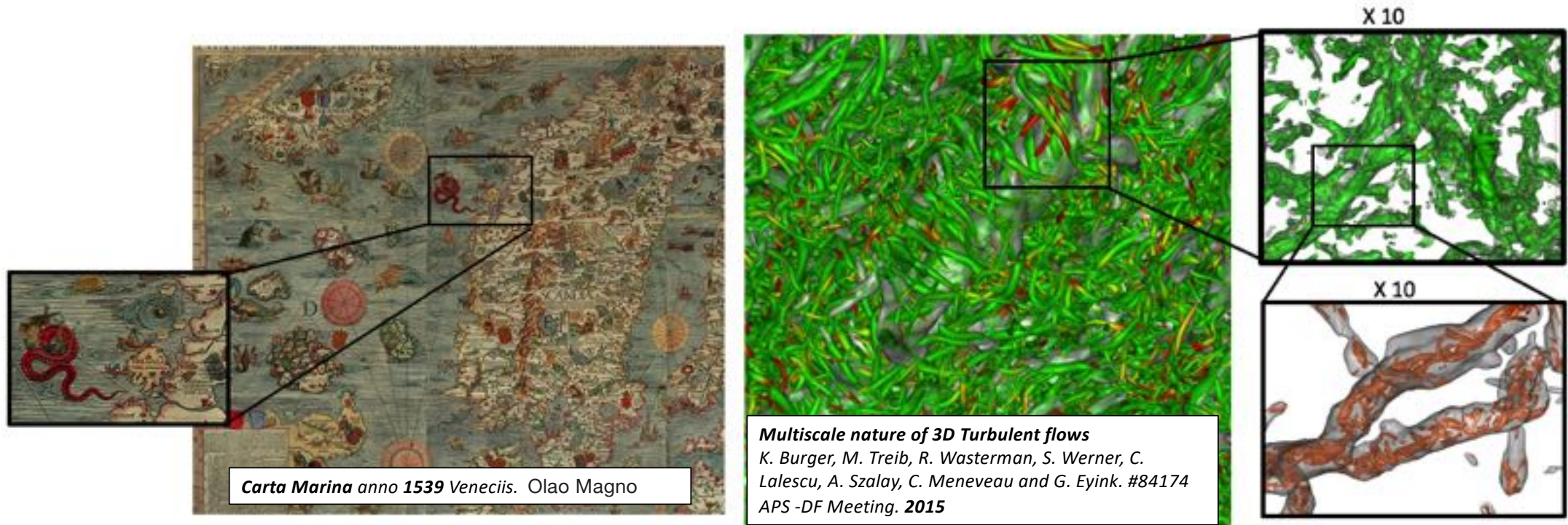


Machine Learning Tools for PDEs Solvers & PDEs Tools for Machine Learning Loops

Luca Biferale Dept. Physics & INFN - University of Rome 'Tor Vergata' biferale@roma2.infn.it

ELLIS-ESA whorshop on CQ and ML for huge data analysis and EO, May 27 2021



CREDITS: M. Buzzicotti, F. Bonaccorso (Univ. Tor Vergata, Rome-IT); A. Mazzino (Univ. Genova, IT); P. Clark di Leoni (JHU, USA)



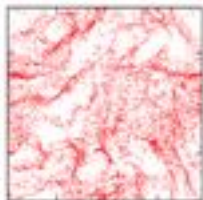
COMPLEX FLUIDS & COMPLEX FLOWS

$$\left\{ \begin{array}{l} \partial_t v + v \cdot \partial v = -\partial P + \nu \partial^2 v + F(B, B) + g\theta + \sum_i c_0(u_i, v) \delta(r - r_i) + f \\ \partial_t \theta + v \cdot \partial \theta = \chi \partial^2 \theta \longleftarrow \text{temperature} \\ \partial_t B + v \cdot \partial B = B \cdot \partial v + \chi \partial^2 B \longleftarrow \text{magnetic field} \\ \partial \cdot v = 0 \\ + \text{boundary conditions} + \text{initial conditions} \end{array} \right.$$

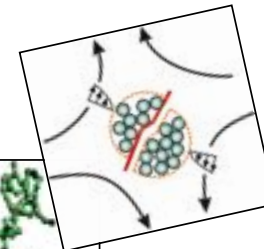
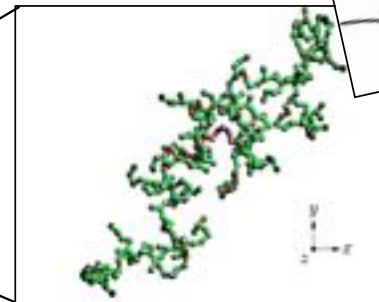
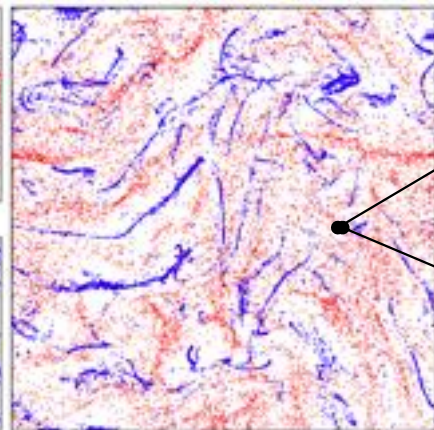
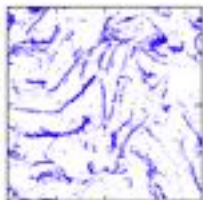
small particles/colloidal aggregates:
Stokes drag, added mass, lift force, etc...

$$\dot{u}_i(t) = -\frac{1}{\tau}(u_i - v) + \beta D_t v$$

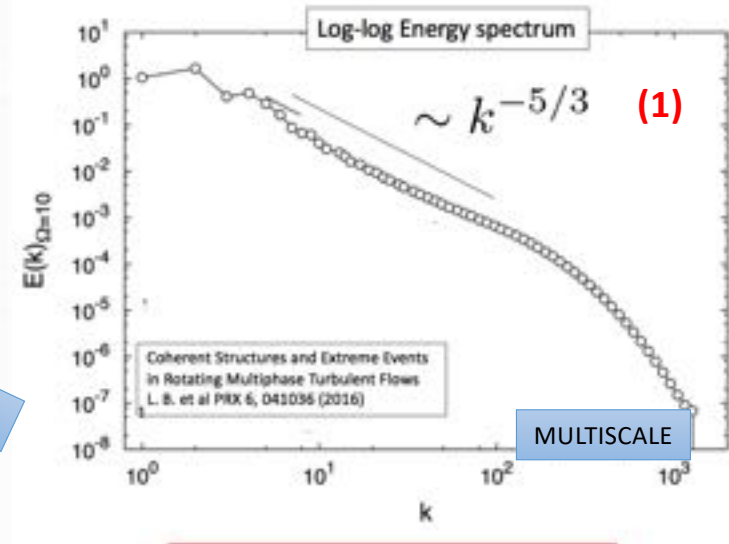
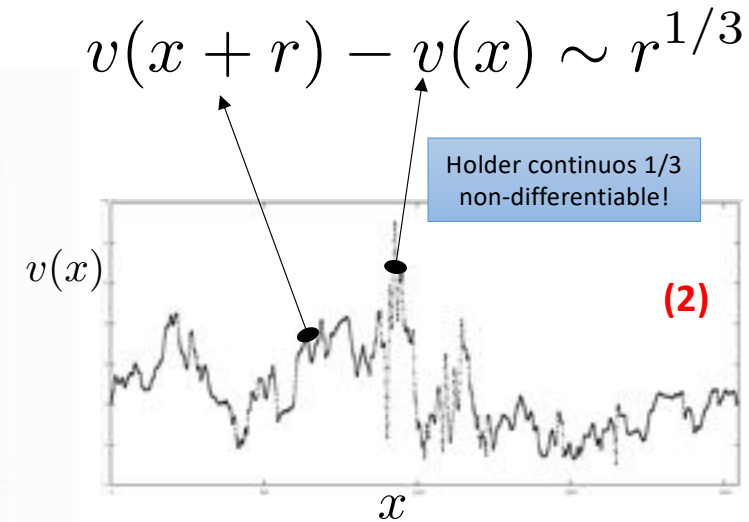
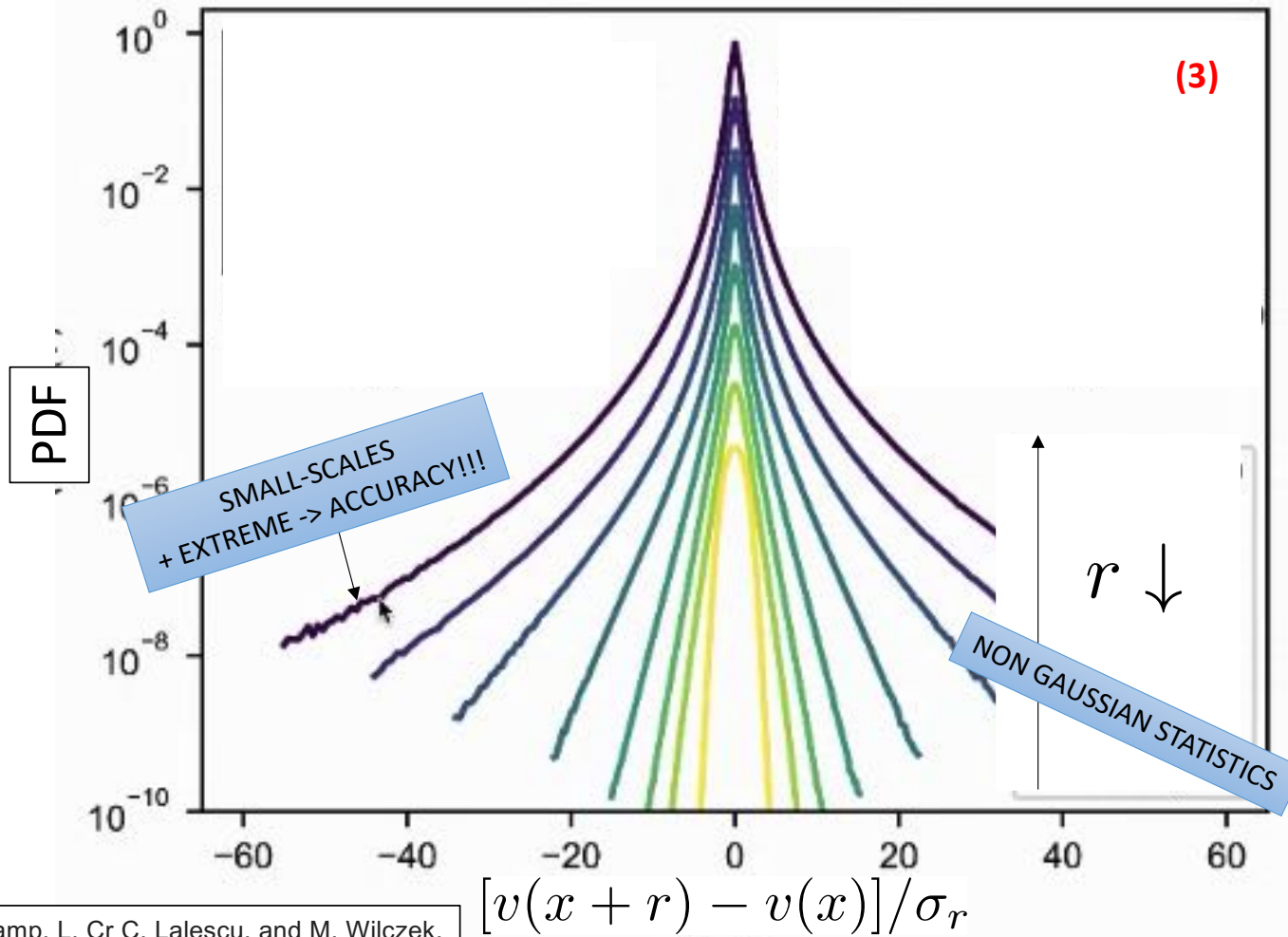
HEAVY
PARTICLES



