Journal of Mathematical Modeling
Vol. 14, No. 1, 2026, pp. 191-206. Research Article

Design of an optimal learning sliding mode controller for linear systems

Tahereh Azizpour, Majid Yarahmadi*

Department of Mathematics and Computer Science, Lorestan university, Lorestan, Iran

Email(s): azizpour.ta@fs.lu.ac.ir, yarahmadi.m@lu.ac.ir

Abstract. In this paper, a new learning robust controller based on the sliding mode control method and reinforcement learning approach is designed for a class of single-input single-output linear systems with a relative degree of uncertainty r. For this purpose, a hybrid controller that including equivalent controller and learning robust controller, is designed. The proposed controller guarantees asymptotic stability, sliding condition, finite reaching time, elimination of chattering phenomenon and tracking of desired output in an optimal approach. For online approximation of the value function and design of an optimal policy, a new robust optimal learning controller is designed. For analytical facilitation and stability analysis, three theorems are proved and a new algorithm is designed. Finally, a simulation example is presented to demonstrate the advantages of the proposed method. The simulation results show the optimality of the controls, the elimination of chattering phenomenon and the tracking of the desired output.

Keywords: Control, sliding mode control, optimal control, reinforcement learning.
AMS Subject Classification 2010: 34A34, 65L05.

1 Introduction

Since the 1990s, the control of systems subject to external disturbances and uncertainties has attracted growing interest due to its theoretical and practical significance [18]. Designing robust controllers for such systems has always been a major challenge in control engineering. Among the various existing control strategies, Sliding Mode Control (SMC) has been recognized as an effective and reliable method for electronic control systems due to its robustness against uncertainties and disturbances, fast dynamic response, ease of implementation, and insensitivity to parameter variations and system dynamics in the presence of perturbations [6, 10].

*Corresponding author
Received: 18 March 2025/ Revised: 20 September 2025/ Accepted: 4 October 2025
DOI: 10.22124/jmm.2025.30152.2694

However, conventional SMC suffers from well-known drawbacks, particularly chattering and the generation of high-gain control signals. Chattering is mainly caused by the discontinuous nature of the control law, which produces undesirable oscillations around the sliding surface [14].

To alleviate chattering, several strategies have been proposed, such as replacing the switching function with continuous approximations [7], including saturation functions, hyperbolic tangent functions [1], support vector machines [14], and fuzzy logic systems [16]. While these techniques can reduce or even eliminate chattering in control signals and state variables, they may also introduce steady-state errors, require complex tuning, or produce high-frequency oscillations if parameters are not carefully selected. Moreover, methods based on support vector machines demand large datasets and high computational resources, and fuzzy logic–based SMC entails complicated stability and convergence analyses.

The output tracking problem is one of the most important and active research topics in control theory [12]. As discussed in [8], the design of tracking controllers aims not only to asymptotically align the system output with a reference signal but also to preserve robust dynamic performance in the presence of uncertainties. Achieving optimal tracking performance [27] requires minimizing quadratic performance indices and guaranteeing stability under system uncertainties. In this regard, combining SMC with optimal control has been investigated as a promising solution [17].

Due to their strong learning and optimization capabilities [9], Adaptive Dynamic Programming (ADP) methods [5] have been extensively studied for solving optimal control problems and the Hamilton-Jacobi-Bellman (HJB) equation. The ADP-based approaches have found applications in path tracking, zero-sum games, systems with uncertainty, and actuator saturation. Reinforcement Learning (RL) [21], often regarded as the data-driven counterpart of ADP, has also been applied in various control scenarios [25]. For instance, [4, 23] developed optimal control schemes using neural networks and RL algorithms for unconstrained feedback control problems, demonstrating the capability of ADP/RL in enhancing performance under complex, dynamic conditions.

Recently, sliding mode-based Integral RL (IRL) event-triggered control has emerged as a leading approach for solving robust control problems in uncertain nonlinear systems [11]. This method integrates IRL with sliding mode concepts to solve HJB equations online [13] without requiring full knowledge of the system dynamics, and applies event-triggering to reduce computational burden-ensuring asymptotic stability via Lyapunov-based proofs. The proposed controllers have demonstrated effective chattering suppression, high tracking accuracy, and adaptability to time-varying environments. Additionally, event-triggered integral sliding mode optimal tracking control has been developed by integrating ADP with SMC to handle uncertain nonlinear systems-significantly reducing update frequency and enhancing robustness while maintaining near-optimal performance [27, 28]. Moreover, event-triggered IRL-based ADP schemes have been proposed for systems with input saturation, achieving guaranteed convergence, chattering reduction, and low communication overhead [24]. Motivated by recent advances, this paper proposes a hybrid control scheme that combines sliding mode control (SMC) with optimal control for output tracking in uncertain linear systems. The hybrid approach aims to ensure robust performance while minimizing chattering around the sliding surface. To achieve this, an IRL-based algorithm is developed to approximate the value function and determine the optimal control policy online. The paper is organized as follows:

In section 2, the system and problem formulation will be described. The control rules are designed in Section 3. The analysis of the stability of the system is provided in Section 4. In Section 5, the *RL* algorithm is designed for approximating the value function. To demonstrate the efficiency and advantage of the proposed control, two simulation experiments are presented in Section 6. Section 7 provides numerical performance comparisons with other state-of-the-art methods, and finally, the conclusion is stated in Section 8.

