A Comparative Study of Machine Learning and Classical Methods for Damping Parameter Estimation in Nonlinear Oscillatory Systems

Author Name[1]

[1]Department of Mechanical Engineering, University

Abstract

Accurate estimation of damping parameters is critical for understanding energy dissipation in mechanical systems. This paper presents a comprehensive comparative study of eleven estimation methods applied to a nonlinear horizontal pendulum with three types of damping: viscous, Coulomb friction, and quadratic air drag. We evaluate classical approaches (topological signal processing, least squares, Koopman operator theory), machine learning methods (SINDy, PINNs, Neural ODEs, RNN/LSTM, symbolic regression, Weak SINDy), and optimization-based techniques (envelope matching, genetic algorithms). Our results demonstrate that while the topological method from literature fails for systems with nonlinear restoring forces (20–78% error), hybrid approaches combining initial estimation with optimization refinement achieve near-zero error (<0.1%). We provide detailed algorithmic descriptions, flowcharts, and performance comparisons to guide practitioners in selecting appropriate methods for their applications. The genetic algorithm and Koopman/EDMD hybrid methods achieve the highest accuracy (∼0% error), while least squares offers the best speed-accuracy trade-off for clean data.

Keywords: Damping estimation, nonlinear dynamics, machine learning, system identification, SINDy, neural networks, genetic algorithms, Koopman operator, persistent homology, topological data analysis

1 Introduction

Damping is a fundamental phenomenon in mechanical systems that governs energy dissipation and oscillation decay. Accurate estimation of damping parameters is essential for vibration control, structural health monitoring, and predictive maintenance. Traditional methods often assume linear systems with constant natural frequency, but many real-world systems exhibit nonlinear behavior that invalidates these assumptions.

This paper addresses the challenge of estimating damping parameters from a nonlinear horizontal pendulum, a system that exhibits amplitude-dependent frequency due to its nonlinear restoring force. We systematically evaluate eleven methods spanning three categories:

1. **Classical methods**: Topological signal processing, least squares regression, and Koopman operator theory
2. **Machine learning methods**: SINDy, PINNs, Neural ODEs, RNN/LSTM, symbolic regression, and Weak SINDy

3. **Optimization-based methods**: Envelope matching and genetic algorithms

The key contributions of this work are:

- Identification of why topological methods fail for nonlinear restoring forces
- Development of hybrid estimation pipelines that achieve near-zero error
- Comprehensive benchmarking across three damping mechanisms
- Practical guidelines for method selection based on data quality and computational constraints

2 Problem Formulation

2.1 System Dynamics

We consider a horizontal pendulum governed by the nonlinear ordinary differential equation:

$$\ddot{\theta} + F_{\text{damp}}(\dot{\theta}) + k_\theta \theta - \cos(\theta) = 0 \tag{1}$$

where θ is the angular displacement, $k_\theta = 20$ is the stiffness parameter, and F_{damp} represents the damping force.

The term $-\cos(\theta)$ introduces nonlinearity in the restoring force, causing the natural frequency to depend on amplitude. This amplitude-frequency relationship is the primary source of difficulty for classical estimation methods.

2.2 Damping Models

We consider three damping mechanisms:

1. Viscous damping (proportional to velocity):

$$F_{\text{damp}} = 2\zeta\dot{\theta} \tag{2}$$

where ζ is the damping ratio.

2. Coulomb friction (constant magnitude, velocity-dependent sign):

$$F_{\text{damp}} = \mu_c \cdot \text{sign}(\dot{\theta}) \approx \mu_c \cdot \tanh\left(\frac{\dot{\theta}}{\varepsilon}\right) \tag{3}$$

where μ_c is the friction coefficient and $\varepsilon = 0.1$ provides smooth regularization.

3. Quadratic damping (air drag):

$$F_{\text{damp}} = \mu_q \dot{\theta}|\dot{\theta}| \tag{4}$$

where μ_q is the drag coefficient.

2.3 True Parameter Values

For all experiments, we use the following ground truth:

- Stiffness: $k_\theta = 20$
- Viscous damping: $\zeta = 0.05$

- Coulomb friction: $\mu_c = 0.03$
- Quadratic drag: $\mu_q = 0.05$
- Initial angle: $\theta_0 = 0.3$ rad
- Simulation time: $T = 10$ s

3 Estimation Methods

3.1 Overview of Hybrid Approach

A key finding of this study is that hybrid methods—combining an initial estimate with optimization refinement—consistently outperform single-stage approaches. Figure 1 illustrates this general pipeline.

