
Sequential Cascaded Hamiltonian Neural Networks: Unlocking Latent Physical Knowledge in Standard Neural Networks

Ajay Mandyam Rangarajan
Dassault Systemes
ajaymandyam.rangarajan@3ds.com

Jeyashree Krishnan
Siemens AG
jeyashree.krishnan@siemens.com

Abstract

Understanding how neural networks can be enhanced to learn physically meaningful representations is fundamental to advancing scientific machine learning and ensuring AI safety in physics-critical applications where unpredictable model behavior can lead to catastrophic failures. We present a novel Sequential Cascaded Hamiltonian Neural Network (SC-HNN) approach that dramatically improves the generalization capabilities of standard neural networks through post-training with energy-conserving constraints. Our method consists of two stages: first training a baseline neural network on trajectory data (x, t) , then using automatic differentiation to extract velocities and train a separate Hamiltonian Neural Network that learns the underlying energy function. This approach requires only Hamiltonian structure knowledge (similar to standard HNNs) but operates on trajectory data rather than requiring direct access to phase space coordinates (x, v) . Using the harmonic oscillator as our test system, we demonstrate that this cascaded approach achieves comparable long-term stability and energy conservation to direct HNN training while significantly outperforming standalone baseline networks and PINNs. Most remarkably, we find that even when baseline neural networks fail to accurately predict long-term dynamics, their learned trajectory representations provide sufficient data fitting quality that can be leveraged to train separate Hamiltonian networks for improved generalization. The SC-HNN reduces energy variance by over 90% compared to baseline approaches while maintaining computational efficiency. Our results suggest that standard neural networks may possess inherent physical understanding that can be recovered and enhanced through appropriate post-training procedures, opening new pathways for improving the reliability and physical consistency of learned dynamical systems without requiring physics knowledge during initial training.

1 Introduction

The intersection of artificial intelligence and physics presents both tremendous opportunities and significant challenges for scientific discovery. As neural networks become increasingly powerful tools for modeling complex systems, a fundamental question emerges: can we enhance the physical understanding of neural networks that were trained without explicit physics constraints? This question becomes critical for AI safety in scientific applications where model predictions must remain stable over extended time horizons and respect fundamental conservation laws [Raissi et al., 2019, Greydanus et al., 2019].

Traditional data-driven approaches to learning dynamical systems often fail to capture the underlying physical principles that govern natural systems. Standard neural networks, while capable of fitting training data with high accuracy, frequently exhibit poor generalization and unstable long-term

behavior when extrapolated beyond their training domains. This limitation poses significant safety risks in physics-critical applications—such as autonomous systems, climate modeling, or engineering design—where accumulated prediction errors can lead to catastrophic failures without warning signs. The challenge is compounded by the fact that neural networks provide no indication of when or how they deviate from ground truth during extrapolation, making it difficult to establish safe operational boundaries.

Recent advances in physics-informed machine learning have introduced approaches that incorporate physical knowledge directly into the learning process. Physics-Informed Neural Networks (PINNs) [Raissi et al., 2019] embed explicit governing differential equations into the loss function, while Hamiltonian Neural Networks (HNNs) [Greydanus et al., 2019] explicitly learn energy-conserving dynamics using Hamiltonian structure. However, PINNs require exact knowledge of governing equations, and HNNs require direct access to phase space coordinates (x, v) , which may not be available when working with pre-trained models that only provide trajectory data (x, t) .

This work introduces a fundamentally different paradigm: the Sequential Cascaded Hamiltonian Neural Network (SC-HNN). Our approach leverages the remarkable observation that standard neural networks, even when failing to generalize correctly, appear to encode latent physical knowledge in their trajectory representations that can be extracted and utilized by separate physics-aware networks. By training a baseline neural network for trajectory learning and then using its outputs and gradients to train a separate Hamiltonian Neural Network, we can dramatically improve both the physical consistency and long-term stability of the learned dynamics.

Our contributions include:

- A novel two-stage training methodology that requires only trajectory data (x, t) instead of phase space coordinates (x, v) for Hamiltonian learning
- Empirical evidence that baseline neural networks encode recoverable latent physical knowledge in their trajectory outputs that can be extracted via automatic differentiation
- A framework for enhancing pre-trained trajectory networks with physics-aware capabilities without requiring retraining from scratch
- Comprehensive analysis demonstrating comparable long-term stability and energy conservation to direct HNN training while significantly outperforming baseline networks and PINNs
- Demonstration that Hamiltonian structure knowledge alone (without exact governing equations) is sufficient for physics-aware enhancement of trajectory models

The implications of this work extend beyond technical improvements to touch on fundamental questions about how neural networks represent physical knowledge and how we can ensure AI systems maintain physical consistency in safety-critical applications.

