Article

Speed Optimization Control of a Permanent Magnet Synchronous Motor Based on TD3

Zuolei Hu [1][iD]**, Yingjie Zhang** [1,*]**, Ming Li** [2] **and Yuhua Liao** [3]

[1] College of Information Science and Engineering, Hunan University, Changsha 410000, China; huzuolei@hnu.edu.cn
[2] School of Information and Electrical Engineering, Hunan University of Science and Technology, Xiangtan 411201, China; minglee@hnust.edu.cn
[3] Department of Mechanical, Aerospace and Civil Engineering, the University of Manchester, Manchester M13 9PL, UK; liaoyuhuahua@outlook.com
[*] Correspondence: zhangyj@hnu.edu.cn

Abstract: Permanent magnet synchronous motors (PMSMs) are widely used in industrial automation and electric vehicles due to their high efficiency and excellent dynamic performance. However, controlling PMSMs presents challenges such as parameter variations and system nonlinearities. This paper proposes a twin delayed deep deterministic policy gradient (TD3)-based energy-saving optimization control method for PMSM drive systems. The TD3 algorithm uses double networks, target policy smoothing regularization, and delayed actor network updates to improve training stability and accuracy. Simulation experiments under two operating conditions show that the TD3 algorithm outperforms traditional proportional–integral (PI) controllers and linear active disturbance rejection control (LADRC) controllers in terms of reference trajectory tracking, q-axis current regulation, and speed tracking error minimization. The results demonstrate the TD3 algorithm's effectiveness in enhancing motor efficiency and system robustness, offering a novel approach to PMSM drive system control through deep reinforcement learning.

Keywords: PMSM; TD3; optimization control; energy-saving

1. Introduction

PMSMs have gained significant traction in industrial automation and electric vehicle applications due to their exceptional efficiency, high power density, and superior dynamic performance [1,2]. The PMSM drive system represents a typical dual-timescale system. Among the most efficient control designs for PMSMs is the cascade structure, which incorporates a rapid inner loop responsible for armature current control while employing a slower outer loop intended to adjust angular velocity by generating appropriate current reference signals. However, achieving the desired control performance for PMSM applications presents challenges, including issues such as parameter variations, external disturbances, and system nonlinearity. In reference [3], it is pointed out that the permanent magnetic flux linkage can vary by 20% of its nominal value, while the stator resistance can vary by 200% of its nominal value. In practical applications, researchers often aim to employ a stable and highly accurate control strategy that can effectively achieve a rapid response and robustness in the presence of uncertain parameters. However, conventional PI controllers prove inadequate for tracking the outer loop velocity when motor parameter uncertainties arise. Consequently, researchers have explored diverse advanced control techniques to address the velocity servo problem, encompassing model predictive control (MPC), robust

control, adaptive control, fuzzy control, disturbance rejection control, sliding mode control, prediction-based model-free control, and deep learning and reinforcement learning-based controls. Paper [4] proposes a robust model predictive current control method based on nonlinear extended state observation to enhance the control performance of PMSM in the presence of parameter variations. Paper [5] presents a robust adaptive model predictive speed control method based on a recurrent neural network to tackle the speed control problem of PMSMs under the conditions of mismatched parameters.

Paper [6] presents a universal control framework that utilizes an observer to estimate both the system state and disturbances, while establishing a predictive current controller based on an enhanced system model. Paper [7] introduces a motor-parameter-free model predictive voltage control strategy for PMSM drive systems. Its fundamental concept aims to reduce dependence on motor parameters in PMSM control, thus enhancing the robustness of this control strategy. Papers [8–12] employ the MPC algorithm to mitigate torque fluctuations in PMSM drive systems induced by interturn faults (ITFs). It simplifies the complexity of control methods for PMSM drives, considering ITF by incorporating an adaptive compensation current approach.

In the field of robust adaptive control, the compensator based on the extended state observer proposed in reference [13] effectively addresses the issue of excessively high switching gains required for disturbance rejection. Reference [14] introduced a system transformation method that effectively converts a PMSM system with current constraints into an unconstrained system, thereby streamlining the controller design process. Reference [15] proposed an observer scheme utilizing neural networks along with a sensorless robust optimal control approach to address the speed and current tracking challenges in partially unknown PMSM systems under disturbances and saturation voltages. The paper [16] integrates the adaptive integral sliding mode method and employs a self-regulation approach to adjust the amplitude of the sliding mode function as well as compensate for load disturbances, thereby enhancing the dynamic performance of the system. The paper [17–19] utilizes deep reinforcement learning (DRL) to solve the control problem of PMSM. By introducing artificial intelligence algorithms into the traditional parameter optimization process, a DRL model is constructed that can automatically optimize and adjust parameters in different application scenarios, thereby achieving optimal control effects in various environments.

In the domain of prediction-based control strategies, reference [20] introduces an advanced model-free active disturbance rejection dead-zone predictive current control approach specifically designed for PMSM. This approach is based on a data-driven methodology, aiming to address the issue of parameter mismatching in dead-zone prediction current control and improve the performance of PMSM control systems. The paper [21] introduces a model-free predictive current control drive system that incorporates an extended Kalman filter to address performance degradation in model predictive control due to variations in motor parameters. The paper [22] introduces a model-free predictive current control strategy for the drive system of PMSMs in electric vehicles. This innovative approach effectively mitigates performance limitations caused by inaccurate inertia estimation through real-time dynamic adjustment of inertia parameters. Paper [23] presents a novel speed control strategy that combines an adaptive speed controller with a radial basis function neural network for precise speed regulation of PMSM. This approach effectively mitigates the impact of parameter uncertainties and load variations on system performance. The paper [24] utilizes a linear–nonlinear switching active disturbance rejection control strategy to design speed and current controllers for PMSM in servo systems, aiming to improve the disturbance rejection performance of PMSM speed and current controllers. Paper [25] presents an optimal tracking control strategy for PMSM systems characterized

by partially unknown dynamics, voltage saturation, and varying speed and current. By integrating an advanced feedforward control input, conventional velocity and current tracking challenges are redefined as novel optimal control problems within a cascaded framework. Experimental results demonstrate that both tracking and approximation errors are uniformly bounded.

