KeAi
CHINESE ROOTS
GLOBAL IMPACT

International Journal of Innovation Studies

ADVANCED MACHINE LEARNING FRAMEWORKS FOR OPTIMIZATION AND INTELLIGENT CONTROL IN LARGE-SCALE ENGINEERING SYSTEMS

Dr.Madduru Sambasivudu[1], Thota Bhavani[2], Dr. Suresh Palarimath[3] , Dr. Rahul Kumar[4] , Akansh Garg[5]

[1]Associate Professor, Computer Science and Engineering, Malla Reddy College of Engineering and Technology samba.siva57@gmail.com

[2]M.Tech Student, Department Of Computer Science and Engineering, Malla Reddy College Engineering and Technology thotabhavani53@gmail.com

[3]Lecturer, College of Computing and Information Sciences, University of Technology and Applied Sciences Salalah, Dhofar, Salalah, Sultanate of Oman, suresh.palarimath@utas.edu.om

[4]Assistant Professor, Amity School of Business, Amity University Jharkhand, Ranchi,. Jharkhand, rkumar5@rnc.amity.edu

[5] Director Array Research Pvt Ltd, 7505264391akg@gmail.com

Abstract

Large-scale engineering systems spanning power grids, industrial manufacturing networks, autonomous transportation infrastructures, and smart city platforms present optimization and control challenges of a complexity and dimensionality that fundamentally exceed the capacity of classical analytical methods, necessitating the development of advanced machine learning frameworks capable of learning adaptive control policies from high-dimensional system data while guaranteeing the safety, stability, and robustness properties that safety-critical engineering deployments demand. This paper presents a comprehensive unified framework integrating deep reinforcement learning, Gaussian process-based Bayesian optimization, physics-informed neural networks, and multi-agent coordination architectures into a hierarchical intelligent control system evaluated across three large-scale engineering benchmarks: a 118-bus power grid with 54 controllable generation units, a 12-stage chemical process plant with coupled nonlinear dynamics, and a 200-node urban traffic network with mixed autonomous and human-driven vehicles. The proposed Hierarchical "Adaptive Machine Learning Control (HAMLC)" framework achieves mean control performance improvements of 23.4 percent over model predictive control baselines, 31.7 percent over classical proportional-integral-derivative control, and 12.1 percent over single-algorithm machine learning benchmarks, while maintaining constraint violation rates below 0.3 percent across all test scenarios. Formal stability guarantees derived from Lyapunov function approximation theory are validated computationally, and the framework demonstrates robust performance degradation under sensor noise, communication delays, and partial observability conditions representative of real-world large-scale engineering deployments. The integration of physics-informed neural network components is shown to reduce the training data requirements by 67 percent compared to purely data-driven baselines, addressing a critical bottleneck for practical deployment in engineering systems where operational data collection is constrained by safety and cost considerations.

Keywords: Deep Reinforcement Learning, Bayesian Optimization, Physics-Informed Neural Networks, Intelligent Control, Large-Scale Systems, Multi-Agent Systems, Power Grid Control, Process Control, Lyapunov Stability, Engineering Optimization

I. INTRODUCTION

The optimization and control of large-scale engineering systems represent one of the most demanding application domains for advanced computational intelligence, combining the high-dimensional state spaces, nonlinear and partially observable dynamics, tight safety constraints, and real-time decision requirements that individually challenge state-of-the-art methods and collectively define a problem class for which no single existing approach is fully adequate [1]. Industrial process control systems managing hundreds of coupled variables with millisecond response requirements, power grids balancing supply and demand across thousands of nodes under stochastic renewable generation, and smart transportation networks coordinating thousands of autonomous agents through complex interaction dynamics all exemplify the class of large-scale engineering control problems for which advanced machine learning frameworks offer transformative potential alongside formidable technical challenges [2].

Classical optimal control methods linear quadratic regulators, model predictive control, and their nonlinear extensions provide theoretically rigorous frameworks with well-understood stability and performance guarantees but require accurate analytical system models whose derivation from first principles is infeasible for many large-scale engineering systems, and whose computational tractability deteriorates rapidly with state space dimensionality [3]. Data-driven system identification approaches can partially address the modeling challenge but introduce model uncertainty that classical control theory handles conservatively, leading to performance penalties that erode the value of adaptive approaches. The integration of machine learning into control systems has been pursued through multiple paradigms model-based reinforcement learning learning dynamics models from data, model-free reinforcement learning directly optimizing control policies through environment interaction, and hybrid approaches combining learned models with classical control structures each offering different tradeoffs between sample efficiency, performance, and safety guarantees [4].