LIGHT
PARTICLES

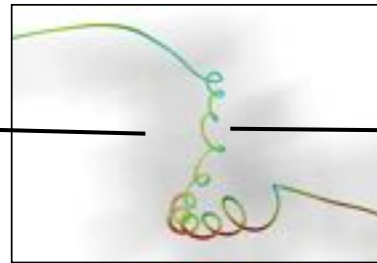
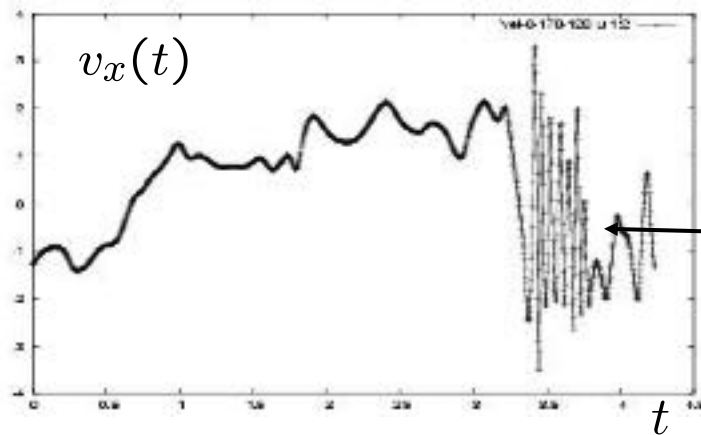


- 1) MULTI-SCALE PHYSICS: BILLIONS OF DEGREES OF FREEDOM
- 2) ROUGH NON-DIFFERENTIABLE FIELDS
- 3) NON-GAUSSIAN STATISTICS

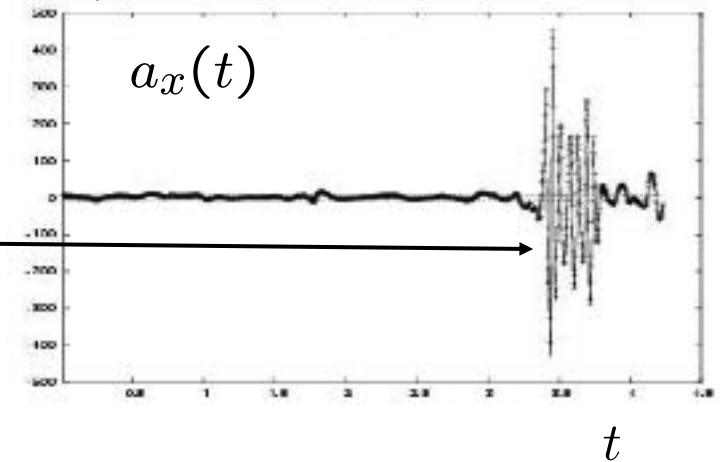


Bentkamp, L, Cr C. Lalescu, and M. Wilczek.
"Nature communications 10.1 (2019): 1-8.

EXTREME EVENTS



L.B. et al Particle trapping in three-dimensional fully developed turbulence
Physics of Fluids 17 (2), 021701 (2005)



Exascale Deep Learning for Climate Analytics. Gordon Bell Award 2018 T. Kurth et al.

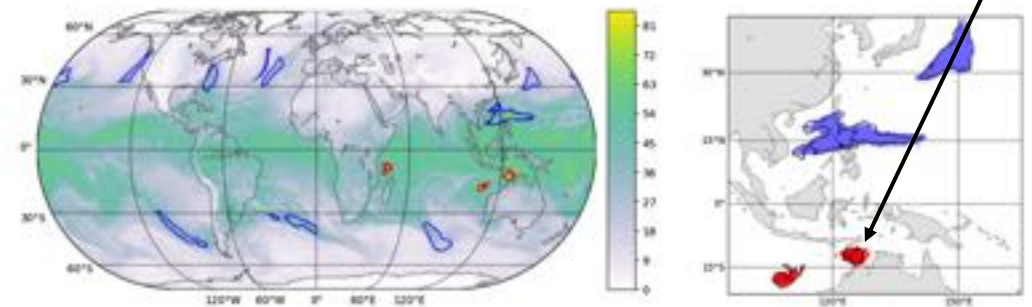


Fig. 7: Segmentation results from modified Deeplabv3+ network. Atmospheric rivers (ARs) are labeled in blue, while tropical cyclones (TCs) are labeled in red.

BAMS Meeting Summary

Machine Learning for Earth System Observation and Prediction

Massimo Bonavita, Rossella Arcucci, Alberto Carrassi, Peter Dueben, Alan J. Geer,
Bertrand Le Saux, Nicolas Longépé, Pierre-Philippe Mathieu, and Laure Raynaud

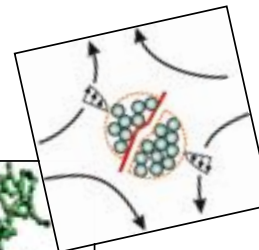
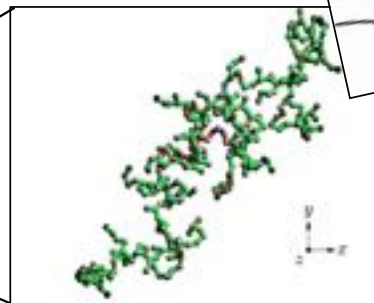
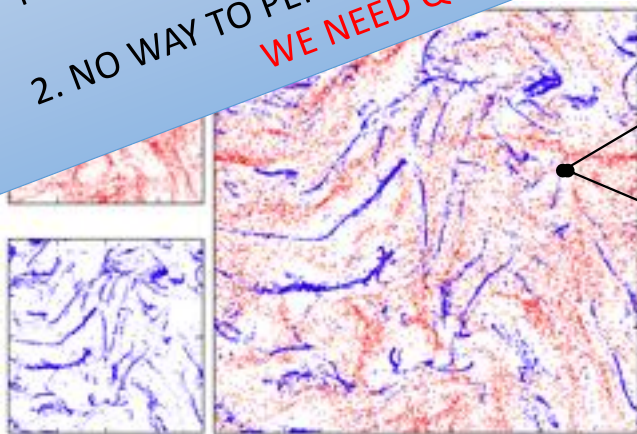
First ECMWF-ESA Workshop on Machine Learning for Earth System Observation
and Prediction

$$\left\{ \begin{array}{l} \partial_t v + v \cdot \partial v = -\partial P + \nu \partial^2 v + F(B, B) + g\theta + \sum_i c_0(u_i, v) \delta(r - r_i) + f \\ \partial_t \theta + v \cdot \partial \theta = \chi \partial^2 \theta \leftarrow \text{temperature} \\ \partial_t B + v \cdot \partial B = B \cdot \partial v + \chi \partial^2 B \leftarrow \text{magnetic field} \\ \partial \cdot v = 0 \\ + \text{boundary conditions + initial conditions} \end{array} \right.$$

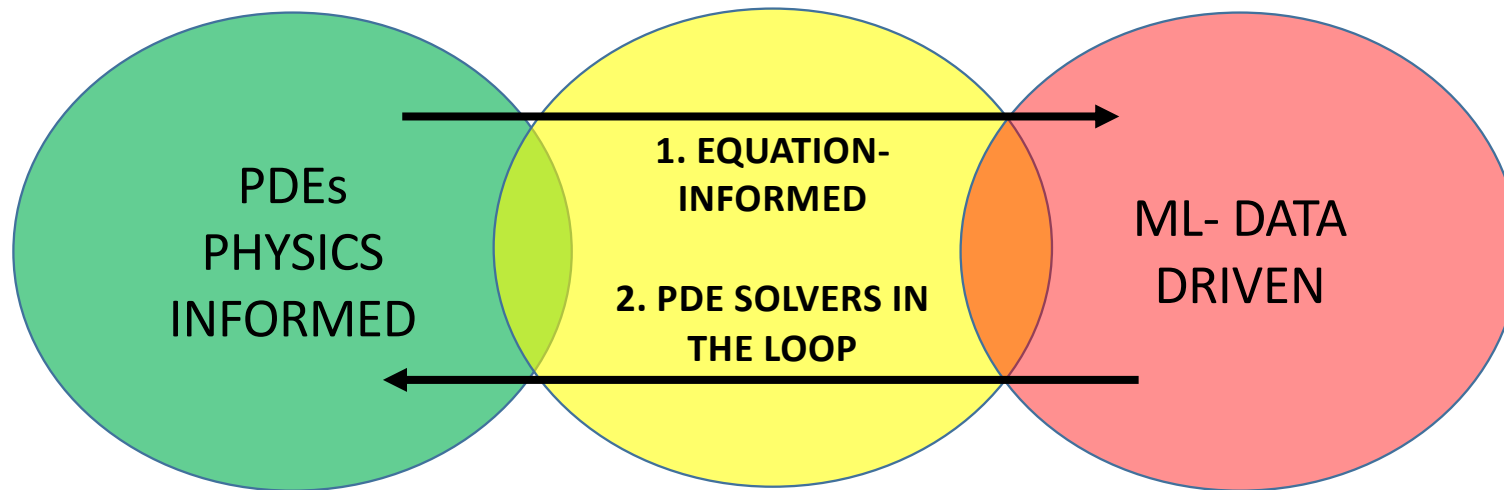
$$\left\{ \begin{array}{l} \frac{du_i(r_i, t)}{dt} = -\rho_f |u_i - v| (u_i - v) \\ + \rho_f \left(\frac{Dv}{Dt} - \frac{Du_i}{Dt} \right) \end{array} \right.$$

gates:
mass, lift force, etc...

1. NO WAY TO PREDICT STATISTICS FOR MEAN PROFILES OR EXTREME EVENTS FROM EoM
(out of equilibrium: no Gibbs distributions!)
2. NO WAY TO PERFORM DIRECT NUMERICAL SIMULATIONS FOR REALISTIC PROBLEMS
- WE NEED QUANTITATIVE MODELS AND TOOLS TO MODEL!**



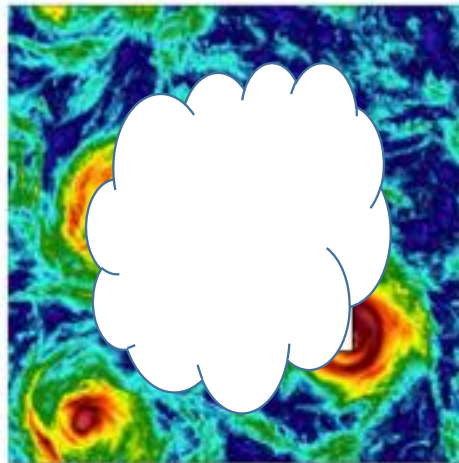
1. ML TOOLS AUGMENTED BY PDEs SOLUTIONS: DATA ASSIMILATION, FLOW RECONSTRUCTION, INPAINTING, SUPER-RESOLUTION
2. PDEs MODELING AUGMENTED by ML: LARGE-EDDY-SIMULATIONS, MODEL REDUCTION, HOMOGENEIZATION



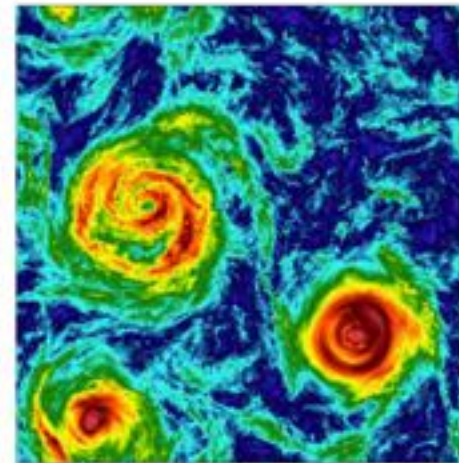
1. DATA ASSIMILATION, FLOW RECONSTRUCTION, INPAINTING, SUPER-RESOLUTION

ML - DATA
DRIVEN

EQUATION FREE
GENERATIVE-ADVERSARIAL-NETWORK



????????



