



UNIVERSITY OF CALIFORNIA
MERCED

The Interplay Between Control Theory and Optimization

Control Theory For Machine Learning

Osama F. Abdel Aal

University of California, Merced
Department of Mechanical Engineering
Ph.D. in Mechanical Engineering

Nov, 2025



UNIVERSITY OF CALIFORNIA
MERCED

The Interplay Between Control Theory and Optimization

Control Theory For Machine Learning

Osama Fuad Abdel Aal

Student No. 401003037

Supervisor: Prof. YangQuan Chen
Professor in Mechanical Engineering, University of California, Merced

Committee Members: Dr. Ricardo de Castro
Professor of Mechanical Engineering, University of California, Merced

Dr. Francisco Danzi
Professor of Mechanical Engineering, University of California, Merced

Dr. Jairo Viola
Control engineer, Lam Research

University of California, Merced
Department of Mechanical Engineering
Ph.D. in Mechanical Engineering

Resume

Education

- 2022 – 2026** **Ph.D. Mechanical Engineering, University of California, Merced.**
Thesis title: The Interplay between Control Theory and Optimization Algorithms – Application to Machine Learning.
Description: Integer-order and fractional-order control theory for optimization and machine learning algorithms – analysis and design.
- 2014 – 2017** **M.Sc. Mechanical Engineering, University of Jordan.**
Thesis title: Nonlinear Vibration and Stability Analysis of a Beam with an Arbitrary Rise.
Description: The nonlinear in-plane resonance behavior of an imperfect Euler–Bernoulli beam on an elastic foundation was analyzed. Steady-state responses and their stability were studied using force–frequency curves obtained through the Harmonic Balance and Multiple Scales methods.
Grade: Excellent.
- 2006 – 2011** **B.Sc. Mechatronics Engineering, University of Jordan.**
Thesis title: Elevator Traffic Analysis and Simulation Software.
Awards: Awarded at the Fourth Engineering Creativity Forum, Jordan Engineers Association, October 2011.
Grade: Very Good (Ranked among top 5).

Experience

- 2022 – 2025 **Teaching Assistant**, Department of Mechanical Engineering, University of California, Merced.
- *Facilitated lab sessions and problem-solving sessions.*
 - *Supported instruction for core engineering courses including Circuits Theory (ENGR-065) and Mechanical Vibrations (ME140).*
- 2012 – 2022 **Lecturer and Lab Engineer**, Mechatronics Engineering Department, University of Jordan, Amman, Jordan.
- *Taught and prepared study materials for courses including Mechanical Vibrations, Dynamics, Robotics, Statics and Strength of Materials, Electrical Machines, MATLAB, Control Lab, Mechatronics Systems Design Lab, and Power Electronics.*
 - *Incorporated practical aspects into courses to reinforce theoretical learning.*
 - *Performed administrative duties including ABET accreditation work and documentation.*
 - *Designed and developed experiments and coordinated laboratory research tasks.*
- 2011 – 2012 **Lecturer and Lab Engineer**, Automotive Engineering Department, Al-Zarqa Private College, Al-Zarqa, Jordan.
- *Created course materials and lectures for Automotive Engineering students.*
 - *Developed tests and instructional materials to promote learning objectives.*
 - *Incorporated practical examples into coursework to enhance understanding.*
 - *Utilized diverse teaching methods, including lectures, presentations, and class activities.*

Publications

1. Al-Sharif, L., Abdel Aal, O. F., Abu Alqumsan, A. M., & Abuzayyad, M. A. (2015). The HARint Space: A Methodology for Compliant Elevator Traffic Designs. *Building Services Engineering Research and Technology*, 36(1), 34–50.
2. Al-Sharif, L., Abdel Aal, O. F., & Abu Alqumsan, A. M. (2013). Evaluating the Elevator Passenger Average Travelling Time Under Incoming Traffic Conditions Using Analytical Formulae and the Monte Carlo Method. *Elevator World*, 61(6), 110–123.
3. Al-Sharif, L., Abu Alqumsan, A. M., & Abdel Aal, O. F. (2013). Automated Optimal Design Methodology of Elevator Systems Using Rules and Graphical Methods (The HARint Plane). *Building Services Engineering Research and Technology*, 34(3), 275–293.
4. Aal, O. F. Abdel, Aal, M., & Al Qaisia, A. (2022). Nonlinear Dynamic Response of a Stiffened Imperfect Beam Under Primary Resonance Excitation. *Journal of Applied Nonlinear Dynamics*, 11(03), 667–702.
5. Hemeimat, R., & Aal, O. F. Abdel. (2024). Improving the Baggage Handling System (BHS) at Queen Alia International Airport by Using Rockwell's Arena

