Title

Ensemble of Numerics-Informed Neural Networks with Embedded Hamiltonian Constraints for Modeling Nonlinear Structural Dynamics

ISBN

978-3-031-04085-6

Authors

Najera-Flores, David A
Todd, Michael D

Publication Date

2023

DOI

10.1007/978-3-031-04086-3_4

Peer reviewed

Ensemble of Numerics-Informed Neural Networks with Embedded Hamiltonian Constraints for Modeling Nonlinear Structural Dynamics

David A. Najera-Flores[1, 2], and Michael D. Todd[1]

[1] Department of Structural Engineering
University of California San Diego

[2] ATA Engineering, Inc., San Diego, CA

ABSTRACT

Data-driven machine learning models are useful for modeling complex structures based on empirical observations, bypassing the need to generate a physical model where the physics is not well known or readily otherwise model-able. One disadvantage of purely data-driven approaches is that they tend to perform poorly in regions outside the original training domain. To mitigate this limitation, physical knowledge about the structure can be embedded in the model architecture via the model topology or numerical constraints in the formulation. For large-scale systems, relevant physics, such as the system state matrices, may be expensive to compute. One way around this problem is to use scalar functionals, such as energy, to constrain the network to operate within physical bounds. We propose a neural network framework based on Hamiltonian mechanics to enforce a physics-informed structure to the model. The Hamiltonian framework allows us to relate the energy of the system to the measured quantities (e.g., accelerations) through the Euler-Lagrange equations of motion. In this work, the potential, kinetic energy, and Rayleigh damping terms are each modeled with a multilayer perceptron. Auto-differentiation is used to compute partial derivatives and assemble all the relevant equations, including computing the generalized inertia matrix by forming the Hessian of the kinetic energy with respect to the generalized coordinates. Moreover, a Bayesian approach is used to estimate model-form error to predict domain shifts in the data and enable model correction. The network incorporates a numerics-informed loss function via the residual of a multi-step integration term, allowing the ensemble of networks to be time-integrated with new initial conditions and an arbitrary external force after it has been trained. The approach is demonstrated on simple exemplars, such as a two degree-of-freedom (DOF) damped oscillator with cubic nonlinearities.

Keywords: Nonlinear dynamics, Hamiltonian mechanics, Euler-Lagrange, neural networks

INTRODUCTION

Data-driven machine learning models are useful for modeling complex structures based on empirical observations, bypassing the need to generate a physical model where the physics is not well known or readily otherwise model-able. In this work, we impose a mathematical structure on a data-driven approach to improve robustness of data-driven models and reduce the need for large amounts of training data. We use an ensemble of neural networks constrained according to Hamiltonian mechanics to ensure that the networks are learning physically interpretable representations. The Hamiltonian framework allows us to relate the energy of the system to the measured quantities (e.g., accelerations) through the Euler-Lagrange equations of motion. This work demonstrates the approach by predicting the response of a two degree-of-freedom (DOF) damped oscillator with cubic nonlinearities.

NEURAL NETWORK ARCHITECTURE

The framework is composed of an ensemble of neural networks that are communicating with each other. Figure *1* shows the general architecture of networks that was used. We begin by assuming that the observations may not be in an inertial coordinate

system and that, in general, they do not correspond to a set of generalized coordinates. The initial coordinates are denoted by r, and the generalized coordinates are denoted by q.

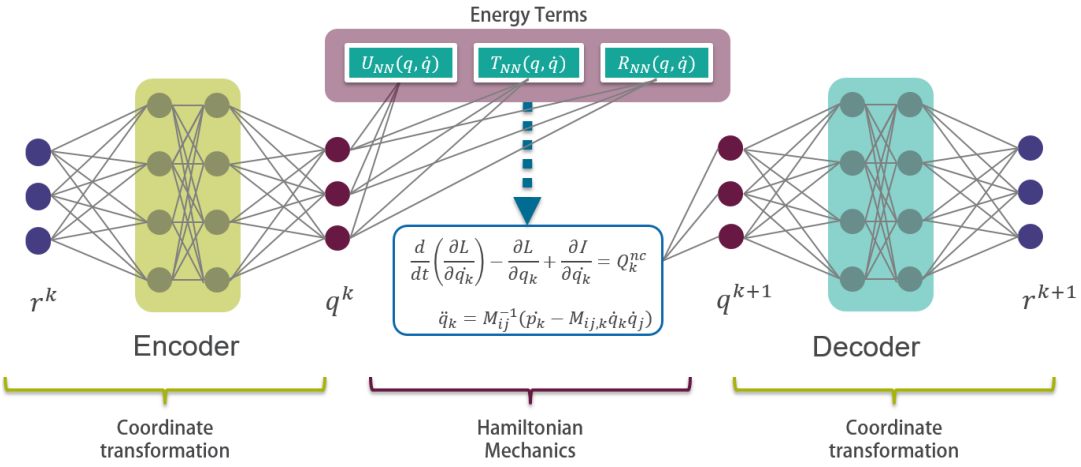

Figure 1. Ensemble of neural network architecture with Hamiltonian constraints.

