Learning to Control PDEs with Differentiable Physics

Anonymous Author(s)
Affiliation
Address
email

1 Introduction

Understanding physical environments is a key requirement for machine learning applications such as autonomous agents and robots [9, 2]. It is typically of vital importance to not only understand the unperturbed physical behavior but also anticipate how the environment reacts to an agent interacting with it [15, 7]. We consider partial differential equations (PDEs) as the most fundamental description of physical systems. The language of PDEs is general enough to describe every physical theory, from quantum mechanics and general relativity to turbulent flows [14]. Existing machine learning methods that deal with agents learning to interact with their environments have often focused on reinforcement learning [11, 6], but for high-dimensional environments, the computational cost of exploring the state space puts severe limits on the number of interaction parameters with which the agent can influence the physical system [10].

Meanwhile, progress has been made in utilizing differentiable solvers to find solutions to high-dimensional optimization problems [15, 5, 13]. Yet existing methods are still computationally expensive and thus limited to short time frames. We combine differentiable physics with deep learning to represent solution manifolds rather than computing single solutions via optimization. In this way, trained models can interact with a physical environment using a large number of interaction parameters, and inference times are orders of magnitude faster than with classic optimization algorithms. Here the use of differentiable physics is key for a robust learning of the complex spaces of behavior encoded by the model PDEs.

We employ a fully differentiable Eulerian PDE solver that can solve a large class of PDEs with analytic gradients. By fully integrating the numerical solver into the training process, neural networks can learn how to optimally control a physical system given an initial state and a target state. We further demonstrate that long time frames can be handled via a specialized architecture and evaluation scheme that separates the learning of physical behavior for different time scales. The resulting technique uses multiple neural networks, sharing the same architecture, and enables the inference of solutions to an optimal control problem for a sequence of length n in time $\mathcal{O}(n)$.

2 Differentiable PDE solvers

Let $\boldsymbol{u}(\boldsymbol{x}, t)$ be described by a PDE that can be explicitly solved forward in time, i.e. time and space derivatives do not mix. The PDE can then be written as

$$\frac{\partial \boldsymbol{u}}{\partial t} = \mathcal{P}\left(\boldsymbol{u}, \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{x}}, \frac{\partial^2 \boldsymbol{u}}{\partial \boldsymbol{x}^2}, ..., \boldsymbol{y}(t)\right) \tag{1}$$

where $\mathcal{P}$ models the physical behavior of the system and $\boldsymbol{y}(t)$ denotes any external factors that can influence the system. A classic solver can move the system forward in time via Euler steps:

$$\boldsymbol{u}(t_{i+1}) = \text{Solver}[\boldsymbol{u}(t_i), \boldsymbol{y}(t_i)] = \boldsymbol{u}(t_i) + \Delta t \cdot \mathcal{P}\left(\boldsymbol{u}(t_i), ..., \boldsymbol{y}(t_i)\right) \tag{2}$$

The square brackets indicate that Solver is a functional rather than a function, i.e. it takes full fields as input. Each step moves the system forward by a time increment Δt. Repeated execution produces a trajectory $u(t)$ that is a solution to the PDE.

This functionality for time advancement by itself is not well-suited to solve optimization problems, since gradients can only be approximated by finite differencing in these solvers. For high-dimensional

Submitted to the ICLR 2020 Workshop "Integration of Deep Neural Models and Differential Equations".

or continuous systems, this method becomes computationally expensive because a full trajectory needs to be computed for each optimizable parameter. Differentiable solvers resolve this issue by solving the adjoint problem [12] via analytic derivatives. The adjoint problem computes the same mathematical expressions while working with lower-dimensional vectors. A differentiable solver can efficiently compute the derivatives with respect to any of its inputs, i.e. $\partial \boldsymbol{u}(t_{i+1})/\partial \boldsymbol{u}(t_i)$ and $\partial \boldsymbol{u}(t_{i+1})/\partial \boldsymbol{y}(t_i)$. This allows for gradient-based optimization of inputs or control parameters of the simulation over an arbitrary number of time steps. The adjoint method is also used by most machine learning frameworks, where it is more commonly known as reverse mode differentiation [16, 4].