2 Problem Statement

In order to design a robust control law that ensures stability and tracking performance, it is essential to first establish a mathematical model of the dynamic system under consideration. In this work, we consider a class of continuous-time linear systems commonly encountered in control applications, where the state dynamics are influenced by both control inputs and external disturbances [19]. These systems are often subject to parametric uncertainties, unmodeled dynamics, and external perturbations, which must be taken into account during the controller design process. The general form of such a system is given by [22]:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + d(t), \\ y(t) &= Cx(t), \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ and $y \in \mathbb{R}$ are the state, the control input and the controlled output, respectively. Also, $d \in \mathbb{R}^n$ denotes unstructured uncertainties in the system, which can be described as mismatch uncertainties, external disturbances, and/or unmodeled dynamics. Note that $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ and $C^T \in \mathbb{R}^n$ are time-invariant matrices.

In this paper, a single-input single-output (SISO) is considered. The controlled output $y(t) \in \mathbb{R}$ is scalar, and the sliding surface $s(t)$ is defined based on the scalar tracking error $e(t)$. Consequently, the control input $u(t)$ is a scalar. Also, the finite-horizon cost functional, the weighting factors Q and R are taken as non-negative and positive scalars, respectively.

Assumption 1. *The pair of A and B are controllable and* $rank(B) = m$.

Assumption 2. *Pair A and C are observable.*

Assumption 3 ([22])**.** *It is assumed that the following upper bound of the uncertainties and their derivatives are known:*

$$\| d(t) \| \leq \eta_1, \quad \| \dot{d}(t) \| \leq \eta_2, ..., \quad \| d^{(r)}(t) \| \leq \eta_r, \qquad \eta_1, \eta_2, ..., \eta_r > 0, \tag{2}$$

where $\| \,.\, \|$ denotes the $2-$norm of a vector and r, is the relative degree of the system. In other words, if the relative degree of the system is r we have

$$CA^iB = 0, \quad i = 0, 1, ..., r-1, \quad CA^rB \neq 0. \tag{3}$$

Notice 1. Controlling systems with external disturbances and unmodeled frequencies requires designing a robust and stable control law. For this purpose, designing an optimal performance evaluation model, eliminating the chattering phenomenon, reducing control effort, and ensuring system stability are key challenges. Table 1 shows the main approaches and novel ideas of this paper in overcoming the issues and challenges.

Table 1: The novelties and contributions of this paper

Issues and Challenges	Ideas and Novelties	Related Sections
Controlling of the system with external disturbances and unmolded frequencies.	Design of robust control (SMC).	Section 2.
Elimination of chattering phenomenon caused by discontinuities in the SMC control law.	Design of a hybrid control that includes equivalent control and reinforcement learning control to eliminate chattering phenomenon.	Section 3 and equations (6)-(13).
System stability analysis.	Stability guaranteed by two theorems.	Section 4, Theorems 1, 2 and 3 and Definition 2.
Difficulty of analytical solution of HJB equation.	The HJB solution is approximated online via the PI algorithm.	Section 5.
Application and implementation.	Presenting an algorithm and simulation.	Section 5, Algorithm 1, and Section 6.

Consider the sliding surface $s(x,t)$ as follows:

$$s(x,t) = \sum_{i=0}^{r} \lambda_i e^{(r-i)} + \lambda_{r+1} \int_0^t e(t)dt, \quad \lambda_0 = 1, \lambda_i > 0, \tag{4}$$

where $e(t) = y(t) - y_d(t)$ and $y_d(t)$ are the tracking error and the desired output, respectively. Also, $e^{(r)}(t)$ represents the $r-$th time derivative of $e(t)$ and $\lambda_i, i = 0, ..., r+1$ are the positive constant parameters. With simple calculations, one can obtain

$$\dot{s}(x,t) = \sum_{i=0}^{r+1} \lambda_i e^{(r+1-i)}, \quad \lambda_0 = 1. \tag{5}$$

For the given relative degree of the system and considering the nominal system (1), the error derivatives can be calculated as follows:

$$e^{(i)} = CA^i x - y_d^{(i)}, CA^i B = 0, \quad 0 \leq i < r.$$

Therefore, $e^{(r+1)} = CA^r \dot{x} - y_d^{(r+1)} = CA^r x + CA^r Bu - y_d^{(r+1)}$, $CA^r B \neq 0$. Thus, $\dot{s}(x,t)$ can be rewritten for the nominal system (1) as follows:

$$\dot{s}(x,t) = CA^{r+1} x(t) + CA^r Bu(t) - y_d^{(r+1)} + \sum_{i=1}^{r+1} \lambda_i e^{(r+1-i)}.$$

Considering the uncertainties and defining that $\delta(t) = C(d(t) + A\dot{d}(t) + \cdots + A^{r-1} d^{(r)}(t))$, then

$$\dot{s}(x,t) = CA^{r+1} x(t) + CA^r Bu(t) + \delta(t) - y_d^{(r+1)} + \sum_{i=1}^{r+1} \lambda_i e^{(r+1-i)}.$$

To achieve optimal control in finite horizon, we define the finite horizon integral cost function as follows:

$$J(s(0)) = \int_0^{t_f} \left(s^T(t)Qs(t) + u^T(t)Ru(t) - \gamma^2 \parallel d(t) \parallel^2\right) dt,$$

where Q and R are positive semidefinite and positive definite matrices, respectively. Our goal is to design a control that can track the desired output of the system, in the presence of the uncertainties in system (1). For this purpose, a hybrid control will be presented. This control consists of an equivalent controller and a RL controller.