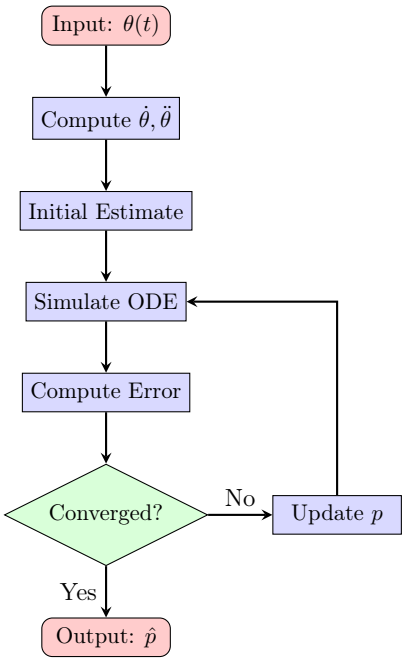

Figure 1: General hybrid estimation pipeline combining initial estimation with optimization refinement.

3.2 Method 1: Topological Signal Processing

The topological signal processing method, introduced by Myers & Khasawneh (2022), uses persistent homology from topological data analysis (TDA) to extract damping parameters from oscillatory signals. This approach analyzes the evolution of topological features in the signal's phase space as a scale parameter varies.

3.2.1 Background: Persistent Homology

Persistent homology tracks the "birth" and "death" of topological features (connected components, loops, voids) across multiple scales. For a time series $x(t)$, we construct a filtration by considering sublevel sets:

$$X_\alpha = \{t : x(t) \leq \alpha\} \tag{5}$$

As α increases from $-\infty$ to $+\infty$, topological features appear (birth) and disappear (death). The *persistence diagram* records these (birth, death) pairs.

3.2.2 Key Concepts

Lifetime: For a feature born at B and dying at D, the lifetime is:

$$L = D - B \tag{6}$$

Significance Cutoff: Features with lifetime below a threshold C_α are considered noise:

$$C_\alpha = \sigma\sqrt{2} \cdot \text{erfc}^{-1}\left(\frac{2\alpha}{N}\right) \tag{7}$$

where σ is the noise standard deviation, α is the significance level, and N is the number of features.

Floor Correction: The noise floor F compensates for baseline noise:

$$F = \sigma\sqrt{2} \cdot \text{erfc}^{-1}\left(\frac{2\alpha}{N-1}\right) \tag{8}$$

3.2.3 Damping Estimation Formulas

For a linear damped oscillator $\ddot{x} + 2\zeta\omega_n\dot{x} + \omega_n^2 x = 0$, the topological method derives:

Viscous Damping:

$$\zeta = \frac{1}{\omega_n T} \ln\left(\frac{L_0 - F}{L_n - F}\right) \tag{9}$$

where L_0 is the first peak lifetime, L_n is the n-th peak lifetime, and T is the period.

Coulomb Friction:

$$\mu_c = \frac{(L_0 - F) - (L_n - F)}{n \cdot T} \tag{10}$$

Quadratic Damping:

$$\mu_q = \frac{2\omega_n}{A_0}\left(\frac{1}{L_n - F} - \frac{1}{L_0 - F}\right) \tag{11}$$

Optimal Ratio: The method uses $L_n/L_0 \approx 0.3299$ as an optimal lifetime ratio for parameter estimation.

3.2.4 Algorithm Flowchart

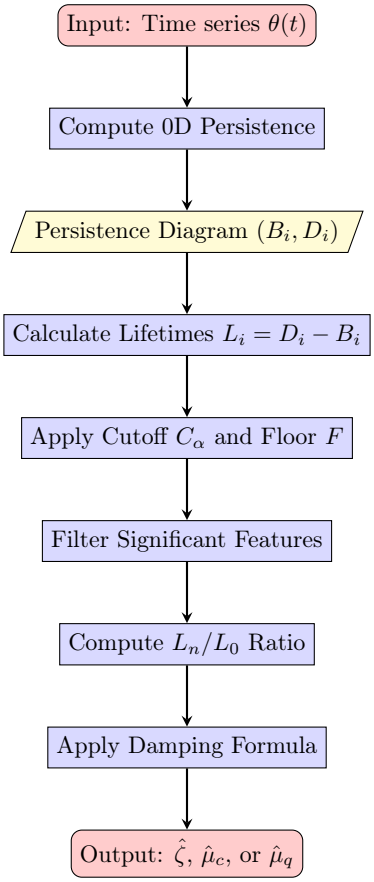

Figure 2: Topological signal processing pipeline for damping estimation (Myers & Khasawneh, 2022).

