Neural Latent Dynamics Models

Nicola Farenga
Politecnico di Milano
nicola.farenga@mail.polimi.it

Stefania Fresca
Politecnico di Milano
stefania.fresca@polimi.it

Andrea Manzoni
Politecnico di Milano
andrea1.manzoni@polimi.it

Abstract

We introduce Neural Latent Dynamics Models (NLDMs), a neural ordinary differential equations (ODEs)-based architecture to perform end-to-end nonlinear latent dynamics discovery, without the need to include any inductive bias related to either the underlying physical model or the latent coordinates space. The effectiveness of this strategy is experimentally tested in the framework of reduced order modeling, considering a set of problems involving high-dimensional data generated from nonlinear time-dependent parameterized partial differential equations (PDEs) simulations, where we aim at performing extrapolation in time, to forecast the PDE solution out of the time interval and/or the parameters range where training data were acquired. Results highlight NLDMs' capabilities to perform low-dimensional latent dynamics learning in three different scenarios.

1 Introduction

The progress in the field of data-driven model discovery has been recently pushed by the adoption of deep learning-based methods (Raissi [2018], Champion et al. [2019]), by relying on autoencoders architectures to perform nonlinear dimensionality reduction; this latter represents a key step in view of learning low-dimensional latent coordinates systems. Recent advancements in the field of reduced order modeling also started to consider deep learning methods, with the introduction of hybrid techniques (Fresca and Manzoni [2022]), combining standard methods such as proper orthogonal decomposition (POD), together with autoencoders, allowing to overcome the main limitations of traditional projection-based techniques, related to their poor performances when modeling nonlinear phenomena. Classical model discovery methods usually require assumptions on the involved coordinates, prior knowledge about the underlying physical model, typically enforced via specific formulations of the loss function, and/or libraries of candidate functions and their derivatives to perform sparse regression on (Rudy et al. [2017], Chen et al. [2021]). All these assumptions allow to shed light on model interpretability, but might be difficult to fulfill in real-world scenarios due to data scarcity, or because they imply infeasible computational loads.

The concept of neural ODEs (Chen et al. [2018], Kidger [2022]) represents a bridge between dynamical systems and deep learning models, naturally leading to the formulation of latent continuous dynamics discovery frameworks. In this contribution we propose Neural Latent Dynamics Models (NLDMs), a neural ODEs-based continuous-time architecture, to perform end-to-end latent dynamics learning, as first-order ODEs, without the need to incorporate any inductive bias. We test the proposed architecture in reduced order modeling settings, applying the framework to multi-dimensional, linear and nonlinear, time-dependent parameterized PDEs, such as a 1D Burgers' equation, a 2D heat equation, and a 3D nonlinear parameterized elastodynamics problem.

DLDE Workshop in the 36th Conference on Neural Information Processing Systems (NeurIPS 2022).

2 Methods

In this section we describe the general formulation of Neural Latent Dynamics Models (NLDMs) and the architecture for their employment in a reduced order modeling setting.

2.1 Neural Latent Dynamics Models

NLDMs combine autoencoders architecture together with neural ODEs to learn the latent dynamics of the observed time-dependent data. Recently, multiple time-series modeling approaches combining autoencoders and neural ODEs have been proposed, like generative latent ODEs (Chen et al. [2018]) and ODE-RNNs (Rubanova et al. [2019]). Differently from these strategies, our approach is non-variational, due to the need to perform deterministic dimensionality reduction as a parallel end goal, and tries to decouple the dimensionality reduction stage from the model discovery one, by defining an architecture made by *(i)* an encoder, *(ii)* a latent neural ODE, and *(iii)* a decoder. The proposed architecture is presented in a general setting, since it can also be employed to perform standard time-series modeling, by using the latent dynamics module only for time advancements, and disregarding any interest in performing latent model discovery.

The general NLDMs framework employs three learnable functions, an encoder $\mathcal{E}$, a decoder $\mathcal{D}$, and a latent dynamics function f depending on a vector of parameters θ, to model the observable high-dimensional state $\mathbf{u}(t) \in \mathbb{R}^m$ evolution, given the initial state $\mathbf{u}(0)$, as follows:

$$\mathbf{z}(0) = \mathcal{E}(\mathbf{u}(0)), \tag{1}$$

$$\mathbf{z}(t) = \mathbf{z}(0) + \int_0^t f(\mathbf{z}(\tau), \tau; \theta) d\tau \quad t \in [0, T], \tag{2}$$