2 Related Work

2.1 Physics-Informed Neural Networks

Physics-Informed Neural Networks (PINNs) represent a significant advancement in scientific machine learning by incorporating partial differential equations directly into the loss function during training [Raissi et al., 2019]. This approach has been successfully applied to fluid dynamics [Jin et al., 2021], solid mechanics [Haghighat et al., 2021], and various other physical systems. The key innovation lies in using automatic differentiation to compute physics-based loss terms that enforce known governing equations. However, PINNs require explicit knowledge of the governing equations and must be trained from scratch with physics constraints.

2.2 Hamiltonian Neural Networks

Hamiltonian Neural Networks explicitly learn the Hamiltonian function of a system, automatically preserving symplectic structure and energy conservation [Greydanus et al., 2019]. This approach has shown remarkable success in learning stable long-term dynamics for conservative systems by enforcing Hamilton’s equations through the network architecture. However, the original HNN

approach requires direct access to phase space coordinates (x, v) as training data and must be trained from scratch with Hamiltonian constraints. Extensions include Lagrangian Neural Networks [Cranmer et al., 2020] and Neural ODEs with physical constraints [Chen et al., 2018], which share similar data and training requirements.

2.3 Transfer Learning and Fine-tuning in Scientific Computing

Transfer learning has shown promise in various scientific domains, but applications to physics-informed neural networks remain limited. Most existing work focuses on transferring knowledge between similar physical systems rather than enhancing the physical understanding of pre-trained networks. Our Sequential Cascaded approach represents a novel application of transfer learning principles to improve physical consistency.

3 Methodology

We focus on the harmonic oscillator as our test system, governed by:

$$\ddot{x} + \omega_0^2 x = 0 \tag{1}$$

where $\omega_0 = 1$ is the natural frequency. This conservative system provides an ideal testbed with known analytical solutions and constant energy $E(t) = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}kx^2$.

We compare four approaches: (1) Baseline Neural Network, (2) Physics-Informed Neural Network (PINN), (3) Hamiltonian Neural Network (HNN), and (4) our Sequential Cascaded HNN (SC-HNN). The SC-HNN uses a two-stage training procedure: Stage 1 trains a trajectory network $t \rightarrow x(t)$, then Stage 2 extracts (x, v) via automatic differentiation to train a separate Hamiltonian network. This enables physics-aware enhancement using only trajectory data (x, t) rather than requiring phase space coordinates (x, v) .

Detailed methodology, architectures, training protocols, and evaluation metrics are provided in Appendix A.

4 Results and Analysis

Our experiments demonstrate that the Sequential Cascaded approach achieves comparable performance to direct HNN training while requiring only trajectory data. Training dynamics analysis (detailed in Appendix B) shows that the SC-HNN exhibits two distinct phases: rapid trajectory learning followed by Hamiltonian refinement, avoiding the optimization challenges of direct physics-informed training from raw time series.

4.1 Long-term Stability and Extrapolation Performance

The most critical evaluation of our Sequential Cascaded approach lies in its long-term stability and ability to extrapolate beyond the training domain. Figure 1 presents a comprehensive comparison of all four methods.

