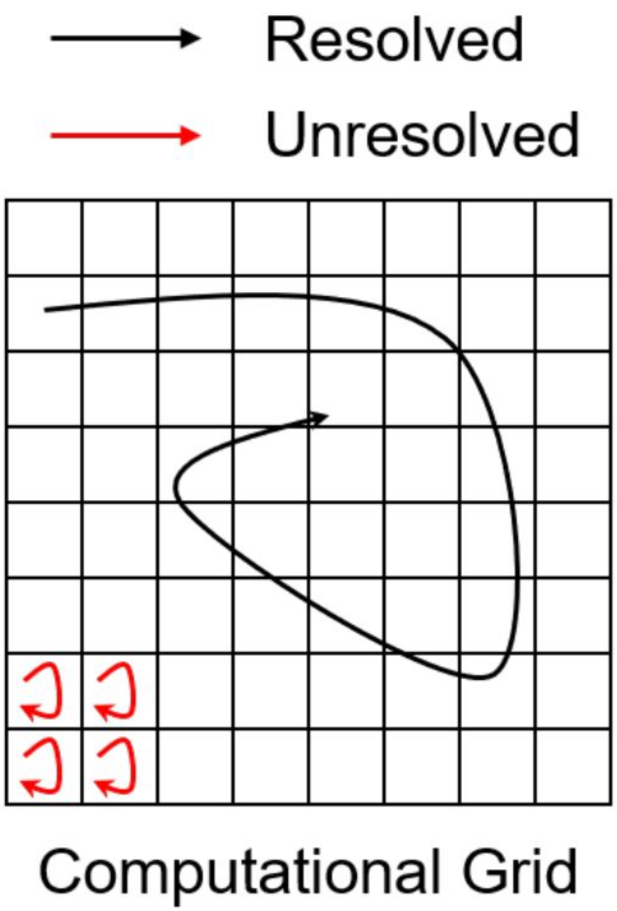
PDE-Constrained Learning of Data-Driven Turbulence Models

Romit Maulik,
Assistant Professor, The Pennsylvania State University,
Faculty Affiliate, Argonne National Laboratory.
Email: rmaulik@psu.edu

PennState

Large eddy simulations (LES)

LES attempts to reduce resolution requirement by modeling unresolved eddies.



- Resolved scales from 'grid-filtered' Navier-Stokes.
- Effect of unresolved scales from turbulence model.

Resolved and unresolved scales interact nonlinearly

$$\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j} = -\frac{\partial P}{\partial x_i} + \frac{1}{Re}\frac{\partial^2 u_i}{\partial x_j \partial x_j}$$

Model choice affects statistics significantly

Model assessment?

- Compare LES statistics to fully resolved Navier-Stokes (DNS).
- Compare LES statistics to experimental observations.

Closure modeling for LES

LES models are applied as a source term to the 'grid-filtered' Navier-Stokes equations.

$$\frac{\partial \bar{u}_i}{\partial t} + \frac{\partial \bar{u}_i \bar{u}_j}{\partial x_j} = -\frac{\partial \bar{P}}{\partial x_i} + \frac{1}{Re}\frac{\partial^2 \bar{u}_i}{\partial x_j \partial x_j} + \frac{\partial \tau_{ij}^R}{\partial x_j}$$

Structural closures

Source term approximated as:

$$\tau_{ij}^* = \bar{u}_i \bar{u}_j - \widetilde{u_i^* u_j^*},$$

$$\tilde{u}_i = G * u_i.$$

Layton scale-similarity model:

$$\tau_{ij}^* = \bar{u}_i \bar{u}_j - \widetilde{\bar{u}_i \bar{u}_j}.$$

Approximate deconvolution (AD)[2]:

$$u_i^{*0} = \bar{u}_i,$$

$$u_i^{*i} = u_i^{*i-1} + \left(\bar{u}_i - G * u_i^{*i-1}\right), \quad i = 1, 2, 3, \ldots, Q.$$

Key limitation: Definition of G.

Functional closures

Source term approximated as:

$$\tau_{ij}^R = 2\nu_e \bar{S}_{ij},$$

Standard Smagorinsky model:

$$\nu_e = (C_s \bar{\Delta})^2 |\bar{S}|,$$

Dynamic Smagorinsky model[3]:

$$C_s^2 = \frac{1}{2}\frac{\mathbb{L}_{kl}^R \bar{S}_{kl}}{\mathbb{M}_{mn} \bar{S}_{mn}}$$

$$\mathbb{L}_{ij}^R = 2C_s^2 \left(\tilde{\bar{\Delta}}^2 |\tilde{\bar{S}}| \tilde{\bar{S}}_{ij} - \bar{\Delta}^2 \widetilde{|\bar{S}_{ij}| \bar{S}_{ij}}\right) = 2C_s^2 \mathbb{M}_{ij}.$$

Key limitation: Definition of $\tilde{\bar{\Delta}}$ and test-filter.

[2] S. Stolz and N. A. Adams. In: Phys. Fluids 11 (1999), pp. 1699-1701.

[3] M. Germano et al. In: Phys. Fluids 3 (1991), pp. 1760-1765.

Data-driven closure modeling

- There is some evidence for the universality of smaller scales in particular flow classes.[5]
- Let coarse-grained Navier-Stokes evolve large-scale structures and let data-driven ideas model sub-grid scales.
- Build generalizable closures through physics-discerning machine learning formulations across different classes.

$$\frac{\partial \bar{u}_i}{\partial t} + \frac{\partial \bar{u}_i \bar{u}_j}{\partial x_j} = -\frac{\partial \bar{P}}{\partial x_i} + \frac{1}{Re} \frac{\partial^2 \bar{u}_i}{\partial x_j \partial x_j} + \frac{\partial \tilde{\tau}_{ij}^R}{\partial x_j}$$

Try and predict $\tilde{\tau}_{ij}^R$ using a machine learning framework - three-dimensional turbulence

$$\frac{\partial \bar{\omega}}{\partial t} + J(\bar{\omega}, \bar{\psi}) = \frac{1}{Re} \nabla^2 \bar{\omega} + \tilde{\Pi}$$

Try and predict $\tilde{\Pi}$ using a machine learning framework - two-dimensional turbulence

[5] Vreman In: Phys. Fluids, 16, 3670-3681.

What is a desirable data-driven closure

1. Should require **no direct numerical simulation (DNS) flow-fields.** Should work in **data-limited scenarios.**
2. Should require **no knowledge of *true* low-pass spatial filter.** (Impossible to know anyway for anything realistic).
3. When possible, it **should use derived quantities.**

Strategy: **PDE-constrained optimization - "A-posteriori learning"**

Some (not all) recent work:

- Chen, Xianyang, Jiacai Lu, and Grétar Tryggvason. "Finding closure models to match the time evolution of coarse grained 2D turbulence flows using machine learning." Fluids 7, no. 5 (2022): 154.
- List, Björn, Li-Wei Chen, and Nils Thuerey. "Learned turbulence modelling with differentiable fluid solvers: physics-based loss functions and optimisation horizons." Journal of Fluid Mechanics 949 (2022): A25.
- Sirignano, Justin, Jonathan F. MacArt, and Jonathan B. Freund. "DPM: A deep learning PDE augmentation method with application to large-eddy simulation." Journal of Computational Physics 423 (2020): 109811.

Other desiderata:

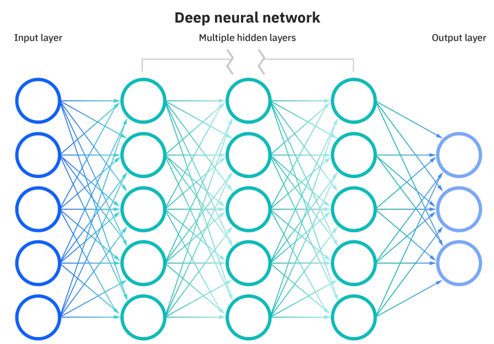
4. Should be compatible with **unstructured, anisotropic, potentially time-varying, adaptive grids.**
5. We wish to **avoid the redevelopment of a forward solver.**
6. **Amenable to numerical analysis** to identify **sources of error**.
7. **Quantify uncertainty** during deployment.

A-posteriori turbulence modeling

Background: Neural ordinary differential equations

Given snapshots of *u* and an assumption of data being generated from autonomous systems - our goal is to identify *f* in:

$$\frac{du}{dt} = f(u(t))$$



With snapshots of training data

$$\{u_0, u_1, \ldots, u_M\}$$

Using a loss-function as follows for various \tau:

