thesis

by H H

Submission date: 10-Jun-2026 12:05PM (UTC+0530)
Submission ID: 2980306334
File name: Thesis.pdf (2.09M)
Word count: 6553
Character count: 37451

CHAPTER 1: INTRODUCTION

1.1 Overview

Numerous industries, including robotics, aerospace, automotive engineering, industrial automation, manufacturing facilities, and electronic systems, heavily rely on control systems. Implementing control systems is essential for proper operation of many contemporary systems and processes. The development of intelligent and automated systems has accelerated during the past few decades. Automated and intelligent control systems must be capable of performing their functions effectively regardless of any disruptions or environmental changes. Therefore, efficient control systems must account for uncertainties, disturbances, and other non-linearities. Research studies often adopt an inverted pendulum as the basis of experiments due to its simplicity of construction although it is difficult to stabilize [1]. It is an established fact that the inverted pendulum system is inherently unstable especially when it is positioned upright. It will be difficult for the pendulum to stay stable because of any disturbance [2].

Rotating inverted pendulum (RIP) is another advanced form of inverted pendulum systems. The inverted pendulum is suspended by a shaft. This shaft is termed as a rotary arm and it would move due to the impact brought about by the motor [3]. The rotary arm moves horizontally as the pendulum moves vertically thereby giving rise to the creation of a non-linear dynamic system. The rotary inverted pendulum is highly recommended as an excellent platform for studying numerous engineering problems. The designing of self-balancing robots, spacecraft attitude control, missile guidance systems, and walking robots are among such problems.

1.2 A brief description of the system

Fig 1: Rotary Inverted Pendulum [4]

The rotary arm is subjected to an action by the motor such that the pendulum swings in the vertical direction due to the effects of gravity. The system has two degrees of freedom. Rotary arm provides one degree of freedom, while the pendulum provides the other degree of freedom. The system exhibits non-linear and coupled properties as a result of the coupling between these two variables. The control system's primary goal is to keep the inverted pendulum stable. In this context, it is crucial to remember that the pendulum's inverted posture is intrinsically unstable [5]. In order to maintain balance, the controller must continuously apply the proper torque to the arm.

1.3 Challenges in a control of non-linear Systems

Non-linear control systems are more difficult to design than linear ones. Controllers designed based on linear model analysis work well only close to the equilibrium point. In case of disturbance, the controller performance may degrade. The pendulum must be stabilised by indirect manipulation of the rotating arm as it is not directly controlled by a force. Disturbances, friction, sensor noise, and parameter changes all have a significant impact on this system. Thus, all the aforementioned uncertainties need to be considered [4].

1.4 Motive of research

- This research project focuses on design and analyse a controller based on a hybrid approach using the method of sliding mode control and reinforcement learning to stabilise the rotating inverted pendulum.
- The second objective is to analyse non-linear dynamics of the system. Additionally, performance analysis for the controller under investigation forms another key purpose of the study.

CHAPTER 2: LITERATURE REVIEW

2.1. Existing work

A rotary inverted pendulum system has shown to be a standard system for analysis of a variety of control strategies as a result of its non-uniformity. The PID controller is among the oldest and most basic controllers that have been designed for inverted pendulum system's stabilisation [6]. It is one of the simplest controllers because of its easy mathematical analysis and design. The PID controller can only stabilize a certain region of the inverted pendulum system because of its non-linearity.

State feedback control methods have also been implemented for inverted pendulum stabilization. These methods use state variables such as angular displacement and velocity of pendulum and rotational arm to generate control actions. State feedback controllers generally provide better dynamic performance compared to simple output feedback methods because they use more information about system behaviour [7].

Although there are many benefits associated with the classical approach, it is noted that classical control designs depend on the linearising of dynamic models of the system. Considering the fact that the behaviour of a rotary inverted pendulum is highly non-linear and coupled, the use of classical controllers may not work well under varying operating conditions [4]. This has encouraged many researchers to investigate new non-linear and intelligent control designs.

Some of the commonly applied control methodologies comprise of PID control, Root locus, Frequency response and pole placement control. They were capable of stabilizing the system near the equilibrium point but were ineffective in achieving non-linear control [6,9,10]. Later researchers have explored the use of modern control approaches including Linear Quadratic Regulator (LQR), control via adaptive, fuzzy logic and neural networks for rotary inverted pendulums [11,12]. Modern control approaches have shown higher effectiveness than traditional controllers. Sliding Mode Control (SMC) is considered one of the most efficient non-linear control techniques for ensuring stability of the RIP. It has been established that sliding mode control is an effective approach that performs better in terms of convergence speed, disturbance resistance, and reduced steady state errors compared to traditional controllers [8].

2.2. Research gaps

Although there have been advances made in inverted pendulum control, some issues remain in research. The following are the issues

- Classical and modern control techniques rely extensively on mathematical modelling and linear approximations. The effectiveness of such methods might be unreliable when applied under non-linear conditions.
- Sliding Mode Control is robust yet plagued by the drawbacks of chattering and poor tuning. The selection of sliding surface parameters is generally difficult and relies on trial-and-error techniques.

Hybrid control algorithms are not yet optimized for real-time experimental implementation and thus, further investigation is needed regarding the hybrid control.

2.3. Research objective

The research objectives can be mentioned as follows:

- The first objective is to implement control of RIP using the SMC strategy.
- The second objective is to further integrate the Sliding Mode Controller with a Reinforcement Learning Algorithm platform to observe the effectiveness of the hybrid approach

CHAPTER 3: SYSTEM MODELLING OF RIP

3.1. Characteristics

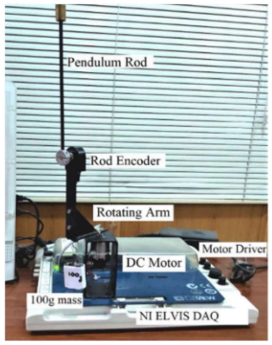

Fig 2 : Rotary inverted pendulum with actuators and data acquisition system [15]

A rotary inverted pendulum generally includes components namely, a rotary arm, a pendulum, a DC servo-motor, motor driver for running the servo-motor, sensors such as encoders for measuring the angle of rotation of pendulum and rotary arm of RIP, current sensors for measuring the current to control the servo-motor torque, DAQ devices for data acquisition and supporting mechanical structure. A rotary arm is mounted onto a DC servo-motor that generates horizontal rotational movement [4]. A pendulum rod is fixed at the end of the rotary arm in a way that makes pendulum movement possible in the vertical plane. Two types of rotations mentioned here are the factors that make the system strongly non-linear. The motion of a pendulum is determined by the generated torque at the rotary arm whereas the latter is influenced by pendulum dynamics [8].

