


Article

Neural Adaptive H_∞ Sliding-Mode Control for Uncertain Nonlinear Systems with Disturbances Using Adaptive Dynamic Programming

Yuzhu Huang * and Zhaoyan Zhang 

College of Electronic and Information Engineering, Hebei University, Baoding 071002, China; zhangzhaoyan@hbu.edu.cn

* Correspondence: hyzweed@163.com

Abstract: This paper focuses on a neural adaptive H_∞ sliding-mode control scheme for a class of uncertain nonlinear systems subject to external disturbances by the aid of adaptive dynamic programming (ADP). First, by combining the neural network (NN) approximation method with a nonlinear disturbance observer, an enhanced observer framework is developed for estimating the system uncertainties and observing the external disturbances simultaneously. Then, based on the reliable estimations provided by the enhanced observer, an adaptive sliding-mode controller is meticulously designed, which can effectively counteract the effects of the system uncertainties and the separated matched disturbances, even in the absence of prior knowledge regarding their upper bounds. While the remaining unmatched disturbances are attenuated by means of H_∞ control performance on the sliding surface. Moreover, a single critic network-based ADP algorithm is employed to learn the cost function related to the Hamilton–Jacobi–Isaacs equation, and thus, the H_∞ optimal control is obtained. An updated law for the critic NN is proposed not only to make the Nash equilibrium achieved, but also to stabilize the sliding-mode dynamics without the need for an initial stabilizing control. In addition, we analyze the uniform ultimate boundedness stability of the resultant closed-loop system via Lyapunov’s method. Finally, the effectiveness of the proposed scheme is verified through simulations of a single-link robot arm and a power system.

Keywords: adaptive dynamic programming; disturbance observer; neural networks; optimal control; sliding-mode control



Citation: Huang, Y.; Zhang, Z. Neural Adaptive H_∞ Sliding-Mode Control for Uncertain Nonlinear Systems with Disturbances Using Adaptive Dynamic Programming. *Entropy* **2023**, *25*, 1570. <https://doi.org/10.3390/e25121570>

Academic Editor: Tao Zhao

Received: 17 October 2023

Revised: 16 November 2023

Accepted: 17 November 2023

Published: 22 November 2023



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

1. Introduction

Within the last few decades, multifarious robust control design theories and methods have been proposed for uncertain nonlinear systems [1]. As one of the most efficient and widely used control methods, sliding-mode control (SMC) has garnered significant attention by reason of its simplicity, order reduction and inherent robustness against the matched uncertainties [2]. The classical SMC approach is to exert a discontinuous control to drive the system states onto a prescribed sliding manifold or surface [3]. As long as the sliding surface is reached, the system will become immune from the matched uncertainties and input disturbances. To remove the reaching phase, an integral SMC was developed by using the integral sliding manifold, including an integral term, which can enable the system states to reach and remain on the sliding manifold from the beginning [4–6]. Although towards a wide variety of actual systems, the relevant uncertainties and disturbances can be assumed to be matched in the design of control systems, there are also many physical systems, such as permanent magnet synchronous motors [7], underactuated aerial vehicles and robotic systems [8] directly affected by unmatched disturbances. Lately, several new approaches involving the integral SMC have been proposed to stabilize various systems with unmatched disturbances [9–13]. Among these methods, it is worth noticing that in [12,13], the impact of the separated unmatched disturbances would not be amplified

after choosing a suitable projection matrix in a sliding manifold and were attenuated by the combination of the integral SMC with H_∞ control theories. This provides a feasible and effective way to handle the unmatched disturbances and helps explore the relationships between integral SMC and H_∞ control in nonlinear system control design.

In many instances, we expect the control policy not just to make the closed-loop system stable, but to possess certain optimality by minimizing the user-defined cost. For nonlinear systems, the settlement of associated optimal control problems requires solving the Hamilton–Jacobi–Bellman (HJB) equation. While considering H_∞ optimal control, based on the dissipativity theory, it can be formulated as an L_2 -gain control problem, which involves solving the Hamilton–Jacobi–Isaacs (HJI) equation [14]. However, the analytical solutions of both HJB and HJI equations are very hard or even impossible to obtain directly because of their inherent nonlinearities [15]. In recent years, a class of neural network (NN) and reinforcement learning (RL)-based intelligent optimization and control methods, referred to as adaptive dynamic programming (ADP), is becoming more and more striking and shows great application potential in solving various optimization problems, and effectively conquers the “curse of dimensionality” [15,16]. By now, many researchers have employed ADP to tackle a variety of optimal control problems for both discrete-time (DT) [17–22] and continuous-time (CT) systems [23–28]. Moreover, how to combine ADP with other robust methods to achieve better performance and stronger robustness for uncertain nonlinear systems is becoming a new research focus [29,30].

Recently, Modares et al. [31] proposed an online integral RL algorithm that incorporates a non-quadratic discounted cost function to address the constrained-input optimal tracking problem. Luo et al. [32] described an NN-based off-policy learning algorithm within the actor-critic framework to deal with the associated HJI equation, and this algorithm was later extended to find the near-optimal H_∞ tracking control solution in [33]. Nevertheless, the influences of potential system or modeling uncertainties were not taken into account in the design. Wang et al. [34] introduced a robust neuro-optimal control approach for input-affine nonlinear systems with both matched and state-dependent uncertainties. They achieved this by redesigning the cost function and selecting a suitable feedback gain, whereas the upper bound function of uncertainties is needed for redesigning the cost function to suppress these uncertainties. Mitra et al. [35] presented an optimal SMC scheme for the single-input cascade nonlinear systems with matched bounded disturbances. Fan et al. [36] investigated an adaptive actor-critic-based integral SMC strategy for CT nonlinear systems with unknown terms and input disturbances, where the initial stabilizing control requirement in the learning was quite stringent and limiting in practical applications. Qu et al. [37] developed an adaptive H_∞ optimal SMC method in the presence of actuator faults and unmatched disturbances using the ADP algorithm, and further explored the optimal guaranteed cost SMC for constrained-input uncertain systems by formulating an auxiliary system and redefining the utility function [38]. Based on [37], combined with event-triggered mechanisms, Yang et al. [39] provided an event-triggered integral SMC design for nonlinear control-affine systems by leveraging the ADP technique. Note that these methods mentioned above rely on the availability of upper bounds for matched or unmatched disturbances, which may cause over-design and thus leads to an over-conservative control scheme. Additionally, in real-world scenarios, determining precise upper bounds of external disturbances is often a challenging task.

Inspired by the works mentioned earlier, we propose an adaptive neural H_∞ SMC scheme for uncertain nonlinear systems subject to external disturbances using the ADP algorithm. Based on the enhanced observer system composed of the NN identifier and nonlinear disturbance observer (DO), an integral SMC is developed to counteract the impacts of the system uncertainties and the separated matched disturbances, as well as unknown approximation errors, without requiring prior knowledge of their upper bounds. While on the sliding manifold, the remaining unmatched disturbances are attenuated by H_∞ optimal control solved by the single-network ADP algorithm. Moreover, the uniform

ultimate boundedness stability of the resultant closed-loop system are guaranteed via the Lyapunov approach.

The principal contributions of this study can be enumerated as follows. First, unlike other existing schemes [34–39], based on the enhanced observer system, the proposed approach makes the designed sliding-mode controller independent from the relevant upper bounds of uncertainties and disturbances, which renders the implementation much easier and more practical and removes the assumption that the upper bounds need to be known in advance. Second, compared with the algorithms presented in [36,37], our approach can deal with both unknown nonlinear terms and unmatched external disturbances, where the single-network ADP is utilized to approximate an H_∞ optimal control. Unlike typical actor–critic–disturbance network architectures, the single critic network structure may bring a simpler implementation, lower calculation amount, and avoid the numerical approximate errors arising from actor and disturbance networks. Third, we introduce an updated law for the critic NN, which not only achieves the Nash equilibrium, but also ensures the stability of the sliding-mode dynamics without the need for an initial stabilizing control in the learning.

The remainder of this paper is arranged as follows. Section 2 outlines the problem formulation and provides some necessary preliminaries. Section 3 describes the design of an integral SMC based on the enhanced observer system. Section 4 presents the application of the single-network ADP to obtain H_∞ optimal control for the sliding-mode dynamics, along with stability analysis. Simulations of the robotic arm and a power system are given in Section 5, followed by a summary of this study in Section 6.

2. Problem Formulation

Consider the following uncertain perturbed nonlinear system as

$$\dot{x} = f(x) + \Delta f(x) + (g(x) + \Delta g(x))u + d, \quad (1)$$

where the state vector $x \in \mathbb{R}^n$ is measurable, $u \in \mathbb{R}^m$ is the control input, $f(x) \in \mathbb{R}^n$ and $g(x) \in \mathbb{R}^{n \times m}$ are the known system drift and input dynamics, respectively; $\Delta f(x)$ and $\Delta g(x)$ denote uncertain nonlinear terms that refer to either the inherent characteristics of the system or modeling uncertainties, while $d \in \mathbb{R}^n$ represents the unknown external disturbances. Moreover, it is assumed that the system uncertainties $\Delta f(x)$ and $\Delta g(x)$ satisfy the matched condition, i.e., $\Delta f(x) + \Delta g(x)u = g(x)w(x, u)$, then the system (1) is rewritten in the form of

$$\dot{x} = f(x) + g(x)u + g(x)w(x, u) + d \quad (2)$$

with $w(x, u)$ being the bounded lumped uncertain term. Let $\Omega \subseteq \mathbb{R}^n$ be a compact set, and suppose that $f(x) + g(x)u$ is Lipschitz continuous over Ω with $f(0) = 0$. Besides, $d \in L_2[0, \infty]$ and its derivative \dot{d} is bounded such that $\|\dot{d}\| \leq d_M$ with $d_M > 0$. To avoid any confusion, $\|\cdot\|$ denotes the 2-norm of a vector or the Frobenius norm of a matrix hereafter, unless otherwise specified.

Assumption 1. *The input matrix $g(x)$ has a full column rank and is norm bounded with $g_M > 0$, that is, $\|g(x)\| \leq g_M$ for any x . Moreover, the resulting left pseudoinverse $g^+(x) \in \mathbb{R}^{m \times n}$ is given by $g^+(x) = (g^T(x)g(x))^{-1}g^T(x)$, which is bounded by $\|g^+(x)\| \leq b_M$, where b_M, g_M are known positive constants.*

Based on Assumption 1, d is then decomposed into the matched and unmatched components through the projection of d onto the input matrix $g(x)$ as

$$d = g(x)g^+(x)d + (I - g(x)g^+(x))d, \quad (3)$$

where I denotes an identity matrix of appropriate dimensions, and $g^+(x)$ is the left pseudoinverse of $g(x)$. It should be noted that Assumption 1 is somewhat restrictive, which may lessen the applicability scope of the proposed approach to some extent. However, many

real-world physical systems, such as the satellite dynamics, the hypersonic flight vehicle and overhead crane systems, have such a property to make this assumption valid [15,20].

To deal with the uncertain nonlinear system (1) with external disturbances, an enhanced observer system is first constructed for estimating the uncertain terms and observing the unknown disturbances simultaneously. Then, based on the reliable estimations, an integral SMC is developed to counteract the impacts of the system uncertainties and the separated matched disturbances, as well as unknown approximation errors, without requiring prior knowledge of their upper bounds. Meanwhile, the remaining unmatched disturbances are attenuated by H_∞ optimal control on the sliding surface. Moreover, the single-network ADP algorithm is employed to learn the cost function related to the Hamilton–Jacobi–Isaacs equation, and then, the H_∞ optimal control is obtained. What is more, a weight updating law is formulated to ensure both the achievement of Nash equilibrium and the stabilization of sliding-mode dynamics during the learning process.

