Application of Deep Reinforcement Learning in Reconfiguration Control of Aircraft Anti-Skid Braking System

Shuchang Liu [1], Zhong Yang [1], Zhao Zhang [2],*, Runqiang Jiang [2], Tongyang Ren [2], Yuan Jiang [2], Shuang Chen [3] and Xiaokai Zhang [3]

1 College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
2 Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China
3 Electronic Engineering Department, Aviation Key Laboratory of Science and Technology on Aero Electromechanical System Integration, Nanjing 211106, China
* Correspondence: zhangzhao@ciomp.ac.cn

Abstract: The aircraft anti-skid braking system (AABS) plays an important role in aircraft taking off, taxiing, and safe landing. In addition to the disturbances from the complex runway environment, potential component faults, such as actuators faults, can also reduce the safety and reliability of AABS. To meet the increasing performance requirements of AABS under fault and disturbance conditions, a novel reconfiguration controller based on linear active disturbance rejection control combined with deep reinforcement learning was proposed in this paper. The proposed controller treated component faults, external perturbations, and measurement noise as the total disturbances. The twin delayed deep deterministic policy gradient algorithm (TD3) was introduced to realize the parameter self-adjustments of both the extended state observer and the state error feedback law. The action space, state space, reward function, and network structure for the algorithm training were properly designed, so that the total disturbances could be estimated and compensated for more accurately. The simulation results validated the environmental adaptability and robustness of the proposed reconfiguration controller.

Keywords: aircraft anti-skid braking system; actuator faults; reconfiguration control; linear active-disturbance rejection control; deep reinforcement learning; twin delayed deep deterministic policy gradient algorithm

1. Introduction

The aircraft anti-skid braking system (AABS) is an essential airborne utilities system to ensure the safe and smooth landing of aircraft [1]. With the development of aircraft towards high speed and large tonnage, the performance requirements of AABS are increasing. Moreover, AABS is a complex system with strong nonlinearity, strong coupling, and time-varying parameters, and is sensitive to the runway environment [2]. These characteristics make AABS controller design an interesting and challenging topic.

The most widely used control method in practice is PID + PBM, which is a speed differential control law. However, it suffers from low-speed slipping and underutilization of ground bonding forces, making it difficult to meet high performance requirements. To this end, researchers have proposed many advanced control methods to improve the AABS performance, such as mixed slip deceleration PID control [3], model predictive control [4], extremum-seeking control [5], sliding mode control [6], reinforcement Q-learning control [7], and so on. Zhang et al. [8] proposed a feedback linearization controller with a prescribed performance function to ensure the transient and steady-state braking performance. Qiu et al. [9] combined backstepping dynamic surface control with an asymmetric barrier Lyapunov function to obtain a robust tracking response in the presence of disturbance

and runway surface transitions. Mirzaei et al. [10] developed a fuzzy braking controller optimized by a genetic algorithm and introduced an error-based global optimization approach for fast convergence near the optimum point. The above-mentioned works provide an in-depth study on AABS control; however, the adverse effects caused by typical component faults such as actuator faults are neglected. Since most AABS are designed based on hydraulic control systems, the long hydraulic pipes create an enormous risk of air mixing with oil, and internal leakage. Without regular maintenance, it is easy to cause functional degradation or even failure, which raises many security concerns [11,12]. How to ensure the stability and the acceptable braking performance of AABS after actuator faults becomes a key issue.

In order to actually improve the safety and reliability of AABS, the fault probability can be reduced by reliability design and redundant technology on the one hand [13]. However, due to the production factors (cost/weight/technological level), the redundancy of aircraft components is so limited that the system reliability is hard to increase. On the other hand, fault-tolerant control (FTC) technology can be introduced into the AABS controller design, which is the future development direction of AABS and the key technology that needs urgent attention [14]. Reconfiguration control is a popular branch of FTC that has been widely used in many safety-critical systems, especially in aerospace engineering [15,16]. The essence of reconfiguration control is to consider the possible faults of the plant in the controller design process. When component faults occur, the fault system information is used to reconfigure the controller structure or parameters automatically [17]. In this way, the adverse effects caused by faults can be restrained or eliminated, thus realizing an asymptotically stable and acceptable performance of the closed-loop system. A number of common reconfiguration control methods can be classified as follows: adaptive control [18,19], multi-model switching control [20], sliding mode control [21], fuzzy control [22], other robust control [23], etc. In addition, the characteristics of AABS increase the difficulty of accurate modeling, and many nonlinear reconfiguration control methods are complex and relatively hard to apply in engineering. Therefore, it is crucial to design a reconfiguration controller with a clear structure, and which is model-independent, strong fault-perturbation resistant, and easy to implement.

Han retained the essence of PID control and proposed an active disturbance rejection control (ADRC) technique that requires low model accuracy and shows good control performance [24]. ADRC can estimate disturbances in internal and external systems and compensate for them [25]. Furthermore, ADRC has been widely used in FTC system design because of its obvious advantages in solving control problems of nonlinear models with uncertainty and strong disturbances [26–28]. Although the structure is not difficult to implement with modern digital computer technology, ADRC needs to tune a bunch of parameters which makes it hard to use in practice [29]. To overcome the difficulty, Gao proposed linear active disturbance rejection control (LADRC), which is based on linear extended state observer (LESO) and linear state error feedback (LSEF) [30,31]. The bandwidth tuning method greatly reduced the number of LADRC parameters. LADRC has been applied to solve various control problems [32–34].

However, it is well known that a controller with fixed parameters may not be able to maintain the acceptable (rated or degraded) performance of a fault system. For this reason, some advanced algorithms with parameter adaptive capabilities have been introduced by researchers that further improve the robustness and environmental adaptability of ADRC, such as neural networks [35,36], fuzzy logic [37,38], and the sliding mode [39,40]. With the development of artificial intelligence techniques, reinforcement learning has been applied to control science and engineering [41,42], and good results have been achieved. Yuan et al. proposed a novel online control algorithm for a thickener which is based on reinforcement learning [43]. Pang et al. studied the infinite-horizon adaptive optimal control of continuous-time linear periodic systems, using reinforcement learning techniques [44]. A Q-learning-based adaptive method for ADRC parameters was proposed by Chen et al. and has been applied to the ship course control [45].

Motivated by the above observations, in this paper, a reconfiguration control scheme via LADRC combined with deep reinforcement learning was developed for AABS which is subject to various fault perturbations. The proposed reconfiguration control method is a remarkable control strategy compared to previous methods for three reasons:

(1) AABS is extended with a new state variable, which is the sum of all unknown dynamics and disturbances not noticed in the fault-free system description. This state variable can be estimated using LESO. It indirectly simplifies the AABS modeling;

(2) Artificial intelligence technology is introduced and combined with the traditional control method to solve special control problems. By combining LADRC with the deep reinforcement learning TD3 algorithm, the selection of controller parameters is equivalent to the choice of agent actions. The parameter adaptive capabilities of LESO and LSEF are endowed through the continuous interaction between the agent and the environment, which not only eliminates the tedious manual tuning of the parameters, but also results in more accurate estimation and compensation for the adverse effects of fault perturbations;

(3) It is a data-driven robust control strategy that does not require any additional fault detection or identification (FDI) module, while the controller parameters are adaptive. Therefore, the proposed method corresponds to a novel combination of active reconfiguration control and FDI-free reconfiguration control, which makes it an interesting solution under unknown fault conditions.

The paper is organized as follows. Section 2 describes AABS dynamics with an actuator fault factor. The reconfiguration controller is presented in Section 3. The simulation results are presented to demonstrate the merits of the proposed method in Section 4, and conclusions are drawn in Section 5.

2. AABS Modeling

The AABS mainly consists of the following components: aircraft fuselage, landing gear, wheels, a hydraulic servo system, a braking device, and an anti-skid braking controller. The subsystems are strongly coupled and exhibit strong nonlinearity and complexity.

Based on the actual process and objective facts of anti-skid braking, the following reasonable assumptions can be made [46]:

(1)　The aircraft fuselage is regarded as a rigid body with concentrated mass;
(2)　The gyroscopic moment generated by the engine rotor is not considered during the aircraft braking process;
(3)　The crosswind effect is ignored;
(4)　Only the longitudinal deformation of the tire is taken into account and the deformation of the ground is ignored;
(5)　All wheels are the same and controlled synchronously.

2.1. Aircraft Fuselage Dynamics

The force diagram of the aircraft fuselage is shown in Figure 1 and the specific parameters described in the diagram are shown in Table 1.
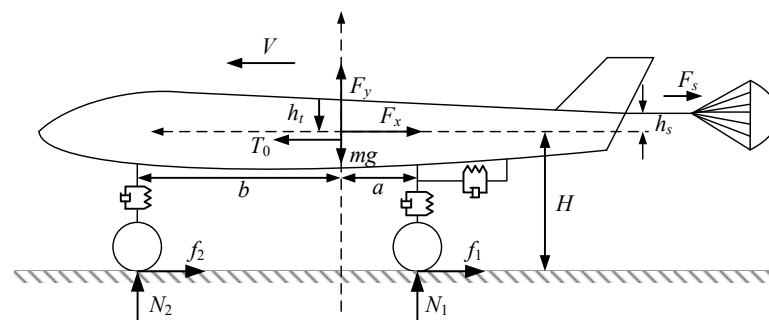
Figure 1. Force diagram of aircraft fuselage.

Table 1. Parameters of aircraft fuselage dynamics.

Name	Description	Value
H	Center of gravity height	
y	Center of gravity height variation	
V	Aircraft speed	
T_0	Engine force	
F_x	Aerodynamic drag	
F_y	Aerodynamic lift	
F_s	Parachute drag	
f_1	Braking friction force between main wheel and ground	
f_2	Braking friction force between front wheel and ground	
N_1	Main wheel support force	
N_2	Front wheel support force	
m	Mass of the aircraft	1761 kg
g	Gravitational acceleration	9.8 m/s^2
h_t	Distance between engine force line and center of gravity	0.1 m
h_s	Distance between parachute drag line and center of gravity	0.67 m
a	Distance between main wheel and center of gravity	1.076 m
b	Distance between front wheel and center of gravity	6.727 m
I	Fuselage inertia	4000 kg·s^2·m
S	Wing aera	50.88 m^2
S_s	Parachute area	20 m^2
C_x	Aerodynamic drag coefficient	0.1027
C_y	Aerodynamic lift coefficient	0.6
C_{xs}	Parachute drag coefficient	0.75
T_0'	Intimal engine force	426 kg
K_v	Velocity coefficient of engine	1 kg·s/m
ρ	Air density	4000 kg·s^2/m^4

The aircraft force and torque equilibrium equations are:

$$\begin{cases} m\dot{V} + F_x + F_s + f_1 + f_2 - T_0 = 0 \\ F_y + N_1 + N_2 - mg = 0 \\ N_2 b + F_s h_s - N_1 a - T_0 h_t - f_1 H - f_2 H = 0 \end{cases} \tag{1}$$

According to the influence of aerodynamic characteristics, we can obtain [46]:

$$\begin{cases} T_0 = T_0' + K_v V \\ F_x = \dfrac{1}{2}\rho C_x S V^2 \\ F_y = \dfrac{1}{2}\rho C_y S V^2 \\ F_s = \dfrac{1}{2}\rho C_{xs} S_s V^2 \\ f_1 = \mu_1 N_1 \\ f_2 = \mu_2 N_2 \end{cases} \tag{2}$$

2.2. Landing Gear Dynamics

The main function of the landing gear is to support and buffer the aircraft, thus improving the longitudinal and vertical forces. In addition to the wheel and braking device, the struts, buffers, and torque arm are also the main components of the landing gear. In this paper, it is assumed that the stiffness of the torque arm is large enough, and the torsional freedom of the wheel with respect to the strut and the buffer is ignored, so the torque arm is not considered.

