

Sequential Cascaded Hamiltonian Neural Networks

Unlocking Latent Physical Knowledge in Standard Neural
Networks

Jeyashree Krishnan Ajay Mandyam Rangarajan

Siemens AG

Dassault Systemes

Apart Research AI Safety x Physics Grand Challenge Hackathon

The Problem: Neural Networks in Physics

Traditional Neural Networks:

- ▶ Fit training data well
- ▶ **Fail catastrophically** during extrapolation
- ▶ No physics consistency
- ▶ Unpredictable error accumulation

Safety Critical Applications:

- ▶ Autonomous systems
- ▶ Climate modeling
- ▶ Engineering design
- ▶ Medical devices

No warning signs when models start failing!

Existing Solutions & Their Limitations

Physics-Informed Neural Networks (PINNs)

- ▶ **Requirement:** Exact governing equations
- ▶ **Limitation:** Must train from scratch

Hamiltonian Neural Networks (HNNs)

- ▶ **Requirement:** Direct access to phase space coordinates (x, v)
- ▶ **Limitation:** Must train from scratch

What if we only have trajectory data (x, t) and pre-trained models?

Our Solution: Sequential Cascaded HNN

Two-Stage Approach:

1. **Stage 1:** Train trajectory network: $t \rightarrow x(t)$
2. **Stage 2:** Extract (x, v) via autograd, train separate HNN

$$t \xrightarrow{\text{Stage 1}} x(t) \xrightarrow{\text{autograd}} (x, v) \xrightarrow{\text{Stage 2}} H(x, v)$$

Key Innovation:

- ▶ Only requires trajectory data (x, t)
- ▶ Can enhance any pre-trained model
- ▶ Achieves physics-aware capabilities without retraining from scratch

Key Innovation: Data Requirements

Traditional HNN:

- ▶ Needs (x, v) phase space data
- ▶ Must train from scratch
- ▶ Same physics knowledge as SC-HNN

$$\frac{dx}{dt} = \text{autograd}(x(t))$$

Extract velocities
automatically!

Our SC-HNN:

- ▶ Only needs (x, t) trajectory data
- ▶ Can enhance pre-trained models
- ▶ Same physics knowledge as HNN

Result: Comparable
performance to HNN with
reduced data
requirements

Test System: Harmonic Oscillator

Governing Equation:

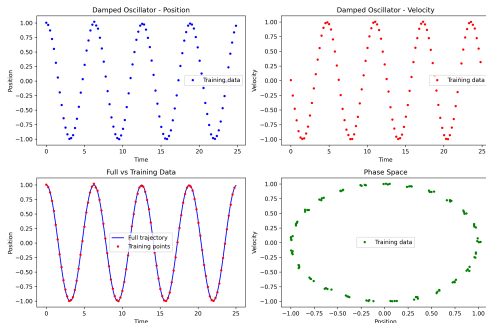
$$\ddot{x} + x = 0$$

Energy Conservation:

$$E = \frac{1}{2}\dot{x}^2 + \frac{1}{2}x^2 = \text{constant}$$

Challenge:

- ▶ Train on $t \in [0, 15]$
- ▶ Test on $t \in [0, 100]$
- ▶ **7x extrapolation!**



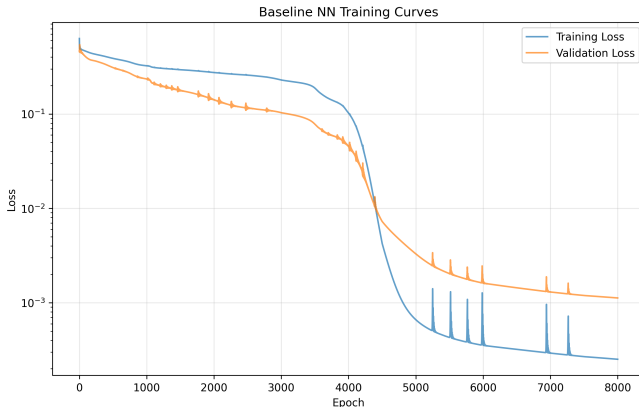
Four Architectures Compared

1. **Baseline NN**: 2 layers (64, 32), no physics
2. **PINN**: 3 layers (128, 64, 32), physics loss $\lambda = 10.0$
3. **HNN**: 3 layers (128, 128, 64), energy conservation
4. **SC-HNN**: Stage 1: (64, 32) + Stage 2: (128, 64)