Despite having 2 degrees of freedom, the system has only 1 actuating component (i.e) the servo-motor [13]. Torque of the servo-motor influences rotary arm movement whereas the pendulum movement is regulated through the rotary arm movement [4]. The RIP system can be called an under-actuated system since it has got degrees of freedom greater than the no of actuators. Two equilibrium points exist in the system: one when the pendulum is pointing downward and one when it is pointing upward. While the lower-pointing equilibrium point is inherently stable, because gravity acts to push it in this direction, the upper equilibrium point is unstable [15]. Control actions are generally divided into two processes called swing-up and stabilization control [1, 3, 4].

Swing-up involves transferring energy into the pendulum to make it rotate around its axis until the pendulum reaches the upper equilibrium point [1]. After the pendulum comes close to the upper equilibrium point, a stabilization controller comes into play. The dynamics of the inverted pendulum are highly susceptible to disturbances, friction, measurement noise, and uncertainty [15]. For this reason, precise modelling is necessary for designing effective controllers.

Applications of rotary inverted pendulum include robotics, aerospace, self-balancing vehicles, and humanoid robots. Several stabilization techniques applied in the RIP are the same as those in various engineering systems such as attitude control of spacecraft or humanoid robotics [15]. The nature of non-linear dynamics combined with instability of the inverted pendulum makes it perfect for testing non-linear and intelligent control methods.

3.2. Assumptions

Some assumptions are made when modelling the rotary inverted pendulum system. These assumptions simplify the modelling of a system.

The first assumption that is generally made concerns the structure of the pendulum and arm. They are rigid bodies, which implies no elastic deformation and structural flexibility when making calculations. Making such an assumption is acceptable when studying systems in laboratories [16].

There is also an assumption about the presence of ideal joint bearing with negligible friction in real conditions. It should be noted that there is some friction but it is usually neglected when carrying out mathematical calculations. It can be taken into account at a later stage.

It is generally assumed that the pendulum rod has constant density, which makes it possible to easily calculate the location of centre of mass and the angular mass using basic formulas.

Effect of air resistance acting on the pendulum and rotary arm is usually neglected due to its relatively insignificant effect on the object's motion [17].

It is often assumed that the motor can generate enough torque instantly upon receiving control signals. Although actuators are subject to saturation, the motor itself should not create any additional complications at the first stage of controller design [7].

During research related to the stabilization process of a pendulum in the neighbourhood of the vertical equilibrium state, researchers may use small angle approximation. This is associated with the ability to simplify equations by excluding trigonometric functions $\sin(\theta)$ and $\cos(\theta)$ [11]. Such an assumption does not work in large amplitude swing-up processes.

It is also assumed that disturbances, whose effects can be neglected, are not taken into consideration when designing models [16].

It should be pointed out that although all these assumptions facilitate modelling, they can lead to modelling error when practically implementing the controllers.

3.3. State-space modelling

State-space modelling is used in the analysis and control of dynamical systems since this method offers a concise mathematical description of system performance. The rotary inverted pendulum system can be modelled using a state vector consisting of angle of rotary arm (with respect to horizontal axis), angle of the pendulum (with respect to vertical axis), angular velocity of rotary arm and angular velocity of pendulum [5]. A state vector can be written as follows:

$$x = \begin{bmatrix} \theta \\ \alpha \\ \dot{\theta} \\ \dot{\alpha} \end{bmatrix} \tag{1}$$

θ – angular rotation of arm

α – angular rotation of pendulum

$\dot{\theta}$ – arm's angular velocity

$\dot{\alpha}$ – pendulum's angular velocity [17].

Torque is the signal input to the rotary arm and it is obtained from the servo-motor.

The continuous-time state-space equations are represented as follows [16]:

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{2}$$

$$y(t) = Cx(t) + Du(t) \tag{3}$$

Under these circumstances, x denotes vector of state, u denotes control input, y denotes vector of output, A denotes matrix of system, B denotes input matrix, C denotes output matrix and D represents matrix of transmission [16].

This matrix D shows us the connection between what we put into a system and what we get out of it without having to think about the state of the system. In a lot of pendulum systems the matrix D is usually zero. This is because what we put into the system does not immediately affect what we get out of it.

Therefore from the above, the state space representation of the system is given by:

$$\begin{bmatrix} \dot{\theta} \\ \dot{\alpha} \\ \ddot{\theta} \\ \ddot{\alpha} \end{bmatrix} = A \begin{bmatrix} \theta \\ \alpha \\ \dot{\theta} \\ \dot{\alpha} \end{bmatrix} + Bu \tag{4}$$

This representation makes it easier to do analysis and simulation of the system. It also helps to use control methods and estimation techniques effectively. In systems we cannot measure all the state variables directly with sensors. So state estimation techniques are implemented to guess the states that cannot be measured.

The error (e) is the difference between current (x) and projected state ($\hat{x}$).

$$e = x - \hat{x} \tag{6}$$

For a control system that uses state-feedback, [3,7]:

$$u = -Kx \tag{7}$$

Here, K refers to the feedback matrix, u refers to the control signal and x refers to the state matrix.

Matrix K is chosen to make the system stable and reduce oscillations [3]. Selecting gain properly will have an impact on how system works. In some control systems we also have a reference input matrix called N. This N matrix is for tracking the reference. So the control input can be represented as :

$$u = -Kx + Nr \tag{8}$$

r - reference input signal

N - reference gain matrix.

The matrix N helps the system follow desired reference values accurately.

CHAPTER 4: SLIDING MODE CONTROL (SMC)

4.1. Fundamentals

SMC is a resilient approach where control input enables the system trajectory to reach and stay within the set system path termed as the sliding surface, thus ensuring stable and reliable performance even under parameter variations. It is applicable for systems that are not stable and can face non-linearities and disturbances [18]. Robotics, automotive systems, aerospace systems etc, are among the fields in which a Sliding Mode Controller can be applied [19].