3. Integral SMC Design Based on the Enhanced Observer System

Recalling the NN universal approximation property, the uncertain term $w(x, u)$ can be represented by a three-layered NN as

$$w(x, u) = W_o^T \sigma(V_o^T \bar{x}) + \varepsilon_o(x), \quad (4)$$

where $W_o \in \mathbb{R}^{l_o \times m}$ and $V_o \in \mathbb{R}^{(n+m) \times l_o}$ denote unknown ideal weight matrices between the output and hidden, and hidden and input layers, respectively; $\bar{x} = [x^T, u^T]^T \in \mathbb{R}^{n+m}$ is the NN input, $\sigma(\cdot) \in \mathbb{R}^{l_o}$ represents the activation function with l_o hidden layer neurons, and $\varepsilon_o(x) \in \mathbb{R}^m$ stands for the NN reconstruction error. To simplify the learning process, only the weights of W_o are adapted online, while V_o is an initialized set with random values and then remains unchanged during the weight updating process [16].

The NN identifier is designed by

$$\dot{\hat{x}} = A\hat{x} + f(x) - Ax + g(x)u + g(x)\hat{W}_o^T \sigma(z) + d, \quad (5)$$

where A is a Hurwitz matrix, \hat{x} is the identifier state, \hat{W}_o is the estimate of W_o , and the activation function $\sigma(z) = \sigma(V_o^T \bar{x})$ with $z = V_o^T \bar{x}$. Since the unknown disturbance term d is needed in (5), inspired by [11], a nonlinear DO is introduced for obtaining \hat{d} , namely, the estimated value of d .

Then, combining the NN identifier with a nonlinear DO, an enhanced observer system is constructed as

$$\begin{aligned} \dot{\hat{x}} &= A\hat{x} + f(x) - Ax + g(x)u + g(x)\hat{W}_o^T \sigma(z) + \hat{d} \\ \dot{\hat{d}}_0 &= -l(x)(f(x) + g(x)u + g(x)\hat{W}_o^T \sigma(z) + d_0 + p(x)), \end{aligned} \quad (6)$$

with $\hat{d} = d_0 + p(x)$, where d_0 is an auxiliary variable, and $p(x)$ is a designed state-dependent function and brings out the gain function $l(x)$ such that $l(x) = (\partial p(x)/\partial x)^T$. Following (6), we have

$$\dot{\hat{d}} = -l(x)\hat{d} + l(x)d + l(x)g(x)\tilde{W}_o^T \sigma(z) + l(x)g(x)\varepsilon_o(x), \quad (7)$$

where $\tilde{W}_o = W_o - \hat{W}_o$ represents the NN weight estimation error. Let $\tilde{x} = x - \hat{x}$ and $\tilde{d} = d - \hat{d}$ be the state and disturbance estimation errors, respectively. Subtracting (5) from (2) and combining with (7), we obtain the coupled error dynamics of (6) as follows:

$$\begin{aligned} \dot{\tilde{x}} &= A\tilde{x} + g(x)\tilde{W}_o^T \sigma(z) + \tilde{d} + g(x)\varepsilon_o(x), \\ \dot{\tilde{d}} &= -l(x)\tilde{d} - l(x)g(x)\tilde{W}_o^T \sigma(z) + \dot{\tilde{d}} - l(x)g(x)\varepsilon_o(x). \end{aligned} \quad (8)$$

Before proceeding, we introduce a common assumption for stability analysis [15,16].

Assumption 2. For the identifier NN, there are known positive constants $\sigma_M, \epsilon_M, W_M$ and V_M in the sense that $\|\sigma(z)\| \leq \sigma_M, \|\epsilon_o(x)\| \leq \epsilon_M, \|W_o\| \leq W_M$ and $\|V_o\| \leq V_M$, respectively.

Lemma 1. Considering the system (2) and the coupled error dynamics (8), let the identifier NN weight \hat{W}_o be updated by

$$\dot{\hat{W}}_o = -\eta_1 \sigma(z) \tilde{x}^T A^{-1} g(x) - \eta_2 (\|\tilde{x}\| + 1) \hat{W}_o, \tag{9}$$

where η_1, η_2 are the positive updating ratios. Moreover, we select parameter matrices A, P and gain function $l(x)$ to satisfy

$$P^T P - l(x) - l^T(x) + l(x)g(x)g^T(x)l^T(x) \leq -\rho I \tag{10}$$

with $\rho > 0$. Then all the estimation errors \tilde{x}, \tilde{d} , and \tilde{W}_o are uniformly ultimately bounded (UUB).

Proof. Consider the Lyapunov function candidate given by

$$L_1 = \frac{1}{2} \tilde{x}^T P \tilde{x} + \frac{1}{2} \tilde{d}^T \tilde{d} + \frac{1}{2} \text{tr}\{\tilde{W}_o^T \tilde{W}_o\}, \tag{11}$$

where $L_{11} = \tilde{x}^T P \tilde{x} / 2 + \tilde{d}^T \tilde{d} / 2, L_{12} = \text{tr}\{\tilde{W}_o^T \tilde{W}_o\} / 2$, and $P = P^T$ is positive definite, which together with some matrix $\Lambda > 0$ satisfies $A^T P + PA = -\Lambda$ for the Hurwitz matrix A . By taking the time derivative of L_{11} and substituting the coupled error dynamics (8), we can obtain

$$\begin{aligned} \dot{L}_{11} = & \frac{1}{2} \tilde{x}^T (A^T P + PA) \tilde{x} + \tilde{x}^T P \tilde{d} + \tilde{x}^T P g(x) \epsilon_o(x) + \tilde{x}^T P g(x) \tilde{W}_o^T \sigma(z) - \tilde{d}^T l(x) \\ & \times g(x) \tilde{W}_o^T \sigma(z) - \frac{1}{2} \tilde{d}^T (l(x) + l^T(x)) \tilde{d} - \tilde{d}^T l(x) g(x) \epsilon_o(x) + \tilde{d}^T \dot{\tilde{d}}. \end{aligned} \tag{12}$$

Based on Assumption 2, together with Young’s inequality, it follows:

$$\begin{aligned} \dot{L}_{11} \leq & -\frac{1}{2} \tilde{x}^T \Lambda \tilde{x} + \frac{1}{2} \tilde{x}^T \tilde{x} + \frac{1}{2} \tilde{d}^T (P^T P - l(x) - l^T(x) + l(x)g(x)g^T(x)l^T(x)) \tilde{d} \\ & + \tilde{d}^T \dot{\tilde{d}} + \tilde{x}^T P g(x) \tilde{W}_o^T \sigma(z) + \tilde{x}^T P g(x) \epsilon_o(x) + \sigma_M^2 \|\tilde{W}_o\|^2 + \epsilon_M^2. \end{aligned} \tag{13}$$

Considering (10), (13) is rewritten as

$$\begin{aligned} \dot{L}_{11} \leq & -\frac{1}{2} \tau \tilde{x}^T \tilde{x} - \frac{1}{2} \rho \tilde{d}^T \tilde{d} + \tilde{x}^T P g(x) \tilde{W}_o^T \sigma(z) + \tilde{x}^T P g(x) \epsilon_o(x) \\ & + \tilde{d}^T \dot{\tilde{d}} + \sigma_M^2 \|\tilde{W}_o\|^2 + \epsilon_M^2, \end{aligned} \tag{14}$$

where $\tau = \lambda_{\min}(\Lambda) - 1 > 0$ ensured by properly selecting positive definite matrix Λ and its minimum eigenvalue $\lambda_{\min}(\Lambda)$.

Combining with (9), \dot{L}_{12} is derived as

$$\dot{L}_{12} = \text{tr}\{\eta_1 \tilde{W}_o^T \sigma(z) \tilde{x}^T A^{-1} g(x) + \eta_2 \tilde{W}_o^T \|\tilde{x}\| \hat{W}_o + \eta_2 \tilde{W}_o^T \hat{W}_o\}. \tag{15}$$

With the inequality $\text{tr}\{\tilde{W}_o^T \hat{W}_o\} \leq \|W_o\|^2 / 2 - \|\tilde{W}_o\|^2 / 2$, (15) becomes

$$\dot{L}_{12} \leq \text{tr}\{\eta_1 \tilde{W}_o^T \sigma(z) \tilde{x}^T A^{-1} g(x)\} + \text{tr}\{\eta_2 \tilde{W}_o^T \|\tilde{x}\| \hat{W}_o\} + \frac{\eta_2}{2} \|W_o\|^2 - \frac{\eta_2}{2} \|\tilde{W}_o\|^2.$$

Note that the relationship $\text{tr}\{A^T B\} = B^T A$ for all $A \in \mathbb{R}^n, B \in \mathbb{R}^n$ and the inequality $\text{tr}\{\tilde{W}_o^T (W_o - \tilde{W}_o)\} \leq W_M \|\tilde{W}_o\| - \|\tilde{W}_o\|^2$, we can have

$$\begin{aligned} \dot{L}_{12} \leq & \eta_1 \sigma_M g_M \|\tilde{x}\| \|A^{-1}\| \|\tilde{W}_o\| + \eta_2 W_M \|\tilde{x}\| \|\tilde{W}_o\| - \eta_2 \|\tilde{x}\| \|\tilde{W}_o\|^2 + \frac{\eta_2}{2} \|W_o\|^2 \\ & - \frac{\eta_2}{2} \|\tilde{W}_o\|^2. \end{aligned} \tag{16}$$

By combining (14) and (16) and taking their norms, one can derive an upper bound for $\dot{L}_1(t)$ as

$$\begin{aligned} \dot{L}_1 \leq & -\frac{1}{2} \tau \|\tilde{x}\|^2 + \left(g_M \varepsilon_M \|P\| + (g_M \sigma_M \|P\| + \eta_1 g_M \sigma_M \|A^{-1}\| + \eta_2 W_M) \|\tilde{W}_o\| - \eta_2 \right. \\ & \left. \times \|\tilde{W}_o\|^2 \right) \|\tilde{x}\| - \frac{1}{2} \rho \|\tilde{d}\|^2 + d_M \|\tilde{d}\| + \frac{\eta_2}{2} \|W_o\|^2 - \frac{\eta_2 - 2\sigma_M^2}{2} \|\tilde{W}_o\|^2 + \varepsilon_M^2. \end{aligned} \tag{17}$$

Select $\eta_2 \geq 2\sigma_M^2$ and complete the square with respect to $\|\tilde{W}_o\|$, then (17) becomes

$$\begin{aligned} \dot{L}_1 \leq & -\frac{1}{2} \tau \|\tilde{x}\|^2 - \frac{1}{2} \rho \|\tilde{d}\|^2 + \left(g_M \varepsilon_M \|P\| - \eta_2 (\|\tilde{W}_o\| - \Theta_1)^2 + \eta_2 \|\Theta_1\|^2 \right) \|\tilde{x}\| \\ & + d_M \|\tilde{d}(t)\| + \Theta_2, \end{aligned} \tag{18}$$

where

$$\Theta_1 = \frac{g_M \sigma_M \|P\| + \eta_1 g_M \sigma_M \|A^{-1}\| + \eta_2 W_M}{2\eta_2}, \quad \Theta_2 = \frac{\eta_2 \|W_o\|^2 + 2\varepsilon_M^2}{2}.$$

Define

$$e_{xd} = \begin{bmatrix} \tilde{x} \\ \tilde{d} \end{bmatrix}, \quad E_o = \frac{1}{2} \begin{bmatrix} \tau I & 0 \\ 0 & \rho I \end{bmatrix}$$

and $B_o = [g_M \varepsilon_M \|P\| + \eta_2 \|\Theta_1\|^2, d_M]$, we can further derive

$$\dot{L}_1 \leq -\lambda_{\min}(E_o) \|e_{xd}\|^2 + \|B_o\| \|e_{xd}\| + \Theta_2. \tag{19}$$

Therefore, we can conclude that $\dot{L}_1 < 0$ only if $\|e_{xd}(t)\|$ satisfies

$$\|e_{xd}\| > \frac{\|B_o\|}{2\lambda_{\min}(E_o)} + \sqrt{\frac{\|B_o\|^2}{4\lambda_{\min}^2(E_o)} + \frac{\Theta_2}{\lambda_{\min}(E_o)}}.$$

Furthermore, according to the Lyapunov extension theorem [16], when the inequality (10) holds by selecting proper matrices, we can infer that all the estimation errors \tilde{x} , \tilde{d} , and \tilde{W}_o are UUB. □

Remark 1. The gain function matrix $l(x)$ is an important design parameter that can be chosen as linear or nonlinear functions. When the form of system function $g(x)$ is simple, it can be easy to find the function $l(x)$ that satisfies the inequality (10) by substituting appropriate functions into (10). However, if the form of system function $g(x)$ is complex, the trial and error method is employed to select appropriate function $l(x)$ that meets the inequality (10). Although there is no universal design procedure for designing $l(x)$, experience has shown that it is not difficult to find a suitable $l(x)$ for specific applications [36,37].