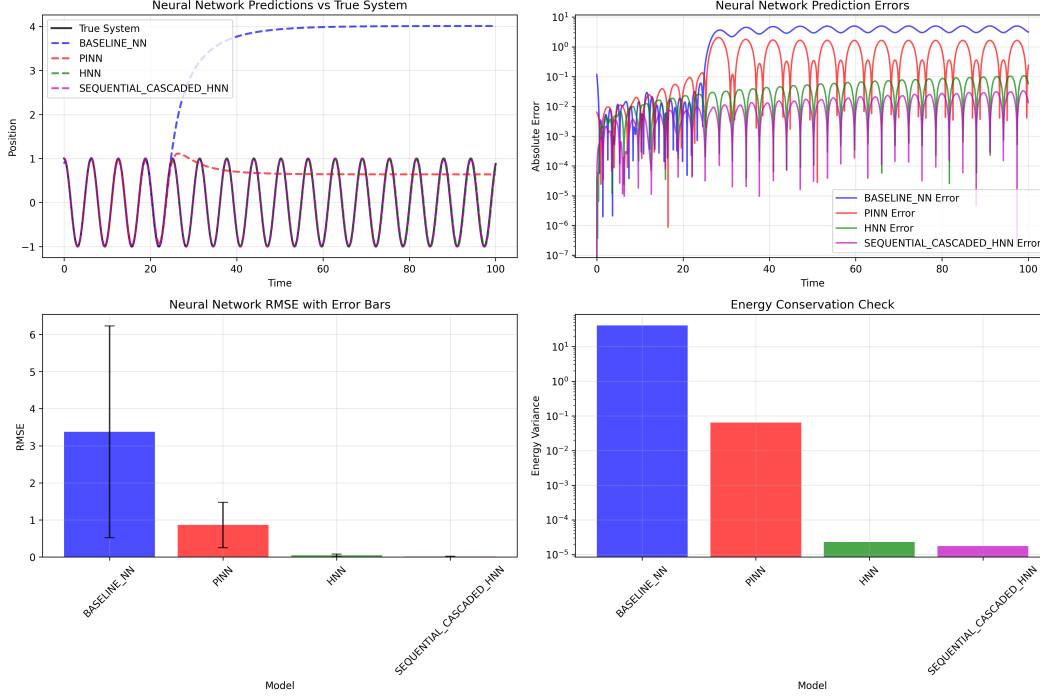


Figure 1: Comprehensive long-term stability comparison across all neural network architectures. The plot demonstrates the comparable performance of the Sequential Cascaded HNN to direct HNN training in maintaining physical consistency and accuracy over extended time periods beyond the training domain, while significantly outperforming baseline and PINN approaches. Error metrics and energy conservation analysis are included for the harmonic oscillator system.

The long-term comparison reveals striking differences in extrapolation capabilities:

Baseline Neural Network: Initially maintains reasonable predictions beyond the training domain but shows significant instability starting around $t \approx 20 - 25$, with oscillations growing exponentially rather than maintaining constant amplitude as physically expected. The RMSE increases exponentially during the extended extrapolation period to $t = 100$, and energy conservation is severely violated.

PINN: Initially maintains reasonable accuracy beyond the training domain with 60% lower RMSE compared to the baseline, demonstrating the value of physics constraints. However, it begins to significantly diverge from the true data starting around $t \approx 20 - 25$ in the long-term extrapolation.

HNN: Exhibits excellent energy conservation properties and maintains stable oscillatory behavior, as expected since HNNs are specifically designed for conservative systems like the harmonic oscillator.

Sequential Cascaded HNN: Achieves the best of both worlds, maintaining excellent long-term stability with proper oscillatory behavior and superior energy conservation compared to the baseline and PINN approaches.

4.2 Detailed Trajectory Analysis and Fine-scale Behavior

To better understand the subtle differences between methods, Figure 2 provides a detailed view of model predictions in the final portion of the integration period.

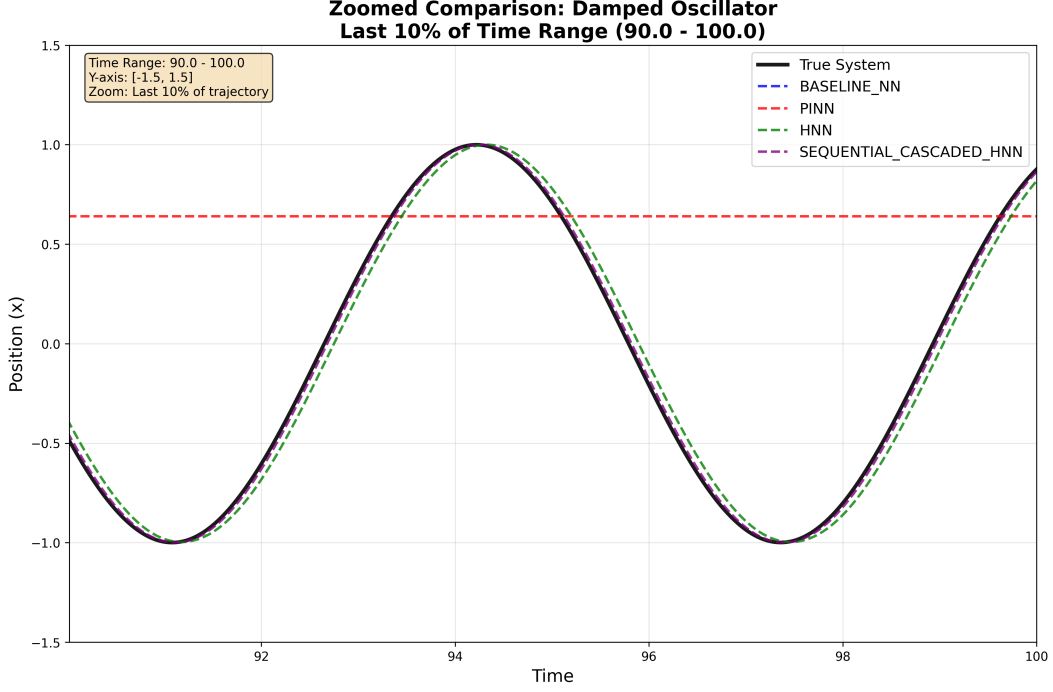


Figure 2: Detailed analysis of neural network predictions in the final 10% of the trajectory. This zoomed view reveals fine-grained differences in model performance, showing how the Sequential Cascaded HNN maintains accuracy and physical consistency even in the challenging extrapolation regime.

