American Journal of Data Science and Machine Learning
ISSN: - applied

Learning Based Optimal Control for Nonlinear Robotic Systems: Integrating Dynamic Programming, Analytical Rigid Body Models, and Diffusion Policies

[1] Dr. Ingrid S. Johansen
[1] Department of Mechanical and Process Engineering ETH Zurich Switzerland

Received: 19th Dec 2025 | Received Revised Version: 23th Dec 2025 | Accepted: 30th Dec 2025 | Published: 18th Jan 2026

Volume 02 Issue 01 2026 | Crossref DOI: 10.64917/ajdsml/V02I01-003

Abstract

The convergence of optimal control theory, robotics, and machine learning has produced a new generation of methodologies capable of addressing high dimensional nonlinear dynamical systems. Classical dynamic programming, rooted in the foundational work of Bellman, provides a principled framework for sequential decision making but suffers from the curse of dimensionality. Trajectory optimization methods such as Differential Dynamic Programming and modern numerical techniques enable efficient local optimization but remain sensitive to initialization and model inaccuracies. Concurrently, advances in rigid body dynamics libraries, including the development of analytical derivatives and fast forward and inverse dynamics algorithms, have significantly reduced computational overhead in robot modeling. In parallel, machine learning has introduced nonparametric representations, Gaussian process based dynamic programming, deep neural approximations for stochastic control, and diffusion based policy learning frameworks that challenge conventional paradigms.

This article presents a comprehensive theoretical integration of these traditions. Drawing strictly from the referenced works, it synthesizes trajectory based nonparametric value representation, random state sampling in dynamic programming, Gaussian process control, analytical rigid body dynamics, machine learning assisted model predictive control, diffusion policy learning, and neural approximations of Pontryagin based optimality conditions. The research develops a unified perspective in which analytical physics based modeling and data driven learning are not competing alternatives but complementary layers within a hierarchical optimal control architecture.

The methodology elaborates a conceptual framework combining exact rigid body derivatives from the Pinocchio library with learning based policy representations trained via stochastic optimal control matching and diffusion models. A detailed theoretical comparison between classical feedback control design and modern deep learning approximations is conducted. The results section provides an extensive descriptive analysis of how these combined approaches improve scalability, generalization, and closed loop robustness in robotic manipulators and aerial platforms such as micro quadrotors and lightweight robotic arms.

The discussion critically examines computational complexity, generalization across domains, and the epistemological implications of replacing explicit dynamic programming recursions with learned approximations. The article concludes by outlining a path toward interpretable, scalable, and computationally efficient learning based optimal control for complex robotic systems.

Keywords: Optimal control, Dynamic programming, Rigid body dynamics, Diffusion policy, Machine learning control, Trajectory optimization.

© 2025 Dr. Ingrid S. Johansen. This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). The authors retain copyright and allow others to share, adapt, or redistribute the work with proper attribution.

Cite This Article: Dr. Ingrid S. Johansen. 2026. Learning Based Optimal Control for Nonlinear Robotic Systems: Integrating Dynamic Programming, Analytical Rigid Body Models, and Diffusion Policies. American Journal of Data Science and Machine Learning 2, 01, 12-16. https://doi.org/10.64917/ajdsml/V02I01-003

1. Introduction

Optimal control theory has historically provided a rigorous mathematical framework for determining control policies that minimize cumulative cost in dynamical systems. The conceptual origin of this framework can be traced to Bellman's formulation of dynamic programming, which formalized the principle of optimality and introduced recursive value function computation as the backbone of sequential decision processes (Bellman, 1957). Bellman's insights established a theoretical structure that unified discrete and continuous decision making, yet the practical implementation of dynamic programming rapidly encountered the exponential growth of computational complexity as state dimensionality increased.

The curse of dimensionality is not merely a computational inconvenience but a structural barrier. When robotic systems evolved from simple planar manipulators to high degree of freedom articulated platforms and aerial vehicles, the dimensional explosion rendered exact dynamic programming intractable. For this reason, researchers turned to approximate methods and trajectory based approaches. Jacobson and Mayne introduced Differential Dynamic Programming as a second order trajectory optimization method that iteratively improves local control sequences by approximating value functions along trajectories (Jacobson and Mayne, 1970). Later, Betts provided a comprehensive survey of numerical trajectory optimization techniques, emphasizing direct and indirect methods for solving large scale nonlinear optimal control problems (Betts, 1998).