To effectively handle both system uncertainties and external disturbances, we propose a compound H_∞ optimal SMC scheme that combines the integral SMC with H_∞ control theories. This compound controller is formulated as

$$u = u_d + u_c, \tag{20}$$

where u_d represents the discontinuous control designed to steer the system trajectories towards and maintain them on the sliding surface, thereby eliminating the effects of

matched uncertainties and disturbances. u_c denotes the continuous control derived to guarantee the system stability and achieve near-optimal performance under the remaining unmatched disturbances on sliding surfaces.

Accordingly, we define the integral sliding surface as follows:

$$s(x) = S_0(x) - S_0(x_0) - \int_0^t G(x)(f(x) + g(x)u_c)dv, \tag{21}$$

where x_0 denotes the initial state, $S_0(x) \in \mathbb{R}^m$ and $G(x) = \partial S_0(x)/\partial x \in \mathbb{R}^{m \times n}$. Moreover, it follows from Assumption 1 that a suitable matrix $G(x)$ can be found such that the product $G(x)g(x)$ is invertible.

Taking the time derivative of $s(x)$ as

$$\dot{s}(x) = G(x)(g(x)u_d + g(x)w(x, u) + d). \tag{22}$$

By incorporating the valid estimators \hat{d} and \hat{W}_o , u_d is devised as

$$u_d = - (G(x)g(x))^{-1} \left(G(x)\hat{d} + G(x)g(x)\hat{W}_o^T \sigma(z) + \mu \text{sgn}(s) + \frac{G(x)G^T(x)s}{\|s^T G(x)\|} \zeta \right), \tag{23}$$

where $\mu > 0$, $\text{sgn}(s) \in \mathbb{R}^m$ is the sign function, and $\zeta \in \mathbb{R}$ is generated by

$$\dot{\zeta} = \kappa \|s^T G(x)\| \tag{24}$$

with $\kappa > 0$. In particular, it is noted that ζ is designed to tackle the unknown bounds of the approximation errors arisen from the estimated terms \hat{d} and \hat{W}_o .

Considering the specific implementation of \hat{d} and \hat{W}_o in (23), we define $\zeta_e = \hat{d} + g(x)\hat{W}_o\sigma(z) + g(x)\varepsilon_o(x)$ to represent the approximation errors. Based on the previous analysis and the boundedness of $g(x)$, ζ_e is bounded as $\|\zeta_e\| \leq \zeta_M$ for an unknown positive constant ζ_M . To estimate ζ_M , we design ζ as defined in (24), and the estimation error is calculated as $\tilde{\zeta} = \zeta_M - \zeta$.

Theorem 1. *Considering system (2) with the sliding surface (21), the discontinuous control u_d is devised by (23) with the adaptive law (24), then it can guarantee the convergence of the sliding surface s to zero from the beginning.*

Proof. Choose the positive definite Lyapunov function candidate as

$$L_s = \frac{1}{2}s^T s + \frac{1}{2}\kappa^{-1}\tilde{\zeta}^2.$$

Along with the system (2), $\dot{L}_s(t)$ is derived as

$$\begin{aligned} \dot{L}_s &= s^T \dot{s} - \kappa^{-1}\tilde{\zeta}\dot{\zeta} \\ &= s^T G(x)(g(x)u_d + g(x)w(x, u) + d) - \kappa^{-1}\tilde{\zeta}\dot{\zeta}. \end{aligned} \tag{25}$$

Substituting (23) and (24) into (25), we can have

$$\begin{aligned} \dot{L}_s &= s^T G(x)(\hat{d} + g(x)\hat{W}_o^T \sigma(z) + g(x)\varepsilon_o(x)) - \mu s^T \text{sgn}(s) - \|s^T G(x)\|\zeta - \tilde{\zeta}\|s^T G(x)\| \\ &= s^T G(x)\zeta_e - \mu s^T \text{sgn}(s) - \|s^T G(x)\|\zeta - \tilde{\zeta}\|s^T G(x)\|. \end{aligned}$$

Using $\|\zeta_e\| \leq \zeta_M$ and the estimation error $\tilde{\zeta} = \zeta_M - \zeta$ yields

$$\begin{aligned} \dot{L}_s &\leq \|s^T G(x)\|\zeta_M - \|s^T G(x)\|\zeta - \tilde{\zeta}\|s^T G(x)\| - \mu s^T \text{sgn}(s) \\ &\leq -\mu s^T \text{sgn}(s). \end{aligned} \tag{26}$$

Thus, it is shown from (26) that $\dot{L}_s \leq -\mu \|s\|_1 < 0$ for any $s \neq 0$, where $\|s\|_1$ denotes the vector 1-norm. This means the asymptotic stability and convergence of sliding mode motion $s(x) = 0$ can be guaranteed. Moreover, according to (21), the sliding surface $s(x_0) = 0$ when $t = 0$, which implies that the system states start on the sliding surface, thus avoiding the need for a separate reaching phase. \square

From Theorem 1, it is clear that the stable sliding motion $s(x) = 0$ exists from the initial time; that is, for all $t \geq 0$, $s(x) = 0$ and $\dot{s}(x) = 0$. Moreover, the equivalent control method is utilized to obtain the sliding-mode dynamics. Combining $\dot{s}(x) = 0$ with (3) and (22), the equivalent control can be derived as

$$u_{\text{deq}} = -(G(x)g(x))^{-1}G(x)(I - g(x)g^+(x))d - g^+(x)d - w(x, u). \tag{27}$$

Then, substitute u_{deq} into (2), the sliding-mode dynamics without matched uncertain term and disturbance component is

$$\dot{x} = f(x) + g(x)u_c + \Gamma(x)d_u, \tag{28}$$

where $\Gamma(x) = I - g(x)(G(x)g(x))^{-1}G(x)$, $d_u = (I - g(x)g^+(x))d$ is the unmatched component of the external disturbance in (3). In order to reduce the influence of multiplier matrix $\Gamma(x)$ and minimize the unmatched disturbance $\Gamma(x)d_u$, an optimal projection matrix $G^*(x)$ within $\Gamma(x)$ is provided in the following Lemma.

Lemma 2. *Considering nonlinear system (2) with Assumption 1, the optimal projection matrix $G^*(x)$ is selected as $G^*(x) = g^+(x)$, which not only minimizes the norm $\|\Gamma(x)d_u\|$, but also makes the relation $\Gamma(x)d_u = d_u$ hold.*

Proof. The proof can refer to Theorem 1 in [12]. \square

As a result, with the relation $\Gamma(x)d_u = d_u$, we can express (28) as

$$\dot{x} = f(x) + g(x)u_c + d_u, \tag{29}$$

which means that the discontinuous control u_d in (23) can fully counteract the impacts of the matched uncertainties and disturbances.

Notice that in (20), u_c aims not only to suppress the remaining unmatched disturbances on sliding surface, but also to achieve a near-optimal performance for sliding-mode dynamics (29). This formulation can be seen as a nonlinear H_∞ optimal control problem, which is known to be challenging to solve directly. In the following, we will demonstrate how to find an approximate H_∞ optimal control solution by using the single-network ADP algorithm.

4. H_∞ Control Design for Sliding-Mode Dynamics

Considering (3) and (29), the sliding-mode dynamics is represented as

$$\dot{x} = f(x) + g(x)u_c + k(x)d, \tag{30}$$

with $k(x) = I - g(x)g^+(x)$. Since $g(x)$ and $g^+(x)$ are bounded, it follows that the function $k(x)$ is also bounded by $\|k(x)\| \leq k_M$ with $k_M > 0$.

For attenuating the remaining unmatched disturbances $k(x)d$, the corresponding H_∞ control problem of sliding-mode dynamics is established, which aims to seek a feedback control u_c to stabilize the system and achieve L_2 -gain no larger than γ , that is,

$$\int_0^\infty (x^T Q x + u_c^T R u_c) dv \leq \gamma^2 \int_0^\infty d^T d dv, \tag{31}$$

where Q and R are positive definite matrices with appropriate dimensions, and $\gamma > 0$ refers to the level of the disturbance attenuation. Based on [32,33], by treating the disturbance d as the other system input, we can reframe the H_∞ optimal control problem for system (30) as a two-player zero-sum game with the following infinite-horizon cost function:

$$V(x) = \int_t^\infty (x^T Qx + u_c^T R u_c - \gamma^2 d^T d) dv. \tag{32}$$

Assuming that $V(x) \in C^1$, the Hamiltonian function with the associated admissible control pair (u_c, d) is defined as

$$H(x, \nabla V, u_c, d) = x^T Qx + u_c^T R u_c - \gamma^2 d^T d + (\nabla V)^T (f(x) + g(x)u_c + k(x)d) \tag{33}$$

with $\nabla V = \partial V(x)/\partial x$. From Bellman’s optimality principle, it follows that the optimal cost function $V^*(x)$ satisfies the HJI equation

$$0 = \min_{u_c} \max_d H(x, \nabla V^*, u_c, d) \tag{34}$$

with $\nabla V^* = \partial V^*(x)/\partial x$. Moreover, according to the zero-sum game theory [16], we have the following Nash condition

$$\min_{u_c} \max_d H(x, \nabla V^*, u_c, d) = \max_d \min_{u_c} H(x, \nabla V^*, u_c, d), \tag{35}$$

which ensures the existence of saddle point (u_c^*, d^*) of the HJI Equation (34). Then, applying the stationary condition, one can derive the optimal control u_c^* and worst disturbance d^* as

$$u_c^* = -\frac{1}{2} R^{-1} g^T(x) \nabla V^*, \tag{36}$$

$$d^* = \frac{1}{2\gamma^2} k^T(x) \nabla V^*. \tag{37}$$

By substituting (36) and (37) into (33), the HJI equation associated with ∇V^* becomes

$$\begin{aligned} 0 = & x^T Qx + (\nabla V^*)^T f(x) - \frac{1}{4} (\nabla V^*)^T g(x) R^{-1} g^T(x) \nabla V^* \\ & + \frac{1}{4\gamma^2} (\nabla V^*)^T k^T(x) k(x) \nabla V^*. \end{aligned} \tag{38}$$

Due to the highly nonlinear nature of the relevant HJI equation, obtaining its analytical solution is extremely difficult, if not impossible. To overcome this challenge, we propose an online optimal algorithm that learns the solution of the HJI equation and achieves H_∞ optimal control. This is accomplished through the use of single-network ADP, where only one critic network, implemented by NN, is adopted to approximate the cost function V^* related to (38). Therefore, by using the critic NN with l_c neurons, V^* is represented over a set Ω as follows:

$$V^*(x) = W_c^T \sigma_c(x) + \varepsilon_c(x) \tag{39}$$

with the ideal weight vector $W_c \in \mathbb{R}^{l_c}$ being unknown, the vector of activation functions $\sigma_c(x) \in \mathbb{R}^{l_c}$ and the reconstruction error $\varepsilon_c(x)$. Meanwhile, we have the gradient vector

$$\nabla V^* = (\nabla \sigma_c)^T W_c + \nabla \varepsilon_c \tag{40}$$

with $\nabla \sigma_c = \partial \sigma_c(x)/\partial x$ and $\nabla \varepsilon_c = \partial \varepsilon_c(x)/\partial x$.

