Real-time optimal control with shallow recurrent decoder networks

Matteo Tomasetto
Department of Mechanical Engineering
Politecnico di Milano, Milan, Italy
`matteo.tomasetto@polimi.it`

Andrea Manzoni
MOX - Department of Mathematics
Politecnico di Milano, Milan, Italy
`andrea1.manzoni@polimi.it`

J. Nathan Kutz
Department of Applied Mathematics and Electrical and Computer Engineering
University of Washington, Seattle, WA, USA
`kutz@uw.edu`

Abstract

Controlling dynamical systems in real-time across multiple scenarios is critical to enable adaptive control strategies, ensuring stability and efficiency. However, parametric optimal control problems require several system simulations to tailor optimal actions in response to varying scenarios, which are often computationally demanding – or even intractable – due to the high-dimensionality of spatio-temporal dynamics. In this work, we exploit SHallow REcurrent Decoder networks-based Reduced Order Modeling (SHRED-ROM) to synthesize a real-time policy for high-dimensional and parametric dynamics, relying solely on limited state sensor readings. After training the model on few optimal examples given by an expert demonstrator, as typically considered in imitation learning, SHRED-ROM mimics the expert behavior with effective distributed control actions in real-time and in new scenarios, mitigating the curse of dimensionality. The performance of the proposed policy is finally assessed on a challenging density control test case.

Introduction

Accurate simulators of dynamical systems are fundamental to predict state evolutions, guide parametric analyses, quantify model uncertainties, and design control strategies. Many physical systems are described in terms of Partial Differential Equations (PDEs) modeling the spatio-temporal evolution of the state $y : \Omega \times [0, T] \to \mathbb{R}$ in the domain Ω and in the time interval $[0, T]$. Solving PDEs requires numerical methods such as, e.g., finite element and spectral methods [7, 10]. In essence, through a discretization of the state variable $y(\mathbf{x}, t_k) \to \mathbf{y}(t_k) \in \mathbb{R}^{N_y}$, with N_y the number of spatial degrees of freedom and $t_1, ..., t_{N_t}$ a uniform grid over $[0, T]$, the PDE turns into a (possibly nonlinear) high-dimensional system of equations to be solved [10]. The high-dimensionality of the resulting system entails a demanding – or even prohibitive – computational burden. The computational bottleneck becomes even more severe when considering optimal control problems parametrized by a vector of scenario parameters $\boldsymbol{\mu} \in \mathbb{R}^p$ due to their intrinsically iterative nature, which requires multiple simulations of the system to be controlled [6, 13], and due to the necessary control adaptations in response to variations in the underlying scenario. Traditional numerical methods are therefore not suitable for solving high-dimensional and parametric optimal control problems related to safety-critical applications with strict timing requirements, such as autonomous vehicles, robotics, plasma control, and aerospace, as delays in the control computations may imply a significant loss of performance and robustness.

To speed up the resolution of PDEs, projection-based Reduced Order Models (ROMs) have been widely utilized in the literature [1, 2, 11]. Given the matrix collecting the state snapshots $\{\mathbf{y}_k^{\boldsymbol{\mu}_i} = \mathbf{y}(t_k, \boldsymbol{\mu}_i)\}_{k=1,\dots,N_t}^{i=1,\dots,N_p}$ for N_t time instants and N_p parameter values, it is possible to reduce the data dimensionality through its Singular Values Decomposition (SVD), also known as Proper Orthogonal

1st Workshop on Differentiable Systems and Scientific Machine Learning @ EurIPS 2025.

Decomposition (POD) [2, 5, 7, 11]. Specifically, the first r left singular vectors are the directions of maximum variability in the data, and represent an optimal (in a least-square or statistical sense) basis where to project the state snapshots and Galerkin-project the PDE. Doing so, one can retrieve a r-dimensional system to be solved at every time step, which is computationally tractable whether $r \ll N_y$. The same rationale can be extended to systems of optimality conditions, speeding up the resolution of distributed control problems [11]. However, projection-based ROMs require complete knowledge of the underlying physics and are limited by the linearity assumption, lacking accuracy and efficiency when dealing with nonlinear and convective phenomena. To mitigate these limitations, Deep Learning-based ROMs (DL-ROMs) have been proposed as efficient nonlinear, non-intrusive and data-driven alternatives [3, 4]. While being faster, more accurate and more flexible than projection-based ROMs, DL-ROMs still require full parametric knowledge to infer the state evolution, are uninformed to the actual system behavior, and typically require demanding hyperparameter tuning.

