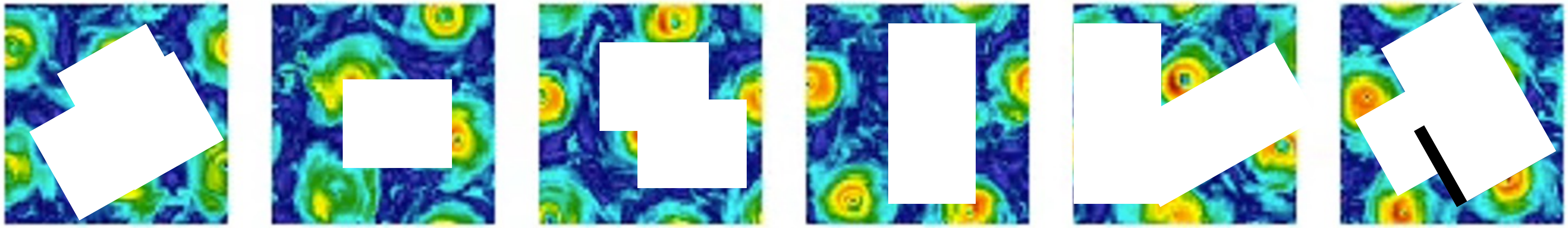
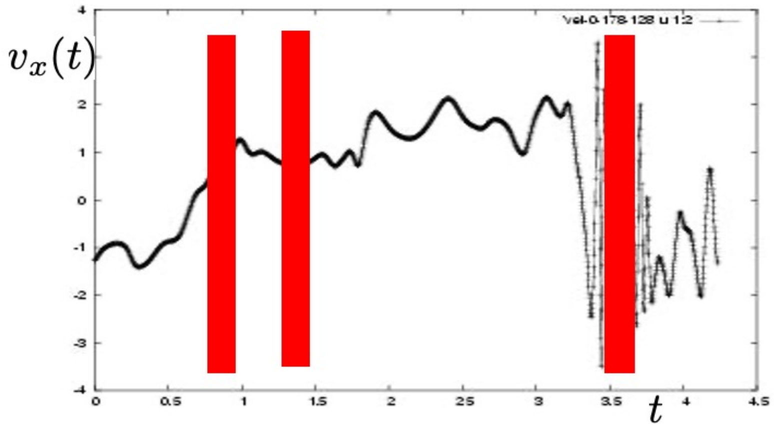
Complex Fluids and Complex Flows Group
Dept. Physics & INFN - University of Rome 'Tor Vergata'
biferale@roma2.infn.it
https://biferale.web.roma2.infn.it/



Rotation rate? | Viscosity? | Reynolds number? | Boundary comditions? | Forcing properties? | Aspect ratio?



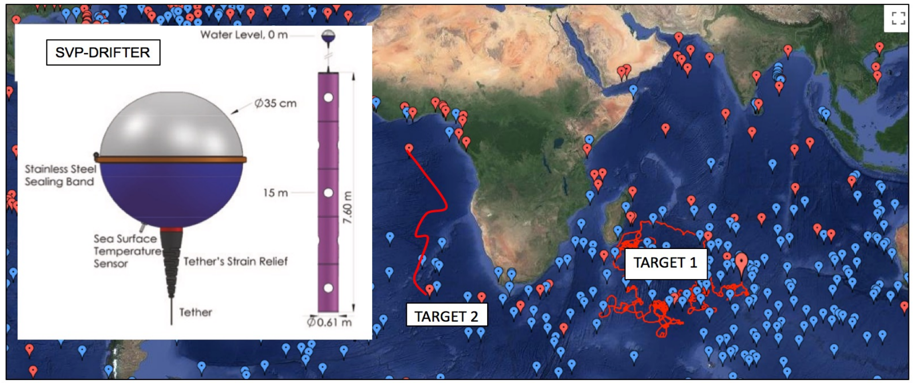




FIG. 1: INSTANTANEOS WORLDWIDE DISTRIBUTION OF DRIFTERS FROM THE GLOBAL DRIFTER MAP PROGRAM [www0]. WE HIGHLIGTHED IN RED TWO POTENTIAL TARGETS: (1) KEEP THE PROBES INSIDE A GIVEN REGION OR (2) MINIMISING THE NAVIGATION TIME AMONG TWO END-POINTS (ZERMELO PROBLEM). INSET: A SKETCH OF THE DRIFTER WITH THE LONG DROUGE AT 15M DEPTH.

