Nonlinear Manifold Least-Squares Petrov-Galerkin with Reduced Over Collocation and Teacher-Student training

Francesco Romor
Advisor: Prof. Gianuigi Rozza
Co-advisor: Dr. Giovanni Stabile

RAMSES, SISSA

17 December 2021

Nonlinear ROMs: Kolmogorov n-width

PDE	FOM	ROM
$\forall \boldsymbol{\mu} \in \mathcal{P},\ U \in \mathcal{U}$	$\mathcal{P} \subset \mathbb{R}^{n_{\text{train}}}, \tilde{U} \in X_h$	$\mathcal{P} \subset \mathbb{R}^{n_{\text{test}}}, \bar{U} \in X_{ROM}$
$\begin{cases} \mathcal{G}_{\boldsymbol{\mu}}(U(\boldsymbol{\mu}, t, \mathrm{x})) = 0 \\ \mathcal{B}_{\boldsymbol{\mu}}(U(\boldsymbol{\mu}, t, \mathrm{y})) = 0 \end{cases}$	$\begin{cases} \tilde{G}_{\boldsymbol{\mu},h,\Delta t}(\tilde{U}_{\boldsymbol{\mu}}) = 0 \\ \tilde{B}_{\boldsymbol{\mu},h,\Delta t}(\tilde{U}_{\boldsymbol{\mu}}) = 0 \end{cases}$	$\begin{cases} \bar{G}_{\boldsymbol{\mu},h,\Delta t}(\bar{U}_{\boldsymbol{\mu}}) = 0 \\ \bar{B}_{\boldsymbol{\mu},h,\Delta t}(\bar{U}_{\boldsymbol{\mu}}) = 0 \end{cases}$

Kolmogorov n-witdh

Let $(\mathcal{P}, \|\cdot\|)_{\mathcal{P}}$ and $(\mathcal{U}, \|\cdot\|_{\mathcal{U}})$ be complex Banach spaces, and $K \subset\subset \mathcal{P}$ compact. If $L : K \subset\subset \mathcal{P} \to \mathcal{U}$ is the solution map, we define the Kolmogorov n-width of $L(K) \subset \mathcal{U}$,

$$d_n(L(K))_{\mathcal{U}} = \inf_{\substack{W \subset\subset \mathcal{U} \\ \dim(W)=n}} \max_{v \in K} \min_{w \in W} \|L(v) - w\|_{\mathcal{U}}. \tag{1}$$

Nonlinear ROMs: Kolmogorov n-width decay

Theorem 1, [Cohen and DeVore, 2016]

Suppose $L : O \subset \mathcal{P} \to \mathcal{U}$ is a holomorphic mapping from an open set into $\mathcal{U}$ and L is uniformly bounded on O. If $K \subset O$ is a compact subset of $\mathcal{P}$, then for any $s > 1$ and $t < s - 1$,

$$\sup_{n\geq 1} n^s d_n(K)_{\mathcal{P}} < \infty \Rightarrow \sup_{n\geq 1} n^t d_n(L(K))_{\mathcal{U}} < \infty \tag{2}$$

$$-\mathrm{div}(\boldsymbol{\mu}\nabla u) = f \quad d_n(\mathcal{M}) \lesssim \exp^{-n}$$ [Babuška et al., 2007],

$$u^3 - \mathrm{div}(\exp(\boldsymbol{\mu})\nabla u) = f \quad d_n(\mathcal{M}) \lesssim n^{-t}$$ [Cohen and DeVore, 2016],

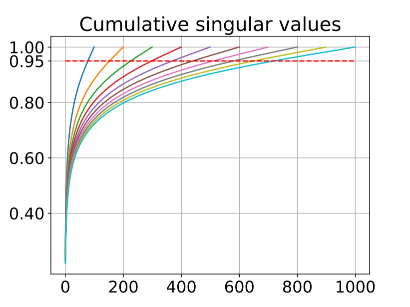
$$\partial_t u - \mu\partial_x u = 0 \quad d_n(\mathcal{M}) \gtrsim n^{-\frac{1}{2}}$$ [Ohlberger and Rave, 2015],

$$\partial^2_{tt} u - \mu\partial^2_{xx} u = 0 \quad d_n(\mathcal{M}) \gtrsim n^{-\frac{1}{2}}$$ [Greif and Urban, 2019].