$$\hat{\mathbf{u}}(t) = \mathcal{D}(\mathbf{z}(t)). \tag{3}$$

Steps (1) and (3) regard the dimensionality reduction task, via the encoder and the decoder, respectively, while step (2) represents the integration of the latent dynamics function f, linked to the time evolution of the latent state $\mathbf{z}(t) \in \mathbb{R}^n$, with $n \ll m$. Operating with discrete observations, the integration step (2) is solved numerically via iterative methods, considering a discretization $t_0, ..., t_N$ of the time interval

$$\mathbf{z}(t_0), ..., \mathbf{z}(t_N) = \texttt{odesolve}(f, \mathbf{z}(t_0), (t_0, ..., t_N)), \tag{4}$$

with the corresponding formulation of the decoding step (3) that reads as

$$\hat{\mathbf{u}}(t_i) = \mathcal{D}(\mathbf{z}(t_i)), \quad \forall i = 0, ..., N. \tag{5}$$

The training procedure consists of minimizing a loss function given by sum of the squared L^2-errors between the observed and learned trajectories,

$$\mathcal{L} = \sum_{i=1}^{M} \sum_{j=0}^{N^*} \|\mathbf{u}(t_j^i) - \hat{\mathbf{u}}(t_j^i)\|_2^2, \tag{6}$$

given a dataset of $N + 1$ sequential observations $\{t_i, \mathbf{u}(t_i)\}_{i=0}^N$, from which M different sub-trajectories $\{\mathbf{u}(t_0^j), ..., \mathbf{u}(t_{N^*}^j)\}_{j=1}^M$ of length $N^* + 1$ have been sampled, with $N^* \leq N$.

2.2 NLDMs for Reduced Order Modeling

Regarding the way NLDMs can be exploited in a reduced order modeling setting, our approach builds up from Fresca and Manzoni [2022], since our aim is to keep a modular structure of the overall model, leading to more flexibility when dealing with high-dimensional data, allowing to eventually employ pre-trained sub-modules. Considering a generic nonlinear parameterized time-dependent PDE, the associated full-order model (FOM) obtained after space discretization can be expressed in the form of a nonlinear parameterized dynamical system that reads as

$$\begin{cases} \mathrm{M}(\mu)\dot{\mathbf{u}}_h(t; \mu) = \mathbf{f}(t, \mathbf{u}_h(t; \mu); \mu), & t \in (0, T) \\ \mathbf{u}_h(0; \mu) = \mathbf{u}_0(\mu) \end{cases}, \tag{7}$$

where $\mathbf{u}_h(t; \mu) \in \mathbb{R}^{N_h}$ represents the time-dependent parameterized solution, $\mathrm{M}(\mu) \in \mathbb{R}^{N_h \times N_h}$ is the mass matrix, $\mathbf{f} : (0, T) \times \mathbb{R}^{N_h} \times \mathcal{P} \to \mathbb{R}^{N_h}$ is a nonlinear function encoding the system dynamics.

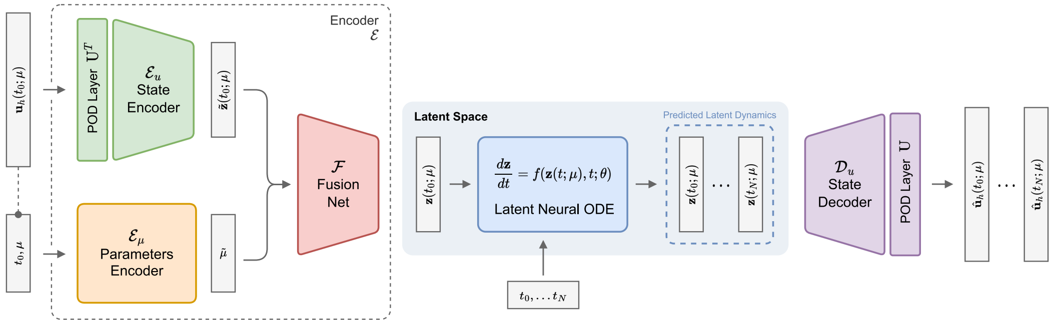

Figure 1: A schematic view of the NLDM architecture employed in a reduced order modeling setting, highlighting the structure of the encoder module $\mathcal{E}$, made by three sub-modules (a state encoder $\mathcal{E}_u$, a parameters encoder $\mathcal{E}_\mu$, and a fusion net $\mathcal{F}$) to produce the parameterized encoding of the initial latent state $\mathbf{z}(t_0; \mu)$.

$\mathcal{P}$ is the input parameters set to which μ belongs, while N_h denotes the FOM dimension, related to the space discretization parameter $h > 0$. Formulation (7) justifies the search for a nonlinear latent dynamics describing the evolution of $\mathbf{u}_h(t; \mu)$.