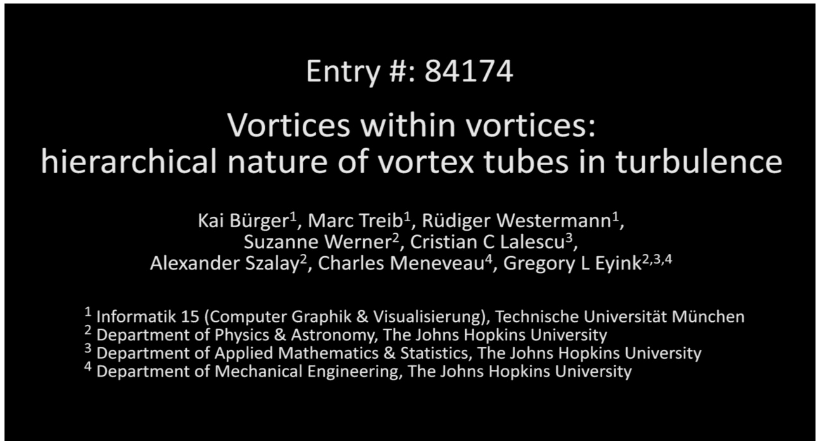
# Machine-learning and equations-informed tools for generation and augmentation of turbulent data.

**Artificial Intelligence and the Uncertainty challenge in Fundamental Physics**
**Paris 2023**

**CREDITS: T. LI, M. BUZZICOTTI, F. BONACCORSO, S. CHEN. M. WAN**

1. Short introduction to Eulerian Turbulence

2. Data-driven and Equation-Informed tools for Eulerian Turbulence:
   a. Linear Principal Orthogonal Decomposition
   b. Deterministic Generative Adversarial Networks
   c. Diffusion Models
   e. Nudging
   f. Physics Informed NN




# TURBULENCE OR TURBULENCES?

MASS X ACCELERATION = INTERNAL FORCES + EXTERNAL FORCES

EULERIAN

$$\rho[\partial_t v + v \cdot \partial v] = -\partial p + \nu \Delta v + \boldsymbol{g}\theta + F(\boldsymbol{B}, \boldsymbol{B}) + 2\Omega \times v + \hat{F}_{mech}$$

$$\partial_t \theta + \boldsymbol{v} \cdot \boldsymbol{\partial} \theta = \chi \partial^2 \theta$$ ← TEMPERATURE

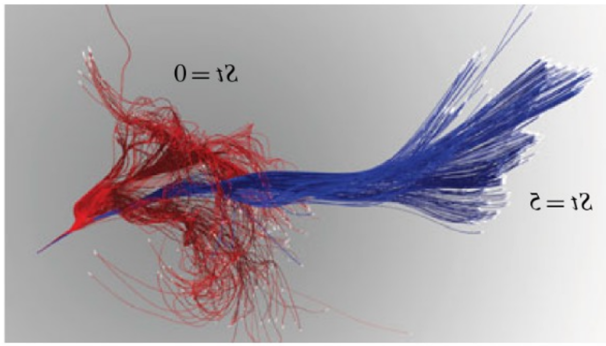
$$\partial_t \boldsymbol{B} + \boldsymbol{v} \cdot \boldsymbol{\partial} \boldsymbol{B} = \boldsymbol{B} \cdot \boldsymbol{\partial} \boldsymbol{v} + \chi \partial^2 \boldsymbol{B}$$ ← MAGNETIC FIELD

$$\partial \cdot v = 0$$

+ BOUNDARY CONDITIONS: (2D, 3D, THIN/THICK LAYERS ETC...)

