www.ijarp.com ISSN 2456-9992 Page: 01-12

International Journal Advanced Research Publications

Research Article

Volume: 01
Issue: 03

HIERARCHICAL PROXIMAL POLICY OPTIMIZATION WITH MULTI-LAYER ADAPTATION FOR SEMI-ACTIVE SUSPENSION CONTROL: A NOVEL REINFORCEMENT LEARNING APPROACH FOR ENHANCED VEHICLE DYNAMICS

[*1]Satyendra Kumar Shukla, [2]Altaf Ali, [3]Mohd Umar Abdullah

[1]Assistant Professor, Department of Mechanical Engineering, FoET, Khwaja Moinuddin Chishti Language University, Lucknow.

[2,3] Student of Mechanical Engineering, FoET, Khwaja Moinuddin Chishti Language University, Lucknow.

Article Received: 31 October 2025, Article Revised: 20 November 2025, Published on: 11 December 2025

***Corresponding Author: Satyendra Kumar Shukla**

Assistant Professor, Department of Mechanical Engineering, FoET, Khwaja Moinuddin Chishti Language University, Lucknow. DOI: https://doi-doi.org/101555/ijrpa.3261

ABSTRACT

This research introduces a novel Hierarchical Proximal Policy Optimization with Multi-Layer Adaptation (HPPO-MLA) algorithm designed for semi-active suspension systems in modern vehicles. The primary objective was to develop a control framework that simultaneously optimizes ride comfort, vehicle stability, and safety constraints under varying road conditions. The methodology employed a sophisticated four-layer hierarchical architecture comprising perception, adaptation, optimization, and safety layers, integrated with proximal policy optimization techniques. The experimental setup involved comprehensive simulations using a high-fidelity quarter-car model validated against real-world data, with performance evaluation conducted across ISO 8608 road classifications A-F. Results demonstrated that HPPO-MLA achieved a 35.2% improvement in ride comfort compared to traditional PID control, reduced RMS body acceleration by 28.7% versus skyhook controllers, and maintained suspension travel within safety limits with 94.3% reliability. The algorithm exhibited superior adaptation capabilities, converging within 50 training episodes compared to 200+ episodes required by conventional reinforcement learning approaches. The study concluded that HPPO-MLA represents a significant advancement in intelligent suspension

control technology, offering robust performance across diverse operating conditions while maintaining computational efficiency suitable for automotive embedded systems.

KEYWORDS: Hierarchical reinforcement learning, semi-active suspension, proximal policy optimization, vehicle dynamics, ride comfort optimization, adaptive control

INTRODUCTION

The automotive industry faces persistent challenges in balancing conflicting objectives within suspension system design, particularly the simultaneous optimization of passenger comfort, vehicle stability, and road holding capability. Semi-active suspension systems, characterized by their electronically controlled variable damping capabilities, have emerged as the preferred solution for modern vehicles due to their favorable compromise between performance enhancement and energy efficiency. These systems offer the potential to adapt to varying road conditions without the substantial energy consumption penalties associated with fully active suspension configurations. However, conventional control methodologies, including Proportional-Integral-Derivative (PID) controllers, skyhook strategies, and their hybrid derivatives, demonstrate fundamental limitations in addressing the complex, multi-objective nature of suspension control under dynamically changing operating conditions.

Recent advancements in reinforcement learning have introduced transformative possibilities for adaptive control systems, with algorithms demonstrating remarkable capabilities in learning optimal control policies through environmental interaction without requiring explicit system modeling. Despite these advancements, contemporary reinforcement learning approaches applied to suspension control manifest several critical shortcomings that limit their practical implementation and effectiveness. Traditional deep deterministic policy gradient (DDPG) and soft actor-critic (SAC) methods exhibit prohibitive sample inefficiency, necessitating extensive training periods that undermine their viability for industrial applications where development time and computational resources represent significant constraints. Furthermore, most reinforcement learning-based suspension controllers conspicuously lack explicit safety constraint mechanisms, potentially leading to suspension travel violations, component fatigue, and premature system degradation during operation. The absence of formal safety guarantees presents a fundamental barrier to the adoption of these approaches in safety-critical automotive applications.

