Workshop on Mathematical Machine Learning and Application, 14-16 December 2020, CCMA

Nonlinear Reduced Order Modelling of Parametrized PDEs using Deep Neural Networks

Theoretical Analysis and Numerical Results

Franco N.R. [a], Manzoni A. [a], Zunino P. [a]

[a] MOX – Modeling and Scientific Computing – Department of Mathematics – Politecnico di Milano (Italy),

POLITECNICO MILANO 1863

MODELLISTICA E CALCOLO SCIENTIFICO MOX MODELING AND SCIENTIFIC COMPUTING

Abstract. We consider the problem of approximating the parameter-to-state map of a parameter dependent PDE by means of Deep Neural Networks. The latter is a currently active area of research, e.g. [1, 2, 3, 4, 5, 6], whose development is mainly motivated by the limitations of the classical approaches such as the Reduced Basis method [7, 8]. In particular, these are known to encounter substantial difficulties in transport-dominated problems and under the presence of highly localized nonlinear terms. Here, we tackle this kind of problems by exploiting the intrinsic nonlinearity of neural networks and propose a Deep Learning based Reduced Order Model [9]. Our construction finds its theoretical fundations in a generalized version of the Kolmogorov *n*-width [10, 11].

Background and theoretical fundations

General setting

We are given a parameter space Θ, a Hilbert state space H and a parametrized PDE, e.g.

$$\begin{cases} -\text{div}(\sigma(\mu)\nabla u(\mu)) + \mathbf{b}(\mu)\cdot\nabla u(\mu) = f(\mu) & \text{in } \Omega \\ u(\mu) = 0 & \text{in } \partial\Omega \end{cases}$$

and we are interested in approximating the parameter-to-state map, $\mu \to u(\mu)$.

Each evaluation of u can be extremely expensive: Full Order Models (FOM) are not suitable for the purpose.

Reduced Order Modelling

In particular, in our case we investigate the use of neural networks, and propose a Deep Learning based Reduced Order Model (**DL-ROM**).

Dimensionality reduction

N_h: FOM dimension, n: ROM dimension

By their very definition, Deep Neural Networks (**DNNs**) introduce nonlinearities in the model. To make the most out of it, we base our construction on generalizations of the Kolmogorov n-width.

Linear projection (e.g. POD)

- At the base of several ROMs
- Founded on the notion of Kolmogorov n-width

$$d_n(\mathcal{S}) := \inf_{\mathbf{V}\in\mathbb{R}^{N_h\times n}} \sup_{w\in\mathcal{S}} ||w - \mathbf{V}\mathbf{V}^T w||$$

Nonlinear reduction

- Promising alternative
- Inspired to nonlinear generalizations of the Kolmogorov n-width

$$\delta_n(\mathcal{S}) := \inf_{\substack{\Psi'\in\mathcal{C}(\mathcal{S},\mathbb{R}^n) \\ \Psi\in\mathcal{C}(\mathbb{R}^n,\mathbb{R}^{N_h})}} \sup_{w\in\mathcal{S}} ||w - \Psi(\Psi'(w))||$$