An autoencoder is used to learn coordinate transformations between these two sets of coordinates. The encoder is first used to map the initial coordinates to the generalized coordinates, and a decoder is used to map them back to the original reference frame. After the transformation is learned, the latent space functions as the generalized coordinate system. During this step, it would be possible to reduce the number of dimensions but that has not been explored yet. For now, the dimension of r and q is assumed to be the same. Once in latent space, three separate networks are used to learn the energy terms; the potential (U), and kinetic energy (T) terms, and the Rayleigh damping (I) are each modeled with a multi-layer perceptron (MLP), as shown below.

$$U(\mathbf{q}) = U_{NN}(\mathbf{q}, \dot{\mathbf{q}}), \quad T(\mathbf{q}) = T_{NN}(\mathbf{q}, \dot{\mathbf{q}}), \quad I(\mathbf{q}) = R_{NN}(\mathbf{q}, \dot{\mathbf{q}})$$

The Lagrangian term is obtained from the potential and kinetic energy: $L(\mathbf{q}, \dot{\mathbf{q}}) = T_{NN}(\dot{\mathbf{q}}) - U_{NN}(\mathbf{q})$

From this, the equations of motion are enforced in the canonical coordinate system by first obtaining the system's momenta for each DOF, and constructing the system's Hamiltonian.

$$H(\mathbf{p}, \dot{\mathbf{q}}) = \mathbf{p} \cdot \dot{\mathbf{q}} - L(\mathbf{q}, \dot{\mathbf{q}})$$

The generalized force term is derived from d'Alembert's principle of virtual work [1], which results in projections of the force in the original coordinate system onto the partial derivative of the displacements in the original coordinate system with respect to the generalized displacements.

$$Q^{nc} = \sum_{i=1}^{N} F_i \cdot \frac{\partial D_{NN_i}}{\partial q_k}$$

The Euler-Lagrange equations of motion are then enforced by including the residual as a loss term:

$$Loss_{EL} = \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_k}\right) - \frac{\partial L}{\partial q_k} + \frac{\partial \mathrm{I}}{\partial \dot{q}_k} - Q^{nc}\right)^2$$

Furthermore, the network also enforces a time-derivative check by incorporating a multi-step integration term, allowing the ensemble of networks to be time-integrated with new initial conditions and an arbitrary external force after it has been trained.

$$Loss_F = \left\{ \sum_{j=0}^{s} \alpha_j \begin{pmatrix} \boldsymbol{q_{k+j}} \\ \dot{\boldsymbol{q}}_{\boldsymbol{k+j}} \end{pmatrix} - h \sum_{j=0}^{s} \beta_j \begin{pmatrix} \dot{\boldsymbol{q}}_{\boldsymbol{k+j}}(\boldsymbol{t}_{k+j}) \\ \ddot{\boldsymbol{q}}_{\boldsymbol{k+j}}(\boldsymbol{t}_{k+j}) \end{pmatrix} \right\}^2$$

RESULTS AND DISCUSSION

The approach is demonstrated on a two DOF damped oscillator with cubic nonlinearities. Each of the MLPs had four hidden layers, each with eight hidden units. A swish activation function was used for all the networks. The ensemble of networks were all trained simultaneously using a single response realization to a random force. The ensemble of networks was implemented using Tensorflow 2.5. Figure 2 shows the masses' response to a new random excitation, along with the predictions for the Lagrangian. The response was obtained by performing direct time integration on the trained network with scipy.integrate, and the results are presented after 6,000 time steps to show that the response does not degrade with time. The trained network was then subjected to sinusoidal excitation at varying amplitudes. The response at steady-state was obtained via direct time integration until transients decayed. The steady-state response was then plotted as a function of excitation amplitude to obtain nonlinear forced-response curves, which are shown in Figure 2 (right). As illustrated, the network was able to capture the stiffening effect typical of the cubic spring system.

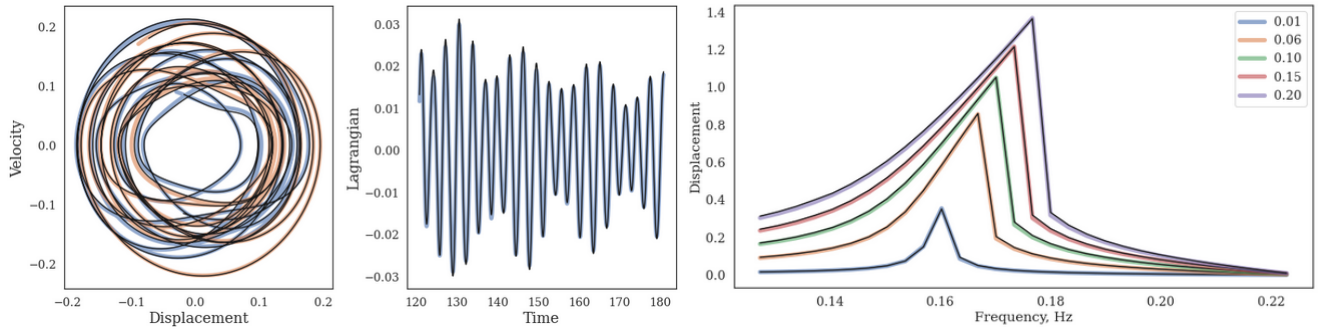

Figure 2. Phase portrait (left) and Lagrangian (center) time history showing the comparison between actual response (color) vs. neural network prediction (black). Nonlinear forced response curves for varying excitation amplitudes (right).

CONCLUSION

This work introduced an ensemble of neural networks that has been constrained by the mathematical and physical structure of Hamiltonian mechanics. Preliminary results show that the approach is robust and does not suffer from excessive error accumulation. It has also been found that limited data may be sufficient for training the network to be used outside of its training regime. These results are a direct consequence of the physical and numerical constraints, including the time integration residual loss term. Future work will extend the method to larger systems and to be fully Bayesian for use in a probabilistic reasoning framework to inform structural health diagnosis and prognosis.

ACKNOWLEDGEMENTS

This work was funded by Sandia National Laboratories, Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.

REFERENCES

[1] Cline, D. (2019). Variational Principles in Classical Mechanics: Revised Second Edition. River Campus Libraries.

[2] Abadi M. et al. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. Software available from tensorflow.org.