In this work, we employ shallow recurrent decoders [8, 12, 14] to synthesize a policy capable of predicting (possibly distributed) control actions in multiple scenarios, relying solely on sparse state sensors, while being agnostic to parameter values.

Shallow recurrent decoder networks-based reduced order modeling

SHallow REcurrent Decoder networks-based Reduced Order Modeling (SHRED-ROM) [8, 12, 14] aims at reconstructing high-dimensional and parametric spatio-temporal fields $\mathbf{y}_k^{\boldsymbol{\mu}}$ for $k = 1, ..., N_t$ in real-time and in multiple scenarios starting from limited sensor measurements, which are often available in real-world applications. Differently from state-of-the-art sensing strategies, SHRED-ROM exploits the past history of L sensor readings, i.e. $\mathbf{s}_{k-L:k}^{\boldsymbol{\mu}} = \{\mathbf{s}(t_{k-L}, \boldsymbol{\mu}), ..., \mathbf{s}(t_k, \boldsymbol{\mu})\}$, where $\mathbf{s}(t_{k-L}, \boldsymbol{\mu}) = \mathbf{0}$ if $k \leq L$. Specifically, SHRED-ROM combines a Long Short Term Memory network (LSTM) f_T and a Shallow Decoder Network (SDN) f_X to encode the temporal history of sensor data and reconstruct the spatio-temporal quantity of interest in multiple scenarios, that is

$$\mathbf{y}_k^{\boldsymbol{\mu}} \approx \tilde{\mathbf{y}}_k^{\boldsymbol{\mu}} = f_X(f_T(\mathbf{s}_{k-L:k}^{\boldsymbol{\mu}})) \quad \text{for } k = 1, ..., N_t.$$

Compressive training strategies based on, e.g., POD may be considered to compress the decoder output, enabling model training at laptop level computing with minimal hyperparameter tuning [8, 12]. Beyond reconstructing state data from its own measurements, it is possible to reconstruct one quantity from sensors monitoring a coupled field [8, 12, 14]. Taking advantage of this property, we here propose a strategy to synthesize parametric policies in the context of imitation learning, enabling real-time distributed feedback control strategies across multiple scenarios in the low-data limit.

Real-time control of distributed parametric systems with SHRED-ROM

Parametric optimal control problems can be formulated as (discrete, for the sake of simplicity) PDE-constrained optimizations in the form

$$J(\mathbf{y}, \mathbf{u}, \boldsymbol{\mu}) \to \min \text{ s.t. } \mathbf{G}(\mathbf{y}, \mathbf{u}, \boldsymbol{\mu}) = \mathbf{0}$$

where $J(\mathbf{y}, \mathbf{u}, \boldsymbol{\mu}) \in \mathbb{R}$ is the loss function to minimize by optimally designing the control $\mathbf{u}(t, \boldsymbol{\mu}) \in \mathbb{R}^{N_u}$, while $\mathbf{G}(\mathbf{y}, \mathbf{u}, \boldsymbol{\mu}) = \mathbf{0}$ stands for the discrete governing equation in implicit form. The optimal action can be computed through the Karush-Kuhn-Tucker system of optimality conditions

$$\begin{cases} \nabla_{\mathbf{y}} J(\mathbf{y}, \mathbf{u}, \boldsymbol{\mu}) + (\partial_{\mathbf{y}} \mathbf{G}(\mathbf{y}, \mathbf{u}, \boldsymbol{\mu}))^\top \boldsymbol{\lambda} = \mathbf{0} & \text{(adjoint equation)} \\ \nabla_{\mathbf{u}} J(\mathbf{y}, \mathbf{u}, \boldsymbol{\mu}) + (\partial_{\mathbf{u}} \mathbf{G}(\mathbf{y}, \mathbf{u}, \boldsymbol{\mu}))^\top \boldsymbol{\lambda} = \mathbf{0} & \text{(optimality condition)} \\ \mathbf{G}\left(\mathbf{y}, \mathbf{u}, \boldsymbol{\mu}\right) = \mathbf{0} & \text{(state equation)} \end{cases}$$