3 Design of control laws

Consider a hybrid control law $u(t)$ as follows:

$$u(t) = u_1(t) + u_2(t), \tag{6}$$

where $u_1(t)$ and $u_2(t)$ are the equivalent control in the SMC method and the RL control signal, respectively. Now, the equivalent control $u_1(t)$ is obtained using the nominal system and $\dot{s} = 0$:

$$u_1(t) = \frac{1}{CA^rB}\left[y_d^{(r+1)} - \sum_{i=1}^{r} \lambda_j e^{(r-i+1)} - CA^{r+1}x\right]. \tag{7}$$

Also, the reaching state control $u_2(t)$ is designed, via an optimization problem solving. Suppose that the value function of this problem is as follows:

$$v(s(t)) = \int_t^{t_f} \left(s^T(\tau)Qs(\tau) + u_2(\tau)^T Ru_2(\tau) - \gamma^2 \parallel d(\tau) \parallel^2\right) d\tau, \tag{8}$$

so that

$$v^*(s(0)) = \min_{u_2} \max_{d} J(s(0), u_2, d).$$

In other words, $v^*(s(0))$ is the optimal value function. The main problem is to find the optimal control u_2^* and disturbance policies d^* such that

$$u_2^* = arg\min J(s(0), u_2, d), \quad d^* = arg\max J(s(0), u_2, d). \tag{9}$$

Definition 1. *The system (1) has L_2-gain less than or equal to γ if the following disturbance attenuation condition is satisfied for all $d \in L_2[0,\infty)$ with $x(0) = 0$:*

$$\frac{\int_t^\infty \parallel z(\tau) \parallel^2 d\tau}{\int_t^\infty \parallel d(\tau) \parallel^2 d\tau} \le \gamma^2, \tag{10}$$

where $\parallel z(\tau) \parallel^2 = s^TQs + u_2^TRu_2$, $d(t)$ is the disturbance input, and γ is the amount of attenuation level.

It is assumed that γ in (8) satisfies $\gamma \geq \gamma^*$, where γ^* is the smallest γ which satisfies the disturbance attenuation condition [2].
By differentiating the value function (8) with respect to t, we have the following Bellman equation:

$$s^T Qs + u_2^T Ru_2 - \gamma^2 \parallel d \parallel^2 + \nabla v^T \dot{s} = 0, \tag{11}$$

where $\nabla v^T(s) = \frac{\partial v(s)}{\partial s}$. The terminal condition for this partial differential equation is $v(t_f) = 0$. A solution to (11) is the value function $v(s)$ for the feedback control $u_2 = u_2(v(s))$ and the disturbance $d = d(v(s))$. Based on (11), the Hamiltonian function is defined as:

$$H(s, \nabla v, u_2, d) = s^T(t)Qs(t) + u_2(t)^T Ru_2(t) - \gamma^2 \parallel d(t) \parallel^2 + \nabla v^T(s)\dot{s}(t). \tag{12}$$

By applying the stationary conditions at the equilibrium point, the optimal control and disturbance are obtained as follows:

$$u_2^*(t) = -\frac{1}{2R} CA^r B \nabla v(s), \qquad d^*(t) = \frac{1}{2\gamma^2} C^T \nabla v(s). \tag{13}$$

On the other hand, $\frac{\partial^2 H}{\partial u_2^2} = 2R > 0$ and $\frac{\partial^2 H}{\partial d^2} = -2\gamma^2 < 0$ so, u_2 and d are the minimum and maximum points of the Hamiltonian function, respectively.

4 System stability analysis

In this section, stability and sliding condition are analyzed based on the proposed controllers. For this purpose, consider the following theorems.

Theorem 1. *Suppose that $v^*(s)$ is a positive semi-definite function such that it solves the HJB equation (12). Also, suppose that u_1, u_2, and d are determined by equations (7) and (13), then the optimal system with equations (1), (6), and (8) is asymptotically stable.*

Proof. First, the optimal control and disturbance are substituted into the Bellman equation (11). Therefore

$$s^T Qs + \frac{1}{4R^2}(CA^r B \nabla v(s))^T R (CA^r B \nabla v(s)) - \frac{1}{4\gamma^2}(C^T \nabla v(s))^T (C^T \nabla v(s)) + \nabla v^T \dot{s} = 0. \tag{14}$$

Since the attenuation condition in Definition 1 is satisfied, the *HJB* equation (14) has a positive semi-definite solution $v^*(s)$ [2]. Now, consider Lyapunov function $V(t)$ as follows:

$$V(t) = \int_t^{t_f} \left(s^T(\tau)Qs(\tau) + u_2^T(\tau)Ru_2(\tau) - \gamma^2 \parallel d(\tau) \parallel^2 \right) d\tau. \tag{15}$$

By selecting γ which satisfies the attenuation condition, we have $V(t) \geq 0$ and $V(t) = 0$ if and only if $s(t) = 0$. Now, by differentiating $V(t)$ with respect to t, we obtain

$$\frac{dV}{dt} = -\left(s^T(t)Qs(t) + u_2^T(t)Ru_2(t) - \gamma^2 \parallel d(t) \parallel^2 \right) \leq 0,$$

where $\dot{V}(s) = 0$ if and only if $s = 0$. Therefore, (1) will be asymptotically stable. □

In the next step, the sliding condition is examined.

Definition 2. *The sliding condition is defined as follows:*

$$\dot{s}(t)sgn(s(t)) \leq -\eta. \tag{16}$$

where η is a positive constant [6].