By combining (36), (37) and (40), it is easy to get

$$u_c^* = -\frac{1}{2}R^{-1}g^T(x)((\nabla\sigma_c)^T W_c + \nabla\varepsilon_c), \tag{41}$$

$$d^* = \frac{1}{2\gamma^2}k^T(x)((\nabla\sigma_c)^T W_c + \nabla\varepsilon_c). \tag{42}$$

Substituting (41) and (42) into (33), the HJI equation becomes

$$\begin{aligned} 0 &= H(x, \nabla V^*, u_c^*, d^*) \\ &= x^T Qx + W_c^T \nabla\sigma_c f(x) - \frac{1}{4}W_c^T \nabla\sigma_c D(\nabla\sigma_c)^T W_c - \varepsilon_{\text{HJI}}, \end{aligned} \tag{43}$$

where $D = g(x)R^{-1}g^T(x) - k(x)k^T(x)/\gamma^2$, and the approximate error ε_{HJI} is defined as $\varepsilon_{\text{HJI}} = -(\nabla\varepsilon_c)^T f(x) + W_c^T \nabla\sigma_c D \nabla\varepsilon_c / 2 + (\nabla\varepsilon_c)^T D \nabla\varepsilon_c / 4$ due to the NN reconstruction error. Furthermore, taking into account $\|k(x)\| \leq k_M$ and $\|g(x)\| \leq g_M$, we can infer that there exists a positive constant D_M in the sense that $\|D\| \leq D_M$.

Because W_c in (39) is unknown, the critic NN with the estimated weights approximates the cost function in the form of

$$\hat{V}(x) = \hat{W}_c^T \sigma_c(x), \tag{44}$$

where \hat{W}_c denotes the estimated values of W_c . In addition, we can obtain

$$\nabla \hat{V} = (\nabla\sigma_c)^T \hat{W}_c. \tag{45}$$

By using (36), (37) and (45), the approximate forms of (41) and (42) are derived as

$$\hat{u}_c = -\frac{1}{2}R^{-1}g^T(x)(\nabla\sigma_c)^T \hat{W}_c, \tag{46}$$

$$\hat{d}_w = \frac{1}{2\gamma^2}k^T(x)(\nabla\sigma_c)^T \hat{W}_c. \tag{47}$$

Then, incorporating (46) and (47) into (43), we have the approximate Hamiltonian as follows:

$$H(x, \hat{W}_c, \hat{u}_c, \hat{d}_w) = x^T Qx + \hat{W}_c^T \nabla\sigma_c f(x) - \frac{1}{4}\hat{W}_c^T \nabla\sigma_c D(\nabla\sigma_c)^T \hat{W}_c. \tag{48}$$

Subtracting (43) from (48), the corresponding Hamiltonian error is defined as

$$e_c = H(x, \hat{W}_c, \hat{u}_c, \hat{d}_w) - H(x, \nabla V^*, u_c^*, d^*) = H(x, \hat{W}_c, \hat{u}_c, \hat{d}_w).$$

To effectively approximate the cost function, one needs to adjust the critic NN weight \hat{W}_c in a manner that minimizes the Hamiltonian error e_c . To this end, it is common practice to train the critic NN by minimizing the squared residual error E_c , where $E_c = e_c^T e_c / 2$. The traditional weight updating laws of critic NN based on gradient descent method can only minimize the squared error, but cannot provide any guarantee for the stability of the resulting system during the learning phase.

However, in practice, the stability is one fundamental requirement of system, and a prerequisite for achieving other higher performance. Thus, not just for minimizing the residual error, but also to guarantee the system stability and eliminate the need for an initial stabilizing control, a weight updating law is developed for the critic NN as follows:

$$\begin{aligned} \dot{\hat{W}}_c &= -\alpha \frac{\phi}{(\phi^T \phi + 1)^2} e_c + \frac{\alpha}{4} \nabla\sigma_c D(\nabla\sigma_c)^T \hat{W}_c \frac{\phi_1^T}{\phi_s} \hat{W}_c - \alpha(F_2 - F_1 \phi_1^T) \hat{W}_c \\ &\quad + \frac{\beta}{2} \Sigma(x, \hat{u}_c, \hat{d}_w) \nabla\sigma_c D \nabla J_a, \end{aligned} \tag{49}$$

where α and β are the positive updating ratios, $\phi = \nabla\sigma_c(f(x) - D(\nabla\sigma_c)^\top \hat{W}_c/2)$, $\phi_1 = \phi/(\phi^\top\phi + 1)$, $\phi_s = \phi^\top\phi + 1$, F_1 and F_2 represent design parameter matrices with suitable dimensions, $J_a(x)$ is a Lyapunov function candidate provided in Assumption 4, and the index operator $\Sigma(x, \hat{u}_c, \hat{d}_w)$ is given by

$$\Sigma(x, \hat{u}_c, \hat{d}_w) = \begin{cases} 0, & \text{if } \dot{J}_a(x) = (\nabla J_a)^\top (f(x) + g(x)\hat{u}_c + k(x)\hat{d}_w) < 0 \\ 1, & \text{otherwise} \end{cases} \quad (50)$$

with $\nabla J_a = \partial J_a(x)/\partial x$.

Remark 2. Note that in (49), the first term is designed by the normalized gradient descent method for minimizing the residual error. The second term has a well-designed form for ensuring the system’s stability, which is derived from the Lyapunov stability analysis. The last term is an additional adjustment term that works or not depends on the index operator $\Sigma(x, \hat{u}_c, \hat{d}_w)$, which is selected based on the derivative of $J_a(x)$ along the sliding-mode dynamics (30), namely, $\dot{J}_a(x) = (\nabla J_a)^\top (f(x) + g(x)\hat{u}_c + k(x)\hat{d}_w)$. Once the system dynamics may become unstable, this results in $\dot{J}_a(x) \geq 0$, then $\Sigma(x, \hat{u}_c, \hat{d}_w) = 1$ and the last term in (49) is activated. Moreover, based on the negative gradient direction of $\dot{J}_a(x)$, i.e., $-\partial((\nabla J_a)^\top (f(x) - D\nabla\sigma_c^\top \hat{W}_c/2))/\partial \hat{W}_c$, the last term is designed to reinforce the training process of the critic NN until the system dynamics become stable. This also eliminates the need for an initial stabilizing control, compared with [35–37,39], where the stabilizing control is required for initialization; however, in practical applications, finding an initial stabilizing control is quite challenging.

Remark 3. Based on [14–16], it is necessary to satisfy the persistence of excitation (PE) requirement for updating the weights of critic NN, which enhances its ability to explore the state space and is indispensable for the weights to converge to their desired ones. To fulfill the PE requirement, a probing noise is injected into the control input [15], which may cause the instability problem during the online learning. As a result, it is important to design the last term in (49) for stabilizing the resulting system, especially when the probing signal is injected.

The schematic structure of the proposed H_∞ SMC scheme is illustrated in Figure 1. As shown in Figure 1, this structure consists of two main modules: the H_∞ optimal learning module and the enhanced observer module. It should be noted that, based on the deduced sliding-mode dynamics, the learning module can operate independently. However, the original system and the observer module rely on the compound control input u , which includes the approximate H_∞ optimal control \hat{u}_c obtained from the learning module. Consequently, it is necessary to first run the learning module to obtain the approximate optimal control \hat{u}_c during the implementation process.

Considering (43), together with $\tilde{W}_c = W_c - \hat{W}_c$, (48) is represented as

$$e_c = -\tilde{W}_c^\top \nabla\sigma_c \left(f(x) - \frac{1}{2}D(\nabla\sigma_c)^\top \tilde{W}_c \right) + \frac{1}{4}\tilde{W}_c^\top \nabla\sigma_c D(\nabla\sigma_c)^\top \tilde{W}_c + \varepsilon_{HJI}. \quad (51)$$

By means of the relation $\dot{\tilde{W}}_c = -\dot{\hat{W}}_c$ and incorporating (51) into (49), we obtain

$$\begin{aligned} \dot{\tilde{W}}_c = & -\alpha \frac{\phi_1}{\phi_s} \left(\tilde{W}_c^\top \phi - \frac{1}{4}\tilde{W}_c^\top \nabla\sigma_c D(\nabla\sigma_c)^\top \tilde{W}_c - \varepsilon_{HJI} \right) - \frac{\alpha}{4} \nabla\sigma_c D(\nabla\sigma_c)^\top \tilde{W}_c \frac{\phi_1^\top}{\phi_s} \tilde{W}_c \\ & + \alpha(F_2 - F_1\phi_1^\top)\tilde{W}_c - \frac{\beta}{2}\Pi(x, \hat{u}_c, \hat{d}_w)\nabla\sigma_c D\nabla J_a. \end{aligned} \quad (52)$$

Next, the main stability theorem is presented, but before that, one basic common assumption for the critic NN is introduced [16], and the other assumption for the sliding-mode dynamics is also needed, which has been used in [34,38].

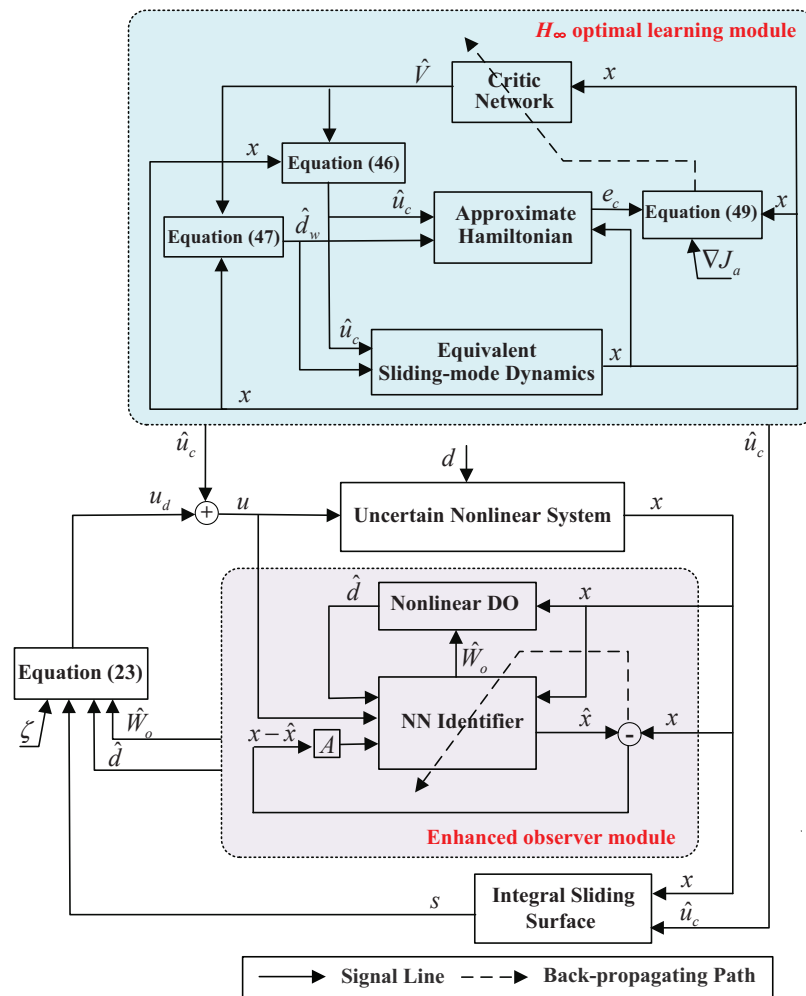


Figure 1. The schematic of the adaptive H_∞ SMC scheme.

Assumption 3. For the critic NN, there exist known positive constants σ_{cM} , σ_{dM} , ε_{cM} , ε_{dM} and W_{cM} such that $\|\sigma_c(x)\| \leq \sigma_{cM}$, $\|\nabla\sigma_c\| \leq \sigma_{dM}$, $\|\varepsilon_c(x)\| \leq \varepsilon_{cM}$, $\|\nabla\varepsilon_c\| \leq \varepsilon_{dM}$ and $\|W_c\| \leq W_{cM}$, respectively. Moreover, the approximation error ε_{HJI} is bounded above by $\varepsilon_H > 0$, namely, $\|\varepsilon_{HJI}\| \leq \varepsilon_H$.

