TDHNN: Time-Dependent Hamiltonian Neural Networks

Shaan A. Desai[1,2] **David Sondak**[2] **Marios Mathiakis**[2]
Machine Learning Research Group
University of Oxford [1]
Institute of Applied Computational Science
Harvard University [2]

Abstract

Deep networks embedded with physically-informed priors demonstrate remarkable performance in accurately learning and predicting non-linear dynamical systems. In particular, energy-conserving networks designed to exploit the Hamiltonian formalism show strong and consistent performance in learning autonomous systems that depend implicitly on time. Here, we extend this work to include an explicit time-dependence that generalizes the learning to include non-autonomous dynamical systems. We achieve this generalisation by embedding the port-Hamiltonian formalism into our neural network. We show that such a system can learn complex forced and damped dynamics, including the chaotic duffing equation, as well as maintain strong performance in settings that have no explicit time-dependence.

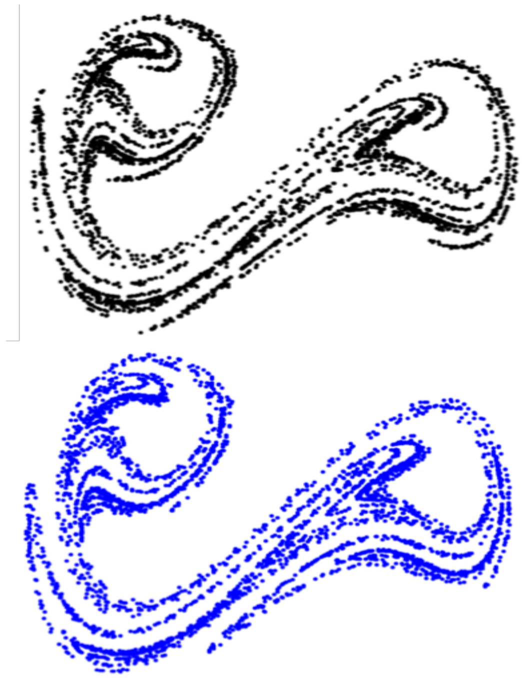

Figure 1: Poincaré Section of a duffing oscillator in a chaotic regime. Top is the ground truth and bottom is the predicted. Predictions are made using TDHNN which is trained on 2000 points.

1 Introduction

Neural networks, as universal function approximators, have shown resounding success across a host of domains. However, their performance in learning physical systems has often been limited. New research aimed at *scientific machine learning* - a branch that tackles scientific problems with domain-specific ML, is paving a way to address these challenges. It has been shown that prior theoretical information embedded in networks, such as Hamiltonian mechanics [?] demonstrate a significant performance uplift in learning. This excitement has spurred others to work with Lagrangians, ODEs and even graphs in order to tackle the learning of dynamical systems. Despite their widespread adoption, a major bottleneck of many of the existing methods is the lack of an explicit-time dependence as is evident across a host of forced dynamical systems. The most general form of Hamilton's equations, includes an explicit time dependence term. We show that the addition of this term, coupled with a few intuitive regularizations can induce networks to learn from both autonomous and non-autonomous settings. We extensively benchmark this addition across multiple datasets and consistently find the inclusion to be of benefit. Furthermore, we emphasise that the constraint is an easy plug-and-play addition to existing networks and illustrate how existing networks such as HNN, Symp ODEN and Hnets benefit from its inclusion.

Proceedings of the 24th International Conference on Artificial Intelligence and Statistics (AISTATS) 2021, San Diego, California, USA. PMLR: Volume 130.

2 Background

2.1 Hamiltonian Neural Networks

Recently, [?] demonstrated that dynamic predictions through time can be improved using Hamiltonian Neural Networks (HNNs) which endow models with a Hamiltonian constraint. The Hamiltonian is an important representation of a dynamical system because it is one of two well-known approaches that generalizes classical mechanics. The Hamiltonian $\mathcal{H}$ is a scalar function of position $\mathbf{q} = (q_1, q_2, ..., q_M)$ and momentum $\mathbf{p} = (p_1, p_2, ..., p_M)$. It is a powerful representation because it allows us to obtain the time derivatives of the inputs $(\dot{q}, \dot{p})$ by simply differentiating the Hamiltonian with respect to its inputs (see Eqn. 1.)

$$\frac{\mathrm{d}\mathbf{q}}{\mathrm{d}t} = \frac{\partial \mathcal{H}}{\partial \mathbf{p}}, \quad \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}} \tag{1}$$