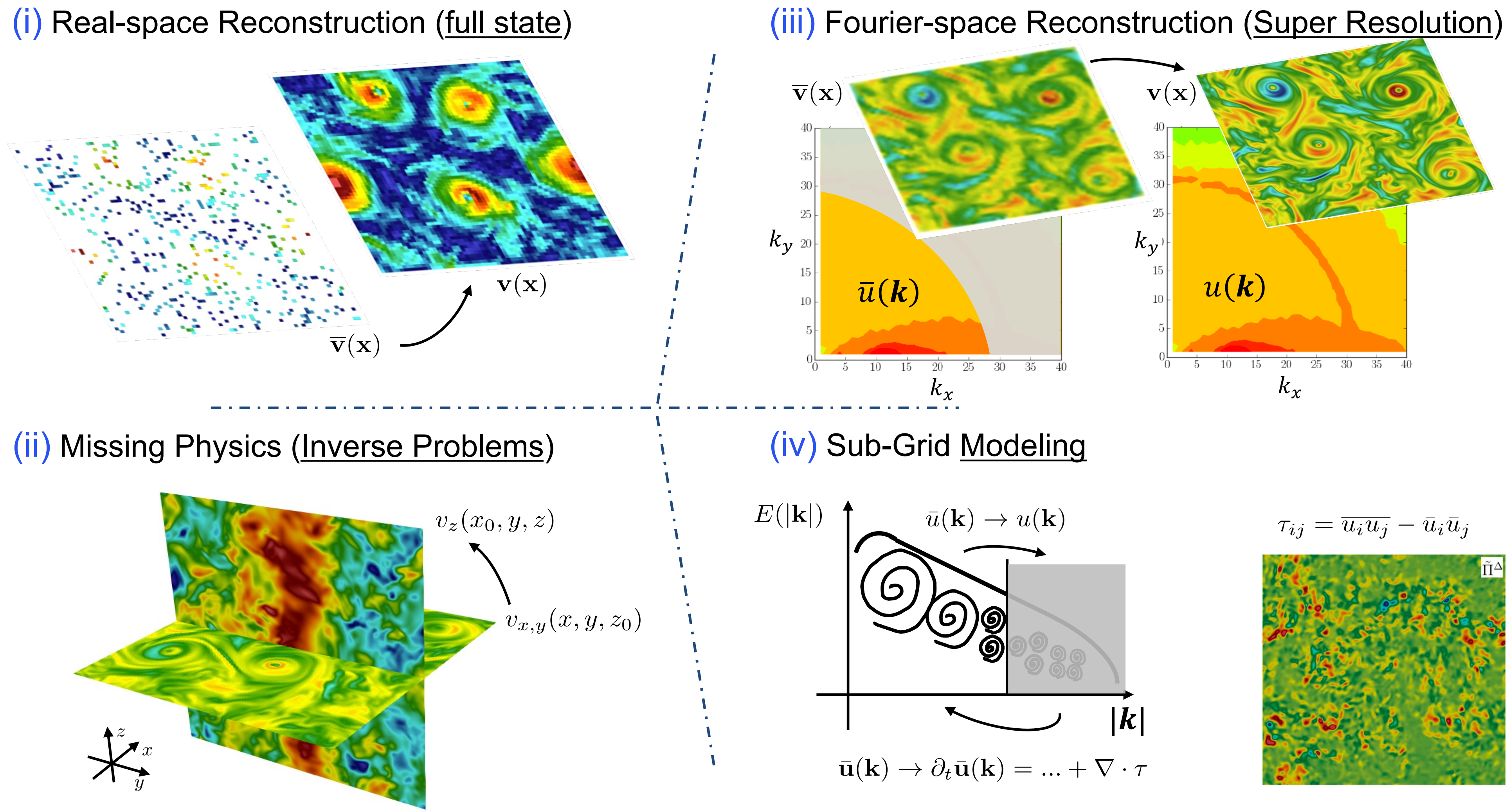
LAGRANGIAN

$$\dot{X}(t) = v(X(t), t)$$



A. Alexakis, L. B. , Cascades and transitions in turbulent flows. Phys. Rep. 767, 1–101 (2018)

# RECONSTRUCTION OF MISSING INFORMATION
# FEATURES RANKING: QUALITY AND QUANTITY OF DATA

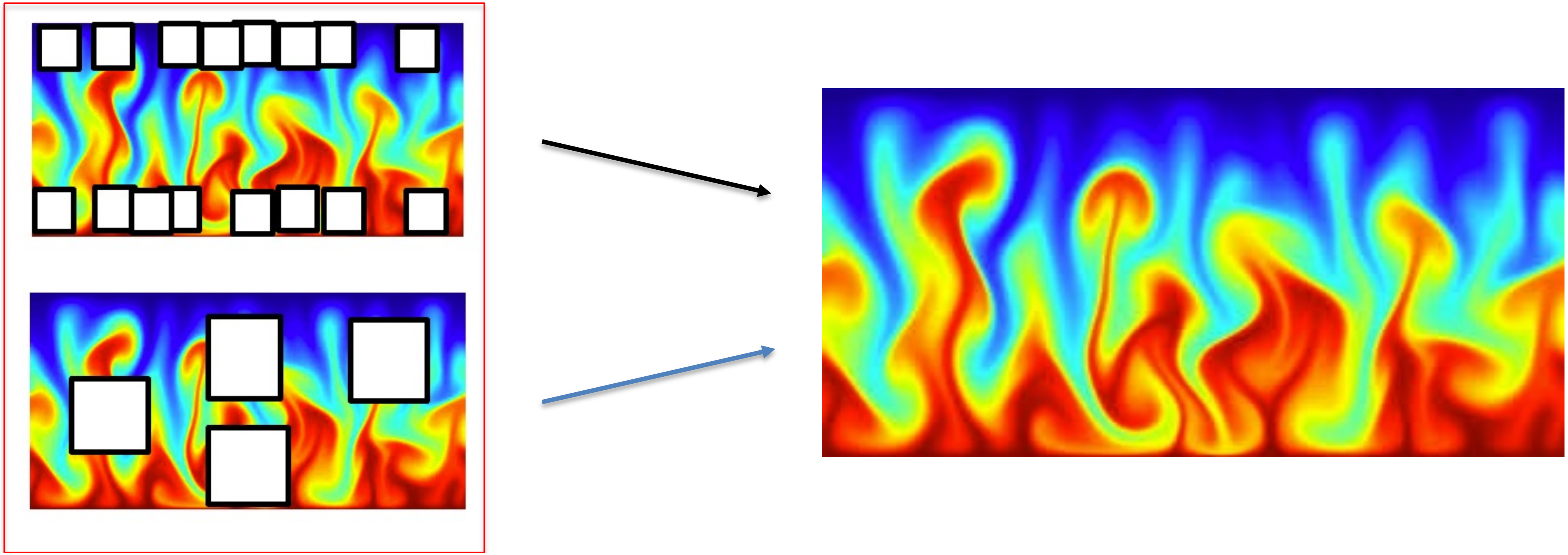