Parallel to developments in optimal control, robotics experienced a revolution in hardware design and modeling fidelity. The KUKA DLR lightweight robot arm represented a milestone in compliant manipulation and torque controlled architectures, providing a research platform that demanded real time optimal control solutions (Bischoff et al., 2010). At the modeling level, Carpentier and Mansard developed analytical derivatives for rigid body dynamics algorithms, enabling efficient gradient based optimization (Carpentier and Mansard, 2018). The Pinocchio library further consolidated these advances by offering a fast and flexible implementation of rigid body dynamics and their derivatives, which are essential for trajectory optimization and model predictive control (Carpentier et al., 2019; Carpentier et al., 2015–2021).

Despite these advances, model inaccuracies, sensor noise, and environmental uncertainty limited purely analytical approaches. Machine learning introduced new paradigms. Atkeson and Morimoto proposed nonparametric representations of policies and value functions through trajectory based memory methods, demonstrating that stored experiences could approximate dynamic programming solutions without explicit parametric forms (Atkeson and Morimoto, 2002). Later, Atkeson and Stephens explored random sampling of states to mitigate the dimensionality problem in dynamic programming (Atkeson and Stephens, 2007). Deisenroth, Rasmussen, and Peters developed Gaussian process dynamic programming, which incorporated probabilistic modeling of system dynamics and uncertainty quantification into control design (Deisenroth et al., 2009).

The last decade witnessed a deeper integration between machine learning and optimal control. Han and E introduced deep learning approximations for stochastic control, showing that neural networks could approximate value functions and policies in high dimensional systems (Han and E, 2016). E, Han, and Zhang framed machine learning assisted modeling as a general scientific paradigm, arguing that data driven components can complement physics based models (E et al., 2021). In the context of model predictive control, E, Han, and Long highlighted the potential of machine learning to enhance real time predictive control strategies (E et al., 2022). Domingo Enrich and collaborators introduced stochastic optimal control matching, bridging generative modeling and control theory (Domingo Enrich et al., 2024). Bottcher and colleagues demonstrated that neural networks can learn to satisfy Pontryagin type optimality conditions, a perspective termed AI Pontryagin (Bottcher et al., 2022). More recently, diffusion models have been adapted for policy learning, enabling expressive action distributions in visuomotor tasks (Chi et al., 2025; Hegde et al., 2023).

The convergence of these strands reveals a literature gap. While individual contributions address either analytical modeling, classical control, or machine learning, a comprehensive theoretical synthesis that integrates rigid body analytical derivatives, dynamic programming approximations, stochastic optimal control matching, and diffusion based policies for nonlinear robotic systems remains underdeveloped. The present article addresses this gap by constructing a unified conceptual architecture grounded strictly in the cited works.

2. Methodology

The methodology presented here is conceptual and theoretical, synthesizing existing frameworks into a unified learning based optimal control architecture. The approach consists of five interlocking components: analytical rigid body modeling, trajectory based optimal control, nonparametric and probabilistic value representation, deep neural approximations of stochastic control, and diffusion based policy generation.

The first component concerns system modeling. Rigid body dynamics are modeled using algorithms that compute forward and inverse dynamics efficiently, as implemented in the Pinocchio library (Carpentier et al., 2015–2021; Carpentier et al., 2019). Analytical derivatives of these algorithms provide exact gradients of dynamics with respect to states and controls (Carpentier and Mansard, 2018). These derivatives are critical for gradient based trajectory optimization methods such as Differential Dynamic Programming (Jacobson and Mayne, 1970) and for numerical trajectory optimization approaches surveyed by Betts (Betts, 1998). The methodology assumes that robotic platforms such as the KUKA DLR lightweight arm (Bischoff et al., 2010) or micro quadrotors (Bouabdallah et al., 2004) can be described by such rigid body models.

The second component employs trajectory optimization as an initial policy generator. Differential Dynamic Programming iteratively refines a nominal trajectory by locally approximating the value function along the trajectory and updating control inputs accordingly (Jacobson and Mayne, 1970). This provides a locally optimal control sequence under the assumption of model accuracy. In parallel, feedback control principles as described by Franklin, Powell, and Emami Naeini are used to ensure closed loop stability around optimized trajectories (Franklin et al., 2002). This hybrid of feedforward optimal control and feedback stabilization forms the baseline controller.

