka and Kai Zenger

Received: 5 September 2023

Revised: 8 November 2023

Accepted: 16 November 2023

Published: 21 November 2023



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

1. Introduction

The parafoil unmanned aerial vehicle (UAV) is a type of flexible aircraft. A conventional parafoil comprises a canopy, parachute, and payload, enabling it to carry out tasks such as airdrops and aircraft recovery. With the development of autonomous aircraft technology, powered parafoils appeared, incorporating thrust devices into the traditional parafoil design, thus enhancing their endurance capabilities [1]. Hence, powered parafoils possess the capability to execute precise tasks, such as stationary airdrops, and effective trajectory tracking control is paramount for mission success. Nevertheless, the challenges of trajectory tracking control encompass wind disturbances, underactuation, coupling, and unexpected dynamics.

A mathematical model is necessary for analyzing the motion characteristics of a parafoil system given the limitations imposed by actual flight tests, which involve substantial preparation work, time, and expenses. Currently, there are numerous ways to express models of powered parafoils. From a dynamics standpoint, this includes three-degree-of-freedom (DOF) modeling [2], four-DOF modeling [3], six-DOF modeling [4], eight-DOF modeling [5], and nine-DOF modeling [6]. Among them, the eight-DOF model takes into account the slew, sway, yaw, heave, pitch, and roll motions, as well as the relative pitch

and relative yaw motion between the parafoil canopy and the payload, which effectively describe the motion state of the actual parafoil [7].

Trajectory tracking control of powered parafoils is a challenging task. The main difficulties arise from three aspects. Firstly, the actual model is more intricate than the constructed model, and there are unmodeled dynamics. Secondly, parafoil trajectory control is achieved by controlling the steering gear during flight, while altitude control is achieved through thrust, resulting in coupling and underactuation between these two control processes. Thirdly, during flight, the system is susceptible to wind disturbances, potentially leading to severe loss of control. Taking into account the issues mentioned above, research has been conducted. However, most of the literature only addresses the problem of path following [8,9]. In our opinion, the main difference between implementing trajectory tracking and path following lies in the design of the guidance law. The guidance law plays a crucial role in expressing the tracking error in alternative forms and effectively addressing the challenges of underactuation. However, in path following control, the guidance law typically overlooks the consideration of forward tracking errors [10]. For example, Sun et al. [11] formulated the guidance law as a sliding mode surface and employed linear extended state observer (LESO) to decouple horizontal path following and height tracking. Subsequently, they accomplished a three-dimensional path following control by utilizing sliding mode controllers (SMCs). Similarly, Li et al. [12] also utilized an SMC for path following control. Guo et al. [13] utilized the barrier Lyapunov function and backstepping method, and developed an adaptive path-following controller. Zheng et al. [14] introduced a horizontal path following guidance law that utilized the lateral tracking error and yaw angle, enabling control of lateral tracking and the yaw angle through flap deflection. However, the study overlooked the parafoil's sideslip angle. There are very few research results on parafoil trajectory tracking control. Li et al. [15] transformed tracking errors into guidance commands for the yaw angle and glide slope angle and employed PID controllers to achieve trajectory tracking control. For a powered parafoil without forward thrust, we proposed a new guidance law and initially realized trajectory tracking control [16]. When selecting a controller, apart from the aforementioned SMCs and PID controllers, other control methods are available, such as LADRC. LADRC is a control method that has been developed based on PID. It operates independently of model information, possesses inherent decoupling capabilities, and is straightforward to implement in engineering applications [17,18]. As a result, LADRC has demonstrated favorable outcomes in path-following control [19–21].

Parameter selection is a vital aspect of the controller configuration process as it directly impacts the tracking accuracy of the controller. Researchers commonly rely on manual tuning of controller parameters, which presents challenges in achieving optimal system performance. Consequently, there has been a continuous emergence of various optimization algorithms aimed at addressing this issue. For example, the heuristic algorithm, exemplified by particle swarm optimization (PSO) [22] and the genetic algorithm [23], is employed to optimize a set of fixed parameters of the controller [24,25]. Fuzzy control and neural networks are utilized to obtain adaptive controller parameters [26,27]. However, fuzzy control heavily relies on model information, while neural networks face challenges in making appropriate decisions based on state changes. Therefore, deep reinforcement learning (DRL) algorithms [28] that do not rely on model information and can make autonomous decisions have been promoted. DRL is a kind of algorithm that combines the computational capabilities of neural networks with the decision-making abilities of RL. Its primary objective is to train optimal decision making through continuous interactions between the agent and the environment. Currently, in the field of DRL, several algorithms have been developed, including Deep Q-Networks (DQNs) [29] and Deep Deterministic Policy Gradient (DDPG) [30]. DQNs are well suited to discrete action spaces, whereas DDPG effectively handles continuous action spaces. The Twin Delayed Deep Deterministic Policy Gradient (TD3) algorithm was developed based on DDPG with enhanced stability [31]. Current applications of the TD3 algorithm in motion control can be delineated into two main categories. The first category entails employing the TD3 algorithm

directly as a controller to optimize control variables [32,33]. The second category involves integrating the TD3 algorithm with the designed controller to enhance its intelligence and performance [34]. Comparatively, the second approach exhibits noticeable improvements in training efficiency and contributes to the advancement of intelligent control systems.

Motivated by prior research, this paper presents a TD3-optimized LADRC method for the trajectory control of a powered parafoil system. The main contributions of this paper are summarized as follows:

- A guidance law for the trajectory tracking of a powered parafoil is proposed, which effectively addresses the underactuation of the system.
- Based on the guidance law, two LADRC controllers aiming to achieve 3D trajectory tracking control are designed to address challenges stemming from wind disturbances and the coupling between horizontal trajectory and altitude.
- The TD3 algorithm is employed to acquire the controller's real-time parameters by leveraging the powered parafoil's flight states. The TD3-optimized LADRC was compared against the traditional LADRC control method using simulation results, demonstrating the effectiveness of the proposed approach.

This paper is organized as follows: Section 2 establishes the eight-DOF model for the powered parafoil system. Section 3 introduces the guidance law and outlines the design process of the LADRC for horizontal trajectory and altitude control. The design process of the TD3-optimized LADRC is presented in Section 4. Section 5 presents the simulation results, and Section 6 concludes the paper.

2. Dynamic Modeling Introduction of Powered Parafoil

To provide a more precise depiction of the parafoil's motion characteristics, this paper adopts an eight-DOF dynamic model [5]. The model encompasses the parafoil canopy's slew, sway, yaw, heave, pitch, and roll movements, along with the relative pitch and yaw motions between the parafoil canopy and the payload. To facilitate the modeling of the parafoil system, three coordinate systems are introduced: the ground coordinate system, denoted as $O_d x_d y_d z_d$; the parafoil coordinate system, denoted as $O_s x_s y_s z_s$; and the payload coordinate system, denoted as $O_w x_w y_w z_w$. These coordinate systems are illustrated in Figure 1.

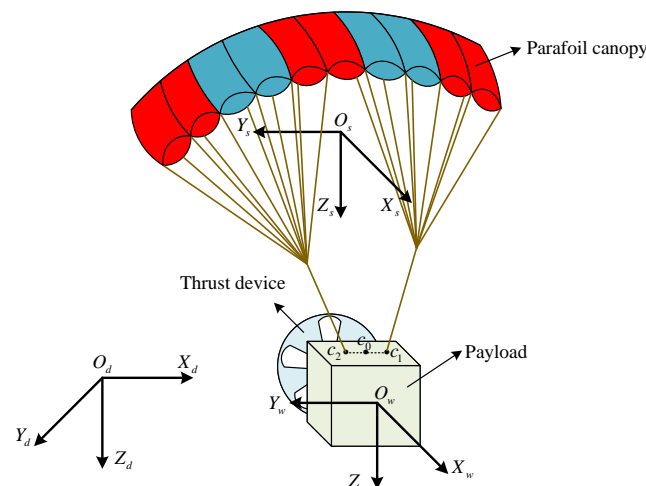


Figure 1. Schematic diagram of parafoil system coordinates.