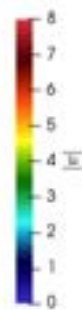
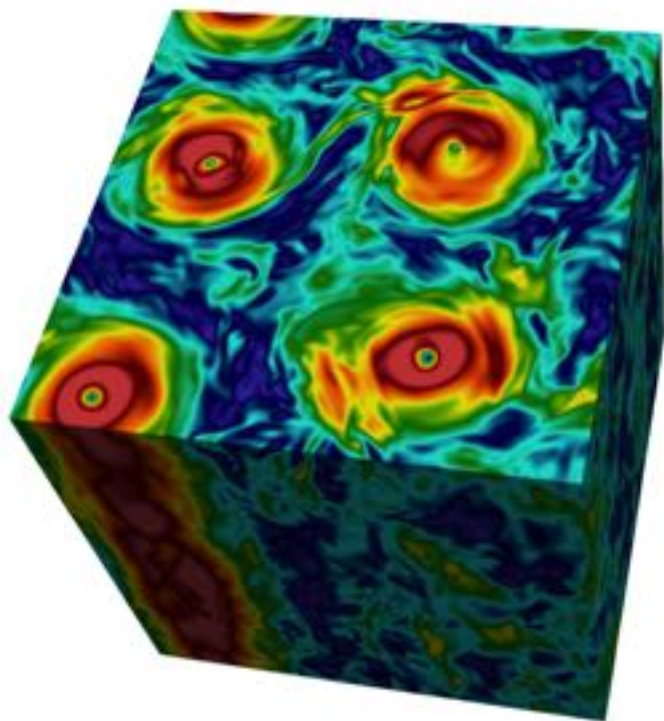
EQUATIONS
BASED

NUDGING

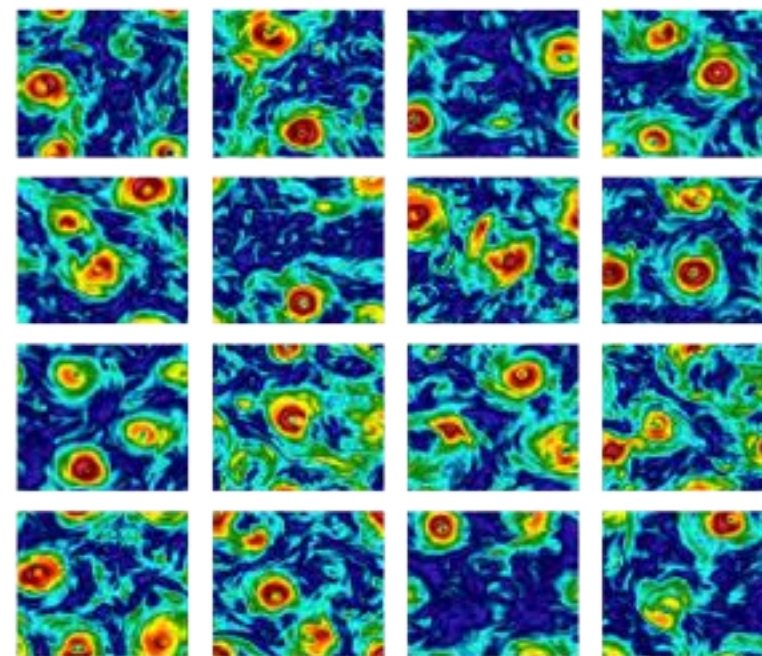
$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} + 2\Omega \hat{x}_3 \times \mathbf{v} = -\nabla p + \nu \nabla^2 \mathbf{v} + \gamma(1 - \hat{M})_{x_3} \odot (\mathbf{v} - \mathbf{v}_{\text{ref}})$$

$$\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} + 2\Omega \hat{x}_3 \times \mathbf{v} = -\nabla p + \nu \nabla^2 \mathbf{v} + \mathbf{f}$$

4096x4096x4096 collocation points

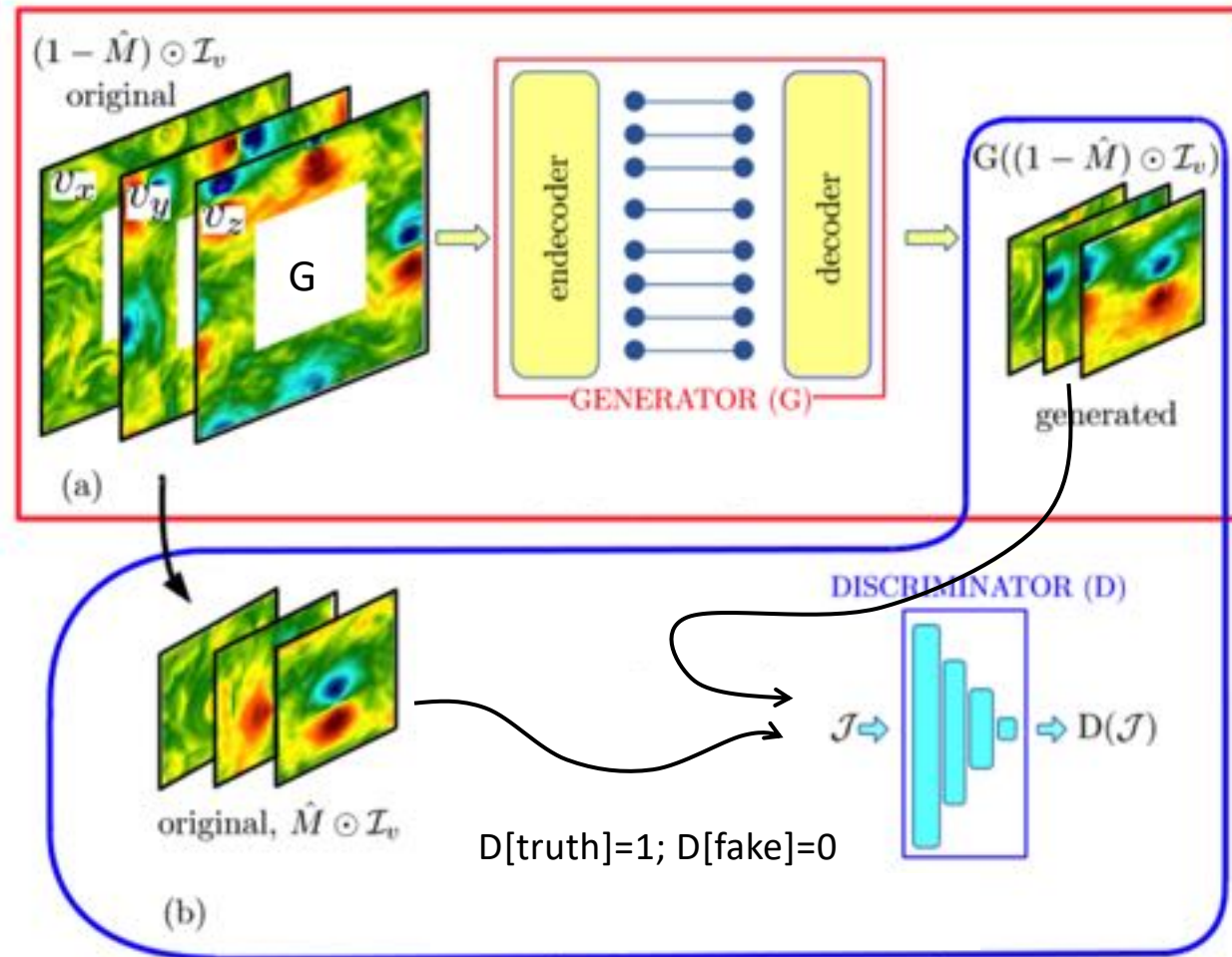


EXTRACT 300K
2D FRAMES
AT 64X64
FOR TRAINING
AND VALIDATION



3D TURBULENCE UNDER ROTATION

GENERATIVE ADVERSARIAL NETWORK: CONTEXT ENCODER



Reconstruction of turbulent data with deep generative models for semantic inpainting from TURB-Rot database

M. Buzzicotti, F. Bonaccorso, P. Clark Di Leoni, and L. B.

Phys. Rev. Fluids 6, 050503, May 2021

MINIMIZE:

$$\mathcal{L}_G = \int_G dx \langle (v_{orig} - v_{gen})^2 \rangle$$

MAXIMIZE:

$$\mathcal{L}_{ADV} = \langle (\log(D[v_{orig}]) + (1 - \log(D[v_{gen}]))) \rangle$$

$$\mathcal{L}_{TOT} = \mathcal{L}_G + \lambda \mathcal{L}_{ADV}$$

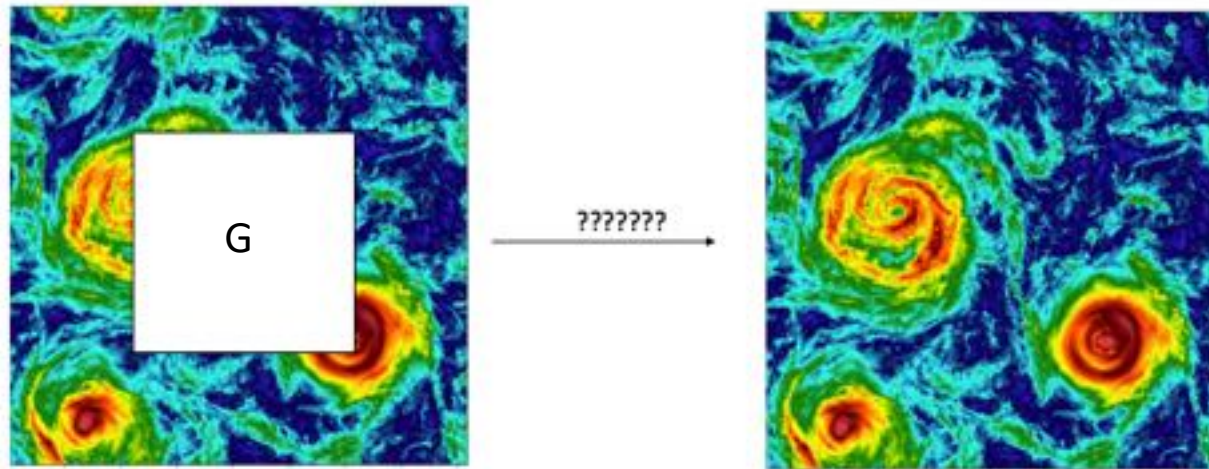
**DATA DRIVEN
NO-EQUATIONS**

- [3] Deepak Pathak, Philipp Krahenbuhl, Jeff Donahue, Trevor Darrell, and Alexei A Efros. Context encoders: Feature learning by inpainting. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2536–2544, 2016.