-
- Simulation Software. *American Journal of Operational Research*, 14(1), 1–11.
6. Aal, O. F. Abdel, Ozbek, N. S., & Chen, Y. (2025). Optimization Processes Designed with Controllable Convergence Performance - A Unified Review of Designs. Submitted to *IEEE/CAA Journal of Automatica Sinica (JAS)*.
 7. Aal, O. F. Abdel, Ozbek, N. S., Viola, J., & Chen, Y. (2025). Controlled Optimization with a Prescribed Finite-Time Convergence Using a Time Varying Feedback Gradient Flow. Under review in *Automatica*.
 8. Aal, O. F. Abdel, Ozbek, N. S., Viola, J., & Chen, Y. (2025). Data-Driven Control Design Using Bode's Ideal Cut-off Filter in the Virtual Reference Feedback Tuning Framework. Under review in *International Journal of Adaptive Control and Signal Processing*.
 9. Aal, O. F. Abdel, Ozbek, N. S., Viola, J., & Chen, Y. (2025). Beyond Nesterov: Dynamical Systems Perspective with Time-Dependent Inertia and Conformable Bregman Flow. Under review in *International Journal of Optimization and Control: Theories & Applications*.
 10. Khasawneh, H. J., Abdelaal, O., Al Saaideh, M. I., & Abo-Hammour, Z. S. (2019). Optimal Lead Compensator for Two-Loop Control System of Linear DC Motor. In *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)*, 634–639. IEEE.
 11. Aal, O. F. Abdel, Aal, M., & Al-Masaed, M. R. (2021). Robust Control of a Proportional Solenoid Valve. In *2021 9th International Conference on Control, Mechatronics and Automation (ICCMA)*, 72–77. IEEE.
 12. Aal, O. F. Abdel, Viola, J., & Chen, Y. (2023). Bode's Ideal Cut-off Based Virtual Reference Feedback Tuning Controller Design. In *2023 International Conference on Fractional Differentiation and Its Applications (ICFDA)*, 1–5. IEEE.
 13. Aal, O. F. Abdel, Viola, J., & Chen, Y. (2024). Fractional-Order Low-Frequency Compensators for Processes with Ramp-Like Disturbances. In *2024 20th IEEE/ASME International Conference on Mechatronic, Embedded Systems and Applications (MESA)*, 1–6. IEEE.
 14. Aal, O. F. Abdel, Viola, J., & Chen, Y. (2024). Fractional Order Euler-Lagrange Model for Accelerated Gradient Methods. *IFAC-PapersOnLine*, 58(12), 466–471. Elsevier.
 15. Aal, O. F. Abdel & Hemeimat, R. (2024). Data-Driven Based Decoupled Control Scheme for Interacting Quadruple Tank Process. In *2024 12th International Conference on Systems and Control (ICSC)*, 149–154. IEEE.
 16. Aal, O. F. Abdel, Ozbek, N., Chen, Y., Cao, S., & Chen, Y. (2025). 15 Ways to Apply Fractional Calculus in Gradient Descent Optimization Methods. Accepted in the *International Conference on Fractional Differentiation and Its Applications (ICFDA)*.
 17. Aal, O. F. Abdel, Ozbek, N., Viola, J., & Chen, Y. (2025). Finite Time Convergence of Continuous Time Accelerated Gradient Methods. *International Conference on Optimization and Applications (ICOA 2025)*.
 18. Aal, O. F. Abdel, Cao, S., & Chen, Y. (2025). A Digital Twin Behavioral Matching