In the vector control system of a three-phase PMSM, the conventional PI regulator is widely adopted as the speed controller due to its simplicity and robustness. However, the PMSM exhibits nonlinear dynamics and strong coupling among multiple variables. In the presence of external disturbances or variations in internal parameters, the traditional PI control method struggles to meet stringent control requirements. To improve the dynamic performance of the PMSM speed regulation system, it is crucial to implement a control strategy that remains insensitive to external disturbances and parameter changes while ensuring rapid response and high accuracy. Furthermore, achieving high-performance control for a PMSM requires precise rotor position and speed information within the magnetic field-oriented vector control framework. The use of mechanical sensors for this purpose, however, increases system cost, size, and weight, and imposes strict operating environment constraints. Sensorless control technology addresses these challenges by monitoring electrical signals within the motor windings and employing advanced algorithms to accurately estimate rotor position and speed, thereby enhancing the robustness and reliability of the PMSM vector control system. This paper introduces an energy-optimized speed control algorithm based on the TD3 for PMSM drive systems. Based on the traditional double closed-loop control structure for PMSM speed and current, the TD3 algorithm is employed to train an intelligent agent aimed at optimizing the inner-loop current controller. In designing the reward function for this intelligent agent, both the speed tracking error of the PMSM and the energy consumption of the controller are taken into account. The performance of this optimized controller is then compared with that of PI and LADRC controllers. The main contributions are summarized as follows.

- The TD3-based optimal control reduces the difficulty of designing a speed tracking controller for nonlinear PMSM.
- Adding energy consumption optimization control to the traditional control objective of steadiness, accuracy, and speed effectively improves the efficiency of the motor.
- The better generalization of the algorithm enables the motor to exhibit better control performance under different operating conditions.

The rest of this article is organized as follows. In Section 2, the paper details the modeling process of PMSM and clearly defines the control objectives. In Section 3, we introduce a PMSM speed tracking control algorithm utilizing TD3 and present an in-depth analysis of both the foundational environment and the implementation process of this algorithm. Section 4 analyzes and discusses the experimental results in detail. Finally, Section 5 provides a summary of the article.

2. Description of the Control Problem

2.1. Control Object Model

The motion equation of PMSM can be described as follows:

$$J\frac{d\omega_m}{dt} = T_e - T_L - B\omega_m \tag{1}$$

$$T_e = \frac{3}{2}p_n i_q \left[i_d (L_d - L_q) + \varphi_f \right] \tag{2}$$

Among them, ω_m, J, B, p_n, T_e and T_L represent the mechanical angular velocity, moment of inertia, damping coefficient, number of pole pairs, electromagnetic torque, and load torque of the motor, respectively. To facilitate controller design in a synchronous rotating coordinate system, $d - q$ for PMSM models are commonly established. The stator voltage equation can be expressed as Equation (3):

$$\begin{cases} u_d = Ri_d + \frac{d}{dt}\varphi_d - \omega_e\varphi_q \\ u_q = Ri_q + \frac{d}{dt}\varphi_q + \omega_e\varphi_d \end{cases} \tag{3}$$

The stator magnetic flux equation is given by Equation (4):

$$\begin{cases} \varphi_d = L_d i_d + \varphi_f \\ \varphi_q = L_q i_q \end{cases} \tag{4}$$

The substitution of Equation (4) into Equation (3) yields the stator voltage equation as shown (5):

$$\begin{cases} u_d = Ri_d + L_d\frac{d}{dt}i_d - \omega_e L_q i_q \\ u_q = Ri_q + L_q\frac{d}{dt}i_q + \omega_e\left(L_d i_d + \varphi_f\right) \end{cases} \tag{5}$$

Among them, u_d, u_q are the components of the stator voltage on the $d - q$ axis, while i_d, i_q are the components of the stator current on the $d - q$ axis; R represents the stator resistance. φ_d, φ_q are the X-axis components of the stator magnetic flux. ω_e represents the angular velocity of electricity, L_d, L_q are the inductance components of the $d - q$ axis, and φ_f represents the magnetic flux of a permanent magnet. In addition, it is important to pay attention to the relationship between variables in Equation (6) when constructing a PMSM simulation model.

$$\begin{cases} \omega_e = p_n\omega_m \\ N_r = \frac{30}{\pi}\omega_m \\ \theta_e = \int \omega_e dt \end{cases} \tag{6}$$

where ω_m represents the mechanical angular velocity of the motor (rad/s), and N_r represents the motor speed ($r/\min$). The common methods of traditional vector control include $i_d = 0$ control and maximum torque current ratio control. The control method where the dq-axis inductance of surface-mounted PMSM is equal, i.e., $L_d = L_q$ and $i_d = 0$, is mainly applicable to three-phase surface-mounted PMSMs. In this case, Equation (2) can be rewritten as Equation (7):

$$T_e = \frac{3}{2}p_n i_q \varphi_f \tag{7}$$

For surface-mounted three-phase PMSMs, $i_d = 0$ control and the maximum torque current ratio control are equivalent.

2.2. Speed Loop Control

Assuming the motor starts with no load, i.e., $T_L = 0$, when adopting control strategy $i_d = 0$, the active damping is defined as in Equation (8):

$$i_q = i'_q - B_a\omega_m \tag{8}$$

The combination of Equations (1) and (2) yields Equation (9), thereby establishing a comprehensive relationship.

$$\frac{d\omega_m}{dt} = \frac{1.5p_n\varphi_f}{J}\left(i'_q - B_a\omega_m\right) - \frac{B}{J}\omega_m \tag{9}$$

Repositioning the poles of Equation (9) at the desired closed-loop bandwidth β and applying a Laplace transform derives the transfer function representing motor speed in relation to shaft current as shown in Equation (10).

$$\omega_m(s) = \frac{1.5 p_n \varphi_f / J}{s + \beta} i'_q(s) \tag{10}$$

The β represents the desired bandwidth of the speed control loop. The coefficient B_a of active damping, as derived from the comparison between Equations (9) and (10), is presented in Equation (11).

$$B_a = \frac{\beta J - B}{1.5 p_n \varphi_f} \tag{11}$$

The speed loop controller can be mathematically represented as Equation (12) when employing a conventional PI controller.

$$i_q* = \left(K_p + \frac{K_i}{s}\right)(\omega_m* - \omega_m) - B_a \omega_m \tag{12}$$

where $K_p = \frac{\beta J}{1.5 p_n \varphi_f}$, $K_i = \beta K_p$.