Assumption 4. Considering the sliding-mode dynamics (30) with the optimal control pair (u_c^*, d^*) in (36) and (37), let $J_a(x)$ be a smooth, radially unbounded and positive definite Lyapunov candidate that satisfies $\dot{J}_a(x) = (\nabla J_a)^T (f(x) + g(x)u_c^* + k(x)d^*) < 0$. Moreover, it is assumed that a positive definite matrix $\Psi(x)$ makes $(\nabla J_a^*)^T \Psi(x) \nabla J_a = x^T Qx + u_c^{*T} R u_c^* - \gamma^2 d^{*T} d^*$ hold. Then, one can derive

$$(\nabla J_a)^T (f(x) + g(x)u_c^* + k(x)d^*) = -(\nabla J_a)^T \Psi(x) \nabla J_a. \tag{53}$$

Remark 4. Note that the plausibility of Assumption 4 depends on the boundedness of optimal sliding-mode dynamics, which is usually assumed to be bounded by a function of system state x . For more details, refer to [34,38]. Furthermore, it is impossible to solve (53) directly for getting the form of $J_a(x)$. Based on [34], one can obtain $J_a(x)$ by selecting an appropriate form, such as a quadratic polynomial.

Theorem 2. Considering the sliding-mode dynamics (30) and its associated cost function (32), the control input and disturbance policy are designed by (46) and (47), respectively, along with the critic weight updating law as given by (49). Then, both the sliding-mode state x and the weight

estimation error \tilde{W}_c are ensured to be UUB. Furthermore, the obtained control input \hat{u}_c can be proven to converge to a neighborhood of the optimum control u_c^* with a small adjustable bound.

Proof. Consider the following Lyapunov function candidate

$$L = \frac{1}{2} \tilde{W}_c^T \alpha^{-1} \tilde{W}_c + \beta_1 J_a(x),$$

where $\beta_1 = \beta/\alpha > 0$. By calculating the time derivative of L along the sliding-mode dynamics (30), we have

$$\dot{L} = \tilde{W}_c^T \dot{\tilde{W}}_c + \beta_1 (\nabla J_a)^T (f(x) + g(x)\hat{u}_c + k(x)\hat{d}_w). \tag{54}$$

Substituting (52) into (54) and making some adjustments, one can get

$$\begin{aligned} \dot{L} = & -\tilde{W}_c^T \phi_1 \phi_1^T \tilde{W}_c + \beta_1 (\nabla J_a)^T (f(x) + g(x)\hat{u}_c + k(x)\hat{d}_w) + \frac{1}{4} \tilde{W}_c^T \nabla \sigma_c D (\nabla \sigma_c)^T W_c \frac{\phi_1^T}{\phi_s} \tilde{W}_c \\ & - \frac{1}{4} \tilde{W}_c^T \nabla \sigma_c D (\nabla \sigma_c)^T W_c \frac{\phi_1^T}{\phi_s} W_c - \frac{1}{4} \tilde{W}_c^T \nabla \sigma_c D (\nabla \sigma_c)^T \tilde{W}_c \frac{\phi_1^T}{\phi_s} W_c \\ & + \tilde{W}_c^T \frac{\phi_1^T}{\phi_s} \varepsilon_{HJI} - \frac{\beta}{2} \Sigma(x, \hat{u}_c, \hat{d}_w) \tilde{W}_c^T \nabla \sigma_c D \nabla J_a \\ & + \tilde{W}_c^T F_2 \tilde{W}_c - \tilde{W}_c^T F_1 \phi_1^T \tilde{W}_c. \end{aligned} \tag{55}$$

Using $\hat{W}_c = W_c - \tilde{W}_c$, the last two terms in (55) become

$$\tilde{W}_c^T F_2 \hat{W}_c - \tilde{W}_c^T F_1 \phi_1^T \hat{W}_c = \tilde{W}_c^T F_2 W_c - \tilde{W}_c^T F_2 \tilde{W}_c - \tilde{W}_c^T F_1 \phi_1^T W_c + \tilde{W}_c^T F_1 \phi_1^T \tilde{W}_c. \tag{56}$$

Defining $Y = [\tilde{W}_c^T \phi_1, \tilde{W}_c^T]^T$, and substituting (55) into (56), it can be rewritten as

$$\begin{aligned} \dot{L} = & -Y^T M Y + Y^T \delta + \beta_1 (\nabla J_a)^T (f(x) + g(x)\hat{u}_c + k(x)\hat{d}_w) \\ & - \frac{\beta}{2} \Sigma(x, \hat{u}_c, \hat{d}_w) \tilde{W}_c^T \nabla \sigma_c D \nabla J_a, \end{aligned} \tag{57}$$

where

$$\begin{aligned} M = & \begin{bmatrix} I & -\frac{\nabla \sigma_c D (\nabla \sigma_c)^T W_c}{4\phi_s} - \frac{F_1}{2} \\ -\frac{\nabla \sigma_c D (\nabla \sigma_c)^T W_c}{4\phi_s} - \frac{F_1}{2} & F_2 \end{bmatrix}, \\ \delta = & \begin{bmatrix} \frac{1}{\phi_s} \varepsilon_{HJI} \\ -\frac{\nabla \sigma_c D (\nabla \sigma_c)^T W_c}{4\phi_s} + F_2 W_c - F_1 \phi_1^T W_c \end{bmatrix}. \end{aligned}$$

With Assumption 3 in mind, and recalling the boundedness of ϕ_1 and D , in particular $\|\phi_1\| < 1$ and $\|D\| \leq D_M$, we can infer that there exists a positive constant δ_M in the sense that $\|\delta\| \leq \delta_M$. For guaranteeing $M > 0$, the appropriate parameters F_1 and F_2 need to be selected in design. Then, one can upper bound \dot{L} as follows:

$$\begin{aligned} \dot{L} \leq & -\lambda_{\min}(M) \|Y\|^2 + \delta_M \|Y\| + \beta_1 (\nabla J_a)^T (f(x) + g(x)\hat{u}_c + k(x)\hat{d}_w) \\ & - \frac{\beta_1}{2} \Sigma(x, \hat{u}_c, \hat{d}_w) \tilde{W}_c^T \nabla \sigma_c D \nabla J_a \end{aligned} \tag{58}$$

with $\lambda_{\min}(M)$ being the minimum eigenvalue of M .

According to (50), there are two cases to consider: $\Sigma(x, \hat{u}_c, \hat{d}_w) = 0$ and $\Sigma(x, \hat{u}_c, \hat{d}_w) = 1$ for (58) in the following analysis.

Case 1: For $\Sigma(x, \hat{u}_c, \hat{d}_w) = 0$, it follows from (50) that $\dot{J}_a(x) < 0$, i.e., $(\nabla J_a)^T \dot{x} < 0$, which, together with the PE condition, can ensure that there exists a positive constant ϱ such that $\|\dot{Z}\| > \varrho$. This implies that $(\nabla J_a)^T \dot{x} < -\varrho \|\nabla J_a\| < 0$. Then, (58) becomes

$$\begin{aligned} \dot{L} &\leq \beta_1 (\nabla J_a)^T \dot{x} - \lambda_{\min}(M) \|Y\|^2 + \delta_M \|Y\| \\ &< -\beta_1 \varrho \|\nabla J_a\| - \lambda_{\min}(M) \left(\|Y\| - \frac{\delta_M}{2\lambda_{\min}(M)} \right)^2 + \frac{\delta_M^2}{4\lambda_{\min}(M)}. \end{aligned} \tag{59}$$

Focus on (59), only if the following inequalities:

$$\|\nabla J_a\| > \frac{\delta_M^2}{4\lambda_{\min}(M)\varrho} \triangleq \mathcal{A}_1$$

or

$$\|Y\| > \frac{\delta_M^2}{\lambda_{\min}(M)}$$

hold, then $\dot{L} < 0$. Moreover, based on the relation $\|Y\| \leq \sqrt{\|\phi_1\|^2 + 1} \|\tilde{W}_c\|$ with $\|\phi_1\| < 1$, we can derive

$$\|\tilde{W}_c\| > \frac{\delta_M^2}{\sqrt{2}\lambda_{\min}(M)} \triangleq \mathcal{B}_1.$$

Case 2: For $\Sigma(x, \hat{u}_c, \hat{d}_w) = 1$, in light of (41) and (42), by adding and subtracting $\beta_1 (\nabla J_a)^T D \nabla \varepsilon_c / 2$ into (58), we can derive

$$\begin{aligned} \dot{L} &\leq -\lambda_{\min}(M) \left(\|Y\| - \frac{\delta_M}{2\lambda_{\min}(M)} \right)^2 + \frac{\delta_M^2}{4\lambda_{\min}(M)} + \beta_1 (\nabla J_a)^T (f(x) + g(x)u_c^* \\ &\quad + k(x)d^*) + \frac{\beta_1}{2} (\nabla J_a)^T D \nabla \varepsilon_c. \end{aligned} \tag{60}$$

Then, using (53) in Assumption 4, and recalling the boundedness of D and $\nabla \varepsilon_c$, (60) is upper bounded as

$$\dot{L} \leq -\lambda_{\min}(M) \left(\|Y\| - \frac{\delta_M}{2\lambda_{\min}(M)} \right)^2 - \frac{\beta_1}{2} \lambda_{\min}(\Psi) \|\nabla J_a\|^2 + \Phi, \tag{61}$$

where $\Phi = \delta_M^2 / (4\lambda_{\min}(M)) + \beta_1 D_M^2 \varepsilon_{dM}^2 / (8\lambda_{\min}(\Psi))$, $\lambda_{\min}(\Psi)$ denotes the minimum eigenvalue of $\Psi(x)$. Hence, provided the following inequalities:

$$\|\nabla J_a\| > \sqrt{\frac{2\Phi}{\beta_1 \lambda_{\min}(\Psi)}} \triangleq \mathcal{A}_2$$

or

$$\|Y\| > \sqrt{\frac{\Phi}{\lambda_{\min}(M)}} + \frac{\delta_M}{2\lambda_{\min}(M)}$$

hold, one has $\dot{L} < 0$. Further, by the relation $\|Y\| \leq \sqrt{2} \|\tilde{W}_c\|$, we have

$$\|\tilde{W}_c\| > \sqrt{\frac{\Phi}{2\lambda_{\min}(M)}} + \frac{\delta_M}{2\sqrt{2}\lambda_{\min}(M)} \triangleq \mathcal{B}_2.$$

To sum up, for both Case 1 and Case 2, with proper parameters F_1 and F_2 satisfying $M > 0$, the inequality $\|\nabla J_a\| \geq \max\{\mathcal{A}_1, \mathcal{A}_2\} = \bar{\mathcal{A}}$ or $\|\tilde{W}_c\| \geq \max\{\mathcal{B}_1, \mathcal{B}_2\} = \bar{\mathcal{B}}$ holds, then, we have $\dot{L} < 0$. From the Lyapunov extension theorem [16], it is found that $\|\nabla J_a\|$ and $\|\tilde{W}_c\|$ are bounded by $\bar{\mathcal{A}}$ and $\bar{\mathcal{B}}$, respectively. Based on Assumption 4, the Lyapunov candidate $J_a(x)$ is radially unbounded, which implies that the boundedness of $\|\nabla J_a\|$

leads to the boundedness of the system state $\|x\|$. In particular, $\|x\|$ is bounded by $\bar{A}_x = \max\{\mathcal{A}_{1x}, \mathcal{A}_{2x}\}$, where \mathcal{A}_{1x} and \mathcal{A}_{2x} are determined by \mathcal{A}_1 and \mathcal{A}_2 , respectively. So far, we can conclude that both x and \tilde{W}_c are guaranteed to be UUB.