4.2. Mathematical Equations

SMC generates control signal that drives the system towards the pre-defined path [20]..

The sliding surface of SMC can be represented as:

$$s = \lambda e + \dot{e} \tag{9}$$

e - tracking error

λ - design parameter

s - sliding surface.

The control signal of SMC can be represented as:

$$u = -k \operatorname{sgn}(s) \tag{10}$$

k - switching gain,

$\operatorname{sgn}(s)$ - signum function of the sliding surface [22].

4.3. Reaching Phase and Sliding Phase

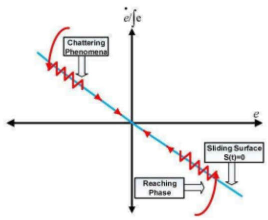

Fig 3: Reaching and Sliding phase in a Sliding Mode Controller [42]

SMC operates in two phases namely, reaching and sliding phase [19,21]. With support of initial action, the system trajectory moves towards predefined surface. This is called reaching phase [22]. Once the system gets to the pre-defined surface, it is termed as sliding phase. After that, behaviour of system will adhere to the pre-defined path.

4.4. Advantages and Disadvantages of SMC

Table 1 : Merits and demerits of a Sliding Mode Control

<table>
<tr><th>Advantages</th><th>Disadvantages</th></tr>
<tr><td>Some of the key advantages of applying SMC include resilience to the effects of disturbances and parameter uncertainties. Sliding Mode Control allows obtaining adequate results when facing uncertainty and disturbances [24,25].</td><td>SMC is associated with a drawback of chattering. This is when the system moves back and forth quickly due to frequent switching thus, reducing the efficiency of the control signal [25].</td></tr>
<tr><td>The behaviour of sliding mode surface is independent of the initial states [26].</td><td rowspan="2">It is sensitive to noise [27].</td></tr>
<tr><td>The algorithm of control allows achieving the point of equilibrium relatively quickly compared to other techniques [23].</td></tr>
</table>

CHAPTER 5: REINFORCEMENT LEARNING

5.1. Fundamentals of Reinforcement Learning

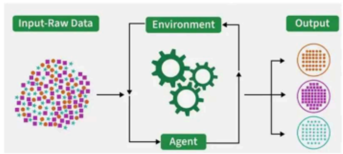

Fig 4 : Reinforcement Learning Architecture [29]

Reinforcement Learning (RL) refers to a method where an agent is employed to make decisions by interacting with the environment. Ultimate goal of RL is to design the best possible control policies that maximizes total rewards. Reinforcement learning does not rely on labelled data sets or a mathematical model for training algorithms but it uses trial and error approach to learn and operate accordingly [28].

RL agent detects the present environmental scenario and decides the right action to be taken [28]. Once an action is completed, the RL agent is rewarded by the system (or environment) and then it moves to a new state. Subsequently, the learning algorithm adjusts itself based on the obtained reward.

The growing popularity of RL has led to its successful applications in various areas. In particular, researchers apply RL in robotics, autonomous driving, automation processes and other non-linear control systems.

RL techniques can be classified into different groups like Value-based techniques and Actor-critic techniques [31].

5.2. Value-Based Techniques

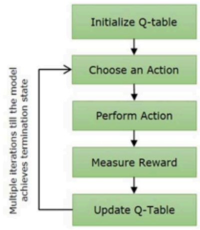

Fig 5 : Block diagram of Q-learning method [32]

Value-based reinforcement learning techniques calculate the projected total rewards from different state-action pairs. These techniques use value functions to evaluate the long-term benefit of actions. Value-based approaches have been popular due to simplicity and theoretical justification.

Q-learning is a widely accepted value-based learning method. Q-learning uses a Q-value function that estimates the projected future reward for performing a certain task in a specific state. One advantage of Q-learning is that it doesn't need prior understanding of the system's behaviour. The optimal actions are decided by interacting with the environment multiple times. However, Q learning is inefficient for high-dimensional continuous state space problems since Q-value lookup tables need to be constructed [33].

Table 2 : Q value look up tables [34]

Q-Table	
State-Action	**Q-value**
-	0
-	0
-	0
-	0
-	0
-	0

Deep Q-Networks (DQN), which was proposed later to solve this problem, incorporates the Q-learning algorithm needed to approximate the Q-functions using deep neural networks. It enables the algorithm to deal with complex and continuous environments [33].

Although DQN is an effective algorithm, the instability issue still exists in training due to the correlation in the samples collected and changes in target values [35].

Two solutions have been found to address these problems: Experience Replay and Target Network. Experience replay stores the experiences obtained in the memory pool for sampling at random when training. The target network is a separate network used to stabilise learnings by providing fixed target Q-values during updates.

As a solution, DQN was developed, which applies Q-learning and neural networks [36]. DQN utilizes neural networks that enable an algorithm to adapt to a complex continuous environment.

5.3. Actor-Critic Methods

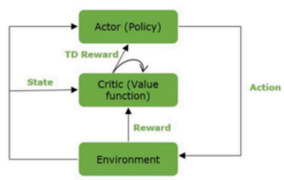

Fig 6: Actor critic Method of Reinforcement Learning [37]

Actor-critic methods use the benefits of value-based and policy-based reinforcement learning methods. Actor and critic are the 2 major components [38]. The actor module selects actions or decisions which are determined by existing policies. The critic verifies by analysing actor's actions using either the value function or reward signals. The critic provides evaluations of the actor's actions to the actor. The actor, then refines its policy depending upon the received feedback.

Actor-critic techniques reduce the high variance associated with pure policy-gradient techniques. At the same time, the ability to operate effectively in continuous action spaces is still retained.

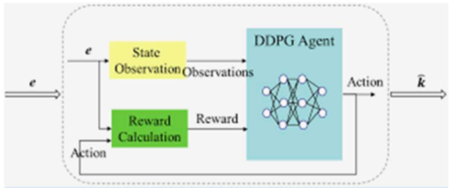

Fig 7: DDPG technique [46]

Among the popular actor-critic techniques, there is a popular method named as Deep Deterministic Policy Gradient (DDPG). This method uses neural networks (NN) with deterministic policy gradient [39,40]. Many researchers have successfully applied DDPG in robotic manipulation and self-driving vehicles. The algorithm performs well in continuous action spaces.