NUDGING: AN EQUATION-INFORMED UNBIASED TOOL TO ASSIMILATE AND RECONSTRUCT TURBULENCE DATA/PHYSICS BY ADDING A DRAG TERM AGAINST PARTIAL FIELD MEASUREMENTS

C.C. Lalescu, C. Meneveau and G.L. Eyink. Synchronization of Chaos in Fully Developed Turbulence. Phys. Rev. Lett. 110, 084102 (2013)
 A.Farhat, E. Lunasin, and E.S. Titi. Abridged Continuous Data Assimilation for the 2d Navier-Stokes Equations Utilizing Measurements of Only One Component of the Velocity Field. J. Math. Fluid Mech. 18(1), 1 (2016)
 Patricio Clark Di Leoni, Andrea Mazzino, and L.B. Synchronization to Big Data: Nudging the Navier-Stokes Equations for Data Assimilation of Turbulent Flows Phys. Rev. X 10, 011023 (2020)

EQUATIONS
BASED



FULLY PHYSICS
COMPLIANT

NEED HUGE
COMPUTATIONAL
RESOURCES

$$\mathbf{v}_N = G[\mathbf{v}_{true}]$$

$$\mathbf{v}_{true}$$

$$\begin{cases} \partial_t \mathbf{v} + \mathbf{v} \cdot \partial_x \mathbf{v} + \partial_x P - \nu \Delta \mathbf{v} = 2\mathbf{v} \times \boldsymbol{\Omega} + \mathcal{S}\mathbf{v} + \alpha g \hat{\mathbf{z}} T + \mathcal{F} - N(\mathbf{v}_N - \mathbf{v}) \\ \partial_x \mathbf{v} = 0 \end{cases}$$

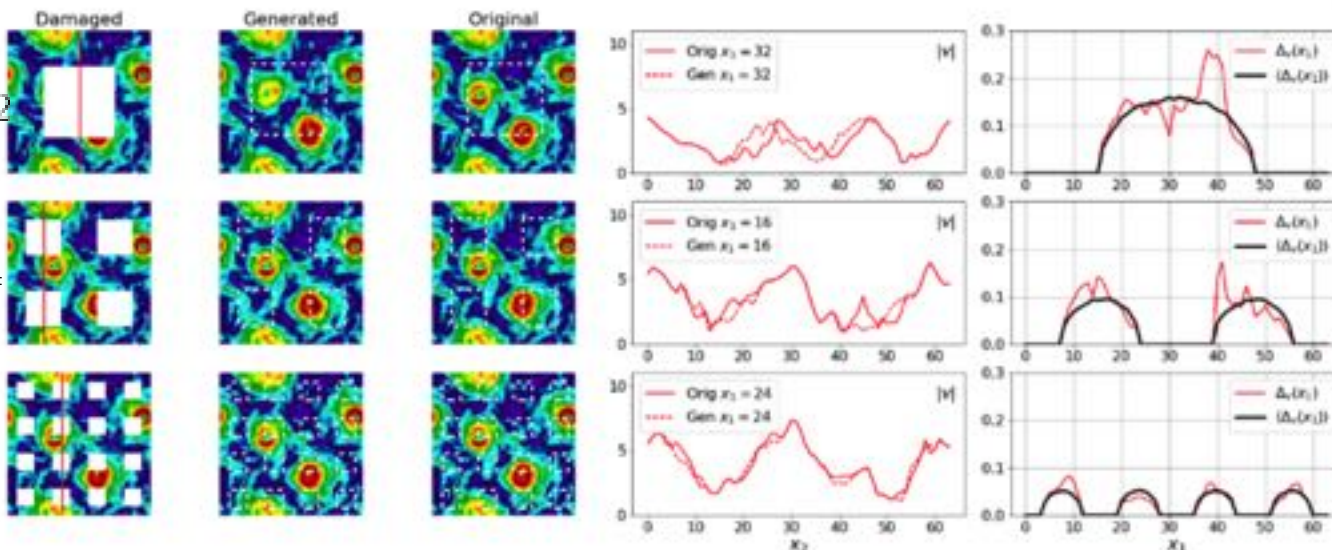
NO NEED TO TRAIN!! NAVIER AND STOKES DID THE JOB FOR YOU: ONE CONF IS ENOUGH

FEATURES RANKING

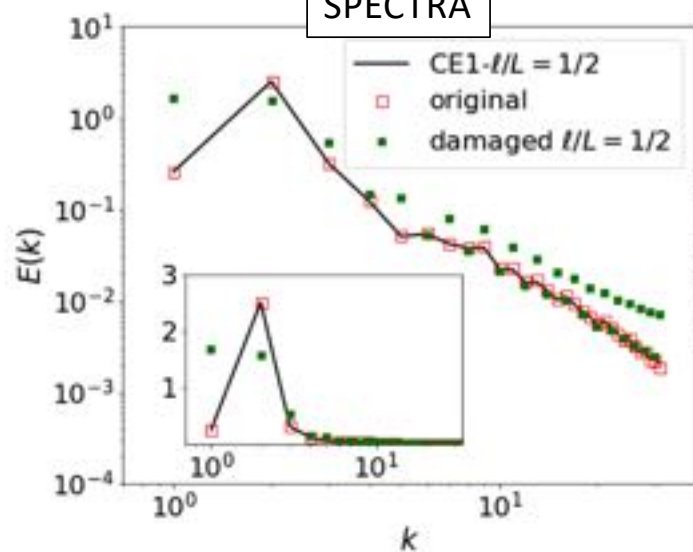
$\ell/L = 1/2$

$\ell/L = 1/4$

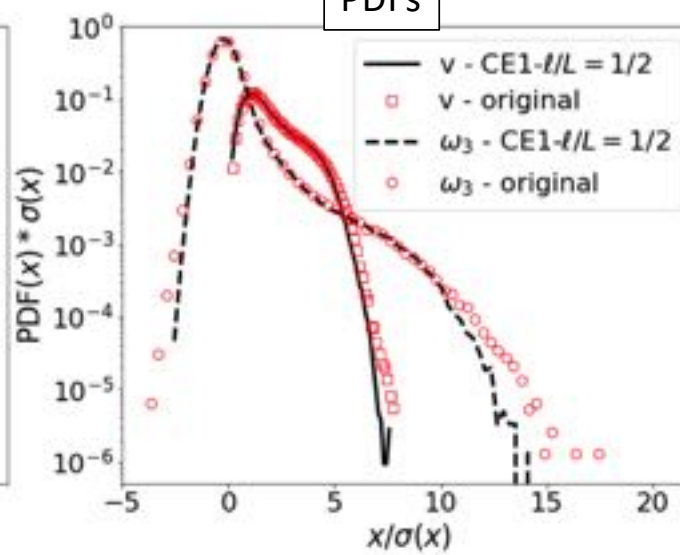
$\ell/L = 1/8$



SPECTRA

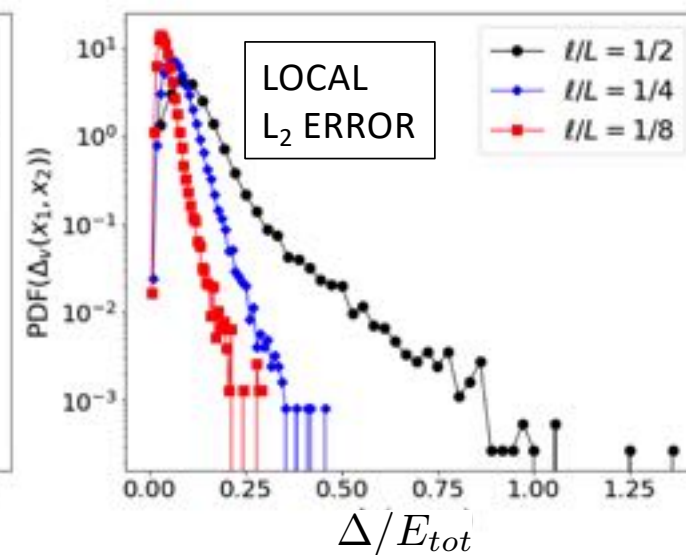


PDFs



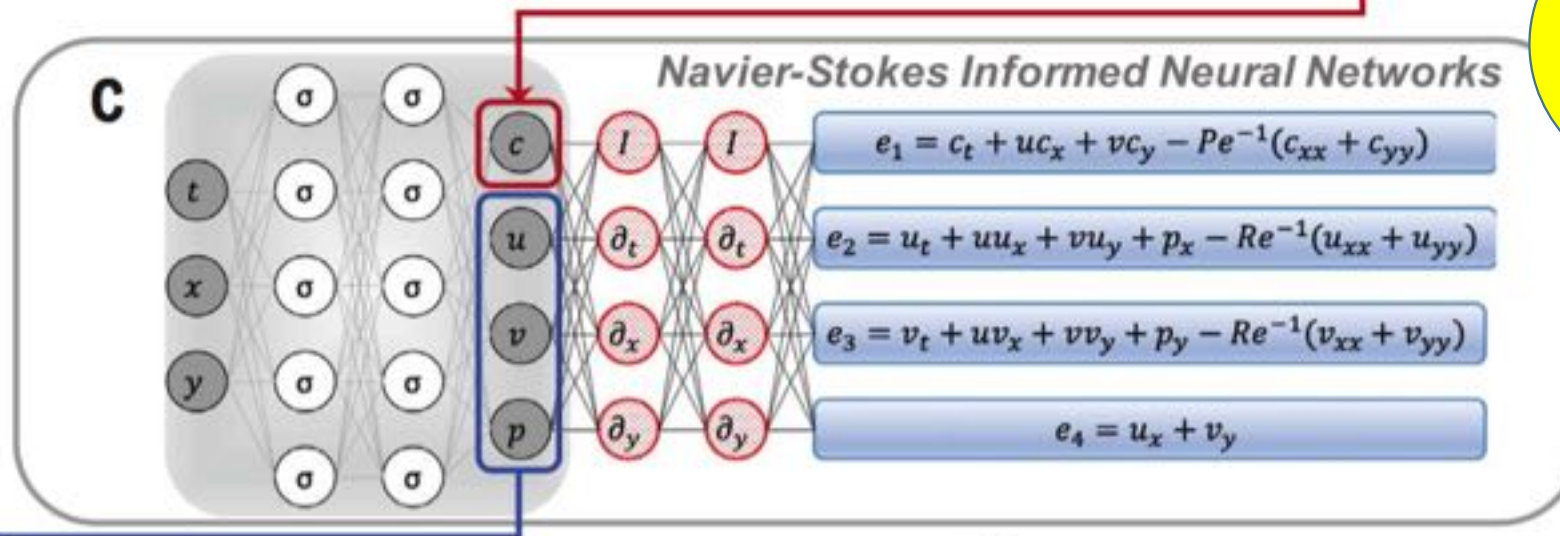
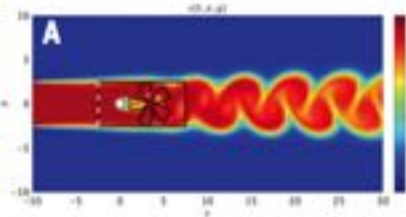
$$\Delta = (v_{true}(x, y) - v_{gen}(x, y))^2$$