- Framework via SPSA. Accepted and presented in *21st Int. Conference on Mechatronic, Embedded Systems and Applications (MESA), ASME/IEEE*.
19. Aal, O. F. Abdel, Ozbek, N., & Chen, Y. (2025). Enhanced Dynamic Control Strategy with Fractional-Order Sliding Mode Control for PV Power Plant Performance under Changing Operating Conditions. Accepted and presented in *21st MESA, ASME/IEEE*.
 20. Aal, O. F. Abdel, Ozbek, N., & Chen, Y. (2025). A Survey on Finite-, Fixed-, and Prescribed-Time Convergent Optimization Algorithms: Control-Theoretic Perspective. Accepted and presented in *21st MESA, ASME/IEEE*.
 21. Aal, O. F. Abdel, Ozbek, N., Viola, J., & Chen, Y. (2025). Proximal Dynamic Framework for Equilibrium Problems With Prescribed-Time Guarantees. Submitted to *2026 American Control Conference (ACC), IEEE*.
 22. Aal, O. F. Abdel, Viola, J., & Chen, Y. (2025). Optimal Control-Based Strategies for Gradient Descent Step Size Design. Submitted to *2026 ACC, IEEE*.

Abstract

Increasing integration of data-driven intelligence and automation has transformed the modern landscape of control and optimization, demanding frameworks that are simultaneously adaptive, interpretable, and robust. This thesis explores the interplay between control theory and optimization algorithms—two fields that, despite being historically distinct, share deep conceptual and methodological connections. By weaving together recent advances in machine learning, digital twins, and robust control analysis, the proposed research aims to establish principled approaches to both (i) learning for control, where optimization-based learning enhances the design and performance of controllers, and (ii) control for learning, where control-theoretic tools provide rigorous foundations for analyzing and improving optimization algorithms that underpin modern machine learning and artificial intelligence (AI).

The first dimension, Learning for Control, addresses the challenges of designing intelligent controllers in complex, uncertain, and dynamic environments. A central theme is the data-driven discovery of control laws, where system dynamics and feedback mechanisms are inferred directly from data rather than exclusively from first-principle models. Building on this foundation, the thesis investigates self-optimizing control (SOC) as a paradigm for smart control engineering. SOC leverages high performance real-time optimization algorithms—both gradient-based and derivative-free—to adjust low-level control parameters or setpoints dynamically, subject to periodic tasks and performance specifications. Complementary strategies such as run-to-run control and iterative learning control (ILC) will be explored as repetitive frameworks that naturally align with SOC objectives. Crucially, the role of digital twins is emphasized: by enabling real-time analytics, exhaustive scenario testing, and continuous system updates, digital twins provide the informational backbone for SOC. This creates a hierarchical structure where SOC serves as the upper layer of intelligence, translating insights from digital twins and industrial AI analytics into adaptive, robust, and “smart” control actions. The thesis contributes novel architectures and algorithms that close the loop between data, optimization, and control, ensuring responsiveness, reliability, and efficiency in next-generation engineering systems.

The second dimension, Control for Learning, reverses the perspective by leveraging control theory to advance the theory and practice of optimization—the computational engine at the heart of machine learning and AI. Optimization algorithms can be interpreted as dynamical systems, and their behavior can thus be analyzed and

shaped using control-theoretic principles. Key tools include Lyapunov-based methods to establish and certify convergence guarantees, as well as techniques for accelerating convergence through structured feedback design. Building on these insights, the thesis investigates the design of optimization algorithms with finite-time or fixed-time convergence, inspired by non-asymptotic control strategies. Dissipativity theory offers another promising lens, enabling the characterization of energy-like quantities in optimization dynamics to inform both stability analysis and novel algorithmic architectures. Furthermore, robust control perspectives will be applied to examine convergence under uncertainty, noise, and model misspecification, thereby enhancing the resilience of learning algorithms. This dual focus—on analysis and design—positions control theory not merely as a tool for understanding existing algorithms but as a systematic foundation for the discovery of new classes of optimization methods with provable and tunable properties.