The third component introduces nonparametric and probabilistic approximations to overcome dimensionality. Atkeson and Morimoto's trajectory based nonparametric representation stores observed trajectories and interpolates value estimates for new states (Atkeson and Morimoto, 2002). Random sampling of states, as explored by Atkeson and Stephens, reduces computational burden by focusing evaluation on representative subsets of the state space (Atkeson and Stephens, 2007). Gaussian process dynamic programming extends this by modeling transition dynamics probabilistically and propagating uncertainty through value updates (Deisenroth et al., 2009). In this framework, uncertainty is not treated as noise to be ignored but as structured information influencing policy selection.

The fourth component incorporates deep learning approximations for stochastic optimal control. Han and E proposed representing value functions with deep neural networks trained to satisfy stochastic control equations (Han and E, 2016). This approach addresses high dimensionality by leveraging universal function approximation capabilities. Bottcher and colleagues extended this by demonstrating that neural networks can implicitly learn Pontryagin type optimality conditions, effectively internalizing adjoint dynamics and control gradients (Bottcher et al., 2022). The methodology interprets these networks as differentiable surrogates for classical control laws.

Neural network training stability is influenced by architectural and activation choices. Glorot and Bengio analyzed initialization strategies to mitigate vanishing gradients in deep feedforward networks (Glorot and Bengio, 2010). Clevert, Unterthiner, and Hochreiter introduced exponential linear units to accelerate convergence and improve robustness (Clevert et al., 2016). These considerations are integrated into the network design for value and policy approximation.

The fifth component integrates diffusion based policy learning. Diffusion Policy, as introduced by Chi and collaborators, models action generation as a diffusion process conditioned on visual and state observations (Chi et al., 2025). Hegde and colleagues further demonstrated that latent diffusion models can generate behaviorally diverse policies (Hegde et al., 2023). In this methodology, diffusion models are trained on trajectories generated by the trajectory optimization and Gaussian process controllers, thereby inheriting optimal structure while gaining expressive stochasticity.

Stochastic optimal control matching, as developed by Domingo Enrich and colleagues, aligns generative models with optimal control distributions by minimizing discrepancies between controlled trajectory distributions and learned generative samples (Domingo Enrich et al., 2024). Kernel maximum mean discrepancy, originally proposed for structured biological data integration, provides a statistical tool for matching distributions (Borgwardt et al., 2006). This technique ensures that diffusion generated policies remain consistent with optimal control objectives.

Machine learning assisted model predictive control, as articulated by E, Han, and Long, serves as the real time deployment mechanism (E et al., 2022). Neural approximations of value functions are embedded within predictive control loops, updating control actions based on current state estimates while leveraging learned models. This hybridization maintains interpretability through physics based constraints while achieving scalability through learned approximations.

3. Results

The integrated framework yields several conceptual findings when applied to nonlinear robotic systems. First, the combination of analytical derivatives and learning based approximations significantly reduces computational latency compared to pure dynamic programming. Analytical gradients from Pinocchio accelerate local trajectory optimization, while neural approximations eliminate the need for exhaustive state space discretization.

Second, probabilistic modeling via Gaussian processes improves robustness under model uncertainty. By explicitly modeling transition uncertainty, the controller anticipates deviations and selects actions that hedge against variance, a property absent in purely deterministic dynamic programming (Deisenroth et al., 2009).

Third, diffusion based policies demonstrate superior expressiveness in multi modal tasks. For example, in manipulation scenarios involving the KUKA DLR arm, multiple feasible grasp trajectories may exist. Diffusion models capture this diversity, generating distinct yet cost consistent action sequences (Chi et al., 2025; Hegde et al., 2023). This contrasts with classical optimal control, which typically converges to a single local minimum.

Fourth, neural approximations of Pontryagin optimality exhibit scalability to higher dimensional systems. Bottcher et al. showed that networks can internalize necessary conditions for optimality (Bottcher et al., 2022). When integrated with machine learning assisted model predictive control, these approximations enable real time updates in complex manipulators (Hu et al., 2025).

Fifth, stability analysis grounded in classical feedback theory ensures that learned components do not destabilize the system. Embedding neural policies within feedback loops as described by Franklin et al. maintains bounded error dynamics (Franklin et al., 2002).

4. Discussion

The theoretical synthesis presented reveals both strengths and limitations. The integration of analytical derivatives with neural approximations preserves physical consistency while enabling scalability. However, dependence on accurate rigid body models may limit performance in unmodeled contact dynamics. While Gaussian processes quantify uncertainty, their computational complexity grows with data size, necessitating sparse approximations.

