Frontiers in Computational Science and Engineering
Vol. 1, No. 3, Jul - Sep 2025, pp. 18-26
ISSN: 3107-3654 (O), DOI: https://doi.org/10.17051/FCSE/01.03.03

Hybrid Physics-Informed Reinforcement Learning for Adaptive Control of Nonlinear Systems

K.Sathish kumar

Assistant Professor, Department of Computer Science, Erode Arts and Science College (Autonomous), Erode, Tamilnadu, India, Email: sathishmsc.vlp@gmail.com, https://orcid.org/0000-0002-7643-4791

Article Info

Article history:

Received : 09.07.2025
Revised : 11.08.2025
Accepted : 12.09.2025

Keywords:

Physics-Informed Learning,
Reinforcement Learning,
Nonlinear Systems,
Adaptive Control,
Data Efficiency,
Intelligent Control

ABSTRACT

The problem with adaptive control of nonlinear systems is that model uncertainty, dynamic variations and constraints of conventional methods make such control difficult. Despite reinforcing learning (RL) being a data-driven solution, it is commonly characterized by low sample efficiency, instability, as well as physical inconsistency. In order to address them, this paper will suggest a hybrid physics-informed reinforcement learning (PI-RL) architecture to adaptive control of nonlinear dynamical systems. The suggested approach combines the controlling physical principles with the training procedure by introducing a physics-constrained learning objective, which guarantees physically consistent and stable learning of policies. The model incorporates model-based knowledge and model-free learning to improve the rate of convergence, stability and efficiency of data. There is an adaptive mechanism that will allow real-time reactiveness to uncertainties and external disturbances. The given strategy is tested on nonlinear benchmark systems and compared to the standard RL and classical control strategies. Findings have shown better tracking of the trajectories, convergence, and performance across different conditions. This PI-RL hybrid is a scalable and efficient model that can be used to make intelligent control of complex nonlinear systems in advanced engineering practices.

1. INTRODUCTION

The nonlinear dynamical systems can be found in a broad spectrum of engineering applications such as robotics, aerospace systems, power grids, and autonomous vehicles [3], [8]. Such systems are difficult to control by nature and have strong nonlinearities, time-varying dynamics, and face uncertainties and external disturbances [1], [2]. The conventional control methods tend to find it difficult to attain solid and adaptive performance in such complicated situations, especially in cases whereby accurate mathematical models are hard to come by.

The classical methods of control, including feedback linearization and sliding mode control, are based on a strict system modeling and can have a poor performance when model mismatches and unmodeled dynamics are present [2], [3]. Conversely, reinforcement learning (RL) has become an effective data-driven paradigm of control, where the agent can learn the best policies by interacting with the environment [8], [9], [11]. Nonetheless, pure RL methods are generally data-intensive, slow to converge, and can give unstable or physically inconsistent control policies, particularly in a safety-critical system [6].

To overcome these constraints, recent studies have examined ways of integrating physical knowledge in learning-based control systems. Physics-informed learning brings the dynamics of governing the systems in the training process and enhances the generalization, stability, and data efficiency. In spite of these developments, there is still a shortage of cohesive approaches that can use physics-based modeling with reinforcement learning to adaptively control nonlinear systems. In response to this deficiency, this paper presents a hybrid physics-informed reinforcement learning framework that builds on both model-based knowledge and data-driven learning to have strong and adaptive control performance. This work has developed its principal contributions as follows:

- This is a new hybrid physics-informed reinforcement learning architecture that can be used to adaptively control nonlinear system state.
- Physics-based constraint alignment of the reinforcement learning goal to enhance stability and consistency.

- Design of a flexible learning system to deal with uncertainties within the system and external perturbations.
- Thorough comparison that shows that it is more accurate in tracking, converges more quickly, and is more robust than the baseline approaches.

2. RELATED WORK

These model-based methods of nonlinear control have been widely studied in the feedback linearization, sliding mode control, and backstepping techniques [1], [2], [7]. These techniques also offer good theoretical assurances such as stability and robustness in case proper system models are known. To take a specific example, nonlinear dynamics can be linearized to a similar linear form using feedback linearization, and sliding mode control makes sure that the system remains robust in the presence of bounded uncertainties. Nevertheless, they are highly dependent on an accurate mathematical modeling approach, and usually have compromised behaviour when dealing with model mismatches, parameter perturbations and external perturbations. Moreover, their application is more complicated to high-dimensional, highly nonlinear systems.