Deep reinforcement learning has demonstrated remarkable control performance in simulated engineering environments, with notable successes including AlphaGo's game-playing performance, Google DeepMind's data center cooling optimization, and various robotics manipulation benchmarks [5]. However, the translation of these laboratory successes to real-world large-scale engineering deployment faces multiple barriers: the sample inefficiency of deep RL algorithms requiring millions of environment interactions that are costly or dangerous to obtain in real engineering systems, the absence of formal stability guarantees for neural network control policies, the challenge of safe exploration in environments where constraint violations may cause physical damage or safety incidents, and the brittleness of learned policies to distribution shifts between training and deployment conditions [6].

Bayesian optimization provides a complementary approach to engineering system optimization characterized by principled uncertainty quantification, efficient exploration of high-dimensional parameter spaces, and compatibility with expensive black-box objective functions typical of engineering design problems [7]. Gaussian process surrogate models at the core of Bayesian optimization enable sample-efficient optimization through uncertainty-driven acquisition functions but face computational scaling challenges for the large evaluation sets and high-dimensional spaces characteristic of large-scale engineering problems. Physics-informed neural networks neural networks trained with physics equation residuals as regularization terms alongside data-fitting objectives offer a path to sample-efficient learning by incorporating domain knowledge from engineering governing equations directly into the learning architecture [8].

This paper develops the Hierarchical Adaptive Machine Learning Control framework as a principled integration of these complementary approaches, assigning control responsibilities to framework components according to their relative strengths: Bayesian optimization for sample-efficient high-level parameter tuning, physics-informed neural networks for dynamics model learning exploiting governing equation knowledge, deep reinforcement learning for adaptive policy optimization given learned dynamics models, and multi-agent coordination architectures for managing the distributed decision-making inherent in large-scale systems. Formal stability analysis connects the learned Lyapunov function approach of the stability guarantee module to established Lyapunov stability theory, providing rigorous performance bounds under identified assumptions. The framework is evaluated on three large-scale engineering benchmarks spanning power systems, chemical process, and transportation domains that collectively test the generality and robustness of the proposed approach [9], [10].

II. OBJECTIVES

Objective 1: To develop the HAMLC framework architecture integrating Bayesian optimization, physics-informed neural networks, deep reinforcement learning, and multi-agent coordination into a coherent hierarchical intelligent control system for large-scale engineering applications.

Objective 2: To derive and validate formal stability guarantees for the HAMLC framework through learned Lyapunov function approximation, establishing theoretical foundations for safety-critical engineering deployment.

Objective 3: To evaluate framework control performance across three large-scale engineering benchmarks—power grid, chemical process, and urban traffic—demonstrating performance improvements over classical and single-algorithm machine learning baselines.

Objective 4: To quantify the sample efficiency benefits of physics-informed neural network integration compared to purely data-driven dynamics learning, establishing the practical value of physics knowledge incorporation for data-constrained engineering deployment.

Objective 5: To assess framework robustness under realistic operational conditions including sensor noise, communication delays, component failures, and partial observability, validating deployment feasibility in real large-scale engineering environments.

III. RELATED WORKS

The application of reinforcement learning to engineering control problems has a history extending to early tabular Q-learning applications to scheduling and inventory management, but the modern deep reinforcement learning era was inaugurated by the DQN work of Mnih et al. [11], which demonstrated human-level performance on Atari video games using convolutional neural network function approximators for Q-value estimation. Subsequent algorithmic developments including deep deterministic policy gradient [12] for continuous action spaces, proximal policy optimization [13] for stable on-policy learning, and soft actor-critic [14] for maximum entropy reinforcement learning with improved sample efficiency have progressively extended deep RL applicability to the continuous high-dimensional control problems characteristic of engineering systems. The specific application to power system control has been explored by Yan and Xu [15], who demonstrated deep RL-based voltage control in distribution networks, and by Diao et al. [16], who applied deep RL to emergency frequency regulation in power grids with renewable generation uncertainty.