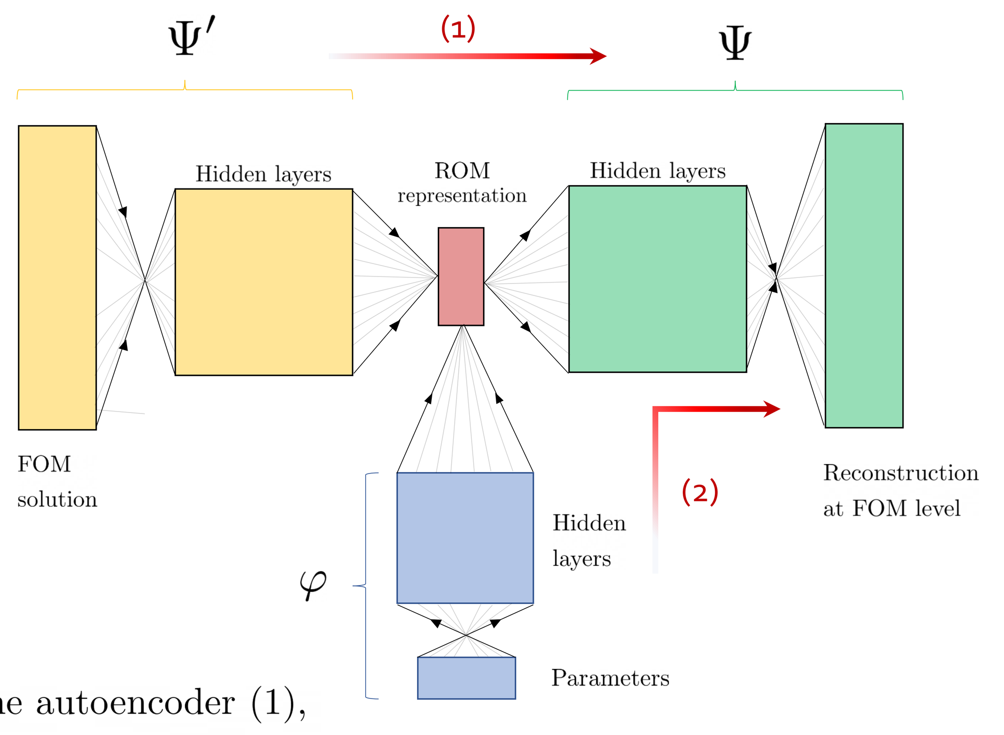
DL-ROM Design

Network architecture

- Autoencoder $\Psi\circ\Psi'$ with input-output dimension N_h and latent dimension n
- Deep Feed Forward NN φ with input dimension p and output dimension n

Activation type: leaky-ReLU

The parameter-to-state map u is then approximated by $\Phi := \Psi\circ\varphi$.



The training is done in two stages: first the autoencoder (1), then we train φ (2). The second phase may involve Ψ as well.

On the latent dimension

The autoencoder yields a low-dimensional representation of the *solution manifold* $\mathcal{S} := \{u(\mu)\}_{\mu\in\Theta} \subset H$. To obtain fast and light ROMs we set $n = n_{\min}(\mathcal{S})$, where:

$$n_{\min}(\mathcal{S}) := \min\ \{m \mid \delta_m(\mathcal{S}) = 0\}.$$

In general, even if the PDE depends on p parameters, one may have $n_{\min}(\mathcal{S}) \neq p$. However,

Theorem 1. Θ compact, u Lipschitz in $\mu \implies n_{\min}(\mathcal{S}) \le 2p+1$.

Theorem 2. Θ compact set with nonempty interior, u continuous in $\mu \implies n_{\min}(\{(\mu, u(\mu))\ |\mu\in\Theta\ \}) = p$.

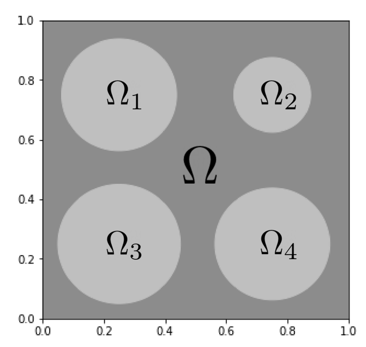
Numerical Experiments

Description

Parametrized 2D Steady Advection-Diffusion

We consider a steady advection-diffusion equation on the unit square. The equation depends on $p = 7$ scalar parameters, $\mu \in \Theta \subset \mathbb{R}^7$.

$$\begin{cases} -\text{div}(\sigma(\mu)\nabla u(\mu)) + \mathbf{b}(\mu)\cdot\nabla u(\mu) = f(\mu) & \text{in } \Omega \\ u(\mu) = 0 & \text{in } \partial\Omega \end{cases}$$



Spatial domain Ω of the differential problem. The conductivity field σ is piecewise constant and can change parametrically in he four highlighted circular subdomains, $\{\Omega_i\}_{i=1}^4$.

Where,

- $\sigma(\mu) := \frac{1}{10} + \sum_{i=1}^4 \mu_i \mathbf{1}_{\Omega_i}$ ⟵ piecewise constant conductivity field
- $f(\mu)$ Dirac delta centered at (μ_6, μ_7) ⟵ strongly localized force
- $\mathbf{b}(\mu) = C\cdot(\cos\mu_5,\ \sin\mu_5)^T$ ⟵ convective field with parametrized direction and fixed intensity

We consider two separate cases: $C = 0.5$ and $C = 40$.

Data and ROM design

Training data: 9000 high-fidelity snapshots obtained using P1-FEM on a 211×211 mesh. → $N_h = 44521$

Test data: 1000 snapshots (as above).

Autoencoder architecture: latent dimension $n = 7$.

Two alternatives are explored: **AE1** reconstructs $u(\mu)$ while **AE2** works with $(\mu, u(\mu))$.

Implementation: Python 3, esp. FEniCS and Pytorch.