Theorem 2. *Consider the control system (1) and the sliding surface (4). If u_2^* and d^* are optimal control and disturbance signals given in equation (13), then the sliding condition will be guaranteed.*

Proof. It is clear that if s and $\dot{s}$ have different sign, then the sliding condition $s\dot{s}$ will be satisfied. Now according to the equation (15), we have

$$\frac{dV}{dt} = 2\int_t^{t_f} Qs(\tau)d\tau \times \frac{ds}{dt},$$

and also can conclude

$$\nabla V = 2\int_t^{t_f} Qs(\tau)d\tau, \tag{17}$$

where $\nabla V = \frac{\partial V}{\partial s}$. On the other hand

$$\nabla V = \frac{dV}{dt} \times \frac{1}{\frac{ds}{dt}}. \tag{18}$$

Since V is a Lyapunov function for the stable system (1), $\frac{dV}{dt} < 0$ for $s \neq 0$. Therefore, according to (18)

$$sgn(\dot{s}) = -sgn(\nabla V). \tag{19}$$

Now, according to the equations (17)-(19), we have: if $s > 0$ then $\nabla V > 0$ so $sgn(\dot{s}) < 0$ and if $s < 0$ then $\nabla V < 0$ so $sgn(\dot{s}) > 0$. So in general, we can say that for $s \neq 0$, s and $\dot{s}$ have different signs. Therefore by selecting $\eta = max(\mid s\dot{s} \mid)$, the sliding condition holds, and this completes the proof. □

Theorem 3. *Let the sliding surface function $s(t)$ satisfy the sliding condition (16), then the system state reaches the sliding surface $s(t) = 0$ in finite time.*

Proof. We consider two cases based on the sign of the sliding surface value.
Case 1: $s(t) > 0$, then according to Theorem 2 and for a given $\eta = max(|\dot{s}s|)$, one can conclude

$$\dot{s}(t) \leq -\eta. \tag{20}$$

By integrating both sides of (20) from time 0 to t_{reach}, we obtain

$$\int_0^{t_{\text{reach}}} \dot{s}(t)\,dt \leq -\int_0^{t_{\text{reach}}} \eta\,dt$$
$$s(t_{\text{reach}}) - s(0) \leq -\eta \cdot t_{\text{reach}}.$$

Since $s(t_{\text{reach}}) = 0$, we conclude

$$t_{\text{reach}} \leq \frac{s(0)}{\eta}. \tag{21}$$

Case 2: Suppose $s(t) < 0$, similarly, we have

$$\dot{s}(t) \geq \eta \Rightarrow t_{\text{reach}} \leq -\frac{s(0)}{\eta}. \tag{22}$$

In both cases, the system reaches the sliding surface in finite time, bounded by

$$t_{\text{reach}} \leq \frac{|s(0)|}{\eta}, \tag{23}$$

this completes the proof. □

5 Design of learning optimal controller based on the integral RL

As is well known, the optimal policies into (13) are based on solving the $v^*(s)$ from the HJB equation. On the other hand, in most cases there is no analytical solution for this equation. According to [3, 15], the solution of HJB equation can be approximated online using the Policy Iteration (PI) algorithm. In the following, the algorithm based on Integral RL for solving the HJB equation will be described.
The value function (8) can be rewritten as follows:

$$\begin{aligned} v(t) &= \int_t^{t+T} r(s(\tau), u_2(\tau), d(\tau))d\tau + \int_{t+T}^{t_f} r(s(\tau), u_2(\tau), d(\tau))d\tau \\ &= \int_t^{t+T} r(s(\tau), u_2(\tau), d(\tau))d\tau + v(s(t+\tau)), \end{aligned}$$

where T is the length of time interval and $r(s, u_2, d) = s^T Qs + u_2^T Ru_2 - \gamma^2 \parallel d \parallel^2$. Suppose that the value function can be approximated by a linear combination of basis functions ϕ with a weight vector W, that is

$$v(s) = W^T \phi(s). \tag{24}$$

Using the approximated value function by equation (24), the policy iteration algorithm, which consists of two steps: the evaluation policy and the policy improvement, will be used to find the optimal policy. The online PI algorithm is presented in Algorithm 1.

6 Simulation example

Example 1. To verify the effectiveness of the control strategy proposed in this paper, we apply it to a Turntable Flight Simulation System. This is a servo system and its model is usually expressed as follows [17]:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -\frac{K_m C_e}{JR_e} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ K_u \frac{K_m}{JR_e} \end{bmatrix} u(t) + d(t), \quad y(t) = Cx(t), \quad C = [1 \quad 0], \tag{25}$$

Algorithm 1: Online PI Algorithm.

Step1. Initialize the admissible policy $u_2^{(1)}$ and $d^{(1)}$ and also initialize $W^{(0)}$.
Step2. For each iteration k, find the value function $v^{(k)}(t)$ by

$$v^{(k)}(t) = \int_{t_1}^{t_1+T} r(s(\tau), u_2(\tau), d(\tau))\, d\tau + W^{(k-1)^T} \phi(s(t_1+T)).$$

Step3. Update the weight vector $W^{(k)}$ according to the estimated $v^{(k)}(s)$ using the least-squares method,

$$W^{(k)}\phi(s) = v^k(s).$$

Step4. Update policies $u_2^{(k+1)}$ and $d^{(k+1)}$ using the relation (13):

$$u_2^{(k+1)} = -\frac{1}{2R}\left(\nabla v^{(k)} CA^r B\right), \quad d^{(k+1)} = \frac{1}{2\gamma^2}\left(\nabla v^{(k)} C\right).$$

Step5. Repeat steps $2-4$ until convergence.

where $x_1(t) = \theta(t)$ is the angular position of the turntable, $x_2(t) = \dot{\theta}(t)$ is the angular velocity, K_u is the amplification factor of the PVM, R_e is the armature resistance, K_m is the motor torque coefficient, C_e is the electromotive force constant, and J is the moment of inertia. Also, $u(t)$ and $d(t)$ represent the control input and the external disturbances, respectively. Suppose the reference signal is given by:

$$\begin{bmatrix} \dot{z}_1(t) \\ \dot{z}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 4 \\ -5 & -1 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix}, \quad y_d(t) = C_d z(t), \qquad C_d = [1 \quad 5]. \tag{26}$$