where $\boldsymbol{\lambda}(t, \boldsymbol{\mu}) \in \mathbb{R}^{N_\lambda}$ is the discrete adjoint vector. The coupling of state and control in the KKT system allows SHRED-ROM to map sparse state measurements into the corresponding distributed control actions in multiple scenarios, thus retrieving the policy

$$\mathbf{u}_k^{\boldsymbol{\mu}} = \mathbf{u}(t_k, \boldsymbol{\mu}) \approx \tilde{\mathbf{u}}_k^{\boldsymbol{\mu}} = f_X(f_T(\mathbf{s}_{k-L:k}^{\boldsymbol{\mu}})) \quad \text{for } k = 1, ..., N_t.$$

To train SHRED-ROM in a supervised manner, we consider a few optimal examples given by an expert demonstrator, as typically considered in imitation learning. Note that the need for optimal training data may represent a limiting factor, especially when dealing with synthetic data, and alternative strategies are required whenever such examples are not available. Figure 1 provides a graphical summary of the proposed feedback control strategy. Thanks to SHRED-ROM sensor efficiency, as well as the independence on parameter values, limited state sensor readings are all you need to control high-dimensional and parametric systems in real-time.

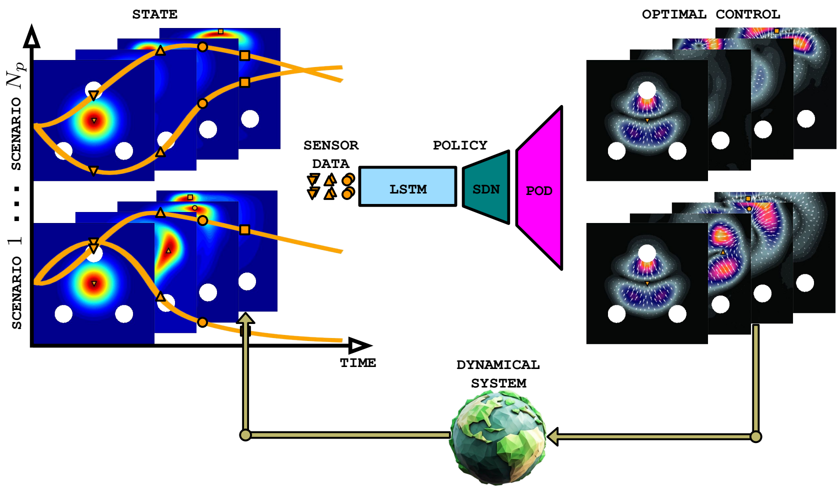

Figure 1: *Real-time optimal control with SHRED-ROM*. We exploit expert demonstrations in N_p scenarios to train the sequence (LSTM) and decoder (SDN) blocks of SHRED-ROM. Compressive training strategies based on POD are considered to compress the decoder output. After training, the synthesized policy provides distributed optimal controls in multiple scenarios unseen during training, relying solely on the history of sparse state sensor readings.

Numerical results

To assess the performance of the proposed controller, we consider a high-dimensional and parametric density control problem. Starting from a Gaussian density centered in the middle of the square $[-1,1]^2$ with variance 0.05, the density y evolves up to $T = 50$ seconds according to the advection-diffusion PDE

$$y_t + \nabla \cdot (-\eta \nabla y + \mathbf{v} y + \mathbf{u} y) = 0$$

with homogeneous Neumann boundary conditions and viscosity $\eta = 0.001$. The density is therefore transported by a parametric fluid flow with velocity $\mathbf{v}(\boldsymbol{\mu})$, whose dynamics is modeled via steady Navier-Stokes equations, with kinematic viscosity $\nu = 1.0$ and no-slip boundary conditions on external walls. The fluid velocity implicitly depends on three rotating cylinders in the domain, whose constant velocities are regarded as scenario parameters $\boldsymbol{\mu}$.