M. Buzzicotti. "Data reconstruction for complex flows using AI: recent progress, obstacles, and perspectives." Europhysics Letters, EPL 142 23001 (2023).

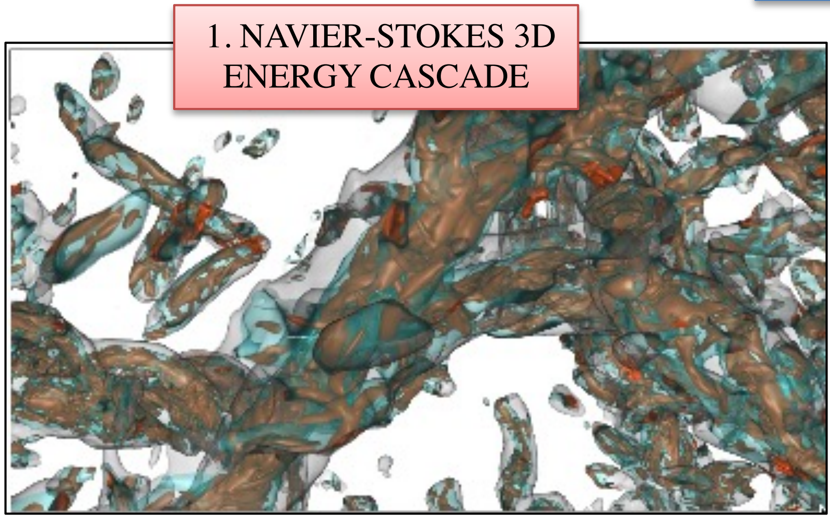
# WHY ?

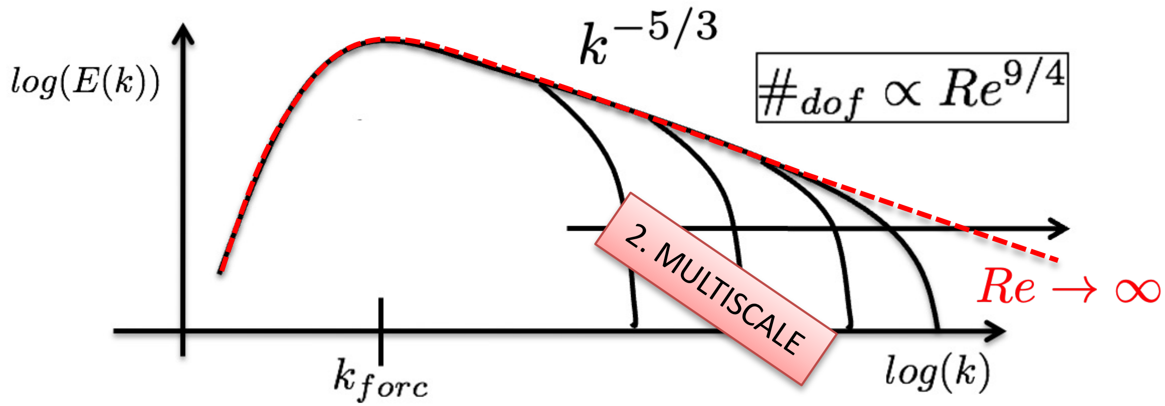
**1. GENERATION OF MISSING INFORMATION (INPAINTING/SUPER-RESOLUTION)**
**2. FEATURES RANKING: WHAT IS THE BEST METRIC TO RECONSTRUCT TURBULENT DATA?**