The simulation parameters are selected as follows:

$$\begin{aligned} &R_e = 7.77, \quad K_m = 6, \quad C_e = 1.2, \quad J = 0.6, \quad K_u = 11, \quad Q = 2, \quad R = 7 \\ &\lambda_1 = \lambda_2 = 20, \quad dt = 0.01, \quad t \in [0,6], x_0 = [3 \quad 0.5]^T, z_0 = [0.1 \quad 0]^T. \end{aligned}$$

Since $CB = 0$ and $CAB \neq 0$, the degree of the system is $r = 1$. Therefore

$$s(t) = \dot{e}(t) + \lambda_1 e(t) + \lambda_2 \int_0^t e(\tau) d\tau, \tag{27}$$

$$\dot{s}(t) = CA^2 x(t) + CABu(t) - \ddot{y}_d + \lambda_1 \dot{e}(t) + \lambda_2 e(t), \tag{28}$$

$$u_1(t) = \frac{1}{CAB}\left(\ddot{y}_d - \lambda_1 \dot{e}(t) - \lambda_2 e(t) - CA^2 x(t)\right). \tag{29}$$

The RL interval length, T, is chosen to be 0.1. Also, the convex function $\phi(s) = s^2$ is used to approximate the value function. Next, according to Algorithm 1 the simulation results are given in the following figures. As can be seen in Figures 1 and 2, the optimal controllers u_1 and u_2 are smooth and without any chattering. Figures 3 and 4 show the output tracking signal of the desired output and the disturbance, respectively. The optimal performance is clearly visible in the compromise between the disturbance and the controller.