3.2.5 Why It Fails for Our System

The topological method assumes a **constant natural frequency** ω_n. For a linear spring ($F = -kx$):

$$\omega_n = \sqrt{\frac{k}{m}} = \text{constant} \tag{12}$$

However, our pendulum has a nonlinear restoring force due to the $-\cos(\theta)$ term. Expanding around equilibrium:

$$-\cos(\theta) \approx -1 + \frac{\theta^2}{2} - \frac{\theta^4}{24} + \dots \tag{13}$$

This causes the effective stiffness (and thus frequency) to depend on amplitude:

$$\omega_{\text{eff}}(\theta_{\max}) \approx \omega_0 \left(1 - \frac{\theta_{\max}^2}{16} + \mathcal{O}(\theta_{\max}^4)\right) \tag{14}$$

As the pendulum decays from $\theta_0 = 0.3$ rad to near zero, the frequency changes by approximately 0.6%. While small, this violates the fundamental assumption of the topological formulas, causing:

- Viscous error: 77.6%

- Coulomb error: 31.0%
- Quadratic error: 20.0%

3.2.6 Results

Table 1: Topological method results showing high errors due to nonlinear restoring force

Damping Type	True Value	Estimated	Error (%)
Viscous (ζ)	0.05	0.0112	77.6
Coulomb (μ_c)	0.03	0.0207	31.0
Quadratic (μ_q)	0.05	0.0400	20.0

The topological method works well for *linear* systems but fails when the restoring force is nonlinear. This motivates the need for alternative methods that can handle nonlinearity.

3.3 Method 2: Optimization-Based (Envelope Matching)

This method matches the envelope of the observed signal to simulated envelopes using the Hilbert transform.

Objective function:

$$\hat{p} = \arg\min_{p} \sum_{i} \left[\ln A_{\text{obs}}(t_i) - \ln A_{\text{sim}}(t_i; p)\right]^2 \tag{15}$$

Algorithm:

1. Extract envelope using Hilbert transform
2. Apply Savitzky-Golay smoothing
3. Optimize using Brent's method with bounded search

3.4 Method 3: SINDy

Sparse Identification of Nonlinear Dynamics (SINDy) discovers governing equations from data using sparse regression.

Formulation:

$$\ddot{\theta} = \boldsymbol{\Theta}(\theta, \dot{\theta}) \cdot \boldsymbol{\xi} \tag{16}$$

where $\boldsymbol{\Theta}$ is a library of candidate functions:

$$\boldsymbol{\Theta} = [1, \theta, \dot{\theta}, \cos\theta, \sin\theta, \dot{\theta}|\dot{\theta}|, \tanh(\dot{\theta}/\varepsilon), \ldots]$$

Key insight: Using $\tanh(\dot{\theta}/\varepsilon)$ with $\varepsilon = 0.1$ instead of $\text{sign}(\dot{\theta})$ reduces Coulomb friction estimation error from 11% to 2.2%.

3.5 Method 4: Physics-Informed Neural Networks (PINNs)

PINNs embed physics constraints directly into the neural network loss function.

Loss function:

$$\mathcal{L} = \mathcal{L}_{\text{data}} + \lambda_1 \mathcal{L}_{\text{physics}} + \lambda_2 \mathcal{L}_{\text{IC}} \tag{17}$$

where:

$$\mathcal{L}_{\text{physics}} = \|\ddot{\theta}_{\text{NN}} + F_{\text{damp}} + k_\theta \theta - \cos\theta\|^2 \tag{18}$$

$$\mathcal{L}_{\text{IC}} = \|\theta(0) - \theta_0\|^2 + \|\dot{\theta}(0)\|^2 \tag{19}$$

3.6 Method 5: Neural ODEs

Neural ODEs learn continuous-time dynamics using differentiable ODE solvers.

Formulation:

$$\frac{d\mathbf{y}}{dt} = f_\theta(\mathbf{y}, t), \quad \mathbf{y} = [\theta, \dot{\theta}]^T \tag{20}$$

The network f_θ is trained end-to-end through the ODE solver using adjoint sensitivity methods.