LOCAL L2 ERROR





2D flow past a cylinder: low complexity



PDE SOLVERS
FOR
AUGMENTED
DATA-
ASSIMILATIONS



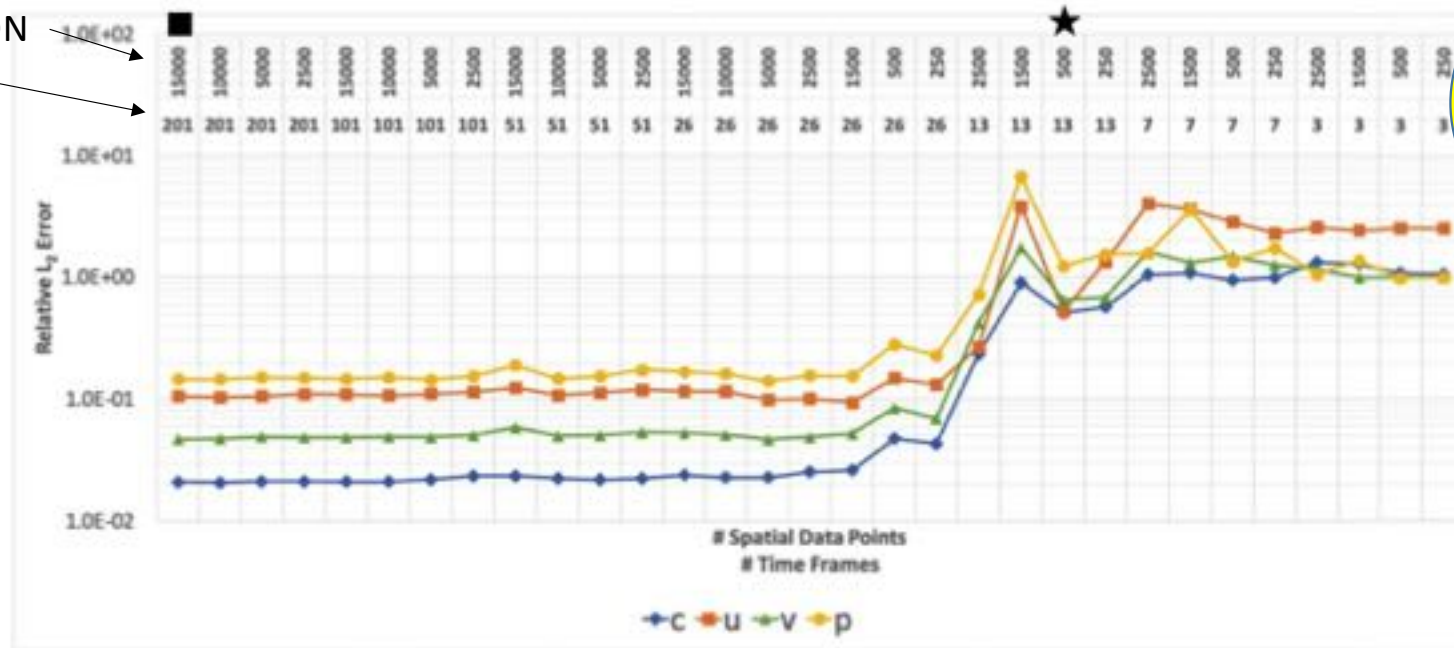
$$MSE = \frac{1}{N} \sum_{n=1}^N |c(t^n, x^n, y^n, z^n) - c^n|^2 + \sum_{i=1}^4 \frac{1}{M} \sum_{m=1}^M |e_i(t^m, x^m, y^m, z^m)|^2$$

ML-TRAINED ON A SPARSE SPATIO+TEMPORAL DATASET
FOR CONCENTRATION -> INFER VELOCITY + PRESSURE
-> BACK PROPAGATE FOR GRADIENTS (**AUTOMATIC DIFFERENTIATION**)->
NAVIER-STOKES

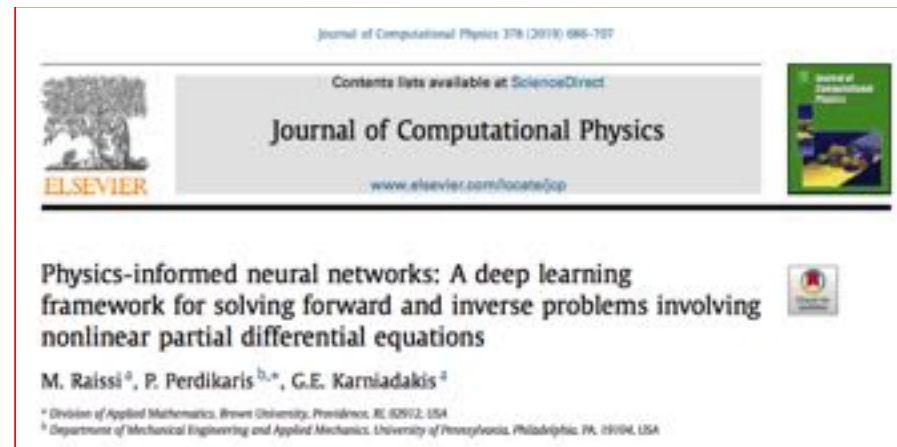
MEAN L2 ERROR

$$MSE = \frac{1}{N} \sum_{n=1}^N |c(t^n, x^n, y^n, z^n) - c^n|^2 + \sum_{i=1}^5 \frac{1}{M} \sum_{m=1}^M |e_i(t^m, x^m, y^m, z^m)|^2$$

PDE SOLVERS
FOR
AUGMENTED
DATA-
ASSIMILATIONS

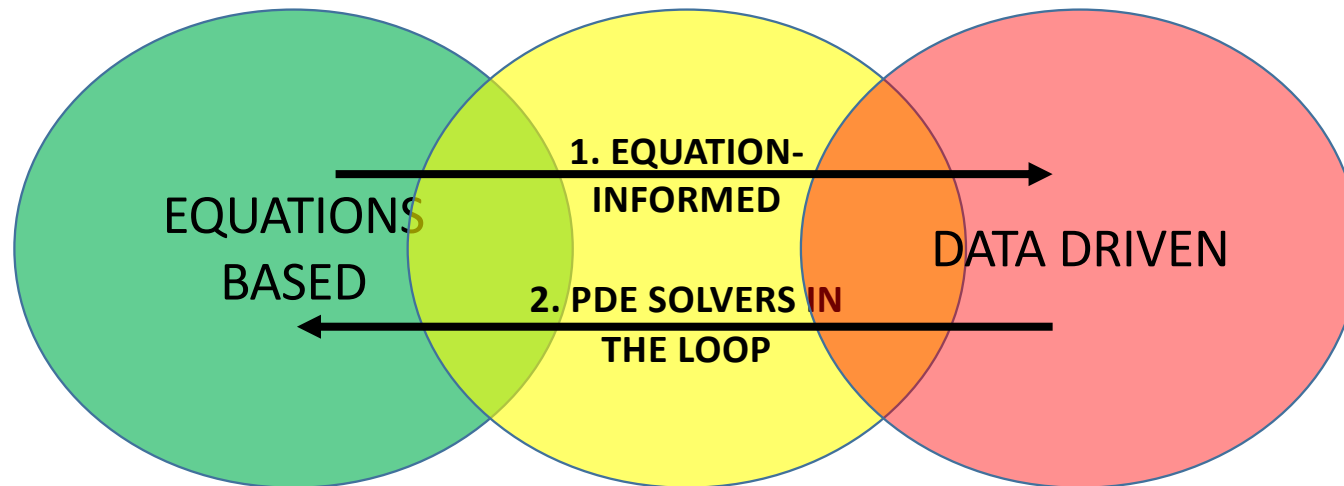


ML-TRAINED ON A SPARSE SPATIO+TEMPORAL DATASET
FOR CONCENTRATION -> INFER VELOCITY + PRESSURE
-> BACK PROPAGATE FOR GRADIENTS -> NAVIER-STOKES



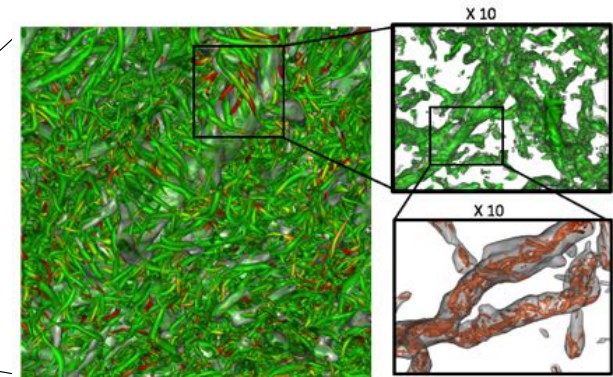
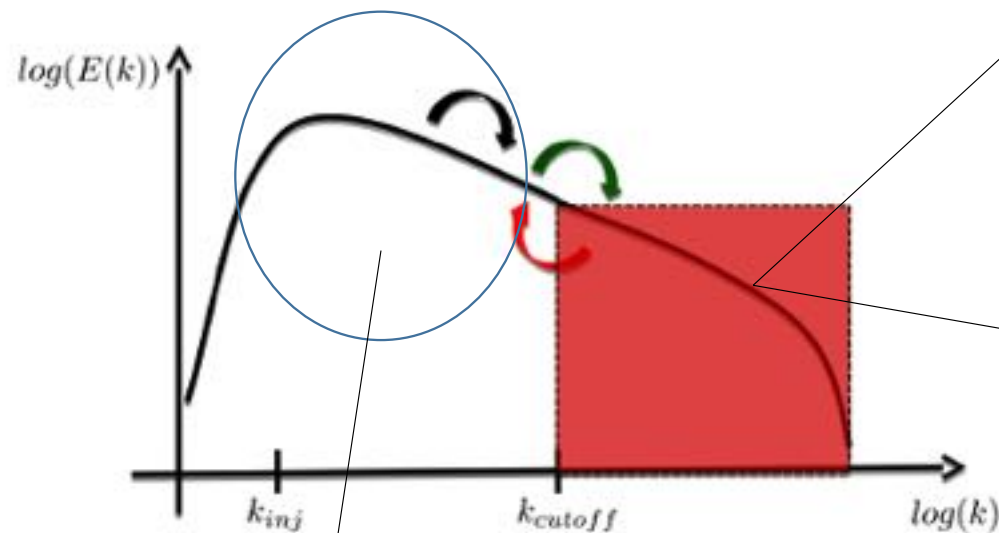
Here, we should underline an important distinction between this line of work and existing approaches in the literature that elaborate on the use of machine learning in computational physics. The term *physics-informed machine learning* has been also recently used by Wang et al. [17] in the context of turbulence modeling. Other examples of machine learning approaches for predictive modeling of physical systems include [18–29]. All these approaches employ machine learning algorithms like support vector machines, random forests, Gaussian processes, and feed-forward/convolutional/recurrent neural networks merely as *black-box* tools. As described above, the proposed work aims to go one step further by revisiting the construction of “custom” activation and loss functions that are tailored to the underlying differential operator. This allows us to open the black-box by understanding and appreciating the key role played by automatic differentiation within the deep learning field. Automatic differentiation in general, and the back-propagation algorithm in particular, is currently the dominant approach for training deep models by taking their derivatives with respect to the parameters (e.g., weights and biases) of the models. Here, we use the exact same automatic differentiation techniques, employed by the deep learning community, to physics-inform neural networks by taking their derivatives with respect to their input coordinates (i.e., space and time) where the physics is described by partial differential equations. We have empirically observed that this structured approach introduces a regularization mechanism that allows us to use relatively simple feed-forward neural network architectures and train them with small amounts of data. The effectiveness of this simple idea may be related to the remarks

1. DATA ASSIMILATION, FLOW RECONSTRUCTION, INPAINTING, SUPER-RESOLUTION
2. LARGE-EDDY-SIMULATIONS, MODEL REDUCTION, HOMOGENEIZATION



$$\partial_t v + \nabla(v \otimes v) = -\nabla p + \nu \Delta v$$

$$\bar{v}(x, t) \equiv \int_{\Omega} dy \, G(|x - y|) v(y, t) = \sum_{k \in \mathbb{Z}^3} G(k) \hat{v}(k, t) e^{ikx}$$