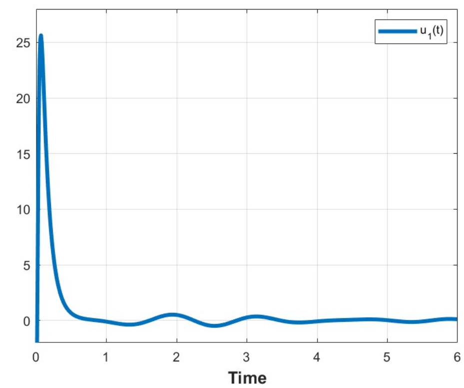

Figure 1: Control function $u_1(t)$

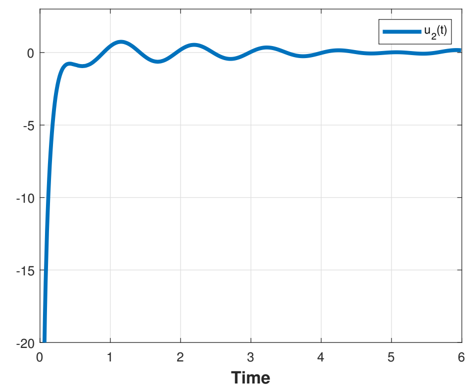

Figure 2: Control function $u_2(t)$

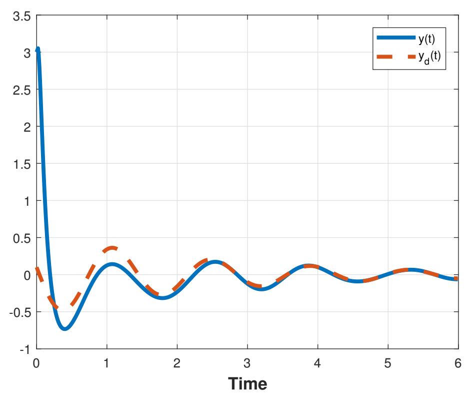

Figure 3: Output $y(t)$ and desired output $y_d(t)$

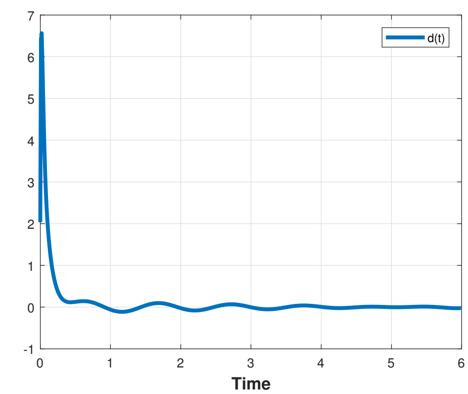

Figure 4: Function $d(t)$

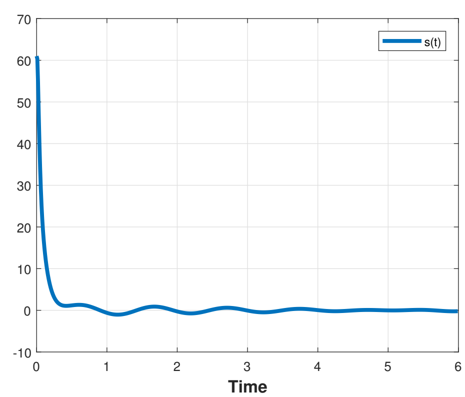

Figure 5: Sliding function $s(t)$

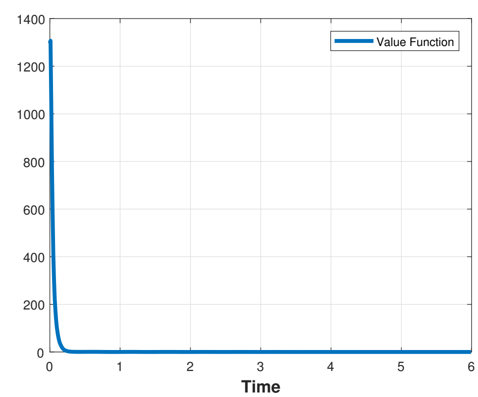

Figure 6: Value function

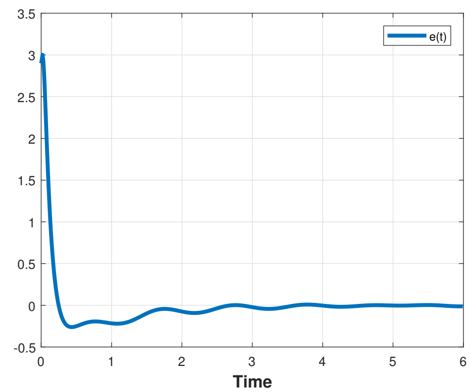

Figure 7: Error function $e(t)$

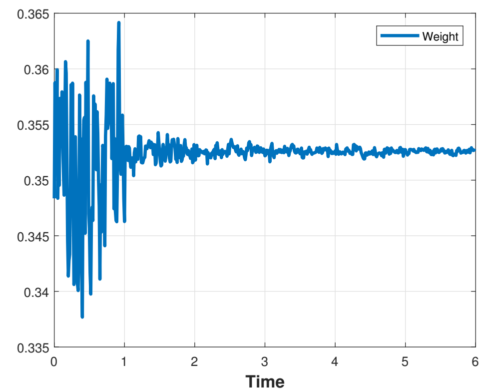

Figure 8: Weight function

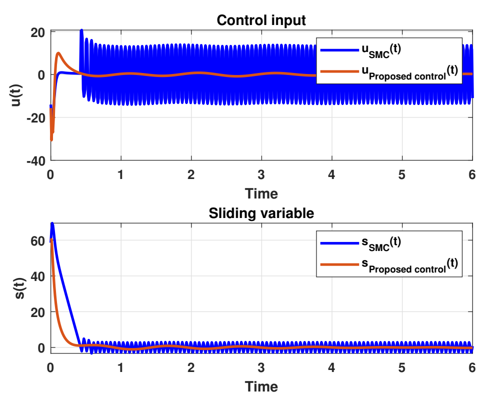

Figure 9: Control functions and sliding functions.

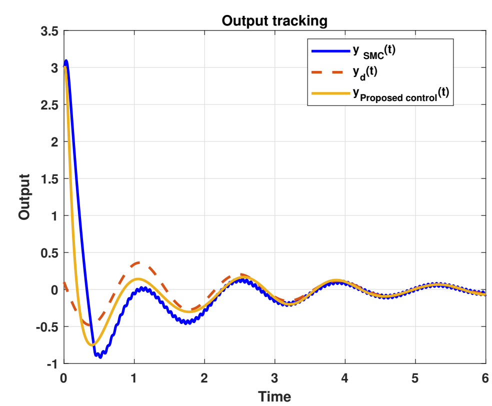

Figure 10: Outputs $y(t)$

Figures 5 and 6 also depict the rapid convergence of the sliding function to zero and the minimization of the value function, respectively. Figures 7 and 8 show the tracking error and the process of learning the weights of the basis functions to approximate the value function, respectively.

7 Numerical performance comparison

In the following, the proposed hybrid controller is compared with conventional sliding mode controllers. First, we consider sliding mode control as follows:

$$u(t) = u_{eq}(t) - Ksign(s(t)). \tag{30}$$

The simulation was performed for $\eta = 5$ and the results are shown in Figures 9 and 10. Furthermore, to eliminate the chattering, we replace the switching function (sign(.)) in traditional SMC with the smooth function tanh(.), i.e.

$$u(t) = u_{eq}(t) - Ktanh(s(t)). \tag{31}$$

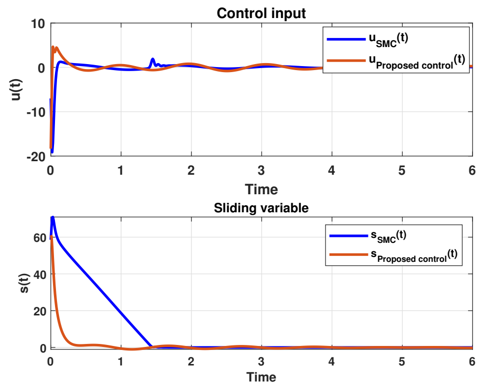

Figure 11: Control functions and sliding functions

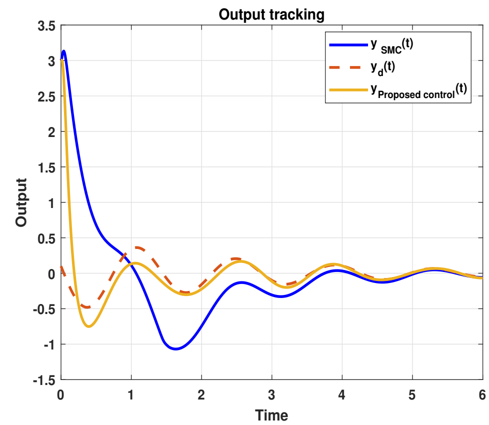

Figure 12: Outputs $y(t)$

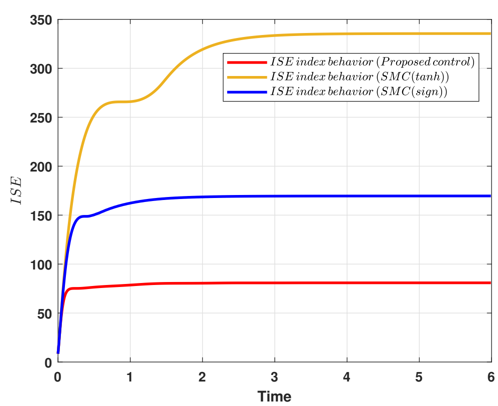

Figure 13: ISE performance index for SMC(sign), SMC(tanh) and proposed controls

Table 2: Comparison of control methods

Method	**Control interval**	**ISE**
SMC (sign)	$[-24.9490,\ 20.7437]$	169.612
SMC (tanh)	$[-18.3179,\ 1.2935]$	335.478
Proposed control	$[-18.3344,\ 4.6595]$	80.8881

The results are shown in Figures 11 and 12. In order to compare the numerical performances, the performances of the proposed method are compared with the performances of the three methods presented in [20] and [26] according to ISE index, where $ISE(t) = \int_0^t e^2(\tau)d\tau$. The results of these comparisons are shown in Table 2. According to Figure 13 and Table 2, the proposed controller can eliminate chattering with a reasonable control effort and track the desired output with ISE much smaller than the others.