Bayesian optimization for engineering design and control parameter tuning has been developed from the foundational Gaussian process optimization framework of Jones, Schonlau, and Welch [17] through numerous extensions addressing the challenges of high-dimensional optimization, noisy objective function evaluation, and batch acquisition for parallel evaluation. Snoek, Larochelle, and Adams [18] demonstrated the effectiveness of Bayesian optimization for neural network hyperparameter tuning, establishing the approach as a practical tool for machine learning engineering. The application to control system parameter optimization has been explored by Berkenkamp et al. [19], who developed safe Bayesian optimization with high-probability constraint satisfaction guarantees particularly relevant for engineering control contexts where parameter configurations that violate safety constraints cannot be evaluated experimentally.

Physics-informed neural networks were introduced by Raissi, Perdikaris, and Karniadakis [8], who demonstrated that neural networks trained with physics equation residuals alongside data-fitting objectives can learn solutions to partial differential equations with significantly fewer data points than purely data-driven approaches, while inheriting the generalization properties of physics-based models. Extensions to engineering control and state estimation include physics-informed Kalman filtering, physics-constrained optimization, and physics-regularized neural ordinary differential equations for dynamics learning [20]. The integration of physics knowledge into reinforcement learning through model-based approaches with physics-constrained world models has been explored by Lutter, Ritter, and Peters [21], who demonstrated improved sample efficiency and generalization for robotic control through physics model regularization.

Multi-agent reinforcement learning for large-scale system control addresses the decentralized decision-making structure inherent in systems with many interacting subsystems, where centralized optimization is computationally intractable and where communication constraints prevent perfect information sharing across agents [22]. Cooperative multi-agent RL frameworks including QMIX [23], MADDPG [24], and mean-field RL have been applied to multi-robot coordination, distributed sensor networks, and smart grid demand response. The stabilizing effect of agent communication and the design of communication protocols that maximize control performance under bandwidth constraints are active research areas with direct relevance to large-scale engineering control where subsystem coordination is essential but communication infrastructure is limited [25].

Stability guarantees for neural network control systems have been pursued through multiple approaches including verification of neural network Lyapunov functions by Chang, Roohi, and Gao [26], who demonstrated that neural networks can learn valid Lyapunov functions for nonlinear control systems when trained with Lyapunov condition satisfaction as a training objective, and control barrier function approaches by Ames et al. [27] that impose safety constraint satisfaction through optimization modifications to nominal control policies. The combination of learned Lyapunov functions with deep RL training to obtain stability-guaranteed adaptive control policies is a developing research area that this paper advances through the HAMLC stability module design.

IV. METHODOLOGY

4.1 HAMLC Framework Architecture

The HAMLC framework is organized as a three-level hierarchy. The strategic level employs Bayesian optimization with a Gaussian process surrogate to tune high-level control objective weights and constraint tolerances on a daily or shift-length timescale, using the upper confidence bound acquisition function with automatic relevance determination kernels adapted to the dimensional structure of engineering objective functions. The tactical level employs a physics-informed neural network world model trained with both operational data and governing equation residuals (conservation laws, thermodynamic constraints, electrical network equations as appropriate to the benchmark domain) to provide a computationally efficient differentiable dynamics model for model-based deep RL policy optimization. The operational level executes control actions at the system sampling rate using the deep RL policy conditioned on the tactical-level world model predictions, with multi-agent coordination handled through a communication-efficient mean-field approximation that scales linearly in the number of agents rather than exponentially as in exact cooperative MARL approaches.

TABLE I: HAMLC Framework Component Summary

Framework Level	Component	Algorithm	Update Timescale	Primary Responsibility
Strategic	Hyperparameter Optimizer	Gaussian Process BO	Daily/Shift	Objective weights, constraint tolerances
Tactical	World Model	Physics-Informed NN (PINN)	Hourly	Dynamics learning, state prediction
Tactical	Policy Optimizer	Soft Actor-Critic	Hourly	Control policy training
Operational	Controller	Multi-Agent RL Policy	Milliseconds	Real-time action execution

Framework Level	Component	Algorithm	Update Timescale	Primary Responsibility
Safety	Stability Monitor	Lyapunov NN Verifier	Continuous	Constraint monitoring, safe fallback

4.2 Physics-Informed Neural Network World Model

The PINN world model learns the system dynamics f(x,u) from a combination of operational data D = {(xt, ut, xt+1)} and physics residual losses L_phys encoding the governing equations of the engineering domain. For the power grid benchmark, physics residuals encode AC power flow equations (Kirchhoff's current and voltage laws) and generator dynamics (swing equation, automatic voltage regulator dynamics). For the chemical process benchmark, residuals encode material and energy balances and kinetic rate equations. For the traffic benchmark, residuals encode the Lighthill-Whitham-Richards traffic flow model. The combined training loss L = alpha*L_data + beta*L_phys + gamma*L_reg balances data fitting with physics constraint satisfaction, with hyperparameters alpha, beta, gamma tuned by the strategic-level Bayesian optimizer.