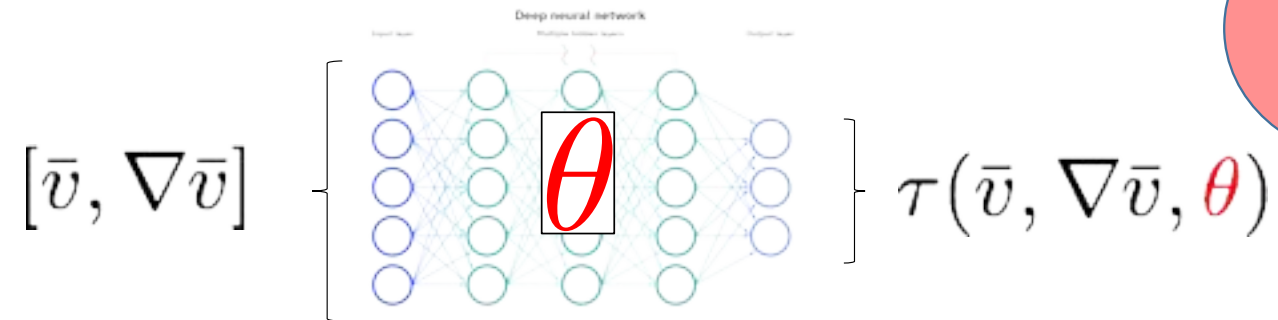


NEED TO MODEL NON-GAUSSIAN
& MULTI-SCALE PHYSICS !!!

$$\partial_t \bar{v} + \nabla \cdot (\bar{v} \otimes \bar{v}) = -\nabla \bar{p} + \nabla \cdot \tau^{\Delta}(v, v) + \nu \Delta \bar{v}$$

$$\tau_{ij}^{\Delta}(v, v) = \overline{v_i v_j} - \bar{v}_i \bar{v}_j \xrightarrow{\text{??????}} \tau(\bar{v}, \bar{v})$$

DATA DRIVEN
NO-
EQUATIONS



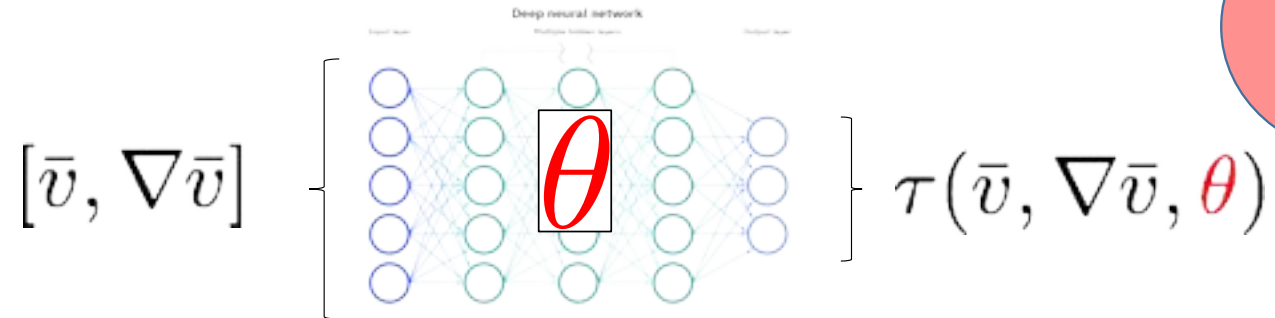
$$\partial_t \bar{v} + \nabla(\bar{v} \otimes \bar{v}) = -\nabla \bar{p} + \nabla \tau(\bar{v}, \nabla \bar{v}, \theta) + \nu \Delta \bar{v}$$

$$\tau_{i,j}(\bar{v}, \bar{v}) = \bar{\nu}_{eff}(\nabla \bar{v})(\nabla_i \bar{v}_j + \nabla_j \bar{v}_i) + \dots$$

EQUATIONS
BASED

APRIORI SUPERVISED TRAINING

DATA DRIVEN
NO-
EQUATIONS



$$\text{COST: } L(\theta) = \int dt \int dx ||\tau_{true}(v, v) - \tau(\bar{v}, \nabla \bar{v}, \theta)||_2$$

$$\text{GRADIENT DESCENT: } \theta_{k+1} = \theta_k - \alpha \partial_{\theta} L(\theta_k)$$

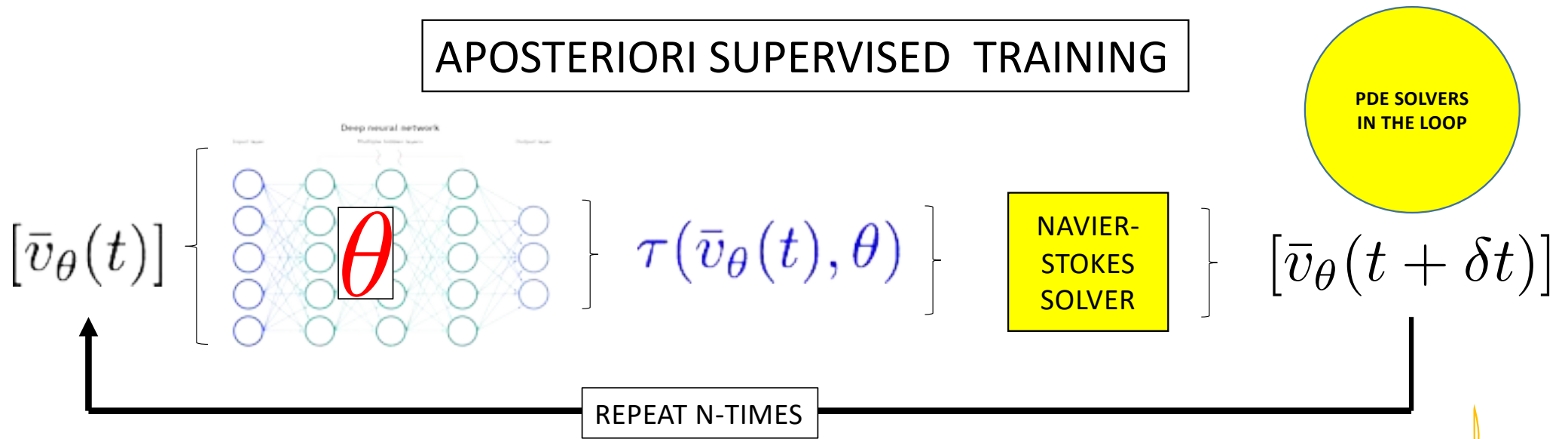
$$\text{CALCULATION GRADIENTS OF COST: } \partial_{\theta} L(\theta) = \partial_{\tau} L(\theta) \partial_{\theta} \tau$$

**EASY: USUAL BACK
PROPAGATION (ML NATIVE) ->
EXTREMELY FAST**

 **1-step in time optimization -> no dynamics, no control of the impact on the PDE evolution/stability**

See for a review: ML for Fluid Mechanics. S. L. Brunton, B.R. Noack and P. Koumoutsakos. Annu. Rev. Fluid Mech (2020) 52, 477. TC 390

APOSTERIORI SUPERVISED TRAINING

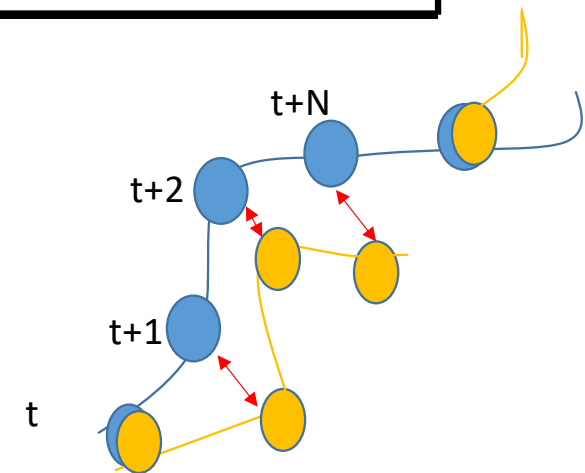


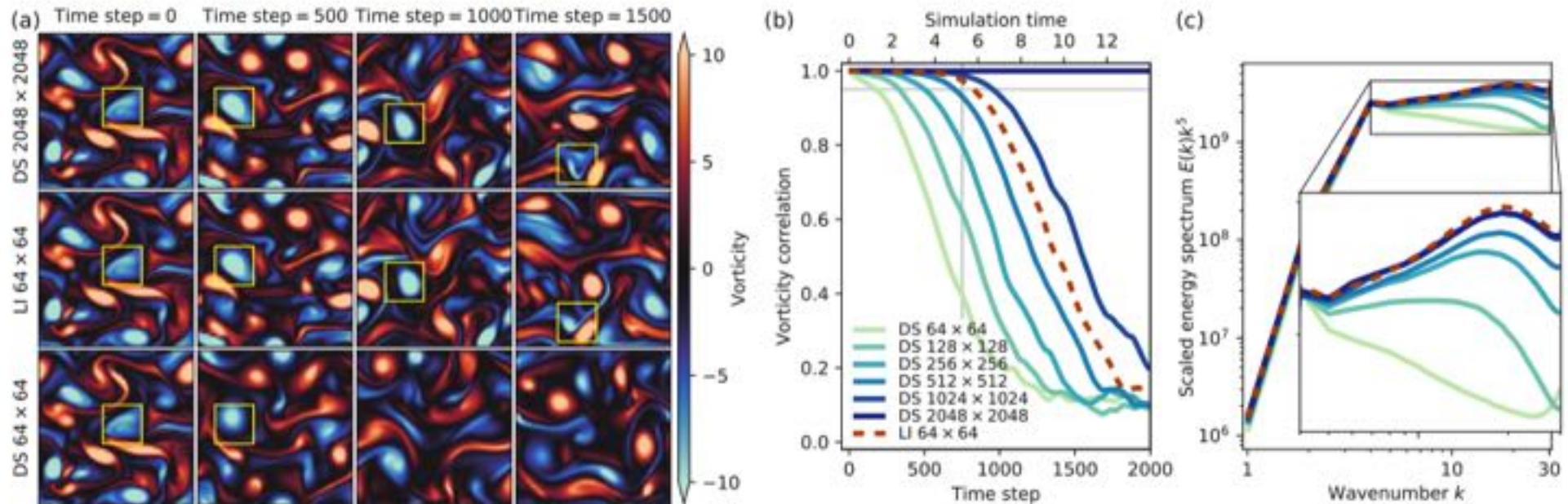
$$L(\theta) = D[v(\cdot), \bar{v}_\theta(\cdot)] = \int dx \int_0^T dt ||v(t) - \bar{v}_\theta(t)||_2$$

$$\theta_{k+1} = \theta_k - \alpha \partial_\theta L(\theta_k) \rightarrow \partial_\theta L = \partial_{\bar{v}} L \partial_\theta \bar{v}$$

PROBLEM !!!!!

$$\partial_t \bar{v}_\theta + \nabla(\bar{v}_\theta \otimes \bar{v}_\theta) = -\nabla \bar{p}_\theta + \nabla \tau(\bar{v}, \theta) + \nu \Delta \bar{v}_\theta$$