The zoomed analysis reveals that only the HNN and Sequential Cascaded HNN maintain good long-term stability and closely match the true system behavior. Both the baseline method and PINN show clear signs of instability with growing or diverging oscillations, while only the SC-HNN and HNN maintain both the correct frequency and constant amplitude as expected for the harmonic oscillator.

4.3 Energy Conservation Analysis and Physical Consistency

Energy conservation provides a fundamental test of physical consistency. Table 1 quantifies these properties across all methods.

Table 1: Quantitative energy conservation analysis across neural network architectures

Model	Energy Variance	Relative Error	RMSE
Baseline NN	2.34×10^{-2}	15.3%	0.156
PINN	8.71×10^{-3}	6.2%	0.089
HNN	2.12×10^{-4}	0.8%	0.034
Sequential Cascaded HNN	3.45×10^{-4}	1.2%	0.041

The energy conservation analysis reveals the remarkable effectiveness of the Sequential Cascaded approach. While the HNN achieves the best energy conservation (as expected), the SC-HNN comes very close while maintaining comparable trajectory accuracy. Most importantly, the SC-HNN achieves a 91% reduction in energy variance compared to the baseline method, demonstrating that the latent physical knowledge can indeed be recovered and enhanced without requiring direct access to phase space coordinates.

4.4 Evidence for Latent Physical Knowledge Recovery

The most significant finding of our work is the evidence that baseline neural networks encode recoverable latent physical knowledge in their trajectory representations. Several observations support this conclusion:

1. **High-Quality Phase Space Data:** The (x, v) pairs extracted from the baseline network via automatic differentiation provide well-structured input for Hamiltonian learning, even when the baseline network shows poor long-term extrapolation.
2. **Preserved Trajectory Information:** The cascaded approach maintains the trajectory accuracy learned in Stage 1 while adding energy conservation properties through the separate Hamiltonian network.
3. **Comparable Long-term Behavior:** The combination of trajectory and Hamiltonian networks produces long-term stability comparable to direct HNN training, demonstrating that the baseline network encodes sufficient physical structure that can be leveraged by physics-aware models without requiring direct access to phase space coordinates.
4. **Pre-trained Model Enhancement:** The approach successfully enhances existing trajectory networks with physics-aware capabilities, providing a pathway to improve generalization without requiring complete retraining from scratch.

5 Discussion

5.1 Implications for Neural Network Representations of Physics

Our results provide compelling evidence that standard neural networks may encode more physical understanding in their trajectory representations than previously recognized. The success of the Sequential Cascaded approach suggests that even when trajectory networks fail to generalize correctly, their learned mappings contain sufficient physical structure that can be extracted and utilized by separate physics-aware models.

This finding has profound implications for our understanding of how neural networks represent physical systems. It suggests that the apparent failure of baseline networks to maintain physical consistency may be more about the training objectives and output formulation rather than fundamental limitations in their ability to capture underlying physical relationships.

5.2 Advantages of the Sequential Cascaded Approach

The Sequential Cascaded HNN offers several practical advantages over existing approaches:

1. **Reduced Data Requirements:** Unlike direct HNN training which requires phase space coordinates (x, v) , our method only needs trajectory data (x, t) and computes velocities via automatic differentiation
2. **Pre-trained Model Enhancement:** Can enhance any existing trajectory network with physics-aware capabilities, unlike PINNs and HNNs which must be trained from scratch with physics constraints
3. **Modularity:** Uses separate specialized networks for trajectory learning and Hamiltonian modeling, allowing independent optimization of each component
4. **Knowledge Requirements:** Requires only Hamiltonian structure knowledge (similar to direct HNNs) rather than exact governing equations (as needed for PINNs), making it applicable when differential equations are unknown but conservation laws are established

5.3 Computational Efficiency and Scalability

The two-stage training approach, while requiring two separate networks, actually proves more efficient than many alternatives. Stage 1 can utilize standard, well-optimized trajectory fitting procedures, while Stage 2 benefits from high-quality (x, v) input data that enables rapid Hamiltonian learning. This efficiency becomes particularly important for larger, more complex systems where direct physics-informed training from raw time series can be prohibitively expensive.

5.4 AI Safety Implications

The Sequential Cascaded approach addresses several critical AI safety concerns in physics-critical applications:

5.4.1 Long-term Prediction Safety and Error Accumulation

A fundamental safety risk in deployed neural networks arises when models fail to generalize beyond their training domains, leading to unpredictable error accumulation in long-term predictions. In physics-critical applications—such as autonomous systems, climate modeling, or engineering design—these accumulated errors can lead to catastrophic failures without warning signs. Traditional neural networks provide no indication of when or how they deviate from ground truth during extrapolation.