First of all, based on Figure 1, we can observe that a powered parafoil system comprises a parafoil canopy, a payload, parafoil ropes, and a thrust device. During flight, the paracord can adjust its direction by pulling it either to the left or right. The thrust device, powered by the propeller, generates the necessary upward lift for the parafoil. Hence, compared to the conventional unpowered parafoil, the powered parafoil investigated in this study

exhibits enhanced flexibility. Subsequently, we will briefly introduce the dynamic equation of the powered parafoil.

By analyzing the forces acting on the canopy and the payload, the following momentum equation can be obtained,

$$\begin{cases} \frac{\partial P_s}{\partial t} + W_s \times P_s = F_s^{aero} + F_s^G + F_s^t \\ \frac{\partial P_w}{\partial t} + W_w \times P_w = F_w^{aero} + F_w^G + F_w^t + F_w^{th} \end{cases} \quad (1)$$

where the subscripts s and w refer to the parafoil canopy and the payload, respectively, indicating the variables associated with each component. P_s and P_w represent the momentums, which are calculated as the product of the mass and velocity of the canopy or payload, respectively. In addition, the right side of the equation represents the force exerted on both the canopy and the payload. The canopy experiences the combined effects of aerodynamic force F_s^{aero} , gravity F_s^G , and the tension in the connecting ropes F_s^t . Similarly, the payload undergoes these three forces, but it is further influenced by the propeller's thrust F_w^{th} . The velocity vectors are denoted as $V_s = [u_s, v_s, w_s]^T$ and $V_w = [u_w, v_w, w_w]^T$, respectively. Similarly, the angular velocity vectors are represented by $W_s = [p_s, q_s, r_s]^T$ and $W_w = [p_w, q_w, r_w]^T$. For the specific parameters of the model in this paper, please refer to [5,35]. Equation (2) is the momentum moment equation corresponding to Equation (1),

$$\begin{cases} \frac{\partial H_s}{\partial t} + W_s \times H_s + V_s \times P_s = M_s^{aero} + M_s^G + M_s^f + M_s^t \\ \frac{\partial H_w}{\partial t} + W_w \times H_w = M_w^{aero} + M_w^f + M_w^t \end{cases} \quad (2)$$

where the moment of momentum H is obtained by multiplying the moment of inertia matrix with the angular velocity.

Indeed, by utilizing Equations (1) and (2), we can derive the dynamic equation of the powered parafoil system. This equation encompasses the state variables V_s , V_w , W_s and W_w , along with the relative yaw angle ψ_r and relative pitch angle θ_r existing between the canopy and the payload. Furthermore, the system comprises two control variables: flap deflection, expressed as u_1 ($-10 \leq u_1 \leq 10$ cm), and thrust, expressed as u_2 ($0 \leq u_2 \leq 120$ N). u_1 and u_2 represent the variables for flap deflection and thrust, respectively. The acceptable range for these variables is determined by the capabilities of the steering gear and thrust motor. The saturation value is intentionally set to safeguard the equipment. Flap deflection alters the flight direction, while thrust generates lift for the parafoil, enabling climbing. It should be pointed out that $u_1 < 0$ means pulling down the left paracord, and $u_1 > 0$ means pulling down the right paracord.

In general, the dynamic equation of the powered parafoil can be briefly described by Equation (3)

$$\dot{x}_s = f(x_s, u_1, u_2) \quad (3)$$

with $x_s = [V_w^T, W_w^T, V_s^T, W_s^T, \psi_r, \theta_r]^T$.

Based on the kinetic equations, the kinematic equation can be further obtained,

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} \cos\theta\cos\psi & \cos\theta\sin\psi & -\sin\theta \\ \sin\phi\sin\theta\cos\psi - \cos\phi\sin\psi & \sin\phi\sin\theta\sin\psi + \cos\phi\cos\psi & \sin\phi\cos\theta \\ \cos\phi\sin\theta\cos\psi + \sin\phi\sin\psi & \cos\phi\sin\theta\sin\psi - \sin\phi\cos\psi & \cos\phi\cos\theta \end{bmatrix}^T V_s \quad (4)$$

where (x, y, z) are the position coordinates of the parafoil in the earth's coordinate system. Additionally, the roll, pitch, and yaw angles are represented by ϕ ($\phi \in [-\pi, \pi]$), θ ($\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$), and ψ ($\psi \in [-\pi, \pi]$), respectively. Equation (5) establishes the relationship between these Euler angles and the angular velocity. Furthermore, it should be noted that

the pitch angle of the parafoil will not reach $\pm 90^\circ$ during flight, so there is no need to consider the issue of singular values.

$$\begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} 1 & \sin \phi \tan \theta & \cos \phi \tan \theta \\ 0 & \cos \phi & -\sin \phi \\ 0 & \sin \phi / \cos \theta & \cos \phi / \cos \theta \end{bmatrix} W_s \quad (5)$$

The above is an introduction to the model of the powered parafoil system, while the system's flight process is elucidated through dynamic equations. Furthermore, it is noteworthy that the flight state of the parafoil can be dynamically changed by implementing both the steering gear and the thrust device. As a result, this capability enables the parafoil to adeptly accomplish a wide range of predetermined trajectory tasks with remarkable flexibility. However, during the control process, when comparing the built simulation model with the actual system, the simulation model may contain uncertain factors, such as unmodeled dynamics, internal parameter perturbations, and external environmental interferences. Consequently, it becomes imperative for the designed control system to address these challenges effectively.

3. Trajectory Tracking Controller Design Based on LADRC

The trajectory tracking of a powered parafoil system faces an underactuation problem, requiring the utilization of two control variables to facilitate movement along the three coordinate axes, as shown in Figure 2. Hence, it becomes imperative to employ the guidance law before controller design to surmount the challenge posed by underactuation.

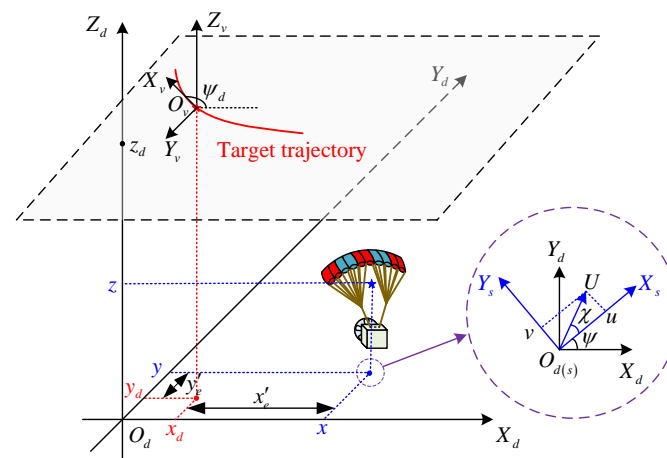


Figure 2. Schematic diagram of trajectory tracking for a powered parafoil system.

3.1. Guidance Law Design for Horizontal Trajectory

In Figure 2, (x, y, z) represents the current position of the parafoil system, while $(x_d(\omega), y_d(\omega), z_d(\omega))$ denotes the target position, where ω is the trajectory variable. By considering (x_d, y_d, z_d) as the origin, we establish a target coordinate system denoted as $O_v X_v Y_v Z_v$. Then, the horizontal trajectory tracking errors x_e and y_e in $O_v X_v Y_v Z_v$ can be expressed as

$$\begin{cases} x_e = (x - x_d) \cos(\psi_d) + (y - y_d) \sin(\psi_d) \\ y_e = -(x - x_d) \sin(\psi_d) + (y - y_d) \cos(\psi_d) \end{cases} \quad (6)$$

where $\psi_d(\omega) = \arctan(\dot{y}_d(\omega) / \dot{x}_d(\omega))$.

In this paper, the guidance law for the horizontal trajectory is designed as

$$\psi_g = \psi_d + \arctan(-\alpha y_e) - \chi \quad (7)$$

where α is a positive constant, ψ_g is the guided yaw angle, and $\chi = \arctan(v/u)$ is the sideslip angle. Then, the following theorem arises.

Theorem 1. *Under the premise that the actual yaw angle ψ tracks the guided yaw angle ψ_g very well, the horizontal trajectory tracking error can gradually converge to 0 with the desired forward speed*

$$u_d = U \cos(\psi + \chi - \psi_d) + kx_e \quad (8)$$

where $k > 0$ and $U = \sqrt{u^2 + v^2}$.