The buffer can be reasonably simplified as a mass-spring-damping system [46], and the force acting on the aircraft fuselage by the buffer can be described as:

$$\begin{cases} N_1 = K_1 X_1 + C_1 \dot{X}_1^2 \\ N_2 = K_2 X_2 + C_2 \dot{X}_2^2 \end{cases} \tag{3}$$

$$\begin{cases} X_1 = a + y \\ X_2 = -b + y \end{cases} \tag{4}$$

whose parameters are shown in Table 2.

Table 2. Parameters of the buffer.

Name	Description	Value
X_1	Main buffer compression	
X_2	Front buffer compression	
K_1	Main buffer stiffness coefficient	42,529
K_2	Front buffer stiffness coefficient	2500
C_1	Main buffer damping coefficient	800
C_2	Front buffer damping coefficient	800

Due to the non-rigid connection between the landing gear and the aircraft fuselage, horizontal and angular displacements are generated under the action of braking forces. However, the struts are cantilever beams, and their angular displacements are very small and negligible. Therefore, the lateral stiffness model can be expressed by the following equivalent second-order equation:

$$\begin{cases} d_a = \dfrac{-\dfrac{f_1}{K_0}}{\dfrac{1}{W_n^2}s^2 + \dfrac{2\xi}{W_n}s + 1} \\ d_V = \dfrac{d}{dt}(d_a) \end{cases} \tag{5}$$

whose parameters are shown in Table 3.

Table 3. Parameters of the landing gear lateral stiffness model.

Name	Description	Value
d_a	Navigation vibration displacement	Please see Equation (5)
d_V	Navigation vibration speed	Please see Equation (5)
K_0	Dynamic stiffness coefficient	536,000
ξ	Dynamic stiffness coefficient	0.2
W_n	Equivalent model natural frequency	60 Hz

2.3. Wheel Dynamics

The force diagram of the main wheel brake is shown in Figure 2.
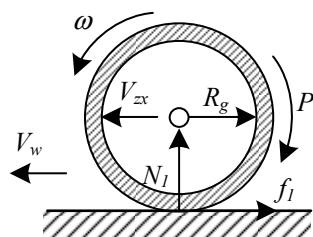
Figure 2. Force diagram of the main wheel.

It can be seen that during the taxiing, the main wheel is subjected to a combined effect of the braking torque M_s and the ground friction torque M_j. Due to the effect of the lateral stiffness, there is a longitudinal axle velocity V_{zx} along the fuselage, which is superimposed by the aircraft velocity V and the navigation vibration velocity d_V. The dynamics equation of the main wheel is [46]:

$$\begin{cases} \dot{\omega} = \dfrac{M_j - M_s}{J} + \dfrac{V_{zx}}{R_g} \\ V_w = \omega R_g \\ V_{zx} = V + d_V \\ R_g = R - N k_\sigma \\ M_j = \mu N R_g n \end{cases} \tag{6}$$

whose parameters are shown in Table 4.

Table 4. Parameters of the main wheel.

Name	Description	Value
ω	Main wheel angular velocity	
$\dot{\omega}$	Main wheel angular acceleration	
V_w	Main wheel line speed	
R_g	Main wheel rolling radius	
N	Radical load	
J	Main wheel inertia	1.855 kg·s^2·m
R	Wheel free radius	0.4 m
k_σ	Tire compression coefficient	1.07×10^{-5} m/kg
n	Equivalent model natural frequency	4

During the braking, the tires are subjected to the braking torque that keeps the aircraft speed always greater than the wheel speed, that is $V > V_w$. Thus, the slip ratio λ is defined to represent the slip motion ratio of the wheels relative to the runway. For the main wheel, using V_{zx} instead of V to calculate λ can avoid false brake release due to landing gear deformation, thus effectively reducing the landing gear walk situation [46]. The following equation is used to calculate the slip rate in this paper:

$$\lambda = \frac{V_{zx} - V_w}{V_{zx}} \tag{7}$$

The tire–runway combination coefficient is related to many factors, including real-time runway conditions, aircraft speed, slip rate, and so on. A simple empirical formula called 'magic formula' developed by Pacejka [47] is widely used to calculate and can be expressed as follows:

$$\mu(\lambda, \tau_j) = \tau_1 \sin(\tau_2 \text{arctg}(\tau_3 \lambda)) \tag{8}$$

where $\tau_j (j = 1, 2, 3)$, τ_1, τ_2, τ_3 are peak factor, stiffness factor, and curve shape factor, respectively. Table 5 lists the specific parameters for several different runway statuses [48].

Table 5. Parameters of the runway status.

Runway Status	τ_1	τ_2	τ_3
Dry runway	0.85	1.5344	14.5
Wet runway	0.40	2.0	8.2
Snow runway	0.28	2.0875	10

2.4. Hydraulic Servo System and Braking Device Modeling

Due to the complex structure of the hydraulic servo system, in this paper, some simplifications have been made so that only electro-hydraulic servo valves and pipes are considered. Their transfer functions are given as follows:

$$
\begin{cases}
M(s) = \dfrac{K_{sv}}{\dfrac{s^2}{\omega_{sv}^2} + \dfrac{2\zeta_{sv}s}{\omega_{sv}} + 1} \\
L(s) = \dfrac{K_p}{T_p s + 1}
\end{cases}
\tag{9}
$$

whose parameters are shown in Table 6.

It should be noted that the anti-skid braking controller should realize both braking control and anti-skid control. To this end, there is an approximately linear relationship between the brake pressure P and the control current I_c, which can be described as follows:

$$
P = -I_c M(s) L(s) + P_0 \tag{10}
$$

where $P_0 = 1 \times 10^7$ Pa.

The braking device serves to convert the brake pressure into brake torque, which is calculated as follows:

$$
M_s = \mu_{mc} N_{mc} P R_{mc} \tag{11}
$$

whose parameters are shown in Table 6.

Table 6. Parameters of the hydraulic servo system.

Name	Description	Value
K_{sv}	Servo valve gain	1
ω_{sv}	Servo valve natural frequency	17.7074 rad/s
ζ_{sv}	Servo valve damping ratio	0.36
K_p	Main wheel rolling radius	1
T_p	Pipe gain	0.01
μ_{mc}	Friction coefficient of brake material	0.23
N_{mc}	Number of friction surfaces	4
R_{mc}	Effective brake friction radius	0.142 m

The hydraulic servo system, as the actuator of AABS, is inevitably subject to some potential faults. Problems such as hydraulic oil mixing with air, internal leakage, and vibration seriously affect the efficiency of the hydraulic servo system [49]. Therefore, in this paper, the loss of efficiency (LOE) is introduced to represent a typical AABS actuator fault, which is characterized by a decrease in the actuator gain from its nominal value [26]. In the case of an actuator LOE fault, the brake pressure generated by the hydraulic servo system deviates from the commanded output expected by the controller. In other words, one instead has:

$$
P_{fault} = k_{LOE} P \tag{12}
$$

where P_{fault} represents the actuator actual output, and $k_{LOE} \in (0, 1]$ refers the LOE fault factor.

Remark 1. *$n\%$ LOE is equivalent to the LOE fault gain $k_{LOE} = 1 - n/100$, $k_{LOE} = 1$ indicates that the actuator is fault-free.*

Remark 2. *Note that if the components do not always have the same characteristics as those of fault-free, it is necessary to establish the fault model. This not only provides an accurate model for the next reconfiguration on controller design, but also ensures that the adverse effects caused by fault perturbation can be effectively observed and compensated for.*

Thus, Equation (11) can be rewritten as follows:

$$M_s' = \mu_{mc} N_{mc} P_{fault} R_{mc} \tag{13}$$

where M_s' is the actual brake torque.

Remark 3. *As can be seen from the entire modeling process described above, AABS is nonlinear and highly coupled. The actuator fault leads to a sudden jump in the model parameters with greater internal perturbation compared to the fault-free case. Meanwhile, external disturbances such as the runway environment cannot be ignored.*

3. Reconfiguration Controller Design

3.1. Problem Description

Despite the aircraft having three degrees of freedom, only longitudinal taxiing is focused on in AABS. In this paper, AABS adopted the slip speed control type [48], that is, the braked wheel speed V_ω was used as the reference input, and the aircraft speed V was dynamically adjusted by the AABS controller to achieve anti-skid braking. According to Section 2, the AABS longitudinal dynamics model can be rewritten as follows:

$$\ddot{V} = f(V, \dot{V}, \varpi_{out}, \varpi_f) + b_v u \tag{14}$$

where $f(\cdot)$ is the controlled plant dynamics, ϖ_{out} represents the external disturbance, ϖ_f is an uncertain term including component faults, b_v is the control gain, and u is the system input.

Let $x_1 = V$, $x_2 = \dot{V}$. Set $f(V, \dot{V}, \varpi_{out}, \varpi_f)$ as the system generalized total perturbation and extend it to a new system state variable, i.e., $x_3 = f(V, \dot{V}, \varpi_{out}, \varpi_f)$. Then the state equation of System (14) can be obtained:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_3 + b_v u \\ \dot{x}_3 = h(V, \dot{V}, \varpi_{out}, \varpi_f) \end{cases} \tag{15}$$

where x_1, x_2, x_3 are system state variables, and $h(V, \dot{V}, \varpi_{out}, \varpi_f) = \dot{f}(V, \dot{V}, \varpi_{out}, \varpi_f)$.

Assumption 1. *Both the system generalized total perturbation $f(V, \dot{V}, \varpi_{out}, \varpi_f)$ and its differential $h(V, \dot{V}, \varpi_{out}, \varpi_f)$ are bounded, i.e.,* $\begin{cases} \left| f(V, \dot{V}, \varpi_{out}, \varpi_f) \right| \leq \sigma_1 \\ \left| h(V, \dot{V}, \varpi_{out}, \varpi_f) \right| \leq \sigma_2 \end{cases}$ *, where σ_1, σ_2 are two positive numbers.*

For System (14), affected by the total perturbation, a LADRC reconfiguration controller was designed next to restrain or eliminate the adverse effects, thus realizing the asymptotic stability and acceptable performance of the closed-loop system.

3.2. LADRC Controller Design
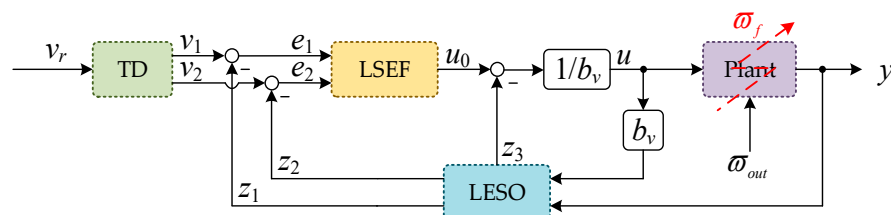
The control schematic of the LADRC is shown in Figure 3.