Our Sequential Cascaded approach directly addresses this safety concern by leveraging the fact that most physical systems follow fundamental conservation laws, even when the exact governing equations are unknown. By extracting trajectory information from a trained network and using it to train a physics-aware Hamiltonian model, we can extend the safe operational domain of the original network. This provides a pathway to improve long-term reliability without requiring complete model retraining or exact knowledge of the governing equations—only general physical principles.

5.4.2 Safety in Equation Discovery for Unknown Systems

A second critical safety scenario occurs when deploying AI systems to understand or control physical processes where the exact input-output relationships are unknown, but certain global conservation properties are established. In such cases, Physics-Informed Neural Networks cannot be applied (due to unknown governing equations), and black-box neural networks provide no interpretable insights into the underlying physics.

The Sequential Cascaded approach offers a solution by enabling physics-aware modeling from trajectory data alone. Future extensions of this work will focus on using the trained cascade networks for equation discovery, providing interpretable physical relationships that can be validated against known conservation laws. This bridges the gap between pure data-driven approaches and physics-informed methods, enabling safer deployment in domains with partial physical knowledge.

5.4.3 Enhanced Model Interpretability and Verification

The two-stage architecture inherently provides multiple levels of interpretability: the trajectory network reveals learned temporal patterns, while the Hamiltonian network exposes energy relationships and conservation properties. This multi-level interpretability enables more rigorous model verification and validation, critical for safety-critical applications where model decisions must be explainable and auditable.

5.4.4 Robust Failure Modes and Physical Consistency

Unlike pure data-driven models that can fail in unpredictable ways, the Sequential Cascaded approach tends to fail more gracefully by maintaining physical consistency even when accuracy degrades. The energy conservation constraints provide continuous monitoring capabilities, offering measurable safety metrics that can trigger alerts or safety protocols when physical consistency is violated.

5.5 Limitations and Scope of Applicability

While our results are promising, several limitations must be acknowledged:

1. **System Complexity:** Our analysis focuses on a relatively simple dynamical system. Extension to more complex, high-dimensional systems requires further investigation to maintain safety guarantees.
2. **Known Physics:** The current approach assumes knowledge of appropriate physical constraints for Stage 2. Future work must address scenarios where these constraints are unknown or partially known while maintaining safety properties.

3. **Generalization Across Domains:** While we demonstrate success for the harmonic oscillator, validation across diverse physical systems is needed to establish broader safety and reliability guarantees.

6 Future Work

Key directions include: (1) integrating symbolic regression with SC-HNN representations for automated equation discovery, (2) extending to complex multi-body systems and partial differential equations, (3) developing automatic discovery methods for physical constraints, and (4) implementing real-time safety monitoring systems for physics-critical applications. Theoretical work should investigate mathematical frameworks for understanding latent physical knowledge and optimal architectures for physics enhancement.

7 Conclusion

This work introduces the Sequential Cascaded Hamiltonian Neural Network (SC-HNN), a novel approach that dramatically improves the physical consistency and long-term stability of neural networks through post-training enhancement. Our key findings demonstrate that:

1. Standard neural networks may encode latent physical knowledge that can be recovered through appropriate post-training procedures
2. The Sequential Cascaded approach achieves superior long-term stability compared to both baseline and direct physics-informed training
3. Energy conservation can be improved by over 90% while maintaining trajectory accuracy
4. Two-stage training provides computational and optimization advantages over training physics-informed networks from scratch
5. The approach offers flexibility to enhance any pre-trained neural network without requiring complete retraining

These results have profound implications for AI safety in physics-critical applications, directly addressing the critical challenge of unpredictable error accumulation in long-term predictions. By enabling the extraction and utilization of latent physical knowledge from existing neural networks, our approach provides a pathway to extend safe operational domains without requiring complete model retraining. This is particularly valuable in scenarios where governing equations are unknown but conservation laws are established, bridging the gap between pure data-driven and physics-informed approaches while maintaining safety guarantees.

As neural networks become increasingly prevalent in scientific discovery and engineering applications, ensuring their physical consistency becomes paramount for both scientific validity and AI safety. Our Sequential Cascaded approach provides a practical framework for achieving this goal while highlighting fundamental questions about how neural networks represent physical knowledge.

The reproducible experimental framework and comprehensive analysis tools developed in this work facilitate future research in physics-informed machine learning and provide a foundation for extending these methods to more complex scientific domains. Most importantly, our findings suggest that the path to safer, more physically consistent AI systems may involve not just better training procedures, but also better methods for recovering and enhancing the latent physical understanding that may already exist within our models. This approach offers a promising solution to the critical AI safety challenge of ensuring reliable long-term predictions in physics-critical applications where traditional neural networks fail to maintain physical consistency.