Taken together, the proposed work highlights a bidirectional synthesis between control theory and optimization. On one side, learning and optimization empower controllers to adapt in real-time, guided by data and enabled by digital infrastructure such as digital twins. On the other hand, control-theoretic methodologies ground optimization algorithms in a rigorous dynamical framework, enabling the design of faster, more robust, and theoretically principled learning methods. This thesis argues that bridging these two perspectives is essential to advancing both fields: it offers a pathway toward smart, self-optimizing control systems in engineering practice, while simultaneously contributing to control-inspired optimization at the core of AI. The anticipated outcomes include new theoretical frameworks, algorithmic innovations, and practical case studies, ultimately advancing the vision of systems that are not only controlled intelligently but that also learn and optimize intelligently within their environments.

An additional dimension explored in this thesis is the integration of fractional calculus into the analysis and design of optimization algorithms. Fractional-order operators, by extending differentiation and integration to non-integer orders, capture memory and hereditary properties that conventional integer-order dynamics cannot. This thesis proposes novel fractional gradient descent schemes, where fractional derivatives replace or augment classical gradients, potentially reshaping optimization trajectories. Such approaches are expected to enhance performance in several respects: improving exploration in nonconvex landscapes by smoothing oscillations, accelerating convergence near critical points, and offering greater robustness against noise and irregular objective geometries. By embedding fractional-order dynamics into optimization, the work aims to provide a principled mechanism for balancing exploration and exploitation, thereby extending the frontier of control-inspired optimization algorithmic design.

Keywords: Learning for Control, Control Theory for Machine Learning, Optimization Algorithms, Fractional Calculus, Data-Driven Control, Self-Optimizing Control.

Contents

<i>List of Figures</i>	x
<i>List of Tables</i>	1
1 Foundations, Motivation, and Research Context	2
1.1 Motivation	3
1.2 From Control to Learning: Foundations Toward a System-Theoretic Framework for Algorithms	4
1.2.1 Context and Vision	4
1.2.2 Conceptual Foundations: <i>CodeMorphics</i> \Rightarrow <i>DynOptics</i> — from <i>Computational recipes</i> forms to evolving <i>dynamical flows</i> .	5
1.3 Historical Overview and Emerging Interplay: From Control Principles and Optimization to Renewed Interest in Large-Scale ML	7
1.4 Optimization as a Dynamical System: Bridging Control Theory and Machine Learning	9
1.4.1 Continuous-Time Optimization through the Lens of Dynamical Systems	14
1.4.2 The Control-Theoretic Architecture of Optimization: Feedback, Robustness, and Dissipativity	17
1.4.3 Passivity and Stability of Learning Algorithms	23
1.4.4 Adaptive Control Based Strategies	26
1.4.5 Optimal Control as Algorithm Design: Bridging Dynamics and Computation	28
1.4.6 Applications to Deep Learning	30
2 Learning for Control: From Data to Controllers with VRFT	32
2.1 Introduction	32
2.2 Reference model design for VRFT-based control	34
2.3 Bode’s Ideal Cut-Off Low-Pass Filter	36
2.4 VRFT Framework	38
2.5 Simulation and Experimental Results	39
2.5.1 Simulation-Based Illustration	39
2.5.2 Experiment-Based illustration	41