However, DDPG suffers from instability and overestimation bias during training. These limitations encouraged the creation of a technique named as Twin Delayed Deep Deterministic Policy Gradient (TD3).

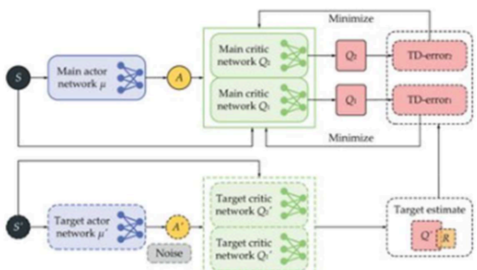

Fig 8: TD3 technique [47]

TD3 enhances process of learning by means of applying double Q-learning, delayed updating, and policy smoothing. It is obvious that results are significantly better for TD3 compared to DDPG [41].

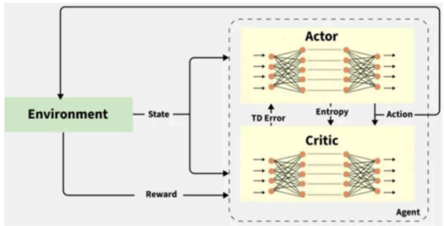

Fig 9: SAC method [48]

There exists another well-known actor-critic model that is named as Soft Actor-Critic (SAC). SAC offers stable learning and high robustness [38].

Currently, actor-critic algorithm is one of the most efficient methods for non-linear control problems. Their ability to handle continuous state-action spaces, actor-critics fit rotary inverted pendulum stabilization well.

CHAPTER 6: ENHANCEMENT OF SLIDING MODE CONTROL USING REINFORCEMENT LEARNING

6.1. Motivation for a Hybrid Approach

Non-linear control strategies are crucial in the modern non-linear systems such as the rotary inverted pendulum where uncertainty of operational environment is common. The conventional methods have not been proven to be effective when non-linearities or modelled parameter variations are noticeable.

SMC is robust due to disturbance resistance and tackling uncertainties. The SMC makes the system's trajectory shift towards the chosen sliding surface while keeping system dynamics along it [42]. In this regard, the stability and disturbance rejection capabilities of the SMC are guaranteed.

Nonetheless, there are many disadvantages to using SMC. First of all, there is the problem of chattering due to the presence of discontinuous switching control inputs. Chattering may cause physical wear of actuators, waste of power and even decrease overall system effectiveness [43].

The second disadvantage relates to the selection of appropriate SMC parameters and surface coefficient which proves rather challenging. The improper selection of the mentioned parameters may result in such problems as low convergence rate or lack of stability [42].

Reinforcement Learning has gained much popularity in recent times as an intelligent control technique due to its capability to learn through interactions. In recent studies, RL-based controllers were found to show high-quality performance for robotics applications, autonomous vehicles, and non-linear dynamic systems. RL and SMC combination may result in a controller with superior features [42]. Within the combined approach, SMC contributes robustness and stability while RL ensures adaptability and parameters adjustment.

In rotary inverted pendulum systems, the hybrid approach appears to be very beneficial due to the non-linearity, under-actuation, and sensitivity to external disturbances.

Combining learning capabilities with robust control makes the stabilization process more effective. A hybrid controller can thus, help achieve high robustness, adaptability, learning effectiveness, and real-time control performance [44].

6.2. Function of RL in Gain Parameter Tuning

Parameter tuning is a very important factor in the design of SMC. Performance of SMC greatly depends upon choosing appropriate sliding surfaces and switching gains [42].

Historically, switching gain was chosen by trial and error method using expert knowledge. But this approach becomes complex when applied to highly non-linear dynamic systems like rotary inverted pendulum. A small value of switching gain prevents the convergence to the sliding surface. In contrary, very high value might cause increase in chattering.

Adaptive parameter tuning using Reinforcement Learning offers an intelligent way to overcome this challenge. RL agents are able to observe system dynamics and tune parameters based on feedback information.

By interacting with environment, RL agents are capable of learning the best parameter values. While training process, the RL agents evaluate performance based on system stability, tracking errors and control efforts [45].

RL techniques allow changing the value of switching gain dynamically depending on current situation. By doing this, it will be possible to improve disturbance rejection and minimize chattering effect.

RL algorithms can be utilized for optimization of sliding surfaces. Sliding surface design greatly affects the speed of convergence and stabilization. Some studies revealed the superiority of RL methods in comparison to fixed parameters SMC algorithms due to adaptation capabilities.

DDPG and TD3 belong to Deep Reinforcement Learning algorithms that perform best when used for parameter tuning since they operate with continuous values. Parameter tuning in these methods is conducted smoothly [40, 41].

Adaptive parameter tuning using RL leads to better system results, decreases human interference, and increases flexibility of controllers. Thus, RL-SMC is applicable to real control problems.

6.3. Algorithm Design and Workflow

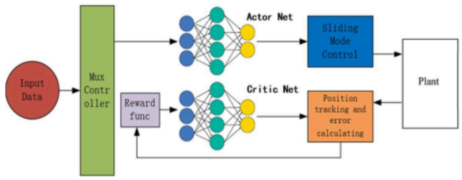

Fig 10: RL-SMC Architecture [42]

The design of an RL-SMC algorithm involves several stages including state observation, action selection, reward evaluation, parameter updating, and control signal generation. Proper workflow design is essential for achieving stable and efficient learning behaviour.

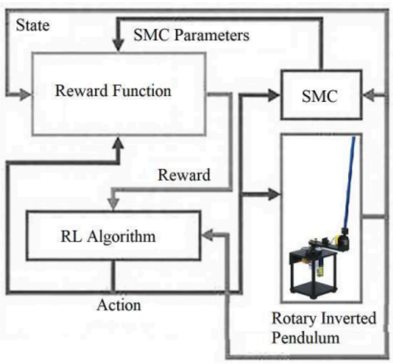

Fig 11: Recommended architecture

The first step begins with defining system’s state vector. In the case of rotary inverted pendulum, the state vector usually contains parameters such as angular rotation and velocity of arm and pendulum.

The second step further proceeds with defining the action space. The actions generated by the RL agent may include control gains, sliding surface gains, or other control inputs.

The reward function is then designed according to control objectives. The reward function usually penalizes tracking error, excessive oscillation, and high control effort.