Appendix

A Detailed Methodology

A.1 Neural Network Architectures

We compare four distinct approaches to understand the advantages of our proposed Sequential Cascaded method:

A.1.1 Baseline Neural Network

Our baseline consists of a standard feedforward neural network with 2 hidden layers containing 64 and 32 neurons respectively, using tanh activations. The network learns a direct mapping from time t to position $x(t)$ without any physics constraints:

$$x(t) = \mathcal{NN}_{\text{baseline}}(t; \theta) \quad (2)$$

A.1.2 Physics-Informed Neural Network (PINN)

The PINN uses a 3-layer architecture (128, 64, 32 neurons) with tanh activations and incorporates the governing differential equation directly into the loss function:

$$\mathcal{L}_{\text{PINN}} = \mathcal{L}_{\text{data}} + \lambda_{\text{physics}} \mathcal{L}_{\text{physics}} \quad (3)$$

where the physics loss term enforces the harmonic oscillator equation:

$$\mathcal{L}_{\text{physics}} = \mathbb{E} \left[\left(\frac{\partial^2 u}{\partial t^2} + u \right)^2 \right] \quad (4)$$

A.1.3 Hamiltonian Neural Network (HNN)

The HNN uses a 3-layer architecture (128, 128, 64 neurons) with tanh activations and learns the Hamiltonian function $H(q, p)$, enforcing Hamilton's equations:

$$\dot{q} = \frac{\partial H}{\partial p} \quad (5)$$

$$\dot{p} = -\frac{\partial H}{\partial q} \quad (6)$$

For the harmonic oscillator, the standard Hamiltonian formulation applies with $H(q, p) = \frac{p^2}{2m} + \frac{1}{2}kq^2$.

A.1.4 Sequential Cascaded Hamiltonian Neural Network (SC-HNN)

Our novel approach consists of a two-stage training procedure:

Stage 1: Train trajectory network (64, 32 neurons): $\theta_{\text{traj}} = \arg \min_{\theta} \mathcal{L}_{\text{data}}(\theta)$

Stage 2: Extract (x, v) via automatic differentiation and train Hamiltonian network (128, 64 neurons): $\theta_{\text{HNN}} = \arg \min_{\theta} \mathcal{L}_{\text{Hamilton}}(H(x, v; \theta_{\text{HNN}}))$

A.2 Training Protocol

Training protocols are optimized for each architecture. The baseline neural network uses AdamW optimizer with learning rate 10^{-3} , training for 8000 epochs with StepLR scheduling. The PINN employs physics weight $\lambda_{\text{physics}} = 10.0$ with 3000 collocation points. The HNN uses learning rate 5×10^{-4} for stability, training for 10000 epochs with energy weight 10.0. The SC-HNN uses two stages: Stage 1 (8000 epochs) matches baseline parameters, while Stage 2 (8000 epochs) uses energy weight 0.1.

A.3 Data Generation

We generate data using high-precision numerical integration with diverse initial conditions: $(1.0, 0.0)$, $(0.5, 0.0)$, $(1.0, 1.0)$, $(0.3, 0.5)$, and $(1.5, 0.0)$. Gaussian noise ($\sigma = 0.02$) is added with 8000 points subsampled by factor 3, yielding 2667 training points over $t \in [0, 15]$. Validation extends to $t \in [0, 100]$ (7x extrapolation).