As visible in the first row of Figure 2, the uncontrolled setting entails significant density dispersion in the domain. Our goal is to design the distributed velocity $\mathbf{u}(t, \boldsymbol{\mu}) \in \mathbb{R}^{N_u}$, with $N_u = 59344$, minimizing density dispersion and boundary collisions in multiple scenarios, i.e. for different combinations of cylinder velocities in the range $[-1, 1]$. To do so, we generate 100 optimal examples through the adjoint method [9] with unitary time step, we split the trajectories into training,validation and test sets with ratio $80{:}10{:}10$, and we train SHRED-ROM to predict the optimal velocity field starting from the state measurements and the coordinates of 1 mobile sensor, placed in the center of the domain at $t = 0$ and passively steered by the underlying transport effect. Alternatively, one can also consider sparse fixed sensors monitoring the state evolution. To speed up training, we reduce the control dimensionality through POD with compression ratio equal to 99.5% ($r = 300 \ll N_u$), yielding a mean relative reconstruction error on test data equal to 3.38%. After training SHRED-ROM with lag $L = 10$, it is possible to deploy the obtained policy in test scenarios unseen during training. As visible in Figure 2, SHRED-ROM can effectively minimize the density dispersion and boundary collisions over time, with performance similar to the target optimal solutions in the test set.

Conclusions

In this work, we propose a parametric feedback control strategy in the low-data limit, agnostic to parameter values. Specifically, in the context of imitation learning, we exploit SHRED-ROM to mimic expert demonstrations and predict distributed control actions in multiple scenarios, relying solely on limited state sensor readings. After training, it is possible to control in closed-loop high-dimensional and parametric systems, designing effective optimal control strategies in real-time.