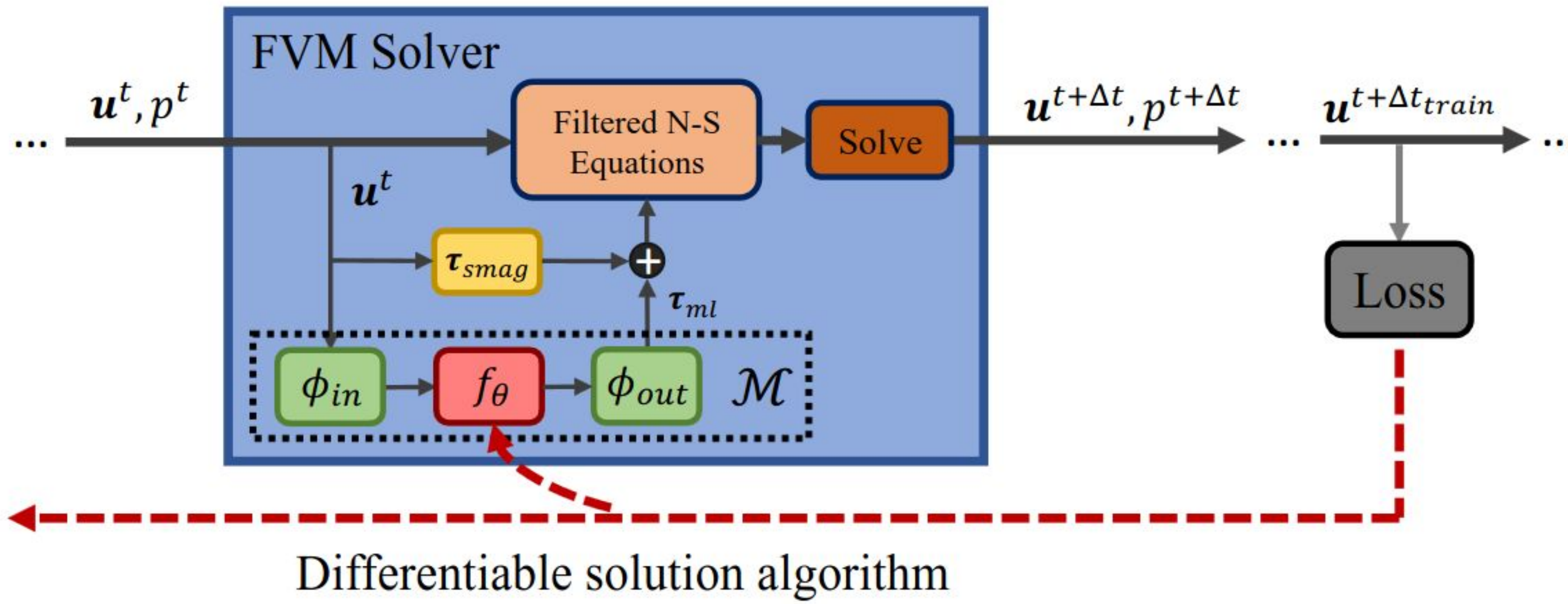
$$\left\langle \sum_{j=1}^{K} ||u_{i+j} - u(t_{i+j})|| \right\rangle$$

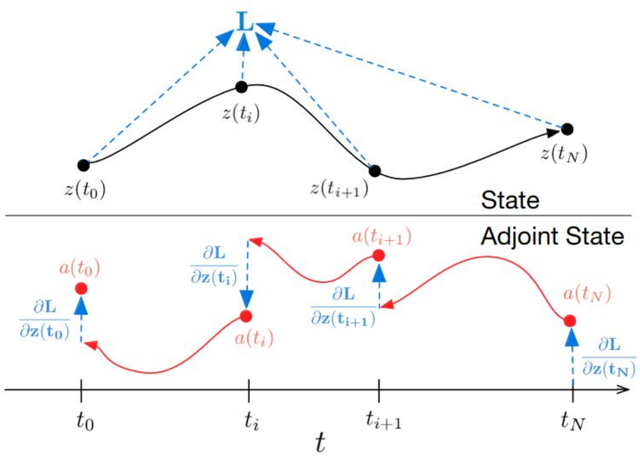
True trajectory

Predicted trajectory

Chen et al., NeurIPS 2018 (Best paper award)
Previously also explored by Kevrekidis and collaborators in early 90s

The *differentiable physics* approach

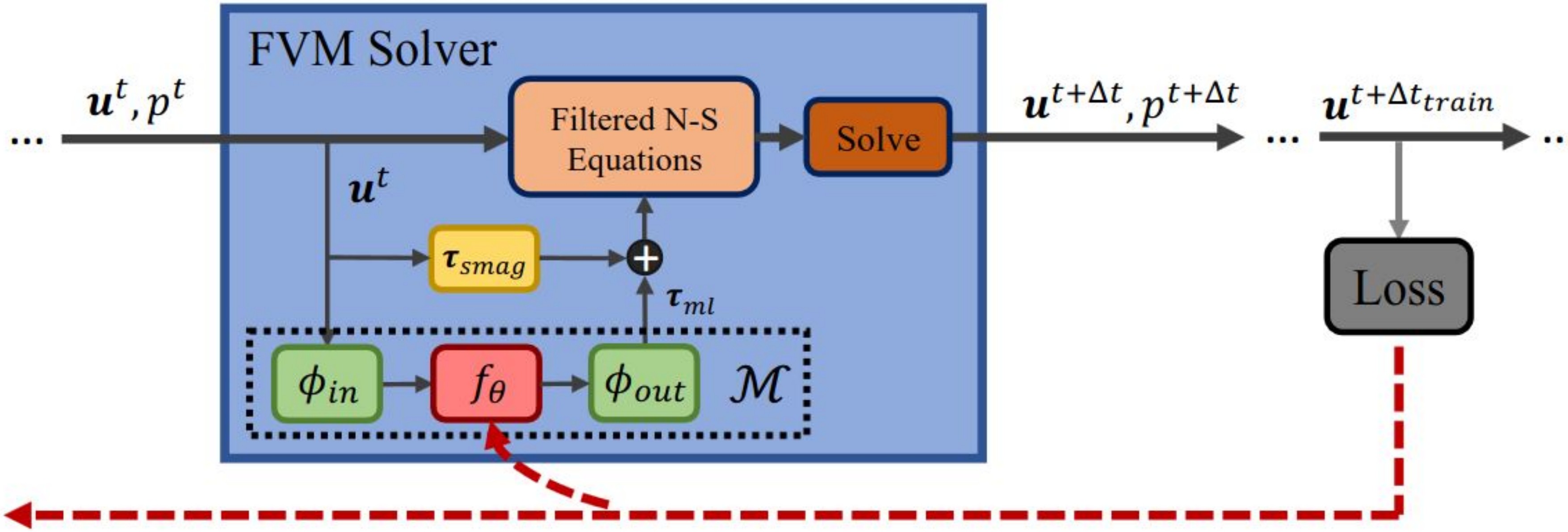




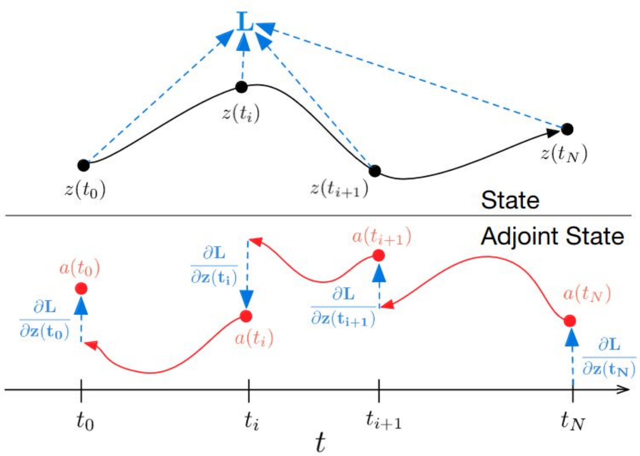
Goal: Leverage rich developments in adjoint-based techniques for numerical solutions of partial differential equations.

V. Shankar, D. Chakraborty, V. Vishwanathan, R. Maulik: Differentiable turbulence I: LES closure as a PDE-constrained optimization, arXiv:2307.03683

The *differentiable physics* approach



Differentiable solution algorithm



$$\hat{\bar{\mathbf{u}}}^0 \dots \hat{\bar{\mathbf{u}}}^t = \mathcal{S}_{ML}(\bar{\mathbf{u}}^0, C_s, \mathcal{M}(\theta)),$$

$$\bar{\mathbf{u}}^0 \dots \bar{\mathbf{u}}^t = G_\Delta \star \mathcal{S}_{DNS}(\mathbf{u}^0),$$

$$\min_{C_s,\theta} \mathcal{L}(\bar{\mathbf{u}}^0 \dots \bar{\mathbf{u}}^t, \mathcal{S}_{ML}(\bar{\mathbf{u}}^0, C_s, \mathcal{M}(\theta))),$$

Note: Target data is merely partial observation of DNS! No filter assumptions.