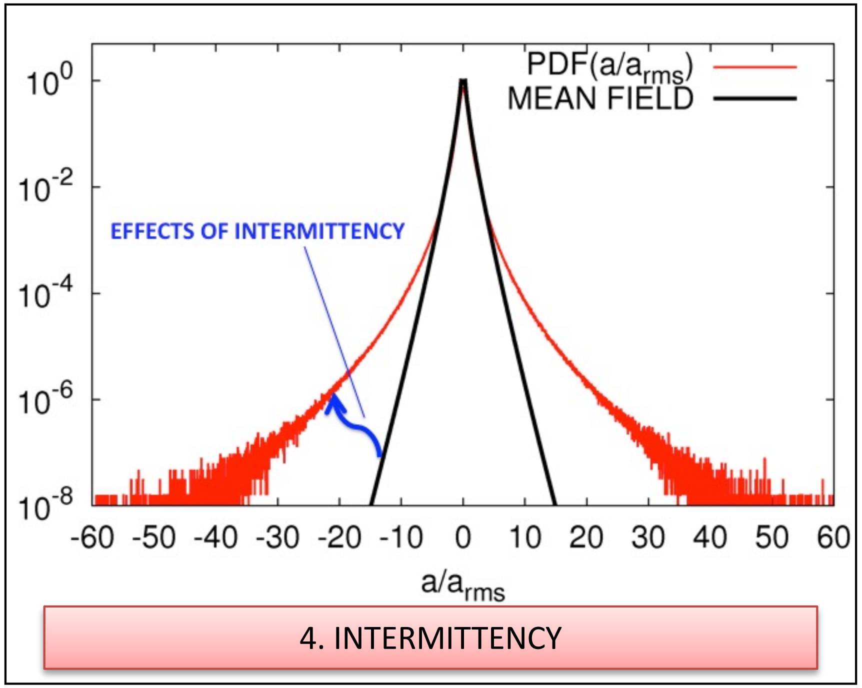
- IT IS MORE DIFFICULT TO RECONSTRUCT SPATIAL OR TEMPORAL DATA?
- HOW MANY DATA/VARIABLES YOU NEED TO SUPPLY FOR PERFECT RECONSTRUCTION (SYNCHRONIZATION-TO-DATA)?
- CAN YOU INFER VELOCITY FIELDS FROM TEMPERATURE SNAPSHOTS AND/OR VICEVERSA?
- IS IT BETTER TO PROVIDE INFORMATION FROM BOUNDARIES OR BULK?
- FROM LARGE OR SMALL SCALES?
- **DO WE NEED (IS IT USEFUL) TO KNOW THE EQUATIONS?**
- **HOW TO COMPARE EQUATIONS-BASED AND EQUATIONS-FREE MODELS?**