A common requirement, while performing dimensionality reduction of time-dependent parameterized PDEs, is to include the information coming from the FOM parameters into the reduced order model (ROM). As shown in Figure 1, in our setting this is achieved by employing a *fusion network* $\mathcal{F}$, to combine the concatenated information, previously encoded by the *state encoder* $\mathcal{E}_u$ and the *parameters encoder* $\mathcal{E}_\mu$, into the initial latent state $\mathbf{z}(t_0; \mu) \in \mathbb{R}^n$, input to the latent neural ODE, where n represents the dimensionality of the reduced manifold. Both the parameters encoder $\mathcal{E}_\mu$ and the fusion network $\mathcal{F}$ are feed-forward neural networks, while the state encoder $\mathcal{E}_u$ and decoder $\mathcal{D}_u$ architectures could differ, depending on the nature of the problem and the complexity of the observed state. Moreover, a hybrid dimensionality reduction technique has been adopted to deal with high-dimensional FOMs states, by employing the POD basis matrix U extracted from a snapshots matrix $\mathrm{S} = [\mathbf{u}_h(t_0; \mu)| \cdots |\mathbf{u}_h(t_N; \mu)]$, through randomized singular value decomposition (SVD).

3 Numerical Experiments

NLDMs capabilities have been tested in a reduced order modeling framework, on datasets coming from PDEs discretizations, since they usually involve complex time and spatial-dependent dynamics.

Burgers' Equation The first test case, inspired by Raissi [2018], consists of modeling the dynamics of the following 1D Burgers' equation

$$u_t + uu_x - \nu u_{xx} = 0, \quad u(0, x) = -\sin(\pi x/8), \quad x \in (-8, 8), \quad t \in [0, 10] \tag{8}$$

discretized with $N_h = 256$ elements in space, over $N_t = 200$ time steps, for a fixed instance of the parameter $\nu = 0.1$. The model's architecture followed the structure depicted in Figure 1, with a latent state dimension $n = 2$, and a latent dynamics function consisting of one hidden layer with 32 units and $\tanh$ activation. The integration was performed by using fifth-order Dormand-Prince adaptive-step solver, from `torchdiffeq` library (Chen [2018]).

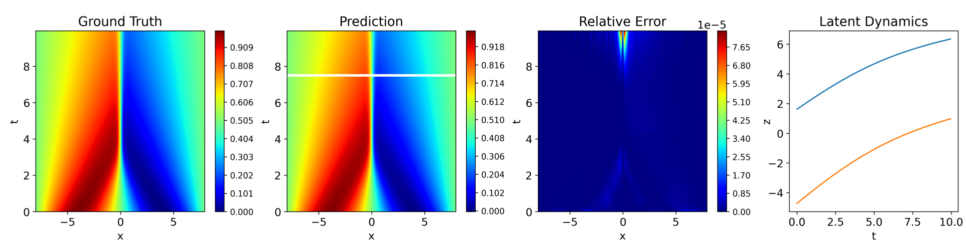

Figure 2: Burgers' equation FOM solution (left), NLDM predicted solution (center left) with relative error (center right), and learned latent dynamics (right). The horizontal line in the prediction plot at $t = 7.5s$ highlights the start of time extrapolation.

As shown in Figure 2, the method is able to capture the dynamics with a relative error of 6.33e-07 in the training time interval [0,7.5)s, and is able to extrapolate in time, on the remaining test interval [7.5,10]s, with a relative error of 4.95e-06. Moreover, time extrapolation capabilities have been tested over a much larger timespan, extending the interval to [0,40]s, achieving an overall relative error of 8.35e-05. The adoption of adaptive-step solvers improve NLDM's stability and extrapolation capabilities over longer time horizons, at the cost of a slower training procedure.

Heat Equation The second test concerns modeling the latent dynamics of a 2D heat equation, for a fixed instance of thermal diffusivity $\alpha = 1$

$$u_t = \alpha \Delta u, \quad u(0, x, y) = \frac{1}{2} e^{-(x^2+y^2)}, \quad (x, y) \in \Omega = [-10, 10]^2, \quad t \in [0, 4] \tag{9}$$

with a FOM dimension of $N_h = 64^2$, and $N_t = 400$. Since this example involves a 2D domain, the state encoder $\mathcal{E}_u$ and decoder $\mathcal{D}_u$ both feature a convolutional architecture. The adopted latent state dimension is $n = 2$, with a latent dynamics function consisting of two hidden layers with 256 units each, and $\tanh$ activation, integrated via forward Euler method.