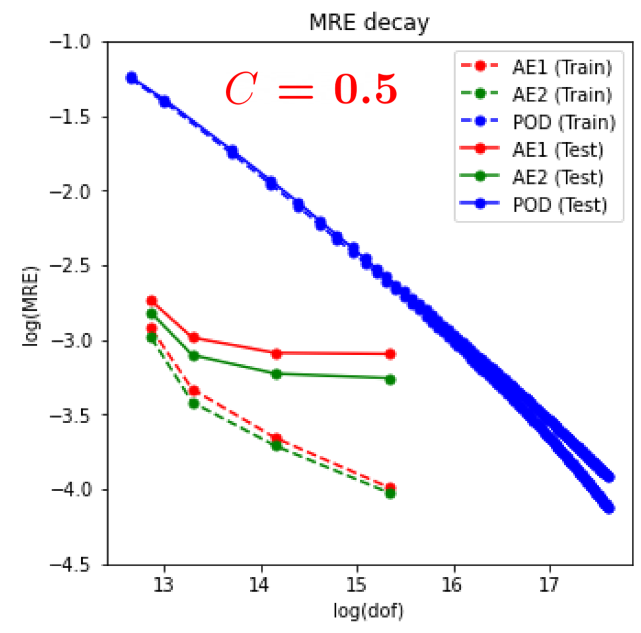
Results

Dimensionality reduction

Below, we compare the autoencoders performance with the POD in terms on **MRE** (Mean Relative Error). To make a meaningfull analysis, we compare models by taking into account their complexities. We measure the latter in terms of **dof** (degrees of freedom): for the autoencoders, it is given by the number of weights and biases; for the POD, it corresponds to the size of $\mathbf{V}$, i.e. the number of basis times N_h. The pictures show the MRE decay in terms of the complexity (loglog scale).

In the tables we focus on the best performing autoencoders. The latent dimension is fixed to 7. Conversely, the reduced dimension (number of basis) needed for the POD to match/surpass the autoencoders is n_{POD}.

In this sense, if dof is the complexity of the network, the autoencoders yield a memory gain of $1 - \frac{\text{dof}}{N_h n_{\text{POD}}}$.



Training set

	n	dof	MRE	n_{POD}	Memory gain
AE1	7	4.62e6	1.85%	840	88%
AE2	7	4.62e6	1.79%	880	90%

Test set

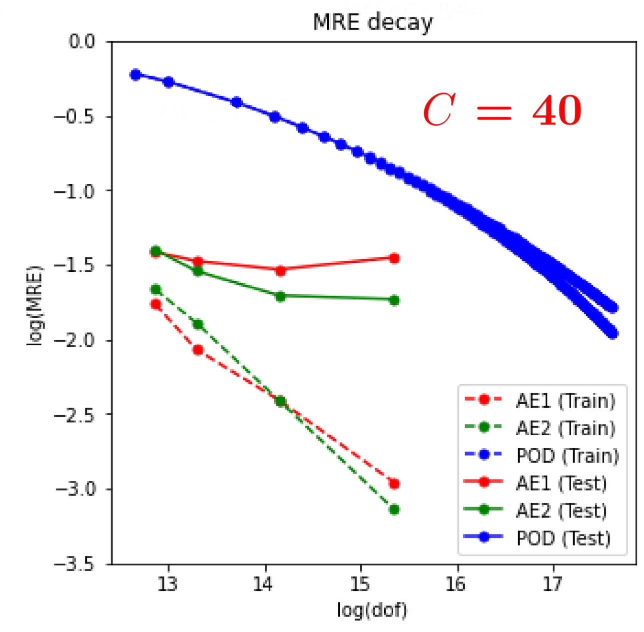
	n	dof	MRE	n_{POD}	Memory gain
AE1	7	4.62e6	4.53%	230	55%
AE2	7	4.62e6	3.85%	290	65%

Training set

	n	dof	MRE	n_{POD}	Memory gain
AE1	7	4.62e6	5.20%	≥ 1000	≥ 90%
AE2	7	4.62e6	4.35%	≥ 1000	≥ 90%

Test set

	n	dof	MRE	n_{POD}	Memory gain
AE1	7	4.62e6	23.40%	420	76%
AE2	7	4.62e6	17.70%	700	86%