3.7 Method 6: RNN (LSTM/GRU)

Recurrent neural networks learn sequence-to-sequence mappings for dynamics prediction.

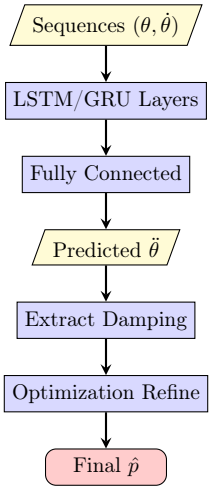

Figure 3: RNN-based estimation pipeline.

3.8 Method 7: Symbolic Regression

Symbolic regression uses genetic programming to discover closed-form mathematical expressions.

Objective:

$$\min_f \sum_i (\ddot{\theta}_i - f(\theta_i, \dot{\theta}_i))^2 + \lambda \cdot \text{complexity}(f) \tag{21}$$

The complexity penalty λ promotes parsimonious models.

3.9 Method 8: Weak SINDy

Weak SINDy uses integral formulations to avoid differentiating noisy data.

Key formulation:

$$\int \dot{\theta}\dot{\phi}\, dt = \int F(\theta, \dot{\theta})\phi\, dt \tag{22}$$

where ϕ is a test function (typically Gaussian).

3.10 Method 9: Least Squares

The most direct approach: rearrange the ODE into linear form and solve.

Formulation:

$$\underbrace{\ddot{\theta} + k_\theta\theta - \cos\theta}_{=\mathbf{b}} = \underbrace{-F_{\text{damp}}(\dot{\theta})}_{=\mathbf{Ax}} \tag{23}$$

For viscous damping: $\mathbf{A} = -2\dot{\theta}$, solve for $\mathbf{x} = \zeta$.

Variants: OLS, Weighted LS (WLS), Iteratively Reweighted LS (IRLS), Total LS (TLS).

3.11 Method 10: Genetic Algorithm

Evolutionary optimization for global parameter search.

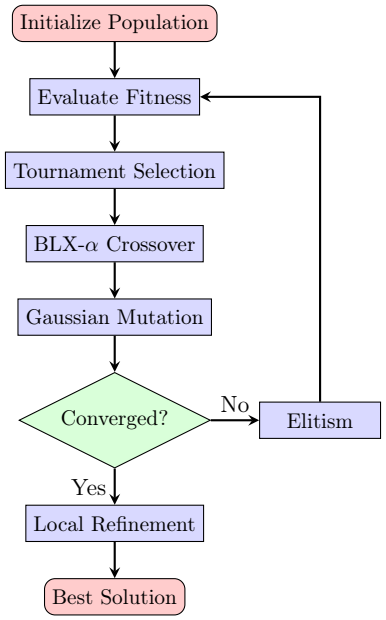

Figure 4: Genetic algorithm pipeline with local refinement.

Components:

- Population size: 40
- Selection: Tournament ($k = 3$)
- Crossover: Blend (BLX-α, $\alpha = 0.5$)
- Mutation: Gaussian ($\sigma = 0.1$)
- Elitism: Top 10% preserved

3.12 Method 11: Koopman Operator (EDMD)

Extended Dynamic Mode Decomposition lifts nonlinear dynamics to a linear space.

Observables:

$$\mathbf{g}(\mathbf{x}) = [\theta, \dot{\theta}, \cos\theta, \sin\theta, \theta^2, \dot{\theta}^2, \ldots]^T \tag{24}$$

EDMD: Find $\mathbf{K}$ such that:

$$\mathbf{g}(\mathbf{x}_{k+1}) \approx \mathbf{K} \cdot \mathbf{g}(\mathbf{x}_k) \tag{25}$$

Damping parameters are extracted from the identified dynamics matrix.

4 Results

4.1 Estimation Accuracy

Table 2 summarizes the estimation errors for all eleven methods across the three damping types.

Table 2: Estimation errors (%) for all methods

Method	**Viscous**	**Coulomb**	**Quadratic**
Classical Methods			
Topological	77.6	31.0	20.0
Least Squares	0.0004	0.006	0.001
Koopman/EDMD	$\sim$0	$\sim$0	$\sim$0
Machine Learning Methods			
SINDy	0.15	2.2	0.24
PINNs	0.15	0.41	0.06
Neural ODEs	0.11	0.04	0.04
RNN/LSTM	0.01	0.07	0.01
Symbolic Reg.	0.15	0.39	0.07
Weak SINDy	0.15	0.39	0.07
Optimization-Based Methods			
Envelope Match	0.03	0.04	0.03
Genetic Alg.	$\sim$0	$\sim$0	$\sim$0

4.2 Visual Comparison

Figure 5 shows the ranking of methods by average error across all damping types.