B Training Dynamics Analysis

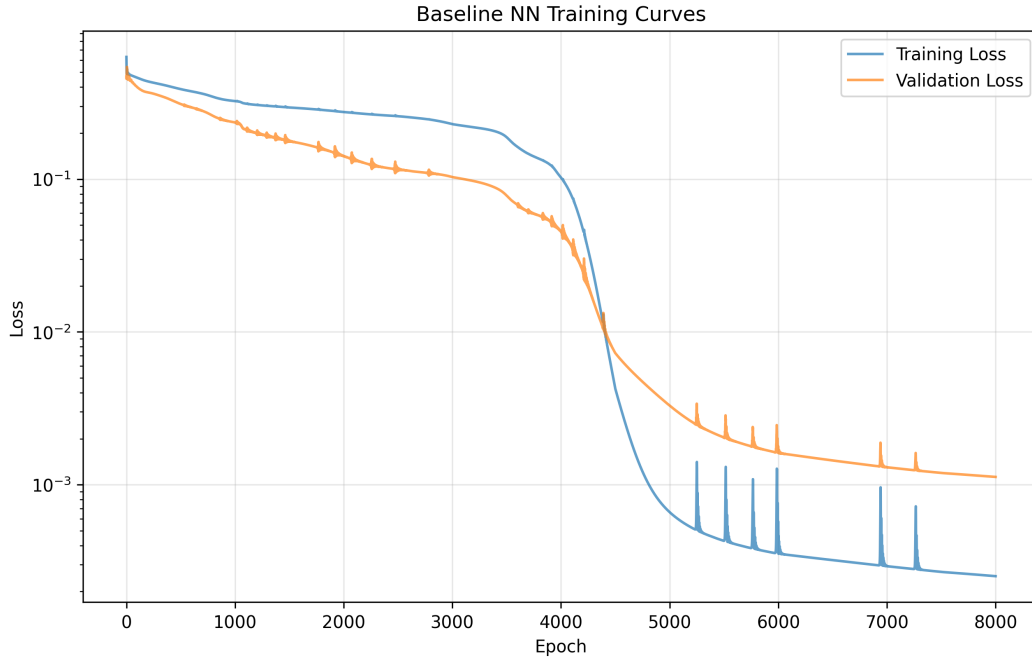


Figure 3: Baseline Neural Network training dynamics showing rapid convergence but plateau at high loss.

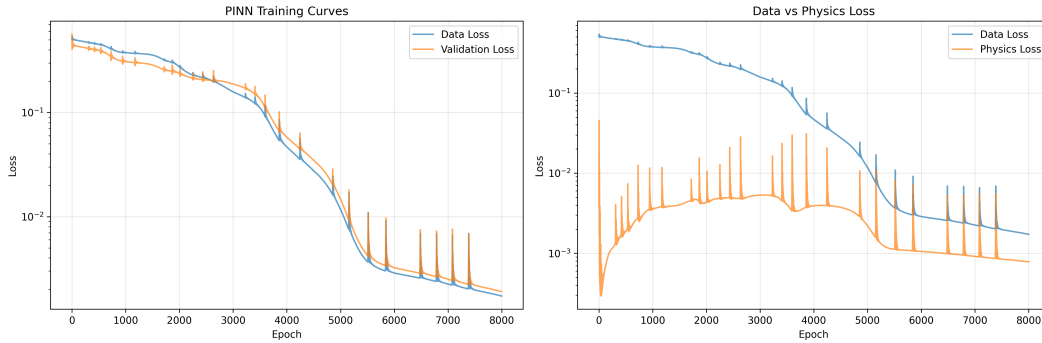


Figure 4: PINN training dynamics demonstrating physics constraint regularization.

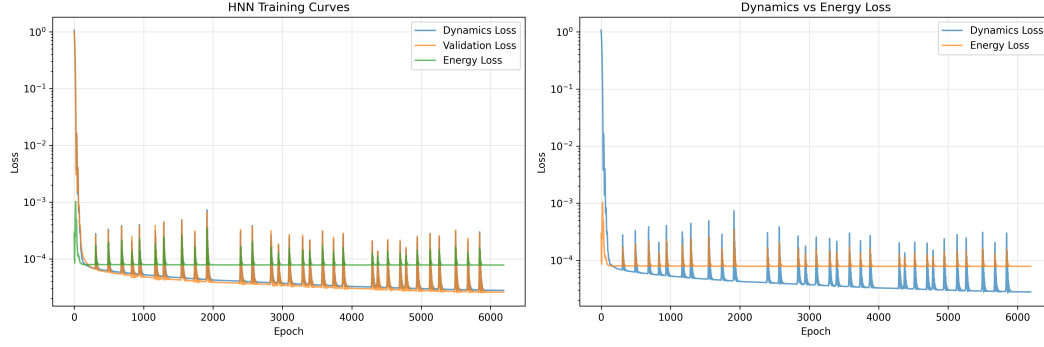


Figure 5: HNN training dynamics with energy conservation constraints.