We make use of this analogy to implement a differentiable PDE solver as a set of mathematical operations within a deep learning framework [1]. We focus on Eulerian rather than Lagrangian methods since they are widely used for a large class of PDEs [14]. All solver operations are implemented in a differentiable manner, i.e. the automatic differentiation tools can chain the derivatives of these operations with built-in machine learning operations to build analytic derivatives for any combination of operations, thus enabling end-to-end training. This toolkit of operations enables the solver to handle a large class of PDEs, including the incompressible Navier-Stokes equations.

3 Learning force-based interactions

Assuming the physical behavior $\mathcal{P}$ is described by a PDE as in Eq. (1), we add a control force $\boldsymbol{F}(t)$ which allows the model to interact with the system:

$$\frac{\partial \boldsymbol{u}}{\partial t} = \mathcal{P}\left(\boldsymbol{u}, \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{x}}, \frac{\partial^2 \boldsymbol{u}}{\partial \boldsymbol{x}^2}, ...\right) + \boldsymbol{F}(t) \tag{3}$$

While the evolution of the complete state $\boldsymbol{u}$ is determined by the above equation, we allow some parts of $\boldsymbol{u}$ to be hidden for the control task. This restriction reflects the fact that it is often not possible to observe the full state of a physical system. When considering a cloud of smoke, for example, the smoke density might be observable while the velocity field cannot be seen directly. Mathematically, we model this restriction by decomposing $\boldsymbol{u}$ into an observable part $\boldsymbol{o}$ and a hidden part $\boldsymbol{h}$ with $\boldsymbol{u} = \boldsymbol{o}(\boldsymbol{u}) \otimes \boldsymbol{h}(\boldsymbol{u})$. Here, $\otimes$ denotes the tensor product, adding all components of the states. The hidden part can include spatial regions of some fields as well as entire fields.

Using the above notation, we define the control task as follows. An initial observable state $\boldsymbol{o}_0$ of the PDE as well as a target state $\boldsymbol{o}^*$ are given. We are interested in a reconstructed trajectory $\boldsymbol{u}^r(t)$ that matches these states at t_0 and t_*, i.e. $\boldsymbol{o}_0 = \boldsymbol{o}(\boldsymbol{u}^r(t_0)), \boldsymbol{o}^* = \boldsymbol{o}(\boldsymbol{u}^r(t_*))$, and requires the least amount of effort over the whole time span. I.e., we aim for minimizing the forces to be applied in terms of their magnitude with:

$$L_{\boldsymbol{F}}[\boldsymbol{u}(t)] = \int_{t_0}^{t_*} |\boldsymbol{F}_{\boldsymbol{u}}(t)|^2 \, dt \tag{4}$$

Taking discrete time steps Δt, the reconstructed trajectory $\boldsymbol{u}^r$ is a sequence of $n = (t_* - t_0)/\Delta t$ states. This problem definition is portrayed in Fig. 1. An initial observation $\boldsymbol{o}_0$ and target observation $\boldsymbol{o}_*$ are given (a). The goal is to reconstruct a trajectory $\boldsymbol{u}^r$ that moves from $\boldsymbol{o}_0$ to $\boldsymbol{o}_*$ in the state space and requires as little force as possible, as shown in (b). The grey lines represent the unperturbed evolution of the physical system. The amount of applied force corresponds to how far the trajectory deviates from the natural evolution in this picture.

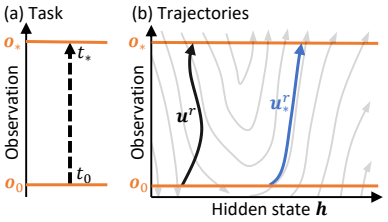

Figure 1: Possible trajectories.