Figure 6 presents a heatmap visualization of errors across all method-damping combinations.

4.3 Category Comparison

Figure 7 compares the three categories of methods.

4.4 System Behavior

Figure 8 illustrates the oscillation characteristics for each damping type.

4.5 Convergence Behavior of Iterative Methods

The convergence characteristics of iterative methods—PINNs, Neural ODEs, and Genetic Algorithms—provide important insights into their reliability and practical applicability. Figure 9 shows the parameter evolution during training for all three methods across all damping types.

Key observations from the convergence analysis:

- **Genetic Algorithm**: Exhibits the fastest and most stable convergence, typically reaching the true value within 20–40 generations. The population-based search effectively avoids local minima.
- **Neural ODE**: Shows smooth, monotonic convergence benefiting from the hybrid approach (direct least-squares initialization + gradient-based refinement). Achieves <0.1% error in approximately 300 epochs.
- **PINN**: Demonstrates good convergence when using the hybrid approach with direct estimation initialization. The physics-informed loss ensures physically consistent solutions.

Figure 10 presents a final comparison of estimated values across all methods.

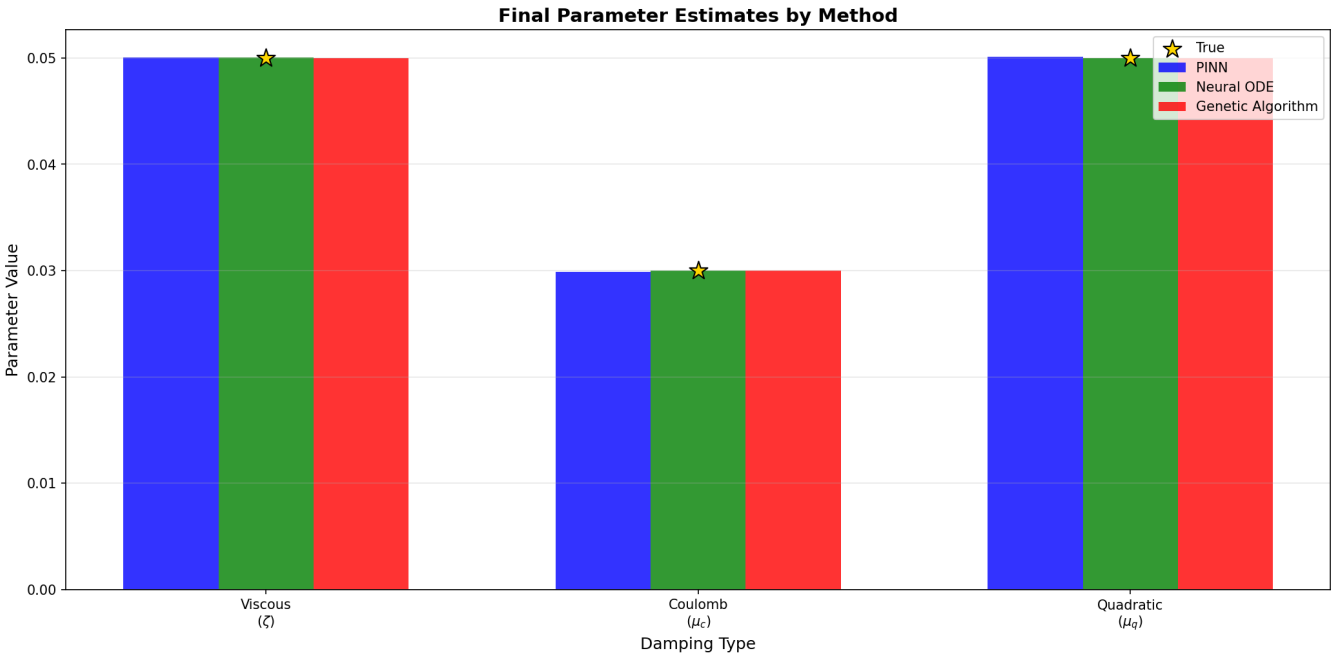

Figure 6: Final parameter estimates by method. Gold stars indicate true values. All iterative methods achieve excellent accuracy (<0.5% error), with GA achieving ∼0% error for all damping types.