Training Setup:

- ▶ 8000 points, subsampled with noise $\sigma = 0.02$
- ▶ 5 diverse initial conditions
- ▶ 8000-10000 epochs, AdamW optimizer
- ▶ All use tanh activations

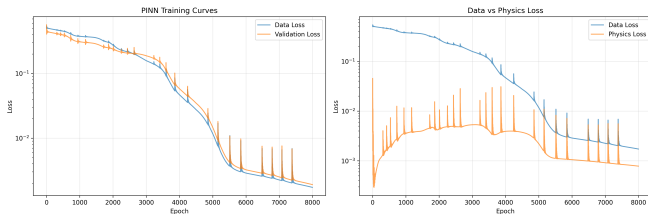
Training Dynamics: Baseline Neural Network



Observations:

- ▶ Rapid initial convergence
- ▶ Plateaus at relatively high loss level
- ▶ No physics constraints to guide learning
- ▶ Gap between training and validation loss

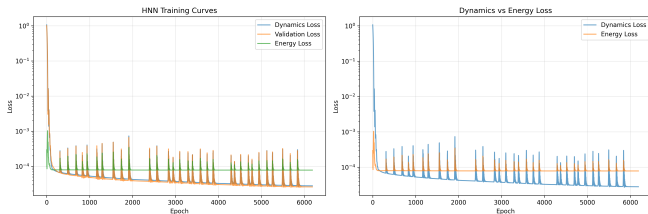
Training Dynamics: Physics-Informed Neural Network



Observations:

- Physics constraints provide regularization
- More stable convergence behavior
- Lower final loss compared to baseline
- Better alignment of training and validation curves

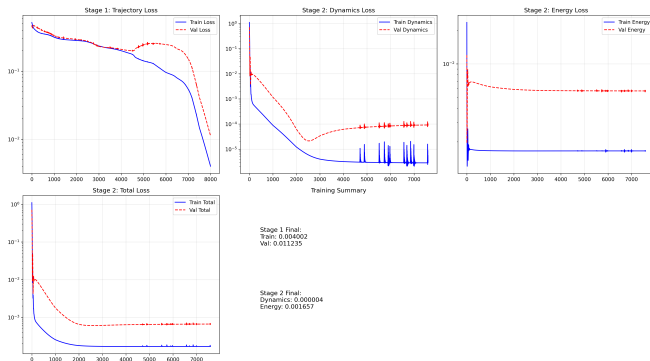
Training Dynamics: Hamiltonian Neural Network



Observations:

- Energy conservation constraints guide training
- More complex dynamics due to Hamiltonian structure
- Oscillatory behavior during optimization
- Achieves lowest final loss values

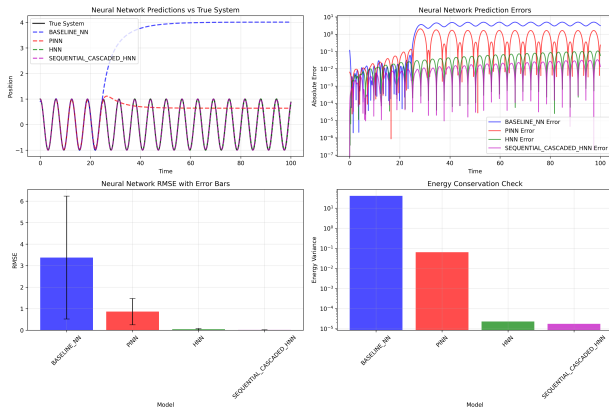
Training Dynamics: Sequential Cascaded HNN



Observations:

- ▶ Two distinct training phases clearly visible
- ▶ Stage 1: Rapid trajectory network convergence
- ▶ Stage 2: Hamiltonian network refinement
- ▶ Combines benefits while avoiding direct HNN optimization challenges