The PINN architecture employs a modified ResNet with 8 hidden layers of 512 units, incorporating physics-inspired activation functions (hyperbolic tangent for smooth differential equation solutions) and separate output heads for different physics domains that are individually constrained to satisfy domain-specific physical bounds. Uncertainty quantification uses a deep ensemble of 10 PINN instances, with the ensemble variance providing an epistemic uncertainty estimate used by the policy optimizer for safe exploration.

4.3 Stability Guarantee Module

The stability guarantee module learns a neural Lyapunov function V(x) simultaneously with the control policy pi(x), trained with the condition that V(x) > 0 for x != x* and the Lyapunov decrease condition nabla V(x) * f(x, pi(x)) < 0 must hold in the operating region. When the stability monitor detects potential Lyapunov condition violations in the predicted future trajectory, it activates a safe fallback controller (classical MPC with conservative constraint tightening) until the RL policy returns to the verified safe region. This provides a formal guarantee that the overall closed-loop system is Lyapunov stable under the conditions verified by the stability module, with the RL policy responsible for performance optimization within the verified safe operating region.

TABLE II: Benchmark System Specifications

Benchmark	System Scale	State Dimension	Action Dimension	Constraint Count	Sampling Rate
Power Grid (IEEE 118-bus)	118 buses, 54 generators	342	54 (generation dispatch)	218 (line, voltage limits)	1 second
Chemical Process (12-stage)	12 stages, 47 unit ops	284	31 (flow, temp setpoints)	94 (safety, quality)	100 ms
Urban Traffic (200-node)	200 intersections, ~8,000 vehicles	1,240	400 (signal timings)	200 (safety clearances)	30 seconds

V. RESULTS AND ANALYSIS

5.1 Control Performance Benchmarking

The HAMLC framework achieves statistically significant performance improvements over all baseline methods across the three engineering benchmarks. On the power grid benchmark, HAMLC reduces operational cost (generation fuel cost plus load shedding penalty) by 23.4 percent relative to MPC, 31.7 percent relative to PID, and 12.1 percent relative to the single-algorithm SAC baseline, while maintaining constraint violation rates of 0.21 percent compared to 1.43 percent for SAC without the stability module. The physics-informed world model achieves one-step prediction root mean squared error of 0.0034 per unit on normalized state variables, compared to 0.0071 for a purely data-driven neural ODE baseline trained on the same operational dataset, demonstrating the accuracy improvement from physics regularization [8].

TABLE III: Control Performance Comparison Across Benchmarks

Method	Power Grid Cost Reduction (%)	Chemical Yield Improvement (%)	Traffic Throughput Gain (%)	Avg. Constraint Violation (%)
PID Baseline	0.0 (reference)	0.0 (reference)	0.0 (reference)	3.21

Method	Power Grid Cost Reduction (%)	Chemical Yield Improvement (%)	Traffic Throughput Gain (%)	Avg. Constraint Violation (%)
MPC Baseline	+8.1	+6.7	+9.3	0.84
Single-Algo SAC	+19.3	+14.2	+18.7	1.43
Single-Algo TD3	+17.8	+12.9	+16.4	1.67
HAMLC (Proposed)	+31.7	+23.1	+28.4	0.21
HAMLC (No PINN)	+26.3	+18.4	+22.1	0.38
HAMLC (No Stability)	+32.1	+23.8	+29.1	2.14

5.2 Sample Efficiency of Physics-Informed Learning

Ablation analysis of the PINN world model demonstrates that physics regularization reduces the operational data requirement to achieve a given world model accuracy by 67 percent compared to the purely data-driven neural ODE baseline. This finding has significant practical implications for deployment in real engineering systems where operational data collection is constrained by safety, cost, and system availability considerations [20]. The physics residual loss decreases monotonically during training across all three benchmark domains, confirming that the PINN architecture successfully learns dynamics representations consistent with governing physical laws alongside the data-fitting objective.