The most significant research gap identified in current literature pertains to the absence of hierarchical reinforcement learning frameworks capable of simultaneously addressing multi-objective optimization, real-time parameter adaptation, and rigorous safety constraint enforcement within semi-active suspension systems. Existing approaches typically employ monolithic network architectures that struggle to effectively manage the hierarchical nature of suspension control objectives, which span different temporal scales and physical domains. This research addresses these critical limitations through the development of a novel Hierarchical Proximal Policy Optimization with Multi-Layer Adaptation (HPPO-MLA) algorithm specifically designed for semi-active suspension control applications. The proposed framework incorporates advanced computational techniques including multi-head attention mechanisms for feature prioritization, recursive least squares estimation for online parameter adaptation, and Pareto front optimization for multi-objective balance, creating a sophisticated control architecture that transcends the capabilities of existing approaches.

MATERIALS AND METHODS

Experimental Framework and Simulation Environment

The experimental investigation employed a high-fidelity quarter-car model as the primary simulation environment, with parameters meticulously selected to represent typical mid-size passenger vehicles while incorporating realistic non-linearities and practical constraints. The simulation framework was developed in MATLAB/Simulink environment (version R2023b) with numerical integration implemented using fourth-order Runge-Kutta methods to ensure accuracy and stability. The vehicle parameters were configured as follows: sprung mass (m_s) varied dynamically within a range of 275-325 kg to simulate changes in passenger and cargo loading during vehicle operation; unsprung mass (m_u) was set at 40 kg with ±5 kg variation to account for differences in brake and wheel assembly configurations; spring stiffness (k_s) implemented progressive characteristics with nominal value of 20,000 N/m increasing by 15% at maximum compression; tire stiffness (k_t) was configured at 200,000 N/m with pressure-dependent variations according to established Pacejka tire models; damping characteristics included asymmetric compression and rebound behavior typical of commercial dampers, with maximum damping coefficient (c_{max}) of 5,000 N·s/m and minimum (c_{min}) of 1,000 N·s/m.

Dataset Generation and Training Protocol

Training datasets were generated using comprehensive road profiles according to ISO 8608 classifications A through F, representing road qualities from very smooth highway surfaces to extremely rough off-road conditions. Additional realistic irregularities including potholes of varying dimensions (0.05-0.15 m depth), speed bumps of standard heights (0.08 m), and random surface defects representing typical road wear patterns were incorporated. The training dataset comprised 50,000 samples across varying vehicle parameters, road conditions, and driving scenarios, with validation and test datasets each containing 10,000 samples to ensure statistically robust evaluation. The training protocol implemented curriculum learning methodology, beginning with smooth road conditions (ISO Class A) and progressively introducing more challenging scenarios to facilitate stable learning convergence.

HPPO-MLA Algorithm Implementation

The Hierarchical Proximal Policy Optimization with Multi-Layer Adaptation algorithm was implemented using a custom neural network architecture developed in PyTorch framework (version 2.0.1). The perception layer employed multi-head attention mechanisms with four attention heads, each processing transformed representations of the state vector through learned linear projections. The adaptation layer implemented modified recursive least squares estimation with forgetting factor $\lambda = 0.98$, enabling continuous parameter estimation for changing vehicle characteristics. The optimization layer utilized Pareto front approximation through adaptive weighting schemes, dynamically adjusting objective weights based on performance gradients with learning rate $\alpha = 0.01$. The safety layer incorporated logarithmic barrier functions with constraint violation penalties increasing exponentially as operational limits were approached.

Training Configuration and Hyper-parameters

The training process employed distributed computation with eight environment instances running in parallel to accelerate data collection. Network architectures were carefully designed with layer sizes determined through systematic ablation studies: perception layer (64 neurons), adaptation layer (32 neurons), optimization layer (48 neurons), and safety layer (16 neurons). Training hyper parameters were configured as follows: learning rate of 3×10^{-4} with cosine annealing schedule reducing to 1×10^{-5} over training; discount factor $\gamma = 0.99$ with hierarchical decomposition applying different factors to each layer; generalized advantage estimation parameter $\lambda = 0.95$; clip parameter $\epsilon = 0.2$ adapting based on policy divergence metrics; batch size of 256 samples employing prioritized experience replay focusing on transitions with high temporal difference error.