Chaotic differentiable physics

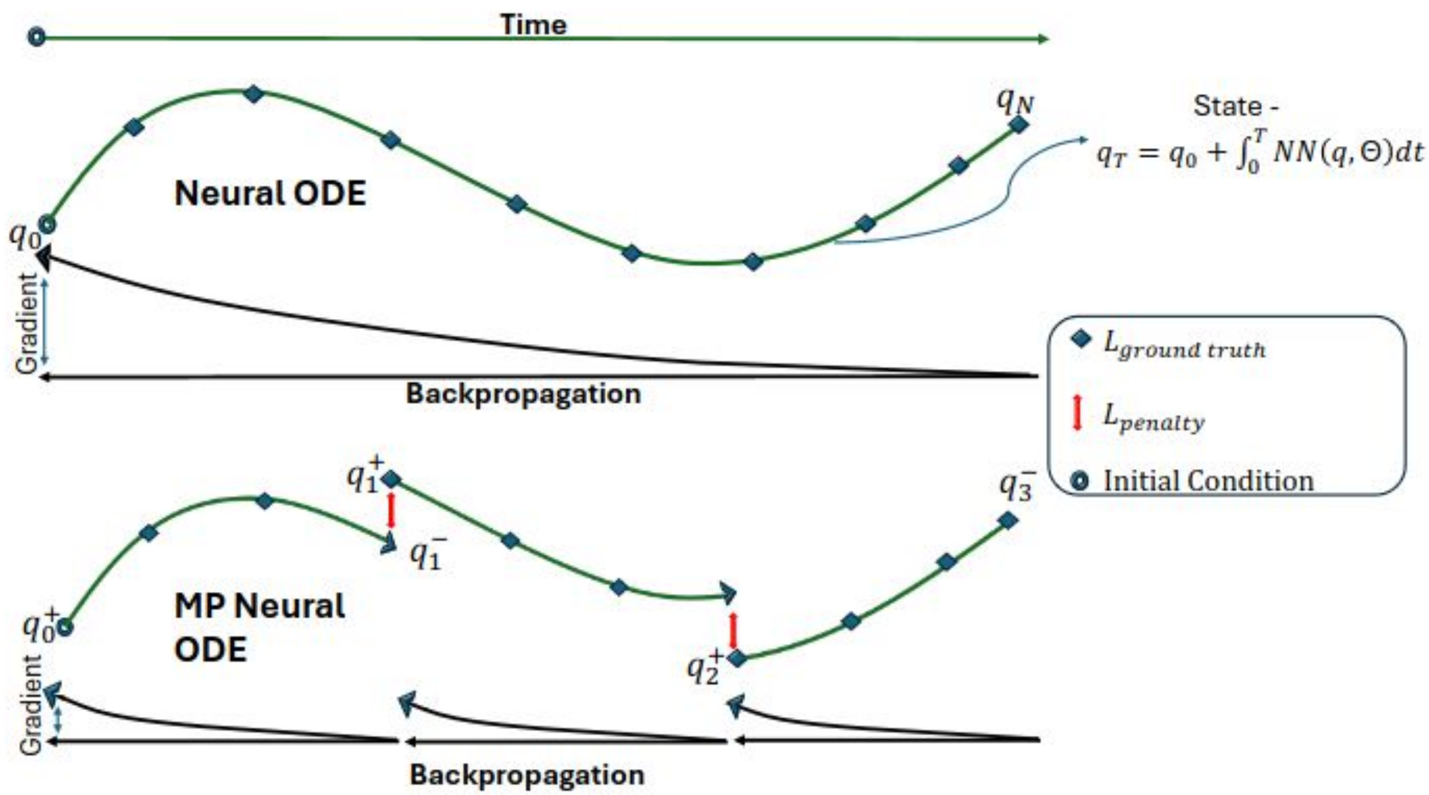
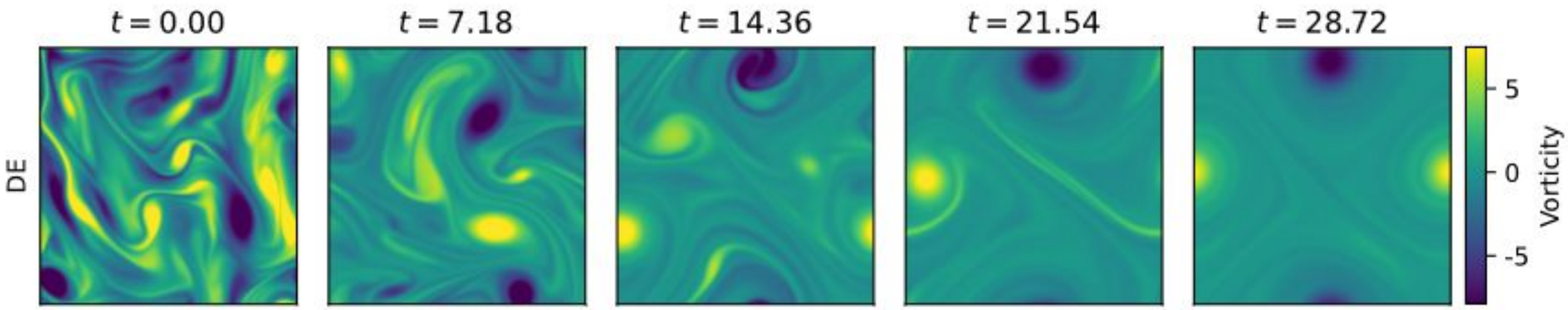


Figure 1: A schematic for multistep penalty based optimization of a data-driven dynamical systems.Discontinuities are introduced into the state evolution as learnable parameters (denoted the 'penalty' term in the loss) during the process of optimization. Over the course of the optimization, the penalty term is driven to zero.

Computing sensitivities of chaotic systems brought from cubic to linear complexity!
Chakraborty, D., Chung, S. W., & Maulik, R. (2024). Divide And Conquer: Learning Chaotic Dynamical Systems With Multistep Penalty Neural Ordinary Differential Equations. arXiv preprint arXiv:2407.00568 (CMAME to appear)

Assessments



Training performance assessed in **a-posteriori deployments** using correlations with randomly sampled-DNS (Re=1000 with primitive formulation). **Note: Training data is merely subsampling a DNS grid (256^2 to 64^2). We don't assume any filter! FDNS = Subsampling (potentially randomly) DNS.**

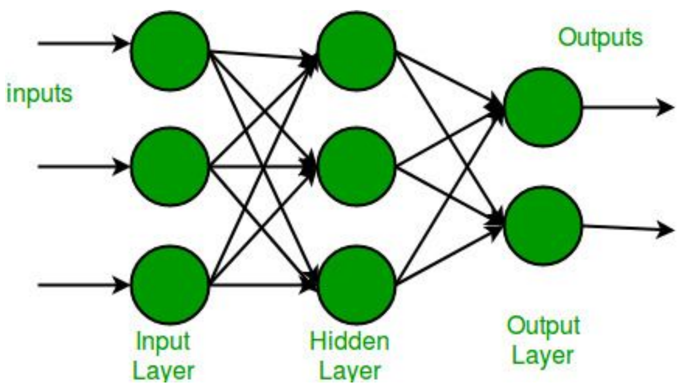
$$\nu_t = (C_s\Delta)^2|\overline{\mathbf{S}}|$$
$$\boldsymbol{\tau}_{smag} = -2\nu_t\overline{\mathbf{S}},$$
$$\hat{\boldsymbol{\tau}} = \boldsymbol{\tau}_{smag} + \boldsymbol{\tau}_{ml},$$
$$\boldsymbol{\tau}_{ml} = \mathcal{M}(\overline{\mathbf{u}}, f_\theta, \phi_{in}, \phi_{out}),$$

$$\phi_{in}(\overline{\mathbf{u}}) = \left\{\overline{\mathbf{S}}^2\right\}, \left\{\overline{\mathbf{R}}^2\right\}$$
$$f_\theta\left(\left\{\overline{\mathbf{S}}^2\right\}, \left\{\overline{\mathbf{R}}^2\right\}\right) = \alpha$$
$$\phi_{out}(\alpha) = \sum_{n=0}^{2} \alpha^{(n)}\mathbf{T}^{(n)},$$

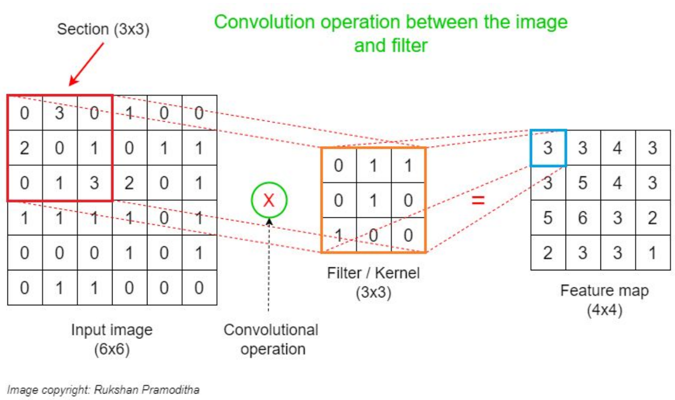
TABLE I. Model names

Name	ϕ_{out}
Linear (LIN)	$\sum_{n=0}^{1} \alpha^{(n)}\mathbf{T}^{(n)}$
Non-linear (NL)	$\sum_{n=0}^{2} \alpha^{(n)}\mathbf{T}^{(n)}$
Non-linear asymmetric (NLA)	$\sum_{n=0}^{3} \alpha^{(n)}\mathbf{T}^{(n)}$
Model-free (MF)	$\alpha\mathbf{I} + \mathbf{D} + \mathbf{A}$
Model-free symmetric (MFS)	$\alpha\mathbf{I} + \mathbf{D}$

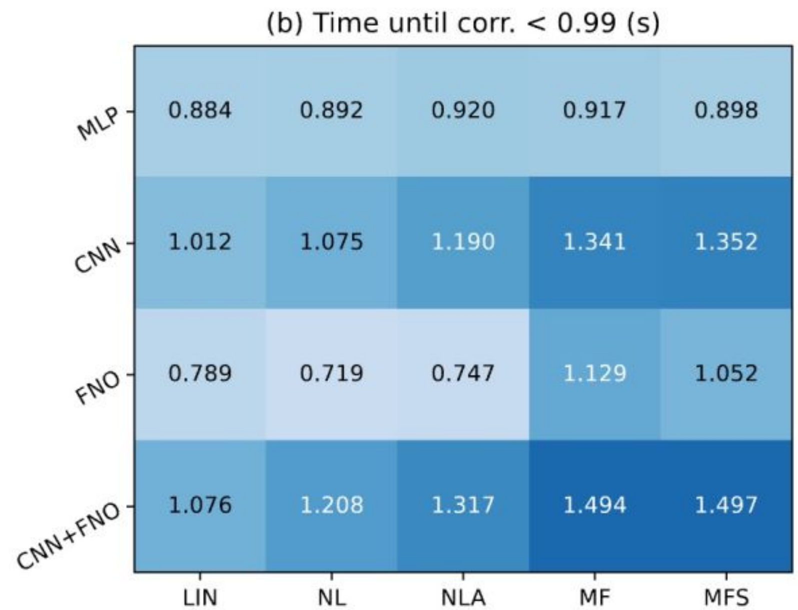
Network architectures



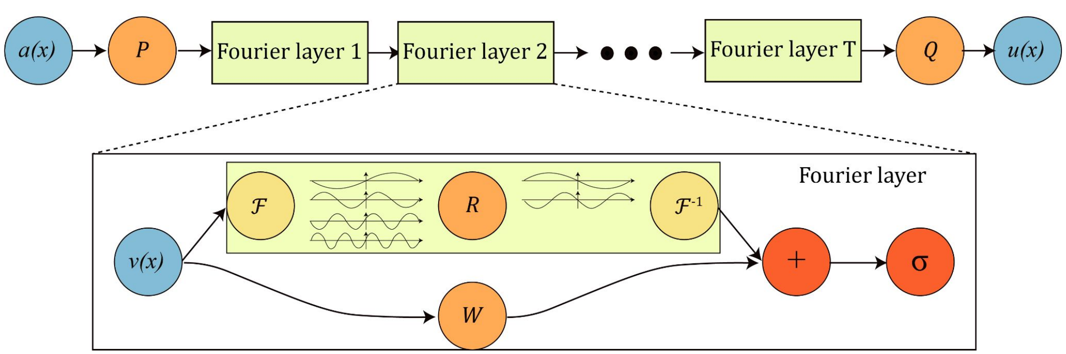
Perfectly local inputs for perfectly local outputs