Proof of Theorem 1. Construct a Lyapunov function as $V = \frac{1}{2}x_e^2 + \frac{1}{2}y_e^2$; then, its derivative is derived as

$$\dot{V} = x_e \dot{x}_e + y_e \dot{y}_e \quad (9)$$

According to Equation (6), the following expression can be derived,

$$\begin{cases} \dot{x}_e = U \cos(\psi + \chi - \psi_d) - u_d + \dot{\psi}_d y_e \\ \dot{y}_e = U \sin(\psi + \chi - \psi_d) - \dot{\psi}_d x_e \end{cases} \quad (10)$$

where the desired forward speed has the following expression:

$$u_d = \dot{\omega} \sqrt{\dot{x}_d^2(\omega) + \dot{y}_d^2(\omega)} \quad (11)$$

With Equation (10), Equation (9) can be rearranged as

$$\dot{V} = x_e U \cos(\psi + \chi - \psi_d) - x_e u_d + y_e U \sin(\psi + \chi - \psi_d) \quad (12)$$

Further, with $\psi = \psi_g$ and substituting Equation (8) into Equation (12), the following inequality can be obtained

$$\begin{aligned} \dot{V} &= -kx_e^2 + y_e U \sin(\arctan(-\alpha y_e)) \\ &= -kx_e^2 - \alpha y_e^2 \\ &\leq 0 \end{aligned} \quad (13)$$

□

Remark 1. *The guidance law presented in Equation (7) effectively merges the lateral error y_e with the yaw angle ψ to achieve convergence of the lateral error. By altering the flight direction, this approach effectively addresses the issue of underactuation. Furthermore, by combining Equation (8) and Equation (11), we can derive the following rules for $\dot{\omega}$:*

$$\dot{\omega} = \frac{U \cos(\psi + \chi - \psi_d) + kx_e}{\sqrt{\dot{x}_d^2(\omega) + \dot{y}_d^2(\omega)}} \quad (14)$$

To conclude, when the yaw angle ψ can accurately track the guided yaw angle ψ_g stated in Equation (7), and the parameter ω satisfies Equation (14), it is possible to achieve horizontal trajectory tracking.

3.2. LADRC Design Process for Trajectory Tracking

Based on the guidance of the above law, it is necessary to develop a corresponding controller that can effectively regulate the steering gear of the parafoil system. This controller is essential for ensuring accurate tracking of the guidance yaw angle. Additionally, another controller is required to calculate the appropriate thrust, enabling the parafoil system to maintain a stable altitude during flight. In this paper, an LADRC controller is chosen for several reasons. One notable advantage of LADRC is its independence from

model information, allowing it to handle unknown disturbances effectively. Moreover, LADRC exhibits inherent decoupling characteristics, effectively addressing the challenge of coupling between horizontal trajectory tracking and altitude tracking. This makes LADRC a suitable choice for the control system design in this study.

3.2.1. Horizontal Trajectory Tracking Controller

LADRC eliminates the need for detailed knowledge about the controlled object, as it solely requires the order relationship between the output and input variables. According to Equation (5), the following expression about ψ can be obtained,

$$\ddot{\psi} = \frac{\sin \phi}{\cos \theta} \dot{q} + \frac{\cos \phi}{\cos \theta} \dot{r} + \frac{\sin \theta \sin 2\phi}{\cos^2 \theta} q^2 - \frac{\sin \theta \sin 2\phi}{\cos^2 \theta} r^2 + \frac{\cos \phi}{\cos \theta} pq - \frac{\sin \phi}{\cos \theta} pr + \frac{2 \sin \theta \cos 2\phi}{\cos^2 \theta} qr \quad (15)$$

Then, combining Equations (3) and (15), we obtain the following second-order system,

$$\ddot{\psi} = f_1(\cdot) + f_2(u_1) \quad (16)$$

where, due to the complexity of the model, we use the abbreviated terms f_1 and f_2 to express the system. This does not impact the design of the controller.

Furthermore, let $y = \psi - \psi_g$; then, there is

$$\ddot{y} = f_1(\cdot) + f_2(u_1) - \ddot{\psi}_g \quad (17)$$

To facilitate the design of the controller, Equation (17) is rearranged to

$$\begin{aligned} \ddot{y} &= f_1(\cdot) + f_2(u_1) - \ddot{\psi}_g - b_{01}u_1 + b_{01}u_1 \\ &= f_h + b_{01}u_1 \end{aligned} \quad (18)$$

where f_h is the total disturbance and b_{01} is an adjustable parameter.

By defining the state as $x_{11} = y$, $x_{12} = \dot{y}$, and $x_{13} = f_h$, the following state space equation can be obtained,

$$\begin{cases} \dot{x}_{11} = x_{12} \\ \dot{x}_{12} = x_{13} + b_{01}u_1 \\ \dot{x}_{13} = \dot{f}_h \\ y = x_{11} \end{cases} \quad (19)$$

Consequently, a full-order LESO is utilized to estimate the above state,

$$\begin{cases} \dot{\hat{x}}_{11} = \hat{x}_{12} + \beta_{01}(y - \hat{x}_{11}) \\ \dot{\hat{x}}_{12} = \hat{x}_{13} + b_{01}u_1 + \beta_{02}(y - \hat{x}_{11}) \\ \dot{\hat{x}}_{13} = \beta_{03}(y - \hat{x}_{11}) \end{cases} \quad (20)$$

where \hat{x}_{11} , \hat{x}_{12} , and \hat{x}_{13} are the estimated values of x_{11} , x_{12} , and x_{13} , respectively. β_{01} , β_{02} , and β_{03} are observer gains, and these three parameters are related to the accuracy of the estimated values. Usually, the pole configuration method is used to configure the observer gain parameters to the pole $-\omega_{o1}$ ($\omega_{o1} > 0$) that can make the system stable, that is, $\beta_{01} = 3\omega_{o1}$, $\beta_{02} = 3\omega_{o1}^2$, and $\beta_{03} = \omega_{o1}^3$. In this case, only the single parameter ω_o must be adjusted to align the observed values with the actual values.

When the observed states are accurate, we have $\hat{x}_{13} \approx f_h$. Then, we can employ a PD control law to mitigate the disturbance,

$$u_1 = \frac{k_{p1}(y_d - \hat{x}_{11}) - k_{d1}\hat{x}_{12} - \hat{x}_{13}}{b_{01}} \quad (21)$$

where k_{p1} and k_{d1} are controller parameters; y_d is the desired value of y . $y_d = 0$. In addition, by substituting Equation (21) into Equation (18), we observe that the estimated disturbance value effectively compensates for the unknown disturbance within the system. This fundamental concept lies at the core of the LADRC algorithm. In addition, the stability analysis of an LADRC control system is provided in the Appendix A.

3.2.2. Altitude Controller

Altitude control is comparatively more straightforward than horizontal trajectory tracking since it does not require the implementation of guidance laws. Nevertheless, the design process for the altitude controller follows a similar approach to that of the horizontal controller, as both require knowledge of the system order.

According to Equation (4),

$$\dot{z} = -u_s \sin \theta + v_s \sin \phi \cos \theta + w_s \cos \phi \cos \theta \quad (22)$$

In fact, according to ref. [5], there exists a relationship between the parafoil canopy's speed and the thrust u_2 , which can be expressed as follows,

$$\begin{cases} \dot{v}_s = \frac{\cos \theta_r \sin \psi_r}{m_s + m_{a,11}} u_2 + \dots \\ \dot{u}_s = \frac{\cos \theta_r \cos \psi_r}{m_s + m_{a,22}} u_2 + \dots \\ \dot{w}_s = \frac{-\sin \theta_r}{m_s + m_{a,33}} u_2 + \dots \end{cases} \quad (23)$$

where the ellipsis represents items not related to u_2 .

Then, further deriving Equation (22), it can be deduced that

$$\begin{aligned} \ddot{z} &= -\dot{u}_s \sin \theta + \dot{v}_s \sin \phi \cos \theta + \dot{w}_s \cos \phi \cos \theta - (u_s \cos \theta + v_s \sin \phi \sin \theta + w_s \cos \phi \sin \theta) \dot{\theta} \\ &\quad - (w_s \cos \theta \sin \phi - v_s \cos \theta \cos \phi) \dot{\phi} \\ &= \left(-\frac{\cos \theta_r \cos \psi_r \sin \theta}{m_s + m_{a,22}} + \frac{\sin \phi \cos \theta \cos \theta_r \sin \psi_r}{m_s + m_{a,11}} - \frac{\sin \theta_r \cos \phi \cos \theta}{m_s + m_{a,33}} \right) u_2 + \dots \end{aligned} \quad (24)$$

As a result, a second-order system can also be obtained:

$$\ddot{z} = f_a + b_{02} u_2 \quad (25)$$

where f_a represents the total unknown disturbance in this system.