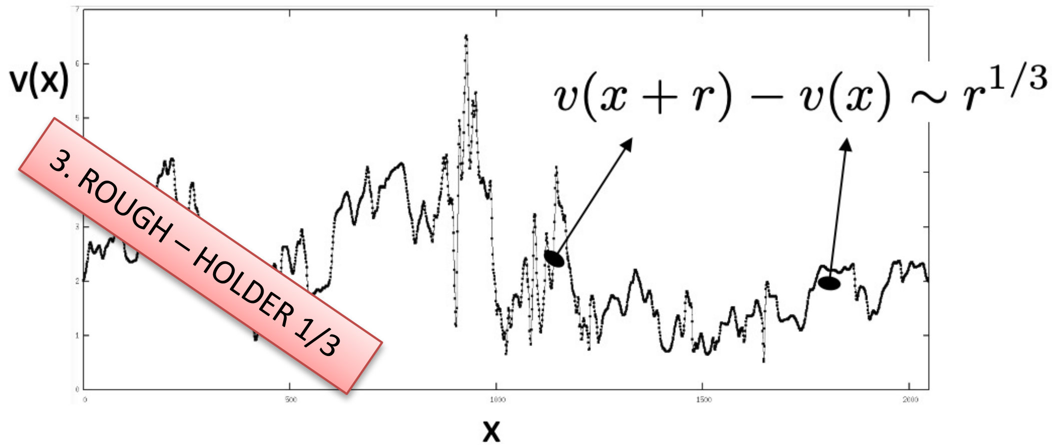


# WHY IS IT TOUGH?

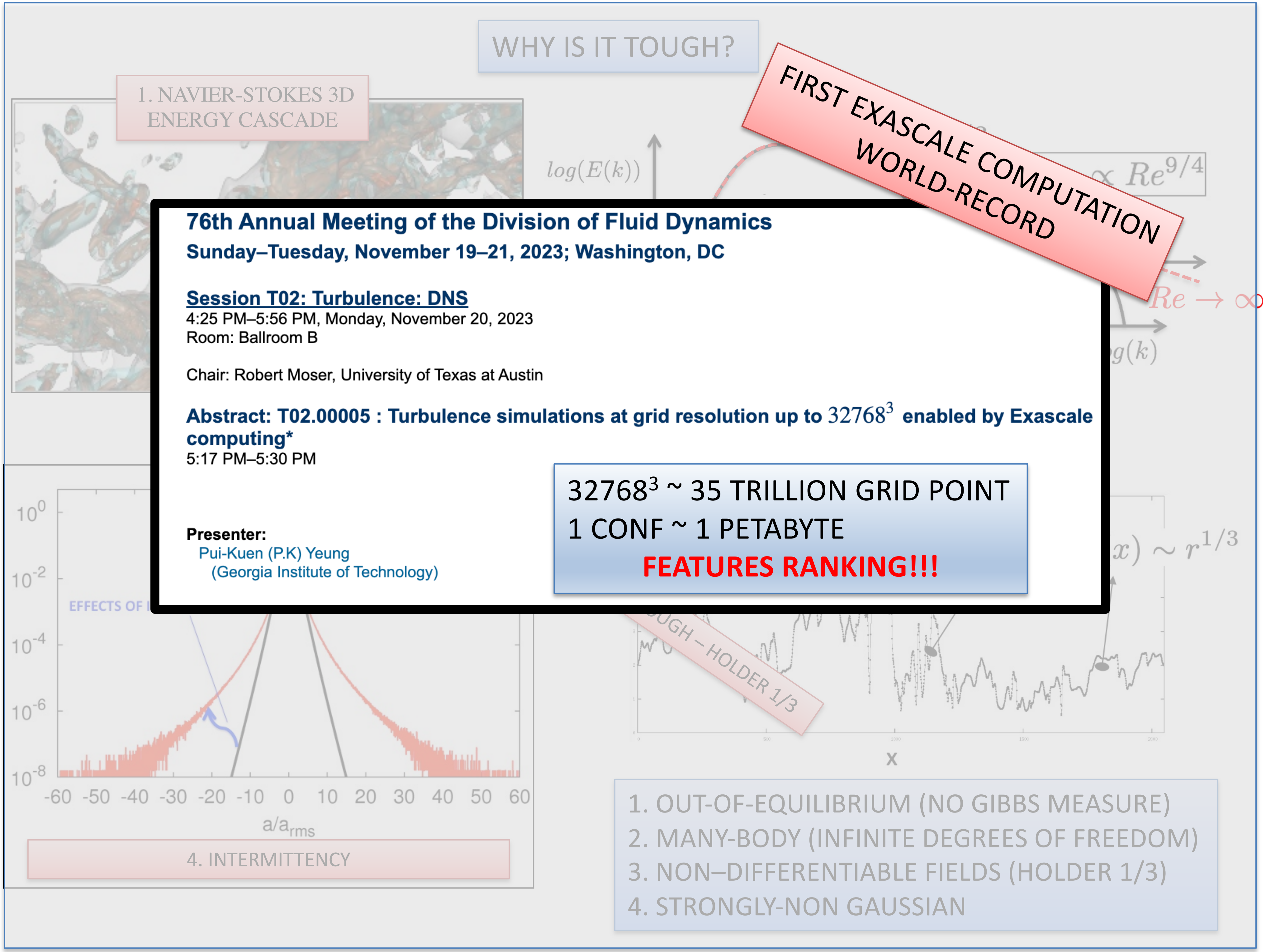


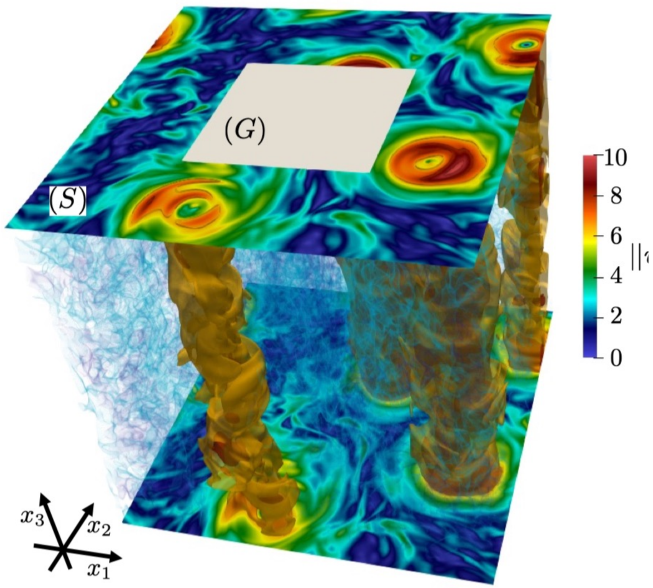

1. OUT-OF-EQUILIBRIUM (NO GIBBS MEASURE)
2. MANY-BODY (INFINITE DEGREES OF FREEDOM)
3. NON–DIFFERENTIABLE FIELDS (HOLDER 1/3)
4. STRONGLY-NON GAUSSIAN



WHY IS IT TOUGH?
1. NAVIER-STOKES 3D ENERGY CASCADE
$log(E(k))$
$\propto Re^{9/4}$
$Re \to \infty$
FIRST EXASCALE COMPUTATION
WORLD-RECORD
76th Annual Meeting of the Division of Fluid Dynamics
Sunday–Tuesday, November 19–21, 2023; Washington, DC
Session T02: Turbulence: DNS
4:25 PM–5:56 PM, Monday, November 20, 2023
Room: Ballroom B
Chair: Robert Moser, University of Texas at Austin