CNN: Iterative convolution operations for non-local influence on predictions - LeNet (Le-Cun)

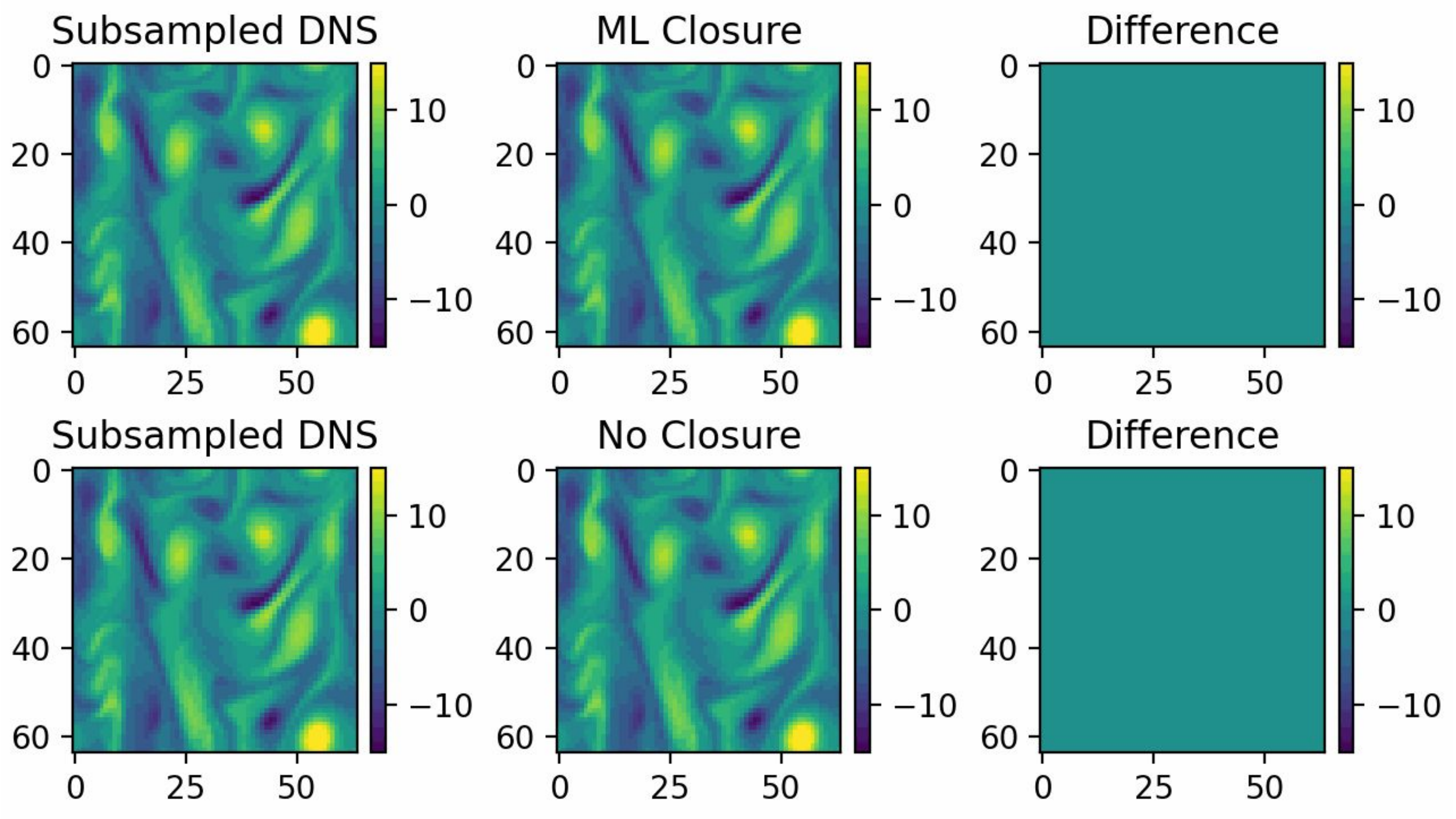


Model-free symmetric superior for training assessments

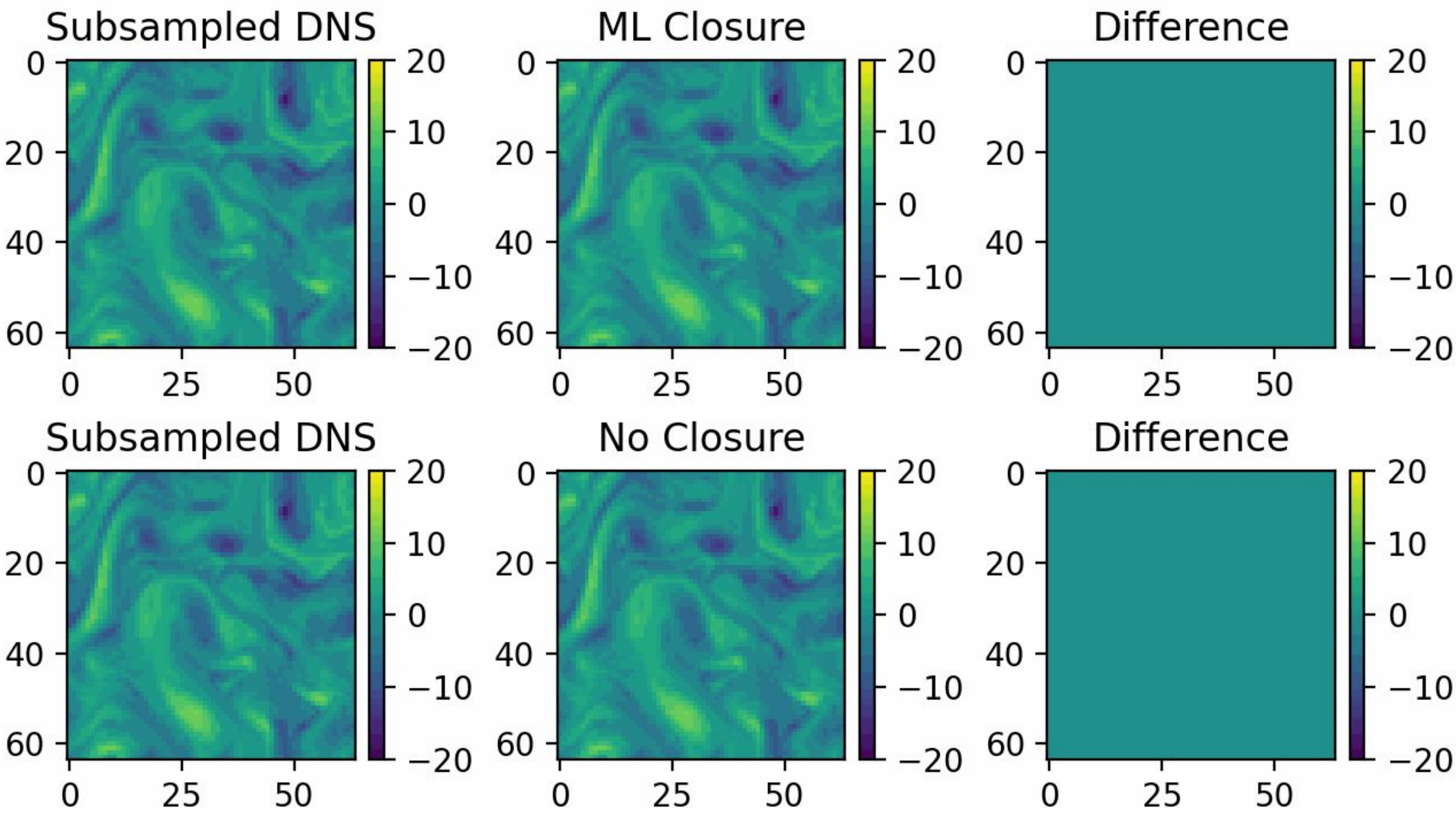


FNO: Fourier Neural Operator that learns a function approximation that is perfectly global