Figure 3. Control schematic of LADRC.

Firstly, the following tracking differentiator (TD) was designed:

$$
\begin{cases}
e(k) = v_1(k) - v_r(k) \\
\mathrm{fh} = \mathrm{fhan}(e(k), v_2(k), r, h) \\
v_1(k+1) = v_1(k) + hv_2(k) \\
v_2(k+1) = v_2(k) + h\mathrm{fh}
\end{cases}
\tag{16}
$$

where v_r is the desired input, v_1 is the transition process of v_r, v_2 is the derivative of v_1, and r and h are adjusted accordingly as filter coefficients. The function $\mathrm{fhan}(\cdot)$ is defined as follows:

$$
\mathrm{fhan}(x_1, x_2, r, h) = -
\begin{cases}
r\mathrm{sgn}(a), |a| > d_0 \\
r\frac{a}{d}, |a| \le d_0
\end{cases}
\tag{17}
$$

We established the following form, LESO:

$$
\begin{cases}
\dot{z}_1 = z_2 - \beta_1(z_1 - v_1) \\
\dot{z}_2 = z_3 - \beta_2(z_1 - v_1) + b_v u \\
\dot{z}_3 = -\beta_3(z_1 - v_1)
\end{cases}
\tag{18}
$$

Selecting the suitable observer gains $(\beta_1, \beta_2, \beta_3)$, LESO then enabled real-time observation of the variables in System (14) [50], i.e., $z_1 \to v_1$, $z_2 \to v_2$, $z_3 \to f(V, \dot{V}, \varpi_{out}, \varpi_f)$. Set

$$
u = \frac{u_0 - z_3}{b_v}
\tag{19}
$$

When z_3 can estimate $f(V, \dot{V}, \varpi_{out}, \varpi_f)$ without error, let LSEF be:

$$
\begin{cases}
e_1 = v_1 - z_1 \\
e_2 = v_2 - z_2 \\
u_0 = k_1 e_1 + k_2 e_2
\end{cases}
\tag{20}
$$

then the system (15) can be simplified to a double integral series structure:

$$
\ddot{V} = (f(V, \dot{V}, \varpi_{out}, \varpi_f) - z_3) + u_0 \approx u_0
\tag{21}
$$

Further, the bandwidth method [50] was used and we could obtain:

$$
\begin{cases}
\beta_1 = 3\omega_o \\
\beta_2 = 3\omega_o^2 \\
\beta_3 = \omega_o^3
\end{cases}
\tag{22}
$$

where ω_o is the observer bandwidth. The larger ω_o is, the smaller LESO observation errors are. However, the sensitivity of the system to noise may be increased, so the ω_o selection requires comprehensive consideration.

Similarly, according to the parameterization method and engineering experience [32], the LSEF parameters can be chosen as:

$$
\begin{cases}
k_1 = \omega_c^2 \\
k_2 = 2\xi\omega_c
\end{cases}
\tag{23}
$$

where ω_c is the controller bandwidth, ξ is the damping ratio, and in this paper $\xi = 1$. Therefore, the parameter tuning problem of LADRC controller was simplified to the observer bandwidth ω_o and controller bandwidth ω_c configuration.

3.3. TD3 Algorithm

TD3 algorithm is an offline RL algorithm based on DDPG proposed in 2015 [51]. This approach adopted a similar method implemented in Double-DQN [52] to reduce the overestimation in function approximation, delaying the update frequency in the actor–network, and adding noises to target the actor–network to release the sensitivity and instability in DDPG. The structure of TD3 is shown in Figure 4.
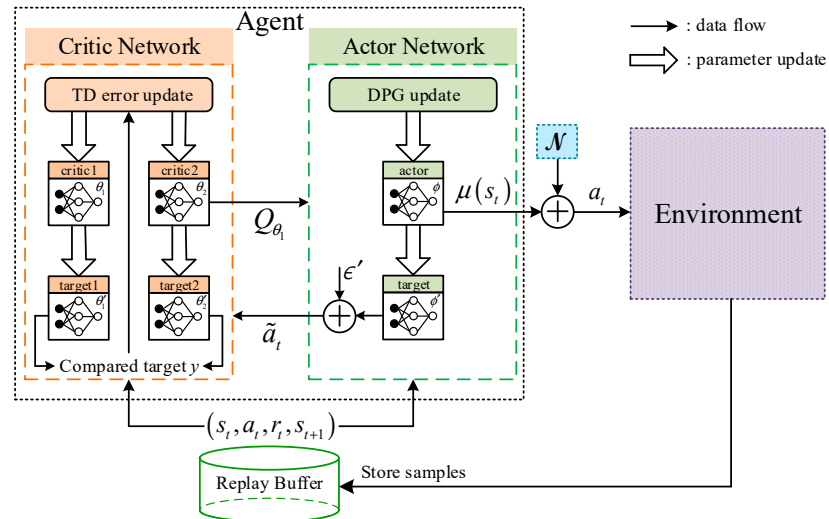

Figure 4. Structure of TD3.

Updating the parameters of critic networks by minimizing loss:

$$L = N^{-1} \sum \left(y - Q_{\theta_i}(s,a) \right)^2 \tag{24}$$

where s is the current state, a is the current action, and $Q_{\theta_i}(s,a)$ stands for the parameterized state-action value function Q with parameter θ_i.

$$y = r + \gamma \min_{i=1,2} Q_{\theta_i'}\left(s', \tilde{a} \right) \tag{25}$$

is the target value of the function $Q_\theta(s,a)$, $\gamma \in [0,1]$ is the discount factor, and the target action is defined as:

$$\tilde{a} = \pi_{\phi'}(s) + \epsilon' \tag{26}$$

where noise ϵ' follows a clipped normal distribution clip $[\mathcal{N}(0,\sigma), -c, c]$, $c > 0$. This implies that ϵ' is a random variable with $\mathcal{N}(0,\sigma)$ and belongs to the interval $[-c, c]$.

The inputs of the actor network are both $Q_\theta(s,a)$ from the critic network and the minibatch form the memory, and the output is the action given by:

$$a_t = \pi_\phi(s_t) + \epsilon \tag{27}$$

where ϕ is the parameter of the actor network, and π_ϕ is the output form the actor network, which is a deterministic and continuous value. Noise ϵ follows the normal distribution $\mathcal{N}(0,\sigma)$, and is added for exploration.

Updating the parameters of the actor–network based on deterministic gradient strategy:

$$\nabla_\phi J(\phi) = N^{-1} \sum \nabla_a Q_{\theta_1}(s,a)\big|_{a=\pi_\phi(s)} \nabla_\phi \pi_\phi(s) \tag{28}$$

TD3 updates the actor–network and all three target networks every d steps periodically in order to avoid a too fast convergence. The parameters of the critic target networks and the actor–target network are updated according to:

$$\begin{cases} \theta_i' \leftarrow \tau\theta_i + (1-\tau)\theta_i' \\ \phi' \leftarrow \tau\phi + (1-\tau)\phi' \end{cases} \tag{29}$$

The pseudocode of the proposed approach is given in Algorithm 1.

Algorithm 1. TD3
1   Initialize critic networks Q_{θ_1}, Q_{θ_2} and actor network π_ϕ with random parameters θ_1, θ_2, ϕ;
2   Initialize target networks $Q_{\theta_1'}$, $Q_{\theta_2'}$ with $\theta_1' \leftarrow \theta_1$, $\theta_2' \leftarrow \theta_2$, and target actor network $\pi_{\phi'}$ with $\phi' \leftarrow \phi$;
3   Initialize replay buffer $\mathcal{R}$;
4   **For every episode:**
5   Initialize state s;
6   **Repeat**;
7   Select action with exploration noise $a \sim \pi(a) + \mathcal{N}(0,\sigma)$;
8   Observe reward r and new state s';
9   Store transition tuple (s, a, r, s') in $\mathcal{R}$;
10   Sample mini-batch of N transitions (s, a, r, s') from $\mathcal{R}$;
11   Attain $\tilde{a} \leftarrow \pi_{\phi'}(s) + \epsilon$, where $\epsilon \sim clip(\mathcal{N}(0,\sigma), -c, c)$;
12   Update critics $\theta_i \leftarrow \min_{\theta_i} N^{-1} \sum (y - Q_{\theta_i}(s,a))^2$;
13   Every d steps:
14   Update ϕ by the deterministic policy gradient:
15   $\nabla_\phi J(\phi) = N^{-1} \sum \nabla_a Q_{\theta_1}(s,a)\big|_{a=\pi_\phi(s)} \nabla_\phi \pi_\phi(s)$;
16   Update target network:
17   $\theta_i' \leftarrow \tau\theta_i + (1-\tau)\theta_i'$
18   $\phi' \leftarrow \tau\phi + (1-\tau)\phi'$;
19   $s \leftarrow s'$;
20   **Until** s reaches terminal state s_T.

3.4. TD3-LADRC Reconfiguration Controller Design

Lack of environment adaptability, poor control performance, and weak robustness are the main shortcomings of parameter-fixed controllers [36]. When a fault occurs, it may not be possible to maintain the acceptable (rated or degraded) performance of the damaged system. Motivated by the above analysis, a reconfiguration controller called TD3-LADRC is proposed in this paper, and its control schematic is shown in Figure 5.
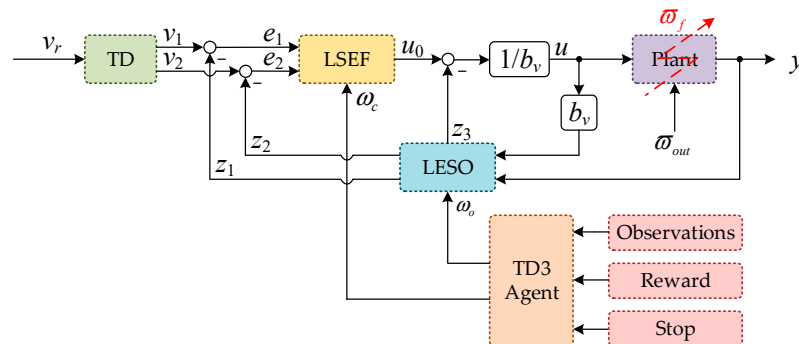

Figure 5. Control schematic of TD3-LADRC.

The deep reinforcement learning algorithm TD3 is introduced to realize the LADRC parameters adaption. The details of each part have been described above. The selection of control parameters is treated as the agent's action a_t, and the response result of the control system s_t is considered as the state, i.e., as follows:

$$\begin{cases} a_t = [\omega_o, \omega_c]^{\mathrm{T}} \\ s_t = s_{obs} = \left[e, \dot{e}, V, \dot{V}\right]^{\mathrm{T}} \end{cases} \tag{30}$$

where $e = V - V_w$, and s_{obs} is the agent observations vector.