The RL agent determines the appropriate action on the basis of the current policy after obtaining the signal's current condition. The selected action is then passed on to the SMC component for generating the control signal.

This process is iterative since it continues until the RL agent learns an optimal policy for stabilization and control. The learning workflow improves continuously through repeated interaction with the environment.

The overall workflow of RL-SMC can therefore be summarized as observation, learning, adaptation, and control. This workflow enables intelligent and robust control of non-linear systems.

6.4. Advantages over Conventional Methods

Hybrid RL-SMC controllers provide several advantages over traditional and standalone control methods. These advantages make RL-SMC suitable for non-linear, uncertain, and under-actuated.

The first advantage of SMC is an increased robustness against disturbances and uncertainties in system's parameters. It guarantees the stabilization of the system despite external perturbations. The second advantage of SMC is adaptive learning [45]. It becomes possible to improve the controller's performance while interacting with surroundings.

In addition, unlike other controllers with fixed values of the parameters, RL-SMC controllers automatically change the settings depending on operational circumstances. It leads to reduced reliance on expert knowledge and manual tuning.

Moreover, RL-SMC controllers can achieve higher stabilization performance and faster convergence.

Another benefit is the ability of the RL-based controller to perform adaptive tuning, which can reduce chattering. Adaptive tuning produces smoother control signals and prevents high-frequency switching oscillations.

RL-SMC schemes also provide enhanced disturbance rejection and fault tolerance, which make these algorithms well suited for practical engineering systems that have uncertain operating conditions [43]. On the whole, the combination of Reinforcement Learning with Sliding Mode Control provides a unified control strategy, which features robustness, adaptability, learning capabilities, and intelligent optimization.

CHAPTER 7: SIMULATION

7.1. Simulation setup

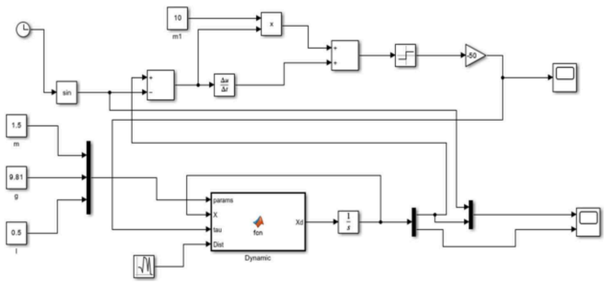

Fig 12: SMC for RIP (Separate setup)

In the above figure, a Simulink model for RIP controlled by SMC has been developed. The parameters to be monitored here are the angular rotation of arm and pendulum. The behaviour of the RIP considering these two parameters can be observed with the help of the Scope block.

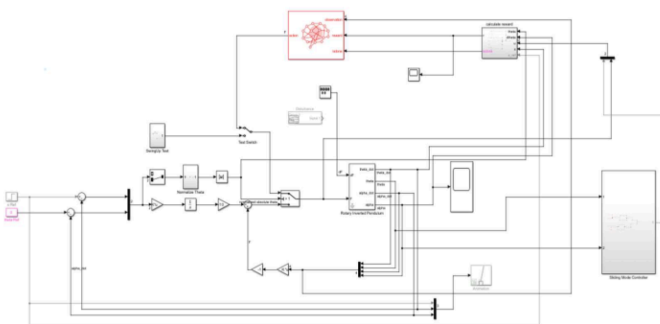

Fig 13: Complete Simulink Model (RL – SMC)

Figure 13 represents the complete Simulink setup for the hybrid Reinforcement Learning – Sliding Mode Control strategy. The setup includes an RL agent that learns from the system states as well as reward signals, a Sliding Mode Controller block, Rotary Inverted Pendulum and a Reward calculation block that provides feedback (also known as reward) to the RL agent based on each action taken by it. The RL algorithm implemented here was Deep Deterministic Policy Gradient (DDPG) and it was set for 100 episodes. The waveforms of the angle of pendulum as well as the angle of rotary arm can be observed with the help of the Scope block.

7.2. Results

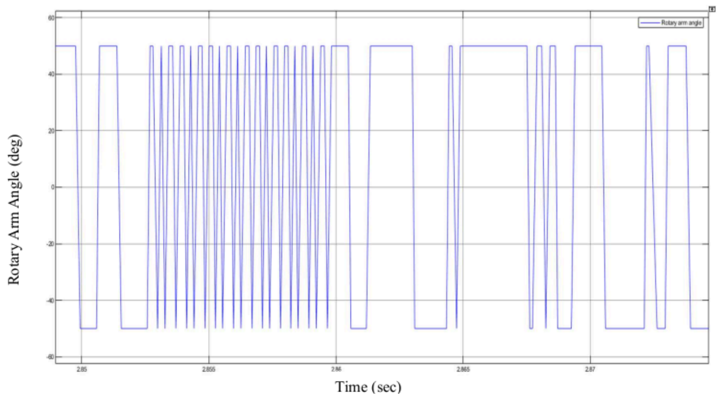

Fig 14: Graph of rotary arm angle of RIP using SMC

From figure 14, waveform of angular rotation of the arm when controlled by Sliding Mode Control (SMC) can be observed. It is clear from the graph that the rotary arm swings very sharply between -50° and +50°. This kind of sharp change in the angular rotation of the arm indicates the quick response of the control strategy to direct the system's trajectory to the desired sliding surface. Oscillations of such types are known as chattering and occur as a result of continuous switching action performed by SMC controllers. However, despite the sharp variations in the swing angle, the controller successfully manages to keep the rotary arm within certain bounds, which keeps the RIP stable.

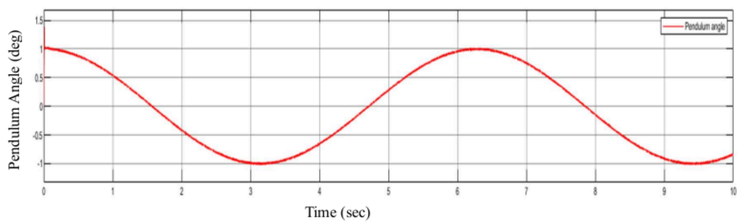

Fig 15: Graph of Pendulum angle of RIP using SMC

Figure 15 shows the waveform of the angle of pendulum upon the implementation of SMC strategy. There is an oscillation ranging from $+1°$ to $-1°$. This depicts that SMC is able to control the response of the pendulum but the waveform fails to settle at a later stage indicating a slight instability of the pendulum. Thus, we can say that SMC can effectively deal with the non-linearity of rotary inverted pendulum yet the limitation of slight stability still remains.