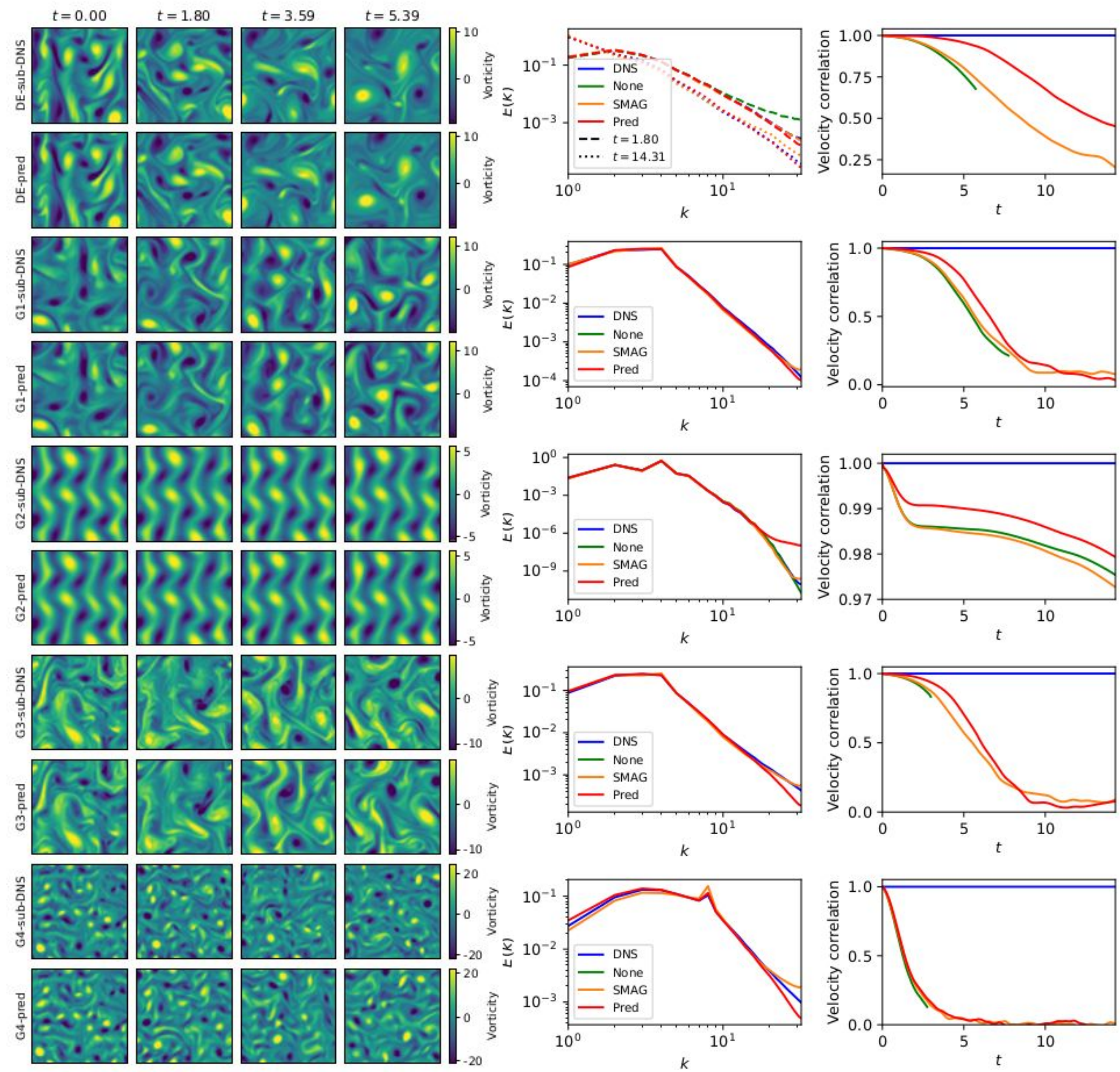
Results - decaying turbulence (Re=8000)



Results - forced turbulence (k=4, Re=1000)



Generalization



Decaying - Re 8000

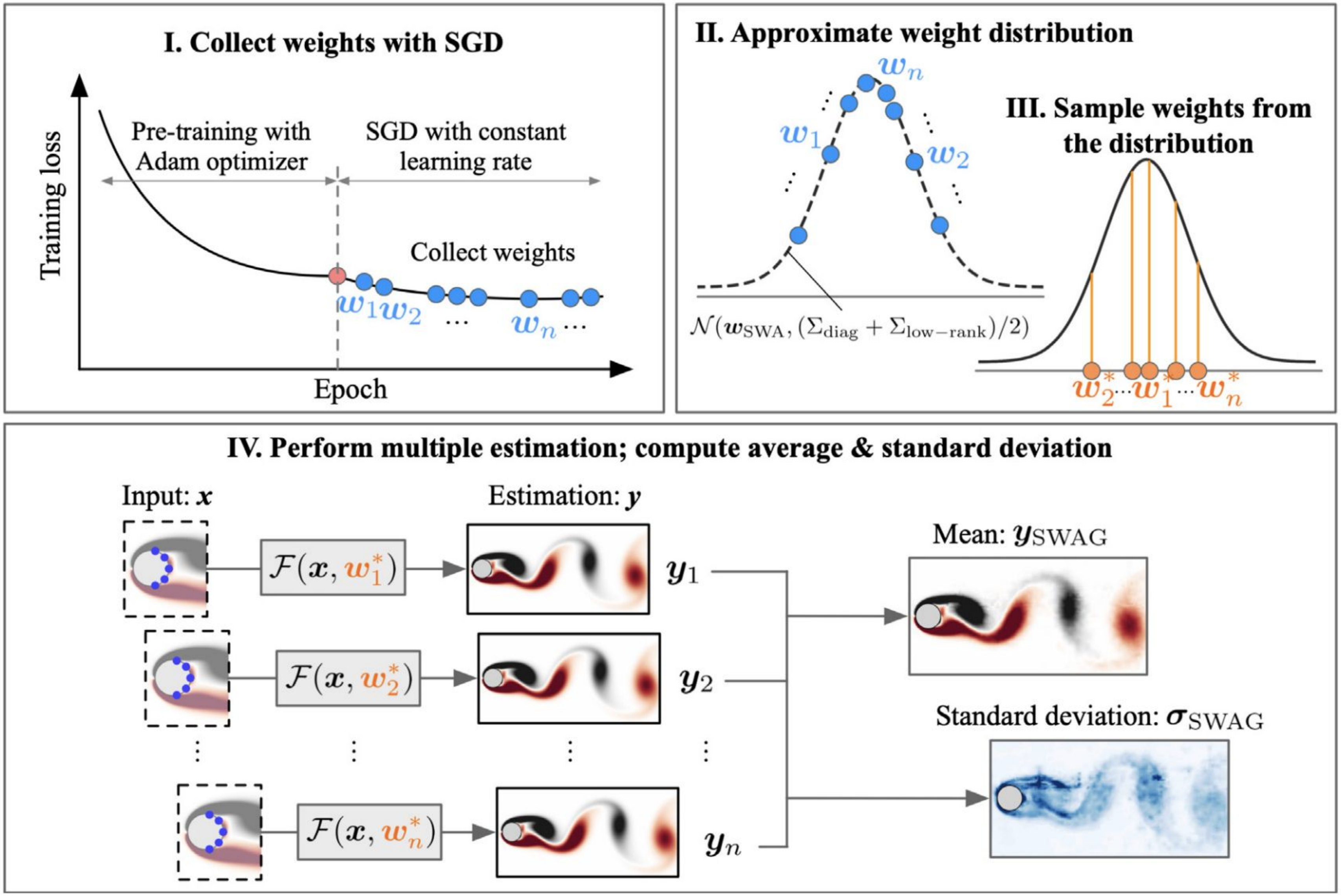
Forced at k=4, Re = 1000

Forced at k=4, Re = 30

Forced at k=4, Re = 8000

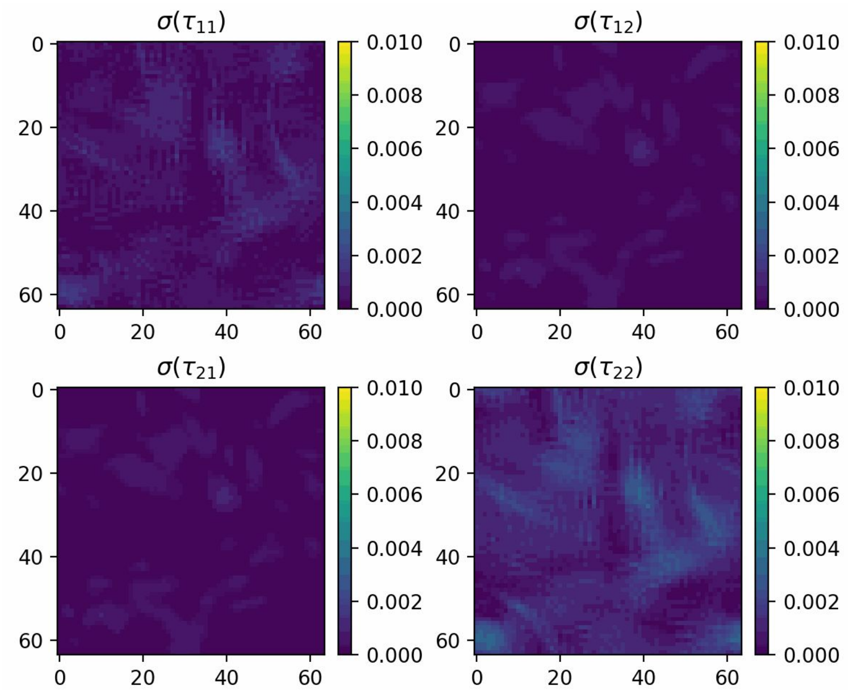
Forced at k=8, Re = 8000

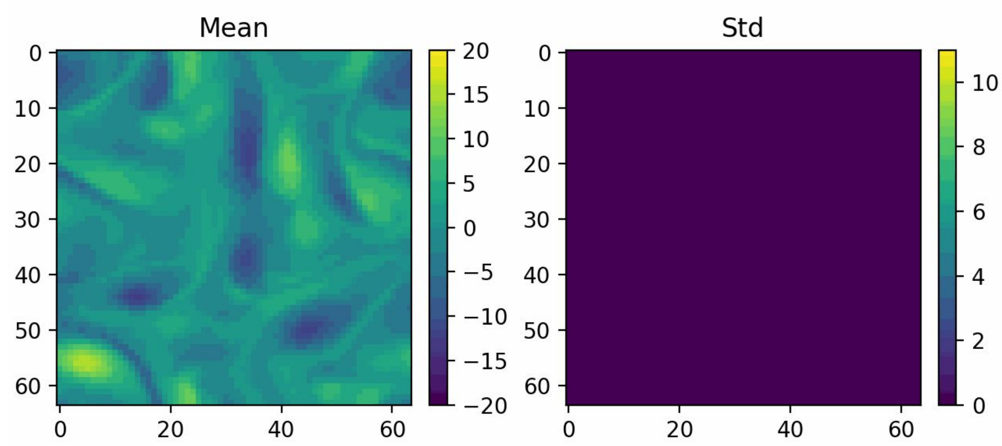
Quantifying uncertainty via deep ensembles