Next, we will prove \hat{u}_c converges to a small neighborhood of u_c^* with an adjustable bound, i.e., $\|\hat{u}_c - u_c^*\| \leq \epsilon_u$. Considering (41) and (46), we have

$$\hat{u}_c - u_c^* = -\frac{1}{2}R^{-1}g^T(x)((\nabla\sigma_c)^T\tilde{W}_c + \nabla\epsilon_c).$$

Noticing that \tilde{W}_c is UUB together with the associated bound $\bar{\mathcal{B}} = \max\{\mathcal{B}_1, \mathcal{B}_2\}$, and invoking $\|g(x)\| \leq g_M$, $\|\nabla\sigma_c\| \leq \sigma_{dM}$, $\|\nabla\epsilon_c\| \leq \epsilon_{dM}$ and boundedness of R , it follows that

$$\|\hat{u}_c - u_c^*\| \leq \frac{1}{2}\lambda_{\max}(R^{-1})g_M(\sigma_{dM}\bar{\mathcal{B}} + \epsilon_{dM}) \triangleq \epsilon_u. \tag{62}$$

□

Remark 5. From the expression of \mathcal{B}_1 and \mathcal{B}_2 , it is seen that $\bar{\mathcal{B}}$ can be kept small with $\lambda_{\min}(M)$ being larger enough. In view of (57), we can enlarge the value of $\lambda_{\min}(M)$ by adjusting the corresponding design parameters F_1 and F_2 . Moreover, we can make the approximate error ϵ_c and its upper bound ϵ_{dM} sufficiently small when the neuron number l_c is large enough. Therefore, we can make the convergence errors ϵ_u in (62) as small as possible in the design.

5. Simulation Results

To validate the effectiveness of the proposed H_∞ optimal SMC scheme, two simulation examples are provided. The first example focuses on a single-link robot arm, while the second example deals with a power system.

5.1. Single-Link Robot Arm

Considering a nonlinear single-link robot arm [23] and its dynamics given by

$$J\ddot{\theta} = -MgL\sin(\theta) - D\dot{\theta} + u + w, \tag{63}$$

where θ is the joint rotation angle of robot arm in radians, u refers to the control torque applied to the joint in Nm, and w denotes the lumped uncertain term. Select the system parameters as follows: the arm length $L = 0.5$ m, the payload mass $M = 1$ kg, the local gravity acceleration $g = 9.81$ m/s². the rotational inertia $J = 1$ kg · m² and the viscous friction $D = 2$ Nm · s/rad. With the system states defined as $x_1 = \theta$ and $x_2 = \dot{\theta}$, and considering the presence of exogenous disturbances, then the dynamics (63) in state-space form can be represented as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -4.905\sin(x_1) - 2x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}(u + w) + d, \tag{64}$$

where d represents the unknown disturbances. Moreover, it is assumed that the initial state is set as $x_0 = [1, -0.5]^T$, the lumped uncertainty term is $w(x, u) = x_2\sin(x_1) + 0.1\sin(x_1)u$, and the disturbance term is chosen as $d = [0.5e^{-t}\sin(t), 0.5\sin(t)]^T$ in the simulation.

The enhanced observer system, consisting of an NN identifier and a nonlinear DO, can be designed as shown in (6), where the identifier NN is selected as a three-layered feed-forward NN with one hidden layer containing six neurons, and the hyperbolic activation function $\tanh(\cdot)$ is utilized. The updating ratios are set as $\eta_1 = 30$ and $\eta_2 = 2.5$, while the weights \hat{W}_o and \hat{V}_o are initialized with random values chosen from the interval $[-0.1, 0.1]$. The initial observer state is set as $\hat{x}_0 = [0.5, 0]^T$. Moreover, based on Lemma 1, select the Hurwitz matrix $A = [-15, 0; 0, -15]$, $p(x) = [10x_1; 10x_2]$ and $l(x) = [10, 0; 0, 10]$ to ensure that the inequality (10) holds. The integral sliding surface function is determined by (21), together with $G(x) = g^+(x) = [0, 1]$ and $S_0(x) = x_2$. Accordingly, the discontinuous SMC

u_d is given by (23) and (24). For the propose of eliminating the chattering phenomenon, an arctangent function $\text{atan}(s/\epsilon)$ with a small positive scalar $\epsilon = 0.005$ is employed to replace the sign function $\text{sgn}(s)$ in (23).

By considering the SMC law u_d , the sliding-mode dynamics can be obtained as

$$\dot{x} = f(x) + g(x)u_c + k(x)d, \tag{65}$$

where $k(x) = I - g(x)g^+(x) = [1, 0; 0, 0]$. We choose the associated cost function as the form of (32), together with $Q = \text{diag}(1, 1)$, $R = 1$ and $\gamma = 1.5$. For the critic NN, the activation function is chosen as $\sigma_c(x) = [x_1^2, x_1x_2, x_2^2, x_1^3x_2, x_1^2x_2^2, x_1x_2^3]^T$, which results in $\hat{W}_c = [\hat{W}_{c1}, \hat{W}_{c2}, \dots, \hat{W}_{c6}]^T$. Select the updating ratios $\alpha = 1, \beta = 0.5$, the design parameters $F_1 = F_2 = 10I, l_c = 6$ and $J_a(x)$ as a quadratic polynomial. Furthermore, the weight vector \hat{W}_c is initialized to zero, which leads to the initial control input of zero. Noticing that the zero initial control cannot make the system (65) stable, it is thus clear that no initial stabilizing control strategy is necessary when implementing the proposed algorithm.

During the learning process, a damped decreasing probing noise is injected into the control input for satisfying PE condition. This noise comprises sinusoids of diverse frequencies and is applied for the first 450 s. Figure 2 shows the trajectories of the critic weights, which eventually converge to $\hat{W}_c = [1.0420, 0.0856, -0.0603, -0.2174, 0.2948, -0.0358]^T$. Figure 3 describes the trajectories of system states in the learning. From Figure 3, one can see that without an initial stabilizing control, the system states stay at or near zero after the probing noise is removed, which indicates that \hat{u}_c generated by the learning module can effectively stabilize the system. With the converged weights, the approximate H_∞ optimal control \hat{u}_c can be calculated by (46).

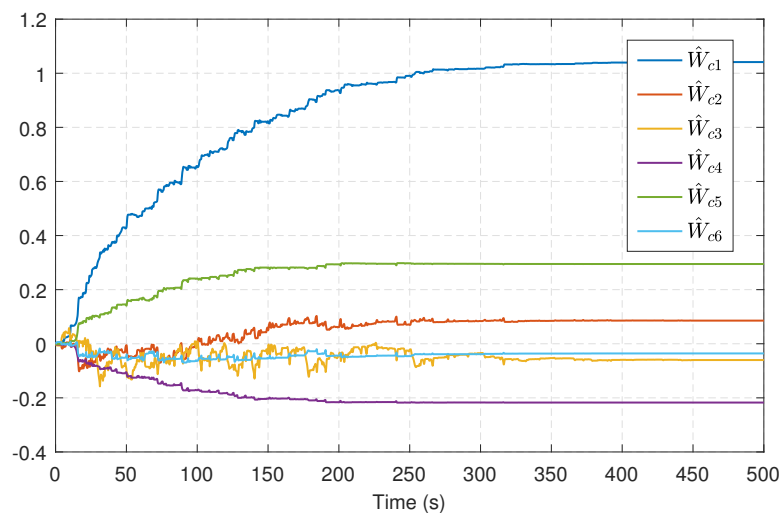


Figure 2. Trajectories of the critic NN weights.

Next, we substitute \hat{u}_c into (21) to obtain an available sliding surface. Subsequently, integrating with the enhanced observer system, the SMC law u_d is implemented by using (23) and (24) with the reliable estimations of uncertainties and disturbances. Figure 4 depicts the estimates of disturbances $d_1 = 0.5e^{-t} \sin(t)$ and $d_2 = 0.5 \sin(t)$, along with small estimation errors. Figure 5 presents the identifications of system states using the identifier NN. It can be observed that the identified states rapidly track the real states, illustrating the effectiveness and efficiency of the identifier NN. Note that the valid estimations \hat{d} and \hat{W}_o are used to design the SMC law u_d , which helps to reduce the sliding-mode gain and alleviate the chattering phenomenon. Figure 6 displays the state trajectories of the robot arm under the compound H_∞ sliding-mode control $u = u_d + \hat{u}_c$. Figure 7 depicts the compound control u , while the H_∞ control \hat{u}_c and the SMC law u_d are given in Figure 8. These results presented in Figures 6–8 confirm that the compound control u successfully

renders the robot arm system stable and exhibits satisfactory performance against both system uncertainties and external disturbances.

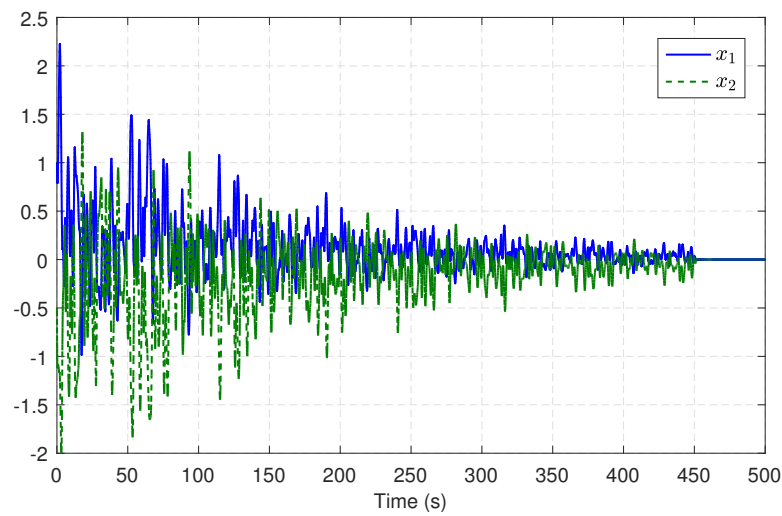


Figure 3. Trajectories of system states in the learning.

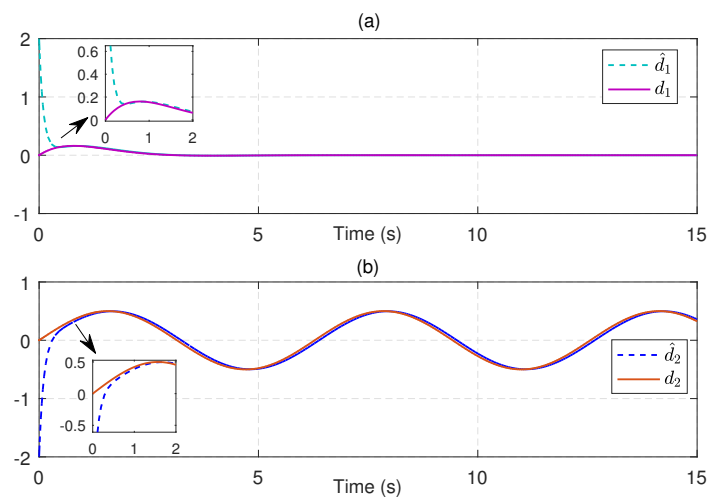


Figure 4. (a) Real disturbance d_1 and its estimation \hat{d}_1 , (b) Real disturbance d_2 and its estimation \hat{d}_2 .

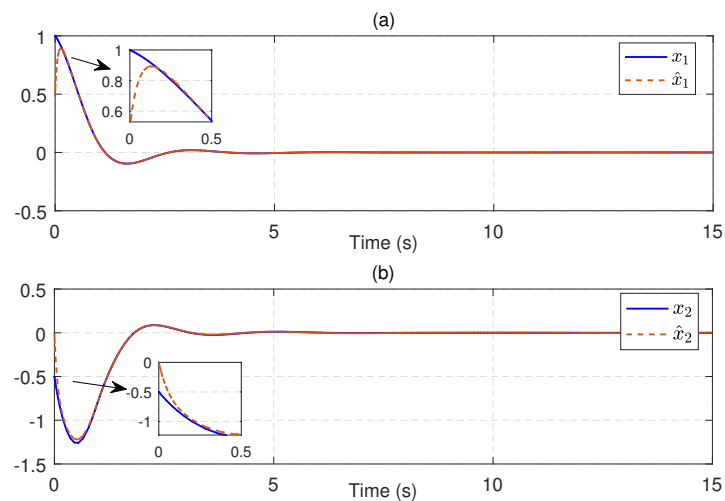


Figure 5. (a) Real state x_1 and identified state \hat{x}_1 , (b) Real state x_2 and identified state \hat{x}_2 .