Next, we can proceed with the sequential design of the LESO and PD control law based on Equations (19)–(21) mentioned earlier. However, the detailed description of this process is beyond the scope of this discussion.

In summary, the trajectory tracking control of a powered parafoil in this paper depends on two LADRC controllers, each requiring the adjustment of four parameters. Hence, the parameters that necessitate adjustment in the control system are as follows: ω_{o1} , k_{p1} , k_{d1} , b_{01} for the horizontal controller, and ω_{o2} , k_{p12} , k_{d2} , b_{02} for the height controller. These parameters will undoubtedly affect the trajectory tracking accuracy. To explore the parameter adjustment rules within the LADRC controller, interested readers are encouraged to consult ref. [36].

4. Optimized LADRC Approach Using TD3

Parameter tuning is indeed a sequential decision-making problem, and DRL algorithms have been proven as an effective way to solve it. The TD3 algorithm is a deep reinforcement learning algorithm designed to handle problems with a continuous action space. This paper focuses on adaptively varying the parameters, making the TD3 algorithm an ideal choice for optimizing the controller parameters. Additionally, it is worth noting

that ref. [37] highlights the significance of ω_0 and b_0 in ensuring the stability of the LESO. Therefore, this paper solely focuses on optimizing the parameters within the control law, namely k_p and k_d .

4.1. The Basics of the TD3 Agent

The TD3 algorithm fundamentally derives the optimal strategy by facilitating continuous interaction between the agent and the environment. This paper uses a powered parafoil system with controllers as the environment. At the same time, the agent functions as the algorithm's cognitive component, responsible for decision-making processes, as shown in Figure 3. Specifically, the agent is situated in a specific state, denoted as $s_t \in S$, at each time step. Depending on the state transition probability $P_{ss'}^a = P[s_{t+1} = s' | s_t = s, a_t = a]$ and the action value $a \in A$, the agent undergoes a transition from the current state s_t to the subsequent state s_{t+1} with a certain probability. Simultaneously, the agent receives an immediate reward $r_{t+1} \in R$ associated with this transition. S and A are the state space and action space, respectively.

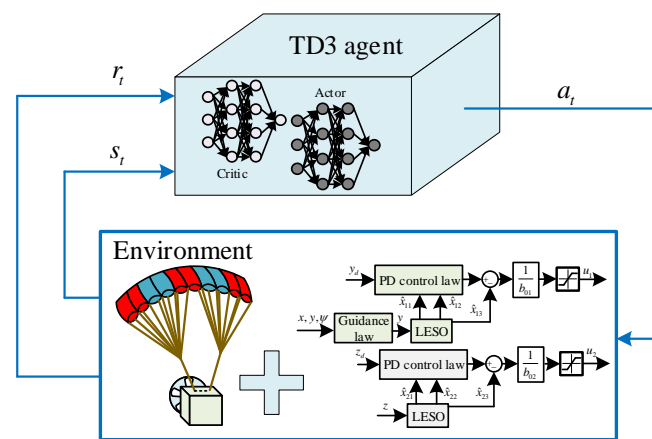


Figure 3. Schematic diagram of the interaction process.

The goal of the TD3 algorithm is to find the optimal strategy. This, of course, necessitates receiving feedback in the form of reward values. Consequently, the optimal strategy is attained by maximizing the expected value of the cumulative reward, which can be expressed as

$$\pi^* = \max_{\pi} E_{\tau \sim \pi} [R_c(\tau)] \quad (26)$$

where $R_c = \sum_{t=0}^{\infty} \gamma^t r_{t+1}$ is the cumulative reward, and the discount factor $\gamma \in (0, 1]$ is employed to represent the importance of future reward values.

The TD3 agent consists of several key components, namely a replay buffer, two critic networks, an actor network, and their respective target critic networks and target actor network, as illustrated in Figure 4. On the one hand, the critic network estimates the Q value expressed in Equation (27) by taking the state and action as inputs and generating the corresponding Q value as output, a common approach in DRL algorithms. On the other hand, the actor network is updated using the policy gradient method, which involves computing the gradient of the expected Q value concerning the actor network parameters. The parameters of the networks are represented by $\theta(\theta')$ for the critic networks and $\phi(\phi')$ for the actor network.

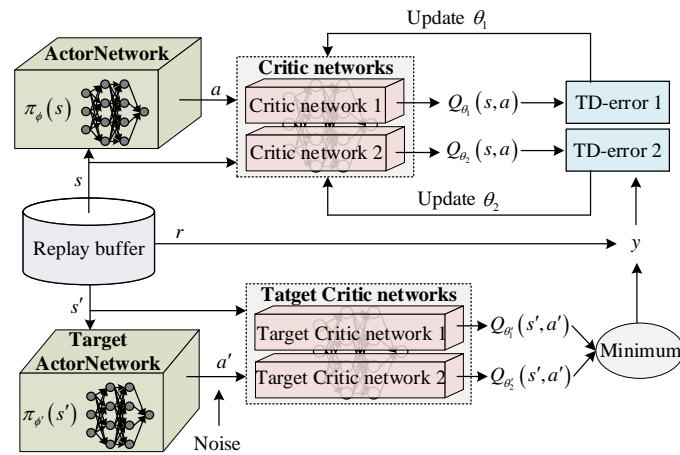


Figure 4. Structure diagram of the TD3 agent.

$$Q_{\pi}(s, a) = E_{\pi}[R_c | s_t = s, a_t = a] \quad (27)$$

The training process of the TD3 algorithm is shown in Algorithm 1.

Algorithm 1 TD3 Algorithm

Initialize critic networks $Q_{\theta_1}(s, a)$, $Q_{\theta_2}(s, a)$ and actor network $\pi_{\phi}(s)$ with random parameters θ_1, θ_2 and ϕ .

Initialize target networks $Q_{\theta'_1}(s, a)$, $Q_{\theta'_2}(s, a)$ and $\pi_{\phi'}(s)$ with weights $\theta'_1 \leftarrow \theta_1, \theta'_2 \leftarrow \theta_2$ and $\phi' \leftarrow \phi$.

Initialize replay buffer \mathcal{D} .

if $t \leq T$ **then**

 Select action with exploration noise $a \sim \pi_{\phi}(s) + \epsilon, \epsilon \sim \mathcal{N}(0, \sigma)$, observe reward r and new states s' .

 Store transition tuple (s, a, r, s') in \mathcal{D} .

 Sample mini-batch of m transitions (s, a, r, s') from \mathcal{D} .

$a' \leftarrow \pi_{\phi'}(s') + \epsilon, \epsilon \sim \text{clip}(\mathcal{N}(0, \sigma'), -c, c)$.

$y \leftarrow r + \gamma \min_{i=1,2} Q_{\theta'_i}(s', a')$.

 Update critics by $\theta_i \leftarrow \arg \min_{\theta_i} \frac{1}{m} \sum (y - Q_{\theta_i}(s, a))^2$.

if $t \bmod \kappa$ **then**

 Update ϕ by the deterministic policy gradient: $\nabla_{\phi} J(\phi) = \frac{1}{m} \sum \nabla_a Q_{\theta_1}(s, a) \Big|_{a=\pi_{\phi}(s)} \nabla_{\phi} \pi_{\phi}(s)$.

 Update target networks by moving average method: $\theta'_i \leftarrow \tau \theta_i + (1 - \tau) \theta'_i, \phi' \leftarrow \tau \phi + (1 - \tau) \phi'$.

end if

end if

4.2. Agent Design for Powered Parafoil System

Based on the introduction to the TD3 agent mentioned earlier, we will proceed with the design of the state space, action space, and reward function of the agent to address the powered parafoil system's trajectory tracking control problem.