Performance Evaluation Metrics

Comprehensive evaluation employed seven meticulously defined metrics: RMS body acceleration (m/s^2) computed over entire episodes as primary comfort indicator; maximum suspension travel (m) measured as worst-case deflection representing safety compliance; RMS tire deflection (m) calculated to assess road holding capability; control effort (N·s) integrated as damping force over time indicating energy consumption; settling time (s) measured after standardized step disturbance inputs; overshoot percentage (%) calculated from transient response to step inputs; safety violation rate (%) computed as percentage of time steps exceeding operational constraints.

Comparative Benchmark Controllers

Performance comparison included eight benchmark controllers representing current state-of-the-art approaches: traditional PID controller with Ziegler-Nichols tuning followed by manual refinement; skyhook controller implementing industry-standard semi-active control strategy; groundhook controller focusing on wheel motion control; mixed skyhook-groundhook hybrid controller; passive suspension with fixed optimal damping establishing baseline performance; DDPG-based reinforcement learning controller with comparable network parameter count; SAC-based controller representing cutting-edge reinforcement learning methodology; constrained PPO controller providing direct comparison to highlight hierarchical decomposition benefits.

RESULTS AND DISCUSSION

Performance Comparison Across Control Methodologies

The comprehensive performance evaluation revealed that HPPO-MLA demonstrated statistically significant superiority across multiple metrics compared to benchmark controllers. The RMS body acceleration metric, representing the primary comfort indicator, showed HPPO-MLA achieving 0.85 m/s^2 with standard deviation of 0.07 across test conditions. This performance represented a 35.2% improvement over traditional PID control at 1.31 m/s^2, a 28.7% improvement over industry-standard skyhook control at 1.12 m/s^2, and a 7.6% improvement over the best learning-based benchmark (SAC) at 0.92 m/s^2. The substantial comfort improvement stemmed from the hierarchical architecture's ability to optimize specifically for comfort while maintaining other objectives through constraint satisfaction rather than weighted trade-offs.

The maximum suspension travel metric, representing the primary safety indicator, showed HPPO-MLA limiting travel to 0.078 meters with standard deviation of 0.005. This performance represented a 30.4% reduction compared to PID control at 0.112 meters, a 17.9% reduction compared to skyhook control at 0.095 meters, and an 8.2% reduction compared to constrained PPO at 0.091 meters. The safety improvement demonstrated the effectiveness of the explicit safety layer in enforcing travel constraints while maintaining performance, addressing a critical limitation of existing approaches that either violate constraints or sacrifice performance to maintain them. The RMS tire deflection metric, representing road holding capability, showed HPPO-MLA achieving 0.0121 meters with standard deviation of 0.001, representing a 25.8% improvement over PID control at 0.0163 meters, a 14.8% improvement over skyhook control at 0.0142 meters, and a 5.5% improvement over SAC at 0.0128 meters.

Table 1: presents comprehensive performance comparison. HPPO-MLA demonstrates superior performance in five of seven metrics.

Metric	**HPPO -MLA**	**PID**	**Skyhook**	**Groundhook**	**Passive**	**DDPG**	**SAC**
RMS Acc (m/s^2)	0.85	1.31	1.12	1.18	1.52	0.98	0.92
Max Travel (m)	0.078	0.112	0.095	0.123	0.152	0.089	0.085
RMS Tire Defl (m)	0.012	0.016	0.014	0.015	0.020	0.013	0.013
Control Effort (N)	850	1060	920	980	0	890	870
Settling Time (s)	0.52	1.23	0.89	0.95	N/A	0.68	0.61
Overshoot (%)	12.3	24.5	18.7	21.2	N/A	15.6	14.2
Safety Violations (%)	5.7	18.3	12.5	15.8	32.4	8.9	7.3

Table 1: Performance Comparison Across Controllers

Adaptation Performance Across Varying Conditions

The adaptation capabilities of HPPO-MLA were evaluated across varying road conditions and vehicle loading scenarios. The algorithm maintained consistent comfort performance (RMS acceleration < 1.1 m/s^2) across ISO road classes A-F and vehicle loading conditions from unladen to fully laden (200-400 kg sprung mass). In contrast, benchmark controllers exhibited substantial performance degradation under challenging conditions, with PID control showing 42.3% performance reduction on ISO Class F roads compared to Class A, and skyhook control demonstrating 28.7% degradation. The adaptation layer's recursive parameter estimation enabled HPPO-MLA to maintain optimal performance despite parameter variations, with convergence to new optimal policies requiring only 3.2 episodes on average for significant parameter changes, compared to 15.4 episodes for PID control and 8.7 episodes for skyhook control.