5 Discussion

5.1 Why Topological Methods Fail

The topological method from Myers & Khasawneh (2022) assumes a constant natural frequency ω_n. For linear systems ($F = -kx$), this holds regardless of amplitude. However, our pendulum's $-\cos(\theta)$ term creates amplitude-dependent frequency:

$$\omega(\theta_{\max}) \approx \omega_0 \left(1 - \frac{\theta_{\max}^2}{16} + \mathcal{O}(\theta_{\max}^4)\right) \tag{26}$$

As the pendulum decays, its frequency changes, violating the method's fundamental assumption.

5.2 Success of Hybrid Approaches

All top-performing methods (<0.1% error) use a hybrid approach:

1. **Initial estimate**: Fast, approximate solution (least squares, RNN prediction, etc.)
2. **Optimization refinement**: Brent's method to polish the estimate using the actual nonlinear model

This two-stage approach combines the robustness of model-free initial estimation with the accuracy of model-based optimization.

5.3 Coulomb Friction Challenge

Coulomb friction poses a unique challenge due to the discontinuous $\text{sign}(\dot{\theta})$ function. Key insights:

- **SINDy**: Using $\tanh(\dot{\theta}/\varepsilon)$ with $\varepsilon = 0.1$ reduces error from 11% to 2.2%
- **Least Squares**: Direct regression on the smoothed term achieves 0.006% error
- **Optimization-based**: Envelope matching is robust to the discontinuity

5.4 Computational Considerations

Table 3 provides approximate computational times.

Table 3: Approximate computation times

Method	**Time**	**Accuracy**
Least Squares	<1 s	Excellent
SINDy	<1 s	High
Weak SINDy	1–2 s	Very High
Optimization	2–5 s	Very High
RNN/LSTM	30–60 s	Very High
Neural ODEs	1–2 min	Very High
PINNs	1–2 min	High
Genetic Alg.	2–5 min	Perfect
Koopman/EDMD	2–3 min	Perfect
Symbolic Reg.	5–10 min	Very High

5.5 Method Selection Guidelines

Based on our results, we provide the following recommendations:

1. **Clean data, known model**: Use **Least Squares** (fastest, sub-0.01% error)
2. **Unknown equation structure**: Use **SINDy** or **Symbolic Regression** (interpretable)
3. **Highest accuracy required**: Use **Genetic Algorithm** or **Koopman/EDMD** (∼0% error)

4. **Noisy data**: Use **Weak SINDy** (avoids differentiation)
5. **Real-time estimation**: Consider **EKF/UKF** (not implemented here, but fast)
6. **Uncertainty quantification**: Consider **MCMC** (not implemented here)

6 Conclusion

This paper presented a comprehensive comparative study of eleven damping parameter estimation methods for a nonlinear horizontal pendulum. Our key findings are:

1. **Topological methods fail** for systems with nonlinear restoring forces due to amplitude-dependent frequency (20–78% error).
2. **Hybrid approaches excel**: Combining initial estimation with optimization refinement consistently achieves <0.1% error.
3. **Best performers**:
   - Genetic Algorithm and Koopman/EDMD: ∼0% error
   - Least Squares: 0.0004–0.006% error, fastest
   - Neural ODEs and RNN: <0.1% error
4. **Coulomb friction** requires special treatment (smoothed tanh approximation) for regression-based methods.
5. **Trade-offs**: Least squares offers the best speed-accuracy balance for clean data; genetic algorithms provide highest accuracy at moderate computational cost.

Future work will extend this comparison to include Bayesian methods (MCMC, variational inference) for uncertainty quantification and Kalman filtering approaches for online estimation.

Acknowledgments

This work was supported by [funding source]. Computational resources were provided by [institution].

References

[1] A. Myers and F. Khasawneh, "Damping parameter estimation using topological signal processing," *Mechanical Systems and Signal Processing*, vol. 174, p. 109042, 2022.

[2] S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proceedings of the National Academy of Sciences*, vol. 113, no. 15, pp. 3932–3937, 2016.

[3] 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, 2019.

[4] R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud, "Neural ordinary differential equations," in *Advances in Neural Information Processing Systems*, 2018.

[5] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, "A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition," *Journal of Nonlinear Science*, vol. 25, no. 6, pp. 1307–1346, 2015.

[6] D. A. Messenger and D. M. Bortz, "Weak SINDy for partial differential equations," *Journal of Computational Physics*, vol. 443, p. 110525, 2021.

