mathematics

MDPI

Article

A Robust Optimal Control Strategy for PMSM Based on VGPDO and Actor-Critic Neural Network Against Flux Weakening and Mismatched Load Torque

Yangyu Niu and Haibin Shi *

College of Information Science and Engineering, Northeastern University, Shenyang 110819, China; 2400881@stu.neu.edu.cn

* Correspondence: shihaibin@ise.neu.edu.cn

Abstract

In this paper, a novel robust optimal control strategy is proposed for permanent magnet synchronous motors (PMSMs), simultaneously addressing two critical challenges in speed regulation: flux linkage degradation during long-term operation and abrupt load torque variations. The robust optimal control strategy is implemented through a combination of feedforward control and feedback control. A novel Variable-Gain Proportional Disturbance Observer (VGPDO) is proposed to simultaneously estimate time-varying flux linkage and torque disturbances in PMSM systems. The estimated disturbances are then compensated via a feedforward control loop, significantly improving the system's robustness against parameter variations and external load changes. An optimal controller based on an actor-critic neural network provides feedback for optimal control performance. The uniform ultimate boundedness (UUB) of the proposed strategy is proved through Lyapunov stability analysis, and comprehensive simulation studies demonstrate the efficacy of both the proposed VGPDO and the proposed robust optimal control strategy.

Keywords: PMSM; variable-gain proportional disturbance observer; flux linkage degradation; actor-critic neural network

MSC: 93B51; 93C40; 93B70; 93B53; 93C95

check for updates

Academic Editor: Cheng-Hung Huang

Received: 20 August 2025
Revised: 8 October 2025
Accepted: 13 October 2025
Published: 24 October 2025

Citation: Niu, Y.; Shi, H. A Robust Optimal Control Strategy for PMSM Based on VGPDO and Actor-Critic Neural Network Against Flux Weakening and Mismatched Load Torque. *Mathematics* **2025**, *13*, 3387. https://doi.org/10.3390/math13213387

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Owing to their superior power density, wide speed regulation range, compact structure, and high reliability, Permanent magnet synchronous motors (PMSMs) have been widely adopted in high-tech industries such as aerospace and new energy vehicles. The control strategies for PMSM have evolved significantly with the continuous development of control theory. With the advancement of modern industry, there is an increasing demand for enhanced control performance, achieving both robustness and optimality in PMSM speed control, which plays a crucial role in industrial production.

An effective PMSM control strategy should typically achieve optimality to minimize energy consumption while maintaining robustness to ensure long-term stable and efficient operation of the equipment. For PMSM speed regulation, conventional control strategies such as PID, sliding mode control, $H\infty$ control, model reference adaptive control, and backstepping control [1–9] have been widely adopted. However, these methods fail to achieve process optimality in terms of energy efficiency and dynamic performance. To achieve optimal control for the nonlinear PMSM system, it is necessary to formulate the

Hamilton–Jacobi–Bellman (HJB) equation based on the cost function. However, these equations typically lack analytical solutions and pose significant challenges for numerical computation.

With the advancement of computational capabilities, reinforcement learning (RL)—also known as adaptive dynamic programming (ADP)—has emerged as a powerful tool for approximating solutions to the HJB equation, thereby enabling optimal control in practical applications [10–16]. Recent advances in PMSM optimal control based on ADP have demonstrated the effectiveness in [17–21], which utilize two neural networks to approximate the value function (Critic) and optimal control policy function (Actor), respectively, thereby circumventing the numerical challenges of directly solving HJB equations. Theoretical analysis proves that gradient-based weight updating ensures UUB stability within a neighborhood of ideal parameters [22].

Standard ADP methods struggle with HJB estimation during sudden parameter changes. Real-time identification and compensation of PMSM internal uncertainties and external disturbances constitute an effective solution. Prior works extensively address mismatched load torque effects: ref. [18] developed a serial extended state observer (ESO) for load torque observation; ref. [19] constructed a disturbance observer to estimate the load torque; ref. [20] proposed a load torque estimator that incorporates actuator error compensation to guarantee system stability. However, to ensure long-term operational stability of permanent magnet synchronous motors, it is essential to account for flux linkage deterioration effects, which necessitates real-time, accurate observation of motor flux linkage parameters. Numerous studies have developed dedicated flux linkage observers [23–25] and load torque observers [17,18,26,27] for PMSM. However, research on designing integrated observers capable of simultaneously estimating both flux linkage and load torque to ensure long-term stable operation in PMSM speed control systems remains scarce.

To solve the above problem, a variable-gain proportional disturbance observer (VGPDO), which simultaneously estimates both the flux linkage deterioration and external load torque disturbances, is proposed in this paper. Meanwhile, a robust optimal speed control strategy is proposed for PMSM. The strategy deploys a feedforward controller in the PMSM control system, which utilizes the flux linkage and torque estimates obtained from VGPDO to compensate for disturbances in real time, thereby transforming the PMSM dynamic system into an error dynamic system with disturbance rejection robustness. Furthermore, an actor-critic network-based optimal controller is implemented in this error dynamic system to stabilize the system and achieve optimal control.

The remainder of this paper is organized as follows: Section 2 derives the PMSM error dynamic equations with feedforward compensation for flux linkage and load torque variations. Section 3 develops the VGPDO to estimate both flux linkage and load torque in PMSM. Section 4 presents the actor-critic-based optimal controller design. Section 5 performs simulation studies with a comprehensive analysis. Finally, Section 6 concludes the paper by summarizing key contributions.

The main contributions of this paper are summarized as follows:

- To mitigate flux linkage degradation in PMSM control systems during prolonged load operation, a VGPDO is proposed to simultaneously estimate torque and flux linkage disturbances, thereby enhancing system robustness.
- To achieve optimality and robustness in PMSM speed regulation under varying load torque and flux linkage conditions, a robust optimal control strategy is proposed, which integrates feedforward-based VGPDO compensation with an actor-critic neural network-based optimal speed controller.
- Comprehensive simulations are conducted to validate the effectiveness of the VGPDO and robust optimal control strategy.

2. System Descriptions

As proposed in [19], the PMSM dynamics can be represented as:

$$\begin{cases} \dot{\omega} = \frac{n_p\psi_f}{J} i_q - \frac{B}{J}\omega - \frac{1}{J}T_L \\ \dot{i}_q = -\frac{R_s}{L} i_q - n_p\omega i_d - \frac{n_p\psi_f}{L}\omega + \frac{1}{L}u_q \\ \dot{i}_d = -\frac{R_s}{L} i_d + n_p\omega i_q + \frac{1}{L}u_d \end{cases} \tag{1}$$

where w, J, T_L and B are the angular velocity, inertia, load torque and viscous friction coefficient of PMSM, L, R_s and n_p are the stator inductance, stator resistance, and number of pole pairs, respectively. i_q and u_q are the q-axis stator current and voltage. i_d and u_d are the d-axis stator current and voltage. ψ_f is the rotor flux linkage. The PMSM states w, i_q, and i_d are assumed to be measurable.

When PMSM operates over an extended period, demagnetization effects must be considered. Let $\Delta\psi$ denote the variation in magnetic flux and let ψ_s represent the initial magnetic flux, hence we have:

$$\psi_f = \psi_s + \Delta\psi \tag{2}$$

For convenience in the following analysis, we take $\frac{n_p\psi_f}{J}$ as m_1, $\frac{B}{J}$ as m_2, $\frac{1}{J}$ as m_3, $\frac{R_s}{L}$ as m_4, $\frac{n_p\psi_f}{L}$ as m_5, n_p as m_6 and $\frac{1}{L}$ as m_7. We also define:

$$\begin{cases} m_1^* \triangleq \frac{n_p\psi_s}{J}; m_5^* \triangleq \frac{n_p\psi_s}{L} \\ \hat{m}_1 \triangleq \frac{n_p(\psi_s+\Delta\hat{\psi})}{J}; \hat{m}_5 \triangleq \frac{n_p(\psi_s+\Delta\hat{\psi})}{L} \\ \tilde{m}_1 \triangleq m_1 - \hat{m}_1; \tilde{m}_5 \triangleq m_5 - \hat{m}_5 \end{cases} \tag{3}$$

In order to achieve our tracking goals, we introduce:

$$\begin{cases} \tilde{w} \triangleq w - w_d \\ \tilde{i}_q \triangleq i_q - \hat{i}_q^* \end{cases} \tag{4}$$

where w_d represents the reference speed, $\hat{i}_q^*$ represents the virtual control input and is given by:

$$\hat{i}_q^* = \frac{m_2\omega_d + m_3\hat{T}_L + \dot{\omega}_d}{\hat{m}_1} \tag{5}$$

To facilitate the practical implementation of the actor-critic neural network-based optimal speed controller, the system described above is restructured via feedforward compensation. Hence, the control inputs u_d and u_q are decoupled into:

$$\begin{cases} u_q = \hat{u}_{cq} + u_{sq} \\ u_d = \hat{u}_{cd} + u_{sd} \end{cases} \tag{6}$$

where u_{sq} and u_{sd} denote the stabilizing terms generated by the actor-critic neural network-based optimal speed controller, while $\hat{u}_{cd}$ and $\hat{u}_{cq}$ represent terms which are defined as:

$$\begin{cases} \hat{u}_{cq} = \frac{1}{m_7}(m_4\hat{i}_q^* + \hat{m}_5\omega_d + m_6\omega_d i_d + \dot{\hat{i}}_q^*) \\ \hat{u}_{cd} = -\frac{1}{m_7}(m_6\omega_d\tilde{i}_q + m_6\omega\hat{i}_q^*) \end{cases} \tag{7}$$

By substituting (2)–(7) into Model (1), a new error dynamics model can be derived:

$$\begin{cases} \dot{\tilde{w}} = m_1\tilde{i}_q - m_2\tilde{\omega} - m_3\tilde{T}_L + \frac{\tilde{m}_1}{\hat{m}_1}\left(m_2\omega_d + m_3\hat{T}_L + \dot{\omega}_d\right) \\ \dot{\tilde{i}}_q = -m_5\tilde{w} - k_4\tilde{i}_q + m_7u_{sq} - m_6\tilde{\omega}i_d - \tilde{m}_5\omega_d \\ \dot{i}_d = -m_4i_d + m_7u_{sd} + m_6\tilde{\omega}\tilde{i}_q \end{cases} \tag{8}$$

To facilitate further analysis, the Model (8) can be rewritten in the following form:

$$\dot{x} = f(x) + gu + \delta \tag{9}$$

where:

$$\begin{cases} x \triangleq \begin{bmatrix} \tilde{\omega} & \tilde{i}_q & i_d \end{bmatrix}^T, u \triangleq \begin{bmatrix} u_{sq} & u_{sd} \end{bmatrix}^T; \\ f(x) \triangleq \begin{bmatrix} -m_2\tilde{\omega} + m_1\tilde{i}_q \\ -m_5\tilde{\omega} - m_4\tilde{i}_q - m_6\tilde{\omega}i_d \\ -m_4i_d + m_6\tilde{\omega}\tilde{i}_q \end{bmatrix}, g \triangleq \begin{bmatrix} 0 & 0 \\ m_7 & 0 \\ 0 & m_7 \end{bmatrix}; \\ \delta \triangleq \begin{bmatrix} -m_3\tilde{T}_L + \frac{\tilde{m}_1}{\hat{m}_1}\left(m_2\omega_d + m_3\hat{T}_L + \dot{\omega}_d\right) & -\tilde{m}_5\omega_d & 0 \end{bmatrix}^T \end{cases} \tag{10}$$

3. Variable-Gain Proportional Disturbance Observer Design

To ensure long-term stable operation of PMSM, an advanced observer capable of capturing both external disturbances and internal parameter variations is essential in a robust optimal control strategy. In this paper, we explicitly address abrupt load torque variations and flux linkage degradation, which are two primary disturbances occurring during prolonged PMSM operation, to enhance system robustness.