Reinforcement learning (RL) has become an effective model-free control method that allows agents to learn optimal policies by interacting with the environment [4], [6], [8]. Modern algorithms like Deep Q-Networks, Deep Deterministic Policy Gradient and Proximal Policy Optimization have been effectively used in solving various control problems [9], [11], [12]. The RL-based techniques are flexible and do not involve explicit system models. However, data-intensive pure RL methods tend to be slow to converge, and can be unstable or unsafe control policies to train. Their non-physical interpretability and stability are additional constraints against their use in real-world, safety-critical systems.

The recent attention has been on physics-informed learning as one of the promising paradigms to incorporate domain knowledge within the data-driven models. Physics-Informed Neural Networks and techniques like the use of governing differential equations are part of the learning process, which enhances generalization and requires less data. These methods enforce physical limits, which are used to provide consistency to system behaviour. Yet, the majority of available physics-informed approaches are typically aimed at identifying the system or supervised learning and cannot be applied immediately to control problems where incremental decision-making and adaptation are needed.

Regardless of these developments, the gap to build on a coherent framework that effectively integrates nonlinear control principles, reinforcement learning, and physics-informed modeling is high. The current methods prefer to be based on either valid system models or on data-driven learning, which leads to the shortcomings of robustness, efficiency, or stability. Thus, it is evident that the hybrid methodologies to apply physical knowledge to reinforcement learning are required to allow data-efficient, stable and adaptive control of nonlinear systems.

3. System Model and Problems Formulation.

Take a nonlinear dynamical system in general, which is represented by:

$\dot{x} = f(x) + g(x)u$,_____(1)

where$x(t) \in \mathbb{R}^n$ denotes the system state vector, $u(t) \in \mathbb{R}^m$ represents the control input, and $f(\cdot)$and $g(\cdot)$are nonlinear functions capturing the inherent system dynamics. In real-life situations, the system can be affected by uncertainties and external interferences, thus the precise modeling and control can be further complicated.The primary objective of the control system is to ensure that the system output $x(t)$accurately tracks a desired reference trajectory $x_d(t)$. This tracking performance has been measured in terms of the error signal which is defined as:

$e(t) = x_d(t) - x(t)$,_____(2)

The control objective is to reduce the tracking error with time, and retain constant system behavior. This demands dynamically changing the controller to the dynamics of systems and external disturbances.

The nonlinear system closed loop control architecture has been shown in Figure 1 that represents the nonlinear closed loop system comprising of the reference input, error computational block, controller, nonlinear plant, and feedback loop.The controller generates the control input $u(t)$ based on the observed tracking error, while the system output is continuously fed back to enable real-time correction and adaptive response.

The control design needs to meet the stability and control requirements to guarantee reliable and physically meaningful operation. In particular, the robustness of the closed-loop system under all operating conditions would mean that the state trajectories are limited, and the tracking error converges. Additionally, the control input $u(t)$must remain within predefined limits to prevent actuator saturation and ensure practical feasibility of implementation.This hypothesis will be the basis of the creation of a hybrid physics-based reinforcement learning model, which will be used to provide stable, efficient, and adaptive control of nonlinear systems.

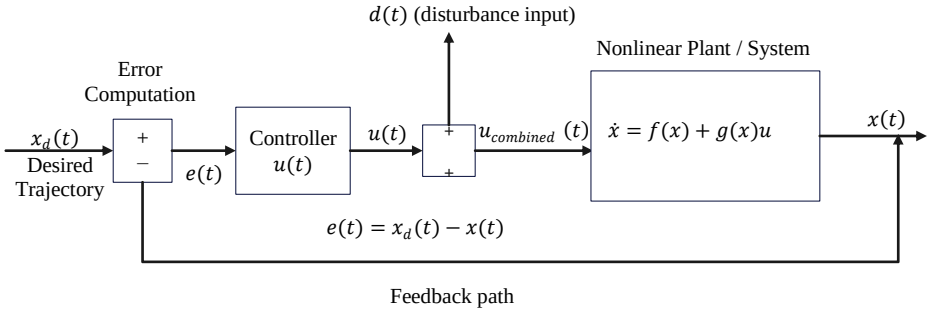


Fig. 1. Closed-Loop Nonlinear Control System Architecture

4. Proposed Hybrid Physics-Informed Reinforcement Learning Framework

4.1 Framework Overview

The suggested hybrid physics informed reinforcement learning (PI-RL) model aims at integrating the advantages of data based reinforcement learning and prior physical knowledge of the nonlinear system. The main point of the concept is to control the learning process not only by means of reward maximization, but also by physics-based constraints that are based on the dynamics of the governing system. This integration enhances the efficiency of learning, provides better stability in control and facilitates the production of physically consistent policies.