As a consequence, it is noted in [?] that by parametrizing the Hamiltonian with a neural network e.g. $H_\theta(\mathbf{q}, \mathbf{p})$ where θ represents a deep neural network, one can easily obtain the system's dynamics by differentiating (via autograd) the Hamiltonian with its inputs. This information allows us to build two 1st-order differential equations which can be used to update the state space, $(\mathbf{q}, \mathbf{p})$. Equation 2 shows this integral, in which we define the symplectic gradient $\mathbf{S} = \left[\frac{\partial \mathcal{H}}{\partial \mathbf{p}}, -\frac{\partial \mathcal{H}}{\partial \mathbf{q}}\right]$:

$$(\mathbf{q}, \mathbf{p})_{t+1} = (\mathbf{q}, \mathbf{p})_t + \int_t^{t+1} \mathbf{S}(\mathbf{q}, \mathbf{p}) \mathrm{d}t \tag{2}$$

However, this is not the only benefit in learning a Hamiltonian. Another key attribute of the Hamiltonian is that the vector field $\mathbf{S}$ is a symplectic gradient meaning $\mathcal{H}$ remains constant as long as state vectors are integrated along $\mathbf{S}$. This result links the Hamiltonian with the total energy of the system such that $\mathcal{H}(\mathbf{q}, \mathbf{p}) = E_{tot}$ for many physical systems. Therefore, the Hamiltonian is a powerful inductive bias that can be utilised to evolve a physical state while maintaining energy conservation.

Although this formalism is compact and powerful, it does not readily generalize to damped or forced system. As such, we refer to port-Hamiltonian systems.

2.2 Port-Hamiltonians

Port-Hamiltonians are a formalism that allow us to incorporate damping and forcing terms. One formalism outlined in [?] shows how to represent a general form for such a system. We extend that work to include time-dependent forcing and eliminate the need for an explicit control input u. This results in the following equation:

$$\begin{bmatrix}\dot{\mathbf{q}} \\ \dot{\mathbf{p}}\end{bmatrix} = \left(\begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{I} & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \mathbf{D}_{\theta_2}(\mathbf{q}) \end{bmatrix} \right) \begin{bmatrix} \frac{\partial \mathcal{H}_{\theta_1}}{\partial \mathbf{q}} \\ \frac{\partial \mathcal{H}_{\theta_1}}{\partial \mathbf{p}} \end{bmatrix} + \begin{bmatrix} 0 \\ \mathbf{g}(\mathbf{q}) \end{bmatrix} F_{\theta_3}(t) \tag{3}$$

where $\mathbf{D}_{\theta_2}(q)$ is the damping term and $g(q)F_{\theta_3}(t)$ is the forcing term.

Given this general formalism, we make some simplifications. Here, instead of using a generalized semi-positive definite damping matrix, we simply aim at learning the lower right term which is most often independent of $\mathbf{q}$. As such, we parametrize $\mathbf{D}$ with a single, scalar learnable parameter. Secondly, for many physical systems, $g(q)$ is also a scalar quantity, so we abstract this scalar learning into $F(t)$ by setting $g(q) = 1$.

3 Method

$$Loss = \left\| \frac{\partial \mathcal{H}_{\theta_1}}{\partial \mathbf{q}} + \frac{\partial \mathbf{p}}{\partial t} \right\| + \left\| \frac{\partial \mathcal{H}_{\theta_1}}{\partial \mathbf{p}} - \frac{\partial \mathbf{q}}{\partial t} \right\| + \alpha_{reg}|F_{\theta_3}(t)| + \beta_{reg}|D_{\theta_2}| \tag{4}$$

To learn the dynamics, we feed in a state-vector $S_t = [\mathbf{q}, \mathbf{p}, \mathbf{t}]$. The first neural-network consists of 3 hidden layers designed to predict H from $[q, p]$ data. The second neural-network consists of a single weight parameter designed to learn the damping coefficient D and the third neural-network consists of 2 hidden layers designed to predict $F(t)$ from t. We use an L2-norm penalty for the predicted state-vectors and an L1-norm for force and damping to encourage sparsity. We do this because we would like our networks to identify classical autonomous systems (which may not have force or time) as well as non-autonomous systems. For our experiments we use 200 hidden layers and find that most activations such as tanh, sin and cos yield comparable results. To benchmark our method, as [?] do, we use a baseline NN that takes in S_t and predicts $[\dot{q}, \dot{p}]$. We also take the straightforward extension of HNN to include time as a variable input.

4 FIRST LEVEL HEADINGS

First level headings are all caps, flush left, bold, and in point size 12. Use one line space before the first level heading and one-half line space after the first level heading.

4.1 Second Level Heading

Second level headings are initial caps, flush left, bold, and in point size 10. Use one line space before the second level heading and one-half line space after the second level heading.

4.1.1 Third Level Heading

Third level headings are flush left, initial caps, bold, and in point size 10. Use one line space before the third level heading and one-half line space after the third level heading.

Fourth Level Heading Fourth level headings must be flush left, initial caps, bold, and Roman type. Use one line space before the fourth level heading, and place the section text immediately after the heading with no line break, but an 11 point horizontal space.