The range of each controller parameter is selected as follows:

$$\begin{cases} \omega_c \in [0, 4] \\ \omega_o \in [100, 200] \end{cases} \tag{31}$$

The reward function plays a crucial role in the reinforcement learning algorithm. The appropriateness of the reward function design directly affects the training effect of the reinforcement learning, which in turn affects the effectiveness of the whole reconfiguration controller. According to the working characteristics of AABS, the following reward function is selected after several attempts to ensure stable and smooth braking:

$$r_t = 1\left(-6 \le \dot{V} \le -4\right) + 0\left(\dot{V} > -4 \,\|\, \dot{V} < -6\right) - 100(V < 2 \,\|\, e > 20) \tag{32}$$

The stop conditions for each training episode are as follows, and one of the three will do:

(1)   The aircraft speed $V < 2$;
(2)   The error between main wheel speed and aircraft speed $e > 20$;
(3)   Simulation time $t > 20$ s.

Remark 4. *TD3, TD, LESO, and LSEF together constitute the TD3-LADRC controller. Compared to normal LADRC, TD3-LADRC realizes the parameter adaption that makes the controller reconfigurable. The robustness and immunity are greatly improved. It can effectively compensate the adverse effects caused by the total perturbations including faults.*

3.5. TD3-LESO Estimation Capability Analysis

In order to prove the stability of the whole closed-loop system, the convergence of TD3-LESO is first analyzed in conjunction with Assumption 1 [53]. Let the estimation errors of TD3-LESO be $\tilde{x}_i = x_i - z_i, i = 1, 2, 3$, and the estimation error equation of the observer can be obtained as:

$$\begin{cases} \dot{\tilde{x}}_1 = \tilde{x}_2 - 3\omega_o \tilde{x}_1 \\ \dot{\tilde{x}}_2 = \tilde{x}_3 - 3\omega_o^2 \tilde{x}_1 \\ \dot{\tilde{x}}_3 = h(V, \dot{V}, \omega_{out}, \omega_f) - \omega_o^3 \tilde{x}_1 \end{cases} \tag{33}$$

Let $\varepsilon_i = \dfrac{\tilde{x}_i}{\omega_o^{i-1}}, i = 1, 2, 3$, then Equation (33) can be rewritten as:

$$\dot{\varepsilon} = \omega_o A_3 \varepsilon + B \frac{h(V, \dot{V}, \omega_{out}, \omega_f)}{\omega_o^2} \tag{34}$$

where $A_3 = \begin{bmatrix} -3 & 1 & 0 \\ -3 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^{\mathrm{T}}$.

Based on Assumption 1 and Theorem 2 in Reference [54], the following theorem can be obtained:

Theorem 1. *Under the condition that $h(V, \dot{V}, \omega_{out}, \omega_f)$ is bounded, the TD3-LESO estimation errors are bounded and their upper bound decrease monotonically with the increase of the observer bandwidth ω_o.*

The proof is given in the Appendix A. Thus, it is clear that there are three positive numbers $v_i, i = 1, 2, 3$, such that the state estimation error $|\widetilde{x}_i| \leq v_i$ holds, i.e., the TD3-LESO estimation errors are bounded, which can effectively estimate the states of the controlled plant and the total perturbation.

3.6. Stability Analysis of Closed-loop System

The closed-loop system consisted of the control laws (19) and (20), and the controlled object (21) is:

$$\ddot{V} = f - z_3 + k_1 e_1 + k_2 e_2 \tag{35}$$

If we defined the tracking errors as $\varepsilon_i = v_i - x_i, i = 1, 2$, then we could attain:

$$\begin{cases} \dot{\varepsilon}_1 = \dot{r}_1 - \dot{x}_1 = r_2 - x_2 = \widetilde{e}_2 \\ \dot{\varepsilon}_2 = \dot{r}_2 - \dot{x}_2 = r_3 - \ddot{V} \\ \qquad = -k_1 \varepsilon_1 - k_1 \widetilde{x}_1 - k_2 \varepsilon_2 - k_2 \widetilde{x}_2 - \widetilde{x}_3 \end{cases} \tag{36}$$

Let $\varepsilon = [\varepsilon_1, \varepsilon_2]^T, \widetilde{x} = [\widetilde{x}_1, \widetilde{x}_2, \widetilde{x}_3]^T$, then:

$$\dot{\varepsilon}(t) = A_\varepsilon \varepsilon(t) + A_{\widetilde{x}} \widetilde{x}(t) \tag{37}$$

where $A_e = \begin{bmatrix} 0 & 1 \\ -k_1 & -k_2 \end{bmatrix}, A_{\widetilde{x}} = \begin{bmatrix} 0 & 0 & 0 \\ -k_1 & -k_2 & -1 \end{bmatrix}.$

By solving Equation (37):

$$\varepsilon(t) = e^{A_\varepsilon t} \varepsilon(0) + \int_0^t e^{A_\varepsilon (t-\tau)} A_{\widetilde{x}} \widetilde{x}(\tau) d\tau \tag{38}$$

Combining Assumption 1, Theorem 1, Theorem 3, and Theorem 4 in the literature [54], the following theorem was proposed to analyze the stability of the closed-loop system:

Theorem 2. *Under the condition that the TD3-LESO estimation errors are bounded, there exists a controller bandwidth ω_c, such that the tracking error of the closed-loop system is bounded. Thus, for a bounded input, the output of the closed-loop system is bounded, i.e., the closed-loop system is BIBO-stable.*

See the Appendix A for proof.

4. Simulation Results

In order to verify the reconfiguration capability and disturbance rejection capabilities of the proposed method, the corresponding simulations are carried out in this section and compared with conventional PID + PBM and LADRC.

The initial states of the aircraft are set as follows:

(1) The initial speed of aircraft landing $V(0) = 72$ m/s;
(2) The initial height of the center of gravity $Hh = 2.178$ m.

To prevent deep wheel slippage as well as tire blowout, the wheel speed was kept following the aircraft speed quickly at first, and the brake pressure was applied only after 1.5 s. The anti-skid brake control was considered to be over when V was less than 2 m/s.

In the experiment, both the critic networks and the actor networks were realized by a fully connected neural network with three hidden layers. The number of neurons in the hidden layer was (50,25,25). The activation function of the hidden layer was selected as the ReLU function, and the activation function of the output layer of the actor network was selected as the tanh function. In addition, the parameters of the actor network and the critic network were tuned by an Adam optimizer. The remaining parameters of TD3-LADRC are shown in Table 7.

Table 7. Parameters of TD3-LADRC.

Name	Value
Control gain b_v	2
$TD - r$	0.001
$TD - h$	1
Discount factor γ	0.99
Actor learning rate	0.0001
Critic learning rate	0.001
Target update rate τ	0.001

Remark 5. *It is noted that the braking time t and braking distance x are selected as the criteria for braking efficiency, and the system stability is observed by slip rate λ.*

The model simulation was carried out in MATLAB 2022a, and the TD3 algorithm was realized through the reinforcement learning toolbox. The simulation time was 20 s, the sampling time was 0.001 s. The training stopped when the average reward reached 12,000. The training took about 6 h to complete. The learning curves of the reward obtained by the agent for each interaction with the environment during the training process are shown in Figure 6.
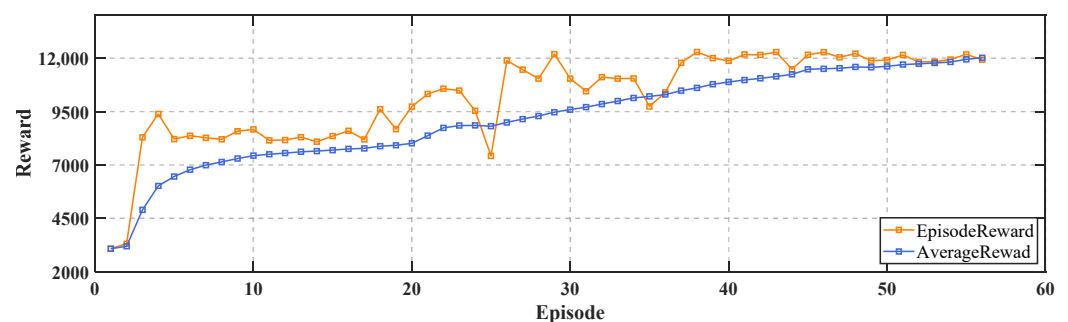

Figure 6. Learning curves.

It can be seen that at the beginning of the training, the agent was in the exploration phase and the reward obtained was relatively low. Later, the reward gradually increased, and after 40 episodes, the reward was steadily maintained at a high level and the algorithm gradually converges.

### 4.1. Case 1: Fault-Free and External Disturbance-Free in Dry Runway Condition

The simulation results of the dynamic braking process for different control schemes are shown in Figures 7 and 8 and Table 8.