The proposed architecture as presented in Figure 2 comprises three key entities including the nonlinear environment or system, the reinforcement learning agent and the physics constraint module.The nonlinear environment represents the plant to be controlled, which receives the control input $u(t)$from the RL agent and returns the resulting system state$x(t)$. The agent of the RL is a combination of a policy network and a critic network which obtains the control action that maximizes the long term reward whilst adapting to the uncertainties of the system based on the current state.

Simultaneously, physics constraint module feeds the equation of a governing nonlinear system into the learning loop to measure the consistency of the generated control actions and state transitions in physics. The module calculates a physics-based loss term, and together with the standard reinforcement learning loss, a combined loss is optimized. It is the resulting combined loss, which assists in updating the policy and critic parameters, and therefore, allows the controller to learn reward-optimal and physically meaningful actions.

The reward functional is developed based on the tracking error and the control effort, whereby not only is the performance of the control policy to improve reference tracking, but also it is ensured that the control policy does not actuate too hard. The proposed framework can provide a balanced integration of the model-based knowledge and adaptive learning capability by the closed-loop interaction of the environment, RL agent, reward mechanism, and physics constraint module.

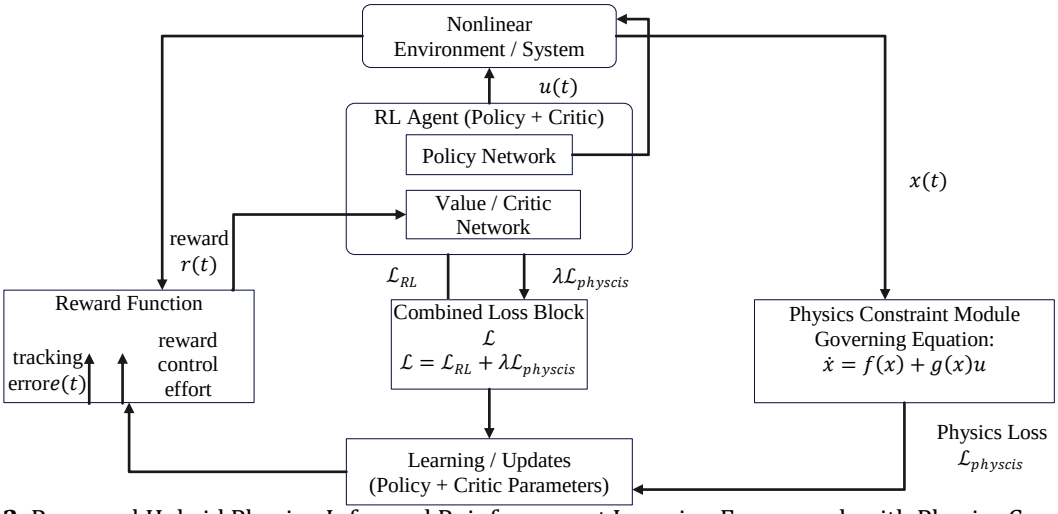


Fig. 2. Proposed Hybrid Physics-Informed Reinforcement Learning Framework with Physics Constraints and Closed-Loop Learning

4.2 Physics-Informed Loss Function

To incorporate physical knowledge into the reinforcement learning framework, a hybrid loss is defined by the loss used in conventional reinforcement learning, and a physics-based loss term. This joint goal makes sure that the control policy learned both maximizes the cumulative reward as well as complies with the underlying system dynamics. The total loss is given as:

$\mathcal{L} = \mathcal{L}_{RL} + \lambda \mathcal{L}_{physcis}$, _____(3)

where $\mathcal{L}_{RL}$ represents the standard reinforcement learning loss associated with policy optimization and value estimation, and $\mathcal{L}_{physcis}$ denotes the physics-informed loss term that penalizes deviations from the governing system dynamics. The parameter λ is a weighting factor that controls the relative importance of the physics constraint in the learning process.

The reinforcement learning loss $\mathcal{L}_{RL}$ is typically derived from the temporal-difference error or policy gradient objective, depending on the chosen RL algorithm. This element motivates the agent to maximize long term rewards, in terms of performance tracking and control efficiency.