2.3. Current Loop Control

The current equation for the $d - q$-axis, which corresponds to Equation (13), can be obtained by rewriting Equation (5).

$$\begin{cases} \frac{d}{dt} i_d = -\frac{R}{L_d} i_d + \frac{L_q}{L_d} \omega_e i_q + \frac{1}{L_d} u_d \\ \frac{d}{dt} i_q = -\frac{R}{L_d} i_d - \frac{1}{Lq} \omega_e \left(L_d i_d + \varphi_f\right) + \frac{1}{L_q} u_q \end{cases} \tag{13}$$

The complete decoupling of i_d and i_q yields Equation (14).

$$\begin{cases} u_{d0} = u_d + \omega_e L_q i_q \\ u_{q0} = u_q - \omega_e \left(L_d i_d + \varphi_f\right) \end{cases} \tag{14}$$

The substitution of Equation (5) into Equation (14) results in the derivation of Equation (15).

$$\begin{cases} u_{d0} = R i_d + L_d \frac{d}{dt} i_d \\ u_{q0} = R i_q + L_q \frac{d}{dt} i_q \end{cases} \tag{15}$$

where u_{d0}, u_{q0} are the d-axis and q-axis voltages after current decoupling, respectively. Using the conventional PI controller and combining with the feed-forward decoupling control strategy, the $d - q$-axis voltage can be obtained as shown in Equation (16).

$$\begin{cases} u_d* = \left(K_{pd} + \frac{K_{id}}{s}\right)(i_d* - i_d) - \omega_e L_q i_q \\ u_q* = \left(K_{pd} + \frac{K_{iq}}{s}\right)(i_q* - i_q) + \omega_e(L_d i_d + \psi_f) \end{cases} \tag{16}$$

where K_{pd}, K_{pq} is the proportional gain of the PI regulator and K_{id}, K_{iq} is the integral gain of the PI regulator.

The block diagram in Figure 1 illustrates the three-phase PMSM vector control employing the method with $i_d = 0$. It is evident from the figure that three-phase PMSM vector control primarily comprises three components, namely a speed loop controller, a current loop controller, and the space vector pulse width modulation (SVPWM) algorithm.
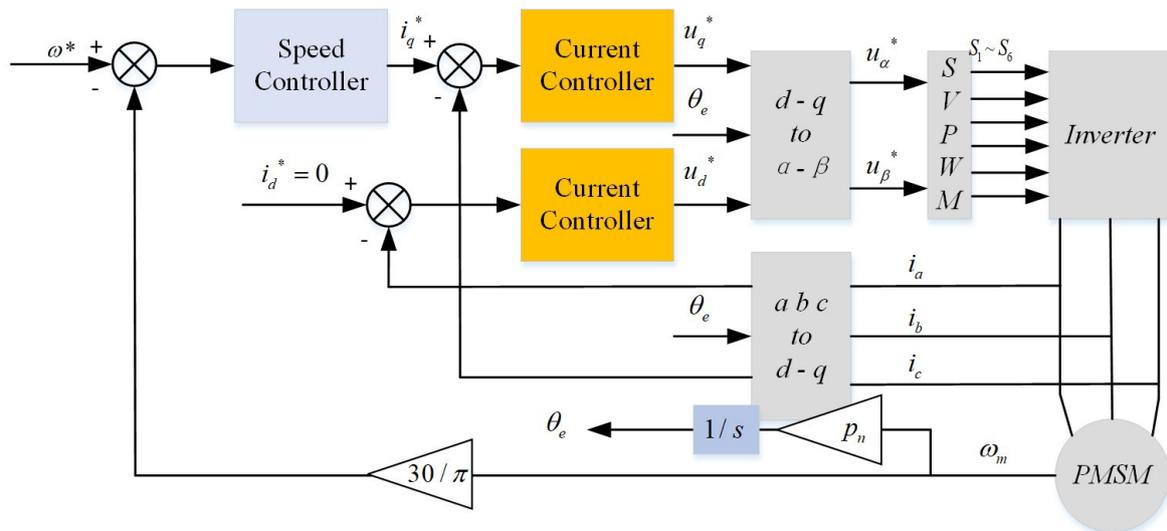
Figure 1. Three-phase PMSM vector control block diagram.

3. TD3 of PMSM

The TD3 algorithm proposes three key enhancement techniques based on the deep deterministic policy gradient (DDPG) algorithm, as delineated below [26]. (1) The double network refers to the utilization of two Critic networks, wherein the smaller one is employed for computing the target value in order to mitigate potential issues related to overestimation. (2) Target policy smoothing regularization incorporates perturbations to the action in the subsequent state while calculating the target value, aiming to enhance the precision of value evaluation. (3) Following multiple updates to the Critic network, subsequent updates are made to the Actor network to enhance training stability. Figure 2 shows the structure of the TD3 algorithm [27].
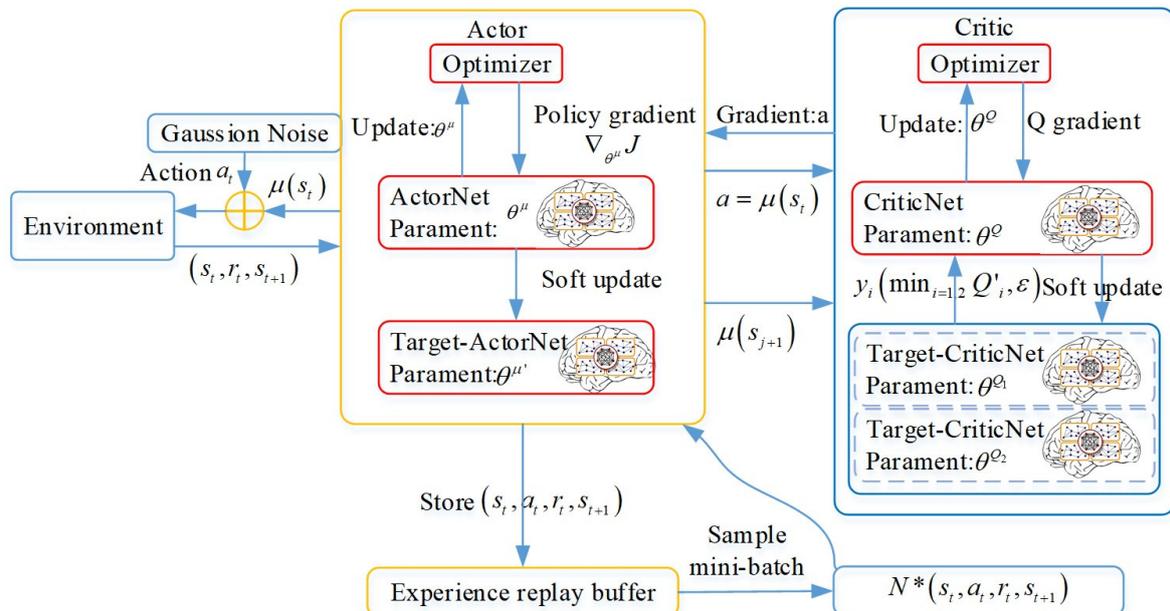
Figure 2. Structure of the TD3 algorithm.