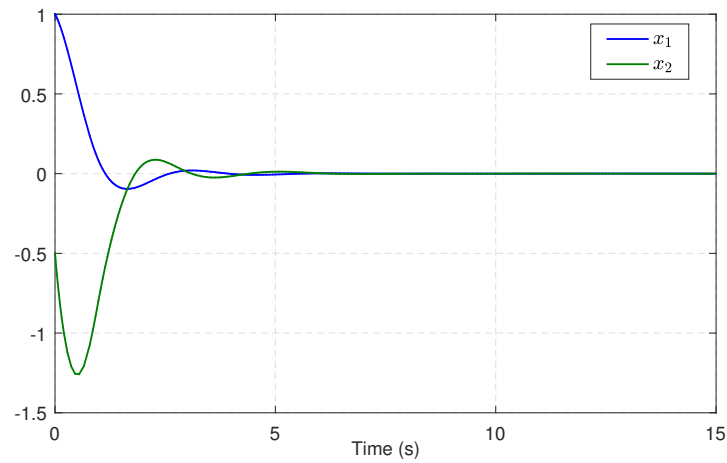


Figure 6. State trajectories of the robotic arm.

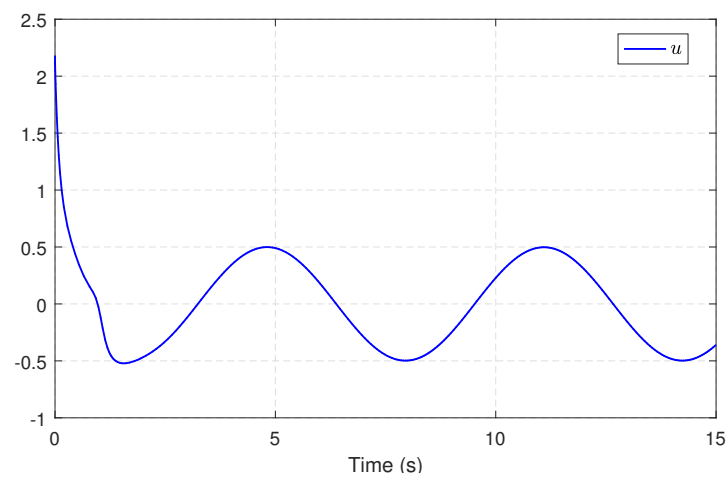


Figure 7. The compound control u .

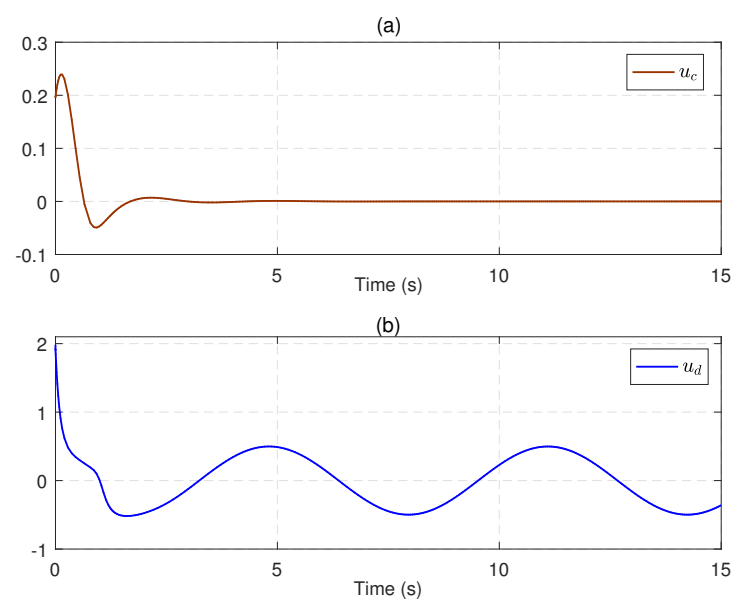


Figure 8. (a) The H_∞ optimal control \hat{u}_c . (b) The SMC law u_d .

5.2. Power Plant System

To further validate the effectivity of the proposed scheme, we consider an electric power system comprised of a gas turbine generator, a system load, and an automatic generation control [34]. To model this system, the incremental frequency deviation Δf_G , the generator output power variation ΔP_m , and the valve position change of the governor Δv are taken into consideration. The control input is represented by the speed change ΔP_c in position deviation. By defining the state vector $x = [\Delta v, \Delta P_m, \Delta f_G]^T \in \mathbb{R}^3$, we can express the reduced power system model in state-space form as

$$\dot{x} = \begin{bmatrix} -\frac{1}{T_g} & 0 & \frac{1}{R_g T_g} \\ \frac{K_t}{T_t} & -\frac{1}{T_t} & 0 \\ 0 & \frac{K_p}{T_p} & -\frac{1}{T_p} \end{bmatrix} x + \begin{bmatrix} \frac{1}{T_g} \\ 0 \\ 0 \end{bmatrix} (u + \vartheta) + d \tag{66}$$

where $g(x) = [1/T_g, 0, 0]^T$, ϑ represents the modeling uncertainty, and d stands for the exterior disturbances. Assume that the uncertain term is $\vartheta = x_2 \sin(x_1)$, and the disturbance term is defined as $d(t) = [\sin(2\pi t)e^{-t}, 0, 0.2 \sin^2(t)e^{-t}]^T$ in the simulation. Let the regulation constant $R_g = 2.5$ Hz/MW, the turbine gain constant $K_t = 1$ s and the generator gain constant $K_p = 120$ Hz/MW. Moreover, the corresponding time constants are set as $T_g = 0.08$ s, $T_t = 0.1$ s and $T_p = 20$ s, respectively.

For estimating the unknown uncertainty and disturbance terms, the enhanced observer system is constructed as (6) with a three-layered feedforward NN containing eight hidden neurons and the Hurwitz matrix $A = [-12, 0, 0; 0, -12, 0; 0, 0, -12]$. The activation function, the initial weights, and the updating ratios are the same as in Section 5.1. Let $p(x) = [10x_1, 0, 10x_3]^T$, $l(x) = [10, 0, 0; 0, 0, 0; 0, 0, 10]$, $G(x) = g^+(x) = [0.08, 0, 0]$ and $S_0(x) = 0.08x_1$. Similarly, an arctangent function $\text{atan}(s/\epsilon)$ is used for designing the SMC law u_d instead of the sign function $\text{sgn}(s)$.

Without the matched uncertainties and disturbances, we can derive the sliding-mode dynamics from (66), wherein $k(x) = [0, 0, 0; 0, 1, 0; 0, 0, 1]$, and the initial state $x_0 = [0.2, -0.2, 0.1]^T$. Let the associated cost function be of the form (32) along with $Q = \text{diag}(1, 1, 1)$, $R = 1$ and $\gamma = 3$. The critic NN is designed as (44) and its corresponding parameters are $\alpha = 15$, $\beta = 0.5$, $\sigma(x) = [x_1^2, x_1x_2, x_1x_3, x_2^2, x_2x_3, x_3^2, x_1^2x_2x_3, x_1x_2^2x_3, x_1x_2x_3^2]^T$ and $\hat{W}_c = [\hat{W}_{c1}, \hat{W}_{c2}, \dots, \hat{W}_{c9}]^T$. Similar to Section 5.1, $J_a(x) = x^T x/2$, the initial weight vector is set to zero, and a similar probing noise is injected into the control input before 550 s. The evolving trajectories of the critic weights are shown in Figure 9, while the trajectories of system states in the learning are depicted in Figure 10. After 550 s, the critic weights converge to $\hat{W}_c = [0.0830, 0.1245, 0.2284, 0.1616, 0.4883, 0.5488, 0.1154, 0.0563, 0.0564]^T$, then we can derive \hat{u}_c using (46) with the converged weights.

Then, we substitute \hat{u}_c into the integral sliding surface (21), and we design the SMC law u_d by (23) and (24). Consequently, the compound control is constructed as $u = u_d + \hat{u}_c$. After simulation, Figure 11 shows the trajectories of the power system states under this compound control for 15 s. Figure 12 presents the compound control u . From Figures 11 and 12, we can conclude that the compound control effectively stabilizes the system states to the equilibrium point, even in the presence of modeling uncertainties and exterior disturbances. These results undeniably demonstrate the viability and efficiency of the proposed approach.

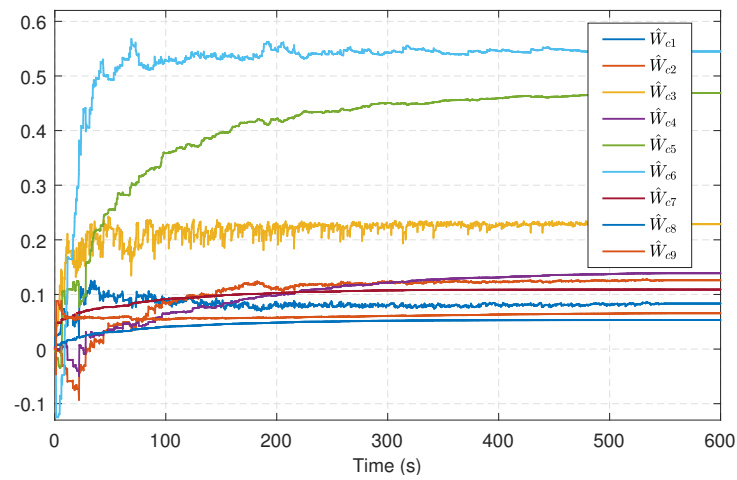


Figure 9. Trajectories of the critic NN weights.

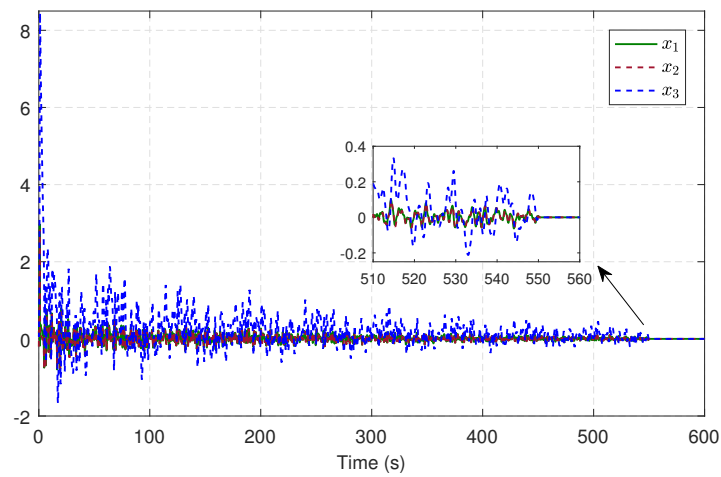


Figure 10. Trajectories of system states in the learning.

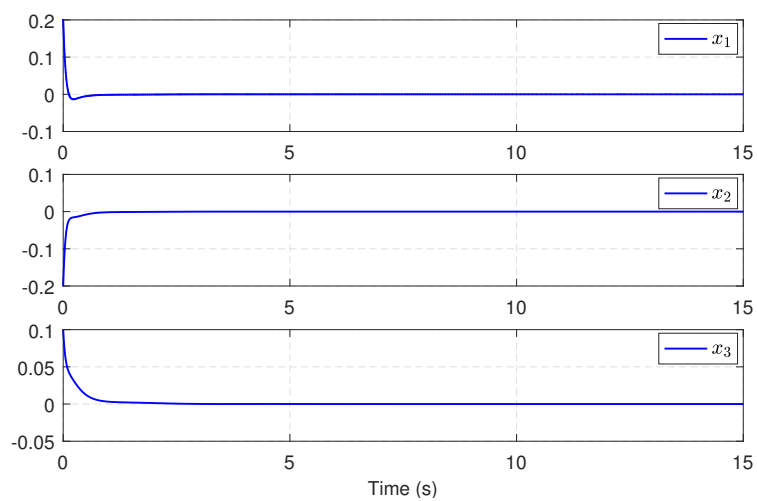


Figure 11. Trajectories of the electric power system.