Abstract: T02.00005 : Turbulence simulations at grid resolution up to $32768^3$ enabled by Exascale computing*
5:17 PM–5:30 PM
Presenter:
Pui-Kuen (P.K) Yeung
(Georgia Institute of Technology)
$32768^3$ ~ 35 TRILLION GRID POINT
1 CONF ~ 1 PETABYTE
FEATURES RANKING!!!
$10^0$
$10^{-2}$
$10^{-4}$
$10^{-6}$
$10^{-8}$
-60 -50 -40 -30 -20 -10 0 10 20 30 40 50 60
$a/a_{rms}$
4. INTERMITTENCY
$\sim r^{1/3}$
HOLDER 1/3
X
1. OUT-OF-EQUILIBRIUM (NO GIBBS MEASURE)
2. MANY-BODY (INFINITE DEGREES OF FREEDOM)
3. NON–DIFFERENTIABLE FIELDS (HOLDER 1/3)
4. STRONGLY-NON GAUSSIAN

grid resolution: $1024^3$

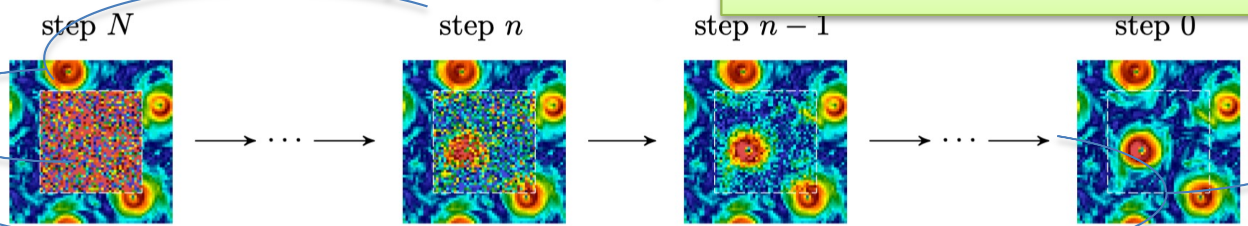
Coherent Structures and Extreme Events in Rotating Multiphase Turbulent Flows.
LB, F. Bonaccorso, I. M. Mazzitelli, et al. Phys. Rev. X **6**, 041036 – 2016