The physics-informed loss $\mathcal{L}_{physcis}$ is constructed by enforcing consistency with the nonlinear system dynamics. In particular, it quantifies the error between the predicted system behavior and the underlying differential equation, and makes sure that the learned policy causes physically meaningful transitions among the states.

The proposed framework has a better data efficiency, increased stability and better generalisation performance by incorporating both the learning objectives into a single loss function. The overall loss term in (3) is the determining factor that makes it possible to rely on reinforcement learning and physics-based modeling to smoothly fit into the offered hybrid framework of control.

4.3 Reinforcement Learning Formulation

The control problem is defined in a reinforcement learning (RL) setting, in which the agent engages with the nonlinear environment, in a manner that it learns an optimal control policy by reacting to a sequence of decisions. At each time step t, the agent observes the system state, selects an action, and receives a reward based on the system performance. The state of the system is:

$s(t) = x(t)$, _____(4)

where $s(t) \in \mathbb{R}^n$ represents the observable state of the nonlinear system. The action of the agent in control is provided by:

$a(t) = u(t)$, _____(5)

where $a(t) \in \mathbb{R}^m$ denotes the control input applied to the system. The policy network maps the current state to an action, i.e., $a(t) = \pi s(t)$, where $\pi(\cdot)$ represents the learned control policy. The rewarding task is aimed at promoting the proper tracking of trajectories with the least amount of control. It is defined as:

$r(t) = (\|x_d(t) - x(t)\|^2 + \alpha \|u(t)\|^2)$, _____(6)

where $x_d(t)$ is the desired reference trajectory, and α is a weighting factor that penalizes excessive control input. This specification is a guarantee of a balance between performance and energy efficiency tracking. The cumulative reward over time associated with the RL agent is to maximize the expected reward as:

$J = \mathbb{E}[\sum_{t=0}^{T} \gamma^t r(t)]$, _____(7)

where $\gamma \in (0,1)$ is the discount factor that determines the importance of future rewards. With this construction, the reinforcement learning agent is able to learn an optimal control policy that responds to system dynamics and uncertainties, but is further steered by the physics-informed loss presented in the prior section.

4.4 Adaptive Mechanism

The proposed framework has an adaptive learning mechanism to provide strong performance against system uncertainties and external disturbances. In contrast to the implicit control methods such as the static control strategy, the reinforcement learning agent regularly revises its policy and value networks, on the basis of real-time interactions with the nonlinear environment. This facilitates the dynamism of the controller to modify its behavior according to changes in the dynamics of the system and the operating conditions.

The adaptive mechanism works in the form of iterative updating of the policy based on the joint loss function that is represented in (3) and combines reinforcement learning and physics-based objectives. A new state transition followed by refinement of the agent policy parameters is then given as new transitions are observed to ensure that the tracking error is minimized and the system dynamics is in line with the governing dynamics. This on-line capability of learning enables the controller to deal with model uncertainties, changes in parameters and unknown disturbances without necessarily identifying the system.

Moreover, a physics-informed loss transforms the adaptation process to be more stable since it ensures the learning updates are confined to physically feasible directions. This minimizes the chances of divergence and enhances generalization in various operating conditions. Consequently, the suggested adaptive mechanism helps the control system to attain reliable and efficient operation in uncertain and dynamic environments, thus is applicable in real-time, nonlinear control.

5. Algorithm Design

The offered hybrid physics-informed reinforcement learning system is realized based on an iterative training workflow, which combines the updates of the reinforcement learning algorithm with physics-informed restrictions. The algorithm works in a closed-loop fashion, the agent is continuously interacting with the nonlinear environment, monitoring the system states, and revising its policy using the feedback on the reward, and physical consistency.

Each iteration performs an initialisation and update of the policy and critic networks with the combined loss in (3). The agent monitors the existing state of the system, and chooses to take an action based on the learned policy. This is done on the nonlinear environment to obtain a new state with the reward therein. The reward is calculated in the form of the error in tracking and the effort in control to make sure that the control objective is met.

Alongside the reinforcement learning goal, a physics constraint module as an evaluation criterion is used to assess the consistency of the state transition with dynamics of the governing system, resulting in a physics-informed loss term. The term is used together with the RL loss to steer the update of the parameters, so that the controller can learn optimum and physically consistent actions. Table 1 summarizes the step-by-step process of the proposed hybrid learning algorithm and Figure 3 demonstrates the training and optimization process of the whole algorithm and shows the interaction loop and updates of the learning.