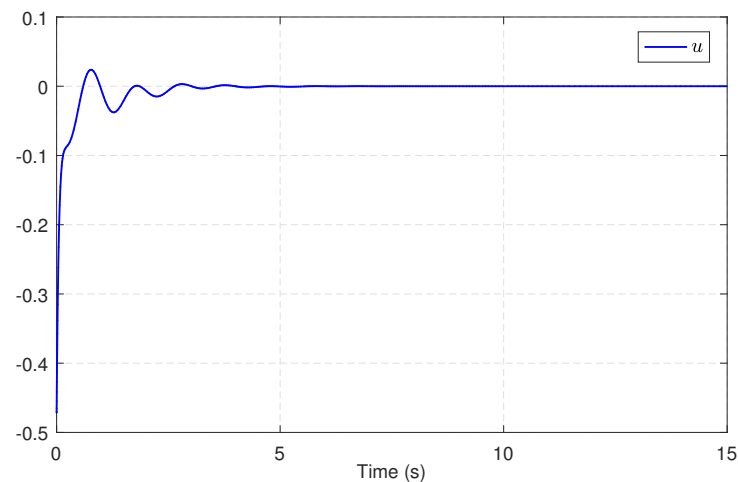


Figure 12. The compound optimal control.

6. Conclusions

In this paper, we develop a neural adaptive H_∞ sliding-mode control scheme for uncertain nonlinear systems subject to external disturbances. Based on the enhanced observer system composed of the NN identifier and nonlinear DO, an integral SMC is designed for suppressing the influences of the uncertain term and the matched disturbance component, as well as unknown approximation errors, with no prior knowledge of their upper bounds. Meanwhile, on the sliding surface, the remaining unmatched disturbances are attenuated using the H_∞ optimal control solved by the single critic network-based ADP algorithm. Furthermore, uniform ultimate boundedness stability of the resultant closed-loop system can be proven by Lyapunov's method. In addition to the theoretical analysis, two simulation examples are provided to further validate the proposed approach. Recently, the growing interest in saving communication resources or reducing the calculation amount of networked control systems makes the event-triggering mechanism gain more and more attention and undergo rapid development. Hence, how to combine the optimal SMC strategy with the event-triggering mechanism for more complex physical systems, not just for control-affine systems, will be our future research topic.

Author Contributions: Y.H. contributed to the conception and design of this study, performed the experiment and the data analysis, and wrote the manuscript; Z.Z. contributed to the results discussion and the writing of the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the Advanced Talents Incubation Program of Hebei University under Grant 521100221049, in part by Hebei Province Higher Education Science and Technology Research Project of China under Grant CXY2023009, in part by Hebei University Research and Innovation Team Project under Grant IT202306, and in part by Baoding Science and Technology Plan Project under Grant 2372P010.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The authors can confirm that all relevant data are included in the article.

Conflicts of Interest: The authors declare that they have no conflict of interest. All authors have approved the manuscript and agreed with submission to this journal.

References

1. Ioannou, P.; Sun, J. *Robust Adaptive Control*; Prentice Hall: Upper Saddle River, NJ, USA, 1996.
2. Utkin, V.; Guldner, J.; Shi, J. *Sliding Mode Control in Electro-Mechanical Systems*; CRC Press: Boca Raton, FL, USA, 2017.
3. Yu, X.; Kaynak, O. Sliding-mode control with soft computing: A survey. *IEEE Trans. Ind. Electron.* **2009**, *56*, 3275–3285.

4. Xu, J.; Guo, Z.; Tong, H. Design and implementation of integral sliding-mode control on an underactuated two-wheeled mobile robot. *IEEE Trans. Ind. Electron.* **2014**, *61*, 3671–3681. [[CrossRef](#)]
5. Chen, L.; Edwards, C.; Alwi, H. Integral sliding mode fault-tolerant control allocation for a class of affine nonlinear system. *Int. J. Robust Nonlinear* **2019**, *29*, 565–582. [[CrossRef](#)]
6. Pan, Y.; Yang, C.; Pan, L.; Yu, H. Integral sliding mode control: Performance, modification, and improvement. *IEEE Trans. Ind. Informat.* **2017**, *14*, 3087–3096. [[CrossRef](#)]
7. Errouissi, R.; Ouhrouche, M.; Chen, W.; Trzynadlowski, A. Robust nonlinear predictive controller for permanent-magnet synchronous motors with an optimized cost function. *IEEE Trans. Ind. Electron.* **2012**, *59*, 2849–2858. [[CrossRef](#)]
8. Huang, J.; Ri, S.; Fukuda, T.; Wang, Y. A disturbance observer based sliding mode control for a class of underactuated robotic system with mismatched uncertainties. *IEEE Trans. Autom. Control* **2019**, *64*, 2480–2487. [[CrossRef](#)]
9. Cui, R.; Chen, L.; Yang, C.; Chen, M. Extended state observer-based integral sliding mode control for an underwater robot with unknown disturbances and uncertain nonlinearities. *IEEE Trans. Ind. Electron.* **2017**, *64*, 6785–6795. [[CrossRef](#)]
10. Wang, Y.; Xie, X.; Chadli, M.; Xie, S.; Peng, Y. Sliding-mode control of fuzzy singularly perturbed descriptor systems. *IEEE Trans. Fuzzy Syst.* **2020**, *29*, 2349–2360. [[CrossRef](#)]
11. Chen, M.; Chen, W. Sliding mode control for a class of uncertain nonlinear system based on disturbance observer. *Int. J. Adapt. Control Signal Process* **2010**, *24*, 51–64. [[CrossRef](#)]
12. Rubagotti, M.; Estrada, A.; Castanos, F.; Ferrara, A. Integral sliding mode control for nonlinear systems with matched and unmatched perturbations. *IEEE Trans. Autom. Control* **2011**, *56*, 2699–2704. [[CrossRef](#)]
13. Castanos, F.; Fridman, L. Analysis and design of integral sliding manifolds for systems with unmatched perturbations. *IEEE Trans. Autom. Control* **2006**, *51*, 853–858. [[CrossRef](#)]
14. Kiumarsi, B.; Vamvoudakis, K.G.; Modares, H.; Lewis, F.L. Optimal and autonomous control using reinforcement learning: A survey. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 2042–2062. [[CrossRef](#)] [[PubMed](#)]
15. Liu, D.; Wei, Q.; Wang, D.; Yang, X.; Li, H. *Adaptive Dynamic Programming with Applications in Optimal Control*; Springer: Cham, Switzerland, 2017.
16. Lewis, F.L.; Liu, D. *Reinforcement Learning and Approximate Dynamic Programming for Feedback Control*; Wiley: Hoboken, NJ, USA, 2013.
17. Ha, M.; Wang, D.; Liu, D. Discounted iterative adaptive critic designs with novel stability analysis for tracking control. *IEEE/CAA J. Autom. Sin.* **2022**, *9*, 1262–1272. [[CrossRef](#)]
18. Wei, Q.; Lewis, F.L.; Liu, D.; Song, R.; Lin, H. Discrete-time local value iteration adaptive dynamic programming: Convergence analysis. *IEEE Trans. Syst. Man Cybern. Syst.* **2018**, *48*, 875–891. [[CrossRef](#)]
19. Wei, Q.; Wang, L.; Lu, J.; Wang, F.Y. Discrete-Time Self-Learning Parallel Control. *IEEE Trans. Syst. Man Cybern. Syst.* **2022**, *52*, 192–204. [[CrossRef](#)]
20. Heydari, A.; Balakrishnan, S. Finite-horizon control-constrained nonlinear optimal control using single network adaptive critics. *IEEE Trans. Neural Netw. Learn. Syst.* **2013**, *24*, 145–157. [[CrossRef](#)] [[PubMed](#)]
21. Lu, J.; Wei, Q.; Wang, Z.; Zhou, T.; Wang, F. Event-triggered optimal control for discrete-time multi-player non-zero-sum games using parallel control. *Inf. Sci.* **2022**, *584*, 519–535. [[CrossRef](#)]
22. Wang, D.; Ren, J.; Ha, M. Discounted linear Q-learning control with novel tracking cost and its stability. *Inf. Sci.* **2023**, *626*, 339–353. [[CrossRef](#)]
23. Zhang, X.; Ni, Z.; He, H. A theoretical foundation of goal representation heuristic dynamic programming. *IEEE Trans. Neural Netw. Learn. Syst.* **2016**, *27*, 2513–2525. [[CrossRef](#)]
24. Huang, Y.; Wang, D.; Liu, D. Bounded robust control design for uncertain nonlinear systems using single-network adaptive dynamic programming. *Neurocomputing* **2017**, *266*, 128–140. [[CrossRef](#)]
25. Yang, X.; Wei, Q. Adaptive critic designs for optimal event-driven control of a CSTR system. *IEEE Trans. Ind. Inform.* **2021**, *17*, 484–493. [[CrossRef](#)]
26. Yang, X.; He, H.; Zhong, X. Approximate dynamic programming for nonlinear-constrained optimizations. *IEEE Trans. Cybern.* **2021**, *51*, 2419–2432. [[CrossRef](#)] [[PubMed](#)]
27. Wen, G.; Niu, B. Optimized tracking control based on reinforcement learning for a class of high-order unknown nonlinear dynamic systems. *Inf. Sci.* **2022**, *606*, 368–379. [[CrossRef](#)]
28. Wang, D.; Qiao, J.; Cheng, L. An approximate neuro-optimal solution of discounted guaranteed cost control design. *IEEE Trans. Cybern.* **2022**, *52*, 77–86. [[CrossRef](#)] [[PubMed](#)]
29. Liu, D.; Xue, S.; Zhao, B.; Luo, B.; Wei, Q. Adaptive dynamic programming for control: A survey and recent advances. *IEEE Trans. Syst. Man Cybern. Syst.* **2020**, *51*, 142–160. [[CrossRef](#)]
30. Wang, D.; Ha, M.; Zhao, M. The intelligent critic framework for advanced optimal control. *Artif. Intell. Rev.* **2022**, *55*, 1–22. [[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. Luo, B.; Wu, H.; Huang, T. Off-policy reinforcement learning for H_∞ control design. *IEEE Trans. Cybern.* **2014**, *45*, 65–76. [[CrossRef](#)]

33. Modares, H.; Lewis, F.L.; Jiang, Z. H_∞ tracking control of completely unknown continuous-time systems via off-policy reinforcement learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *26*, 2550–2562. [[CrossRef](#)]
34. Wang, D.; He, H.; Liu, D. Adaptive critic nonlinear robust control: A survey. *IEEE Trans. Cybern.* **2017**, *47*, 3429–3451. [[CrossRef](#)]
35. Mitra, A.; Behera, L. Continuous-time single network adaptive critic based optimal sliding mode control for nonlinear control affine systems. In Proceedings of the 34th Chinese Control Conference, HangZhou, China, 28–30 July 2015; pp. 3300–3306.
36. Fan, Q.; Yang, G. Adaptive actor-critic design-based integral sliding-mode control for partially unknown nonlinear systems with input disturbances. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *27*, 165–177. [[CrossRef](#)] [[PubMed](#)]
37. Qu, Q.; Zhang, H.; Yu, R.; Liu, Y. Neural network-based H_∞ sliding mode control for nonlinear systems with actuator faults and unmatched disturbances. *Neurocomputing* **2018**, *275*, 2009–2018. [[CrossRef](#)]
38. Zhang, H.; Qu, Q.; Xiao, G.; Cui, Y. Optimal guaranteed cost sliding mode control for constrained-input nonlinear systems with matched and unmatched disturbances. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 2112–2126. [[CrossRef](#)] [[PubMed](#)]
39. Yang, D.; Li, T.; Xie, X.; Zhang, H. Event-triggered integral sliding-mode control for nonlinear constrained-input systems with disturbances via adaptive dynamic programming. *IEEE Trans. Syst. Man Cybern. Syst.* **2019**, *50*, 4086–4096. [[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.