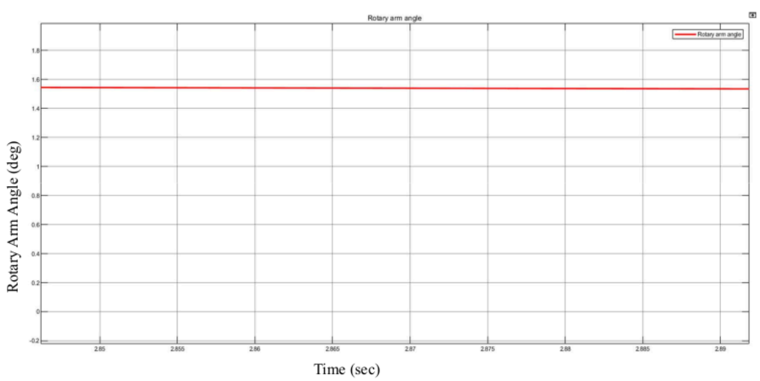

Fig 16: Graph of rotary arm angle of RIP using RL-SMC

Figure 16 represents the waveform of angle of rotary arm. It can be observed that the waveform remains fixed throughout the observed time period, which indicates that rotary arm is at a stable position compared to the SMC strategy. However, this is not practical for a rotary inverted pendulum. In actual operation, the rotary arm keeps moving rapidly to maintain a state of equilibrium for the pendulum.

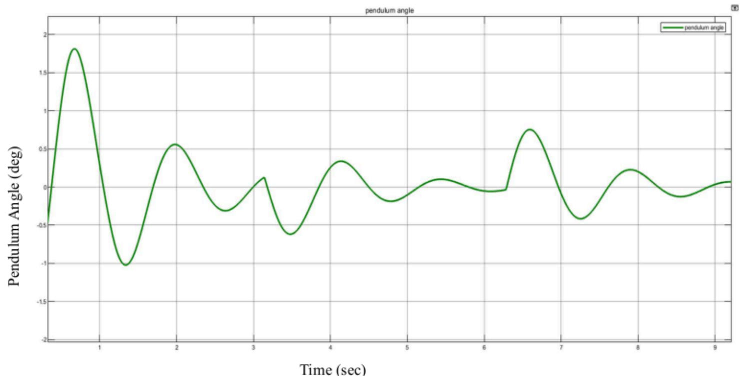

Fig 17: Graph of pendulum angle of RIP using RL-SMC

In figure 17, the waveform of the angle of pendulum is represented with the help of Simulink platform. Initially, there is a significant overshoot indicating an initial deviation from the equilibrium position. The amplitude of oscillations, however, decreases with time indicating the controller is able to stabilise the system. Though slight oscillations are visible, an improvement in the performance is visible compared to the Sliding Mode Controller.

7.3. Discussion of Results

- The Simulation results indicate that both SMC and RL-SMC are capable of generating appropriate control signals but their responses are varied.
- The integration of Reinforcement Learning with Sliding Mode Controller helped in improving damping of oscillations and also pushing the system trajectory closer to the equilibrium position, compared to conventional Sliding Mode Controller.
- The results further demonstrates that RL-SMC can be further improved for obtaining a better performance for non-linear systems and also further increasing the number of episodes.

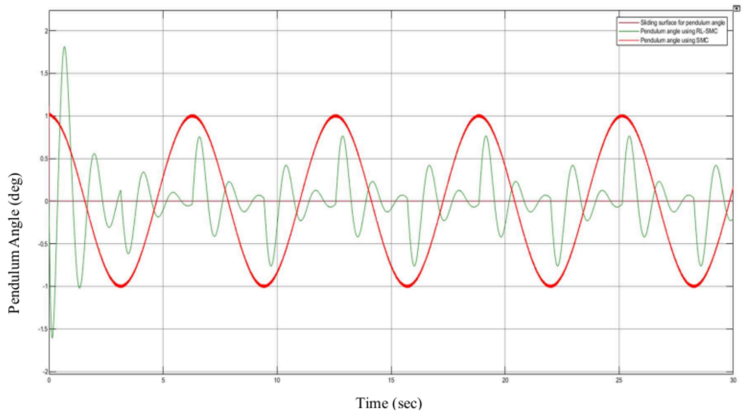

Fig 18: Graph for comparison of SMC vs RL-SMC for pendulum angle

Fig 14 helps us identify the difference in performances demonstrated by the RL-SMC strategy and the difference in parameters have been concluded in Table 3.

Table 3: Comparison of SMC vs RL-SMC for pendulum angle

Parameter	**SMC**	**RL-SMC**	**Improvement demonstrated by RL-SMC**
Rise time	0.7 s	0.3 s	57.1% faster
Peak time	1.5 s	0.5 s	66.7% faster
Overshoot	10%	90%	80% higher than SMC (Not an improvement)
Steady state error	1 deg	0.05 deg	95% reduction
Damping coefficient	0	0.18	Damped System

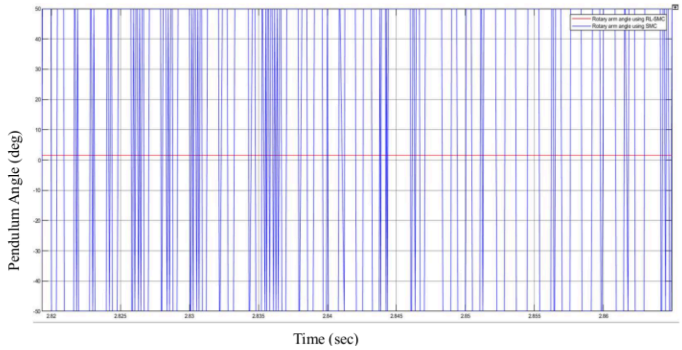

Fig 19: Graph for comparison of SMC vs RL-SMC for rotary arm angle

Fig 15 helps us identify the difference in performances demonstrated by the RL-SMC strategy and the difference in parameters have been concluded in Table 4.