Diffusion models offer expressive policy distributions but require extensive training data, potentially derived from computationally intensive trajectory optimization. Moreover, interpretability remains a concern. Although AI Pontryagin demonstrates that networks can approximate optimality conditions, extracting explicit control laws from deep networks is challenging (Bottcher et al., 2022).

Future research may explore tighter integration between stochastic optimal control matching and analytical adjoint methods, enabling principled regularization of generative policies. Another direction involves extending machine learning assisted modeling frameworks to incorporate structural priors from rigid body dynamics more explicitly (E et al., 2021).

5. Conclusion

This article has developed a unified theoretical framework integrating dynamic programming, trajectory optimization, analytical rigid body modeling, Gaussian process dynamic programming, deep learning approximations of stochastic control, and diffusion based policy learning. Grounded strictly in the referenced works, the study demonstrates that learning based optimal control for nonlinear robotic systems is most effective when analytical physics based insights and machine learning approximations are treated as complementary components rather than competing paradigms. The resulting architecture offers scalability, robustness, and expressive policy generation suitable for advanced robotic platforms.

References

1. Atkeson C and Morimoto J. Nonparametric representation of policies and value functions: A trajectory based approach. Advances in Neural Information Processing Systems. Curran Associates

ISSN: - applied

Inc. 2002.

2. Atkeson C and Stephens B. Random sampling of states in dynamic programming. Advances in Neural Information Processing Systems. Curran Associates Inc. 2007.
3. Bellman R E. Dynamic Programming. Princeton University Press. 1957.
4. Betts J T. Survey of numerical methods for trajectory optimization. Journal of Guidance Control and Dynamics. 1998.
5. Bischoff R et al. The KUKA DLR Lightweight Robot arm a new reference platform for robotics research and manufacturing. ISR 2010 and ROBOTIK 2010. VDE. 2010.
6. Borgwardt K M et al. Integrating structured biological data by kernel maximum mean discrepancy. Bioinformatics. 2006.
7. Bottcher L Antulov Fantulin N and Asikis T. AI Pontryagin or how artificial neural networks learn to control dynamical systems. Nature Communications. 2022.
8. Bouabdallah S Murrieri P and Siegwart R. Design and control of an indoor micro quadrotor. Proceedings of the IEEE International Conference on Robotics and Automation. 2004.
9. Carpentier J and Mansard N. Analytical derivatives of rigid body dynamics algorithms. Robotics Science and Systems. 2018.
10. Carpentier J et al. The Pinocchio C plus plus library: A fast and flexible implementation of rigid body dynamics algorithms and their analytical derivatives. IEEE SICE International Symposium on System Integration. 2019.
11. Carpentier J et al. Pinocchio: Fast forward and inverse dynamics for poly articulated systems. 2015 to 2021.
12. Chi C Xu Z Feng S Cousineau E Du Y Burchfiel B Tedrake R and Song S. Diffusion policy: Visuomotor policy learning via action diffusion. International Journal of Robotics Research. 2025.
13. Clevert D Unterthiner T and Hochreiter S. Fast and accurate deep network learning by exponential linear units. International Conference on Learning Representations. 2016.
14. Coates A Abbeel P and Ng A Y. Learning for control from multiple demonstrations. International Conference on Machine Learning. 2008.
15. Deisenroth M P Rasmussen C E and Peters J. Gaussian process dynamic programming. Neurocomputing. 2009.
16. Domingo Enrich C Han J Amos B Bruna J and Chen R T Q. Stochastic optimal control matching. Advances in Neural Information Processing Systems. 2024.
17. E W Han J and Long J. Empowering optimal control with machine learning: A perspective from model predictive control. IFAC PapersOnLine. 2022.
18. E W Han J and Zhang L. Machine learning assisted modeling. Physics Today. 2021.
19. Franklin G F Powell J D and Emami Naeini A. Feedback Control of Dynamic Systems. Upper Saddle River. 2002.
20. Glorot X and Bengio Y. Understanding the difficulty of training deep feedforward neural networks. International Conference on Artificial Intelligence and Statistics. 2010.
21. Han J and E W. Deep learning approximation for stochastic control problems. arXiv. 2016.
22. Hegde S Batra S Zentner K and Sukhatme G. Generating behaviorally diverse policies with latent diffusion models. Advances in Neural Information Processing Systems. 2023.
23. Hu W Zhao Y E W Han J and Long J. Learning free terminal time optimal closed loop control of manipulators. American Control Conference. 2025.
24. Jacobson D H and Mayne D Q. Differential Dynamic Programming. Elsevier Publishing Company. 1970.