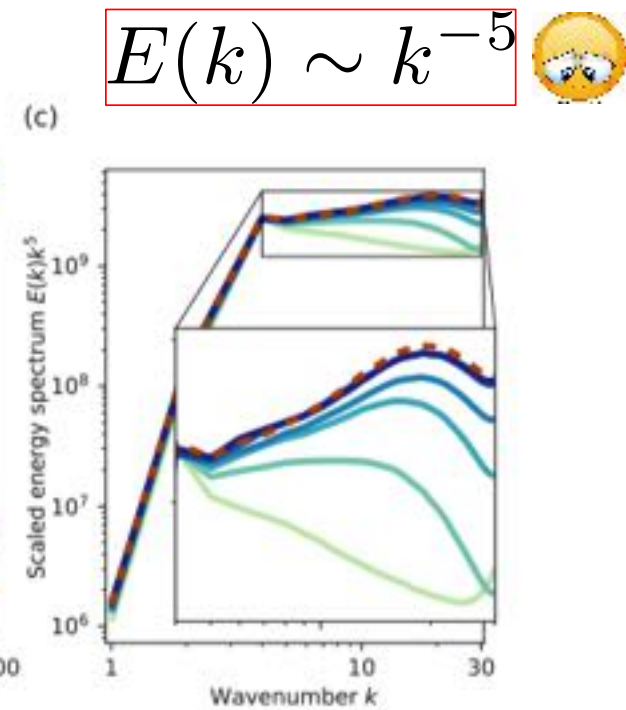
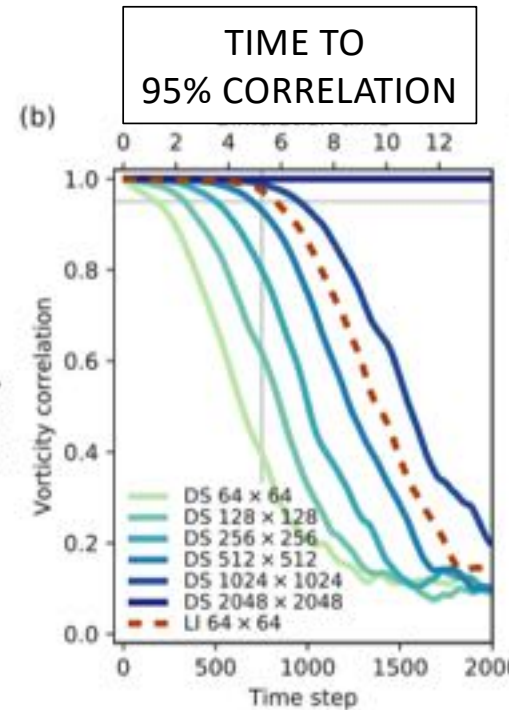
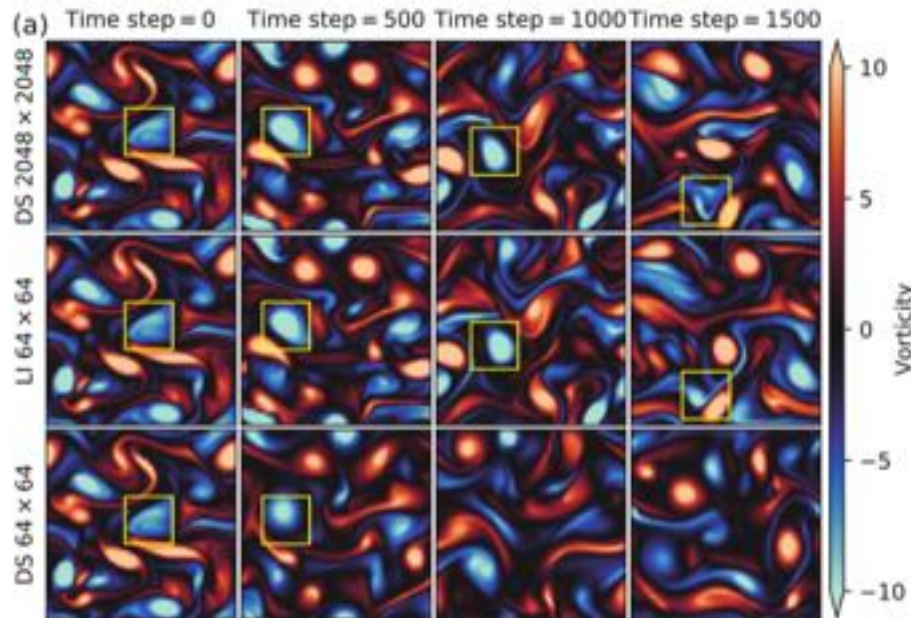
Machine learning accelerated computational fluid dynamics. arXiv:2102.01010v1 [physics.flu-dyn] Jan 2021.

D. Kochkov, J. A. Smith, A. Alieva, Q. Wang, M. P. Brenner, and S. Hoyer



Numerical method for solving the underlying PDEs as a **differentiable program**, with the neural networks and the numerical method written in a framework (JAX) supporting **reverse-mode automatic differentiation**. This allows for end-to-end gradient based optimization, **of the entire algorithm (NSE included)**

Automatic Differentiation in Machine Learning: a Survey. Baydin et al. arXiv:1502.05767v4 [cs.SC] 2018



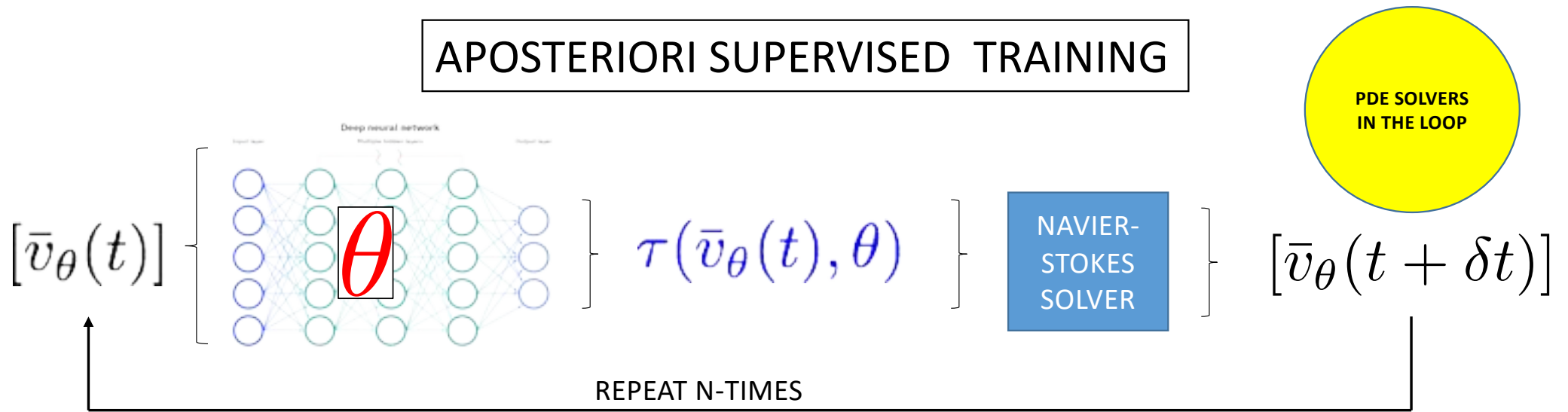
Machine learning accelerated computational fluid dynamics. arXiv:2102.01010v1 [physics.flu-dyn] Jan 2021.
D. Kochkov, J. A. Smith, A. Alieva, Q. Wang, M. P. Brenner, and S. Hoyer

GROUND TRUTH: 2D TURBULENCE 2048x2048
ERROR OF LEARNED INTERPOLATIONS AT 64x64 IS
EQUIVALENT TO STANDARD DISCRETIZATION AT 512x512
GAIN X8 FOR A GIVEN ERROR TOLERANCE

$$\nabla v|_{x=x_i} = \sum_{k=-j}^j \alpha_j v_{i+j} \quad \alpha = (-1, 0, 1)$$

$$\alpha_i \rightarrow \alpha_i(v, \nabla v, \theta)$$

APOSTERIORI SUPERVISED TRAINING



$$L(\theta) = D[v(\cdot), \bar{v}_\theta(\cdot)] = \int dx \int_0^T dt ||v(t) - \bar{v}_\theta(t)||_2$$

ADJOINT BACK PROPAGATION METHOD

$$S(\theta, \lambda) = \int dx \int_0^T dt ||v(t) - \bar{v}_\theta(t)||_2 + \lambda [\partial_t \bar{v}_\theta + \nabla(\bar{v}_\theta \otimes \bar{v}_\theta) + \nabla \bar{p}_\theta - \nabla \tau(\bar{v}, \theta) - \nu \Delta \bar{v}_\theta]$$

USE LAMBDA TO REMOVE ALL DERIVATIVES WRT TO θ AND REMAIN WITH LINEAR EQUATIONS FOR λ

$$\partial_\theta S(\theta, \lambda) = \partial_\theta L(\theta) = - \int dx \int_0^T dt \lambda \partial_\theta \nabla \tau(\bar{v}, \theta)$$

Embedded training of neural-network sub-grid-scale turbulence models J. F. MacArt, J. Sirignano, and J. B. Freund. May 2021
arXiv:2105.01030v1

3D JET DNS: 1024x1280x768 LES 64x80x48 COLLOCATION POINTS

Embedded training of neural-network sub-grid-scale turbulence models Jonathan F. MacArt, Justin Sirignano, and Jonathan B. Freund. May 2021 arXiv:2105.01030v1

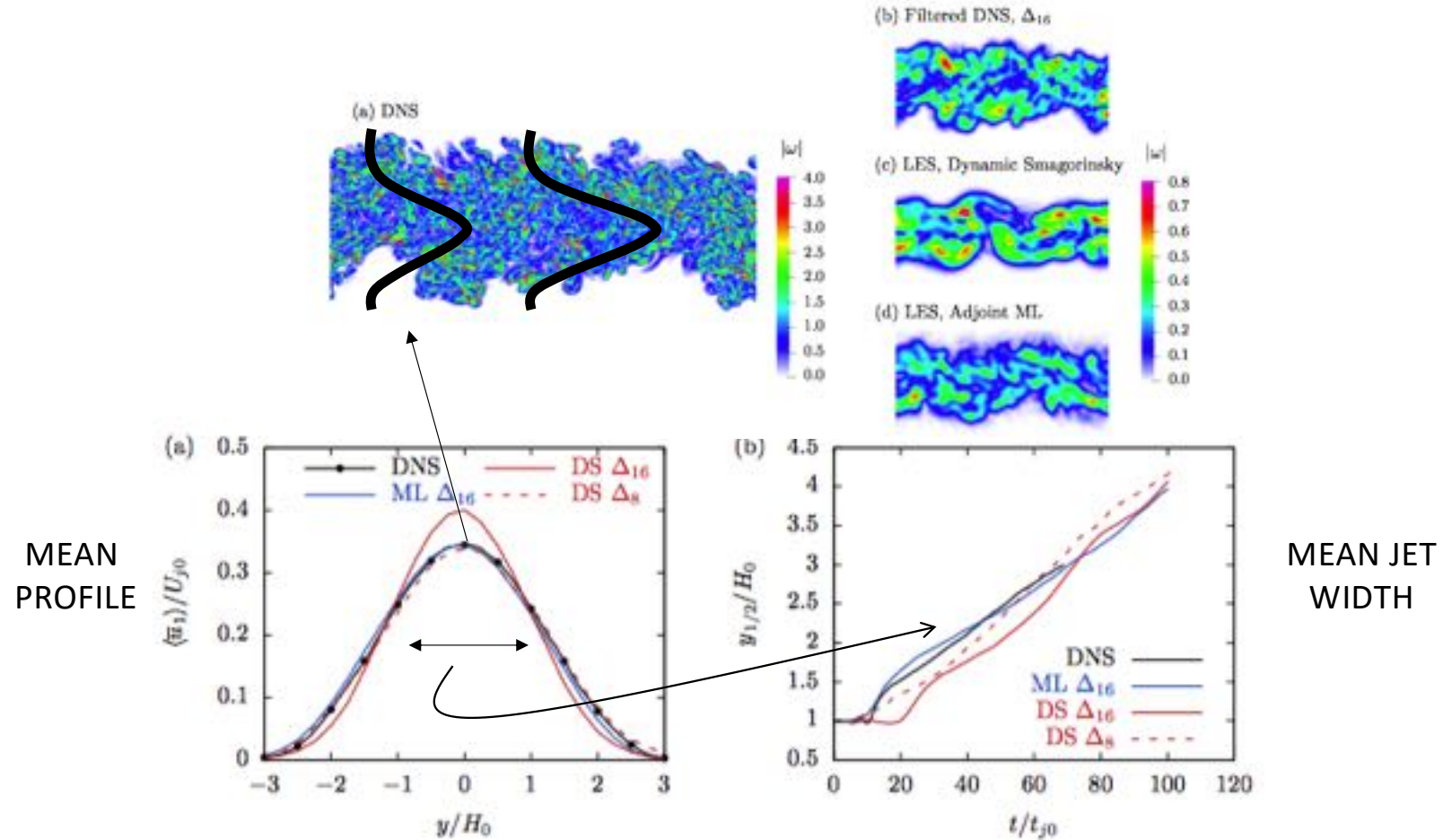


Figure 8: Single jet (case A) in-sample comparison for learning (ML) and dynamics Smagorinsky (DS) models: (a) mean streamwise velocity \bar{u}_1 at $t = 62.5t_{j0}$ and (b) half-width $y_{1/2}$ evolution for the indicated filter sizes. The direct numerical simulation data are included for comparison.