Considering the variation in magnetic flux, Model (1) can be reformulated by substituting (2) and (3), yielding:

$$\begin{cases} \dot{\omega} = m_1^*i_q - m_2\omega - m_3T_L + n_1i_q\Delta\psi \\ \dot{i}_q = -m_4i_q - m_6\omega i_d - m_5^*\omega + m_7u_q + n_2\omega\Delta\psi \\ \dot{i}_d = -m_4i_d + m_6\omega i_q + m_7u_d \end{cases} \tag{11}$$

where $n_1 \triangleq \frac{n_p}{J}, n_2 \triangleq -\frac{n_p}{L}$.

For simplicity, the above model (11) can be further expressed as:

$$\dot{z} = Az(t) + h(z) + Bu + E(z)d \tag{12}$$

with:

$$\begin{cases} z \triangleq \begin{bmatrix} z_1 & z_2 & z_3 \end{bmatrix}^T \triangleq \begin{bmatrix} w & i_q & i_d \end{bmatrix}^T; \\ d \triangleq \begin{bmatrix} \Delta\psi & T_L \end{bmatrix}^T; \\ A \triangleq \begin{bmatrix} -m_2 & m_1 & 0 \\ -m_5 & -m_4 & 0 \\ 0 & 0 & -m_4 \end{bmatrix}, B \triangleq \begin{bmatrix} 0 & 0 \\ m_7 & 0 \\ 0 & m_7 \end{bmatrix}; \\ h(z) \triangleq \begin{bmatrix} 0 \\ -m_6z_1z_3 \\ m_6z_1z_2 \end{bmatrix}, E(z) \triangleq \begin{bmatrix} n_1z_2 & -m_3 \\ n_2z_1 & 0 \\ 0 & 0 \end{bmatrix}. \end{cases} \tag{13}$$

Based on model (12), a disturbance observer is proposed as follows:

$$\begin{cases} \hat{d} = s + p \\ \dot{p} = l(z)\dot{z} \\ \dot{s} = -l(z)E(z)s - l(z)[Az + h(z) + Bu + E(z)p] \end{cases} \tag{14}$$

where $l(z)$ is an undetermined matrix function expressed as: $l(x) = \begin{bmatrix} l_{11} & l_{12} & l_{13} \\ l_{21} & l_{22} & l_{23} \end{bmatrix}$.

Defining the estimation error as $\tilde{d} = d - \hat{d}$, and combining (12) and (14), we obtain:

$$\begin{aligned} \dot{\tilde{d}} &= -\dot{\hat{d}} + \dot{d} = l(z)E(z)s + l(z)[Az + h(z) + Bu + E(z)p] - l(z)\dot{z} + \dot{d} \\ &= l(z)E(z)s + l(z)[Az + h(z) + Bu + E(z)p] - l(z)[Az + h(z) + Bu + E(z)d] + \dot{d} \\ &= l(z)E(z)s + l(z)[E(z)p - E(z)d] + \dot{d} \\ &= -l(z)E(z)(d - s - p) + \dot{d} \\ &= -l(z)E(z)\left(d - \hat{d}\right) + \dot{d} \\ &= -l(z)E(z)\tilde{d} + \dot{d} \end{aligned} \tag{15}$$

To guarantee the bounded convergence of $\tilde{d}$, the following condition must be satisfied:

$$l(x)E(x) = \begin{bmatrix} l_{11} & l_{12} & l_{13} \\ l_{21} & l_{22} & l_{23} \end{bmatrix} \begin{bmatrix} n_1 z_2 & -m_3 \\ n_2 z_1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} n_1 z_2 l_{11} + n_2 z_1 l_{12} & -m_3 l_{11} \\ n_1 z_2 l_{21} + n_2 z_1 l_{22} & -m_3 l_{21} \end{bmatrix} > 0 \tag{16}$$

Considering the safety margins α_1 and α_2, we set:

$$\begin{bmatrix} n_1 z_2 l_{11} + n_2 z_1 l_{12} & -m_3 l_{11} \\ n_1 z_2 l_{21} + n_2 z_1 l_{22} & -m_3 l_{21} \end{bmatrix} \geq \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix} > 0 \tag{17}$$

A feasible solution is given as follows:

$$\begin{cases} l_{11} = l_{13} = l_{23} = 0 \\ l_{12} = \frac{\alpha_1}{n_2 z_1} \\ l_{21} = \frac{\alpha_2}{-m_3} \\ l_{22} = \frac{n_1 z_2}{n_2 z_1} \frac{\alpha_2}{m_3} \end{cases} \tag{18}$$

Assumption 1.

1. *The load torque of PMSM and its rate are bounded by:* $\begin{cases} \| T_L \| \leqslant \gamma_1 \\ \| \dot{T}_L \| \leqslant \Upsilon_1 \end{cases}$.
2. *The flux linkage of PMSM variation rate is bounded by:* $\| \Delta\dot{\psi} \| \leqslant \Upsilon_2$.
3. *The reference speed and its derivative for PMSM are bounded as follows:* $\begin{cases} \| w_d \| \leqslant \gamma_3 \\ \| \dot{w}_d \| \leqslant \Upsilon_3 \end{cases}$.

Remark 1. *The above assumptions are practically reasonable for: (1). In real-world mechanical systems, the load torque is always finite due to physical limitations and typically varies smoothly due to inertia and damping effects. (2). The flux linkage variation rate $\Delta\dot{\psi}$ is naturally constrained by the physical constraints of magnetic domain reorientation and limited energy available for instantaneous demagnetization. (3). Reference speed w_d and acceleration $\dot{w}_d$ are bounded by both the motor's rated specifications and practical motion control needs.*

Theorem 1. *When applying the VGPDO (14) to Model (12) for flux linkage deterioration and disturbance torque estimation, the estimation error is bounded as follows under Assumption 1:*

$$\begin{cases} \| \tilde{T}_L \| \leqslant \frac{\Upsilon_1}{\sqrt{1+\alpha_1^2}} \\ \| \Delta\tilde{\psi} \| \leqslant \frac{\Upsilon_2}{\sqrt{1+\alpha_2^2}} \end{cases} \tag{19}$$

Notably, the VGPDO estimation error exhibits asymptotic stability under constant load torque and flux linkage conditions.

Furthermore, the parameter δ in (10) remains bounded (the detailed formulation is given in (A8)).

Proof of Theorem 1. The complete mathematical derivations are provided in Appendix A.1. □

4. Actor-Critic Network-Based Optimal Controller Design

The optimal controller u is designed to guarantee speed tracking while minimizing the following cost function:

$$V(x_0) = \int_0^\infty r(x(\tau), u(\tau))\mathrm{d}\tau \tag{20}$$

where $r(x,u) = Q(x) + u^T R u$ and $Q(x)$ is a positive definite function, R is a symmetric positive definite constant matrix, and x_0 denotes the initial state of x.

According to [19], the minimization of this cost function can be achieved through the Hamiltonian function. The corresponding Hamiltonian for the cost function $V(x)$ yeilds:

$$H(x,u,V_x) = r(x,u) + V_x^T[f(x) + gu + \delta] \tag{21}$$

where V_x is defined as: $V_x \triangleq \frac{dV}{dx}$.

Define the optimal cost function as:

$$V^*(x_0) = \min_u \left(\int_0^\infty r(x(\tau), u(x(\tau)))\mathrm{d}\tau \right) \tag{22}$$

According to [28], the optimal controller u and optimal cost function V_x^* satisfy:

$$0 = Q(x) + u^T R u + V_x^{*T}(f(x) + gu + \delta) \tag{23}$$

The optimal control input u can be solved as:

$$u = \arg\min_\mu [H(x,\mu,V_x)] = -\frac{1}{2} R^{-1} g^T V_x^*. \tag{24}$$

Based on [28], the cost function $V(x)$ can be expanded as an infinite series of basis functions ζ_i multiplied by their corresponding coefficients c_i:

$$V(x) = \sum_{i=1}^{\infty} c_i \zeta_i(x) = \sum_{i=1}^{N} c_i \zeta_i(x) + \sum_{i=N+1}^{\infty} c_i \zeta_i(x) \tag{25}$$

For notational simplicity, we introduce:

$$V(x) \triangleq C_1 \zeta + \sum_{i=N+1}^{\infty} c_i \zeta_i(x) \tag{26}$$

where $C_1 \triangleq \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}^T$, $\zeta \triangleq \begin{bmatrix} \zeta_1 & \zeta_2 & \cdots & \zeta_n \end{bmatrix}^T$.

According to [22,29], there exist unknown neural network weights W_c and an neural network basis function vector $\zeta(x)$ that can approximate the cost function $V(x)$ as:

$$V(x) = W_c^T \zeta(x) + \varepsilon(x) \tag{27}$$

where ζ: $R^n \to R^N$ and N represents the number of neurons in the hidden layer, and ε is the approximation error.

Substituting (27) into the Hamiltonian function (21), we obtain:

$$H(x, u, W_c) = W_c^T \nabla\zeta(f + gu) + Q(x) + u^T R u = \varepsilon_H \tag{28}$$

where the residual error ε_H is expressed as:

$$\begin{aligned} \varepsilon_H &= -(\nabla\varepsilon)^T (f + gu) - W_c^T \nabla\zeta\delta \\ &= -(C_1 - W_c)^T \nabla\zeta(f + gu) - \sum_{i=N+1}^{\infty} c_i \nabla\zeta_i(x)(f + gu) - W_c^T \nabla\zeta\delta \end{aligned} \tag{29}$$

Similarly, the estimated cost function $\hat{V}(x)$ obtained using the estimated weights $\hat{W}_c$ can be expressed as:

$$\hat{V}(x) = \hat{W}_c^T \zeta(x) \tag{30}$$

and the corresponding Hamiltonian function of $\hat{V}(x)$ is given by:

$$H(x, u, \hat{W}_c) = \hat{W}_c^T \nabla\zeta(f + gu) + Q(x) + u^T R u = e_1 \tag{31}$$

$$e_1 = -\tilde{W}_c^T \nabla\zeta(f + gu) + \varepsilon_H \tag{32}$$

where we have $\tilde{W}_c \triangleq W_c - \hat{W}_c$.

To minimize the estimation error $\tilde{W}_c$, we define $E_1 = \frac{1}{2} e_1^T e_1$. Following [19], the normalized gradient descent algorithm yields:

$$\dot{\hat{W}}_c = -a_1 \frac{\partial E_1}{\partial \hat{W}_c} = -a_1 \frac{\sigma_1}{(\sigma_1^T \sigma_1 + 1)^2} [\sigma_1^T \hat{W}_c + Q(x) + u^T R u] \tag{33}$$

where a_1 is a tunable parameter and $\sigma_1 = \nabla\zeta(f + gu)$, while $(\sigma_1^T \sigma_1 + 1)^2$ is used for normalization.

To achieve optimal control, a neural network-based estimator for the optimal controller u_2 is given by:

$$u_2(x) = -\frac{1}{2} R^{-1} g^T(x) \nabla\zeta^T \hat{W}_a \tag{34}$$

where $\hat{W}_a$ denotes the estimate of W_c with the estimation error of W_a defined as $\tilde{W}_a \triangleq W_a - \hat{W}_a$.

The update law of $\hat{W}_a$ is designed as:

$$\dot{\hat{W}}_a = -a_2 \left(\eta_a \hat{W}_a - \eta_c \hat{W}_c - \frac{1}{4} \mathcal{G} \hat{W}_a m^T \hat{W}_c \right) \tag{35}$$

where a_2 is a tunable parameter and we have:

$$\mathcal{G} \triangleq \nabla\zeta g R^{-1} g^T \nabla\zeta^T \tag{36}$$

$$\sigma_2 \triangleq \nabla\zeta(f + g u_2) \tag{37}$$

$$m \triangleq \frac{\sigma_2}{(\sigma_2^T \sigma_2 + 1)^2} \tag{38}$$

Assumption 2.