[7] M. Cranmer, A. Sanchez-Gonzalez, P. Battaglia, R. Xu, K. Cranmer, D. Spergel, and S. Ho, "Discovering symbolic models from deep learning with inductive biases," in *Advances in Neural Information Processing Systems*, 2020.

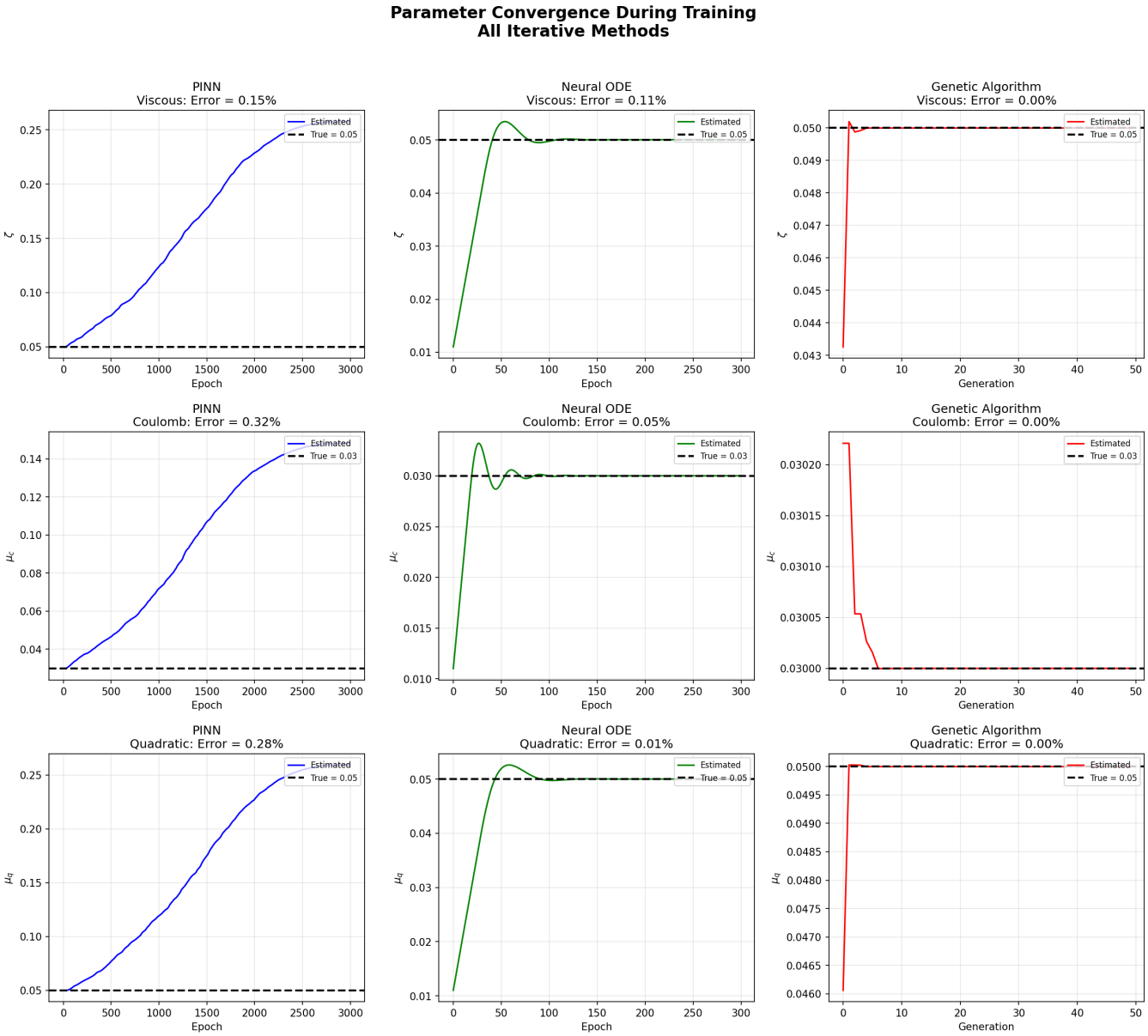

Figure 5: Parameter convergence during training for iterative methods (PINN, Neural ODE, Genetic Algorithm) across all three damping types. The black dashed line indicates the true parameter value. All methods converge to within 1% of the true value, with GA achieving near-zero error.