3	Beyond Nesterov: Dynamical Systems Perspective with Time-Dependent Inertia and Deformable Bregman Flows	48
3.1	Introduction	48
3.2	Preliminaries	51
3.2.1	The Bregman Divergence	51
3.2.2	Convexity	52
3.2.3	The Conformable Derivative	52
3.3	A Differential Equation for Modeling Nesterov’s Accelerated Gradient Method	53
3.4	A Modified Nestrov Model for Accelerated Gradient Methods (AGM): Toward Higher Rates	55
3.4.1	Convergence Rate: A Lyapunov Approach	56
3.4.2	Higher Convergence Rates by Design	57
3.5	Time-Weighted Derivatives: A Variational Model of Accelerated Methods in Optimization	58
3.5.1	Conformable Bregman Flows	59
3.5.2	Convergence Rate	61
3.6	Results and Discussion	62
3.7	Conclusion	63
4	Controlled Optimization Processes with Convergence by Design: A Unified Design	66
4.1	Introduction	66
4.2	Preliminaries and Fundamental Concepts on Convergence Properties	68
4.2.1	Lyapunov Foundations and Definitions of Convergence Algorithms	70
4.3	Breaking the Asymptotic Barrier: A Dynamical systems approach	72
4.4	Exponential Convergence of Classical Gradient Flows	74
4.4.1	Optimization Frameworks with Convergence by Design- A Unified Approach	75
4.4.2	A Taxonomy of Time-Constrained Optimization Flows	80
4.5	Convergence properties of distributed optimization algorithms	82
4.5.1	Finite-Time Distributed Algorithms	85
4.5.2	Fixed-Time Distributed Algorithms	86
4.5.3	Prescribed-Time Distributed Algorithms	86
4.6	Controllable convergence algorithms for neurodynamic optimization	87
4.7	Challenges and Opportunities	89
4.7.1	Discretization	90
4.7.2	Robustness	91
4.7.3	Machine Learning Integration	92
4.8	Concluding Remarks	92

5	Controlled Optimization Processes with Convergence by Design- A control Approach	93
5.1	Finite-Time Convergence of Continuous-Time Accelerated Gradient Methods	94
5.1.1	On the asymptotic convergence of Accelerated Gradient Methods	94
5.1.2	Finite-Time Convergence with Accelerated Gradient Methods . .	96
5.1.3	Numerical illustration	100
5.2	Controlled Optimization Process with a Prescribed Finite-Time Convergence Using a Time-Varying Feedback Gradient Flow	100
5.2.1	Prescribed-Time Convergence of Unconstrained Optimization . .	100
5.2.2	Prescribed-time Convergent Newton’s Method via Time-varying Feedback	105
5.2.3	Prescribed-time Convergence of Primal-Dual Gradient Dynamics (PDGD): Constrained Optimization	106
5.2.4	Numerical examples	108
5.2.5	Concluding Remarks	110
6	Optimal Control–Based Strategies for Gradient Descent Step Size Design	112
6.1	Introduction	112
6.2	Problem Formulation	113
6.3	Optimal Control Meets Optimization: Deriving Gradient Descent via PMP	114
6.4	Optimal Learning Rate via Discrete Pontryagin Principle	116
6.5	Numerical Illustration	118
6.6	Concluding Remarks	120
7	Memory-Driven Learning: Fractional Dynamics in Gradient Descent	121
7.1	Introduction	121
7.2	Preliminaries	123
7.3	Gradient Descent Algorithms	124
7.4	Exploring 15 Fractional Calculus Techniques for Gradient Descent Optimization	125
7.4.1	Fractional Order Updating Law	126
7.4.2	Fractional Order GD Learning Coefficient	127
7.4.3	Fractional Order Gradient	128
7.4.4	Fractional Order Noise in Gradient	129
7.5	Concluding Remarks	130
	<i>Bibliography</i>	132

List of Figures

1.1	From Implementation to Evolution: The Dual Perspectives of Optimization.	6
1.2	Breaking the Asymptotic Barrier of Gradient Flows: A Dynamical Systems Approach (Ozaslan et al., 2024).	17
1.3	The block diagram of the primal-dual optimization-PI dynamics.	26
2.1	The Framework of the Virtual Reference Feedback Tuning Method.	37
2.2	The Frequency Response of the BICO Low Pass Filter.	37
2.3	Results for Example 1.	43
2.4	Comparison between the closed-loop performances based on maximum sensitivity M_r .	43
2.5	<i>ECP Model 210</i> Mass Spring Damper Plant.	44
2.6	Design of Controller of the 2 degrees of Freedom Motion System.	44
2.7	Experimental Setup and Control Scheme For the Rotary Flexible Robotic Joint.	44
2.8	Design of Controller of the Rotary Flexible Joint Robotix System.	45
3.1	Comparison of the classical Nesterov accelerated gradient (NAG) method and the proposed approach with a designed rate of $O(1/t^3)$ on the Sum of Squares function.	63
3.2	Comparison of the classical Nesterov accelerated gradient (NAG) method and the proposed approach with a designed rate of $O(1/t^3)$ on the Trid function.	63
3.3	Comparison of the classical Nesterov accelerated gradient (NAG) method and the proposed approach with a designed rate of $O(e^{-0.5t})$ on the Trid function.	64
3.4	Comparison of the classical Nesterov accelerated gradient (NAG) method and the proposed approach with a designed rate of $O(1/t^4)$ on the 10–dimensional nonconvex Rosenbrock function.	64
3.5	Comparison of the classical Nesterov accelerated gradient (NAG) method and the proposed approach with a designed rate of $O(1/t^4)$ on the 10–dimensional nonconvex Rastrigin function.	64