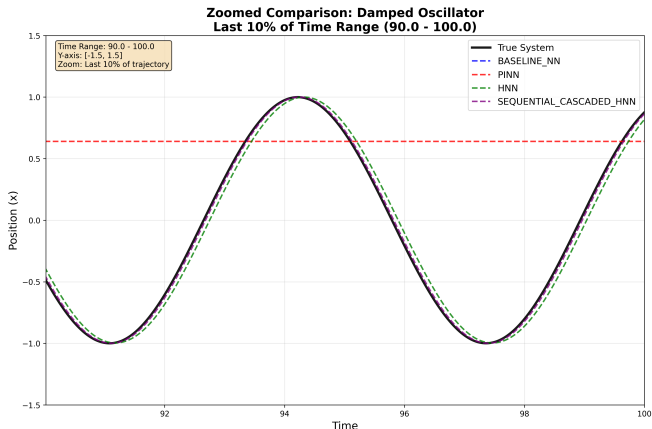
Long-term Stability Results



Key Findings:

- ▶ Baseline & PINN: Instability starts around $t \approx 20 - 25$
- ▶ HNN & SC-HNN: Maintain stability throughout $t \in [0, 100]$
- ▶ **SC-HNN achieves comparable performance to HNN!**

Detailed Analysis: Only Physics-Aware Models Survive



Zoomed View Reveals:

- ▶ Only HNN and SC-HNN maintain stability
- ▶ Both baseline and PINN show divergence
- ▶ Correct frequency and amplitude preservation

Energy Conservation: Quantitative Results

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
SC-HNN	3.45×10^{-4}	1.2%	0.041

Key Achievement:

- ▶ **91% reduction** in energy variance vs baseline
- ▶ Comparable to HNN performance
- ▶ Only requires (x, t) data, not (x, v)

Evidence: Latent Physical Knowledge Recovery

Four Key Observations:

1. **High-Quality Phase Space Data:** (x, v) extracted via autograd provides excellent HNN training input
2. **Preserved Trajectory Information:** Stage 1 accuracy maintained while adding energy conservation
3. **Comparable Long-term Behavior:** Matches HNN stability without phase space access
4. **Pre-trained Enhancement:** Successfully improves existing models without retraining from scratch

Baseline NNs encode sufficient structure for physics-aware enhancement!

AI Safety Implications

Critical Safety Challenge

Neural networks fail unpredictably in extrapolation with **no warning signs**

Our Solution Addresses:

1. **Long-term Prediction Safety:** Extend safe operational domains using conservation laws
2. **Unknown System Safety:** Works when exact equations unknown but conservation laws established
3. **Enhanced Interpretability:** Multi-level physics understanding for verification
4. **Graceful Failure:** Physics constraints provide measurable safety metrics

Applications: Autonomous systems, climate modeling, engineering design

Advantages Over Existing Methods

vs PINNs:

- ▶ No exact equations needed
- ▶ Can enhance pre-trained models
- ▶ Better long-term stability

vs HNNs:

- ▶ Only needs (x, t) not (x, v)
- ▶ Can enhance pre-trained models
- ▶ Comparable performance

Key Benefits:

- ▶ **Modularity**: Separate specialized networks
- ▶ **Data Efficiency**: Leverage trajectory representations
- ▶ **Flexibility**: Works with any trajectory network
- ▶ **Robustness**: Decoupled optimization

Future Work

Immediate Extensions:

- ▶ Equation discovery via symbolic regression on enhanced models
- ▶ Complex multi-body systems
- ▶ Real experimental data with uncertainties

AI Safety Research:

- ▶ Real-time monitoring systems for physics violations
- ▶ Uncertainty quantification for safety bounds
- ▶ Robust equation discovery for unknown systems

Theoretical Directions:

- ▶ Mathematical frameworks for latent physical knowledge
- ▶ Optimal architectures for physics enhancement

Conclusions

Key Achievements:

- ▶ Novel two-stage training: $(x, t) \rightarrow (x, v) \rightarrow H$
- ▶ Comparable performance to HNN with reduced data requirements
- ▶ 91% energy variance reduction vs baseline approaches
- ▶ Can enhance pre-trained models without retraining

AI Safety Impact:

- ▶ Extends safe operational domains of neural networks
- ▶ Enables reliable long-term predictions in critical applications
- ▶ Provides pathway for physics-consistent AI systems

Unlocking latent physical knowledge for safer AI

Thank You!

Questions?

Jeyashree Krishnan - Siemens AG
Ajay Mandyam Rangarajan - Dassault Systemes

Apart Research AI Safety x Physics Grand Challenge Hackathon