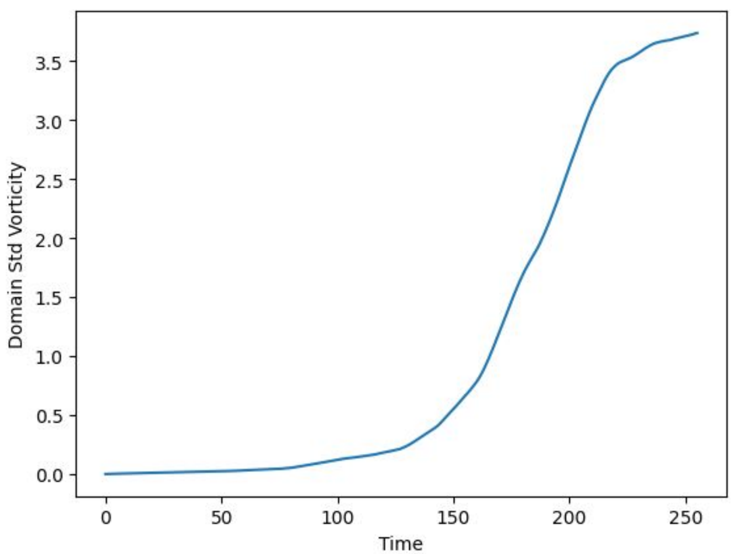
Morimoto, Masaki, Kai Fukami, Romit Maulik, Ricardo Vinuesa, and Koji Fukagata. "Assessments of epistemic uncertainty using Gaussian stochastic weight averaging for fluid-flow regression." *Physica D: Nonlinear Phenomena* 440 (2022): 133454.

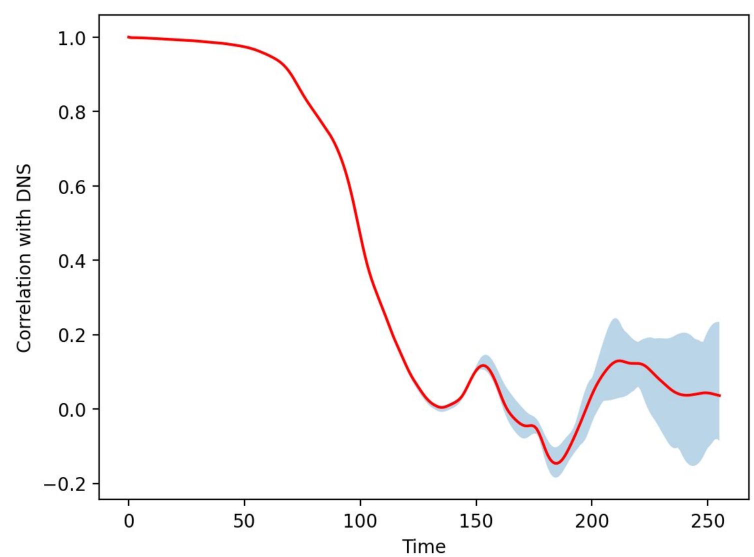
Quantifying uncertainty



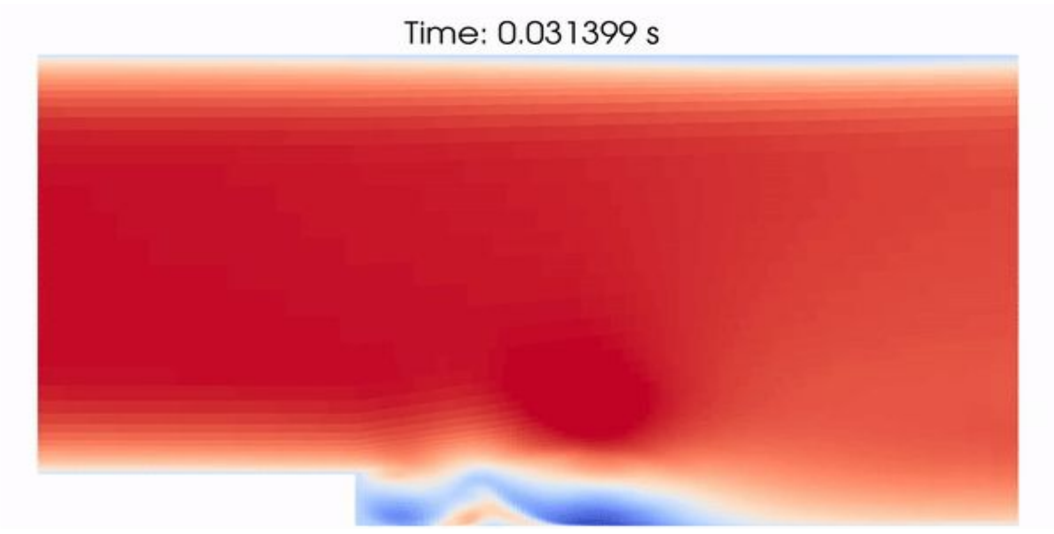


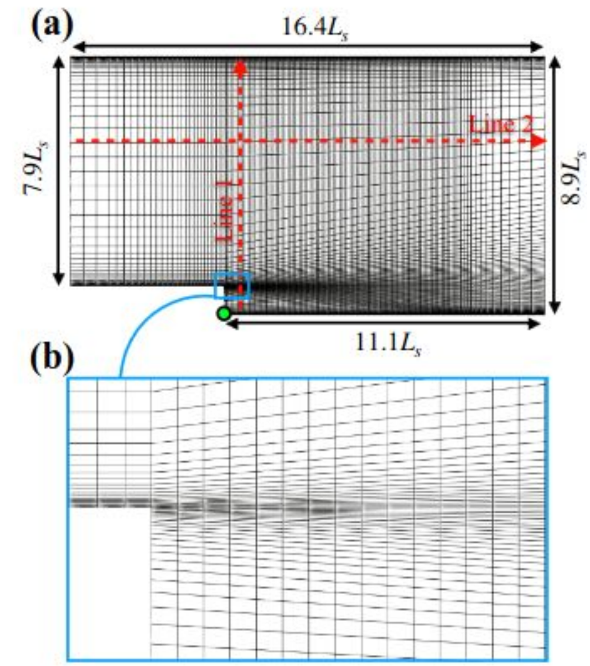
Inflection point in standard deviation of ensemble predictions when correlation with DNS is lost completely.