3.6	Effect of the fractional order ν on convergence behavior compared with classical Nesterov's method on the ill-conditioned 10-dimensional Ellipsoid function. Smaller ν values yield less oscillation and greater stability in the optimization trajectory.	65
4.1	Quantitative illustration of finite-time stability.	72
4.2	Breaking the Asymptotic Barrier: A Dynamical Systems Approach (Ozaslan et al., 2024).	76
5.1	Convergence of normed states of the Sphere function for different initial conditions.	101
5.2	Convergence of normed states of the Trid function for different initial conditions.	101
5.3	Convergence of the normed weights in the continuous-time SVM problem.	102
5.4	Continuous optimization as a Lur'e system with time-varying feedback	102
5.5	Prescribes-time Convergence of the Trid Function.	109
5.6	Comparison between the convergence of the three GFs for the Rosenbrock function starting from various initial conditions, $q_1 = 1.8, q_2 = 2.2, \eta_1 = \eta_2 = 1, T_{p1} = 2, T_{p2} = 4, T_{p3} = 6$	109
5.7	Comparison between the convergence of the three GFs for solving the PDGD for various values of $\kappa, q_1 = 1.8, q_2 = 2.2, \eta_1 = \eta_2 = 1, T_{p1} = 0.5, T_{p2} = 1$	110
5.8	Comparison between the convergence of the three GFs for solving the SVM problem, $q_1 = 1.8, q_2 = 2.2, \eta_1 = \eta_2 = 1, T_p = 5$	110
6.1	Convergence performance on the ill-conditioned Sum of Squares function: PMP-based learning rate achieves the fastest descent.	118
6.2	Convergence performance on the Trid function: PMP-based learning rate achieves the fastest descent.	119
6.3	Convergence performance on the Ackley function: PMP-based learning rate achieves the fastest descent.	119
6.4	Convergence performance on the Rastrigin function: PMP-based learning rate achieves the fastest descent.	119
7.1	The integrator model, embedded in $F(z)$, ensures zero steady-state tracking error for constant reference inputs by applying the internal model principle (IMP). $[F(z)]^{(\alpha_F)}, \alpha_F \in (0, 1)$ can be designed as a FO-Transfer function.	126
7.2	The four entry points of fractional calculus in GD algorithm.	126

List of Tables

1.1	Comparison between existing optimizers in the continuous-time perspective.	15
1.2	Momentum and Gradient Methods as Dynamical Filters.	21
1.3	Unified system representations of optimization algorithms in discrete-time dynamical form.	27
2.1	Comparison of Design Methods Based on IAE and ITAE Criteria.	46
2.2	<i>IAE</i> and <i>TV</i> comparison between various design methods.	46
2.3	Performance evaluation comparison between various design methods of the 2DoF mass-spring-damper system.	47
2.4	Performance evaluation of the obtained closed-loop responses using various reference models.	47
4.1	Taxonomy of Optimization Algorithms with Controlled Convergence Properties.	83
4.2	Summary of Predefined-Time Convergence (PrTC) Mechanisms, Lyapunov Conditions, and Supporting Propositions	84
4.3	Taxonomy of Distributed Optimization Algorithms with Their Controlled Convergence Properties.	88
4.4	Taxonomy Based on Existing Predefined-Time Mechanisms.	89
4.5	Comparison of Selected Neurodynamic Optimization Approaches.	89
7.1	The 15 variants of fractional gradient descent with indexed equations and references.	130