Nonlinear ROMs: discrete point of view SVD

$$\partial_t u - \partial_x u = 0 \;\rightarrow\; \mathrm{SVD}[\mathcal{U}_{train}] = V\Sigma U \;\rightarrow\; \dot{\alpha}_j = \mathrm{v}_j \cdot \mathrm{F}(\textstyle\sum_{i=1}^{r} \alpha_i \mathrm{v}_i)$$

$$\mathcal{U}_{\text{train}} = \begin{bmatrix} 1 & 1 & 1 & 1 & \dots \\ 0 & 1 & 1 & 1 & \dots \\ 0 & 0 & 1 & 1 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}_{n_x \times N_t}$$



The singular values decay depends on

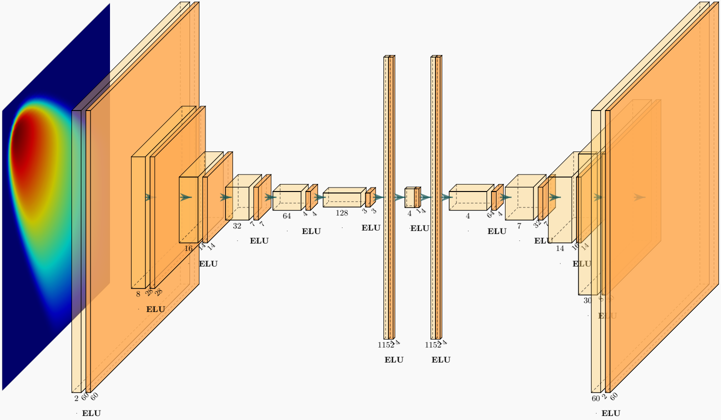
1. mesh refinements, temporal discretization step
2. choice of linear subspaces
3. full-order model Kolmogorov n-witdh decay

Survey of nonlinear ROMs

1. Purely data-driven:
   - ConvAE-LSTM [Mücke et al., 2021],
   - POD-DL-ROM [Fresca and Manzoni, 2021],
   - Multi-level ConvAe [Xu and Duraisamy, 2020].
2. Local and dictionary based ROMs:
   - Model reduction by domain decomposition [Buffoni et al., 2009],
   - ROM-net [Daniel et al., 2020].
3. Non-linear pull-back in reference system:
   - Registration-based methods [Taddei, 2020],
   - Advection-informed methods [Iollo and Lombardi, 2014],
   - ALE for convection-dominated flows [Mojgani and Balajewicz, 2017].
4. Filtering, transforms ...
   - Laplace transform and Contour Integral Methods [Guglielmi et al., 2020]
5. Nonlinear-manifold with shallow autoencoders:
   - NM-LSPG-GNAT with shallow autoencoders [Kim et al., 2020].

Manifold learning with autoencoders, cite geometric ML

The usual POD algorithm can be generalized with an autoencoder. The embedding maps $\psi : \mathbb{R}^d \to \mathbb{R}^r$ and $\phi : \mathbb{R}^r \to \mathbb{R}^d$ are trained with a mean square loss and the Adam algotihm in PyTorch [Paszke et al., 2019].



NM Least-squares Petrov-Galerkin [Lee and Carlberg, 2020]

Let us suppose that $G_{h,\delta t} : \mathcal{P} \times X_h \times X_h^{|I_t|} \to X_h$ is the discrete FOM,

$$G_{h,\delta t}(\boldsymbol{\mu}, \mathrm{U}_h^t, \{\mathrm{U}_h^s\}_{s\in I_t}) = 0. \tag{3}$$

For each discrete time instant t the following nonlinear least-squares problem is solved for the latent state $\mathrm{z}^t \in Z$, with the Levenberg-Marquardt algorithm

$$\mathrm{z}^t = \underset{\mathrm{z}\in\mathbb{R}^r}{\mathrm{argmin}} \; \| G_{h,\delta t}(\boldsymbol{\mu}, \phi(\mathrm{z}), \{\phi(\mathrm{z}^s)\}_{s\in I_t}) \|^2_{X_h}. \tag{4}$$