As can be seen from Figure 7, PID + PBM leads to numerous skids during braking, which may cause serious loss to the tires. In contrast, LADRC and TD3-LADRC not only skid less frequently, but also have shorter braking time and braking distance. Moreover, the control effect of TD3-LADRC is better than LADRC. Figure 8 shows that TD3-LADRC can dynamically tune the controller parameters to accurately observe and compensate for the total disturbances, and thus improve the AABS performance.
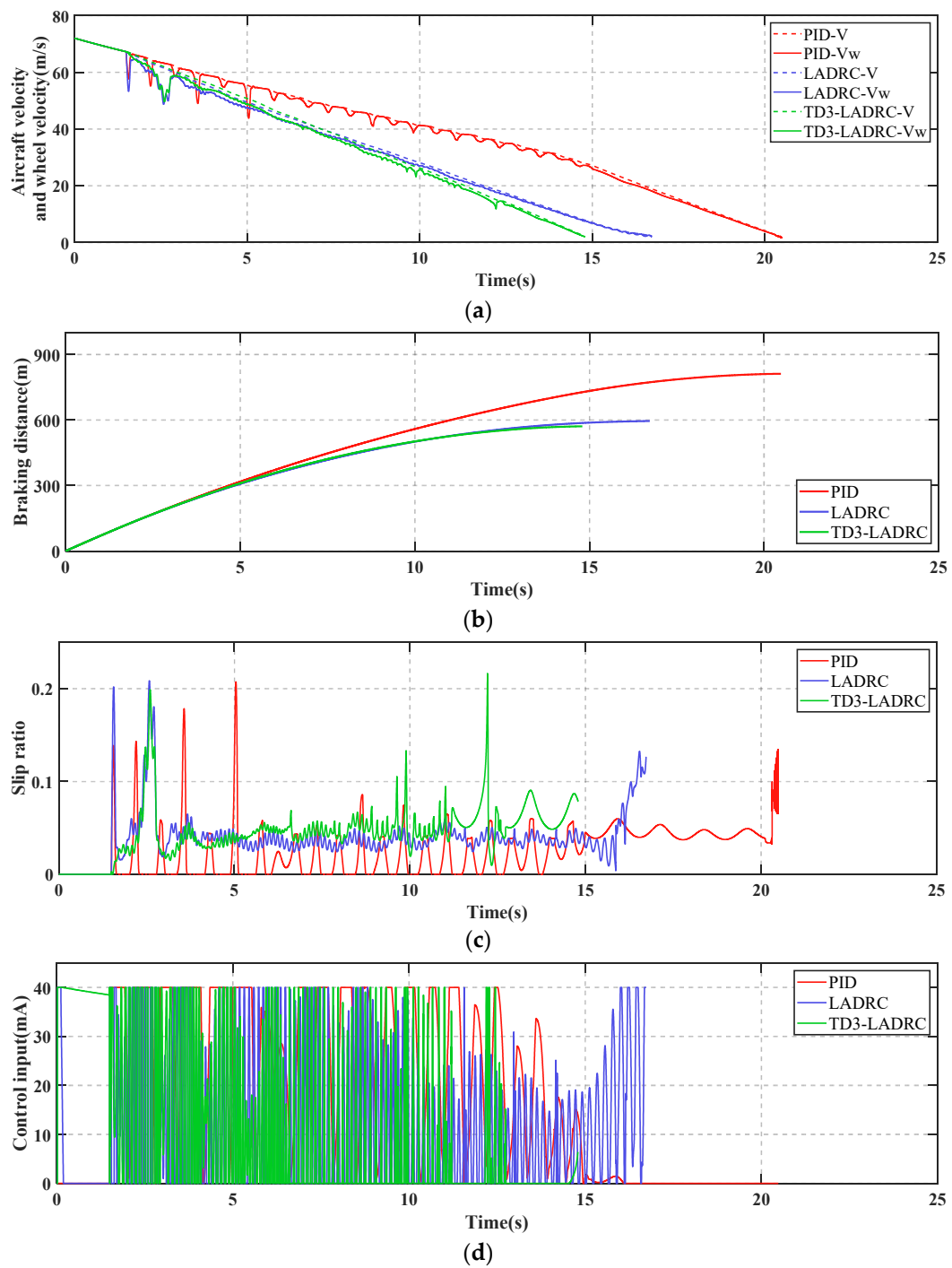
Figure 7. (**a**) Aircraft velocity and wheel velocity; (**b**) breaking distance; (**c**) slip ratio; (**d**) control input.
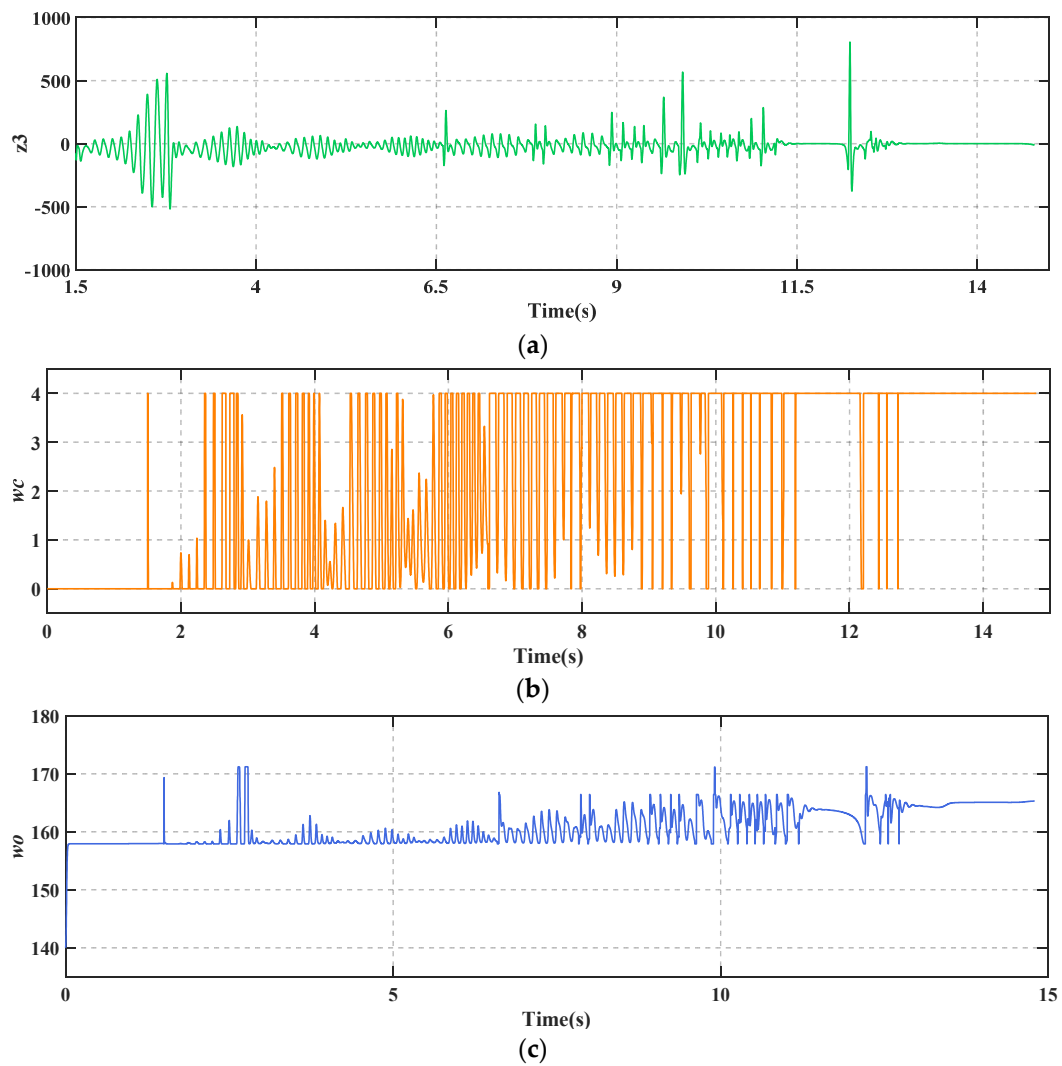
Figure 8. (**a**) Extended state of TD3-LADRC; (**b**) controller bandwidth ω_c; (**c**) controller bandwidth ω_o.
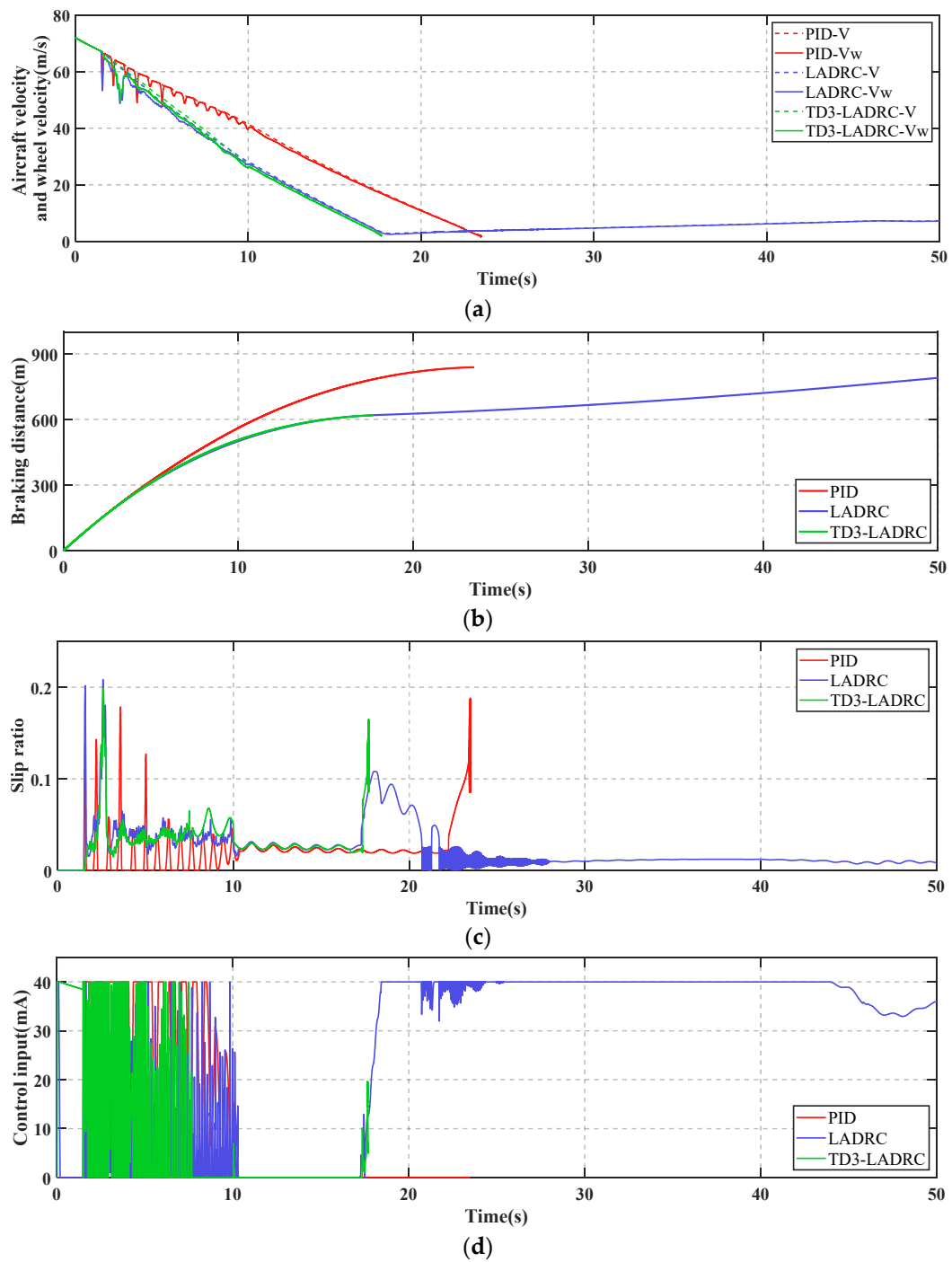
Table 8. AABS performance index.

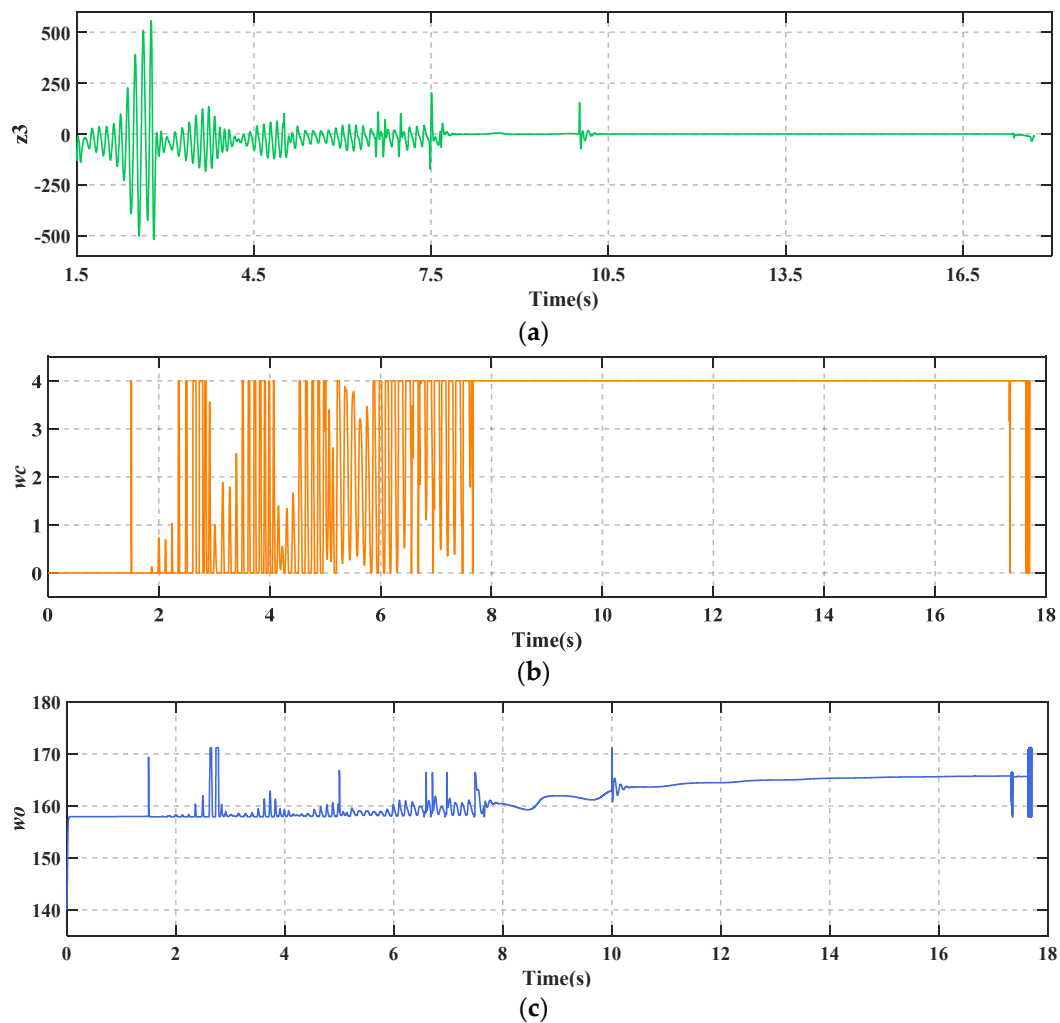| Performance Index	PID + PBM	LADRC	TD3-LADRC
Braking time (s)	20.48	16.73	14.79
Braking distance (m)	811.9	595.46	571.18