The Actor network updates by maximizing the cumulative expected return through a deterministic policy gradient, while both Critic1 and Critic2 networks update by minimizing the error between the evaluation value and the target value using mean squared error. All target networks are updated using an exponential moving average (EMA) soft update

method. During the training phase, a batch of data are sampled from the Replay Buffer with a specific batch size. The Replay Buffer plays a pivotal role in reinforcement learning. It enhances the efficiency of sample utilization and accelerates training speed. Additionally, by mitigating data correlation, it improves both the stability and generalization capability of the model. In this study, the capacity of the Replay Buffer is configured to 10,000. Assuming one sample is denoted as $(s, a, r, s', done)$, the update process for all networks follows. The updating procedure for the Critic1 and Critic2 networks entails employing the Target Actor network to calculate the action $a' = \mu'\left(s' \middle| \theta^{\mu'}\right)$ in state s'. Subsequently, target policy smoothing regularization is implemented, and ε is incorporated into the target action a'. ε represents Gaussian noise, which adheres to a normal distribution. s' and a' denote the state and action at the subsequent time step, respectively. $\mu'(\cdot)$ signifies the target neural network function, while $\theta^{\mu'}$ represents the parameters of this target neural network.

$$a' = a' + \varepsilon$$
$$\varepsilon \sim clip(N(0, \sigma), -c, c) \tag{17}$$

Building upon the concept of dual networks, Equation (18) is employed for the computation of the target value y.

$$y = r + \gamma \min_{i=1,2} Q_i'\left(s', a' \middle| \theta_i^{Q'}\right) \tag{18}$$

The gradient descent algorithm is ultimately employed to minimize the error between the evaluation value and the target value, thereby facilitating parameter updates in both the Critic1 and Actor network update process.

$$y = r + \gamma \min_{i=1,2} Q_i'\left(s', a' \middle| \theta_i^{Q'}\right) \tag{19}$$

After updating the Critic1 and Critic2 networks for d steps, an update of the Actor network is initiated. The Actor network is employed to compute action a_{new} for state s, where $a_{new} = \mu(s|\theta^{\mu})$.

It is crucial to emphasize that there is no necessity to introduce noise after computing the action, as our objective is for the Actor network to update towards the direction of maximum value. Adding noise would be inconsequential in this context. Subsequently, we employ either Critic1 or Critic2 network to assess the state–action pair (s, a_{new}), assuming utilization of the Critic1 network $q_{new} = Q_1(s, a_{new}|\theta^{Q_1})$. The q_{new} optimization is ultimately achieved by employing a gradient ascent algorithm to facilitate updates in the Actor network.

The target network is updated using a soft update method, wherein a learning rate τ is introduced to calculate the weighted average between the old parameters of the target network and their corresponding new parameters. Subsequently, these averaged parameters are assigned to the target network (Equation (20)).

$$\theta^{Q_i'} = \tau\theta^{Q_i} + (1-\tau)\theta^{Q_i'}, (i = 1, 2)$$
$$\theta^{\mu'} = \tau\theta^{\mu} + (1-\tau)\theta^{\mu'} \tag{20}$$

where $\tau \in (0, 1)$, The value of t is usually set to 0.005.

In this paper, the traditional vector control double closed-loop system's current loop control algorithm is upgraded to the TD3 reinforcement learning algorithm, consequently updating Figures 1–3 as illustrated. The agent operates within the current inner loop of the PMSM, regulating the voltage inputs to its d-axis and q-axis. During training, the state quantities selected are the q-axis current error, d-axis current error, integral of the q-axis

current error, integral of the *d*-axis current error, motor speed ω_m, and motor reference speed ω_m*, with the execution action being the *q*-axis input voltage and *d*-axis input voltage. The reward at each time step is

$$r_t = (\alpha_1 * id_{err}^2 + \alpha_2 * iq_{err}^2 + \beta * \sum_j u^j_{t-1}{}^2). \tag{21}$$

Figure 3. The dual closed-loop control structure of PMSM speed and current based on TD3.

Here, α_1, α_2, β are the coefficients of each term, id_{err} is the d-axis current error, iq_{err} is the q-axis current error, and u^j_{t-1} are the actions from the previous time step.

The hyperparameter settings for the training process are shown in Table 1.

Table 1. The hyperparameter settings for the training process.

Parameters	Value
Sample Time	0.0001
Discount Factor	0.99
Experience Buffer Length	10000
Target Smoothing Factor	0.006
Target Update Frequency	10
Mini Batch Size	256

4. Results and Analysis

The parameters of the motor utilized in the simulation presented in this paper are detailed in Table 2. In addition, the PI, LADRC, and TD3 controllers proposed in this article can be implemented in MATLAB/Simulink to control the PMSM. The simulation was performed using MATLAB version R2024b on a personal computer equipped with a Windows 11 64-bit operating system, 16 GB of memory, and a CPU running at 3.7 GHz. The snapshot of the implemented software is shown in Figure 4, as follows.

In this article, simulation experiments and data analysis employ per-unit (*pu*) values. Per-unit values are a dimensionless indicator commonly utilized in power systems to represent the ratio between actual values and reference values. This enables relative comparisons among different systems, electrical quantities, or engineering parameters, facilitating quantitative analysis and research. The reference values used in this article can be found in Table 3.

Table 2. Parameters of PMSM.