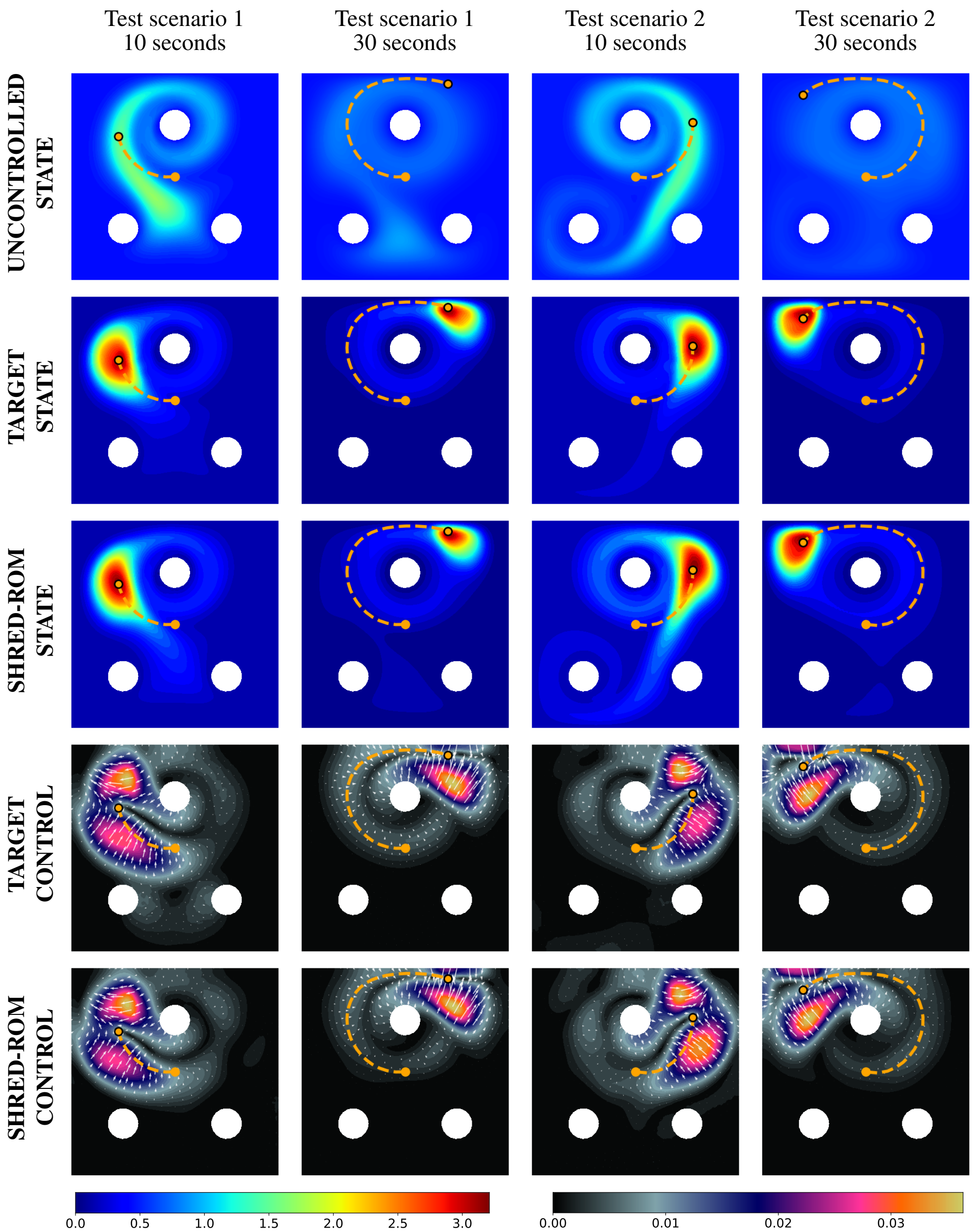

Figure 2: *Density control*. Uncontrolled states (first row), optimal target states (second row), controlled states with SHRED-ROM policy (third row), target optimal controls (fourth row), and SHRED-ROM control predictions (fifth row) in two different test scenarios $\boldsymbol{\mu} = [-0.06, 0.06, -0.70]^\top$ and $\boldsymbol{\mu} = [-0.74, 0.01, 0.83]^\top$ at $t = 10$ and $t = 30$ seconds. The mobile sensor trajectories are depicted in orange. The control velocity fields are represented through vector fields, with the underlying colours corresponding to their magnitude.

Acknowledgements

The work of JNK was supported in part by the US National Science Foundation (NSF) AI Institute for Dynamical Systems (dynamicsai.org), grant 2112085. AM acknowledges the Project Reduced Order Modeling and Deep Learning for the real-time approximation of PDEs (DREAM) (Starting Grant No. FIS00003154), funded by the Italian Science Fund (FIS) - Ministero dell'Università e della Ricerca, the project FAIR (Future Artificial Intelligence Research), funded by the NextGenerationEU program within the PNRR-PE-AI scheme (M4C2, Investment 1.3, Line on Artificial Intelligence) and the project "Dipartimento di Eccellenza" 2023–2027, funded by MUR. MT and AM are members of the Gruppo Nazionale Calcolo Scientifico-Istituto Nazionale di Alta Matematica (GNCS-INdAM).

References

[1] A. C. Antoulas. *Approximation of large-scale dynamical systems*. SIAM, 2005.

[2] P. Benner, S. Gugercin, and K. Willcox. A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM review*, 57(4):483–531, 2015.

[3] 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.

[4] J. Hesthaven and S. Ubbiali. Non-intrusive reduced order modeling of nonlinear problems using neural networks. *Journal of Computational Physics*, 363:55–78, 2018.

[5] P. Holmes, J. L. Lumley, G. Berkooz, and C. W. Rowley. *Turbulence, coherent structures, dynamical systems and symmetry*. Cambridge university press, 2012.

[6] D. E. Kirk. *Optimal control theory: an introduction*. Courier Corporation, 2004.

[7] J. N. Kutz. *Data-driven modeling & scientific computation: methods for complex systems & big data*. Oxford University Press, 2013.

[8] J. N. Kutz, M. Reza, F. Faraji, and A. Knoll. Shallow recurrent decoder for reduced order modeling of plasma dynamics, 2024.

[9] S. K. Mitusch, S. W. Funke, and J. S. Dokken. dolfin-adjoint 2018.1: automated adjoints for fenics and firedrake. *Journal of Open Source Software*, 4(38):1292, 2019.

[10] A. Quarteroni. *Numerical Models for Differential Problems*. Springer Cham, 2017.

[11] A. Quarteroni, A. Manzoni, and F. Negri. *Reduced basis methods for partial differential equations: an introduction*, volume 92. Springer, 2015.

[12] M. Tomasetto, J. P. Williams, F. Braghin, A. Manzoni, and J. N. Kutz. Reduced order modeling with shallow recurrent decoder networks, 2025.

[13] F. Tröltzsch. *Optimal control of partial differential equations: theory, methods, and applications*, volume 112. American Mathematical Soc., 2010.

[14] J. P. Williams, O. Zahn, and J. N. Kutz. Sensing with shallow recurrent decoder networks. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 480(2298):20240054, 2024.