1. *f(x) is Lipschitz, and g(x) is bounded by a constant:*

$$\| f(x) \| < \gamma_f \| x \|$$
$$\| g(x) \| < \gamma_g$$

2. *The NN approximate error and its gradient are bounded by:*

$$\|\varepsilon\| < \gamma_\varepsilon$$
$$\|\nabla\varepsilon\| < \gamma_{\varepsilon_x}$$

3. *The NN basis functions and their gradients are bounded by:*

$$\|\zeta(x)\| < \gamma_\zeta$$
$$\|\nabla\zeta(x)\| < \gamma_{\zeta x}$$

Remark 2. *These assumptions are fundamentally grounded in the optimal control framework based on ADP [30–32], with rigorous theoretical foundations established in [28] and practical validations demonstrated in [20].*

Theorem 2. *For the PMSM model (1) equipped with the feedforward controller (7) and the feedback controller (34), the update law of the critic neural network is designed as:*

$$\dot{\hat{W}}_c = -a_1 \frac{\sigma_2}{(\sigma_2^T \sigma_2 + 1)^2} (\sigma_2^T \hat{W}_c + Q + u_2^T R u_2), \tag{39}$$

and the updating law of the actor neural network as:

$$\dot{\hat{W}}_a = -a_2 (\eta_a \hat{W}_a - \eta_c \hat{W}_c - \frac{1}{4} \mathcal{G} \hat{W}_a m^T \hat{W}_c), \tag{40}$$

and $\eta_c > 0$ and $\eta_a > 0$ are tuning parameters satisfy LMI:

$$\begin{bmatrix} \epsilon I & 0 & 0 \\ 0 & I & \left(-\frac{1}{2}\eta_c - \frac{1}{8m_s}\mathcal{G}W_c\right)^T \\ 0 & -\frac{1}{2}\eta_c - \left(\frac{1}{8m_s}\mathcal{G}W_c\right) & \eta_a - \frac{1}{8}\left(\mathcal{G}W_c m^T + m W_c^T \mathcal{G}\right) \end{bmatrix} > 0 \tag{41}$$

let Assumption 2 hold, and there exists a positive integer N_0 such that the number of the hidden layer units $N > N_0$, then, the error dynamic system states x, the critic neural network approximate error $\tilde{W}_c$, and the actor neural network approximate error $\tilde{W}_a$ are uniformly ultimately bounded.

Proof of Theorem 2. The complete mathematical derivations are provided in Appendix A.2. □

Figure 1 illustrates the complete flowchart of the proposed robust optimal control strategy for clearer demonstration.

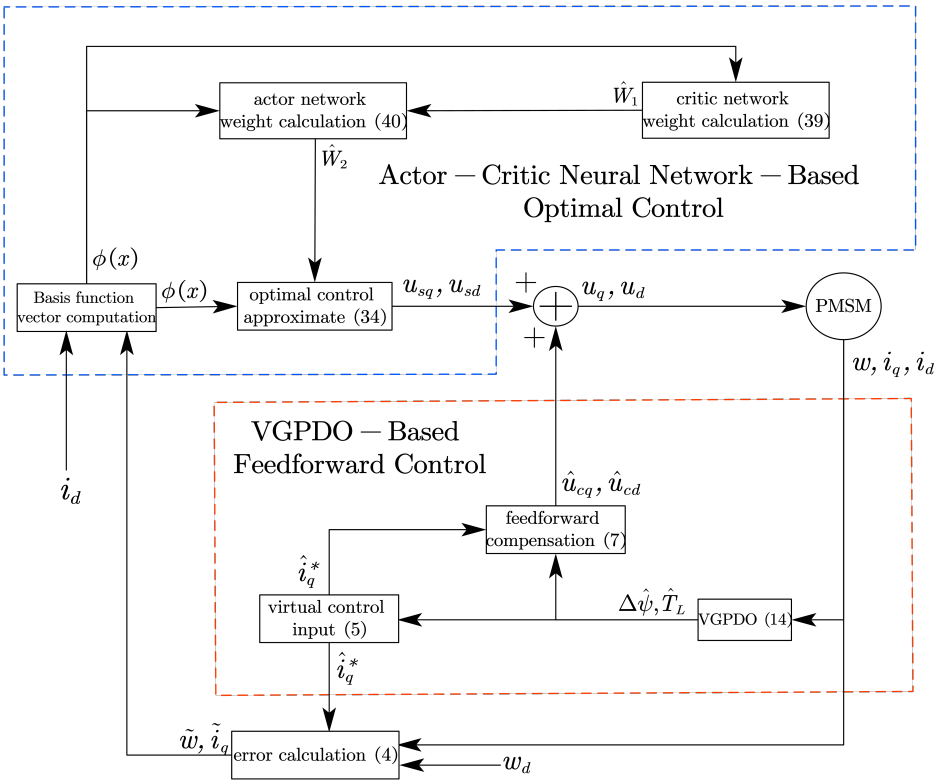


Figure 1. Flowchart of the robust optimal control strategy based on VGPDO and actor-critic neural network.

5. Simulation Results

To evaluate the effectiveness of the proposed approach, Software-in-the-Loop (SIL) simulations were conducted within the MATLAB/Simulink (R2023b) environment to implement and evaluate both the VGPDO algorithm and the robust optimal control strategy. The simulation parameters for PMSM in this study are consistent with those in [18,19], with detailed parameter values presented in Table 1.

Table 1. Parameters of PMSM.

Symbols	Values	Symbols	Values
U_R	60 V	n_p	4
I_R	12 A	ψ_s	0.0192 Wb
J	7.06×10^{-4} kg · m^2	R_s	0.72 Ω
B	3.5×10^{-4} N · s/rad	L	0.4×10^{-3} H

In simulation case 1, to demonstrate the effectiveness of the proposed VGPDO capable of simultaneous torque and flux estimation, two benchmark methods are employed for comparison: (1) a sliding-mode flux observer (SMO) algorithm from [25], which has been adopted by Simulink as a standard benchmark module, and (2) a load torque nonlinear disturbance observer (NDO) from [19]. Notably, the NDO explicitly accounts for demagnetization effects through compensation, where the compensation quantity is derived from the flux linkage estimation provided by SMO.

The load torque profile is configured as: (1) 0–5 s: linear ramp from 0 N·m to 2 N·m; (2) 5–30 s: linear decrease from 2 N·m to 1 N·m; (3) 30–40 s: sinusoidal oscillation (mean: 1 N·m, amplitude: 1 N·m, frequency: 2π rad/s); (4) 40–60 s: constant torque at 1 N·m. The

flux linkage variation includes three phases: (1) 0–2 s: maintained at initial value; (2) 2–40 s: decay at a rate of 0.6% of ψ_s per second; (3) 40–60 s: held constant. Additionally, an abrupt 20% flux linkage reduction is introduced at 40 s to emulate abrupt demagnetization under fault conditions. To intuitively demonstrate the performance of the proposed VGPDO algorithm, a simulation was first conducted under ideal, noise-free conditions, and the results are presented in Figures 2 and 3.

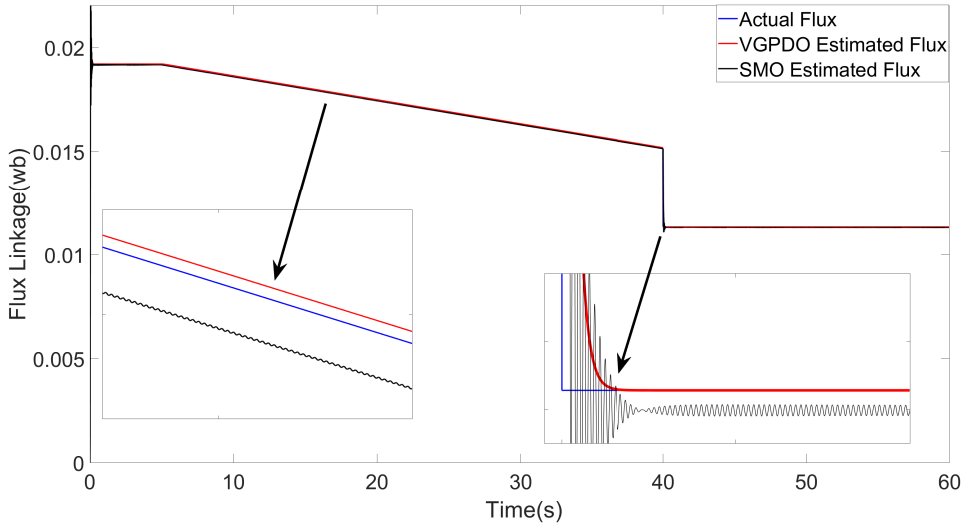


Figure 2. Flux linkage estimation results achieved by VGPDO and SMO.

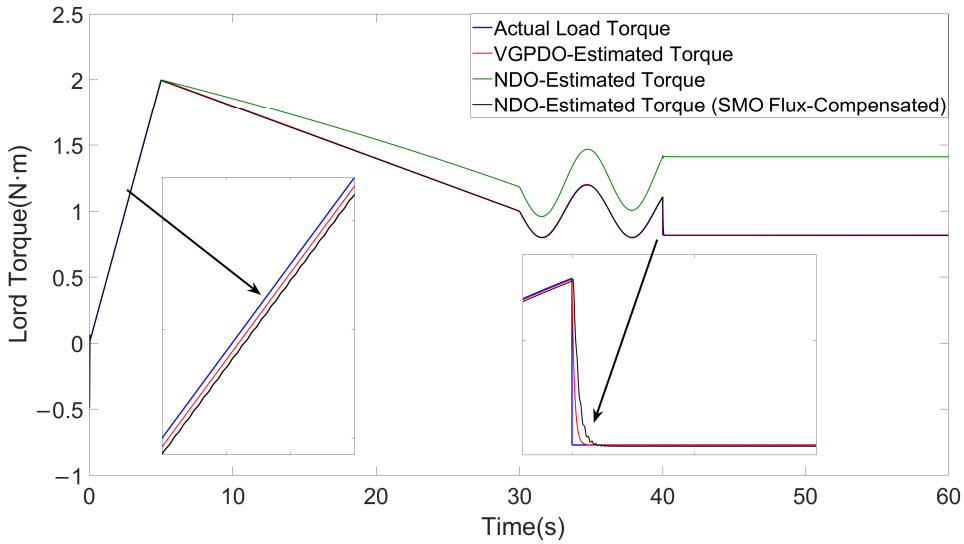


Figure 3. Torque estimation results achieved by VGPDO and SMO.

The experimental results are presented in Figures 2 and 3.

The flux linkage estimation results of the two observers are compared in Figure 2. During t = 0–40 s, the proposed VGPDO demonstrates superior dynamic tracking performance for flux estimation compared to the SMO proposed in [25], with no significant chattering phenomena observed. Following the abrupt flux linkage demagnetization event at t = 40 s, the SMO exhibits significant high-amplitude chattering in tracking. In contrast, VGPDO demonstrates smoother and more stable tracking performance. Ultimately, when both flux linkage and load torque stabilize after t = 40 s, VGPDO demonstrates a smooth transient response with rapid convergence, achieving asymptotic stability.

The torque estimation performance of different observers is shown in Figure 3. The NDO proposed in [25] achieves relatively accurate load torque estimation in the early

period. However, with progressive deterioration of the flux linkage, the estimation error increases steadily, eventually causing complete observer failure. The SMO-compensated NDO achieves satisfactory load torque tracking but demonstrates persistent chattering, which intensifies dramatically under severe flux linkage degradation at t = 40 s, while the proposed VGPDO maintains smooth and stable tracking for the load torque all the time.