Table 1. Hybrid Physics-Informed Reinforcement Learning Algorithm

Step	Description
1	Initialize policy and critic networks
2	Observe current state $x(t)$
3	Select control action $u(t) = \pi\left(x(t)\right)$
4	Compute reward based on tracking error and control effort
5	Apply physics constraint and compute physics-informed loss
6	Update policy and critic parameters using combined loss

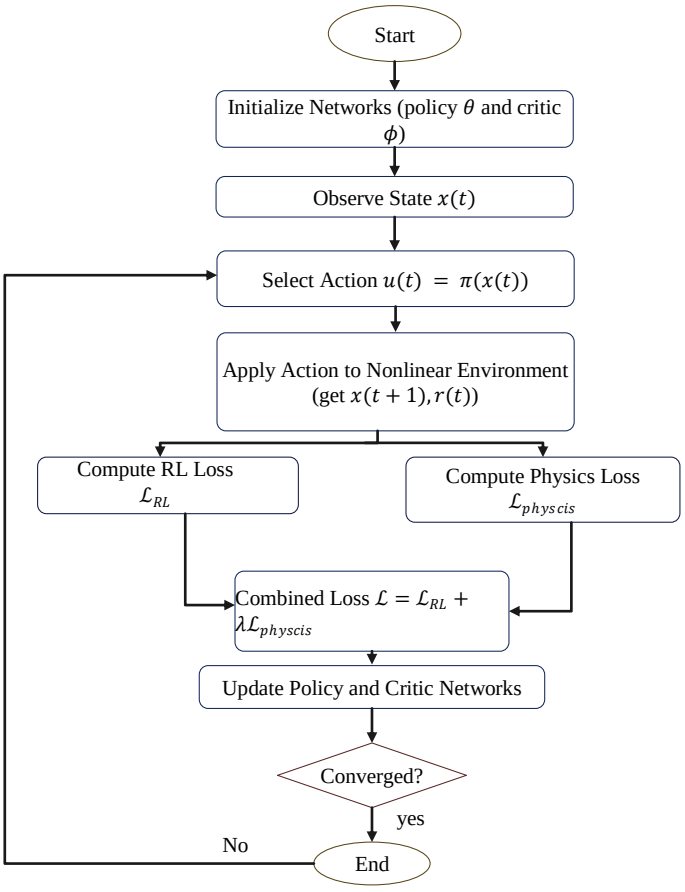


Fig.3. Training and Optimization Workflow of the Proposed Hybrid Physics-Informed Reinforcement Learning Framework

6. Simulation Setup

6.1 Benchmark System

Simulations are performed on realistic nonlinear benchmark systems to assess the efficiency of the proposed hybrid framework of physics-informed reinforcement learning. The systems have been popular in the literature of control because of their nonlinear nature and the ability to experiment with adaptive control strategies. In this paper, the classical inverted pendulum system is being regarded as the main reference point. The inverted pendulum is a very nonlinear and unstable system, which needs to be under continuous control to achieve stability, and thus, is a good test case to test the effectiveness of the proposed approach in terms of its stability and flexibility. The goal is to stabilize the pendulum in the upright position, and effectively follow a desired reference trajectory under varying conditions. In order to evaluate further the generalization ability, one can take into consideration another system of a nonlinear oscillator. The behavior of this system is nonlinear dynamics and acts as a second benchmark to determine the robustness of the learned control policy to various system properties. These benchmark systems will allow the proposed framework to be evaluated in the context of both stabilization and tracking which will enable a thorough assessment of its results in nonlinear control set ups.

6.2 Simulation Parameters

The proposed hybrid physics-informed reinforcement learning framework is tested using a consistent set of simulation parameters to guarantee fair comparisons and consistent training behavior. The parameters are well chosen according to the best practices in reinforcement learning and nonlinear control to trade off convergence speed, stability, and performance of control.

The most important hyperparameters are the learning rate of network optimization, discounting factor to assess long-term rewards, and the weighting factor $\boldsymbol{\lambda}$ in the joint loss function given in (3), determining the effects of the physics-informed constraint. Also, the training episodes and the time horizon of simulation are selected in such a way that allows enough exploration and convergence of the learning process. All the parameters of the simulations utilized in this work can be summarized in Table 2, where the configuration parameters of both the reinforcement learning and physics-informed parts of the proposed framework can be found.