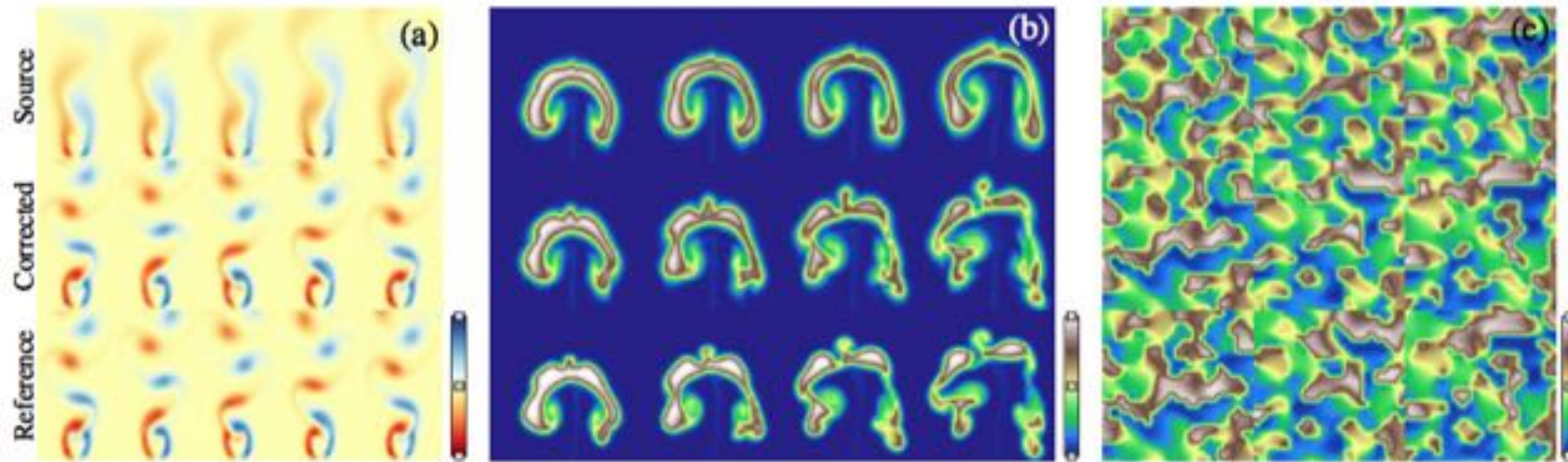


Figure 3: Our PDE scenarios cover a wide range of behavior including (a) vortex shedding, (b) complex buoyancy effects, and (c) advection-diffusion systems. Shown are different time steps (l.t.r.) in terms of vorticity for (a), transported density for (b), and angle of velocity direction for (c).

2) Solver-in-the-Loop: Learning from Differentiable Physics to Interact with Iterative PDE-Solvers.

K, Um, R. Brand, Y. R. Fei, P. Holl. N Thuerey Xiv:2007.00016v2 Jan 2021

$$\partial_t \bar{v} + \nabla(\bar{v} \otimes \bar{v}) = -\nabla \bar{p} + \nabla \tau(\bar{v}, \nabla \bar{v}, \theta) + \nu \Delta \bar{v}$$

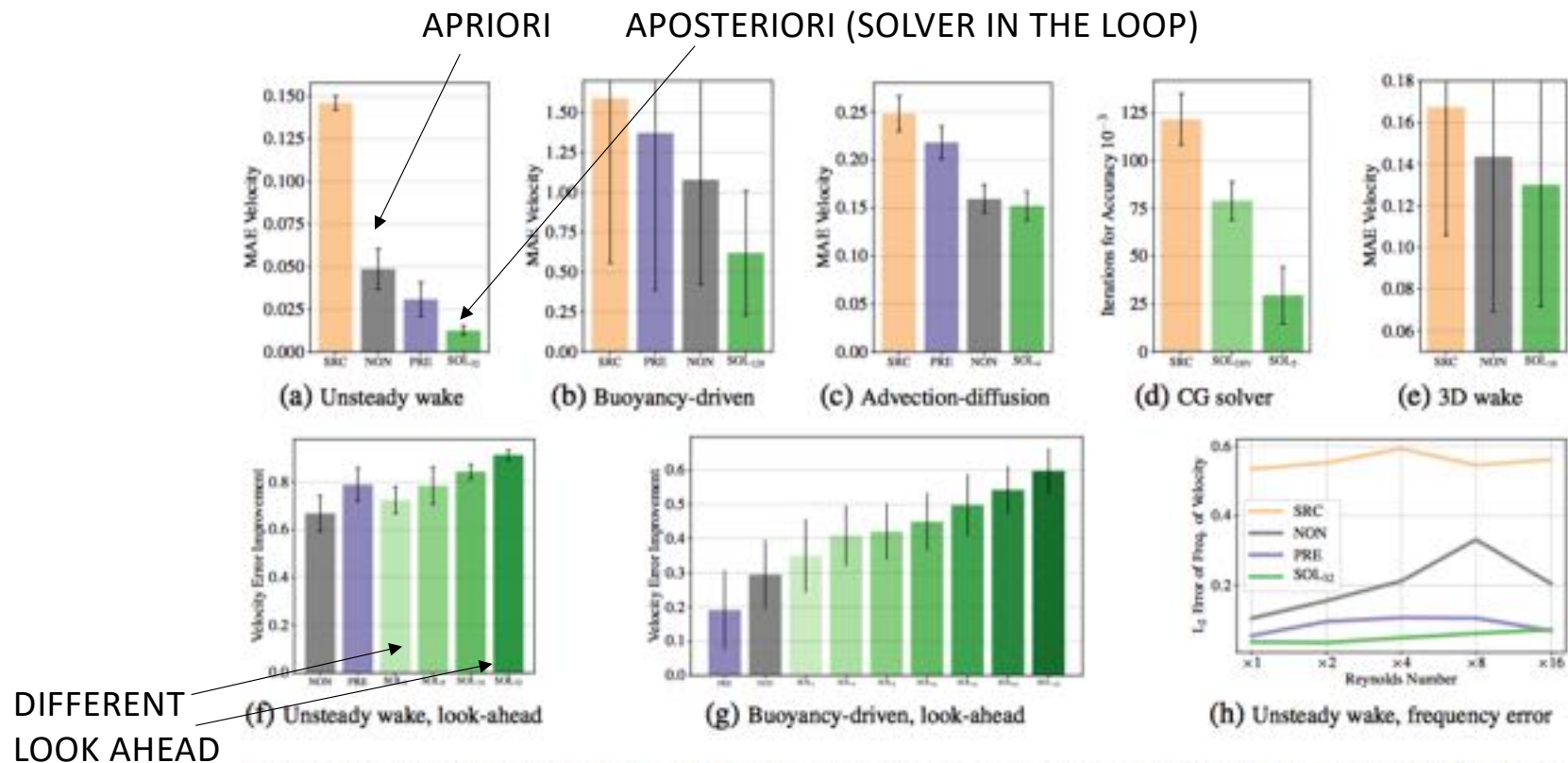


Figure 4: (a)-(e) Numerical approximation error w.r.t. reference solution for unaltered simulations (SRC) and with learned corrections. The models trained with differentiable physics and look-ahead achieve significant gains over the other models. (f,g) Relative improvement over varying look-ahead horizons. (h) A frequency-based evaluation for the unsteady wake flow scenario.

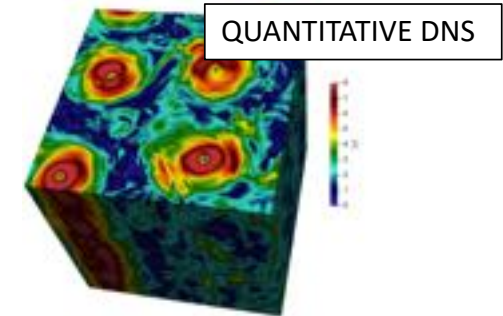
2) Solver-in-the-Loop: Learning from Differentiable Physics to Interact with Iterative PDE-Solvers.
K. Um, R. Brand, Y. R. Fei, P. Holl. N Thuerey Xiv:2007.00016v2 Jan 2021

ARE WE CLOSE TO AI SUPREMACY IN FLUID DYNAMICS?

- WE HAVE NEW TOOLS IN THE BOX
- **NEW APPLICATIONS FOR PDEs SOLVERS AUGMENTED BY MACHINE LEARNING**
- **NEW APPLICATIONS FOR MACHINE LEARNING AUGMENTED BY PDEs**

BUT

- RATE OF PUBLICATIONS IN THE DOMAIN >> RATE OF READING/PEER REVIEWING -> DANGER OF INFLATIONARY ERA
- HUNDREDS OF PAPERS IN THE ARXIVES CITED BY HUNDREDS OF OTHER PAPERS WITHOUT CHECK ON THE RESULTS, **NOT EVEN WRONG!**
- **NEED FOR QUANTITATIVE AI: SCALING, VALIDATION, BENCHMARKS, GENERALISATION, GRAND-CHALLENGES TO ESTABLISH BEST-PRACTISE**
- **NEED FOR PHYSICAL DIMENSIONALISATION: NETWORK STRUCTURE VS PHYSICAL PARAMETERS** (deepness, structure, coding, # training data vs Reynolds, Rayleigh, Time-to-solution, Mach, Mass fraction etc...)
- **NEED FOR INTERDISCIPLINARY COLLABORATIONS: APPLIED SCIENTISTS** (FOR THE GOOD QUESTIONS) + **AI SPECIALISTS** (TO OPEN THE BOX) + **NUMERICAL SCIENTISTS** (FOR THE GOOD DATA)



LORENZ96

WARNING

DESTINATION EARTH

2021-2026 AdG GRANT Smart-TURB
A Physics-Informed Machine-Learning Platform Smart Lagrangian Harness and Control of TURBulence
biferale@roma2.infn.it biferale@gmail.com

WE ARE HIRING
POST-DOCS



WE ARE HIRING
POST-DOCS