The action variables within the agent are unquestionably parameters that require optimization, namely, k_{p1} and k_{d1} for the horizontal controller and k_{p2} and k_{d2} for the altitude controller. Then, the four-dimensional action space can be expressed as

$$\{a_1, a_2, a_3, a_4 \in A | a_1 = k_{p1}, a_2 = k_{d1}, a_3 = k_{p2}, a_4 = k_{d2}\} \quad (28)$$

State variables are pieces of information that the agent can directly acquire from the environment. For the trajectory tracking problem, the trajectory tracking errors directly indicate the controller's performance. Therefore, the state space is constructed as

$$\{s_1, s_2, s_3, s_4 \in S | s_1 = y, s_2 = \dot{y}, s_3 = z - z_d, s_4 = \dot{s}_3\} \quad (29)$$

In Equation (29), the impact of horizontal trajectory tracking is observed through variable y in Equation (17), with an expected value of 0. Based on the state space, it is necessary to design the reward function. The fundamental principle guiding the design of the reward function is to assign a reward when the state value approximates the desired state, otherwise applying a penalty. In this paper, the reward function is designed as follows:

$$r = -|s_1| - 5|s_3| + a + b \quad (30)$$

where

$$a = \begin{cases} 2 \tanh(1/|s_1|), & \text{If } \exists (s_1 \leq 0.5 \& s_2 \leq 0.05) \\ 2 \tanh(1/|s_3|), & \text{If } \exists (s_3 \leq 5 \& s_4 \leq 0.5) \\ 0, & \text{Otherwise} \end{cases} \quad (31)$$

$$b = \begin{cases} -2, & \text{If } \exists s_1 s_2 > 0 \& s_3 s_4 > 0 \\ 0, & \text{Otherwise} \end{cases} \quad (32)$$

In the designed reward function, the terms a and b correspond to the reward and penalty components, respectively. The two terms are employed to encourage minimal trajectory tracking errors.

Once the action space and state space have been established, the number of input and output neurons for both the critic network and actor network of the TD3 agent can be determined. To enhance the visual representation of the network structure, Figure 5 is utilized to present it more intuitively.

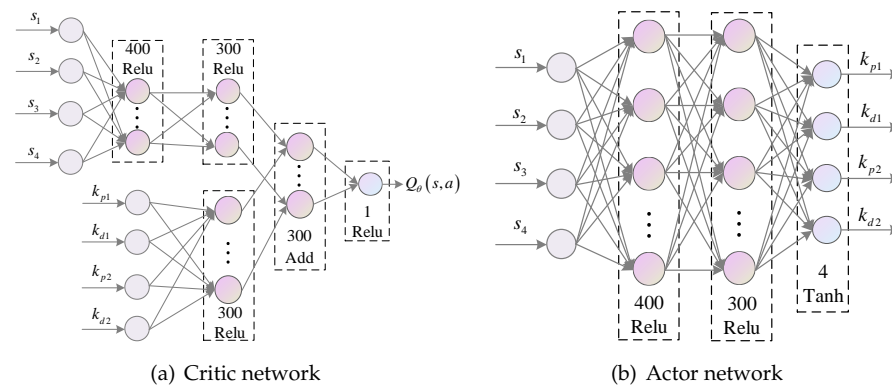


Figure 5. Structure diagram of critic and actor networks.

5. Simulation Results

Table 1 illustrates the dimensions of the powered parafoil investigated in this paper. Numerical simulations were performed in the Matlab environment and executed on a laptop with an i7-10875H CPU. In addition, the hyperparameters involved in the TD3 algorithm were $\gamma = 0.99$ and $\tau = 0.005$. The simulation in this paper was completed on the MATLAB platform. The learning rate of networks was 0.001.

To validate the efficacy of the proposed TD3-LADRC method, simulation verification is performed using the model presented in Section 2. With an initial velocity of the parachute set as $V_0 = [14.9, 0, 2.1]^T$ m/s and $W_0 = [0, 0, 0]^T$ rad/s, the initial position is specified as

$(x_0, y_0, z_0) = (20, 20, 550)$ m. The initial Euler angles are defined as $[\phi_0, \theta_0, \psi_0] = [0, 0, 0]$. In addition, suppose the following formula expresses the target trajectory:

$$\begin{cases} x_d(\omega) = 6\omega \\ y_d(\omega) = 8\omega \\ z_d = 500 \end{cases} \quad (33)$$

Table 1. Physical parameters of the powered 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

Subsequently, the TD3-LADRC control method is employed to achieve the desired trajectory, as depicted in Equation (33). In the control system, the parameter values in the guidance law are set as $\alpha = 0.02$ and $k = 50$. For the horizontal trajectory controller, ω_o and b_0 are chosen as 5 and 0.5, respectively. Likewise, for the altitude controller, the values selected are 15 and 2, respectively. Furthermore, the action space is expressed by Equation (34). Moreover, the following four cases are considered:

- Wind-free: there is no wind during the flight;
- Wind-X: a wind of 2 m/s from the west direction is added at 60 s and lasts until the end of the simulation;
- Wind-Y: a wind of 2 m/s from the south direction is added at 60 s and lasts until the end of the simulation;
- Wind-Z: a wind of 2 m/s from the vertical direction is added at 60 s and lasts until the end of the simulation.

$$\begin{cases} k_{p1} \in [0.005, 0.05] \\ k_{d1} \in [0.1, 0.5] \\ k_{p2} \in [0.05, 0.5] \\ k_{d2} \in [1, 2] \end{cases} \quad (34)$$

The simulation results are shown in Figures 6–11. First and foremost, the results of episode rewards are displayed in Figure 6, which indicate that the TD3 agent can achieve stabilization. For the cases mentioned above, Figure 7 displays the schematic diagrams of both the 3D and 2D trajectories. Additionally, Figure 8 illustrates the progression of the steering gear and thrust obtained by the controller. The corresponding trajectory tracking errors and the reward value change process are depicted in Figure 9. Furthermore, Figure 10 showcases the variations in Euler angles and velocity throughout the flight, while Figure 11 presents the TD3-optimized controller parameters.

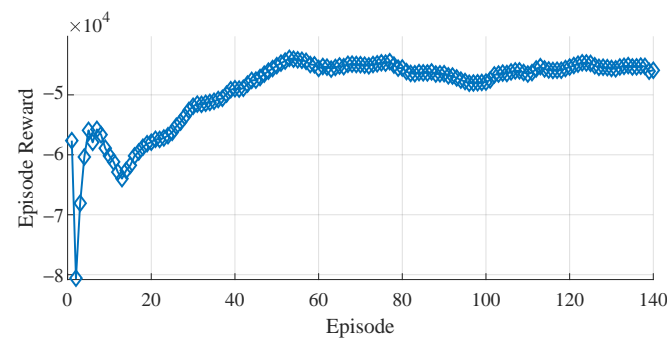


Figure 6. Episode reward for the TD3 agent.

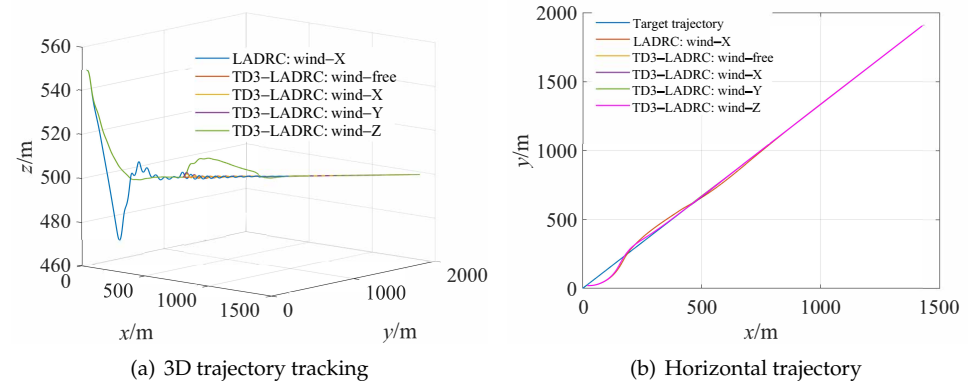


Figure 7. Trajectory performance in different cases.