That is for each time instant the following intermediate solutions $\{\mathrm{z}^{t,k}\}_{k\in\{0,\dots,N(t)\}}$, $\mathrm{z}^{t,0} = \mathrm{z}^{t-1,N(t-1)}$ of the linear system in $\mathbb{R}^r$ are considered,

$$\Big((dG^{t,k-1}d\phi^{t,k-1})^T dG^{t,k-1}d\phi^{t,k-1} + \lambda\, I_d\Big)\delta\mathrm{z}^{t,k} = \tag{5}$$

$$= -(dG^{t,k-1}d\phi^{t,k-1})^T G^{t,k}$$

$$\mathrm{z}^{t,k} = \mathrm{z}^{t,k-1} + \alpha^k\, \delta\mathrm{z}^{t,k} \tag{6}$$

Reduced over-collocation method [Chen et al., 2021]

At each optimization step the reduced variable $z \in \mathbb{R}^r$ is forwarded to $U_h = \phi(z) \in \mathbb{R}^d$. Let $P_{r_h} \in \mathcal{S} = \{P \in \mathbb{R}^{r_h \times d} | P = (e_{i_1}|\dots|e_{i_{r_h}})^T\}$,

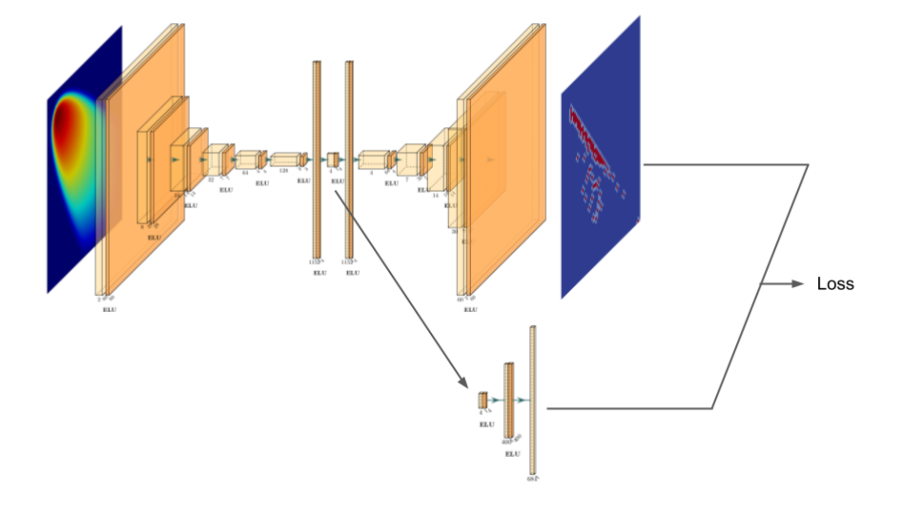
$$\begin{aligned} z^t &= \underset{z\in\mathbb{R}^r}{\operatorname{argmin}} \|P_{r_h} G_{h,\delta t}(\boldsymbol{\mu}, \phi(z), \{\phi(z^s)\}_{s\in I_t})\|^2_{\mathbb{R}^{r_h}} && (7)\\ &= \underset{z\in\mathbb{R}^r}{\operatorname{argmin}} \|P_{r_h} G_{h,\delta t}(\boldsymbol{\mu}, P_{r_h}(\phi(z)), \{P_{r_h}(\phi(z^s))\}_{s\in I_t})\|^2_{\mathbb{R}^{r_h}} && (8)\\ &= \underset{z\in\mathbb{R}^r}{\operatorname{argmin}} \|\tilde{G}_{h,\delta t}(\boldsymbol{\mu}, \tilde{\phi}(z), \{\tilde{\phi}(z^s)\}_{s\in I_t})\|^2_{\mathbb{R}^{r_h}}. && (9) \end{aligned}$$

The nodes/magic points of the over-collocation hyper-reduction method should be defined such that,

$$P_{r_h} = \underset{P\in\mathcal{S}}{\operatorname{argmin}} \max_{(z^t, \{z^s\}_{s\in I_t})\in\mathcal{T}} \|P G_{h,\delta t}(\boldsymbol{\mu}, P(\phi(z^t)), \{P(\phi(z^s))\}_{s\in I_t})\|^2_{\mathbb{R}^{r_h}}, \quad (10)$$

where $\mathcal{T}$ is the space of discrete solution trajectories.