Remark 6. *During the braking process, it is observed that in some instants $\omega_c = 0$. It may not affect the stability of the whole system. On the one hand, the value of ω_c does not change the fact that A_ε is Hurwitz (see Proof of Theorem 2 for details). On the other hand, ω_c is constantly changed by the agent through a continuous interaction with the environment, and in these instants the agent considers $\omega_c = 0$ as optimal, i.e., no anti-skid braking control leads to better braking results.*

4.2. Case 2: Actuator LOE Fault in Dry Runway Condition

The fault considered here assumed a 20% actuator LOE at 5 s and escalated to 40% LOE at 10 s. The simulation results are shown in Figures 9 and 10 and Table 9.

Figure 9. (**a**) Aircraft velocity and wheel velocity; (**b**) breaking distance; (**c**) slip ratio; (**d**) control input.

Figure 10. (**a**) Extended state of TD3-LADRC; (**b**) controller bandwidth ω_c; (**c**) controller bandwidth ω_o.
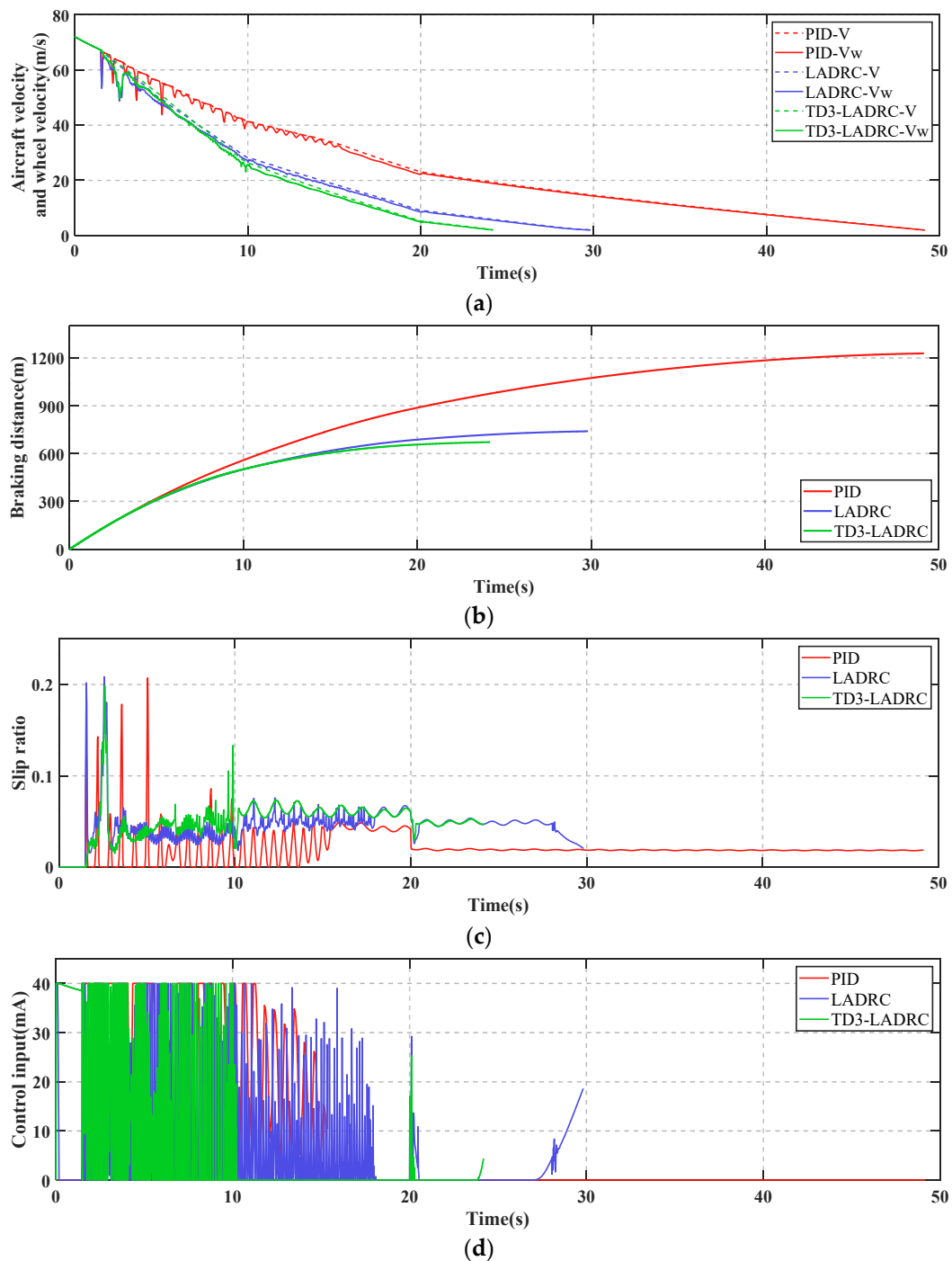
Table 9. AABS performance index.

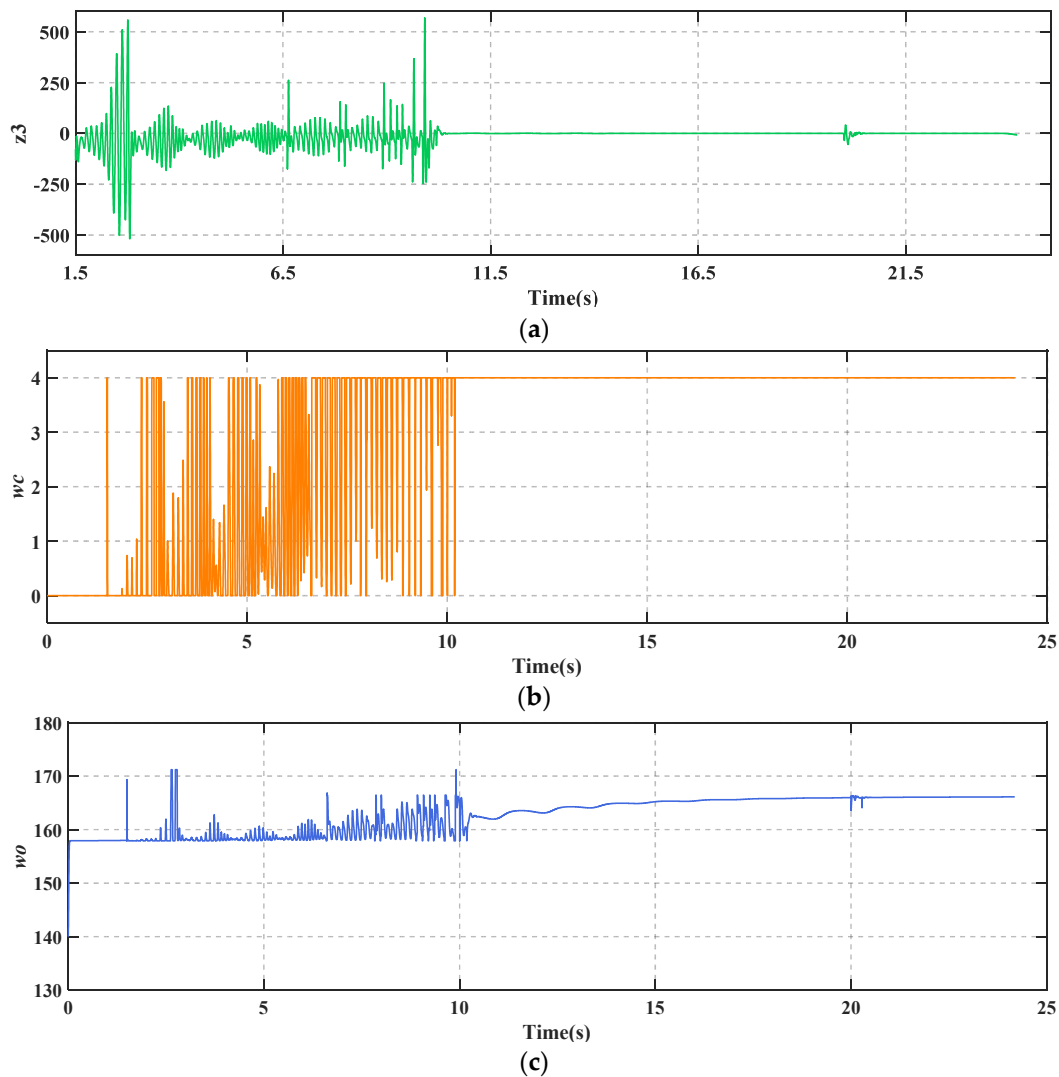| Performance Index	PID + PBM	LADRC	TD3-LADRC
Braking time (s)	23.48	-	17.70
Braking distance (m)	838.46	-	618.12

As can be seen in Figure 9, PID + PBM continuously performed a large braking and releasing operation under the combined effect of fault and disturbance. This makes braking much less efficient and risks dragging and flat tires. In addition, LADRC cannot brake the aircraft to a stop which is not allowed in practice. Figure 9c shows that there is a high frequency of wheel slip in the low-speed phase of the aircraft. In contrast, TD3-LADRC retains the experience gained from the agent's prior training and continuously adjusts the controller parameters online based on the plant states, which ultimately allows the aircraft to brake smoothly. From Figure 10a, it can be seen that the total fault perturbations are estimated fast and accurately based on the adaptive LESO. Overall, TD3-LADRC not only improves the robustness and immunity of the controller in fault-perturbed conditions, but also greatly significantly improves the safety and reliability of AABS.