With $y_d = 0$ and $C = [1,0]$ in Example 1, simulations are repeated and their results are compared with the approach in reference [2]. Figures 14, 15, and 16 show the tracking errors, ISE index behavior, and disturbance functions for the proposed optimal learning controller and the controller in reference [2],

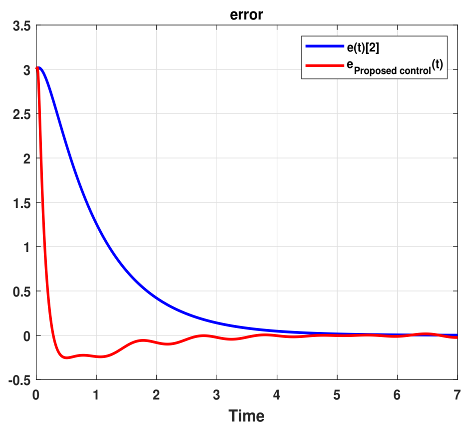

Figure 14: Tracking errors

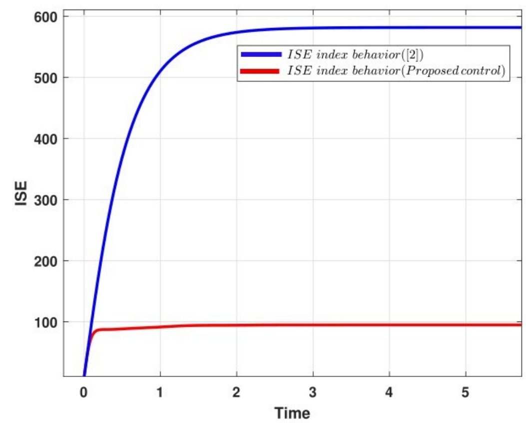

Figure 15: ISE performance indexes

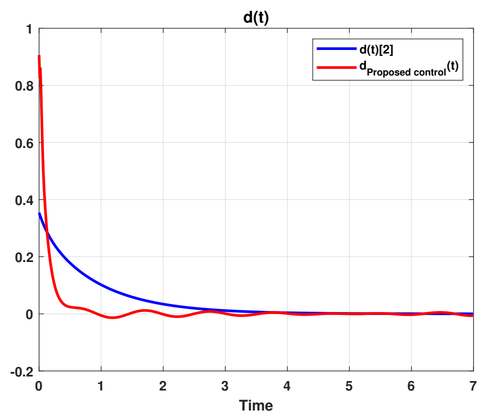

Figure 16: Disturbance functions

Table 3: Comparison of control methods

Method	Control interval	ISE
Method in [2]	$[-1.5, -0.002]$	581.753
Proposed control	$[-0.94853, 3.054]$	95.0197

respectively. To compare the performance of these two controllers, the sum of squared errors were calculated and are given in Table 3. This demonstrates that the proposed control, employing a reasonable control effort, achieves superior tracking performance compared to the method presented in [2]. Furthermore, it significantly reduces the ISE performance index.

Corollary 1. *In classical SMC methods, due to the existence of discontinuous switching function, chattering phenomenon occurs and this chattering is observed in the control signal. In the proposed method, the control input consists of two control signals, u_1 (equivalent controller) and u_2 (RL robust controller). As can be seen in equations ((7) and (13)) and figures(1 and 2), both are continuous. Therefore, the* *chattering phenomenon does not occur.*

8 Conclusion

In this paper, a novel reinforcement learning controller based on sliding mode control method is designed for linear single input -single output class systems with relative uncertainty degree r. The proposed reinforcement learning controller and the external disturbance signal are tuned using an online policy iteration algorithm. For this purpose, an online approximation of a given value function and an optimal policy are calculated in a robust learning approach. Output tracking and filtering of the effects of unmodeled frequencies are guaranteed without any chattering phenomenon. To facilitate analytical analysis and asymptotic stability, three theorems are proved and a new algorithm is designed. In the proposed method, the controller consists of two continuous control signals, the equivalent controller and the reinforcement learning robust controller, which avoids any chattering in the control signals while maintaining optimal output tracking and asymptotic stability. A simulation example demonstrates the advantages of the proposed method well. Moreover, the comparative analysis with conventional sliding mode control methods (sign and tanh) as well as the approach in [2] highlights the superiority of the proposed hybrid controller. The results show that although the proposed controller operates within a slightly wider control range, this increase is mainly due to the stronger disturbances applied in the simulation compared with those considered in [2]. Even under these more challenging conditions, the proposed controller maintains a reasonable control effort and achieves a significantly smaller ISE index (80.88 compared to 169.61 and 335.47 for classical sliding mode control, and 95.01 compared to 581.75 for [2]). In addition, the chattering phenomenon is completely eliminated and the control signals remain smooth. These findings confirm that the proposed learning-based sliding mode controller not only ensures stability and accurate output tracking but also outperforms state-of-the-art techniques in terms of robustness, efficiency, and tracking precision with an acceptable control effort.