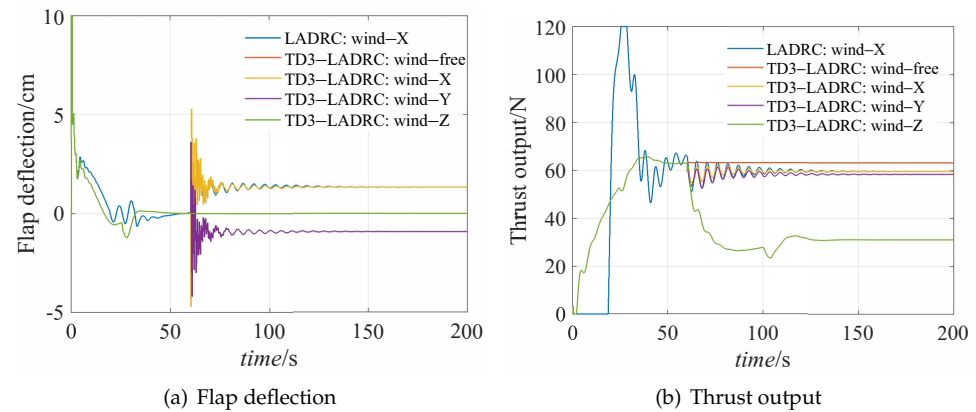


Figure 8. Control variables.

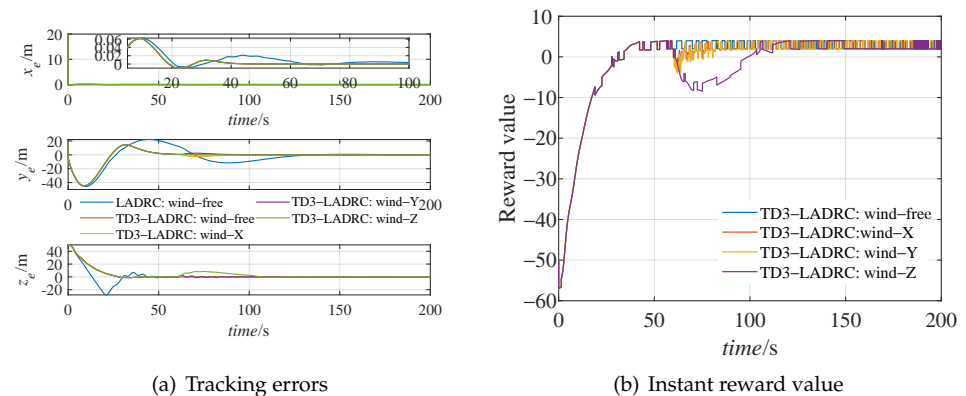


Figure 9. Trajectory tracking errors and the corresponding reward value.

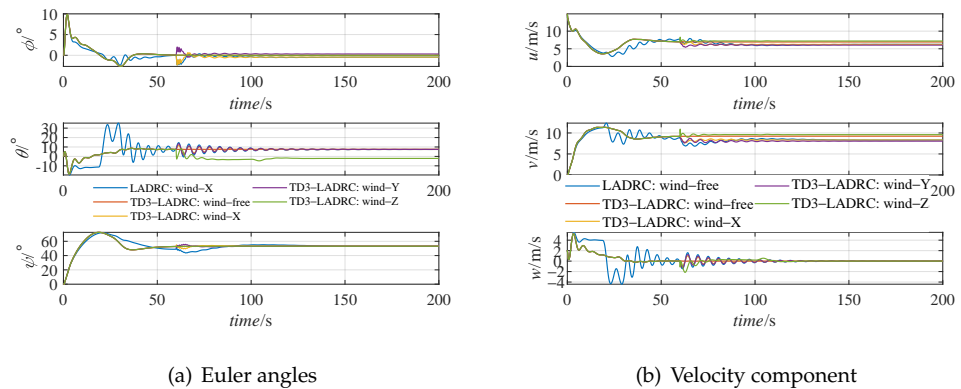


Figure 10. Euler angles and speed change process.

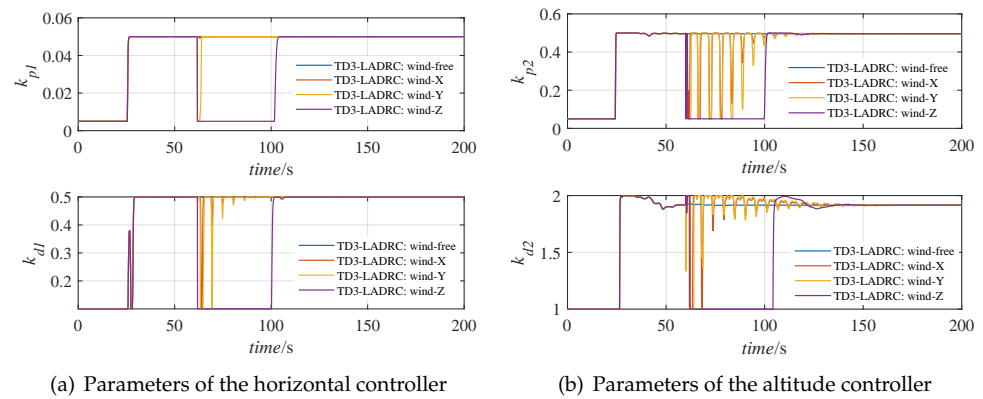


Figure 11. Controller parameters by TD3.

In general, our objective is to test the trajectory tracking effect of the proposed TD3-LADRC in various cases. Additionally, we compared the control effect of the LADRC with fixed parameters in the obtained results. For the fixed parameters, we selected $k_{p1} = 0.005$, $k_{d1} = 0.1$, $k_{p2} = 0.05$, and $k_{d2} = 1$. Overall, the method depicted in the results diagram can successfully accomplish trajectory tracking control. This, to a certain extent, demonstrates the effectiveness of the proposed guidance law and LADRC control method. Specifically, under the same condition of being disturbed by wind in the X direction, it can be found from Figure 7a that the trajectory tracking effect of the yellow curve is better than the blue curve, which shows that the proposed TD3-LADRC method is better than the traditional LADRC method, and this conclusion can also be further verified in Figure 9a. Observing the response of the proposed TD3-LADRC to different disturbances, it can be found that the thrust output and flap deflection will also fluctuate when they are perturbed, and they will eventually stabilize to a constant value, as illustrated in Figure 8. Furthermore, the proposed method can overcome the effects of disturbance and coupling to achieve trajectory tracking control. It can be seen from Figure 10 that the parafoil does not lose control of its attitude during flight and can fly smoothly. The selection of parameters in Figure 11 is determined based on the reward value in Figure 9b. As state changes result in fluctuations in reward values, the controller parameters experience more frequent adjustments in the presence of disturbances compared to disturbance-free scenarios. Furthermore, when subjected to the same disturbance, the trajectory tracking effect optimized by the TD3 algorithm proves superior, as evident from Figure 9a. By examining the change process of the steering gear, thrust output, and controller parameters, as illustrated in Figures 8 and 11, we observe that the system responds differently to various wind disturbances to achieve precise trajectory tracking. This observation further supports the trajectory tracking approach presented in this paper.

6. Conclusions

This paper investigates the trajectory tracking control problem in powered parafoil systems. To address the underactuation, coupling, and wind disturbance challenges during the control process, we propose an intelligent control method based on TD3-optimized LADRC. Initially, an eight-DOF model is established to simulate the actual parafoil system's flight state accurately. Subsequently, we address the underactuation issue by introducing a trajectory tracking guidance law. Moreover, we design two LADRC controllers to achieve decoupling between horizontal trajectory tracking and height control while enabling trajectory tracking control. To further enhance the trajectory tracking process, we employ the TD3 algorithm to obtain real-time parameters for the controller, thereby reducing tracking errors. The proposed method successfully achieves trajectory tracking control under various disturbance conditions, with its effectiveness duly validated.

The proposed method in this paper effectively achieves trajectory tracking control for powered parafoils and is also instructive for the trajectory tracking control of other aircraft. However, during the TD3 optimization process, finding a good balance between the horizontal and height tracking states within the reward function is a significant challenge that warrants further consideration. Assigning excessively high rewards to either state can undermine the feedback loop of the other state during the training process. We intend to continue investigating this issue in our future research. Additionally, in this paper, the configuration of the action space is manually adjusted based on multiple simulations, lacking theoretical solid justification. Addressing this limitation is a crucial aspect that requires dedicated research efforts. Furthermore, this paper did not delve into a theoretical analysis of the actor–critic framework. Some insights on this can be found in refs. [38–40], and we plan to carry out further research in this direction in the future.