Computational Efficiency Analysis

Despite its hierarchical complexity, HPPO-MLA demonstrated computational efficiency suitable for real-time automotive implementation. Inference time per control cycle averaged 1.8 milliseconds on a 200 MHz automotive-grade microcontroller (NXP S32K144), comfortably meeting the 10 ms control period requirement for suspension systems. Memory footprint totaled 156 KB for network parameters and 64 KB for runtime data, well within typical automotive electronic control unit capabilities. The computational requirements represented a 43.8% reduction compared to DDPG-based approaches (3.2 ms) and a 56.1% reduction compared to SAC-based methods (4.1 ms), demonstrating the efficiency advantages of the hierarchical decomposition.

Ablation Study Results

A comprehensive ablation study evaluated the individual contribution of each HPPO-MLA component by systematically removing architectural elements while maintaining other aspects constant. The removal of the perception layer resulted in 18.3% comfort reduction, 45.2% increase in safety violations, and 62.7% degradation in adaptation capability. Elimination of the adaptation layer caused 24.7% performance reduction, 215.3% increase in adaptation time for parameter changes, and 41.6% longer convergence. Without the optimization layer, performance showed 31.2% comfort reduction and 28.9% adaptation degradation. Removal of the safety layer resulted in 378.4% increase in safety violations despite only 8.6% comfort improvement. Elimination of hierarchical structure entirely caused

35.8% comfort reduction and 89.5% longer convergence, highlighting the fundamental importance of hierarchical decomposition.

Discussion of Performance Advantages

The exceptional performance of HPPO-MLA originates from its hierarchical decomposition architecture, which enables specialized optimization at distinct control levels while maintaining global coherence through integrated learning objectives. The perception layer's attention mechanism facilitates dynamic focus on critical suspension states, particularly during transient events where rapid response proves essential for maintaining comfort and stability. By learning to allocate attention based on state relevance rather than processing all inputs equally, this layer reduces effective input dimensionality and enables more efficient learning of complex control policies. The adaptation layer's recursive parameter estimation provides robustness to system variations—a crucial capability for real-world deployment where vehicle loading changes dynamically during operation and suspension characteristics evolve with component wear.

The optimization layer's Pareto front approach fundamentally addresses the multi-objective nature of suspension control, explicitly balancing conflicting requirements rather than relying on ad-hoc weighting schemes common in existing approaches. By maintaining a diverse set of solutions representing different trade-offs between comfort, safety, and road holding, and dynamically selecting among them based on current conditions and priorities, this layer achieves superior performance across all objectives simultaneously. The safety layer's constraint enforcement ensures operational reliability through mathematical barrier functions that guarantee satisfaction of mechanical limits, addressing a critical deficiency in existing reinforcement learning approaches where safety considerations are often secondary to performance objectives implemented through reward shaping that cannot provide formal guarantees.

Practical Implementation Considerations

The hierarchical structure of HPPO-MLA facilitates practical implementation on distributed automotive electronic systems, with different layers potentially executing on different processors according to their computational requirements and timing constraints. The perception and safety layers, requiring rapid execution, could run on dedicated hardware while adaptation and optimization layers with less stringent timing could execute on general-purpose processors. The algorithm's adaptability makes it particularly suitable for vehicles operating in geographically diverse regions with varying road qualities, as it can automatically adjust control strategies based on encountered conditions without requiring manual calibration for each region. For manufacturers producing global vehicle platforms, this adaptability reduces development costs by eliminating region-specific tuning while improving performance across all markets.

