

Section 1 Introduction

December 25, 2025

1 Introduction

1.1 Motivation and Background

In December 2023, Boston Dynamics' Atlas humanoid robot demonstrated unprecedented balance recovery during a push test, stabilizing a double-inverted-pendulum-like configuration (torso + articulated legs) within 0.8 seconds using advanced model-based control. This real-world demonstration highlights the critical need for fast, robust control of inherently unstable multi-link systems—a challenge that has motivated decades of research on the double-inverted pendulum (DIP) as a canonical testbed for control algorithm development.

The DIP control problem has direct applications across multiple domains:

- Humanoid Robotics: Torso-leg balance for Atlas, ASIMO, and bipedal walkers requiring multi-link stabilization
- Aerospace: Rocket landing stabilization (SpaceX Falcon 9 gimbal control resembles inverted pendulum dynamics)
- Rehabilitation Robotics: Exoskeleton balance assistance for mobility-impaired patients with real-time stability requirements
- Industrial Automation: Overhead crane anti-sway control with double-pendulum payload dynamics

These applications share critical characteristics with DIP: inherent instability, underactuation (fewer actuators than degrees of freedom), nonlinear dynamics, and stringent real-time performance requirements (sub-second response). The DIP system exhibits these same properties, making it an ideal testbed for evaluating sliding mode control (SMC) techniques, which promise robust performance despite model uncertainties and external disturbances.

Sliding mode control (SMC) has evolved over nearly five decades from Utkin's pioneering work on variable structure systems in 1977 [ref1] through three distinct eras: (1) **Classical SMC** (1977-1995): Discontinuous switching with boundary layers for chattering reduction [1-6], (2) Higher-Order SMC (1996-2010): Super-twisting and second-order algorithms achieving continuous control action [12-19], and (3) Adaptive/Hybrid SMC (2011-present): Parameter adaptation and mode-switching architectures combining benefits of multiple approaches [20-31]. Despite these advances, comprehensive comparative evaluations across multiple SMC variants remain scarce in the literature, with most studies evaluating 1-2 controllers in isolation rather than providing systematic multi-controller comparisons enabling evidence-based selection.

1.2 Literature Review and Research Gap

Classical Sliding Mode Control: First-order SMC [1,6] establishes theoretical foundations with reaching phase and sliding phase analysis. Boundary layer approaches [2,3] reduce chattering at the cost of approximate sliding. Recent work [45,46] demonstrates practical implementation on inverted pendulum systems but focuses on single controller evaluation.

Higher-Order Sliding Mode: Super-twisting algorithms [12,13] and second-order SMC [17,19] achieve continuous control action through integral sliding surfaces, eliminating chattering theoretically. Finite-time convergence proofs [14,58] provide stronger guarantees than asymptotic stability. However, computational complexity and gain tuning challenges limit adoption.

Adaptive SMC: Parameter adaptation laws [22,23] address model uncertainty through online estimation. Composite Lyapunov functions [ref24] prove stability of adaptive schemes. Applications to inverted pendulums [45,48] show improved robustness but at computational cost.

Hybrid and Multi-Mode Control: Switching control architectures [30,31] combine multiple controllers for different operating regimes. Swing-up and stabilization [ref46] require multiple Lyapunov functions for global stability. Recent hybrid adaptive **STA-SMC** [ref20] claims combined benefits but lacks rigorous comparison.

Optimization for SMC: Particle swarm optimization (PSO) [ref37] and genetic algorithms [ref67] enable automatic gain tuning. However, most studies optimize for single scenarios, ignoring generalization to diverse operating conditions.

Table 1.1: Literature Survey of SMC for Inverted Pendulum Systems (2015-2025)

[TABLE - See Markdown version for details]

Summary Statistics from Survey of 50+ Papers (2015-2025): - Average controllers per study: 1.8 (range: 1-3; only 4percent evaluate 3+ controllers) - Average metrics evaluated: 3.2 (range: 2-5; 85percent focus on settling time/overshoot only) - Studies with optimization: 15percent (3/20 in table; mostly single-scenario PSO) - Studies with robustness analysis: 25percent (5/20; typically ± 10 percent uncertainty only) - Studies with hardware validation: 10percent (2/20; majority simulation-only)

Research Gaps (Quantified):

- Limited Comparative Analysis: Of 50 surveyed papers (2015-2025), 68percent evaluate single controllers, 28percent compare 2 controllers, and only 4percent evaluate 3+ controllers (Table 1.1). No prior work systematically compares 7 SMC variants (Classical, STA, Adaptive, Hybrid, Swing-Up, MPC, combinations) on a unified platform with identical scenarios and metrics—a critical gap for evidence-based controller selection.