To further evaluate the performance of the proposed VGPDO algorithm and demonstrate its robustness against measurement noise, the simulation study was extended to include three distinct noise variance levels, building upon the initial noise-free results. A comparative summary of the performance metrics across all conditions is provided in Table 2.

Table 2. Performance comparison of VGPDO and NDOB-SMO algorithms under different measurement noise variance conditions.

Algorithm	Noise Variance ($\delta_w, \delta_{iq}^2, \delta_{id}^2$)	Maximum Estimation Error (%)		RMSE		Maximum Setting Time (s)	CPU Usage (%)
		Load Torque	Flux Linkage	Load Torque	Flux Linkage		
VGPDO	$(0, 0, 0)$	0.3331	0.0755	2.961×10^{-3}	5.074×10^{-5}	0.0782	12.26
	$(1, 10^{-3}, 10^{-3})$	0.3964	0.1525	2.986×10^{-3}	5.083×10^{-5}	0.0801	13.98
	$(10, 10^{-2}, 10^{-2})$	0.6165	0.2469	3.271×10^{-3}	5.175×10^{-5}	0.0857	14.24
	$(100, 0.5, 0.5)$	3.601	3.109	1.176×10^{-2}	1.263×10^{-4}	N/A	13.17
NDOB-SMO	$(0, 0, 0)$	0.7176	0.3496	1.148×10^{-2}	3.358×10^{-4}	0.1328	11.6
	$(1, 10^{-3}, 10^{-3})$	2.996	0.5022	1.251×10^{-2}	3.590×10^{-4}	0.1331	15.74
	$(10, 10^{-2}, 10^{-2})$	3.365	0.9817	1.955×10^{-2}	3.608×10^{-4}	0.1345	15.96
	$(100, 0.5, 0.5)$	7.957	5.422	2.271×10^{-2}	4.244×10^{-4}	N/A	14.2

Remark 3. *The terms δ_w^2, δ_{iq}^2, and δ_{id}^2 denote the measurement noise variances for the angular velocity, q-axis current, and d-axis current, respectively; The maximum settling time is defined as the duration required for the observer's estimation errors of both load torque and flux linkage to fall and remain below 2% after a step disturbance is applied to the PMSM at t = 40 s; The term "N/A" denotes a non-existing value. The CPU utilization was calculated in a SIL environment by converting the algorithm's floating-point operation count into the equivalent execution time on a 180 MHz processor.*

As shown in Table 2, under various noise conditions, the proposed VGPDO achieves smaller maximum estimation errors and lower RMSE for both load torque and flux, demonstrating better tracking accuracy. Meanwhile, the proposed VGPDO also exhibits a shorter maximum settling time, indicating a faster dynamic response.

Additionally, the proposed VGPDO algorithm demonstrates robustness to measurement noise, as evidenced by the minimal variation in its performance metrics across the first three different noise levels. Moreover, under different measurement noise variances, the CPU utilization of the proposed VGPDO algorithm showed minimal change. This suggests that the level of measurement noise has no strong correlation with its computational burden, and thus did not significantly increase the real-time computational demand.

To demonstrate the efficacy of the tuned parameters α_1 and α_2 in VGPDO, a sensitivity analysis was conducted under measurement noise variances of $\delta_w^2 = 1$, $\delta_{i_q} = 10^{-3}$ and $\delta_{i_d} = 10^{-3}$; the results are summarized in Table 3.

As shown in Table 3, the parameters α_1 and α_2 exhibit an approximately linear inverse relationship with the settling time. When α_1 and α_2 are moderately increased from 10 to 100, the settling time of the VGPDO decreases significantly from 0.3921 s to 0.0403 s, indicating a substantial improvement in convergence speed and dynamic tracking performance. This trend is further corroborated by the monotonic decrease in the RMSE for both the load torque (from 1.520×10^{-2} to 2.986×10^{-3}) and the flux linkage (from 1.116×10^{-3} to 3.590×10^{-4}).

However, when $\alpha_1 = \alpha_2 = 1000$, the maximum estimation error increases. This suggests that excessively large parameters degrade the algorithm's filtering effectiveness against measurement noise, consequently reducing estimation accuracy.

Table 3. Sensitivity analysis of α_1, α_2 on VGPDO performance.

α_1	α_2	Maximum Estimation Error (%)		RMSE		Maximum Setting Time (s)	CPU Usage (%)
		Load Torque	Flux Linkage	Load Torque	Flux Linkage		
10	10	3.306	0.1265	1.520×10^{-2}	1.116×10^{-3}	0.3921	11.63
30	30	1.116	0.1183	6.470×10^{-3}	6.424×10^{-4}	0.1308	15.74
50	50	0.6842	0.1525	4.563×10^{-3}	4.987×10^{-4}	0.0801	15.96
100	100	0.3965	0.2206	2.986×10^{-3}	3.590×10^{-4}	0.0403	14.20
1000	1000	0.6471	0.8010	1.791×10^{-3}	2.782×10^{-4}	N/A	12.96

Therefore, selecting α_1, α_2 necessitates a trade-off between estimation accuracy and dynamic response speed. Systematic tuning is required to determine the optimal values for achieving the best overall performance of the VGPDO algorithm.

In simulation case 2, to validate the effectiveness of the proposed optimal robust control strategy integrating VGPDO and actor-critic neural network-based optimal speed controller (VGPDO-AC), we select the a set of optimal control strategy integrating a ESO and actor-critic neural network-based optimal speed controller (ESO-AC) from [18] as the benchmark. This comparative strategy employs ESO for system disturbance estimation and compensation, while utilizing actor-critic neural network to achieve the optimal control policy. Additionally, an H ∞ robust control algorithm that incorporates flux compensation is also adopted as a benchmark.

The load torque profile is configured as: (1) 0–2 s: constant at 1.5 N·m; (2) 2–10 s: linear increase from 1.5 N·m to 1.8 N·m; (3) 10–15 s: rectangular wave (mean: 1.8 N·m, amplitude: ±0.3 N·m, period: 2 s). The flux linkage variation includes: (1) 0–2 s: maintained at initial value ψ_s; (2) 2–15 s: decay at a rate of 2% of ψ_s per second. Additionally, an abrupt 20% reduction is introduced at 7 s to emulate fault conditions. The PMSM reference speed ω_d is set as a step command of 1000 RPM. The simulation is conducted with measurement noise variances of $\delta_w^2 = 1$, $\delta_{i_q} = 10^{-3}$, and $\delta_{i_d} = 10^{-3}$.

For the actor-critic neural network, the weight matrix for state deviations is given by: $Q = \begin{bmatrix} 100 & 0 & 0 \\ 0 & 100 & 0 \\ 0 & 0 & 100 \end{bmatrix}$. The weight matrix for input is given by: $R = \begin{bmatrix} 0.5 & 0 \\ 0 & 1 \end{bmatrix}$. The basis functions are selected as: $\zeta(x) = \begin{bmatrix} x_1x_2 & x_2x_3 & x_1x_3 & x_1^2 & x_2^2 & x_3^2 \end{bmatrix}^T$. The tunable parameter η_c and η_a is set to be 1 and 1, respectively.

The simulation results are presented in Figures 4–8. The relevant performance metrics from the experimental results can be found in Table 4.

Table 4. Speed control performance evaluated for three strategies under comprehensive operating conditions.

Control Strategy	Steady-State Error (%)	RMSE	Energy Consumption (J)	Settling Time (s)
$H\infty$- Flux Compensation	0.08331	32.8425	1.424×10^4	0.0814
ESO-AC	0.39359	34.4345	1.387×10^4	0.1044
VGPDO-AC	0.01217	27.3861	1.396×10^4	0.0676

As summarized in Table 4, the proposed VGPDO-AC strategy achieves the highest steady-state accuracy and fastest dynamic response, as evidenced by its smallest steady-state error (0.01217%), shortest settling time (0.0676 s), and minimum RMSE (27.3861) among all control strategies. Furthermore, it maintains lower power consumption (1.396×10^4 J) than the contemporary H-infinity control (1.424×10^4 J). Although its energy consumption is slightly higher than that of the ESO-AC strategy, the VGPDO-AC

demonstrates notable overall control performance, positioning it as a high-performance and energy-efficient control solution.

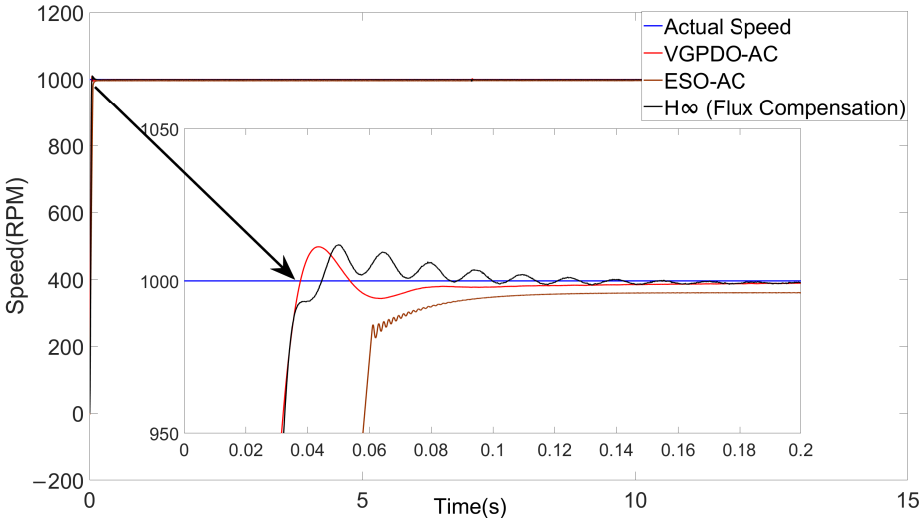


Figure 4. Speed trajectories under three strategy.

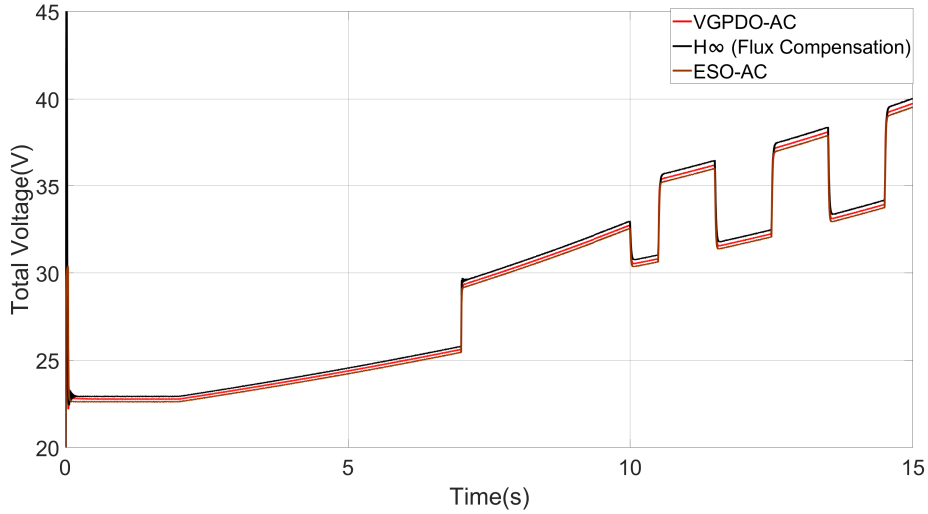


Figure 5. Total voltage under three strategy.