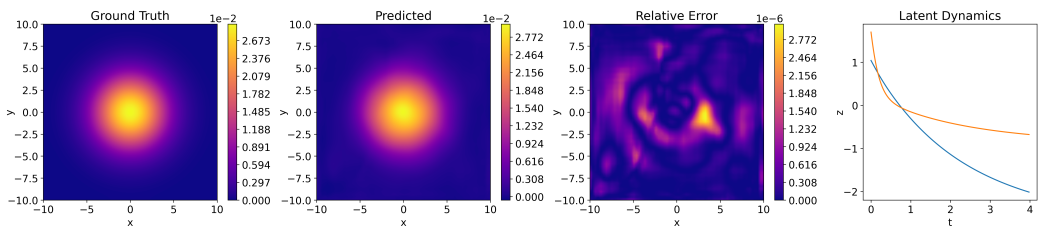

Figure 3: Heat equation FOM solution (left) and NLDM predicted solution for $t = 4s$ (center left), relative error (center right), learned latent dynamics (right).

The NLDM learned a two-dimensional dynamics with a mean relative error of 1.81e-07 on the training [0,3)s interval, and of 1.74e-08 when performing time extrapolation over the interval [3,4]s, whose last step is shown in Figure 3, together with the latent dynamics evolution over the complete interval.

Parameterized Elastodynamics Problem The last application regards a 3D parameterized nonlinear elastodynamics problem for a restrained micro-beam subject to internal piezoelectric actuation force, with FOM dimension $N_h = 7821$, $N_t = 400$ time-steps over the interval [0,25.32]s and 10 instances of the two parameters controlling amplitude and frequency of the piezoelectric oscillating force. Again, latent state dimension is $n = 2$, while the latent dynamics net features one hidden layer with 256 units and $\tanh$ activation, integrated via forward Euler method. The model learned a continuous periodic latent dynamics with a mean relative error of 3.9e-03 on the training interval [0,19)s, and 6.3e-03 when performing time extrapolation considering the whole interval [0,25.32]s, averaged over all the 10 parameter instances. Moreover, when extrapolating out of the training parameters' range, the NLDM achieved a mean relative error of 7.8e-03.

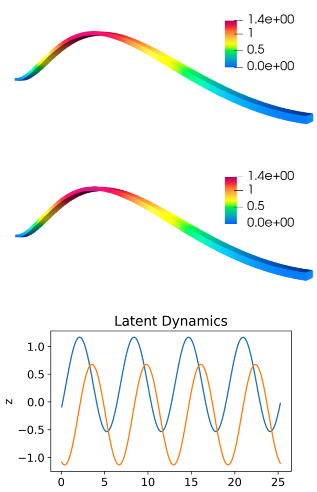

Figure 4: Elastodynamics problem FOM solution (top) and NLDM predicted solution (middle) for $t = 25.32s$ and parameters instance $(-0.9782, 0.25)$, learned latent dynamics (bottom).

4 Conclusions

We introduced NLDMs, a model architecture combining autoencoders and neural ODEs, to perform end-to-end latent dynamics discovery of high-dimensional systems, without the need of prior assumptions or knowledge about the physics governing the problem. We investigated NLDMs effectiveness in the field of reduced order modeling, by defining a specific encoder structure to deal with parameterized problems, demonstrating their suitability for learning low-dimensional dynamics of nonlinear time-dependent parameterized PDEs systems.

References

K. Champion, B. Lusch, J. N. Kutz, and S. L. Brunton. Data-driven discovery of coordinates and governing equations. *Proceedings of the National Academy of Sciences*, 116(45):22445–22451, 2019.

R. T. Q. Chen. `torchdiffeq`, 2018. URL `https://github.com/rtqichen/torchdiffeq`.

R. T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. K. Duvenaud. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.

Z. Chen, Y. Liu, and H. Sun. Physics-informed learning of governing equations from scarce data. *Nature Communications*, 12(1):6136, 2021.

S. Fresca and A. Manzoni. POD-DL-ROM: enhancing deep learning-based reduced order models for nonlinear parametrized PDEs by proper orthogonal decomposition. *Computer Methods in Applied Mechanics and Engineering*, 388:114181, 2022.

P. Kidger. On neural differential equations. *CoRR*, abs/2202.02435, 2022. URL `https://arxiv.org/abs/2202.02435`.

M. Raissi. Deep hidden physics models: Deep learning of nonlinear partial differential equations. *The Journal of Machine Learning Research*, 19(1):932–955, 2018.

Y. Rubanova, R. T. Chen, and D. K. Duvenaud. Latent ordinary differential equations for irregularly-sampled time series. *Advances in neural information processing systems*, 32, 2019.

S. H. Rudy, S. L. Brunton, J. L. Proctor, and J. N. Kutz. Data-driven discovery of partial differential equations. *Science advances*, 3(4):e1602614, 2017.