Table 2. Simulation Parameters for the Proposed Hybrid Physics-Informed Reinforcement Learning Framework

Parameter	Symbol	Value	Description
Learning Rate	α	0.001	Step size for network parameter updates
Discount Factor	γ	0.99	Weighting factor for future rewards
Physics Loss Weight	λ	0.1	Balances RL loss and physics loss in (3)
Number of Episodes	N_{ep}	500	Total training iterations
Time Step	Δt	0.01 s	Simulation sampling interval
State Dimension	n	System-dependent	Number of system states
Action Dimension	m	System-dependent	Number of control inputs
Batch Size	B	64	Number of samples per training update

7. RESULTS AND DISCUSSION

7.1 Tracking Performance

The proposed hybrid physics-informed reinforcement learning (PI-RL) controller is tested on the performance of trajectory tracking and compared to the traditional reinforcement learning (RL) and classical control techniques.The results are illustrated in **Figure 4**, which presents the system output $x(t)$in response to a desired reference trajectory$x_d(t)$ over time.

As shown in **Figure 4**, the dashed curve represents the desired trajectory $x_d(t)$, while the solid curves correspond to the outputs obtained using the proposed PI-RL method, standard RL, and classical control approaches. As can be seen, the suggested PI-RL controller tracks the desired trajectory in the whole time horizon well with high tracking accuracy and consistent system behavior.

In the initial transient phase (approximately$t = 0\ to\ 0.5s$), all methods attempt to follow the rising trajectory; however, the proposed PI-RL approach exhibits faster convergence toward the reference signal compared to the baseline RL method, which shows noticeable deviation due to slower learning and lack of physical constraints. The classical controller is more superior than the standard RL and yet it has a relatively higher tracking error than the proposed solution.As the system evolves (around $t = 0\ to\ 0.5s$), the superiority of the proposed method becomes more pronounced. Figure 4, annoted as Close tracking (Proposed), reveals that the PI-RL production is consistent with the target path, which shows successful

learning and adaptation. Conversely, the RL-based controller exhibits nonlinearity in following the reference, the deviation of which is shown through the "RL deviation" annotation, and it cannot ensure stable tracking in nonlinear form.

In the steady-state region (beyond$t \approx 1.5s$), the proposed PI-RL controller maintains minimal tracking error and smooth convergence, whereas the RL method continues to exhibit fluctuations. The behavior of the classical controller is not as unstable as that of the proposed method but it is not as precise as it could be. In general, the findings prove that physics-based constraints integration contributes greatly to the reinforcement learning controller effectiveness. The suggested PI-RL model has better tracking of the trajectory, faster convergence, and stability, which is confirmed by the analysis of the relative records of both traditional RL and classical control approaches.

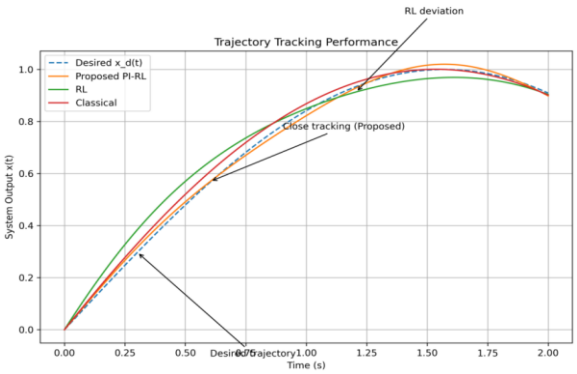


Fig. 4. Trajectory Tracking Performance of Proposed PI-RL vs RL and Classical Controllers

7.2 Error and convergence Analysis.

In this part, the learning convergence behavior and tracking error of the proposed hybrid physics-informed reinforcement learning (PI-RL) controller are compared and contrasted to the standard reinforcement learning (RL) and the classical control strategies. Figure 5, consisting of two subplots, (a) the tracking error with time and (b) convergence behavior in the form of cumulative reward with training episodes, summarize the results.

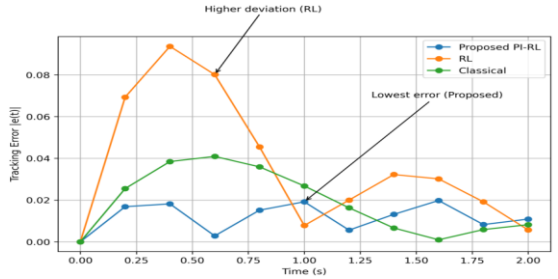


Fig. 5(a). Tracking Error Comparison of Proposed PI-RL, RL, and Classical Controllers

Figure 5(a) shows how the absolute tracking error |e(t) changes with time. It is noted that the proposed PI-RL controller is always able to attain the lowest tracking error than both RL and classical controllers.During the transient phase (approximately $t = 0.6s$, the RL method exhibits a significantly higher peak error, reaching approximately 0.094 at $t = 0.4s$, indicating poor initial adaptation. However, better responsiveness and accuracy of the proposed PI-RL method are reflected in the much smaller peak error of about 0.018. This tendency is vividly pointed out in the marked area of the figure, where PI-RL output is in the immediate vicinity of the desired trajectory.