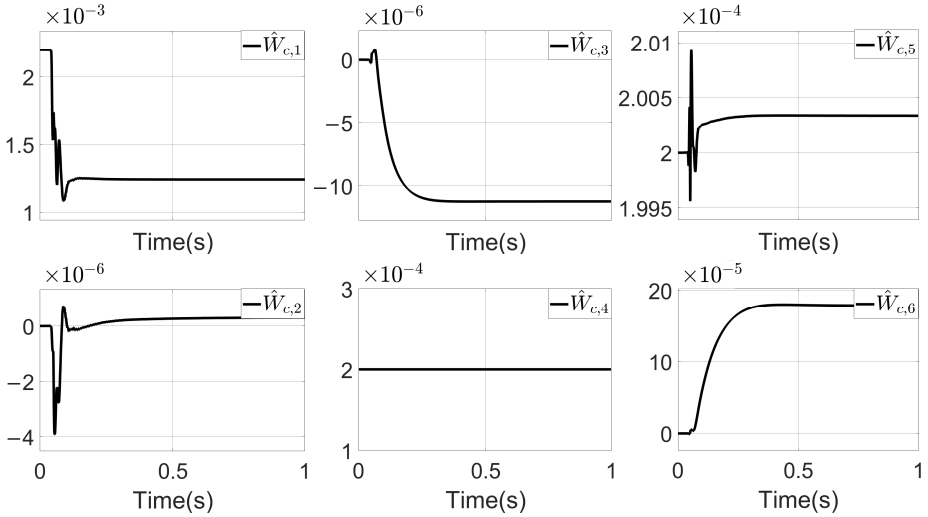


Figure 6. Critic NN weight update trajectory on VGPDO-AC.

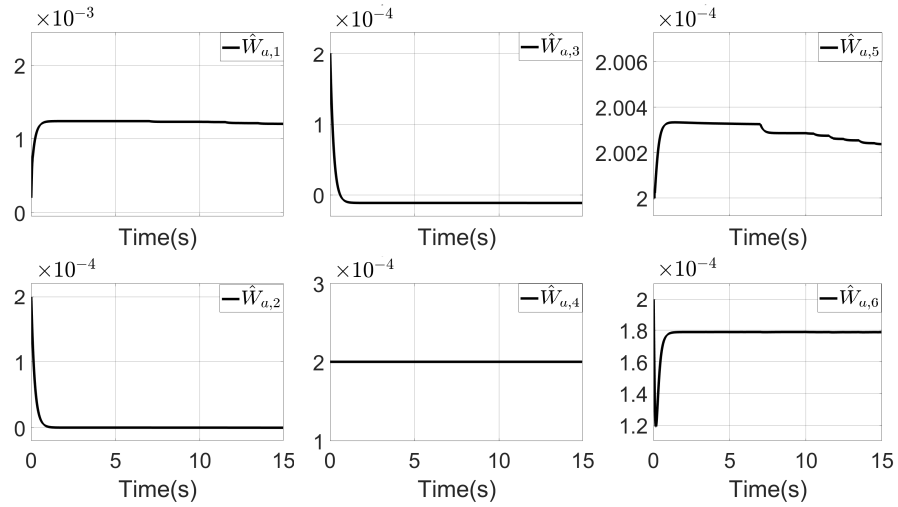


Figure 7. Actor NN weight update trajectory on VGPDO-AC.

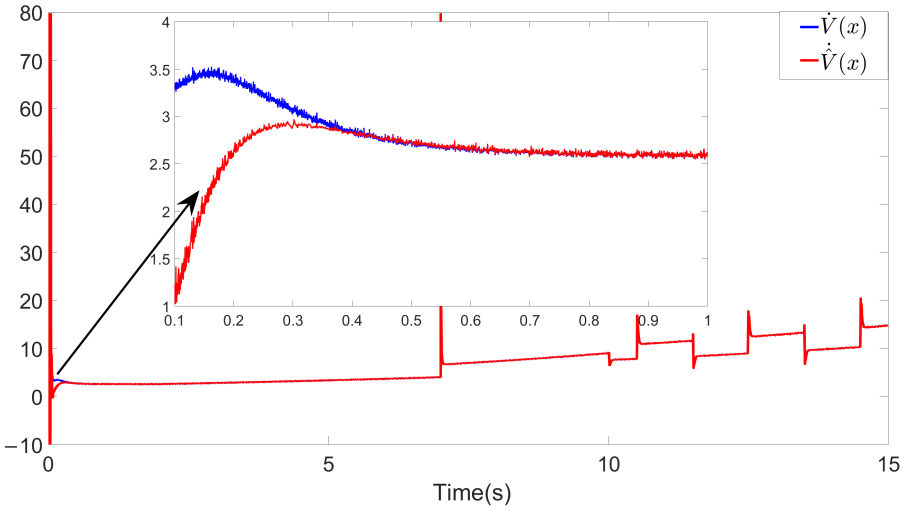


Figure 8. $\dot{V}(x)$ and $\dot{\hat{V}}(x)$ trajectory on VGPDO-AC.

To demonstrate the robustness of the proposed VGPDO-AC control strategy, the dynamic speed responses of the PMSM under flux linkage and load torque step changes are illustrated in Figure 9 and Figure 10, respectively, with the corresponding performance metrics provided in Tables 5 and 6.

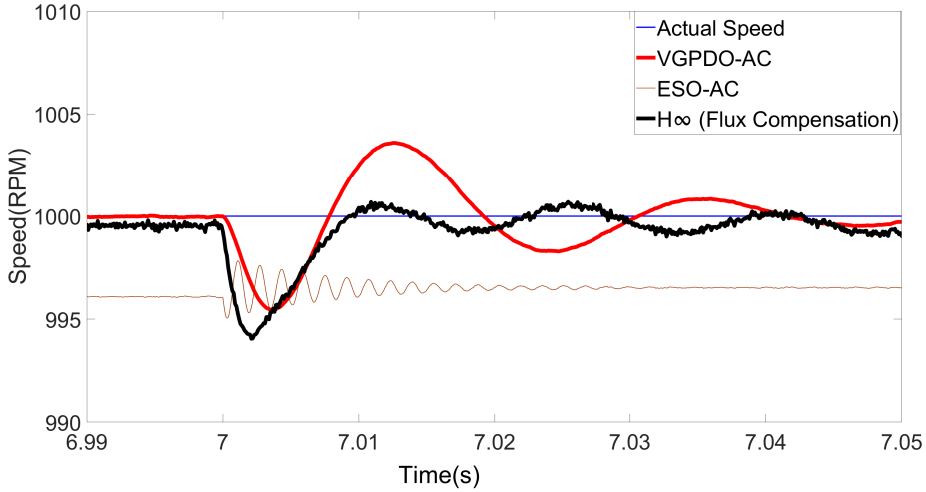


Figure 9. Speed trajectories for three strategies under flux linkage step disturbance.

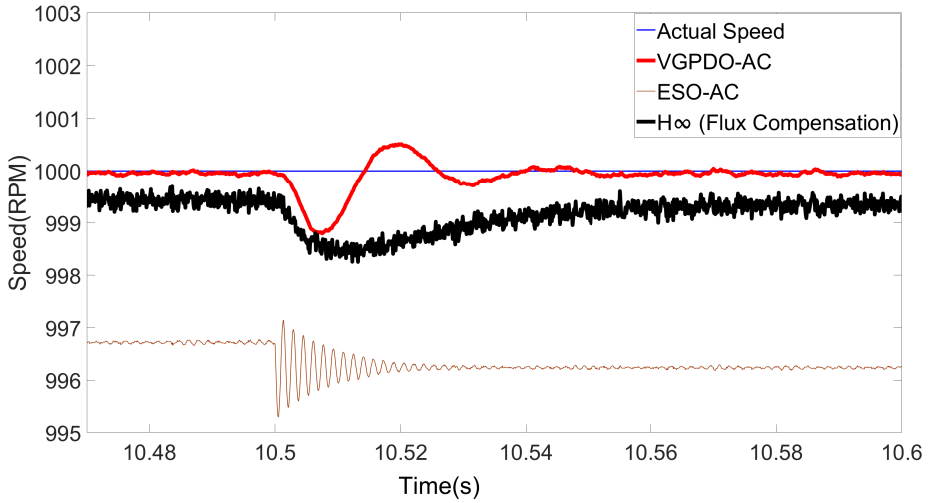


Figure 10. Speed trajectories for three strategies under load torque step disturbance.

Table 5. Speed control performance evaluated for three strategies under flux linkage step disturbance.

Algorithm	Maximum Deviation (%)	RMSE
$H\infty$- Flux Compensation	0.5957	0.6515
ESO-AC	0.495	3.4886
VGPDO-AC	0.4623	0.3582

Table 6. Speed control performance evaluated for three strategies under load torque step disturbance.

Algorithm	Maximum Deviation (%)	RMSE
$H\infty$- Flux Compensation	0.1747	0.6947
ESO-AC	0.3816	3.7403
VGPDO-AC	0.1156	0.1034

As shown in Tables 5 and 6, the proposed VGPDO-AC control strategy achieves both the lowest RMSE and the minimum maximum deviation under flux linkage and load torque disturbances. These results demonstrate that the proposed VGPDO-AC control strategy exhibits enhanced robustness against flux-weakening and abrupt load changes.

Based on the results from simulation case 2 under the step speed command scenario during the first second, a sensitivity analysis of the tunable parameters η_a and η_c is conducted in Tables 7 and 8 to quantify their effects on key performance metrics.

Table 7. Sensitivity analysis of η_a, η_c on PMSM speed control.

η_a	η_c	Speed		
		RMSE	Steady-State Error (%)	Settling Time (s)
0.1	0.1	100.20	0.00972	0.1004
0.5	0.5	96.99	0.01058	0.0874
1	1	92.21	0.01217	0.0689
2	2	90.87	0.0226	0.0443

Remark 4.

- *The settling time of* $\hat{W}_a$ *is defined as the earliest time* $t_{settling}$ *such that for all* $t > t_{settling}$*, we have:* $\frac{\|\hat{W}_a(t)-\hat{W}_c(t)\|}{\|\hat{W}_c(t)\|} \leq 2\%$

- *The percentage deviation of $\hat{W}_a$ is used to evaluate its tracking performance relative to $\hat{W}_c$ and is computed as:*

$$\max_{t \geq t_{steady}} \left\| \frac{\hat{W}_c(t) - \hat{W}_a(t)}{\hat{W}_c(t)} \right\| \times 100\%$$

where t_{steady} denotes the time instant at which the system state enters the steady-state regime, and is is computed as twice the value of $t_{settling}$.

- *The steady-state estimation error on $\dot{\hat{V}}(x)$ is used to evaluate the approximation degree of $\hat{W}_c$ to the actual cost function after $\hat{W}_c$ has converged, and is calculated as:*

$$\max_{t \geq t_{steady}} \left\| \frac{\dot{V}(x(t)) - \dot{\hat{V}}(x(t))}{\dot{V}(x(t))} \right\| \times 100\%$$

As shown in Table 8, the steady-state estimation error of $\dot{\hat{V}}(x)$ remains below 2% with no significant variation. This confirms that the critic neural network weight $\hat{W}_c$ in the proposed VGPDO-AC algorithm achieves an excellent approximation of the cost function $V(x)$. Furthermore, the CPU usage remains within the range of 50% to 60%, demonstrating the computational efficiency of the proposed algorithm. This indicates its potential for practical deployment on processors with 180 MHz clock frequencies at 5 kHz control rates.

Table 8. Sensitivity analysis of η_a, η_c on NN network update performance of VGPDO-AC algorithm.

η_a	η_c	$\hat{W}_a$		Steady-State Estimation Error on $\dot{\hat{V}}(x)$ (%)	CPU Usage (%)
		Percentage Deviation (%)	Settling Time (s)		
0.1	0.1	0.5307	0.6329	1.684	60.329
0.5	0.5	0.3303	0.3048	1.653	59.484
1	1	0.1050	0.2066	1.640	56.649
2	2	0.0129	0.1609	1.807	52.418

Meanwhile, as shown in Tables 7 and 8, as parameters η_a and η_c increase, the dynamic performance of both PMSM speed control and neural network weight updates is improved, evidenced by the monotonic decrease in the RMSE of the speed, the settling time of the speed, percentage deviation on $\hat{W}_a$, settling time of $\hat{W}_a$. This indicates that appropriately increasing η_a and η_c is beneficial for enhancing both the dynamic response rate of PMSM speed control and the convergence rate of $\hat{W}_a$. However, this comes at the trade-off of an increase in the steady-state error of the speed. Therefore, the selection of η_a and η_c should be based on the specific control performance requirements.