Knowledge distillation: teacher-student training

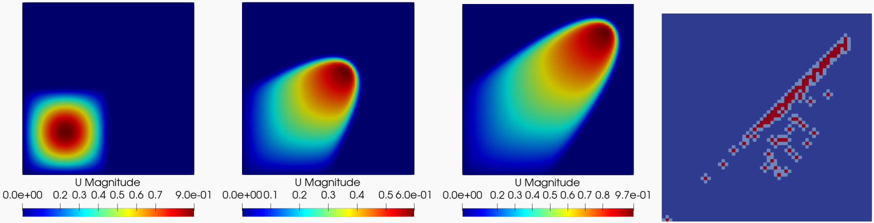


Test case: nonlinear conservation law

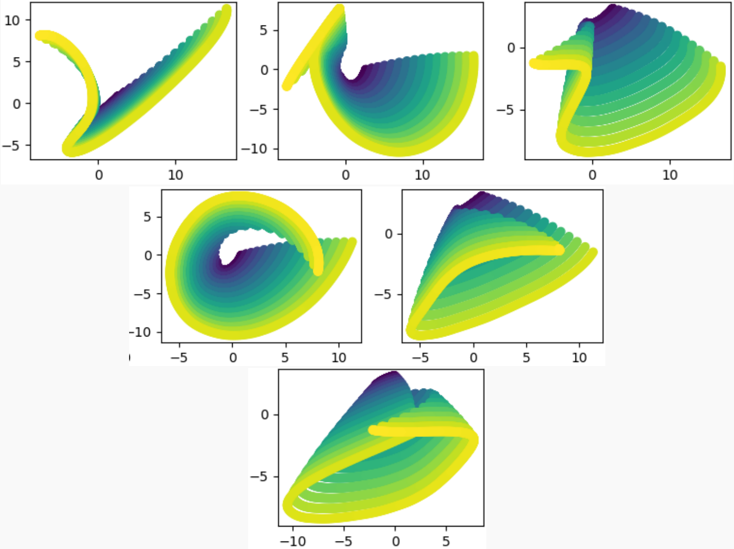
We consider the following parametric nonlinear conservation law model $\Omega = [0, 1]^2$,

$$\begin{cases} \partial_t \mathsf{u} + \frac{1}{2}\nabla \cdot (\mathsf{u} \otimes \mathsf{u}) = \nu \Delta \mathsf{u} \\ \mathsf{u}(0, \mathsf{x}) = \mu * 0.3 * \sin(2\pi x)\sin(2\pi y) * \chi_{[0,0.5]^2} \end{cases}$$

where $\mu \in [0.8, 2]$.The FOM is solved for with the FVM in OpenFoam. The NM-LSPG-ROC method is implemented in ITHACA-FV.



Latent 4-dimensional dynamics



Results

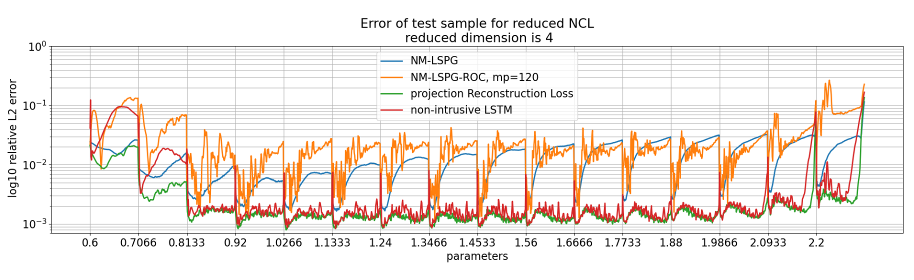


Figure: ConvAE (Convolutional Autoencoder), NM-LSPG-LM (Nonlinear Manifold Least-Squares Petrov-Galerkin Levenberg-Marquardt), NM-LSPG-ROC (Nonlinear manifold Least-Squares Petrov Galerkin Reduced Over-Collocation), LSTM (Long Short-term Memory NN)