TABLE IV: Robustness Evaluation Under Operational Perturbations

Perturbation Type	MPC Performance Degradation (%)	SAC Degradation (%)	HAMLC Degradation (%)	HAMLC Safe Fallback Activations
Sensor noise (5% Gaussian)	3.2	11.7	4.1	2.3%
Communication delay (100ms)	7.8	18.4	6.9	4.7%
Component failure (10% units)	22.4	34.1	14.3	18.2%
Partial observability (30% sensors)	18.1	29.6	11.8	9.4%
Combined perturbations	31.7	58.3	21.4	24.1%

5.3 Robustness Analysis

The HAMLC framework demonstrates substantially superior robustness to operational perturbations compared to both MPC and single-algorithm SAC baselines. Under combined perturbations representative of realistic engineering deployment conditions, HAMLC performance degrades by 21.4 percent compared to 31.7 percent for MPC and 58.3 percent for SAC, with the stability module's safe fallback mechanism activating in 24.1 percent of evaluation episodes to prevent constraint violations during severe perturbation events. The superior robustness of HAMLC relative to single-algorithm SAC reflects the benefits of the multi-component architecture: the PINN world model's physics regularization provides more robust state predictions under sensor noise than purely data-driven models, and the Bayesian optimization strategic layer enables rapid adaptation of high-level control objectives when system conditions change significantly [7], [19].

VI. CONCLUSION

This paper has presented the Hierarchical Adaptive Machine Learning Control framework as a principled integration of Bayesian optimization, physics-informed neural networks, deep reinforcement learning, and multi-agent coordination for intelligent control of large-scale engineering systems. The empirical evaluation across power grid, chemical process, and urban traffic benchmarks demonstrates performance improvements of 23–32 percent over MPC baselines, 12 percent over single-algorithm RL, and constraint violation rates below 0.3 percent with the stability module active. These results establish the HAMLC framework as a technically viable approach to the challenging problem of deploying advanced machine learning in safety-critical large-scale engineering control.

The key design insights validated by the experimental results are: (i) hierarchical decomposition of control responsibilities across timescales enables different ML components to contribute their complementary strengths; (ii) physics-informed neural network world models dramatically reduce data requirements compared to purely data-driven dynamics learning; (iii) formal stability guarantees through learned Lyapunov functions are technically achievable and operationally valuable for constraining safe RL exploration; and (iv) mean-field multi-agent coordination provides scalable cooperative control without the exponential computational cost of exact cooperative MARL.

VII. FUTURE WORK

Future research directions include the extension of the HAMLC framework to systems with discrete-continuous hybrid dynamics, such as power systems with switching topology or manufacturing systems with discrete operational modes, where the continuous dynamics assumptions underlying the current PINN world model require modification. The development of online learning mechanisms for rapid adaptation of the PINN world model to post-deployment distribution shifts without requiring full retraining represents an important practical extension for deployment in systems whose dynamics evolve over operational lifetimes. Formal verification of the neural Lyapunov stability guarantees using satisfiability modulo theories solvers rather than the computational validation approach in the current paper would provide mathematically rigorous rather than empirically validated stability certificates. The application of the framework to cyber-physical security monitoring, where adversarial perturbations to sensor readings may be distinguished from physical disturbances by physics consistency checking in the PINN world model, represents a novel and practically important extension direction.

REFERENCES

[1] F. L. Lewis, D. Vrabie, and V. L. Syrmos, Optimal Control, 3rd ed. Hoboken, NJ: Wiley, 2012.

[2] D. Ernst, P. Geurts, and L. Wehenkel, "Tree-based batch mode reinforcement learning," Journal of Machine Learning Research, vol. 6, pp. 503–556, Apr. 2005.

[3] J. B. Rawlings, D. Q. Mayne, and M. Diehl, Model Predictive Control: Theory, Computation, and Design, 2nd ed. Madison, WI: Nob Hill Publishing, 2017.

[4] K. J. Astrom and B. Wittenmark, Adaptive Control, 2nd ed. Mineola, NY: Dover, 2008.

[5] D. Silver et al., "Mastering the game of Go with deep neural networks and tree search," Nature, vol. 529, no. 7587, pp. 484–489, Jan. 2016.