Author Contributions: Conceptualization, Y.Z., Z.C. and M.S.; methodology, J.T.; software, Y.Z.; validation, Y.Z., Z.F. and J.T.; formal analysis, Y.Z.; investigation, Q.S.; resources, Q.S.; writing—original draft preparation, Y.Z. and J.T.; writing—review and editing, Q.S. and H.S.; visualization, Y.Z.; supervision, Q.S. and H.S.; 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, 62003177 and 62073177) and Key Technologies Research and Development Program of Tianjin (Grant No. 19JCZDJC32800).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to copyright issues with co-developers.

Conflicts of Interest: Author Jin Tao was employed by the company Silo AI. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A

This section takes the LADRC in Section 3.2.1 as an example to prove the stability of the closed-loop system. Firstly, according to the LESO in Equation (20), the estimation error vector is defined as $\varepsilon = [\varepsilon_1, \varepsilon_2, \varepsilon_3]^T$, with $\varepsilon_i = x_{1i} - \hat{x}_{1i}$. Then, the following Theorem A1 can be obtained.

Theorem A1. Assuming that $h = \dot{f}_h$ is bounded, that is, $|h| \leq M_1$, ($M_1 > 0$), then the observer error ε of the LESO is also bounded, that is, $|\varepsilon| \leq M_2$, $M_2 > 0$.

Proof of Theorem A1. With Equations (19) and (20), the state space equation about ε can be derived as:

$$\dot{\varepsilon} = A_1 \varepsilon + B h \quad (A1)$$

$$\text{where } A_1 = \begin{bmatrix} -\beta_{01} & 1 & 0 \\ -\beta_{02} & 0 & 1 \\ -\beta_{03} & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

According to Equation (A1), we have

$$\varepsilon(t) = e^{A_1 t} \varepsilon(0) + \int_0^t e^{A_1(t-\tau)} B h d\tau \quad (\text{A2})$$

Since A_1 is a Hurwitz matrix, there exists an invertible real matrix T that allows the following expression of A_1 :

$$A_1 = T \text{diag}\{-\lambda_1, -\lambda_2, -\lambda_3\} T^{-1} \quad (\text{A3})$$

where $-\lambda_i (\lambda_i > 0), i = 1, 2, 3$, represents the eigenvalues.

Then, we have

$$e^{A_1 t} = T \text{diag}\{e^{-\lambda_1}, e^{-\lambda_2}, e^{-\lambda_3}\} T^{-1} \quad (\text{A4})$$

Further, the following inequality can be derived by m_∞ -norm:

$$\|e^{A_1 t}\|_{m_\infty} \leq \beta e^{-\lambda_1 t} \quad (\text{A5})$$

where β is a constant value.

According to Equation (A5), with the assumption of $|h| \leq M_1$, Equation (A2) can have the following inequality derivation process:

$$\begin{aligned} \|\varepsilon(t)\| &= \left\| e^{A_1 t} \varepsilon(0) + \int_0^t e^{A_1(t-\tau)} B h d\tau \right\| \\ &\leq \|e^{A_1 t} \varepsilon(0)\| + \left\| \int_0^t e^{A_1(t-\tau)} B h d\tau \right\| \\ &\leq \|e^{A_1 t}\|_{m_\infty} \|\varepsilon(0)\| + \int_0^t \|e^{A_1(t-\tau)}\| \|B\| \|h\| d\tau \\ &\leq \beta e^{-\lambda_1 t} \|\varepsilon(0)\| + \frac{M_1 \beta}{\lambda_1} (1 - e^{-\lambda_1 t}) \\ &\leq \beta \|\varepsilon(0)\| + \frac{M_1 \beta}{\lambda_1} \\ &= M_2 \end{aligned} \quad (\text{A6})$$

□

Theorem A2. Assume that the control system tracks a bounded input r and the observer error satisfies

$$\lim_{t \rightarrow \infty} \|\varepsilon(t)\| = 0 \quad (\text{A7})$$

then the tracking errors e can also converge to 0.

Proof of Theorem A2. Let $r_1 = r, r_2 = \dot{r}, r_3 = f'_h$, then the tracking error vector of the controller can be defined as $\xi = [\xi_1, \xi_2]^T$, with $\xi_i = r_i - x_{1i}$. In addition, the PD control law in Equation (21) can have the normal form,

$$u_1 = \frac{k_{p1}(r_1 - \hat{x}_{11}) + k_{d1}(r_2 - \hat{x}_{12}) + (r_3 - \hat{x}_{13})}{b_{01}} \quad (\text{A8})$$

Then, according to Equations (19) and (A8), the following expression can be derived:

$$\begin{cases} \dot{\xi}_1 = \xi_2 \\ \dot{\xi}_2 = -k_{p1}(\xi_1 + \varepsilon_1) - k_{d1}(\xi_2 + \varepsilon_2) - \xi_3 \end{cases} \quad (\text{A9})$$

Furthermore, Equation (A9) can be expressed in another form:

$$\dot{\xi} = A_2 \xi + A_3 \varepsilon \quad (\text{A10})$$

with $A_2 = \begin{bmatrix} 0 & 1 \\ -k_{p1} & -k_{d1} \end{bmatrix}$, and $A_3 = \begin{bmatrix} 0 & 0 & 0 \\ -k_{p1} & -k_{d1} & -1 \end{bmatrix}$.
Similarly, Equation (A10) has the following solution,

$$\xi(t) = e^{A_2 t} \xi(0) + \int_0^t A_3 \varepsilon e^{A_2(t-\tau)} d\tau \quad (\text{A11})$$

A_2 is also a Hurwitz matrix, thus satisfying $\|e^{A_2 t}\|_{m_\infty} \leq \beta e^{-\lambda_1 t}$. As a result, there is the following inequality about $\xi(t)$,

$$\begin{aligned} \|\xi(t)\| &= \left\| \xi^{A_2 t} \xi(0) + \int_0^t A_3 \varepsilon e^{A_2(t-\tau)} d\tau \right\| \\ &\leq \left\| \xi^{A_2 t} \xi(0) \right\| + \left\| \int_0^t A_3 \varepsilon e^{A_2(t-\tau)} d\tau \right\| \\ &\leq \beta e^{-\lambda_1 t} \|\xi(0)\| + \|\varepsilon\| \|A_3\| \beta \int_0^t e^{-\lambda_1(t-\tau)} d\tau \\ &\leq \beta e^{-\lambda_1 t} \|\xi(0)\| + \frac{M_2 \|A_3\| \beta}{\lambda_1} (1 - e^{-\lambda_1 t}) \end{aligned} \quad (\text{A12})$$

In accordance with the assumption presented in Equation (A7), finding the limit of ξ in Equation (A12) leads to the following result,

$$\lim_{t \rightarrow \infty} \|\xi(t)\| = 0 \quad (\text{A13})$$

□

Remark A1. Theorem A1 proves the stability of the linear extended state observer under the premise that the disturbance is bounded, which shows that the observation errors are bounded.

Remark A2. Theorem A2 further establishes that with the utilization of the expanded state observer and PD control law, the tracking error of the closed-loop system is bounded, thus ensuring the stability of the closed-loop system.