- Incomplete Performance Metrics: Survey analysis reveals 85percent of papers evaluate only transient response (settling time, overshoot), while computational efficiency (real-time feasibility) is reported in 12percent, chattering characteristics in 18percent, energy consumption in 8percent, and robustness analysis in 25percent. Multi-dimensional evaluation across 10+ metrics spanning computational, transient, chattering, energy, and robustness categories remains absent from the literature.

- Narrow Operating Conditions: 92percent of surveyed studies evaluate controllers under small perturbations (± 0.05 rad), with only 8percent testing realistic disturbances (± 0.3 rad) or model uncertainty (± 20 percent parameter variation). This narrow scope fails to validate robustness claims—a critical concern for real-world deployment where larger disturbances are common.

- Optimization Limitations: Among the 15percent of papers using PSO/GA optimization, 100percent optimize for single nominal scenarios without validating generalization to diverse perturbations or disturbances. This severe limitation manifests as 50.4x performance degradation when single-scenario-optimized gains face realistic conditions (Section 8.3)—a previously unreported failure mode.

- Missing Validation: While 45percent of papers present Lyapunov stability proofs, only 10percent validate theoretical convergence rates against experimental data. The disconnect between theory (asymptotic/finite-time guarantees) and practice (measured settling times, chattering) limits confidence in theoretical predictions and necessitates rigorous experimental validation of stability claims.

1.3 Contributions

This paper addresses these gaps through:

- Comprehensive Comparative Analysis: First systematic evaluation of 7 SMC variants (Classical, STA, Adaptive, Hybrid, Swing-Up, MPC, combinations) on a unified DIP platform with 400+ Monte Carlo simulations across 4 operating scenarios (Section 6.3), revealing **STA-SMC** achieves 91percent chattering reduction and 16percent faster settling (1.82s vs 2.15s) compared to **Classical SMC** (Section 7).

- Multi-Dimensional Performance Assessment: First 12-metric evaluation spanning 5 categories—computational (compute time, memory), transient (settling, overshoot, rise time), chattering (index, frequency, HF energy), energy (total, peak power), robustness (uncertainty tolerance, disturbance rejection)—with 95percent confidence intervals via bootstrap validation (10,000 resamples) and statistical significance testing (Welch’s t-test, alpha=0.05, Bonferroni correction) across all comparisons (Section 6.2, Section 7).

- Rigorous Theoretical Foundation: Four complete Lyapunov stability proofs (Theorems 4.1-4.4) establishing convergence guarantees—asymptotic (Classical, Adaptive), finite-time (STA with explicit time bound $T \geq 2.1s$ for typical initial conditions), and ISS (Hybrid)—experimentally validated with 96.2percent agreement on Lyapunov derivative negativity (Section 4.5).

- Experimental Validation at Scale: 400-500 Monte Carlo simulations per scenario (1,300+ total trials) with rigorous statistical methods—Welch’s t-test (alpha=0.05), Bonferroni correction (family-wise error control), Cohen’s d effect sizes ($d=2.14$ for STA vs Classical settling time, indicating large practical significance), and bootstrap 95percent CI with 10,000 resamples ensuring robust statistical inference (Section 6.4).

- Critical PSO Optimization Analysis: First demonstration of severe PSO generalization failure—50.4x chattering degradation (2.14 ± 0.13 nominal -vs- 107.61 ± 5.48 realistic) and 90.2percent instability rate when single-scenario-optimized gains face realistic disturbances—and robust multi-scenario PSO solution achieving 7.5x improvement (144.59x -vs- 19.28x degradation) across 15 diverse scenarios (3 nominal, 4 moderate, 8 large perturbations) with worst-case penalty (alpha=0.3) ensuring conservative design (Section 8.3).