Parameters	Symbol	Value
Rated current	I_N	7.2600
Rated torque	T_N	0.3471
Maximum speed	N_{max}	4300
Number of pole pairs	Np	7
Nominal phase resistance	R_s	0.2930
Nominal d-axis inductance	L_d	8.7678×10^{-5}
Nominal q-axis inductance	L_q	7.7724×10^{-5}
Nominal permanent flux	Φ_m	0.0046
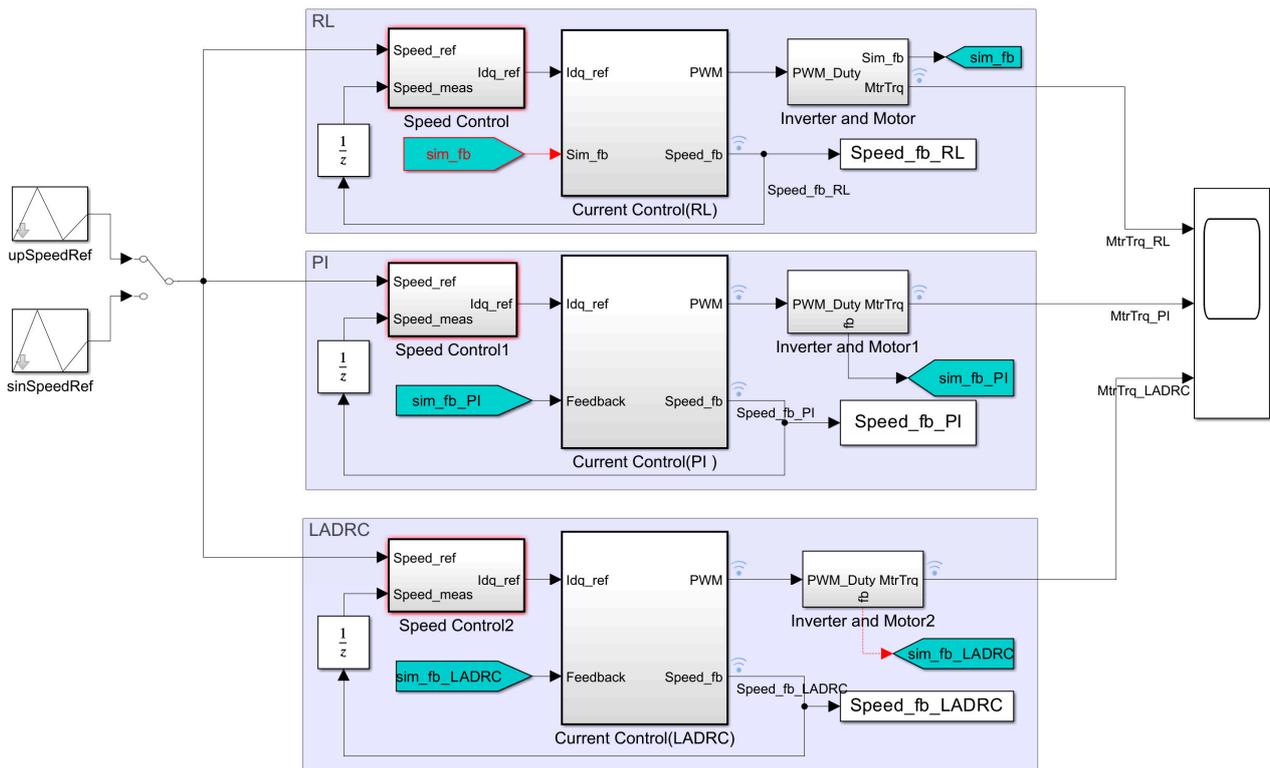
Figure 4. Snapshot of the implemented software.

Table 3. The reference values used in this article.

Name	Value
Speed base (RPM)	3476
Current base (A)	21.4286
Voltage base (V)	13.8564

To evaluate the effectiveness and robustness of the algorithm, simulation experiments were conducted under two representative PMSM operating conditions.

Working condition 1: The PMSM load is constant, the starting moment is given as 0.5 *pu*, and the reference speed trajectory is a sinusoidal signal, $Speed_{ref} = 0.3 * \sin(t) + 0.5$; there is a no-load startup. The experimental results are shown in Figures 5 and 6. In Figure 5, Episode Q0 refers to the value of the initial state estimated by the critic (i.e., Q(S0)). The same applies to Figure 7 below.
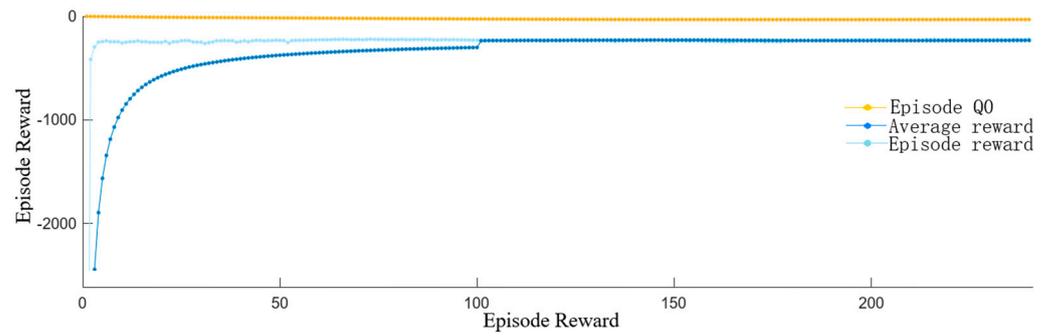
Figure 5. Training results for working condition 1.
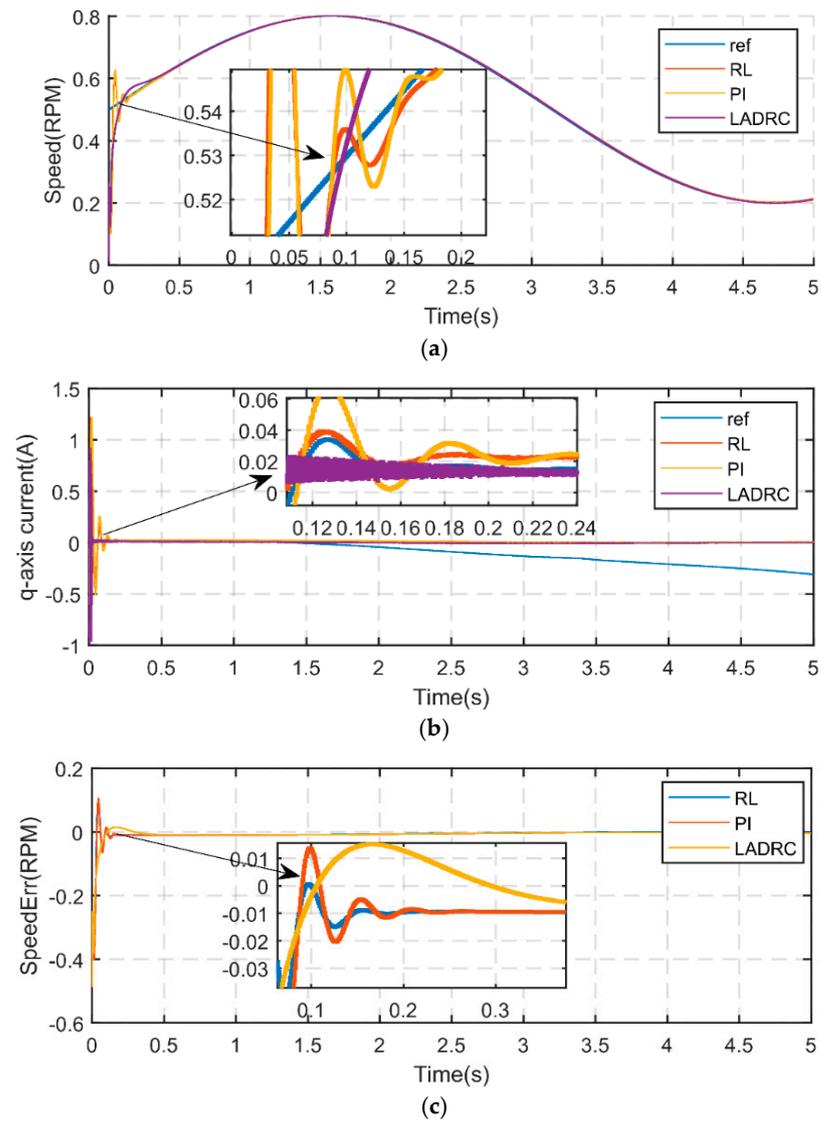
Figure 6. Experimental results of a PMSM operating in working condition 1. (**a**) Rotor speed; (**b**) Q-axis current; (**c**) speed tracking error.
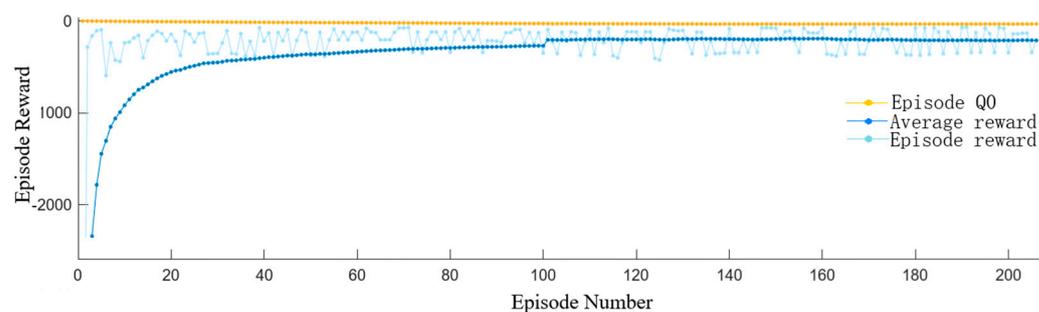
Figure 7. Training results for working condition 2.