4.3. Case 3: Actuator LOE Fault in Mixed Runway Condition

The mixed runway structure is as follows: dry runway in the interval of 0–10 s, wet runway in the interval of 10–20 s, and snow runway after 20 s. The fault considered here assumed a 10% actuator LOE at 10 s. The simulation results are shown in Figures 11 and 12 and Table 10.

Figure 11. (**a**) Aircraft velocity and wheel velocity; (**b**) breaking distance; (**c**) slip ratio; (**d**) control input.

Figure 12. (**a**) Extended state of TD3-LADRC; (**b**) controller bandwidth ω_c; (**c**) controller bandwidth ω_o.

Table 10. AABS performance index.

Performance Index	PID + PBM	LADRC	TD3-LADRC
Braking time (s)	49.14	29.82	24.19
Braking distance (m)	1228.71	739.99	672.03

The deterioration of the runway conditions has resulted in a very poor tire–ground bond. It can be seen from Figure 11 that both braking time and braking distance have increased compared to the dry runway. Figure 12 shows that TD3-LADRC is still able to achieve controller parameters adaption, accurately observe the total fault perturbations, and effectively compensate for the adverse effects. The whole reconfiguration control system adapts well to runway changes. The environmental adaptability of AABS is improved.

5. Conclusions

A linear active disturbance rejection reconfiguration control scheme based on deep reinforcement learning was proposed to meet the higher performance requirements of AABS under fault-perturbed conditions. According to the composition structure and working principle, AABS mathematical model with an actuator fault factor is established.

A TD3-LADRC reconfiguration controller was developed, and the parameters of LSEF and LESO were adjusted online using the TD3 algorithm. The simulation results under different conditions verified that the designed controller can effectively improve the anti-skid braking performance even under faults and perturbations, as well as different runway environments. It successfully strengthened the robustness, immunity, and environmental adaptability of the AABS, thereby improving the safety and reliability of the aircraft. However, TD3-LADRC is so complex that its control effectiveness was verified only by simulations in this paper. The combined effect caused by various uncertainties in practical applications on the robustness of the controller cannot be completely considered. Therefore, in future work, an aircraft braking hardware-in-loop experimental platform is necessary to build, consisting of the host PC, the target CPU, the anti-skid braking controller, the actuators, and the aircraft wheel. The host PC and the target CPU are the software simulation part, while the other four parts are the hardware part.

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

Data Availability Statement: Not applicable.

Appendix A. Proof of Theorems

Proof of Theorem 1. By solving Equation (34) we can attain:

$$\varepsilon(t) = e^{\omega_o A_3 t}\varepsilon(0) + \int_0^t e^{\omega_o A_3(t-\tau)} B \frac{h(V(\tau), \dot{V}(\tau), \omega_{out}, \omega_f)}{\omega_o^2} d\tau \tag{A1}$$

Define $\zeta(t)$ as follows:

$$\zeta(t) = \int_0^t e^{\omega_o A_3(t-\tau)} B \frac{h(V(\tau), \dot{V}(\tau), \omega_{out}, \omega_f)}{\omega_o^2} d\tau \tag{A2}$$

From the fact that $h(V(\tau), \dot{V}(\tau), \omega_{out}, \omega_f)$ is bounded, we have:

$$|\zeta_i(t)| \le \frac{\sigma}{\omega_o^3}\left[\left|\left(A_3^{-1}B\right)_i\right| + \left|\left(A_3^{-1}e^{\omega_o A_3 t}B\right)_i\right|\right] \tag{A3}$$

Because $A_3^{-1} = \begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & -3 \\ 0 & 1 & -3 \end{bmatrix}$, we can attain:

$$\left|\left(A_3^{-1}B\right)_i\right| \le 3 \tag{A4}$$

Considering that A_3 is Hurwitz, there is thus a finite time T_1 so that for any $t \geq T_1$, $i, j = 1, 2, 3$, the following formula holds [54]:

$$\left| \left[e^{\omega_o A_3 t} \right]_{ij} \right| \leq \frac{1}{\omega_o^3} \tag{A5}$$

Therefore, the following formula is satisfied:

$$\left| \left[e^{\omega_o A_3 t} B \right]_i \right| \leq \frac{1}{\omega_o^3} \tag{A6}$$

Finally, we can attain:

$$\left| \left(A_3^{-1} e^{\omega_o A_3 t} B \right)_i \right| \leq \frac{4}{\omega_o^3} \tag{A7}$$

From Equations (A3), (A4), and (A7) we can attain:

$$|\zeta_i(t)| \leq \frac{3\sigma}{\omega_o^3} + \frac{4\sigma}{\omega_o^6} \tag{A8}$$

Let $\varepsilon_{\text{sum}}(0) = |\varepsilon_1(0)| + |\varepsilon_2(0)| + |\varepsilon_3(0)|$, for all $t \geq T_1$, the following formula holds:

$$\left| \left[e^{\omega_o A_3 t} \varepsilon(0) \right]_i \right| \leq \frac{\varepsilon_{sum}(0)}{\omega_o^3} \tag{A9}$$

Form Equation (A1) we can attain:

$$|\varepsilon_i(t)| \leq \left| \left[e^{\omega_o A_3 t} \varepsilon(0) \right]_i \right| + |\zeta_i(t)| \tag{A10}$$

Let $\widetilde{x}_{sum}(0) = |\widetilde{x}_1(0)| + |\widetilde{x}_2(0)| + |\widetilde{x}_3(0)|$, from $\varepsilon_i = \frac{\widetilde{x}_i}{\omega_o^{i-1}}$ and formulas (A8)–(A10), we can attain:

$$|\widetilde{x}_i(t)| \leq \left| \frac{\widetilde{x}_{sum}(0)}{\omega_o^3} \right| + \frac{3\sigma}{\omega_o^{4-i}} + \frac{4\sigma}{\omega_o^{7-i}} = v_i \tag{A11}$$

For all $t \geq T_1, i = 1, 2, 3$, the above formula holds. $\square$

Proof of Theorem 2. According to Equation (37) and Theorem 1, we can attain:

$$\begin{cases} [A_{\widetilde{x}} \widetilde{x}(\tau)]_1 = 0 \\ |[A_{\widetilde{x}} \widetilde{x}(\tau)]_2| \leq k_{sum} v_i = \gamma_l, \forall t \geq T_1 \end{cases} \tag{A12}$$

where $k_{sum} = 1 + k_1 + k_2$, bringing in the controller bandwidth $k_{sum} = 1 + \omega_c^2 + 2\omega_c$, and taking the parameters in this way ensures that A_ε is Hurwitz [54].

Define $\Theta = \begin{bmatrix} 0 & \gamma_l \end{bmatrix}^T$, let $\vartheta(t) = \int_0^t e^{A_\varepsilon(t-\tau)} A_{\widetilde{x}} \widetilde{x}(\tau) d\tau$, then we can attain:

$$|\vartheta_i(t)| \leq \left| \left(A_\varepsilon^{-1} \Theta \right)_i \right| + \left| \left(A_\varepsilon^{-1} e^{A_\varepsilon t} \Theta \right)_i \right|, i = 1, 2 \tag{A13}$$

$$\begin{cases} \left| (A_\varepsilon^{-1} \Theta)_1 \right| = \frac{\gamma_l}{k_1} = \frac{\gamma_l}{\omega_c^2} \\ \left| (A_\varepsilon^{-1} \Theta)_2 \right| = 0 \end{cases} \tag{A14}$$

Consider that A_ε is Hurwitz; thus, there is a finite time T_2 so that for any $t \geq T_2$, $i, j = 1, 2, 3$, the following formula holds [54]:

$$\left| \left[e^{A_\varepsilon t} \right]_{ij} \right| \leq \frac{1}{\omega_c^3} \tag{A15}$$

Let $T_3 = \max\{T_1, T_2\}$, for any $t \geq T_3$, $i = 1, 2$, we can attain:

$$\left|(e^{A_\varepsilon t}\Theta)_i\right| \leq \frac{\gamma_l}{\omega_c^3} \tag{A16}$$

Then we can attain:

$$\left|\left(A_\varepsilon^{-1} e^{A_\varepsilon t}\Theta\right)_i\right| \leq \begin{cases} \dfrac{1+k_2}{\omega_c^2}\dfrac{\gamma_l}{\omega_c^3}, & i = 1 \\[2ex] \dfrac{\gamma_l}{\omega_c^3}, & i = 2 \end{cases} \tag{A17}$$

From Equations (A13), (A14), and (A17) we can attain that for any $t \geq T_3$:

$$|\vartheta_i(t)| \leq \begin{cases} \dfrac{\gamma_l}{\omega_c^2} + \dfrac{(1+k_2)\gamma_l}{\omega_c^5}, & i = 1 \\[2ex] \dfrac{\gamma_l}{\omega_c^3}, & i = 2 \end{cases} \tag{A18}$$

Let $\varepsilon_s(0) = |\varepsilon_1(0)| + |\varepsilon_2(0)|$, then for any $t \geq T_3$:

$$\left|\left[e^{A_\varepsilon t}\varepsilon(0)\right]_i\right| \leq \frac{\varepsilon_s(0)}{\omega_c^3} \tag{A19}$$

From Equation (A12), we can attain:

$$|\varepsilon_i(t)| \leq \left|\left[e^{A_\varepsilon t}\varepsilon(0)\right]_i\right| + |\vartheta_i(t)| \tag{A20}$$

From Equations (A12), (A18)–(A20), we can attain that for any $t \geq T_3$, $i = 1, 2$:

$$|\varepsilon_i(t)| \leq \begin{cases} \dfrac{\varepsilon_s(0)}{\omega_c^3} + \dfrac{k_{sum}v_i}{\omega_c^2} + \dfrac{(1+k_2)k_{sum}v_i}{\omega_c^5}, & i = 1 \\[2ex] \dfrac{k_{sum}v_i + \varepsilon_s(0)}{\omega_c^3}, & i = 2 \end{cases} \tag{A21}$$

$\square$

References

1. Li, F.; Jiao, Z. Adaptive Aircraft Anti-Skid Braking Control Based on Joint Force Model. *J. Beijing Univ. Aeronaut. Astronaut.* **2013**, *4*, 447–452.
2. Jiao, Z.; Sun, D.; Shang, Y.; Liu, X.; Wu, S. A high efficiency aircraft anti-skid brake control with runway identification. *Aerosp. Sci. Technol.* **2019**, *91*, 82–95. [CrossRef]
3. Chen, M.; Liu, W.; Ma, Y.; Wang, J.; Xu, F.; Wang, Y. Mixed slip-deceleration PID control of aircraft wheel braking system. *IFAC-PapersOnLine* **2018**, *51*, 160–165. [CrossRef]
4. Chen, M.; Xu, F.; Liang, X.; Liu, W. MSD-based NMPC Aircraft Anti-skid Brake Control Method Considering Runway Variation. *IEEE Access* **2021**, *9*, 51793–51804. [CrossRef]
5. Dinçmen, E.; Güvenç, B.A.; Acarman, T. Extremum-seeking control of ABS braking in road vehicles with lateral force improvement. *IEEE Trans. Contr. Syst. Technol.* **2012**, *22*, 230–237. [CrossRef]
6. Li, F.B.; Huang, P.M.; Yang, C.H.; Liao, L.Q.; Gui, W.H. Sliding mode control design of aircraft electric brake system based on nonlinear disturbance observer. *Acta Autom. Sin.* **2021**, *47*, 2557–2569.
7. Radac, M.B.; Precup, R.E. Data-driven model-free slip control of anti-lock braking systems using reinforcement Q-learning. *Neurocomputing* **2018**, *275*, 317–329. [CrossRef]
8. Zhang, R.; Peng, J.; Chen, B.; Gao, K.; Yang, Y.; Huang, Z. Prescribed Performance Active Braking Control with Reference Adaptation for High-Speed Trains. *Actuators* **2021**, *10*, 313. [CrossRef]
9. Qiu, Y.; Liang, X.; Dai, Z. Backstepping dynamic surface control for an anti-skid braking system. *Control Eng. Pract.* **2015**, *42*, 140–152. [CrossRef]
10. Mirzaei, A.; Moallem, M.; Dehkordi, B.M.; Fahimi, B. Design of an Optimal Fuzzy Controller for Antilock Braking Systems. *IEEE Trans. Veh. Technol.* **2006**, *55*, 1725–1730. [CrossRef]

11. Xiang, Y.; Jin, J. Hybrid Fault-Tolerant Flight Control System Design Against Partial Actuator Failures. *IEEE Trans. Control Syst. Technol.* **2012**, *20*, 871–886.