When an observable dimension cannot be controlled directly, there may not exist any trajectory $\boldsymbol{u}(t)$ that matches both $\boldsymbol{o}_0$ and $\boldsymbol{o}^*$. This can stem from either physical constraints or numerical limitations. In these cases, we settle for an approximation of $\boldsymbol{o}^*$. To measure the quality of the approximation of the target, we define an observation loss $L^*_{\boldsymbol{o}}$. The form of this loss can be chosen to fit the problem. For our experiments we use the filtered L_2 distance between target and reconstruction:

$$L^*_{\boldsymbol{o}}(\boldsymbol{u}(t_*)) = |B_r(\boldsymbol{o}^*) - B_r\left(\boldsymbol{o}(\boldsymbol{u}(t_*))\right)|^2 \tag{5}$$

where B_r denotes a spatial blur function with a fixed, problem-dependent radius $r \geq 0$. We combine Eqs. 4 and 5 into the objective loss function

$$L[u(t)] = \alpha \cdot L_{\boldsymbol{F}}[\boldsymbol{u}(t)] + \beta \cdot L^*_{\boldsymbol{o}}(u(t_*)), \tag{6}$$

with $\alpha, \beta > 0$. Since our solver is differentiable, L can be used directly to optimize a machine learning model such as a neural network that models $\boldsymbol{u}^r(t), \boldsymbol{o}_*, t \rightarrow \boldsymbol{F}(t)$ with weights $\boldsymbol{w}$. We call this network the control force estimator (CFE).

For a sequence of n frames, $L[\boldsymbol{u}(t)]$ depends on all n states of the trajectory $\boldsymbol{u}(t)$. Thus, for recurrent end-to-end training, n linked copies of the network need to be chained together. When inferring the force, this results in a CFE chain, shown in Fig. 2, that alternates between network and solver execution. When using a CFE chain, the complete sequence needs to be run forward and backward for each optimization step of the model. This is not only slow, it also means that gradients are passed through a potentially long chain of highly non-linear simulation steps. When the reconstruction u^r is close to an optimal trajectory, this is not a problem since the gradients $\Delta \boldsymbol{u}^r$ are small and the operations executed by the solver are differentiable by construction. The solver can therefore be locally approximated by a first-order polynomial and the gradients can be safely backpropagated. For large $\Delta \boldsymbol{u}^r$, such as at the beginning of training, this approximation breaks down, causing the gradients to become highly unstable while passing through the chain.

This workshop paper can only provide a summary of our approach – in the full version [8], we give details on how a divide-and-conquer scheme can be used to resolve this problem so that the feedback from a differentiable solver leads to stable convergence in training. In this version we employ a second model, which predicts the observable state $\boldsymbol{o}^p\left((t_i + t_j)/2\right)$ given two observations. We refer to this model as the observation predictor (OP).

4 Results

We apply our algorithm to two-dimensional fluid dynamic problems, which are highly challenging due to the complexities on the governing Navier-Stokes equations [3] for the velocity field $\boldsymbol{v}$,

$$\mathcal{P}(\boldsymbol{v}, \nabla \boldsymbol{v}) = -\boldsymbol{v} \cdot \nabla \boldsymbol{v} + \nu \nabla^2 \boldsymbol{v} + \nabla p, \tag{7}$$

subject to the hard constraints $\nabla \cdot \boldsymbol{v} = 0$ and $\nabla \times p = 0$, where p denotes pressure and ν the viscosity. In addition, we consider a passive density ρ which moves with the fluid via $\partial \rho / \partial t = -\boldsymbol{v} \cdot \nabla \rho$. We set $\boldsymbol{v}$ to be hidden and ρ to be observable and allow forces to be applied to all of $\boldsymbol{v}$.

We run our tests on a 128×128 grid, resulting in more than 16,000 effective continuous control parameters. We train the OP and CFE networks for two different tasks: reconstruction of natural fluid flows and controlled shape transitions. Example sequences are shown in Fig. 3 and a quantitative evaluation, averaged over 100 examples, is given in Tab. 1. While all divide-and-conquer methods manage to approximate the target state well, there are considerable differences in the amount of force applied. The supervised technique, denoted as *regular*, exerts significantly more force than the differentiable solver based methods, resulting in jittering reconstructions. A prediction refinement scheme (denoted as *refined*) re-evaluates predictions over the course of a sequence. This version produces the smoothest transitions, converging to about half the loss of the regular, non-refined variant. For comparison, we run a classic optimization with hierarchical shooting that computes solutions for single cases, and find that it requires 1500 iterations to compute a control function that our trained model infers almost instantly.