Table 4: Comparison of SMC vs RL-SMC for rotary arm angle

Parameter	**SMC**	**RL-SMC**	**Improvement demonstrated by RL-SMC**
Rise time	0.001 s	0.005 s	400% slower (Not an improvement)
Peak time	0.001 s	0.005 s	400% slower (Not an improvement)
Overshoot	Not applicable (Due to absence of a target value)	Not applicable (Due to absence of a target value)	-
Steady state error	50 deg	2.5 deg	95% reduction

CHAPTER 8: SUMMARY

8.1. Inference

The performance of the hybrid control strategy of Reinforcement Learning assisted Sliding Mode Control of RIP was verified with the help of MATLAB/Simulink. The robustness and adaptive skills of RL-SMC were demonstrated. The simulation results helped in concluding that the robust control strategy was able to maintain stability. The findings also suggest that the hybrid control strategy could be used for other non-linear systems as well

8.2. Future Scope

The future work may focus on the following areas:

- The proposed control strategy of Reinforcement Learning integrated with Sliding Mode Control can be implemented on the hardware setup to understand the working in a practical environment.
- The hybrid approach can be useful for other non-linear and under-actuated bodies such as UAVs (Unmanned Aerial Vehicles) or drones, Walking robots and ASVs (Autonomous Surface Vessels)
- Future works can explore the use of more advanced RL algorithms to improve learning effectiveness and accurate control

REFERENCES

[1] Åström, Karl Johan, and Katsuhisa Furuta. "Swinging up a pendulum by energy control." *Automatica* 36, no. 2 (2000): 287-295.

[2] Spong, Mark W. "The swing up control problem for the Acrobot." *IEEE Control Systems Magazine* 15, no. 1 (1995): 49-55.

[3] Furuta, Katsuhisa, Masaki Yamakita, and Sei Kobayashi. "Swing-up control of inverted pendulum using pseudo-state feedback." *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering* 206, no. 4 (1992): 263-269.

[4] T.-N. Ho and V.-D.-H. Nguyen, "Model-Free Swing-Up and Balance Control of a Rotary Inverted Pendulum using the TD3 Algorithm: Simulation and Experiments", *Eng. Technol. Appl. Sci. Res.*, vol. 15, no. 1, pp. 19316–19323, Feb. 2025.

[5] Hernandez R, Garcia-Hernandez R, Jurado F, "Modeling, Simulation, and Control of a Rotary Inverted Pendulum: A Reinforcement Learning-Based Control Approach." *Modelling.* vol 5, no. 4, pp 1824-1852, 2024.

[6] Prasad, L.B., Tyagi, B. & Gupta, H.O. Optimal Control of Nonlinear Inverted Pendulum System Using PID Controller and LQR: Performance Analysis Without and With Disturbance Input. *Int. J. Autom. Comput.* 11, 661–670 (2014).

[7] Jithin Babu, Elizabeth Vargheese, 2015, Stabilization of Rotary Arm Inverted Pendulum using State Feedback Techniques, *International Journal of Engineering Research & Technology* Volume 04, Issue 07 (July 2015)

[8] Sanjeewa SDA, Parnichkun M. Control of rotary double inverted pendulum system using LQR sliding surface based sliding mode controller. *Journal of Control and Decision*, 2022, 9(1), 89–101.

[9] Muñoz-Poblete, "Pole placement controller applied to a Rotary Inverted Pendulum System. A didactic view," *IEEE International Conference on Automation/XXIII Congress of the Chilean Association of Automatic Control*, pp. 1-6, 2018

[10] R. Banerjee and A. Pal, "Stabilization Of Inverted Pendulum On Cart Based On Pole Placement and LQR," *International Conference on Circuits and Systems in Digital Enterprise Technology*, pp. 1-5, 2018

[11] F. A. Darma, A. D. A. D. Isa, A. M. Abdullahi, L. B. Ila, "Fuzzy logic control of a rotary double inverted pendulum system," *Applications of Modelling and Simulation*, vol. 2, no. 1, pp. 26–33, 2018.

[12] H. Zhang, H. Fang, D. Zhang, X. Luo, and Q. Zou, "Adaptive fuzzy sliding mode control for a 3-dof parallel manipulator with parameters uncertainties," *Complexity*, vol. 2020, pp. 1–16, 2020.

[13] Grasser, Franz, Aldo D'Arrigo, Silvio Colombi, and Alfred C. Rufer. "JOE: A mobile, inverted pendulum." *IEEE Transactions on Industrial Electronics* 49, no. 1 (2002): 107-114.

[14] Saleem O, Mahmood-ul-Hasan K. "Adaptive State-space Control of Under-actuated Systems Using Error-magnitude Dependent Self-tuning of Cost Weighting-factors." *Int J Control Autom Syst*. 2021; 19 (02): 931–941.

[15] Astom KJ, Apkarian J, Karam P, Levis M, J Falcon, *Student Workbook: QNET Rotary Inverted Pendulum Trainer for NI ELVIS.* Ontario: Quanser Inc.; (2011).

[16] N. Lenka, "Modeling and Controller Design for an Inverted Pendulum System," B.Tech thesis, Department of Electrical Engineering, National Institute of Technology Rourkela, 2011

[17] I. Tejado, D. Torres, E. Pérez, B. M. Vinagre, "Physical Modeling based Simulators to Support Teaching in Automatic Control: The Rotary Inverted Pendulum." in *Proceedings of 11th IFAC Symposium on Advances in Control Education*, vol. 49, no. 6, pp. 75-80, 2016

[18] Akhtaruzzaman, Md, and Amir Akramin Shafie. "Modeling and control of a rotary inverted pendulum using various methods, comparative assessment and result analysis." *IEEE International Conference on Mechatronics and Automation*. IEEE, 2010.

[19] Utkin, Vadim I. *Sliding Modes in Control and Optimization*. Springer Science & Business Media, 2013.

[20] Young, Kenneth D., V. I. Utkin, and Ümit Özgüner. "A control engineer's guide to sliding mode control." *IEEE Transactions on Control Systems Technology* 7, no. 3 (1999): 328-342.

[21] Utkin, Vadim I. "Variable structure systems with sliding modes." *IEEE Transactions on Automatic Control* 22, no. 2 (1977): 212-222.

[22] Edwards, Christopher, and Sarah K. Spurgeon. "Sliding mode control: Theory and applications." *IEE Proceedings-Control Theory and Applications* 136, no. 6 (1989): 343-352.