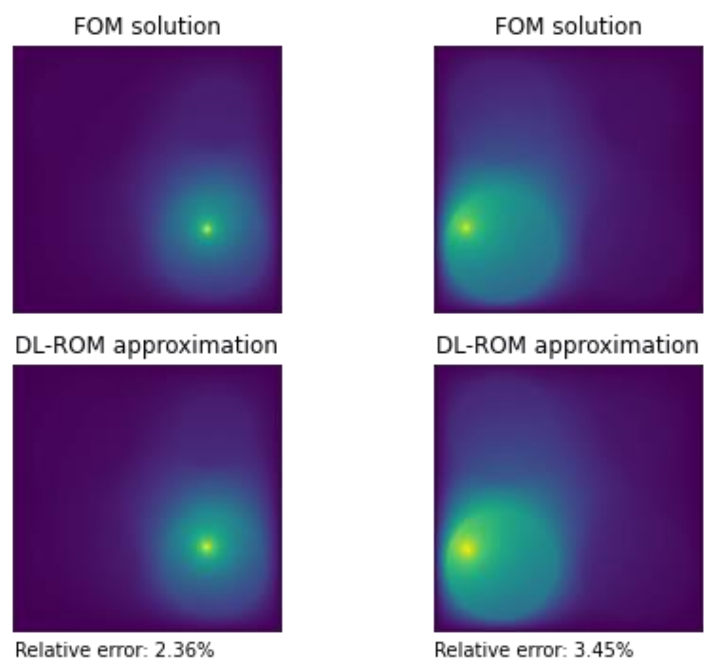
Approximation of the Parametric Map

Separately for each case, $C = 0.5$ and $C = 40$, we pick the best autoencoder in terms of Test MRE. We then train φ according to our model. Numerical results and figures below.

$C = 0.5$

Train MRE = 1.55%, Test MRE = 4.99%

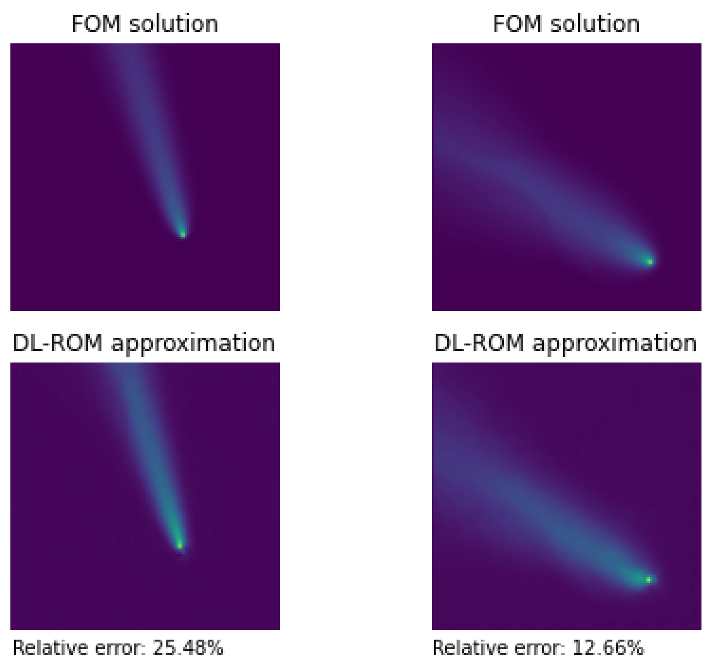
Examples from the **Test** set:



$C = 40$

Train MRE = 5.22%, Test MRE = 20.38%

Examples from the **Test** set:



References

[1] Kutyniok, G., Petersen, P., Raslan, M. and Schneider, R., A Theoretical Analysis of Deep Neural Networks and Parametric PDEs, *arXiv preprint, arXiv:1904.00377* (2019).

[2] Laakmann, F. and Petersen, P., Efficient Approximation of Solutions of Parametric Linear Transport Equations by ReLU DNNs, *arXiv preprint, arXiv:2001.11441* (2020).

[3] Fresca, S., Manzoni, A. and Dede, L., A comprehensive deep learning-based approach to reduced order modeling of nonlinear time-dependent parametrized PDEs, *arXiv preprint, arXiv:2001.04001* (2020).

[4] Schwab, C. and Zech, J., Deep learning in high dimension: Neural network expression rates for generalized polynomial chaos expansions in UQ, *Analysis and Applications*, vol. 17, 1, pp. 19-55 (2019).

[5] Bhattacharya, K., Hosseini, B., Kovachi, N.B., and Stuart, A.M., Model Reduction and Neural Networks for Parametric PDEs, *arXiv preprint, arXiv:2005.03180v1* (2020).

[6] Chen, W., Wang, Q., Hesthaven, J.S. and Zhang, C., Physics-informed machine learning for reduced-order modeling of nonlinear problems, *Journal of Computational Physics preprint* (2020).

[7] Quarteroni, A., Manzoni, A. and Negri F., *Reduced basis methods for partial differential equations - An introduction*, vol. 92 of Unitext. Springer, Cham (2016).

[8] Cohen, A., and DeVore, R., Approximation of high-dimensional parametric PDEs, *Acta Numerica*, pp. 1–159 (2015).

[9] Franco, N.R., Manzoni, A. and Zunino, P., Deep-Learning based Reduced Order Modelling for Parametrized Partial Differential Equations, *In preparation*.

[10] DeVore, R., Howard, R., and Micchelli, C., Optimal Nonlinear Approximation, *Manuscripta Mathematica*, 63, 469 - 478 (1989).

[11] DeVore, R., Nonlinear approximation, *Acta Numerica*, pp. 51–150 (1998).