References

- Li, B.; He, Y.; Han, J.; Xiao, J. A new modeling scheme for powered parafoil unmanned aerial vehicle platforms: Theory and experiments. *Chin. J. Aeronaut.* **2019**, *32*, 2466–2479. [\[CrossRef\]](#)
- Murali, N.; Dineshkumar, M.; WC, A.K. Parafoil trajectory comparison for optimal control and proportional controller. In Proceedings of the 2013 International Conference on Control Communication and Computing (ICCC), Thiruvananthapuram, India, 13–15 December 2013; pp. 227–232.
- Yang, H.; Song, L.; Wang, W.; Huang, J. 4-DOF longitudinal dynamic simulation of powered-parafoil. *J. Beijing Univ. Aeronaut. Astronaut.* **2014**, *40*, 1615–1622.
- 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.; Kang, X.; He, Y. Modeling of powered parafoil based on Kirchhoff motion equation. *Nonlinear Dyn.* **2014**, *79*, 617–629. [\[CrossRef\]](#)
- 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. [\[CrossRef\]](#)
- Hur, G.; Valasek, J. System identification of powered parafoil-vehicle from flight test data. In Proceedings of the AIAA Atmospheric Flight Mechanics Conference and Exhibit, Austin, TX, USA, 11–14 August 2003; p. 5539.
- Feng, L.; Xing, X.; Gong, Q.; Li, Y.; Guo, Y. Trajectory Control of PFC Recovery under ADRC and Improved LOS Guidance Law. In *International Conference on Guidance, Navigation and Control*; Springer Nature: Singapore, 2022; pp. 5498–5507.
- Guo, Y.; Yan, J.; Wu, C.; Chen, M.; Xing, X. Autonomous homing design and following for parafoil/rocket system with high-altitude. *J. Intell. Robot. Syst.* **2021**, *101*, 73. [\[CrossRef\]](#)

10. Viswa, S.; Avijit, B.; Sumeet, S.; Roy, S.; Nikolakopoulos, G. Adaptive control for a payload carrying spacecraft with state constraints. *Control Eng. Pract.* **2023**, *135*, 105515.
11. 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. [\[CrossRef\]](#)
12. Li, Z.; Nan, Y. Optimal Path Planning and Tracking Control Methods for Parafoil. *Appl. Sci.* **2023**, *13*, 8115. [\[CrossRef\]](#)
13. Guo, Y.; Xing, X.; Wu, X.; Wu, C.; Xiao, B. Adaptive path-following control for parafoil dynamic systems with wind disturbance and rate constraint. *Nonlinear Dyn.* **2023**, *111*, 13039–13051. [\[CrossRef\]](#)
14. 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* **2022**, *15*, 435. [\[CrossRef\]](#)
15. 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. [\[CrossRef\]](#)
16. Zheng, Y.; Tao, J.; Sun, Q.; Sun, H.; Chen, Z.; Sun, M.; Xie, G. Sideslip angle estimation based active disturbance rejection 3D trajectory tracking control for powered parafoil system and hardware-in-the-loop simulation verification. *Aerosp. Sci. Technol.* **2023**, *141*, 108497. [\[CrossRef\]](#)
17. Han, J. From PID to active disturbance rejection control. *IEEE Trans. Ind. Electron.* **2019**, *56*, 900–906. [\[CrossRef\]](#)
18. Gao, Z. On the foundation of active disturbance rejection control. *Control Theory Appl.* **2013**, *30*, 1498–1510.
19. Xia, G.; Chu, H.; Shao, Y.; Xia, B. DSC and LADRC Path Following Control for Dynamic Positioning Ships at High Speed. In Proceedings of the 2019 IEEE International Conference on Mechatronics and Automation (ICMA), Tianjin, China, 4–7 August 2019; pp. 39–44.
20. Kang, N.; Han, Y.; Guan, T.; Wang, S. Improved ADRC-Based Autonomous Vehicle Path-Tracking Control Study Considering Lateral Stability. *Appl. Sci.* **2022**, *12*, 4660. [\[CrossRef\]](#)
21. Li, H.; An, X.; Feng, R.; Chen, Y. Motion Control of Autonomous Underwater Helicopter Based on Linear Active Disturbance Rejection Control with Tracking Differentiator. *Appl. Sci.* **2023**, *13*, 3836. [\[CrossRef\]](#)
22. Marini, F.; Walczak, B. Particle swarm optimization (PSO). A tutorial. *Chemom. Intell. Lab. Syst.* **2015**, *149*, 153–165. [\[CrossRef\]](#)
23. Katoch, S.; Chauhan, S.; Kumar, V. A review on genetic algorithm: Past, present, and future. *Multimed. Tools Appl.* **2021**, *80*, 8091–8126. [\[CrossRef\]](#)
24. Sun, X.; Xiong, Y.; Yao, M.; Tang, X.; Tian, X. A unified control method combined with improved TSF and LADRC for SRMs using modified grey wolf optimization algorithm. *ISA Trans.* **2022**, *131*, 662–671. [\[CrossRef\]](#)
25. Abdul-Kareem, A.I.; Hasan, A.F.; Al-Qassar, A.A.; Humaidi, A.J.; Hassan, R.F.; Ibraheem, I.K.; Azar, A.T. Rejection of wing-rock motion in delta wing aircrafts based on optimal LADRC schemes with butterfly optimization algorithm. *J. Eng. Sci. Technol.* **2022**, *17*, 2476–2495.
26. Sun, C.; Liu, C.; Feng, X.; Jiao, X. Visual servoing of flying robot based on fuzzy adaptive linear active disturbance rejection control. *IEEE Trans. Circuits Syst. II Express Briefs* **2021**, *68*, 2558–2562. [\[CrossRef\]](#)
27. Liu, W.; Zhao, T.; Wu, Z.; Huang, W. Linear active disturbance rejection control for hysteresis compensation based on backpropagation neural networks adaptive control. *Trans. Inst. Meas. Control* **2021**, *43*, 915–924. [\[CrossRef\]](#)
28. Arulkumaran, K.; Deisenroth, M.; Brundage, M.; Bharath, A.A. Deep reinforcement learning: A brief survey. *IEEE Signal Process. Mag.* **2017**, *34*, 26–38. [\[CrossRef\]](#)
29. Osband, I.; Blundell, C.; Pritzel, A.; Van Roy, B. Deep exploration via bootstrapped DQN. In Proceedings of the Advances in Neural Information Processing Systems, Barcelona, Spain, 5–10 December 2016; p. 29.
30. Li, S.; Wu, Y.; Cui, X.; Dong, H.; Fang, F.; Russell, S. Robust multi-agent reinforcement learning via minimax deep deterministic policy gradient. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 27 January–1 February 2019; Volume 33, pp. 4213–4220. [\[CrossRef\]](#)
31. Dankwa, S.; Zheng, W. Twin-delayed ddpg: A deep reinforcement learning technique to model a continuous movement of an intelligent robot agent. In Proceedings of the 3rd International Conference on Vision, Image and Signal Processing, Vancouver, BC, Canada, 26–28 August 2019; pp. 1–5.
32. Zhang, H.; Yin, C.; Zhang, Y. Motion planning using reinforcement learning method for underactuated ship berthing. In Proceedings of the 2020 IEEE 16th International Conference on Control & Automation (ICCA), Sapporo, Japan, 9–11 October 2020; pp. 354–359.
33. Wang, Y.; Gao, Z.; Zhang, J.; Cao, X.; Zheng, D.; Gao, Y.; Ng, D.W.; Di Renzo, M. Trajectory design for UAV-based Internet of Things data collection: A deep reinforcement learning approach. *IEEE Internet Things J.* **2021**, *9*, 3899–3912. [\[CrossRef\]](#)
34. Chu, Z.; Sun, B.; Zhu, D.; Zhang, M.; Luo, C. Motion control of unmanned underwater vehicles via deep imitation reinforcement learning algorithm. *IET Intell. Transp. Syst.* **2020**, *14*, 764–774. [\[CrossRef\]](#)
35. Zhu, H.; Sun, Q.; Tao, J.; Chen, Z.; Dehmer, M.; Xie, G. Flexible modeling of parafoil delivery system in wind environments. *Commun. Nonlinear Sci. Numer. Simul.* **2022**, *108*, 106210. [\[CrossRef\]](#)
36. Jin, H.; Song, J.; Lan, W.; Gao, Z. On the characteristics of ADRC: A PID interpretation. *Sci. China Inf. Sci.* **2020**, *63*, 1–3. [\[CrossRef\]](#)
37. Qin, H.; Tan, P.; Chen, Z.; Sun, M.; Sun, Q. Deep reinforcement learning based active disturbance rejection control for ship course control. *Neurocomputing* **2022**, *484*, 99–108. [\[CrossRef\]](#)
38. Bu, X.; Xiao, Y.; Lei, H. An adaptive critic design-based fuzzy neural controller for hypersonic vehicles: Predefined behavioral nonaffine control. *IEEE/ASME Trans. Mechatronics* **2019**, *24*, 1871–1881. [\[CrossRef\]](#)

39. Bu, X.; Qi, Q. Fuzzy optimal tracking control of hypersonic flight vehicles via single-network adaptive critic design. *IEEE Trans. Fuzzy Syst.* **2020**, *30*, 270–278. [[CrossRef](#)]
40. Qi, Q.; Bu, X. Adaptive dynamic programming design for the neural control of hypersonic flight vehicles. *J. Frankl. Inst.* **2021**, *358*, 8169–8192. [[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.