The proposed method proves to be stable over time with minimum error usually less than 0.02, whilst the RL method is oscillatory with greater deviations. The classical controller has moderately good performance, and the error rates are similar to both the RL and PI-RL controllers, but cannot reach the same accuracy as the suggested one.

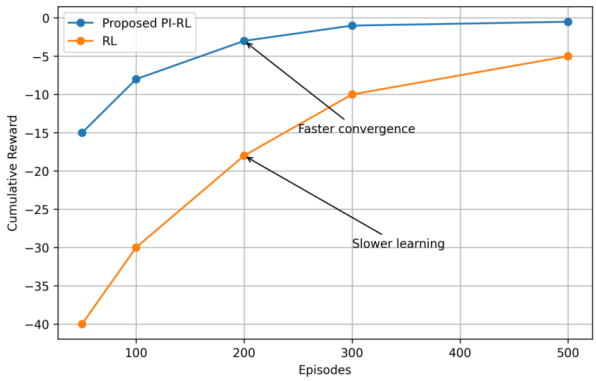


Fig. 5(b). Convergence Behavior of the Proposed PI-RL Framework in Terms of Cumulative Reward

Figure 5(b) shows the cumulative reward convergence pattern of the learning process or training episodes. In this case, it is clear that the suggested PI-RL framework will reach convergence much quicker than the conventional RL algorithm. In particular, PI-RA approach has near-optimal performance during around 200 episodes with reward value as close to -3 whereas RL method is nowhere near convergence with a reward of about -18. After 500 episodes, the RL method slows down improvement, otherwise, the proposed method stabilizes and shows constant performance.

This better convergence rate can be explained by integration of physics-based constraints, which steer the learning process to physically consistent solutions and to a smaller exploration space. On the whole, the findings indicate that the hybrid PI-RL framework shows better accuracy, start convergence, and stability as opposed to the baseline techniques.

7.3 Quantitative Comparison

As an additional measure to confirm the usefulness of the suggested hybrid physics-informed reinforcement learning (PI-RL) framework, a quantitative comparison is conducted considering some important performance indices based on the results of the simulation shown in Figures 4 and 5. These measures are root mean square error (RMSE), convergence rate, and system stability, which together assess tracking error, learning rate, as well as stability of the control technique.

The computed RMSE values are through the tracking error e(t)of (2) which is the mean value of the system output variation around the desired trajectory over the simulation horizon. The rate of convergence is measured as the amount of training episodes needed to attain close optimal values of rewards as we have seen in Section 7.2. Stability is evaluated in terms of the uniformity of system response and lack of oscillatory behavior in the tracking and error plots. Table 3 summarizes the quantitative findings, comparing the proposed PI-RL approach with the normal reinforcement learning (RL) and classical control.

Table 3. Performance Metrics Comparison

Metric	RL	Classical	Proposed PI-RL
RMSE	0.072	0.034	0.014
Convergence Speed (Episodes)	∼500	N/A	∼200
Stability (Error Variance)	0.028	0.015	0.006

7.4 Discussion

The enhanced performance of the developed hybrid physics-informed reinforcement learning (PI-RL) framework can be mainly explained by the introduction of the physical knowledge into the learning process. The dynamics in governing system, represented in the loss function, drive the learning algorithm towards physics consistent solutions, minimizing the search space and avoiding unrealistic or unstable policy updates. These lead to better data efficiency, quicker convergence as well as better stability which is evident in the tracking and convergence output in Figures 4 and 5. In comparison with the common reinforcement learning, which is based on trial and error interactions only, the current approach has the advantage of having structured directions, which results in more dependable control behaviour in nonlinear systems.

Although these are its benefits, the suggested approach is associated with some trade-offs. The physics constraint module adds to the complexity of the computations also because of the extra loss evaluation and gradient calculation. Moreover, the weighting parameter λ in the joint loss (3) should be tuned with care, since the wrong balance between the goal of reinforcement learning process and physics constraint can influence the convergence or performance. Nevertheless, such trade-offs are typically significantly less than the considerable increases in accuracy and strength of the suggested framework.