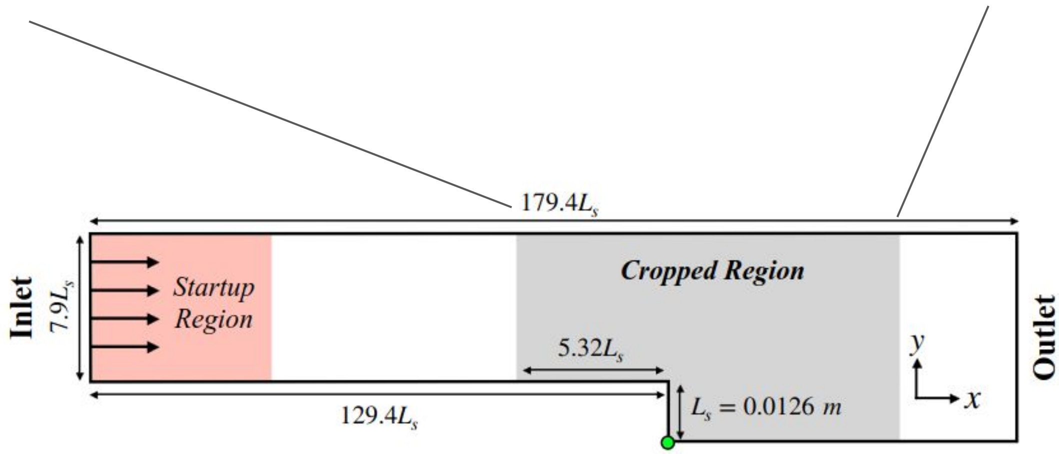




Towards realistic cases

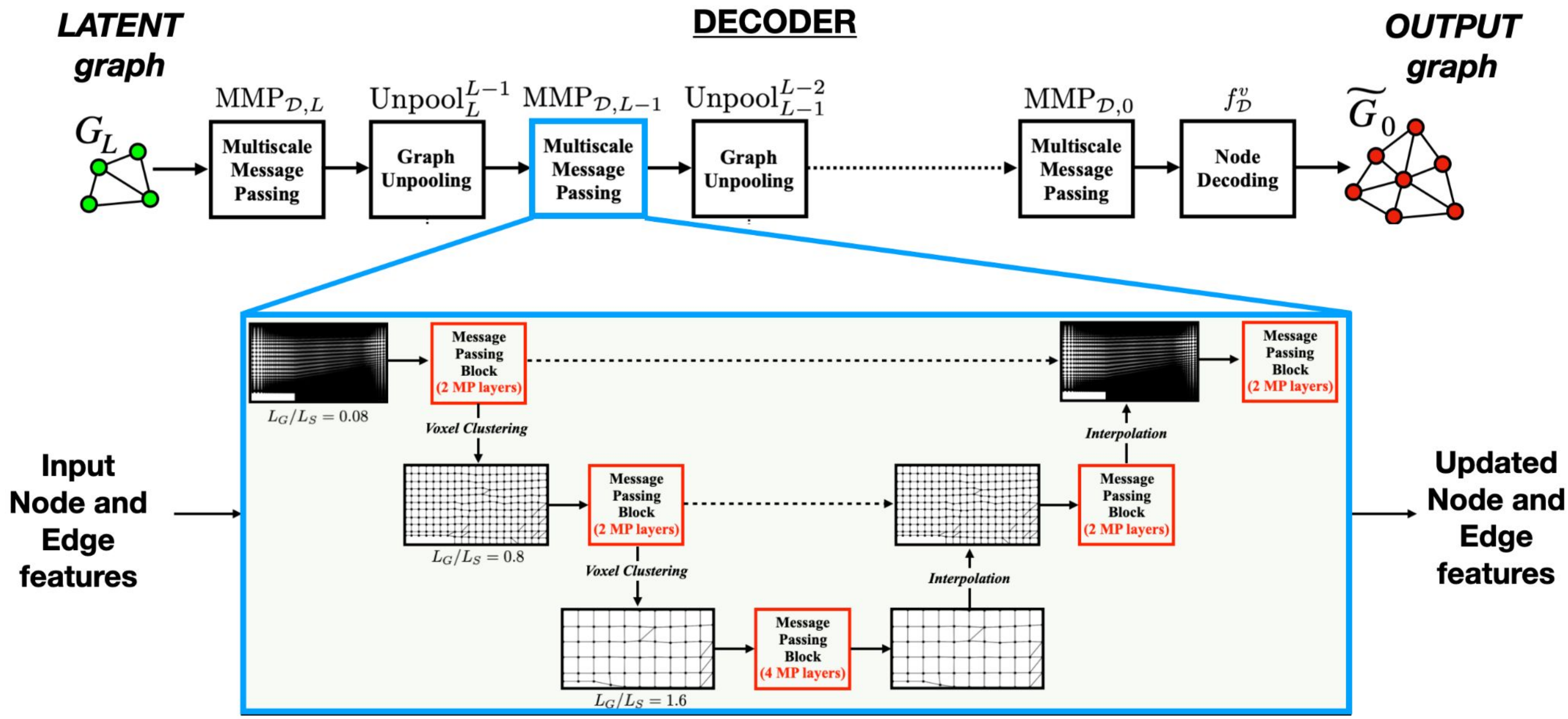






A more realistic test-case for data-driven closure modeling. Characterized by separation, anisotropy, sharp gradients -> not amenable to structured-grid neural net architectures!

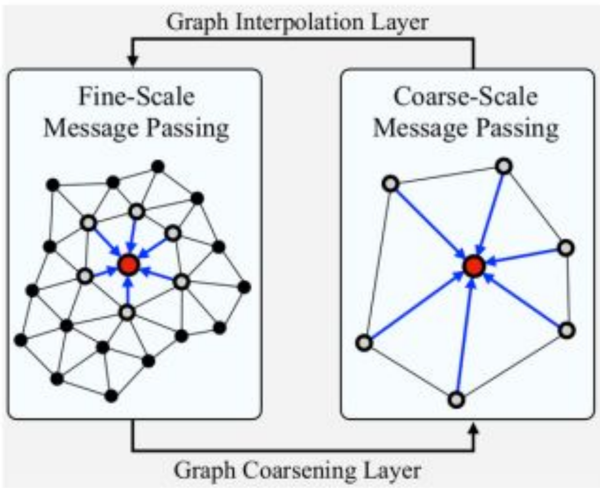
A multiscale graph neural network



Multiscale Graph Neural Network closure for unstructured grids

S. Barwey, V. Shankar, V. Vishwanathan, R. Maulik: Multiscale graph neural network autoencoders for interpretable scientific machine learning, Journal of Computational Physics, 2023

End-to-end differentiable modeling



Turbulence model

$\boldsymbol{\tau}$ →

$$\nabla \bar{u} = 0$$

$$\frac{\partial \bar{\mathbf{u}}}{\partial t} + \nabla \cdot (\bar{\mathbf{u}} \otimes \bar{\mathbf{u}}) = \frac{1}{Re}\nabla^2 \bar{\mathbf{u}} - \frac{1}{\rho}\nabla \bar{p} + \bar{\mathbf{f}} + \nabla \cdot \boldsymbol{\tau}$$

$$J = J(\bar{u})$$

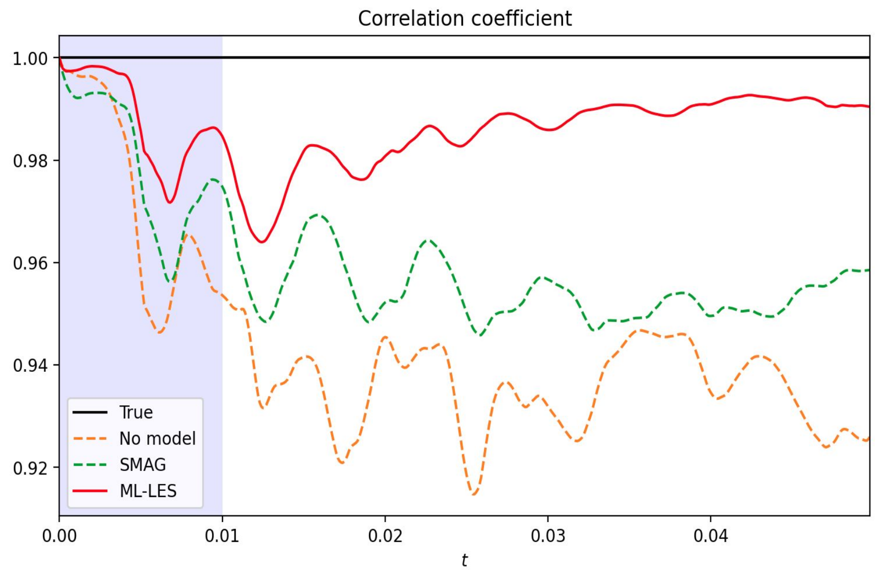
Observation operator

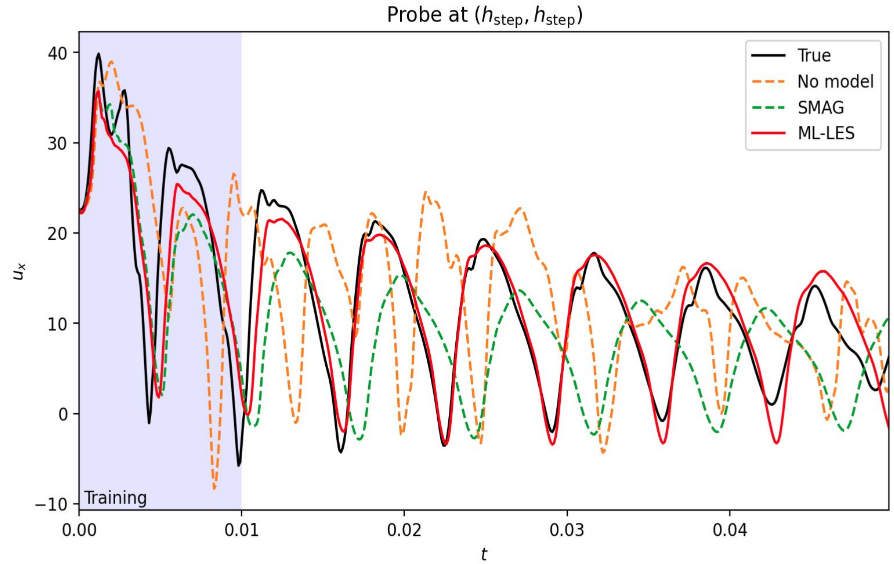
$$\boxed{\frac{\partial J}{\partial \boldsymbol{\theta}}} \xleftarrow{\text{Backpropagation}} \frac{\partial J}{\partial \boldsymbol{\tau}} \xleftarrow{\text{Adjoint model}} J$$

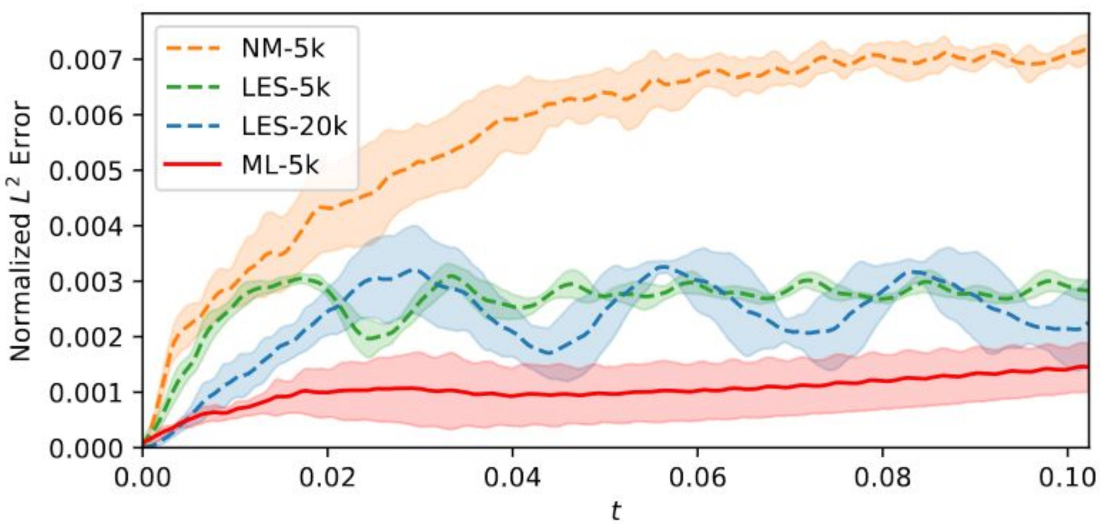
Goes to a stochastic gradient descent type optimizer (here ADAM)