In future research, the proposed hybrid learning sliding mode controller can be extended to linear multi-input multi-output systems. Moreover, integrating the approach with deep reinforcement learning algorithms may further enhance adaptability and performance in complex environments.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

[1] M. Aghababa, M. Akbari, *A chattering-free robust adaptive sliding mode controller for synchronization of two different chaotic systems with unknown uncertainties and external disturbances*, J. Appl. Math. Comput. **218** (2012) 5757–5768.

[2] P. An, M. Liu, Y. Wan, F. Lewis, *Multi-player H_∞ Differential game using on-policy and off-policy reinforcement learning*, Sixteenth IEEE International Conference of Control and Automation **1** (2020) 1137–1142.

[3] R. Avram, X. Zhang X, J. Muse, *Nonlinear adaptive fault-tolerant quadrotor altitude and attitude tracking with multiple actuator faults*, IEEE Trans. Control Syst. Technol. **13** (2018) 701–707.

[4] W. Bai, T. Li, S. Tong, *NN reinforcement learning adaptive control for a class of nonstrict-feedback discrete-time systems*, IEEE Trans. Cybern. **1** (2020) 4573–4584.

[5] D.P. Bertsekas, *Dynamic Programming and Optimal Control*, 4th Edition, Athena Scientific, 2012.

[6] S. Chegini, M. Yarahmadi, *Quantum sliding mode control via error sliding surface*, J. Vib. Control **24** (2018) 5345–5352.

[7] J. Choi, Y. Bakch, *Chattering reduction using improved switching functions in the SMC method for battery chargers of PMDs*, IEEE Access **99** (2024) 182663–182672.

[8] T. Ciment, S.P. Banks, *Nonlinear optimal tracking control with application to super-tankers for autopilot designe*, Automatica **40** (2004) 1845–1863.

[9] N. Elseddeq, S. Elghamrawy, A. Eldesouky, *Optimized robust learning framework based on big data for forecasting cardiovascular crises*, Sci. Rep. **14** (2024) 28224.

[10] K. Guo, H. Zhang, C. Wei, H. Jiang, J. Wang, *Novel sliding mode control of the manipulator based on a nonlinear disturbance observer*, Sci. Rep. **13** (2024) 30656.

[11] C. Jia, X. Li, H. Wang, Z. Song, *Sliding mode-based integral reinforcement learning event triggered control*, Int. J. Control Autom. Syst. **23** (2025) 315–331.

[12] B. Kiumarsi, F.L. Lewis, *Actor-critic-based optimal tracking for partially unknown nonlinear discrete-time systems*, IEEE Trans. Neural Netw. Learn. Syst. **26** (2015) 140–151.

[13] J.Y. Lee, Y. Kim, *Hamilton-Jacobi Based Policy-Iteration via Deep Operator Learning*, Neurocomputing **646** (2024) 130515.

[14] J. Li, H. Su, Y. Zhang, J. Chu, *Chattering free sliding mode control for uncertain discrete time-delay singular systems*, Asian J. Cont. **15** (2013) 260–269.

[15] Y. Li, H. Xia, B. Zhao, *Policy iteration algorithm based faulttolerant tracking control: an implementation on reconfigurable manipulators*, J. Electr. Eng. Technol. **26** (2018) 1740–1751.

[16] T. Niknam, M.H. Khooban, A. Kavousifard, M.R. Soltanpout, *An optimal type II fuzzy sliding mode control design for a class of nonlinear systems*, Nonlinear Dyn. **75** (2013) 75–83.

[17] H. Pang, Q. Yang, *Optimal sliding mode output tracking control for linear systems with uncertainties*, International Conference on Machine Learning and Cybernetics **1** (2010) 942–946.

[18] H. Pouyanfar, S. Effati, A. Mansoori, *Dynamic adaptation strategies for optimal control in unknown linear time-invariant system*, J. Math. Model. **14(1)** (2026) 1–17.

[19] R. Shahnazi, *Robust exponential concurrent learning adaptive control for systems preceded by dead-zone input nonlinearity*, J. Math. Model. **12(2)** (2024) 371–385.

[20] J.J.E. Slotine, W. Li, *Applied Nonlinear Control*, Prentice Hall, 1991.

[21] R. S. Sutton, A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd Edition, MIT Press, 2018.

[22] H. Toshani, M. Farrokhi, *Optimal sliding-mode control of linear systems with uncertainties and input constraints using projection neural network*, J. Optim. Control Appl. Methods **39** (2017) 963–980.

[23] D. Vrabie, F. Lewis, *Neural network approach to continuous-time direct adaptive optimal control for partially unknown nonlinear systems*, Neural Networks **22** (2009) 237–246.

[24] S. Xue, B. Luo, D. Liu, Y. Gao, *Integral reinforcement learning-based event-triggered control with input saturation*, Neural Networks **131** (2020) 144–153.

[25] Y. Yang, K.G. Vamvoudakis, H. Modares, Y. Yin, D. Wunsch, *Hamiltonian-driven hybrid adaptive dynamic programming*, J. IEEE Trans. Syst. Man Cybern. Syst. **51** (2020) 6423 –6434.

[26] M. Yarahmadi, M. Karbassi, *Robust wavelet sliding-mode control via time-variant sliding function*, IEICE Trans. Fundam. Electron. Commun. Comput. Sci. **93** (2010) 1181–1189.

[27] B. Zhang, D.W. Ding, *Optimal tracking and regulation performance of networked control systems with additive colored Gaussian noise and finite bandwidth*, J. Appl. Math. Comput. **501** (2025) 129469.

[28] Y. Zhang, S. Zhang, Y. Wang, D. Liu, *Integral sliding mode-based event-triggered nearly optimal tracking control for uncertain nonlinear systems*, Int. J. Robust Nonlinear Control **34** (2023) 2325–2332.