- Evidence-Based Design Guidelines: Application-specific controller selection matrix (Table 9.1) validated across 1,300+ simulations—**Classical SMC** for embedded systems (18.5 mus compute, 4.8x faster than Hybrid), **STA-SMC** for performance-critical applications (1.82s settling, 91percent chattering reduction, 11.8J energy), Hybrid STA for robustness-critical systems (16percent uncertainty tolerance, highest among all controllers)—enabling systematic controller selection based on quantified performance-robustness tradeoffs (Section 9.1).

- Open-Source Reproducible Platform: Complete Python implementation (3,000+ lines, 100+ unit tests, 95percent coverage) with benchmarking scripts, PSO optimization CLI, HIL integration, and FAIR-compliant data release (seed=42, version pinning, Docker containerization) enabling full reproducibility of all 1,300+ simulation results and facilitating community extensions (GitHub: [REPO LINK]).

1.4 Why Double-Inverted Pendulum?

The double-inverted pendulum (DIP) serves as an ideal testbed for SMC algorithm evaluation due to five critical properties that distinguish it from simpler benchmarks:

- Sufficient Complexity, Bounded Scope
- vs. Single Pendulum: DIP adds coupled nonlinear dynamics (inertia matrix coupling M, M, M; Coriolis forces $\theta\dot{\theta}$) absent in single pendulum, requiring multi-variable sliding surfaces ($\sigma = \lambda\dot{\theta} + \lambda\ddot{\theta} + k\theta + k\dot{\theta}$) and coordinated gain tuning across 4-6 parameters.
- vs. Triple/Quad Pendulum: DIP maintains analytical tractability for Lyapunov analysis (3x3 inertia matrix, 6D state space) while exhibiting representative underactuated challenges. Triple pendulums suffer from explosive state space (9D), 6x6 inertia matrices, and prohibitive computational cost limiting rigorous theoretical treatment.

- Underactuation with Practical Relevance
- 1 actuator, 3 DOF (cart + 2 pendulums): Directly matches humanoid torso-leg systems (1 hip actuator controlling 2-link leg dynamics during single-support phase) and crane anti-sway (1 trolley motor controlling double-pendulum payload from hook + load).
- Balanced difficulty: Single pendulum (1 actuator, 1 DOF) is fully actuated after feedback linearization; higher-order pendulums become impractical for systematic comparison (computational cost scales as $O(n^3)$ for n-pendulum systems).

- Rich Nonlinear Dynamics Stress-Testing Robustness
- Inertia matrix M(q): Configuration-dependent with 12 coupling terms (6 unique due to symmetry), varying by 40-60percent across workspace
- Coriolis matrix C(q,q): Velocity-dependent with centrifugal ($\theta\dot{\theta}^2$) and Coriolis ($\theta\dot{\theta}\dot{\theta}$) terms
- Gravity vector G(q): Strongly nonlinear ($\sin\theta, \sin\theta\dot{\theta}$) with unstable equilibrium requiring active stabilization
- Friction: Asymmetric viscous + Coulomb friction introducing model uncertainty ($\pm 15\%$ typical variation)

These terms stress-test SMC robustness to: (a) parametric uncertainty ($\pm 20\%$ in masses, lengths, inertias), (b) unmodeled dynamics (friction, flexibility), and (c) external disturbances (step, impulse, sinusoidal 0.5-5 Hz).

- Established Literature Benchmark
- 50+ papers (2015-2025) use DIP for SMC evaluation (Table 1.1), enabling direct comparison with prior art and validation of claimed improvements against standardized baseline.
- Standardized initial conditions: ± 0.05 rad (nominal), ± 0.3 rad (realistic) facilitate reproducibility and inter-study comparison.
- Commercial hardware availability: Quanser QUBE-Servo 2, Googol Tech GI03 enable sim-to-real validation (our MT-8 HIL experiments, Section 8.2, Enhancement num3).

- Transferability to Complex Systems

Control insights from DIP generalize to diverse applications:

- Humanoid robots: Balance recovery (Atlas, ASIMO), walking stabilization (bipedal dynamics = DIP during single-support), push recovery
- Aerospace: Multi-stage rocket attitude control (Falcon 9 landing), satellite attitude with flexible appendages
- Industrial: Overhead cranes (double-pendulum payload from hook + load), rotary cranes with boom + payload dynamics
- Rehabilitation: Powered exoskeletons (hip-knee-ankle control = triple pendulum; DIP provides foundational analysis), balance assistance for mobility-impaired patients

The DIP benchmark thus balances theoretical tractability (enabling rigorous Lyapunov proofs), practical relevance (matching real-world underactuated systems), and community standardization (facilitating reproducibility and comparison)—justifying its selection for this comprehensive comparative study over simpler (single pendulum) or more complex (triple+ pendulum) alternatives.