As can be seen from Figure 5, the agent's Episode reward value converges around −310 after approximately 100 episodes. An episode consists of a sequence of time steps. During each time step, the agent performs an action, receives a reward from the environment, and transitions to a new state. This sequence of interactions forms a complete episode. The duration of different episodes can vary significantly.

From Figure 6a, it is evident that the TD3 algorithm of this design achieves faster tracking of the upper reference trajectory speed with minimal overshoot compared to the other controllers. The traditional PI controller exhibits significant overshoot, whereas the LADRC controller requires a longer time to reach a steady state. From Figure 6b, it can be observed that the TD3-based controller stabilizes the q-axis current closer to the reference value more rapidly and with less fluctuation. In contrast, the PI controller shows substantial overshoot, and the LADRC controller experiences greater fluctuations. From Figure 6c, it is clear that the TD3 algorithm, as a reinforcement learning approach, demonstrates the smallest speed tracking error and superior stability.

Working condition 2: The PMSM speed is suddenly increased, the load is constant, the given speed is 0.5 *pu* at the starting moment, and it is suddenly changed to 0.8 *pu* at 1 *s*; there is a no-load startup. The experimental results are shown in Figures 7 and 8.

As can be seen from Figure 7, the agent's Episode reward value converges around −230 after approximately 100 episodes.

From Figure 8a, it is evident that the controller using the LADRC algorithm exhibits the smallest overshoot but requires the longest time to reach a steady state. In contrast, the RL algorithm achieves the fastest return to a steady state with significantly less overshoot compared to the PI algorithm. From Figure 8b, it can be observed that the torque fluctuation caused by a sudden change in rotational speed is also minimized. Furthermore, as shown in Figure 8c, the proposed algorithm demonstrates the smallest overall rotational speed tracking error. Table 4 summarizes the performance comparison of the three algorithms over a 1 s period.

Table 4. Performance comparison under different control algorithms.

Performance Parameters	PI	LADRC	RL
Settlingtime (s)	0.25	0.50	0.18
Risetime (s)	0.03	0.15	0.04
Undershoot (%)	11.25	2.38	5.01
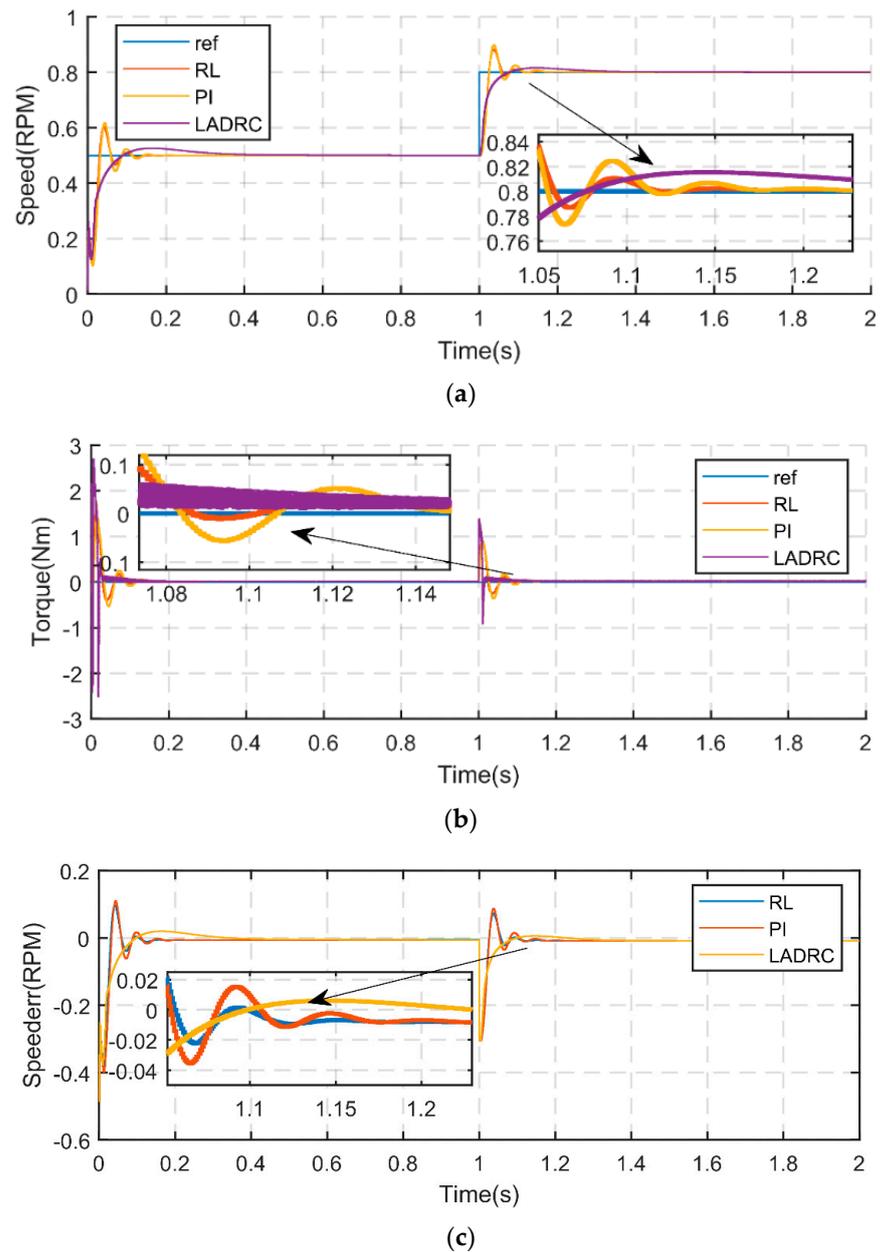
Figure 8. Experimental results of PMSM operating in working condition 2. (**a**) Rotor speed; (**b**) Q-axis current; (**c**) speed tracking error.