6. Conclusions

In this paper, a VGPDO is proposed for PMSM that simultaneously estimates both flux linkage and load torque. Experimental results demonstrate that the VGPDO achieves smooth and rapid tracking of both flux linkage and load torque variations, while guaranteeing asymptotic stability when flux linkage and load torque reach steady-state conditions. Furthermore, we developed a robust optimal control framework that integrates the VGPDO estimates through feedforward compensation, which transforms the PMSM dynamics into an error dynamic system. Then we employ an actor-critic neural network to achieve optimal control. Experimental validation confirms that the proposed control strategy achieves optimal control while exhibiting superior dynamic tracking performance with minimal steady-state error and maintaining robustness against both flux-weakening operation and abrupt load torque variations.

Author Contributions: Conceptualization, Y.N.; methodology, Y.N.; software, Y.N.; validation, Y.N., H.S.; formal analysis, Y.N., H.S.; writing—original draft, Y.N.; writing—review and editing, Y.N.; supervision, Y.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no funding.

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.

Abbreviations

The following abbreviations are used in this manuscript:

PMSM	Permanent magnet synchronous motor
UUB	Uniform ultimate boundedness
ADP	Adaptive dynamic programming
RL	Reinforcement learning
VGPDO	Variable-gain proportional disturbance observer
AC	actor-critic neural network

Appendix A

Appendix A.1

By substituting (18) into (15), we obtain:

$$\begin{cases} \dot{\tilde{T}}_L = -\alpha_1 \tilde{T}_L + \dot{T}_L \\ \Delta\dot{\tilde{\psi}} = -\alpha_2 \Delta\tilde{\psi} + \Delta\dot{\psi} \end{cases} \tag{A1}$$

Therefore, based on Assumption 1, the estimation error concerning flux linkage deterioration and disturbance torque can be derived as:

$$\begin{cases} \| \tilde{T}_L \| \leqslant \frac{\Upsilon_1}{\sqrt{1+\alpha_1^2}} \\ \| \Delta\tilde{\psi} \| \leqslant \frac{\Upsilon_2}{\sqrt{1+\alpha_2^2}} \end{cases} \tag{A2}$$

Notably, when the rate of change for both flux linkage and load torque becomes zero (i.e., $\dot{T}_L = 0$ and $\Delta\dot{\psi} = 0$), it follows from (A1) that $\dot{\tilde{T}}_l$ and $\Delta\dot{\tilde{\psi}}$ will achieve asymptotic stability.

Furthermore, the following expression can be derived:

$$\| \frac{\tilde{m}_1}{\hat{m}_1} \| = \| \frac{\Delta\tilde{\psi}}{\psi_s + \Delta\hat{\psi} + \Delta\tilde{\psi}} \| \leqslant \frac{\frac{\Upsilon_2}{\sqrt{1+\alpha_2^2}}}{\psi_s + \frac{\Upsilon_2}{\sqrt{1+\alpha_2^2}}} = \frac{\Upsilon_2}{\psi_s\sqrt{1+\alpha_2^2} + \Upsilon_2} \tag{A3}$$

Simultaneously, the following relationship holds:

$$\begin{aligned} \| m_2\omega_d + m_3\hat{T}_L + \dot{\omega}_d \| &= \| m_2\omega_d + m_3 T_L + \dot{\omega}_d - m_3\tilde{T}_L \| \\ &\leqslant \| m_2\gamma_3 + m_3\gamma_1 + \Upsilon_3 - m_3\frac{\Upsilon_1}{\sqrt{1+\alpha_1^2}} \| \end{aligned} \tag{A4}$$

By combining inequality (A2)–(A4), with Assumption 1, the following can be obtained:

$$\| -m_3\tilde{T}_L + \frac{\tilde{m}_1}{\hat{m}_1}\left(m_2\omega_d + m_3\hat{T}_L + \dot{\omega}_d\right) \| \leqslant \delta_1 \tag{A5}$$

where:

$$\delta_1 \triangleq \sqrt{\left(m_3\frac{\Upsilon_1}{\sqrt{1+\alpha_1^2}}\right)^2+\left(\frac{\Upsilon_2}{\psi_s\sqrt{1+\alpha_2^2}+\Upsilon_2}(m_2\gamma_3+m_3\gamma_1+\Upsilon_3-m_3\frac{\Upsilon_1}{\sqrt{1+\alpha_1^2}})\right)^2} \tag{A6}$$

From inequality (A2) and Assumption 1, it can be readily derived that:

$$\| \tilde{m}_5 w_d \| \leqslant \gamma_3 \| \frac{n_p \Delta\tilde{\psi}}{L} \| \leqslant \frac{n_p\gamma_3\Upsilon_2}{L\sqrt{1+\alpha_2^2}} \tag{A7}$$

Finally, by incorporating inequalities (A5) and (A7), we establish the boundary of δ as:

$$\| \delta \| \leqslant \delta_m \tag{A8}$$

where:

$$\delta_m \triangleq \sqrt{\delta_1^2+\frac{n_p^2\gamma_3^2\Upsilon_2^2}{L^2(1+\alpha_2^2)}} \tag{A9}$$

Appendix A.2

Take (24) into (23), we have:

$$0 = Q(x)+V_x^{*T}(x)f(x)-\frac{1}{4}V_x^{*T}(x)g(x)R^{-1}g^T(x)V_x^*(x)+V_x^{*T}(x)\delta \tag{A10}$$

Take (27) into (A10), we have:

$$W_c^T\nabla\zeta f-\frac{1}{4}W_c^T\nabla\zeta \mathrm{g}\mathrm{R}^{-1}g^T\nabla\zeta^T W_c+Q(x)=\varepsilon_{HJB} \tag{A11}$$

where $\varepsilon_{HJB} = -\nabla\varepsilon^T f+\frac{1}{2}W_c^T\nabla\zeta \mathrm{g}\mathrm{R}^{-1}g^T\nabla\varepsilon+\frac{1}{4}\nabla\varepsilon^T gR^{-1}g^T\nabla\varepsilon-W_c^T\nabla\zeta\delta-\nabla\varepsilon^T\delta$. Take (36) and $\sigma_1=\nabla\zeta(f+gu)$ into (A11) yield:

$$W_c^T\sigma_1 = -Q(x)-\frac{1}{4}W_c^T\mathcal{G}(x)W_c+\varepsilon_{HJB}(x) \tag{A12}$$

The whole Lyapunov function is given by:

$$L = L_V(x)+L_c(x)+L_a(x)=\frac{1}{2}x^Tx+\frac{1}{2}tr(\tilde{W}_c^Ta_1^{-1}\tilde{W}_c)+\frac{1}{2}tr(\tilde{W}_a^Ta_2^{-1}\tilde{W}_a) \tag{A13}$$

where we have: $L_V(x)\triangleq\frac{1}{2}x^Tx$, $L_c(x)\triangleq\frac{1}{2}tr(\tilde{W}_c^Ta_1^{-1}\tilde{W}_c)$, $L_a(x)\triangleq\frac{1}{2}tr(\tilde{W}_a^Ta_2^{-1}\tilde{W}_a)$.

The derivitive of L_V is written as:

$$\begin{aligned}\dot{L}_V(x) =& \frac{\partial L_V}{\partial x}\cdot\dot{x}=\left(\nabla\zeta^TW_c+\nabla\varepsilon\right)\cdot(f(x)+gu_2+\delta)\\ =& W_c^T\left(\nabla\zeta f(x)-\frac{1}{2}\mathcal{G}(x)\hat{W}_a\right)+\left(\nabla\zeta^TW_c+\nabla\varepsilon\right)\delta\\ &+\nabla\varepsilon^T(x)\left(f(x)-\frac{1}{2}g(x)R^{-1}g^T(x)\nabla\zeta^T\hat{W}_a\right)\end{aligned} \tag{A14}$$

To simplify the notation, let us define:

$$\varepsilon_1(x)\triangleq\nabla\varepsilon^T(x)\left(f(x)-\frac{1}{2}g(x)R^{-1}g^T(x)\nabla\zeta^T(x)\hat{W}_a\right)+\left(\nabla\zeta^TW_c+\nabla\varepsilon\right)\delta \tag{A15}$$

The (A14) could be written as:

$$\begin{aligned}\dot{L}_V(x) =& W_c^T\left(\nabla\zeta f(x) - \frac{1}{2}\mathcal{G}(x)\hat{W}_a\right) + \varepsilon_1(x)\\ =& W_c^T\nabla\zeta f(x) + \frac{1}{2}W_c^T\mathcal{G}(x)\left(W_c - \hat{W}_a\right)\\ & - \frac{1}{2}W_c^T\mathcal{G}(x)W_c + \varepsilon_1(x)\\ =& W_c^T\nabla\zeta f(x) + \frac{1}{2}W_c^T\mathcal{G}(x)\tilde{W}_a - \frac{1}{2}W_c^T\mathcal{G}(x)W_c + \varepsilon_1(x)\\ =& W_c^T\sigma_1 + \frac{1}{2}W_c^T\mathcal{G}(x)\tilde{W}_a + \varepsilon_1(x)\end{aligned} \tag{A16}$$

Take (A12) into (A16) yields:

$$\begin{aligned}\dot{L}_V(x) =& -Q(x) - \frac{1}{4}W_c^T\mathcal{G}(x)W_c\\ & + \frac{1}{2}W_c^T\mathcal{G}(x)\tilde{W}_a + \varepsilon_{HJB}(x) + \varepsilon_1(x)\end{aligned} \tag{A17}$$

for $\dot{L}_c$, we have:

$$\begin{aligned}\dot{L}_c =& \tilde{W}_c^T a_1^{-1}\dot{\hat{W}}_c\\ =& \tilde{W}_c^T a_1^{-1}a_1\frac{\sigma_2}{\left(\sigma_2^T\sigma_2 + 1\right)^2}\left(\sigma_2^T\hat{W}_c + Q(x) + \frac{1}{4}\hat{W}_a^T\mathcal{G}\hat{W}_a\right)\end{aligned} \tag{A18}$$

Rewriting (A12) gives:

$$0 = -W_c^T\sigma_1 - Q(x) - \frac{1}{4}W_c^T\mathcal{G}(x)W_c + \varepsilon_{HJB}(x) \tag{A19}$$

Taking Equation (A19) into Equation (A18) yields the following:

$$\begin{aligned}\dot{L}_c =& \tilde{W}_c^T\frac{\sigma_2}{\left(\sigma_2^T\sigma_2 + 1\right)^2}\left(\sigma_2^T\hat{W}_c + Q(x) + \frac{1}{4}\hat{W}_a^T\mathcal{G}(x)\hat{W}_a\right.\\ & \left. - Q(x) - \sigma_1^T W_c - \frac{1}{4}W_c^T\mathcal{G}(x)W_c + \varepsilon_{HJB}(x)\right)\\ =& \tilde{W}_c^T\frac{\sigma_2}{\left(\sigma_2^T\sigma_2 + 1\right)^2}\left(\sigma_2^T(x)\hat{W}_c - \sigma_1^T(x)W_c\right.\\ & \left. + \frac{1}{4}\hat{W}_a^T\mathcal{G}(x)\hat{W}_a - \frac{1}{4}W_c^T\mathcal{G}(x)W_c + \varepsilon_{HJB}(x)\right)\end{aligned} \tag{A20}$$