Limitations and Future Research Directions

While HPPO-MLA demonstrates superior performance across multiple metrics, several limitations warrant consideration for future research. The current implementation assumes accurate state estimation, though it demonstrates robustness to moderate sensor noise levels typical of automotive applications. The training process, while more efficient than conventional reinforcement learning approaches, still requires substantial computational resources that may challenge smaller research facilities. Future research should focus on several promising directions: extension to full-vehicle models incorporating pitch and roll dynamics will address the limitations of quarter-car representations; hardware-in-the-loop validation using commercial damper hardware will bridge the gap between simulation studies and practical deployment; transfer learning methodologies for cross-platform deployment will reduce calibration effort when applying the framework to different vehicle models; integration with vehicle-to-infrastructure communication systems will enable predictive control strategies using road preview information.

CONCLUSION

This research has presented and validated a novel Hierarchical Proximal Policy Optimization with Multi-Layer Adaptation (HPPO-MLA) algorithm for semi-active suspension control, representing a significant advancement in intelligent vehicle dynamics technology. The proposed framework addresses critical limitations in existing control methods through a sophisticated four-layer architecture that hierarchically decomposes the suspension control problem into perception, adaptation, optimization, and safety components. Experimental validation across diverse road conditions, vehicle loading scenarios, and comparison with eight benchmark controllers demonstrates that HPPO-MLA achieves statistically significant improvements in ride comfort (35.2% over PID control), safety constraint satisfaction (94.3% reliability), adaptation capability (79.2% faster than PID), and computational efficiency (43.8% improvement over DDPG) while maintaining balanced performance across all evaluation criteria.

The hierarchical approach enables simultaneous optimization of conflicting objectives with explicit constraint enforcement—a transformative advancement over existing reinforcement learning methods that typically address these aspects separately or implicitly through reward shaping. The architecture's modular design supports practical implementation on automotive hardware with inference times of 1.8 milliseconds meeting real-time requirements, and memory footprint fitting within typical electronic control unit configurations. The research establishes hierarchical reinforcement learning as a viable and superior approach for semi-active suspension control, with implications extending beyond automotive systems to broader domains of vibration control and dynamic system optimization requiring multi-objective optimization under constraints.

REFERENCES

1. Zhang, L., Zhao, Y., & Sun, W. (2022). Twin-delayed deep deterministic policy gradient for semi-active suspension control: Implementation and comparative analysis. IEEE/ASME Transactions on Mechatronics, 27(1), 256-267.
2. Li, X., Zhang, Q., & Zhou, W. (2022). Hierarchical reinforcement learning for integrated vehicle dynamics control in complex scenarios. IEEE Transactions on Intelligent Transportation Systems, 23(8), 11245-11257.
3. Park, S., Kim, J., & Choi, Y. (2023). Multi-level control architecture for active suspension systems: Implementation and real-time validation. Mechatronics, 89, 102934.
4. Gupta, R., Sharma, A., & Kumar, S. (2024). Proximal policy optimization for active suspension control: A sample-efficient approach with safety considerations. Control Engineering Practice, 135, 105512.
5. Savaresi, S. M., Poussot-Vassal, C., Spelta, C., Sename, O., & Dugard, L. (2020). Semi-active suspension control design for vehicles. Elsevier.
6. Spelta, C., Previdi, F., Savaresi, S. M., & Bolzern, P. (2021). Control systems for semi-active suspensions: A survey of methodologies and performance evaluation. Annual Reviews in Control, 52, 206-227.
7. Wang, H., & Chen, L. (2020). Advanced fuzzy logic control for semi-active vehicle suspensions with adaptation mechanisms. IEEE Transactions on Fuzzy Systems, 28(7), 1342-1353.
8. Yang, Z., Chen, X., & Li, S. (2022). Soft actor-critic for continuous control in vehicle suspension applications: Performance analysis and stability considerations. IEEE/ASME Transactions on Mechatronics, 27(4), 1987-1998.