4.2 Citations, Figure, References

4.2.1 Citations in Text

Citations within the text should include the author's last name and year, e.g., (Cheesman, 1985). Be sure that the sentence reads correctly if the citation is deleted: e.g., instead of "As described by (Cheesman, 1985), we first frobulate the widgets," write "As described by Cheesman (1985), we first frobulate the widgets."

The references listed at the end of the paper can follow any style as long as it is used consistently.

4.2.2 Footnotes

Indicate footnotes with a number[1] in the text. Use 8 point type for footnotes. Place the footnotes at the bottom of the column in which their markers appear, continuing to the next column if required. Precede the footnote section of a column with a 0.5 point horizontal rule 1 inch (6 picas) long.[2]

4.2.3 Figures

All artwork must be centered, neat, clean, and legible. All lines should be very dark for purposes of reproduction, and art work should not be hand-drawn. Figures may appear at the top of a column, at the top of a page spanning multiple columns, inline within a column, or with text wrapped around them, but the figure number and caption always appear immediately below the figure. Leave 2 line spaces between the figure and the caption. The figure caption is initial caps and each figure should be numbered consecutively.

Make sure that the figure caption does not get separated from the figure. Leave extra white space at the bottom of the page rather than splitting the figure and figure caption.

[1] Sample of the first footnote.

[2] Sample of the second footnote.

This figure intentionally left non-blank

Figure 2: Sample Figure Caption

4.2.4 Tables

All tables must be centered, neat, clean, and legible. Do not use hand-drawn tables. Table number and title always appear above the table. See Table 1.

Use one line space before the table title, one line space after the table title, and one line space after the table. The table title must be initial caps and each table numbered consecutively.

Table 1: Sample Table Title

PART	DESCRIPTION
Dendrite	Input terminal
Axon	Output terminal
Soma	Cell body (contains cell nucleus)

5 SUPPLEMENTARY MATERIAL

If you need to include additional appendices during submission, you can include them in the supplementary material file. You can submit a single file of additional supplementary material which may be either a pdf file (such as proof details) or a zip file for other formats/more files (such as code or videos). Note that reviewers are under no obligation to examine your supplementary material. If you have only one supplementary pdf file, please upload it as is; otherwise gather everything to the single zip file.

You must use `aistats2021.sty` as a style file for your supplementary pdf file and follow the same formatting instructions as in the main paper. The only difference is that it must be in a *single-column* format. You can use `supplement.tex` in our starter pack as a starting point. Alternatively, you may append the supplementary content to the main paper and split the final PDF into two separate files.

6 SUBMISSION INSTRUCTIONS

To submit your paper to AISTATS 2021, please follow these instructions.

1. Download `aistats2021.sty`, `fancyhdr.sty`, and

`sample_paper.tex` provided in our starter pack. Please, do not modify the style files as this might result in a formatting violation.

2. Use `sample_paper.tex` as a starting point.
3. Begin your document with

```
\documentclass[twoside]{article}
\usepackage{aistats2021}
```

The `twoside` option for the class article allows the package `fancyhdr.sty` to include headings for even and odd numbered pages.

4. When you are ready to submit the manuscript, compile the latex file to obtain the pdf file.
5. Check that the content of your submission, *excluding* references, is limited to **8 pages**. The number of pages containing references alone is not limited.
6. Upload the PDF file along with other supplementary material files to the CMT website.

6.1 Camera-ready Papers

If your papers are accepted, you will need to submit the camera-ready version. Please make sure that you follow these instructions:

1. Change the beginning of your document to

```
\documentclass[twoside]{article}
\usepackage[accepted]{aistats2021}
```

The option `accepted` for the package `aistats2021.sty` will write a copyright notice at the end of the first column of the first page. This option will also print headings for the paper. For the *even* pages, the title of the paper will be used as heading and for *odd* pages the author names will be used as heading. If the title of the paper is too long or the number of authors is too large, the style will print a warning message as heading. If this happens additional commands can be used to place as headings shorter versions of the title and the author names. This is explained in the next point.

2. If you get warning messages as described above, then immediately after `\begin{document}`, write

```
\runningtitle{Provide here an alternative
shorter version of the title of your
paper}
\runningauthor{Provide here the surnames
of the authors of your paper, all
separated by commas}
```

Note that the text that appears as argument in `\runningtitle` will be printed as a heading in the *even* pages. The text that appears as argument in `\runningauthor` will be printed as a heading in the *odd* pages. If even the author surnames do not fit, it is acceptable to give a subset of author names followed by "et al."

3. The camera-ready versions of the accepted papers are 8 pages, plus any additional pages needed for references.
4. If you need to include additional appendices, you can include them in the supplementary material file.
5. Please, do not change the layout given by the above instructions and by the style file.

Acknowledgements

All acknowledgments go at the end of the paper, including thanks to reviewers who gave useful comments, to colleagues who contributed to the ideas, and to funding agencies and corporate sponsors that provided financial support. To preserve the anonymity, please include acknowledgments *only* in the camera-ready papers.