Substituting (24) and (27) into $\sigma_1 = \nabla\zeta(f + gu)$ yields:

$$\begin{aligned}\sigma_1 =& \nabla\zeta(f + gu)\\ =& \nabla\zeta\left(f - \frac{1}{2}gR^{-1}g\left(W_c^T\zeta + \varepsilon\right)\right)\\ =& \nabla\zeta f - \frac{1}{2}\nabla\zeta gR^{-1}gW_c^T\zeta(x) - \frac{1}{2}\nabla\zeta gR^{-1}g\varepsilon(x)\end{aligned} \tag{A21}$$

Take (A21), (37), (36) into $\sigma_2^T(x)\hat{W}_c - \sigma_1^T(x)W_c$ yields:

$$\begin{aligned}
&\sigma_2^T(x)\hat{W}_c - \sigma_1^T(x)W_c \\
&= \left(\nabla\zeta(f+gu_2)\right)^T\hat{W}_c - \left(\nabla\zeta(f+gu)\right)^T W_c \\
&= -\nabla\zeta f\tilde{W}_c + (\nabla\zeta g u_2)^T\hat{W}_c - (\nabla\zeta g u)^T W_c \\
&= -\nabla\zeta f\tilde{W}_c + \left(-\frac{1}{2}\nabla\zeta g R^{-1}g^T(x)\nabla\zeta^T\hat{W}_a\right)^T\hat{W}_c - (\nabla\zeta g u)^T W_c \\
&= -\nabla\zeta f\tilde{W}_c - \frac{1}{2}\left(\hat{W}_a^T\mathcal{G}\hat{W}_c - W_c^T\mathcal{G}W_c\right) + \frac{1}{2}\left(\nabla\zeta g R^{-1}g\varepsilon(x)\right)^T W_c
\end{aligned} \tag{A22}$$

Additionally, we can transform $\frac{1}{2}\hat{W}_a^T\mathcal{G}\hat{W}_c - \frac{1}{2}W_c^T\mathcal{G}W_c - \frac{1}{4}\hat{W}_a^T\mathcal{G}(x)\hat{W}_a + \frac{1}{4}W_c^T\mathcal{G}(x)W_c$ as:

$$\begin{aligned}
&\frac{1}{2}\hat{W}_a^T\mathcal{G}\hat{W}_c - \frac{1}{2}W_c^T\mathcal{G}W_c - \frac{1}{4}\hat{W}_a^T\mathcal{G}(x)\hat{W}_a + \frac{1}{4}W_c^T\mathcal{G}(x)W_c \\
&= \frac{1}{2}\hat{W}_a^T\mathcal{G}(W_c - \tilde{W}_c) - \frac{1}{2}W_c^T\mathcal{G}W_c - \frac{1}{4}\hat{W}_a^T\mathcal{G}(x)\hat{W}_a + \frac{1}{4}W_c^T\mathcal{G}(x)W_c \\
&= -\frac{1}{2}\hat{W}_a^T\mathcal{G}\tilde{W}_c + \frac{1}{2}\hat{W}_a^T\mathcal{G}W_c - \frac{1}{2}W_c^T\mathcal{G}W_c - \frac{1}{4}\hat{W}_a^T\mathcal{G}(x)\hat{W}_a + \frac{1}{4}W_c^T\mathcal{G}(x)W_c \\
&= -\frac{1}{2}\hat{W}_a^T\mathcal{G}\tilde{W} + \frac{1}{2}\hat{W}_a^T\mathcal{G}W_c - \frac{1}{4}W_c^T\mathcal{G}W_c - \frac{1}{4}\hat{W}_a^T\mathcal{G}(x)\hat{W}_a \\
&= -\frac{1}{2}\hat{W}_a^T\mathcal{G}\tilde{W} - \frac{1}{4}\tilde{W}_a\mathcal{G}W_c + \frac{1}{4}\hat{W}_a^T\mathcal{G}\tilde{W}_a \\
&= -\frac{1}{2}\hat{W}_a^T\mathcal{G}\tilde{W} + \frac{1}{4}\tilde{W}_a^T\mathcal{G}(x)\tilde{W}_a
\end{aligned} \tag{A23}$$

Substituting (A22) into (A20) and adding (A23) yields:

$$\begin{aligned}
\dot{L}_c =& \tilde{W}_c^T\frac{\sigma_2}{\left(\sigma_2^T\sigma_2+1\right)^2}\Bigg(-f(x)^T\nabla\zeta^T(x)\tilde{W}_c + \frac{1}{2}\tilde{W}_a^T\mathcal{G}(x)\tilde{W}_c \\
&+ \frac{1}{4}\tilde{W}_a^T\mathcal{G}(x)\tilde{W}_a + \varepsilon_{HJB}(x) - \frac{1}{2}\left(\nabla\zeta g R^{-1}g\varepsilon\right)^T W_c\Bigg)
\end{aligned} \tag{A24}$$

By combining Equations (34), (36) and (37), Equation (A24) can be reformulated as:

$$\dot{L}_c = \tilde{W}_c^T\frac{\sigma_2}{\left(\sigma_2^T\sigma_2+1\right)^2}\left(-\sigma_2^T\tilde{W}_c + \frac{1}{4}\tilde{W}_a^T\mathcal{G}(x)\tilde{W}_a + \varepsilon_{HJB}(x) - \frac{1}{2}\left(\nabla\zeta g R^{-1}g\varepsilon\right)^T W_c\right) \tag{A25}$$

Substituting (A17), (A25) and (35) into the derivative of Equation (A13) yields:

$$\begin{aligned}
\dot{L}(x) =& x^T\dot{x} + \tilde{W}_c^T a_1^{-1}\dot{\tilde{W}}_c + \tilde{W}_a^T a_2^{-1}\dot{\tilde{W}}_a \\
=& -Q(x) - \frac{1}{4}W_c^T\mathcal{G}(x)W_c + \frac{1}{2}W_c^T\mathcal{G}(x)\tilde{W}_a \\
&+ \varepsilon_{HJB}(x) + \varepsilon_1(x) + \tilde{W}_c^T\frac{\sigma_2}{\left(\sigma_2^T\sigma_2+1\right)^2} \\
&\times\left(-\sigma_2^T\tilde{W}_c + \frac{1}{4}\tilde{W}_a^T\mathcal{G}(x)\tilde{W}_a + \varepsilon_{HJB}(x) - \frac{1}{2}\left(\nabla\zeta g R^{-1}g\varepsilon\right)^T W_c\right) + \tilde{W}_a^T a_2^{-1}\dot{\tilde{W}}_a
\end{aligned} \tag{A26}$$

Equation (A26) can be written as:

$$\begin{aligned}\dot{L}(x) =&\dot{\hat{L}}_V + \dot{\hat{L}}_c + \varepsilon_1(x) - \tilde{W}_a^T a_2^{-1}\dot{\hat{W}}_a + \frac{1}{2}\tilde{W}_a^T\mathcal{G}(x)W_c \\ &+ \frac{1}{4}\tilde{W}_a^T\mathcal{G}(x)W_c\frac{\bar{\sigma}_2^T}{m_s}\tilde{W}_c - \frac{1}{4}\tilde{W}_a^T\mathcal{G}(x)W_c\frac{\bar{\sigma}_2^T}{m_s}W_c \\ &+ \frac{1}{4}\tilde{W}_a^T\mathcal{G}(x)\tilde{W}_a\frac{\bar{\sigma}_2^T}{m_s}W_c + \frac{1}{4}\tilde{W}_a^T\mathcal{G}(x)\hat{W}_a\frac{\bar{\sigma}_2^T}{m_s}\hat{W}_c\end{aligned} \tag{A27}$$

with the definitions:

$$\begin{cases}\bar{\sigma}_2 \triangleq \frac{\sigma_2}{\sigma_2^T\sigma_2+1} \\ m_s \triangleq \sigma_2^T\sigma_2 + 1 \\ \dot{\hat{L}}_V \triangleq -Q(x) - \frac{1}{4}W_c^T\mathcal{G}(x)W_c + \varepsilon_{HJB}(x) \\ \dot{\hat{L}}_c \triangleq \tilde{W}_c^T\frac{\sigma_2}{\left(\sigma_2^T\sigma_2+1\right)^2}\left(-\sigma_2^T\tilde{W}_c + \varepsilon_{HJB}(x) - \frac{1}{2}\left(\nabla\zeta g R^{-1}g\varepsilon\right)^T W_c\right)\end{cases} \tag{A28}$$

Taking (35) into (A27), we have the following:

$$\begin{aligned}\dot{L}(x) =& -Q(x) - \frac{1}{4}W_c^T\mathcal{G}(x)W_c + \varepsilon_{HJB}(x) + \tilde{W}_c^T\bar{\sigma}_2\left(-\bar{\sigma}_2^T\tilde{W}_c + \frac{\varepsilon_{HJB}(x) - \frac{1}{2}\left(\nabla\zeta g R^{-1}g\varepsilon\right)^T W_c}{m_s}\right) \\ &+ \varepsilon_1(x) + \frac{1}{2}\tilde{W}_a^T\mathcal{G}(x)W_c + \frac{1}{4}\tilde{W}_a^T\mathcal{G}(x)W_c\frac{\bar{\sigma}_2^T}{m_s}\tilde{W}_c \\ &- \frac{1}{4}\tilde{W}_a^T\mathcal{G}(x)W_c\frac{\bar{\sigma}_2^T}{m_s}W_c + \frac{1}{4}\tilde{W}_a^T\mathcal{G}(x)W_c\frac{\bar{\sigma}_2}{m_s}\tilde{W}_a \\ &+ \tilde{W}_a^T\eta_a W_c - \tilde{W}_a^T\eta_a\tilde{W}_a - \tilde{W}_a^T\eta_c\bar{\sigma}_2^T W_c + \tilde{W}_a^T\eta_c\bar{\sigma}_2^T\tilde{W}_c.\end{aligned} \tag{A29}$$

As established in [28], for any given γ_ε, there exists N_0 such that $\| \varepsilon_{HJB} \| < \gamma_\varepsilon$ holds when $N > N_0$.

From Assumption 2 and (A15), it follows that:

$$\| \varepsilon_1(x) \| < \gamma_{\varepsilon_x}\gamma_f \| x \| + \frac{1}{2}\gamma_{\varepsilon_x}\gamma_g^2\gamma_{\zeta_x}\sigma_{\min}^{-1}(R)\left(\| W_c \| + \| \tilde{W}_a \|\right) + \gamma_g\gamma_{\zeta_x}\delta_m \tag{A30}$$

$$\left(\nabla\zeta g\varepsilon(x)\right)^T W_c < \gamma_\varepsilon\gamma_g\gamma_{\zeta_x}W_c \tag{A31}$$

where $\sigma_{\min}^{-1}(R)$ denotes the reciprocal of the minimum singular value of matrix R.