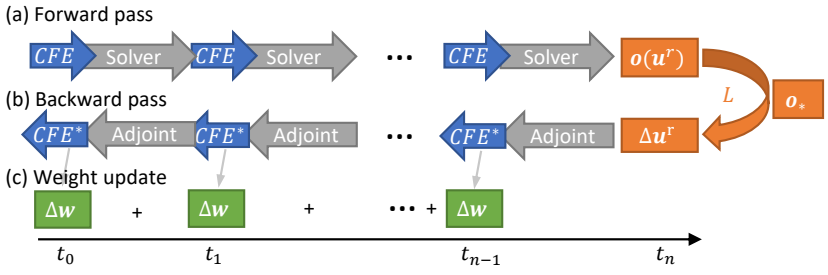

Figure 2: Optimization scheme of a chained force prediction network. (a) The forward pass reconstructs a trajectory by alternating between force estimation and solver execution. (b) For backpropagation, the adjoint problem of the sequence is computed. (c) The weight updates from each time step are accumulated and applied to the model.

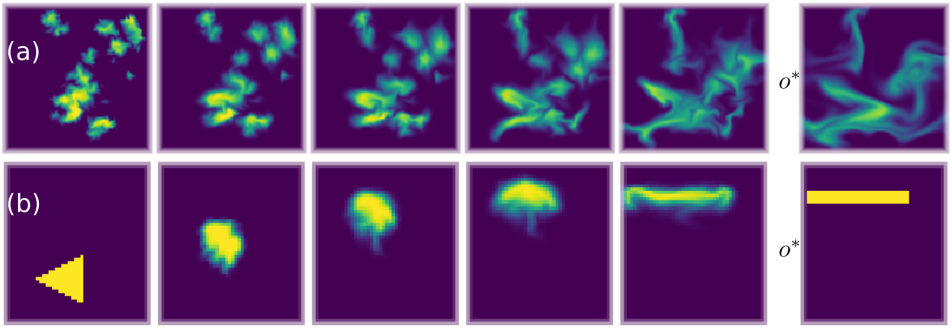

Figure 3: Example reconstructed trajectory from (a) the natural flow test set and (b) the shape test set. The target state $\boldsymbol{o}^*$ is shown on the right.

Table 1: A comparison of methods in terms of final cost for the (a) natural flow setup and (b) the shape transitions. The initial distribution is sampled randomly and evolved to the target state.

Execution	Loss	a) Force $L_{\boldsymbol{F}}$	a) Obs. $L_{\boldsymbol{o}}^*$	b) Force $L_{\boldsymbol{F}}$	b) Obs. $L_{\boldsymbol{o}}^*$
Regular	Supervised	243 ± 11	1.53 ± 0.23	n/a	n/a
Regular	Diff. Physics	22.6 ± 1.1	0.64 ± 0.08	89 ± 6	0.331 ± 0.134
Refined	Diff. Physics	11.7 ± 0.6	0.88 ± 0.11	75 ± 4	0.126 ± 0.010

The next experiment increases the complexity of the fluid control problem by adding obstacles to the simulated domain and limiting the area that can be controlled by the network. An example sequence using this setup is shown in Fig. 4. Here, the goal is to move the smoke from its initial position near the center into one of the three buckets, i.e. separated regions, located at the top. The control forces can only be applied in the peripheral regions, which are outside the visible smoke distribution. Only by synchronizing the 5000 continuous control parameters can a directed velocity field be constructed in the central region. We first infer trajectories using a trained CFE network and predictions that move the smoke into the desired bucket in a straight line. This baseline manages to transfer $89\% \pm 2.6\%$ of the smoke into the target bucket. Next we enable the hierarchical predictions and train the OPs. This version manages to maneuver $99.22\% \pm 0.15\%$ of the smoke into the desired buckets while requiring $19.1\% \pm 1.0\%$ less force.