From Table 3, it is evident that the RL algorithm demonstrates superior adaptability to sudden speed changes. It can rapidly return to a steady state, with the overshoot remaining within an acceptable range.

To verify the stability of the proposed algorithm, torque disturbances with abrupt changes were introduced under two typical operating conditions. In condition 1, a torque load of 1 *N* was applied at 2 s, while in condition 2, the same load was applied at 1.5 s. The experimental results are presented in Figure 9. Under both conditions, the proposed method exhibited robust anti-interference performance.
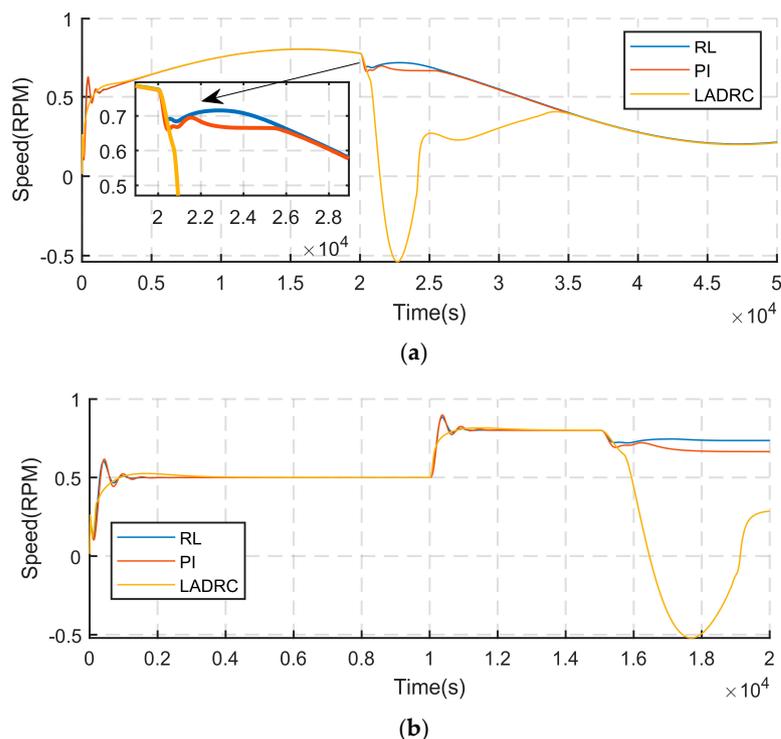
Figure 9. Experimental results of a PMSM operating in working condition 2 with torque disturbances. (**a**) Working condition 1; (**b**) working condition 2.

5. Conclusions

The PMSM drive system is a typical dual-timescale system characterized by control challenges such as parameter variations, external disturbances, and nonlinearities. In this paper, we propose a TD3-based optimal control method aimed at minimizing the energy consumption of PMSM drive systems. We conducted simulation experiments under two typical operating conditions to validate the effectiveness and robustness of the proposed algorithm. At the same time, when designing the algorithm, the energy loss of the controller was taken into account, that is, the control quantity was added to the reward function of the agent. The experimental results demonstrate that the TD3 algorithm surpasses both the traditional PI controller and the linear active disturbance rejection control (LADRC) algorithm in terms of reference trajectory tracking accuracy, *q*-axis current regulation, and speed tracking error minimization. Furthermore, when the PMSM was subjected to torque mutations, the algorithm proposed in this paper exhibited superior control performance. This study offers a novel approach to PMSM drive system control, significantly enhancing motor efficiency and system robustness through the integration of deep reinforcement learning.

Author Contributions: Conceptualization, Y.Z.; methodology, and writing—original draft preparation, Z.H.; software, M.L.; validation, Y.L. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