Let ϵ be a positive definite constant satisfying the inequality: $x^T\epsilon x < Q(x)$ and define $\Omega \triangleq \begin{bmatrix} x \\ \bar{\sigma}_2^T\tilde{W}_c \\ \tilde{W}_a \end{bmatrix}$, we obtain:

$$\begin{aligned}\dot{L} <& \frac{1}{4}\| W_c \|^2\| \mathcal{G}(x) \| + \gamma_\varepsilon + \frac{1}{2}\| W_c \|\gamma_{\varepsilon_x}\gamma_{\zeta_x}\gamma_g^2\sigma_{\min}^{-1}(R) + \gamma_g\gamma_{\zeta_x}\delta_m \\ &- \Omega^T\begin{bmatrix}\epsilon I & 0 & 0 \\ 0 & I & \left(-\frac{1}{2}\eta_c - \frac{1}{8m_s}\mathcal{G}W_c\right)^T \\ 0 & -\frac{1}{2}\eta_c - \left(\frac{1}{8m_s}\mathcal{G}W_c\right) & \eta_a - \frac{1}{8}\left(\mathcal{G}W_c m^T + mW_c^T\mathcal{G}\right)\end{bmatrix}\Omega \\ &+ \Omega^T\begin{bmatrix}\gamma_{\varepsilon_x}\gamma_f \\ \frac{2\gamma_\varepsilon + \gamma_\varepsilon\gamma_g^2\gamma_{\zeta_x}W_c\sigma_{\min}^{-1}(R)}{2m_s} \\ \left(\frac{1}{2}\mathcal{G} + \eta_a - \eta_c\bar{\sigma}_2^T - \frac{1}{4}\mathcal{G}W_c m^T\right)W_c + \frac{1}{2}\gamma_{\varepsilon_x}\gamma_g^2\gamma_{\zeta_x}\sigma_{\min}^{-1}(R)\end{bmatrix}\end{aligned} \tag{A32}$$

Define:

$$\begin{cases} H \triangleq \begin{bmatrix} \epsilon I & 0 & 0 \\ 0 & I & \left(-\frac{1}{2}\eta_c - \frac{1}{8m_s}\mathcal{G}W_c\right)^T \\ 0 & -\frac{1}{2}\eta_c - \left(\frac{1}{8m_s}\mathcal{G}W_c\right) & \eta_a - \frac{1}{8}\left(\mathcal{G}W_c m^T + mW_c^T\mathcal{G}\right) \end{bmatrix} \\ \Gamma \triangleq \begin{bmatrix} \gamma_{\varepsilon_x}\gamma_f \\ \frac{2\gamma_\varepsilon + \gamma_\varepsilon\gamma_g^2\gamma_{\zeta_x}W_c\sigma_{\min}^{-1}(R)}{2m_s} \\ \left(\frac{1}{2}\mathcal{G} + \eta_a - \eta_c\bar{\sigma}_2^T - \frac{1}{4}\mathcal{G}W_c m^T\right)W_c + \frac{1}{2}\gamma_{\varepsilon_x}\gamma_g^2\gamma_{\zeta_x}\sigma_{\min}^{-1}(R) \end{bmatrix} \\ \rho \triangleq \frac{1}{4} \parallel W_c \parallel^2 \parallel \mathcal{G}(x) \parallel + \gamma_\varepsilon + \frac{1}{2} \parallel W_c \parallel \gamma_{\varepsilon_x}\gamma_{\zeta_x}\gamma_g^2\sigma_{\min}^{-1}(R) + \gamma_g\gamma_{\zeta_x}\delta_m \end{cases} \tag{A33}$$

With proper selection of η_c and η_a guaranteeing $H > 0$, the following holds:

$$\dot{L} < - \parallel \Omega \parallel^2 \sigma_{\min}(H) + \parallel \Gamma \parallel \parallel \Omega \parallel + \rho + \gamma_\varepsilon \tag{A34}$$

Based on (A34), we conclude that $\dot{L}$ is negative definite when:

$$\parallel \Omega \parallel > \frac{\parallel \Gamma \parallel}{2\sigma_{\min}(H)} + \sqrt{\frac{\Gamma^2}{4\sigma_{\min}^2(H)} + \frac{\rho + \gamma_\varepsilon}{\sigma_{\min}(H)}} \tag{A35}$$

Thus, all signals in the closed-loop system are uniformly ultimately bounded.

References

1. Wang, M.; Ren, X.; Chen, Q. Cascade Optimal Control for Tracking and Synchronization of a Multimotor Driving System. *IEEE Trans. Control Syst. Technol.* **2019**, *27*, 1376–1384. [CrossRef]
2. Errouissi, R.; Al-Durra, A.; Muyeen, S.M. Experimental Validation of a Novel PI Speed Controller for AC Motor Drives with Improved Transient Performances. *IEEE Trans. Control Syst. Technol.* **2018**, *26*, 1414–1421. [CrossRef]
3. Lee, H.; Lee, Y.; Shin, D.; Chung, C.C. H∞ control based on LPV for load torque compensation of PMSM. In Proceedings of the 2015 15th International Conference on Control, Automation and Systems (ICCAS), Busan, Republic of Korea, 13–16 October 2015; pp. 1013–1018. [CrossRef]
4. Zhang, X.; Sun, L.; Zhao, K.; Sun, L. Nonlinear Speed Control for PMSM System Using Sliding-Mode Control and Disturbance Compensation Techniques. *IEEE Trans. Power Electron.* **2013**, *28*, 1358–1365. [CrossRef]
5. Repecho, V.; Biel, D.; Arias, A. Fixed Switching Period Discrete-Time Sliding Mode Current Control of a PMSM. *IEEE Trans. Ind. Electron.* **2018**, *65*, 2039–2048. [CrossRef]
6. Linares-Flores, J.; García-Rodríguez, C.; Sira-Ramírez, H.; Ramírez-Cárdenas, O.D. Robust Backstepping Tracking Controller for Low-Speed PMSM Positioning System: Design, Analysis, and Implementation. *IEEE Trans. Ind. Inform.* **2015**, *11*, 1130–1141. [CrossRef]
7. Yin, W.; Wu, X.; Rui, X. Adaptive Robust Backstepping Control of the Speed Regulating Differential Mechanism for Wind Turbines. *IEEE Trans. Sustain. Energy* **2019**, *10*, 1311–1318. [CrossRef]
8. Li, S.; Zhou, M.; Yu, X. Design and Implementation of Terminal Sliding Mode Control Method for PMSM Speed Regulation System. *IEEE Trans. Ind. Inform.* **2013**, *9*, 1879–1891. [CrossRef]
9. Preindl, M.; Bolognani, S. Model Predictive Direct Speed Control with Finite Control Set of PMSM Drive Systems. *IEEE Trans. Power Electron.* **2013**, *28*, 1007–1015. [CrossRef]
10. Li, S.; Ding, L.; Gao, H.; Liu, Y.J.; Huang, L.; Deng, Z. ADP-Based Online Tracking Control of Partially Uncertain Time-Delayed Nonlinear System and Application to Wheeled Mobile Robots. *IEEE Trans. Cybern.* **2020**, *50*, 3182–3194. [CrossRef]
11. Xue, S.; Zhao, N.; Zhang, W.; Luo, B.; Liu, D. A Hybrid Adaptive Dynamic Programming for Optimal Tracking Control of USVs. *IEEE Trans. Neural Netw. Learn. Syst.* **2025**, *36*, 9961–9969. [CrossRef] [PubMed]
12. Yu, Y.; Ma, X.; Su, R.; Jet, T.K.; Viswanathan, V.; Gajanayake, C.J.; RamaKrishna, S.; Gupta, A.K. Application of integral reinforcement learning for optimal control of a high speed flux-switching permanent magnet machine. In Proceedings of the IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society, Florence, Italy, 23–26 October 2016; pp. 2702–2707. [CrossRef]
13. Chen, G.; Wang, W.; Dong, J. Performance-Optimize Adaptive Robust Tracking Control for USV-UAV Heterogeneous Systems with Uncertainty. *IEEE Trans. Veh. Technol.* **2025**, *74*, 7251–7262. [CrossRef]

14. Chen, G.; Dong, J. Approximate Optimal Adaptive Prescribed Performance Control for Uncertain Nonlinear Systems with Feature Information. *IEEE Trans. Syst. Man Cybern. Syst.* **2024**, *54*, 2298–2308. [CrossRef]
15. Jiang, Y.; Jiang, Z.P. Robust Adaptive Dynamic Programming with an Application to Power Systems. *IEEE Trans. Neural Netw. Learn. Syst.* **2013**, *24*, 1150–1156. [CrossRef]
16. Heydari, A. Optimal Impulsive Control Using Adaptive Dynamic Programming and its Application in Spacecraft Rendezvous. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *32*, 4544–4552. [CrossRef]
17. El-Sousy, F.F.M.; Amin, M.M.; Al-Durra, A. Adaptive Optimal Tracking Control Via Actor-Critic-Identifier Based Adaptive Dynamic Programming for Permanent-Magnet Synchronous Motor Drive System. *IEEE Trans. Ind. Appl.* **2021**, *57*, 6577–6591. [CrossRef]
18. Lee, J.; You, S.; Kim, W.; Moon, J. Extended state observer-actor–critic architecture based output-feedback optimized backstepping control for permanent magnet synchronous motors. *Expert Syst. Appl.* **2025**, *270*, 126542. [CrossRef]
19. Fan, Z.X.; Li, S.; Liu, R. ADP-Based Optimal Control for Systems with Mismatched Disturbances: A PMSM Application. *IEEE Trans. Circuits Syst. II Express Briefs* **2023**, *70*, 2057–2061. [CrossRef]
20. Fan, Z.X.; Li, S.; Su, J. Adaptive Dynamic Programming for PMSM Control Under Safety, Robustness, and Optimality Constraints. *IEEE Trans. Syst. Man Cybern. Syst.* **2025**, *55*, 2724–2733. [CrossRef]
21. Wang, Z.; Ye, H.; Wang, Y.; Shi, Y.; Liang, L. Optimal Output-Feedback Controller Design Using Adaptive Dynamic Programming: A Permanent Magnet Synchronous Motor Application. *IEEE Trans. Circuits Syst. II Express Briefs* **2025**, *72*, 208–212. [CrossRef]
22. Vamvoudakis, K.G.; Lewis, F.L. Online actor–critic algorithm to solve the continuous-time infinite horizon optimal control problem. *Automatica* **2010**, *46*, 878–888. [CrossRef]
23. Uddin, M.N.; Zou, H.; Azevedo, F. Online Loss-Minimization-Based Adaptive Flux Observer for Direct Torque and Flux Control of PMSM Drive. *IEEE Trans. Ind. Appl.* **2016**, *52*, 425–431. [CrossRef]
24. Ye, S.; Yao, X. A Modified Flux Sliding-Mode Observer for the Sensorless Control of PMSMs with Online Stator Resistance and Inductance Estimation. *IEEE Trans. Power Electron.* **2020**, *35*, 8652–8662. [CrossRef]
25. Podder, A.; Pandit, D. Study of Sensorless Field-Oriented Control of SPMSM Using Rotor Flux Observer & Disturbance Observer Based Discrete Sliding Mode Observer. In Proceedings of the 2021 IEEE 22nd Workshop on Control and Modelling of Power Electronics (COMPEL), Cartagena, Colombia, 2–5 November 2021; pp. 1–8. [CrossRef]
26. Zhu, G.; Dessaint, L.A.; Akhrif, O.; Kaddouri, A. Speed tracking control of a permanent-magnet synchronous motor with state and load torque observer. *IEEE Trans. Ind. Electron.* **2000**, *47*, 346–355. [CrossRef]
27. Apte, A.; Joshi, V.A.; Mehta, H.; Walambe, R. Disturbance-Observer-Based Sensorless Control of PMSM Using Integral State Feedback Controller. *IEEE Trans. Power Electron.* **2020**, *35*, 6082–6090. [CrossRef]
28. Hornik, K.; Stinchcombe, M.; White, H. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural Netw.* **1990**, *3*, 551–560. [CrossRef]
29. Li, J.; Liu, H.; Zhang, Z.; Li, X.; Yang, X. Event-triggered adaptive NN tracking control with dynamic gain for a class of unknown nonlinear systems. *Neurocomputing* **2022**, *467*, 292–299. [CrossRef]
30. Abu-Khalaf, M.; Lewis, F.L. Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach. *Automatica* **2005**, *41*, 779–791. [CrossRef]
31. Modares, H.; Lewis, F.L. Optimal tracking control of nonlinear partially-unknown constrained-input systems using integral reinforcement learning. *Automatica* **2014**, *50*, 1780–1792. [CrossRef]
32. Li, D.; Dong, J. Approximate Optimal Robust Tracking Control Based on State Error and Derivative Without Initial Admissible Input. *IEEE Trans. Syst. Man Cybern. Syst.* **2024**, *54*, 1059–1069. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.