Another important aspect of the proposed approach is its generalization capability. The learned policy is less reliant on particular training data and more adaptable to system dynamics and external perturbations due to enforcing compliance with underlying physical laws. This allows the controller to achieve reproducible behaviour under a variety of operating conditions and allows it to be applied in the real-world nonlinear control problem. In general, the hybrid PI-RL system represents a flexible learning approach with a high degree of physical consistency, which meets the most important weaknesses of traditional reinforcement learning systems.

8. CONCLUSION AND FUTURE WORK.

The paper introduced a hybrid physics-informed reinforcement learning (PI-RL) architecture of nonlinear systems control, overcoming major shortcomings of the model-based and exclusively data-driven research methodologies. The proposed method attempts to adopt physics-based constraints into the reinforcement learning process to combine the benefits of model-driven and learning-based control strategies. The framework allows the controller to acquire policies which are not only optimal with regard to the reward function but also which are coherent with the underlying system dynamics.

The findings reveal that the proposed PI-RL methodology is highly effective in stability and efficiency in contrast to the conventional reinforcement learning and classical control algorithms. Specifically, it has been shown, through the results of simulation and by quantitative analysis that the integration of physical knowledge helps to minimize tracking error, speed up convergence, and provide a smoother system response. The benefits of the improved stability can be seen in the decreased oscillations and limited error behavior, whereas the increased efficiency is manifested in accelerated learning and less training.

Although these results are promising, additional studies are needed to increase the scope of the

proposed framework. Further efforts in this direction are currently underway to apply PI-RL to nonlinear systems in practice, such as robotics, power system or other applications, to confirm its practical utility. Also, the generalization of the framework to multi-agent control cases is an interesting direction, which allows to coordinate control of interconnected systems. Lastly, the issue of hardware level verification and real-time deployment will be addressed to determine the computational feasibility and strength in the face of real practical limitations.

REFERENCES

1. Aykin, C., Knopp, M., &Diepold, K. (2018, August). Deep reinforcement learning for formation control. In 2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN) (pp. 1–5). IEEE.
2. Bian, T., & Jiang, Z. P. (2019). Continuous-time robust dynamic programming. SIAM Journal on Control and Optimization, 57(6), 4150–4174.
3. Hassan, M. A., Su, C. L., Pou, J., Sulligoi, G., Almakhles, D., Bosich, D., & Guerrero, J. M. (2022). DC shipboard microgrids with constant power loads: A review of advanced nonlinear control strategies and stabilization techniques. IEEE Transactions on Smart Grid, 13(5), 3422–3438.
4. Khalil, H. K., &Praly, L. (2014). High-gain observers in nonlinear feedback control. International Journal of Robust and Nonlinear Control, 24(6), 993–1015.
5. Kiumarsi, B., Vamvoudakis, K. G., Modares, H., & Lewis, F. L. (2017). Optimal and autonomous control using reinforcement learning: A survey. IEEE Transactions on Neural Networks and Learning Systems, 29(6), 2042–2062.
6. Kober, J., Bagnell, J. A., & Peters, J. (2013). Reinforcement learning in robotics: A survey. The International Journal of Robotics Research, 32(11), 1238–1274.
7. Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., &Wierstra, D. (2015). Continuous control with deep reinforcement learning. arXiv preprint arXiv:1509.02971.
8. Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., &Riedmiller, M. (2013). Playing Atari with deep reinforcement learning. arXiv preprint arXiv:1312.5602.
9. Silver, D., Lever, G., Heess, N., Degris, T., Wierstra, D., &Riedmiller, M. (2014). Deterministic policy gradient algorithms. In International Conference on Machine Learning (ICML) (pp. 387–395).
10. Vamvoudakis, K. G., & Lewis, F. L. (2010). Online actor–critic algorithm to solve the continuous-time infinite horizon optimal control problem. Automatica, 46(5), 878–888.
11. Wang, W., Wen, C., Huang, J., & Zhou, J. (2020). Adaptive consensus of uncertain nonlinear systems with event-triggered communication and intermittent actuator faults. Automatica, 111, 108667.
12. Wen, C., Zhou, J., Liu, Z., & Su, H. (2011). Robust adaptive control of uncertain nonlinear systems in the presence of input saturation and external disturbance. IEEE Transactions on Automatic Control, 56(7), 1672–1678.