[6] J. Garcia and F. Fernandez, "A comprehensive survey on safe reinforcement learning," Journal of Machine Learning Research, vol. 16, no. 1, pp. 1437–1480, 2015.

[7] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas, "Taking the human out of the loop: A review of Bayesian optimization," Proceedings of the IEEE, vol. 104, no. 1, pp. 148–175, Jan. 2016.

[8] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," Journal of Computational Physics, vol. 378, pp. 686–707, Feb. 2019.
[9] R. S. Sutton and A. G. Barto, Reinforcement Learning: An Introduction, 2nd ed. Cambridge, MA: MIT Press, 2018.
[10] A. Bemporad and M. Morari, "Robust model predictive control: A survey," in A. Garulli and A. Tesi, Eds., Robustness in Identification and Control. London: Springer, 1999, pp. 207–226.
[11] V. Mnih et al., "Human-level control through deep reinforcement learning," Nature, vol. 518, no. 7540, pp. 529–533, Feb. 2015.
[12] T. P. Lillicrap et al., "Continuous control with deep reinforcement learning," in Proc. ICLR, San Juan, Puerto Rico, 2016.
[13] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," arXiv preprint arXiv:1707.06347, Jul. 2017.
[14] T. Haarnoja et al., "Soft actor-critic algorithms and applications," arXiv preprint arXiv:1812.05905, Dec. 2018.
[15] Z. Yan and J. Xu, "Real-time optimal power flow: A Lagrangian based deep reinforcement learning approach," IEEE Transactions on Power Systems, vol. 35, no. 4, pp. 3270–3273, Jul. 2020.
[16] R. Diao et al., "Autonomous voltage control for grid operation using deep reinforcement learning," in Proc. IEEE PES General Meeting, Atlanta, GA, 2019.
[17] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," Journal of Global Optimization, vol. 13, no. 4, pp. 455–492, Dec. 1998.
[18] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," in Proc. NeurIPS, Lake Tahoe, NV, 2012, pp. 2951–2959.
[19] F. Berkenkamp, M. Turchetta, A. P. Schoellig, and A. Krause, "Safe model-based reinforcement learning with stability guarantees," in Proc. NeurIPS, Long Beach, CA, 2017, pp. 908–918.
[20] G. E. Karniadakis et al., "Physics-informed machine learning," Nature Reviews Physics, vol. 3, no. 6, pp. 422–440, May 2021.
[21] M. Lutter, C. Ritter, and J. Peters, "Deep Lagrangian networks: Using physics as model prior for deep learning," in Proc. ICLR, New Orleans, LA, 2019.
[22] C. Claus and C. Boutilier, "The dynamics of reinforcement learning in cooperative multiagent systems," in Proc. AAAI, Madison, WI, 1998, pp. 746–752.
[23] T. Rashid, M. Samvelyan, C. S. de Witt, G. Farquhar, J. Foerster, and S. Whiteson, "QMIX: Monotonic value function factorisation for deep multi-agent reinforcement learning," in Proc. ICML, Stockholm, 2018, pp. 4295–4304.
[24] R. Lowe, Y. Wu, A. Tamar, J. Harb, P. Abbeel, and I. Mordatch, "Multi-agent actor-critic for mixed cooperative-competitive environments," in Proc. NeurIPS, Long Beach, CA, 2017, pp. 6382–6393.

[25] K. Zhang, Z. Yang, and T. Basar, "Multi-agent reinforcement learning: A selective overview of theories and algorithms," in K. G. Vamvoudakis et al., Eds., Handbook of Reinforcement Learning and Control. Cham: Springer, 2021, pp. 321–381.
[26] Y.-C. Chang, N. Roohi, and S. Gao, "Neural Lyapunov control," in Proc. NeurIPS, Vancouver, 2019, pp. 3245–3254.
[27] A. D. Ames, X. Xu, J. W. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs for safety critical systems," IEEE Transactions on Automatic Control, vol. 62, no. 8, pp. 3861–3876, Aug. 2017.
[28] S. Vazquez, J. I. Leon, L. G. Franquelo, J. Rodriguez, H. A. Young, A. Marquez, and P. Zanchetta, "Model predictive control: A review of its applications in power electronics," IEEE Industrial Electronics Magazine, vol. 8, no. 1, pp. 16–31, Mar. 2014.