1.5 Paper Organization

The remainder of this paper is organized as follows:

- Section 2: System model (6D state space, full nonlinear Euler-Lagrange dynamics with inertia matrix $M(q)$, Coriolis $C(q,q)$, gravity $G(q)$) and control objectives (5 formal requirements: asymptotic stability, settling time $\leq 3s$, overshoot $\leq 10\%$, control bounds $|u| \leq 100N$, real-time feasibility $\leq 100\mu s$)
 - Section 3: Controller design for all 7 SMC variants with explicit control law formulations—Classical (boundary layer + saturation, 6 gains), STA (continuous 2nd-order, 6 gains), Adaptive (time-varying gain $K(t)$, 5 parameters), Hybrid Adaptive STA (mode-switching, 4 gains), Swing-Up (energy-based 2-phase), MPC (finite-horizon optimization), and combinations
 - Section 4: Lyapunov stability analysis with 4 complete convergence proofs (Theorems 4.1-4.4) establishing asymptotic stability (Classical, Adaptive), finite-time convergence with explicit time bound (STA, $T \leq 2.1s$), and input-to-state stability (Hybrid)—experimentally validated via Lyapunov derivative monitoring (96.2percent V ≤ 0 confirmation)
 - Section 5: PSO optimization methodology including multi-objective fitness function (4 components: ISE, control effort, slew rate, sliding surface variance), search space design (controller-specific bounds), PSO hyperparameters (40 particles, 200 iterations, $w=0.7$, $c_1=c_2=2.0$), and robust multi-scenario approach (15 scenarios spanning ± 0.05 to ± 0.3 rad perturbations) addressing generalization failure
 - Section 6: Experimental setup detailing Python simulation platform (RK45 adaptive integration, $dt=0.01s$, 100 Hz control loop), 12 performance metrics across 5 categories (computational, transient, chattering, energy, robustness), 4 benchmarking scenarios (nominal ± 0.05 rad, realistic ± 0.3 rad, model uncertainty $\pm 20\%$, disturbances), and statistical validation methods (Welch's t-test, bootstrap 95percent CI with 10,000 resamples, Cohen's d effect sizes)
 - Section 7: Performance comparison results presenting computational efficiency (Classical 18.5 μs fastest, all $\leq 50\mu s$ real-time budget), transient response (STA 1.82s settling best, 16percent improvement), chattering analysis (STA 2.1 index, 91percent reduction vs Classical 8.2), and energy consumption (STA 11.8J optimal)—establishing **STA-SMC** performance dominance and **Classical SMC** computational advantage
 - Section 8: Robustness analysis evaluating model uncertainty tolerance (Hybrid 16percent best, default gains 0percent convergence requiring PSO tuning), disturbance rejection (STA 91percent sinusoidal attenuation, 0.64s impulse recovery), PSO generalization failure (50.4x degradation, 90.2percent instability), and robust PSO solution (7.5x improvement, 94percent degradation reduction)—revealing critical optimization limitations
 - Section 9: Discussion of performance-robustness tradeoffs (3-way analysis: speed vs performance vs robustness), controller selection guidelines (5 decision matrices for embedded/performance/robustness-critical/balanced/research applications), Pareto optimality (STA and Hybrid dominate; Adaptive non-Pareto), and theoretical vs experimental validation (96.2percent Lyapunov agreement, convergence ordering matches theory)
 - Section 10: Conclusions summarizing 6 key findings (STA dominance, robustness tradeoffs, PSO failure, theory validation), 4 practical recommendations (controller selection, PSO mandatory with multi-scenario, real-time feasibility, actuator choice), and 8 future research directions (3 high-priority: robust PSO extensions, complete LT-6 uncertainty analysis, non-SMC benchmarks; 3 medium: data-driven hybrids, higher-order systems; 2 long-term: industrial case studies)
-