[23] C. Edwards and S. Spurgeon, *Sliding Mode Control: Theory and Applications (1st ed)*. London, UK. CRC Press, Taylor and Francis, 1998

[24] Bartolini, Giorgio, "Modern sliding mode control theory." *Lecture Notes in Control and Information Sciences* 375 (2008): 1-42.

[25] Levant, Arie. "Higher-order sliding modes, differentiation and output-feedback control." *International Journal of Control* 76, no. 9-10 (2003): 924-941.

[26] Shtessel, Yuri, Christopher Edwards, Leonid Fridman, and Arie Levant. "Sliding mode control and observation." *International Journal of Robust and Nonlinear Control* 19, no. 15 (2009): 1657-1659.

[27] Utkin, Vadim, Jürgen Guldner, and Jingxin Shi. "Sliding mode control in electro-mechanical systems." *IEEE Transactions on Automatic Control* 45, no. 11 (2000): 2168-2169.

[28] Lewis, Frank L., Draguna Vrabie, and Kyriakos G. Vamvoudakis. "Reinforcement learning and feedback control: Using natural decision methods to design optimal adaptive controllers." *IEEE Control Systems Magazine* 32, no. 6 (2012): 76-105.

[29] https://www.geeksforgeeks.org/machine-learning/what-is-reinforcement-learning/

[30] Modares, Hamidreza, Frank L. Lewis, and Zhong-Ping Jiang. "H∞ tracking control of completely unknown continuous-time systems via off-policy reinforcement learning." *IEEE Transactions on Neural Networks and Learning Systems* 26, no. 10 (2015): 2550-2562.

[31] Fujimoto, Scott, Herke van Hoof, and David Meger. "Addressing function approximation error in actor-critic methods." *International Conference on Machine Learning (ICML)* (2018): 1587-1596.

[32]https://www.tutorialspoint.com/machine_learning/machine_learning_q_learning.htm

[31] Kiumarsi, Bahare, K1yriakos G. Vamvoudakis, Hamidreza Modares, and Frank L. Lewis. "Optimal and autonomous control using reinforcement learning: A survey." *IEEE Transactions on Neural Networks and Learning Systems* 29, no. 6 (2018): 2042-2062.

[32] Sutton, Richard S., and Andrew G. Barto. "Reinforcement learning in artificial intelligence." *IEEE Transactions on Neural Networks* 9, no. 5 (1998): 1054-1054.

[33] Watkins, Christopher J. C. H., and Peter Dayan. "Q-learning." *Machine Learning* 8, no. 3-4 (1992): 279-292.

[34] https://www.geeksforgeeks.org/deep-learning/deep-q-learning/

[35] Alejandro Puente-Castro, Daniel Rivero, Eurico Pedrosa, Artur Pereira, Nuno Lau, Enrique Fernandez-Blanco, "Q-Learning based system for Path Planning with Unmanned Aerial Vehicles swarms in obstacle environments." *Expert Systems with Applications*, Volume 235, 2024

[36] B. Jang, M. Kim, G. Harerimana and J. W. Kim, "Q-Learning Algorithms: A Comprehensive Classification and Applications," in *IEEE Access*, vol. 7, pp. 133653-133667, 2019

[37]https://www.tutorialspoint.com/machine_learning/machine_learning_actor_critic_algorithm.htm

[38] Haarnoja, Tuomas, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor." *International Conference on Machine Learning (ICML)* (2018): 1861-1870.

[39] Ebrahim Hamid Sumiea, Said Jadid Abdulkadir, Hitham Seddig Alhussian, Safwan Mahmood Al-Selwi, Alawi Alqushaibi, Mohammed Gamal Ragab, Suliman Mohamed Fati, “Deep deterministic policy gradient algorithm: A systematic review.” *Heliyon*, Volume 10, Issue 9, 2024

[40] Silver, David, "Deterministic policy gradient algorithms." *International Conference on Machine Learning (ICML)* 2014: 387-395.

[41] Husam A. Neamah, Oscar Agustin Mayorga Mayorga, “Optimized TD3 algorithm for robust autonomous navigation in crowded and dynamic human-interaction environments.” *Results in Engineering*, Volume 24, 2024

[42] S. Mosharafian, S. Afzali, Y. Bao and J. M. Velni, "A Deep Reinforcement Learning-based Sliding Mode Control Design for Partially-known Nonlinear Systems," *2022 European Control Conference (ECC)*, London, United Kingdom, 2022, pp. 2241-2246

[43] Nicola, M.; Nicola, C.-I.; Selișteanu, D. “Improvement of PMSM Sensorless Control Based on Synergetic and Sliding Mode Controllers Using a Reinforcement Learning Deep Deterministic Policy Gradient Agent.” *Energies* 2022

[44] Xu Z, Huang W, Li Z, Hu L, Lu P. “Nonlinear Nonsingular Fast Terminal Sliding Mode Control Using Deep Deterministic Policy Gradient.” *Applied Sciences*. 2021

[45] Tengda Wang, Huanqing Wang, Ning Xu, Liang Zhang, Khalid H. Alharbi, “Sliding-mode surface-based decentralized event-triggered control of partially unknown interconnected nonlinear systems via reinforcement learning”, *Information Sciences*, Volume 641, 2023

[46] Hu W, Yang Y, Liu Z. “Deep Deterministic Policy Gradient (DDPG) Agent-Based Sliding Mode Control for Quadrotor Attitudes.” *Drones*. 2024

[47] Qiu, Y., Zhou, S., Xia, D., Gu, W., Sun, K., Han, G., Zhang, K., Lv, H.: “Local integrated energy system operational optimization considering multi-type uncertainties: A reinforcement learning approach based on improved TD3 algorithm.” *IET Renewable Power Generation* 17, 2236–2256 (2023).

[48] https://www.geeksforgeeks.org/deep-learning/soft-actor-critic-reinforcement-learning-algorithm/

thesis

ORIGINALITY REPORT

1%	1%	%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	folk.ntnu.no Internet Source	<1%
2	www.theses.fr Internet Source	<1%
3	Submitted to University of Bristol Student Paper	<1%
4	Submitted to Imperial College of Science, Technology and Medicine Student Paper	<1%

Exclude quotes	Off	Exclude matches	< 14 words
Exclude bibliography	On		