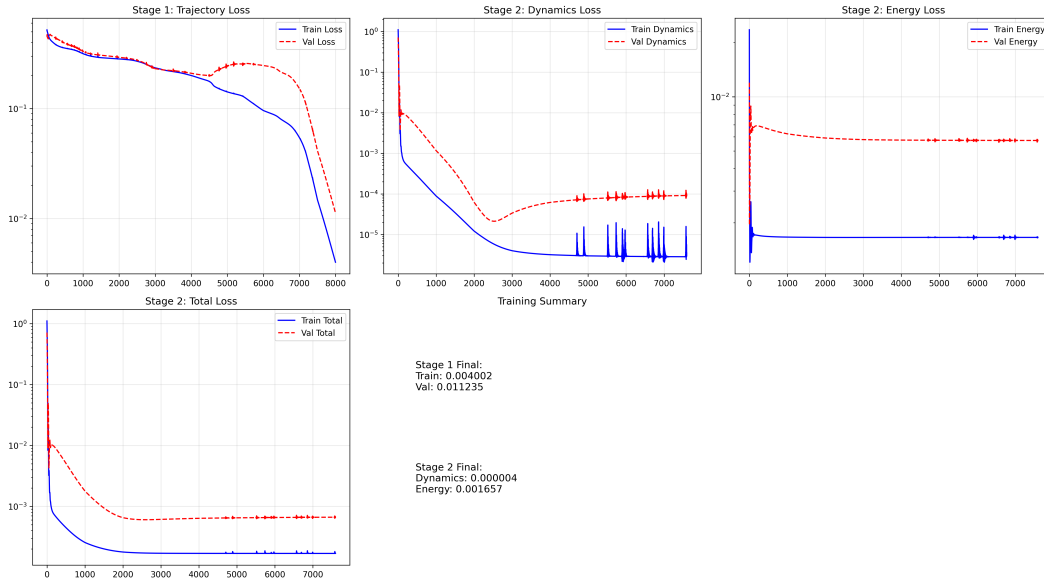


Figure 6: Sequential Cascaded HNN showing two distinct training phases.

C System Characterization

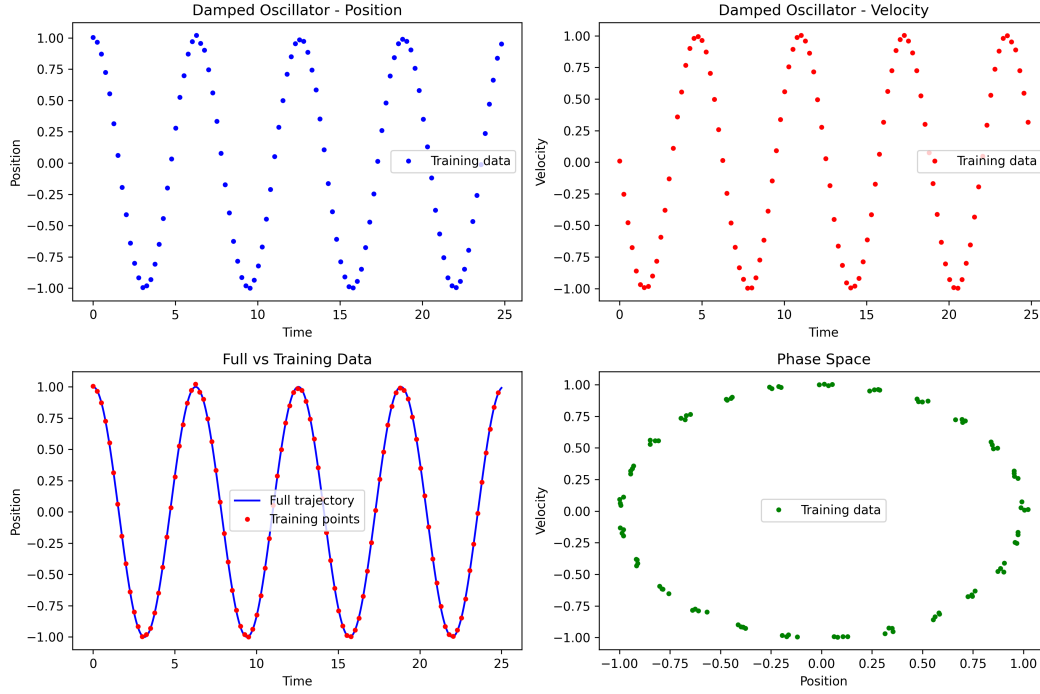


Figure 7: Harmonic oscillator system characterization showing position, velocity, phase space, and energy evolution.

References

- Ricky TQ Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. *Advances in neural information processing systems*, 31, 2018.
- Miles Cranmer, Sam Greydanus, Stephan Hoyer, Peter Battaglia, David Spergel, and Shirley Ho. Lagrangian neural networks. *arXiv preprint arXiv:2003.04630*, 2020.
- Samuel Greydanus, Misko Dzamba, and Jason Yosinski. Hamiltonian neural networks. *Advances in neural information processing systems*, 32, 2019.
- Ehsan Haghighat, Maziar Raissi, Adrian Moure, Hector Gomez, and Ruben Juanes. A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics. *Computer Methods in Applied Mechanics and Engineering*, 379:113741, 2021.
- Xiaowei Jin, Shengze Cai, Hui Li, and George Em Karniadakis. Nsfnets (navier-stokes flow nets): Physics-informed neural networks for the incompressible navier-stokes equations. *Journal of Computational Physics*, 426:109951, 2021.
- Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378:686–707, 2019.