5 Conclusions

We have demonstrated that deep learning models in conjunction with a differentiable physics solver can successfully predict the behavior of complex physical models and learn to control them. The introduction of a hierarchical predictor-corrector architecture allows us to learn to reconstruct long sequences by treating the physical behavior on different time scales separately. Based on these results, we believe that learning differentiable physics has significant potential to provide physical intuition for a wide range of systems that understand and interact with the real world.

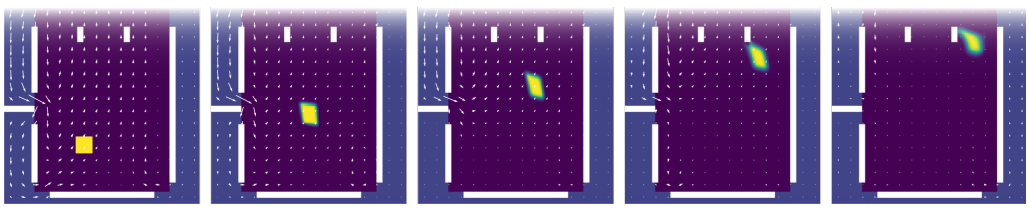

Figure 4: Example indirect control sequence. Obstacles are marked white, control regions in light blue (at left, bottom and right sides). The white arrows indicate the velocity field. The domain is enclosed in a solid box with an open top.

References

[1] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, et al. Tensorflow: A system for large-scale machine learning. In *Symposium on Operating Systems Design and Implementation*, 2016.

[2] Pulkit Agrawal, Ashvin V Nair, Pieter Abbeel, Jitendra Malik, and Sergey Levine. Learning to poke by poking: Experiential learning of intuitive physics. In *Advances in Neural Information Processing Systems*, 2016.

[3] G. K. Batchelor. *An Introduction to Fluid Dynamics*. Cambridge University Press, 1967.

[4] Tian Qi Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems*, 2018.

[5] Filipe de Avila Belbute-Peres, Kevin Smith, Kelsey Allen, Josh Tenenbaum, and J Zico Kolter. End-to-end differentiable physics for learning and control. In *Advances in Neural Information Processing Systems*, 2018.

[6] Chelsea Finn, Ian Goodfellow, and Sergey Levine. Unsupervised learning for physical interaction through video prediction. In *Advances in Neural Information Processing Systems*, 2016.

[7] Nick Haber, Damian Mrowca, Li Fei-Fei, and Daniel LK Yamins. Learning to play with intrinsically-motivated self-aware agents. *arXiv:1802.07442*, 2018.

[8] Philipp Holl, Vladlen Koltun, and Nils Thuerey. Learning to control pdes with differentiable physics. *arXiv*, 2019.

[9] Leslie Pack Kaelbling, Michael L Littman, and Andrew W Moore. Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4:237–285, 1996.

[10] Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *arXiv:1509.02971*, 2015.

[11] Volodymyr Mnih, Adria Puigdomenech Badia, Mehdi Mirza, Alex Graves, Timothy Lillicrap, Tim Harley, David Silver, and Koray Kavukcuoglu. Asynchronous methods for deep reinforcement learning. In *ICML*, 2016.

[12] Lev Semenovich Pontryagin. *Mathematical Theory of Optimal Processes*. John Wiley, 1962.

[13] Connor Schenck and Dieter Fox. SPNets: Differentiable fluid dynamics for deep neural networks. In *Conference on Robot Learning*, 2018.

[14] Gordon D Smith. *Numerical Solution of Partial Differential Equations: Finite Difference Methods*. Oxford University Press, 1985.

[15] Marc Toussaint, Kelsey Allen, Kevin Smith, and Joshua B Tenenbaum. Differentiable physics and stable modes for tool-use and manipulation planning. In *Robotics: Science and Systems*, 2018.

[16] Paul J Werbos. Backwards differentiation in AD and neural nets: Past links and new opportunities. In *Automatic Differentiation: Applications, Theory, and Implementations*, pages 15–34. Springer, 2006.