12. Niksefat, N.; Sepehri, N. A QFT fault-tolerant control for electrohydraulic positioning systems. *IEEE Trans. Control Syst. Technol.* **2002**, *4*, 626–632. [CrossRef]

13. Wang, D.Y.; Tu, Y.; Liu, C. Connotation and research of reconfigurability for spacecraft control systems: A review. *Acta Autom. Sin.* **2017**, *43*, 1687–1702.

14. Han, Y.G.; Liu, Z.P.; Dong, Z.C. *Research on Present Situation and Development Direction of Aircraft Anti-Skid Braking System*; China Aviation Publishing & Media CO., LTD.: Xi'an, China, 2020; Volume 5, pp. 525–529.

15. Calise, A.J.; Lee, S.; Sharma, M. Development of a reconfigurable flight control law for tailless aircraft. *J. Guid. Control Dyn.* **2001**, *24*, 896–902. [CrossRef]

16. Yin, S.; Xiao, B.; Ding, S.X.; Zhou, D. A review on recent development of spacecraft attitude fault tolerant control system. *IEEE Trans. Ind. Electron.* **2016**, *63*, 3311–3320. [CrossRef]

17. Zhang, Y.; Jin, J. Bibliographical review on reconfigurable fault-tolerant control systems. *Annu. Rev. Control* **2008**, *32*, 229–252. [CrossRef]

18. Zhang, Z.; Yang, Z.; Xiong, S.; Chen, S.; Liu, S.; Zhang, X. Simple Adaptive Control-Based Reconfiguration Design of Cabin Pressure Control System. *Complexity* **2021**, *2021*, 6635571. [CrossRef]

19. Chen, F.; Wu, Q.; Jiang, B.; Tao, G. A reconfiguration scheme for quadrotor helicopter via simple adaptive control and quantum logic. *IEEE Trans. Ind. Electron.* **2015**, *62*, 4328–4335. [CrossRef]

20. Guo, Y.; Jiang, B. Multiple model-based adaptive reconfiguration control for actuator fault. *Acta Autom. Sinica* **2009**, *35*, 1452–1458. [CrossRef]

21. Gao, Z.; Jiang, B.; Shi, P.; Qian, M.; Lin, J. Active fault tolerant control design for reusable launch vehicle using adaptive sliding mode technique. *J. Frankl. Inst.* **2012**, *349*, 1543–1560. [CrossRef]

22. Shen, Q.; Jiang, B.; Cocquempot, V. Fuzzy Logic System-Based Adaptive Fault-Tolerant Control for Near-Space Vehicle Attitude Dynamics with Actuator Faults. *IEEE Trans. Fuzzy Syst.* **2013**, *21*, 289–300. [CrossRef]

23. Lv, X.; Jiang, B.; Qi, R.; Zhao, J. Survey on nonlinear reconfigurable flight control. *J. Syst. Eng. Electron.* **2013**, *24*, 971–983. [CrossRef]

24. Han, J. From PID to active disturbance rejection control. *IEEE Trans. Ind. Electron.* **2009**, *56*, 900–906. [CrossRef]

25. Huang, Y.; Xue, W. Active disturbance rejection control: Methodology and theoretical analysis. *ISA Trans.* **2014**, *53*, 963–976. [CrossRef] [PubMed]

26. Guo, Y.; Jiang, B.; Zhang, Y. A novel robust attitude control for quadrotor aircraft subject to actuator faults and wind gusts. *IEEE/CAA J. Autom. Sin.* **2017**, *5*, 292–300. [CrossRef]

27. Zhang, Z.; Yang, Z.; Zhou, G.; Liu, S.; Zhou, D.; Chen, S.; Zhang, X. Adaptive Fuzzy Active-Disturbance Rejection Control-Based Reconfiguration Controller Design for Aircraft Anti-Skid Braking System. *Actuators* **2021**, *10*, 201. [CrossRef]

28. Zhou, L.; Ma, L.; Wang, J. Fault tolerant control for a class of nonlinear system based on active disturbance rejection control and rbf neural networks. In Proceedings of the 2017 36th Chinese Control Conference (CCC), Dalian, China, 26–28 July 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 7321–7326.

29. Tan, W.; Fu, C. Linear active disturbance-rejection control: Analysis and tuning via IMC. *IEEE Trans. Ind. Electron.* **2015**, *63*, 2350–2359. [CrossRef]

30. Gao, Z. Scaling and bandwidth-parameterization based controller tuning. In Proceedings of the 2003 American Control Conference, Denver, CO, USA, 4–6 June 2003; pp. 4989–4996.

31. Gao, Z. Active disturbance rejection control: A paradigm shift in feedback control system design. In Proceedings of the 2006 American Control Conference, Minneapolis, MN, USA, 14–16 June 2006; IEEE: Piscataway, NJ, USA, 2006; pp. 2399–2405.

32. Li, P.; Zhu, G.; Zhang, M. Linear active disturbance rejection control for servo motor systems with input delay via internal model control rules. *IEEE Trans. Ind. Electron.* **2020**, *68*, 1077–1086. [CrossRef]

33. Wang, L.X.; Zhao, D.X.; Liu, F.C.; Liu, Q.; Meng, F.L. Linear active disturbance rejection control for electro-hydraulic proportional position synchronous. *Control. Theory Appl.* **2018**, *35*, 1618–1625.

34. Li, J.; Qi, X.H.; Wan, H.; Xia, Y.Q. Active disturbance rejection control: Theoretical results summary and future researches. *Control Theory Appl.* **2017**, *34*, 281–295.

35. Qiao, G.L.; Tong, C.N.; Sun, Y.K. Study on Mould Level and Casting Speed Coordination Control Based on ADRC with DRNN Optimization. *Acta Autom. Sin.* **2007**, *33*, 641–648.

36. Qi, X.H.; Li, J.; Han, S.T. Adaptive active disturbance rejection control and its simulation based on BP neural network. *Acta Armamentarii* **2013**, *34*, 776–782.

37. Sun, K.; Xu, Z.L.; Gai, K.; Zou, J.Y.; Dou, R.Z. Novel Position Controller of Pmsm Servo System Based on Active-disturbance Rejection Controller. *Proc. Chin. Soc. Electr. Eng.* **2007**, *27*, 43–46.

38. Dou, J.X.; Kong, X.X.; Wen, B.C. Attitude fuzzy active disturbance rejection controller design of quadrotor UAV and its stability analysis. *J. Chin. Inert. Technol.* **2015**, *23*, 824–830.

39. Zhao, X.; Wu, C. Current deviation decoupling control based on sliding mode active disturbance rejection for PMLSM. *Opt. Precis. Eng.* **2022**, *30*, 431–441. [CrossRef]

40. Li, B.; Zeng, L.; Zhang, P.; Zhu, Z. Sliding mode active disturbance rejection decoupling control for active magnetic bearings. *Electr. Mach. Control.* **2021**, *7*, 129–138.
41. Buşoniu, L.; de Bruin, T.; Tolić, D.; Kober, J.; Palunko, I. Reinforcement learning for control: Performance, stability, and deep approximators. *Annu. Rev. Control* **2018**, *46*, 8–28. [CrossRef]
42. Nian, R.; Liu, J.; Huang, B. A review on reinforcement learning: Introduction and applications in industrial process control. *Comput. Chem. Eng.* **2020**, *139*, 106886. [CrossRef]
43. Yuan, Z.L.; He, R.Z.; Yao, C.; Li, J.; Ban, X.J. Online reinforcement learning control algorithm for concentration of thickener underflow. *Acta Autom. Sin.* **2021**, *47*, 1558–1571.
44. Pang, B.; Jiang, Z.P.; Mareels, I. Reinforcement learning for adaptive optimal control of continuous-time linear periodic systems. *Automatica* **2020**, *118*, 109035. [CrossRef]
45. Chen, Z.; Qin, B.; Sun, M.; Sun, Q. Q-Learning-based parameters adaptive algorithm for active disturbance rejection control and its application to ship course control. *Neurocomputing* **2020**, *408*, 51–63. [CrossRef]
46. Zou, M.Y. Design and Simulation Research on New Control Law of Aircraft Anti-skid Braking System. Master's Thesis, Northwestern Polytechnical University, Xi'an, China, 2005.
47. Pacejka, H.B.; Bakker, E. The magic formula tyre model. *Veh. Syst. Dyn.* **1992**, *21*, 1–18. [CrossRef]
48. Wang, J.S. Nonlinear Control Theory and its Application to Aircraft Antiskid Brake Systems. Master's Thesis, Northwestern Polytechnical University, Xi'an, China, 2001.
49. Jiao, Z.; Liu, X.; Shang, Y.; Huang, C. An integrated self-energized brake system for aircrafts based on a switching valve control. *Aerosp. Sci. Technol.* **2017**, *60*, 20–30. [CrossRef]
50. Yuan, D.; Ma, X.J.; Zeng, Q.H.; Qiu, X. Research on frequency-band characteristics and parameters configuration of linear active disturbance rejection control for second-order systems. *Control Theory Appl.* **2013**, *30*, 1630–1640.
51. Lillicrap, T.P.; Hunt, J.J.; Pritzel, A.; Heess, N.; Erez, T.; Tassa, Y.; Silver, D.; Wierstra, D. Continuous control with deep reinforcement learning. *arXiv* **2015**, arXiv:1509.02971.
52. Fujimoto, S.; Hoof, H.; Meger, D. Addressing function approximation error in actor-critic methods. In Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018; pp. 1587–1596.
53. Chen, Z.Q.; Sun, M.W.; Yang, R.G. On the Stability of Linear Active Disturbance Rejection Control. *Acta Autom. Sin.* **2013**, *39*, 574–580. [CrossRef]
54. Zheng, Q.; Gao, L.Q.; Gao, Z. On stability analysis of active disturbance rejection control for nonlinear time-varying plants with unknown dynamics. In Proceedings of the 2007 46th IEEE conference on decision and control, New Orleans, LA, USA, 12–14 December 2007; IEEE: Piscataway, NJ, USA, 2007; pp. 3501–3506.