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

References

1. Lin, X.; Xu, R.; Yao, W.; Gao, Y.; Sun, G.; Liu, J. Observer-Based Prescribed Performance Speed Control for PMSMs: A Data-Driven RBF Neural Network Approach. *IEEE Trans. Ind. Inform.* **2024**, *20*, 7502–7512. [CrossRef]
2. Lin, X.; Wu, C.; Yao, W.; Liu, Z.; Shen, X.; Xu, R. Observer-Based Fixed-Time Control for Permanent-Magnet Synchronous Motors with Parameter Uncertainties. *IEEE Trans. Power Electron.* **2023**, *38*, 4335–4344. [CrossRef]
3. Krishnan, R. *Electric Motor Drives: Modeling, Analysis, and Control*; Prentice-Hall: Upper Saddle River, NJ, USA, 2001.
4. Zhang, Z.; Liu, Y.; Liang, X.; Guo, H.; Zhuang, X. Robust Model Predictive Current Control of PMSM Based on Nonlinear Extended State Observer. *IEEE J. Emerg. Sel. Top. Power Electron.* **2023**, *11*, 862–873. [CrossRef]
5. Yang, C.; Meng, F.; Zhang, H.; Zhao, J.; Wang, H.; Zhou, L. Optimal Coordinated Control for Speed Tracking and Torque Synchronization of Rigidly Connected Dual-Motor Systems. *IEEE/ASME Trans. Mechatron.* **2023**, *28*, 2609–2620. [CrossRef]
6. Li, X.; Tian, W.; Gao, X. A Generalized Observer-Based Robust Predictive Current Control Strategy for PMSM Drive System. *IEEE Trans. Ind. Electron.* **2022**, *69*, 1322–1332. [CrossRef]
7. Wang, Y.; Fang, S.; Hu, J.; Huang, D. A Novel Active Disturbance Rejection Control of PMSM Based on Deep Reinforcement Learning for More Electric Aircraft. *IEEE Trans. Energy Convers.* **2023**, *38*, 1461–1470. [CrossRef]
8. Wang, Y.; Fang, S.; Hu, J. Active Disturbance Rejection Control Based on Deep Reinforcement Learning of PMSM for More Electric Aircraft. *IEEE Trans. Power Electron.* **2023**, *38*, 406–416. [CrossRef]
9. Jiang, X.; Yang, Y.; Fan, M.; Ji, A.; Xiao, Y.; Zhang, X. An Improved Implicit Model Predictive Current Control with Continuous Control Set for PMSM Drives. *IEEE Trans. Transp. Electrif.* **2022**, *8*, 2444–2455. [CrossRef]
10. Xu, B.; Jiang, Q.; Ji, W.; Ding, S. An Improved Three-Vector-Based Model Predictive Current Control Method for Surface-Mounted PMSM Drives. *IEEE Trans. Transp. Electrif.* **2022**, *8*, 4418–4430. [CrossRef]
11. Wang, X.; Wang, Z.; Gu, M.; Xu, Z.; Zou, Z.; Wang, W. Fault-Tolerant Control of Common Electrical Faults in Dual Three-Phase PMSM Drives Fed by T-Type Three-Level Inverters. *IEEE Trans. Ind. Appl.* **2021**, *57*, 481–491. [CrossRef]
12. Sun, Z.; Deng, Y.; Wang, J.; Yang, T.; Wei, Z.; Cao, H. Finite Control Set Model-Free Predictive Current Control of PMSM with Two Voltage Vectors Based on Ultralocal Model. *IEEE Trans. Power Electron.* **2023**, *38*, 776–788. [CrossRef]
13. Ma, Y.; Li, D.; Li, Y.; Yang, L. A Novel Discrete Compound Integral Terminal Sliding Mode Control With Disturbance Compensation for PMSM Speed System. *IEEE/ASME Trans. Mechatron.* **2022**, *27*, 549–560. [CrossRef]
14. Zhang, J.; Ren, W.; Sun, X. Current-Constrained Adaptive Robust Control for Uncertain PMSM Drive Systems: Theory and Experimentation. *IEEE Trans. Transp. Electrif.* **2023**, *9*, 4158–4169. [CrossRef]
15. Tan, L.; Cong, T.; Cong, D. Neural Network Observers and Sensorless Robust Optimal Control for Partially Unknown PMSM with Disturbances and Saturating Voltages. *IEEE Trans. Power Electron.* **2021**, *36*, 12045–12056. [CrossRef]
16. Li, Z.; Wang, F.; Ke, D.; Li, J.; Zhang, W. Robust Continuous Model Predictive Speed and Current Control for PMSM with Adaptive Integral Sliding-Mode Approach. *IEEE Trans. Power Electron.* **2021**, *36*, 14398–14408. [CrossRef]
17. Zhao, J.; Yang, C.; Gao, W.; Zhou, L. Reinforcement Learning and Optimal Control of PMSM Speed Servo System. *IEEE Trans. Ind. Electron.* **2023**, *70*, 8305–8313. [CrossRef]
18. Attestog, S.; Senanayaka, J.; Khang, H.; Robbersmyr, K. Robust Active Learning Multiple Fault Diagnosis of PMSM Drives with Sensorless Control Under Dynamic Operations and Imbalanced Datasets. *IEEE Trans. Ind. Inform.* **2023**, *19*, 9291–9301. [CrossRef]
19. Wang, Y.; Fang, S.; Hu, J.; Huang, D. Multiscenarios Parameter Optimization Method for Active Disturbance Rejection Control of PMSM Based on Deep Reinforcement Learning. *IEEE Trans. Ind. Electron.* **2023**, *70*, 10957–10968. [CrossRef]
20. Wang, Y.; Fang, S.; Huang, D. An Improved Model-Free Active Disturbance Rejection Deadbeat Predictive Current Control Method of PMSM Based on Data-Driven. *IEEE Trans. Power Electron.* **2023**, *38*, 9606–9616. [CrossRef]
21. Luo, L.; Huang, W.; Huang, M.; Fan, Q. Model-Free Predictive Current Control of Sensorless PMSM Drives with Extended Kalman Filter. In Proceedings of the 2023 26th International Conference on Electrical Machines and Systems (ICEMS), Zhuhai, China, 5–8 November 2023; pp. 2369–2374.
22. Wei, Y.; Men, S.; Wei, Y.; Qi, H.; Wang, F. A Model-Free Predictive Current Control for PMSM Driving System of EV with Adjustable Low Inertia. In Proceedings of the 2022 IEEE Transportation Electrification Conference and Expo, Asia-Pacific (ITEC Asia-Pacific), Haining, China, 28–31 October 2022; pp. 1–7.
23. Jie, H.; Zheng, G.; Zou, J.; Xin, X.; Guo, L. Speed Regulation Based on Adaptive Control and RBFNN for PMSM Considering Parametric Uncertainty and Load Fluctuation. *IEEE Access* **2020**, *8*, 190147–190159. [CrossRef]
24. Lin, P.; Wu, Z.; Liu, K.; Sun, X. A Class of Linear–Nonlinear Switching Active Disturbance Rejection Speed and Current Controllers for PMSM. *IEEE Trans. Power Electron.* **2021**, *36*, 14366–14382. [CrossRef]
25. Tan, L.; Pham, T. Optimal Tracking Control for PMSM With Partially Unknown Dynamics, Saturation Voltages, Torque, and Voltage Disturbances. *IEEE Trans. Ind. Electron.* **2022**, *69*, 3481–3491. [CrossRef]

26.   Saglam, B.; Mutlu, F.; Cicek, D.; Kozat, S. Parameter-Free Reduction of the Estimation Bias in Deep Reinforcement Learning for Deterministic Policy Gradients. *Neural. Process. Lett.* **2024**, *56*, 80. [CrossRef]

27.   Yin, F.; Yuan, X.; Ma, Z.; Xu, X. Vector Control of PMSM Using TD3 Reinforcement Learning Algorithm. *Algorithms* **2023**, *16*, 404. [CrossRef]