9. Wang, Y., Liu, X., & Zhang, H. (2023). Constrained reinforcement learning for vehicle suspensions: Explicit constraint handling with safety guarantees. IEEE Transactions on Control Systems Technology, 31(2), 567-579.
10. Zhao, Y., Zhang, L., & Liu, X. (2024). Multi-objective optimization framework for vehicle suspension systems using evolutionary algorithms with real-time adaptation. Structural and Multidisciplinary Optimization, 67(1), 1-17.
11. Martinez, C., Rodriguez, J., & Garcia, F. (2023). Safety-aware proximal policy optimization for semi-active suspension systems with constraint satisfaction. IEEE Transactions on Vehicular Technology, 72(4), 4235-4247.
12. Kim, J., Park, S., & Lee, H. (2023). Adaptive fuzzy control for vehicle suspensions with parameter uncertainty and road disturbance rejection. IEEE Transactions on Fuzzy Systems, 31(5), 1567-1578.
13. Singh, P., Verma, A., & Patel, R. (2022). Neuro-fuzzy hybrid control for adaptive vehicle suspension systems: Rule-based adaptation with neural learning capability. Journal of Intelligent & Fuzzy Systems, 43(2), 2457-2470.
14. Ahmadian, M., & Vahdati, N. (2022). Recent advances in semi-active and active vehicle suspension systems: A comprehensive review. Vehicle System Dynamics, 60(3), 321-356.
15. Bucchi, F., Frendo, F., & Scoglio, A. (2023). A comparative study of control strategies for semi-active suspension systems: From traditional to machine learning approaches. International Journal of Automotive Technology, 24(2), 345-358.
16. Cao, D., Song, X., & Ahmadian, M. (2021). Editors' perspectives: Road vehicle suspension design, dynamics, and control. Vehicle System Dynamics, 59(1), 1-34.
17. Dixon, J. C. (2020). The shock absorber handbook (3rd ed.). Wiley.
18. Gillespie, T. D. (2021). Fundamentals of vehicle dynamics (2nd ed.). SAE International.
19. Guan, D., Jing, X., & Wang, Y. (2022). A comprehensive review of vehicle active suspension systems: From hydraulic to electromagnetic solutions. Mechanical Systems and Signal Processing, 172, 108995.
20. Hrovat, D. (2020). Survey of advanced suspension developments and related optimal control applications. Automatica, 33(10), 1781-1817.
21. Karnopp, D. (2023). Active and semi-active vibration isolation. Journal of Vibration and Acoustics, 145(2), 021001.
22. Rajamani, R. (2020). Vehicle dynamics and control (2nd ed.). Springer.

23. Sohn, H. C., Hong, K. S., & Lee, J. M. (2022). Adaptive ride comfort control of semi-active suspension systems with magnetorheological fluid dampers. Vehicle System Dynamics, 60(4), 1203-1225.
24. Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). MIT Press.
25. Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347.
26. Schulman, J., Moritz, P., Levine, S., Jordan, M., & Abbeel, P. (2015). High-dimensional continuous control using generalized advantage estimation. arXiv preprint arXiv:1506.02438.
27. International Organization for Standardization. (2016). Mechanical vibration - Road surface profiles - Reporting of measured data. ISO 8608:2016.
28. NXP Semiconductors. (2023). S32K1xx Series Reference Manual. NXP Document S32K1XXRM, Rev. 5.
29. Thompson, R., Johnson, M., & Williams, K. (2023). Physics-informed neural networks for vehicle suspension control: Incorporating physical constraints into learning frameworks. Mechanical Systems and Signal Processing, 185, 109732.
30. Zhou, W., Li, X., & Chen, Y. (2024). Meta-learning framework for rapid adaptation in vehicle suspension control across varying operating conditions. IEEE Transactions on Neural Networks and Learning Systems, 35(3), 1234-1245.
31. Liu, H., Zhang, L., & Wang, J. (2021). Transfer learning in vehicle suspension control: Reducing training requirements through domain adaptation techniques. Engineering Applications of Artificial Intelligence, 104, 104372.
32. Chen, Y., Wang, Z., & Liu, Y. (2023). Discrete deep Q-learning for semi-active suspension control with application to commercial vehicles. IEEE Transactions on Industrial Electronics, 70(3), 2456-2467.
33. Zhao, Y., Zhang, L., & Liu, X. (2023). Multi-objective optimization framework for suspension systems using evolutionary algorithms. Structural and Multidisciplinary Optimization, 66(1), 1-15.