References

- Lee, Kookjin, and Kevin T. Carlberg. "Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders." Journal of Computational Physics 404 (2020): 108973.
- Kim, Y., Choi, Y., Widemann, D., Zohdi, T. (2021). A fast and accurate physics-informed neural network reduced order model with shallow masked autoencoder. Journal of Computational Physics, 110841

References I

Babuška, I., Nobile, F., and Tempone, R. (2007).
A stochastic collocation method for elliptic partial differential equations with random input data.
SIAM Journal on Numerical Analysis, 45(3):1005–1034.

Buffoni, M., Telib, H., and Iollo, A. (2009).
Iterative methods for model reduction by domain decomposition.
Computers & Fluids, 38(6):1160–1167.

Chen, Y., Gottlieb, S., Ji, L., and Maday, Y. (2021).
An eim-degradation free reduced basis method via over collocation and residual hyper reduction-based error estimation.
arXiv preprint arXiv:2101.05902.

References II

Cohen, A. and DeVore, R. (2016).
Kolmogorov widths under holomorphic mappings.
IMA Journal of Numerical Analysis, 36(1):1–12.

Daniel, T., Casenave, F., Akkari, N., and Ryckelynck, D. (2020).
Model order reduction assisted by deep neural networks (rom-net).
Advanced Modeling and Simulation in Engineering Sciences, 7(1):1–27.

Fresca, S. and Manzoni, A. (2021).
Pod-dl-rom: enhancing deep learning-based reduced order models for nonlinear parametrized pdes by proper orthogonal decomposition.
arXiv preprint arXiv:2101.11845.

Greif, C. and Urban, K. (2019).
Decay of the kolmogorov n-width for wave problems.
Applied Mathematics Letters, 96:216–222.

References III

Guglielmi, N., López-Fernández, M., and Manucci, M. (2020).
Pseudospectral roaming contour integral methods for convection-diffusion equations.
arXiv preprint arXiv:2012.07085.

Iollo, A. and Lombardi, D. (2014).
Advection modes by optimal mass transfer.
Physical Review E, 89(2):022923.

Kim, Y., Choi, Y., Widemann, D., and Zohdi, T. (2020).
A fast and accurate physics-informed neural network reduced order model with shallow masked autoencoder.
arXiv preprint arXiv:2009.11990.

References IV

Lee, K. and Carlberg, K. T. (2020).
Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders.
Journal of Computational Physics, 404:108973.

Maday, Y. and Rønquist, E. M. (2002).
A reduced-basis element method.
Journal of scientific computing, 17(1):447–459.

Mojgani, R. and Balajewicz, M. (2017).
Arbitrary lagrangian eulerian framework for efficient projection-based reduction of convection dominated nonlinear flows.
In *APS Division of Fluid Dynamics Meeting Abstracts*, pages M1–008.

References V

Mücke, N. T., Bohté, S. M., and Oosterlee, C. W. (2021).
Reduced order modeling for parameterized time-dependent pdes using spatially and memory aware deep learning.
Journal of Computational Science, page 101408.

Ohlberger, M. and Rave, S. (2015).
Reduced basis methods: Success, limitations and future challenges.
arXiv preprint arXiv:1511.02021.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S. (2019).
Pytorch: An imperative style, high-performance deep learning library.

References VI

In Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., and Garnett, R., editors, *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc.

Taddei, T. (2020).
A registration method for model order reduction: data compression and geometry reduction.
SIAM Journal on Scientific Computing, 42(2):A997–A1027.

Xu, J. and Duraisamy, K. (2020).
Multi-level convolutional autoencoder networks for parametric prediction of spatio-temporal dynamics.
Computer Methods in Applied Mechanics and Engineering, 372:113379.