Computed via ground truth trajectories at some sensor locations in flow-field. In this case - a grid that is ~15x finer. But could be experimental observations

Quantitative metrics







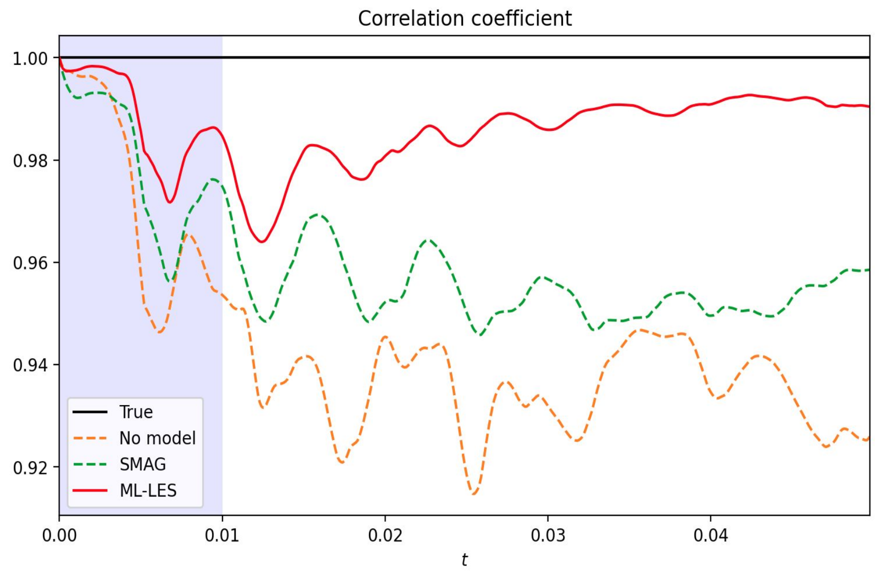
$$L = \frac{1}{T}\sum_{t=0}^{T}\int_{\Omega} ||\bar{\mathbf{u}}_t - \hat{\mathbf{u}}_t||_2^2 \, d\Omega,$$

Training time: 6 hours on 2 V100s (needs optimization)

TABLE I. Train and test cases

Case	$h(m)$	$U_\infty(m/s)$	Re
Train 0	0.0127	32	26,051
Train 1	0.0147	40	37,692
Test 0	0.0127	56	45,590
Test 1	0.0137	48	42,154

Quantitative metrics



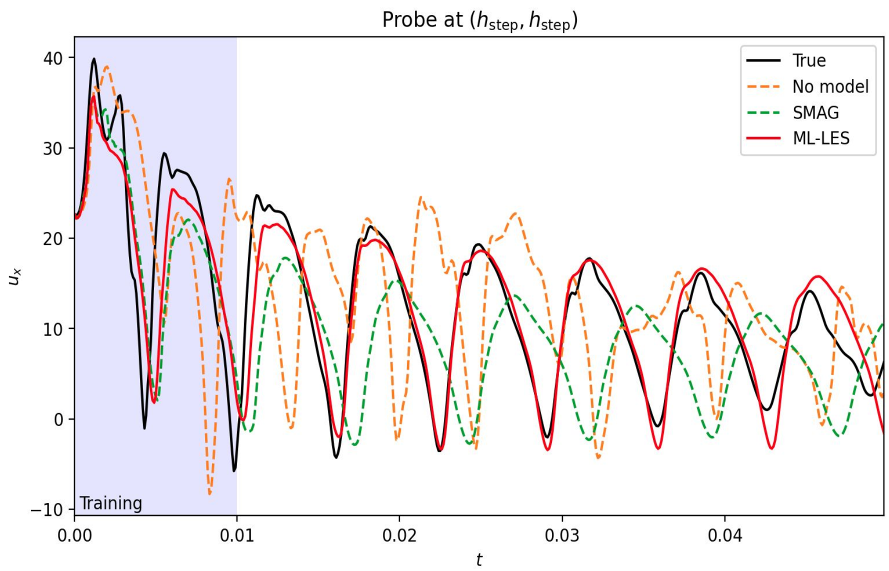


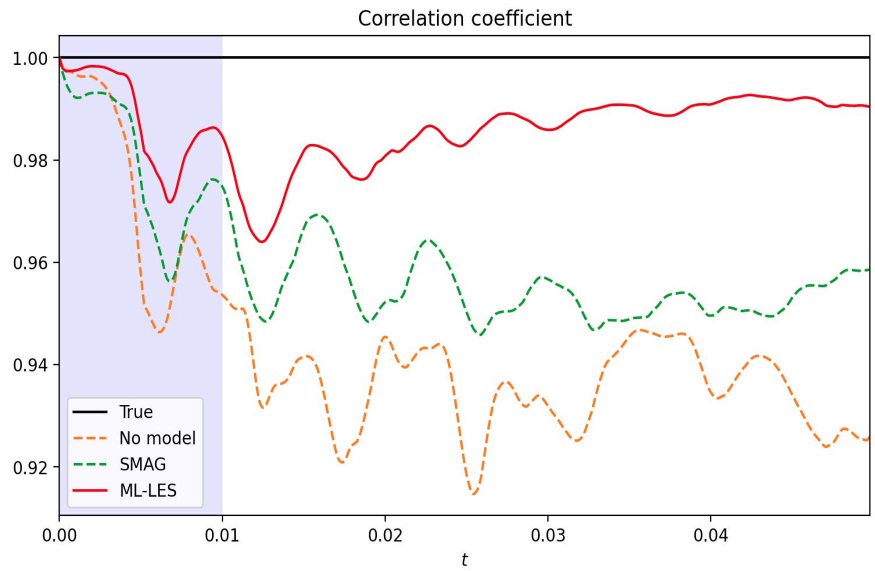
TABLE II. Core/GPU-s per 10^{-4} s of physical time in simulation

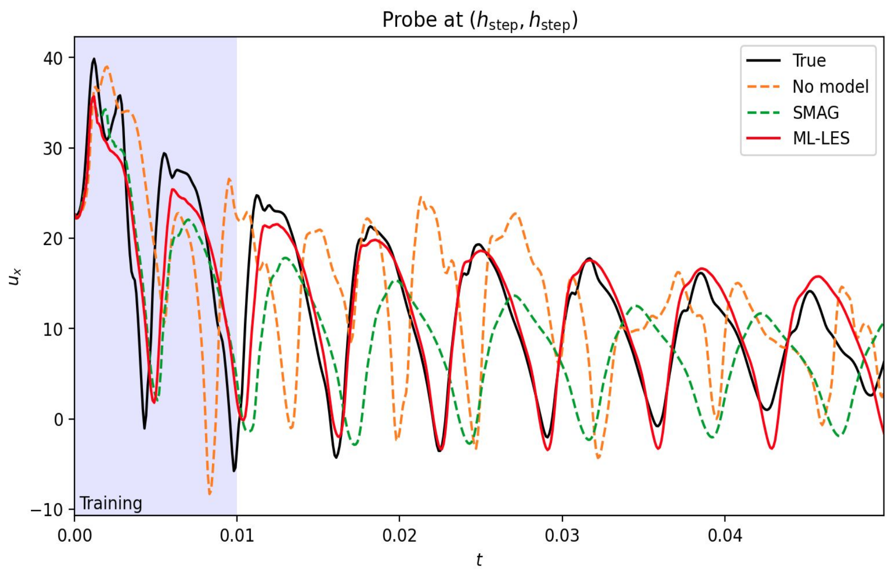
Model	Cost
LES-80k	177 s
LES-20k	7.63 s
LES-5k	0.557 s
ML-5k	0.711 s

$$L = \frac{1}{T}\sum_{t=0}^{T}\int_{\Omega} ||\bar{\mathbf{u}}_t - \hat{\mathbf{u}}_t||_2^2 \, d\Omega,$$

Shankar, Varun, Romit Maulik, and Venkatasubramanian Viswanathan. "Differentiable Turbulence II." *arXiv preprint arXiv:2307.13533* (2023).

What have we learned?





Predictions of a dynamically varying Smagorinsky coefficient from graph multiscale NN

Efforts underway

1. Scaling up cyber-infrastructure for realistic applications.
2. **Multifidelity** data and model fusion during training.
3. Long rollout times with chaotic systems (somewhat solved - happy to discuss more)
4. Theoretical connections to Bayesian inference.

Please reach out for discussions!

Acknowledgements: DOE ASCR/SCIDAC RAPIDS, NSF GRFP

Dr. Shivam Barwey
AETS Named Postdoctoral Fellow
Argonne National Laboratory

Dr. Varun Shankar
NSF GFRP
Carnegie Mellon University

Dibyajyoti Chakraborty
PhD Student
Pennsylvania State University

Haiwen Guan
PhD Student
Pennsylvania State University