3D TURBULENCE UNDER ROTATION

MOCK SATELLITE MEASUREMENTS

DATA-DRIVEN
GENERATIVE ADVERSARIAL NETWORK & DIFFUSION MODELS




$$u_G^{(p)}(x) = \sum_{n=1}^{N_S} b_S^{(n)} \phi_E^{(n)}(x)$$

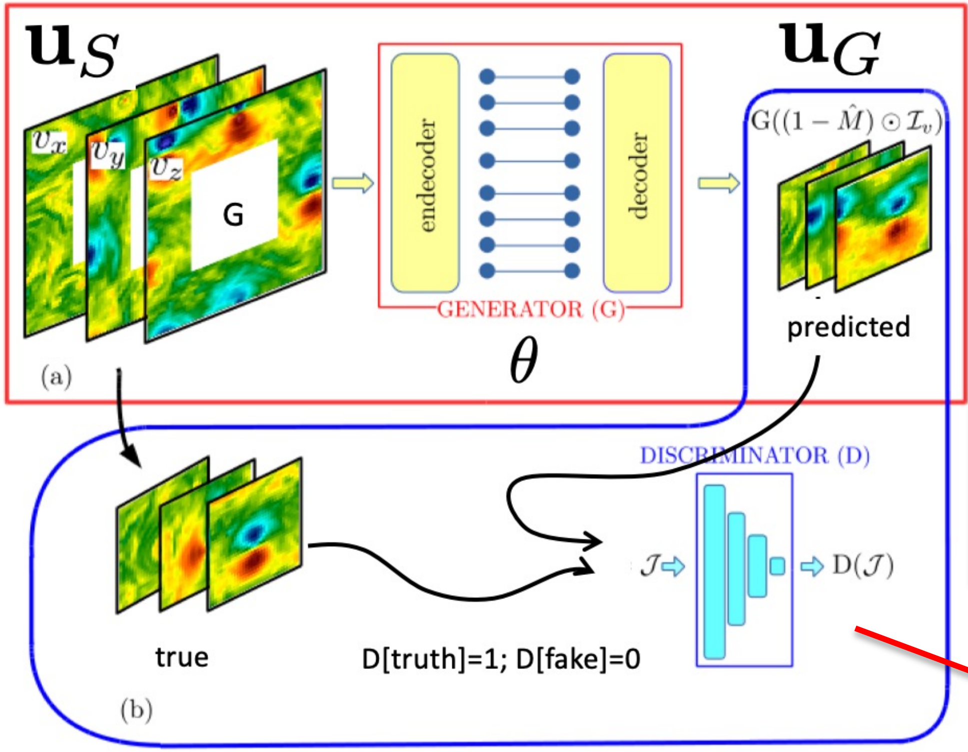
DATA-DRIVEN:
PRINCIPAL COMPONENT ANALYSIS

EQUATIONS-BASED: NUDGING

$$\partial_t \mathbf{u} + (\mathbf{u}\cdot\nabla)\mathbf{u} = -\nabla p + \nu\Delta\mathbf{u} + \mathbf{f}$$

**Context Encoder: features learning by inpainting.** D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrel, A. Efros. Proceed. of the IEEE Conf. on Computer Vision and Pattern Recognition, 2536, (2016)
**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

# NONLINEAR GENERATIVE ADVERSARIAL NETWORK: CONTEXT ENCODER (ACTOR-CRITIC)




MINIMIZE:

$$\mathcal{L}_{GEN} = (1-\lambda_{adv})\mathcal{L}_{\text{MSE}} + \lambda_{adv}\mathcal{L}_{adv},$$

$$\mathcal{L}_{\text{MSE}} = \langle \frac{1}{A(I)} \int_I \|\boldsymbol{u}_G^{(p)}(\boldsymbol{x}) - \boldsymbol{u}_G^{(t)}(\boldsymbol{x})\|^2 \, d\boldsymbol{x} \rangle$$

$$\mathcal{L}_{adv} = \langle \log(1 - D(\boldsymbol{u}_G^{(p)})) \rangle$$

$$= \int p(\boldsymbol{u}_S) \log[1 - D(GEN(\boldsymbol{u}_S))] \, d\boldsymbol{u}_S$$

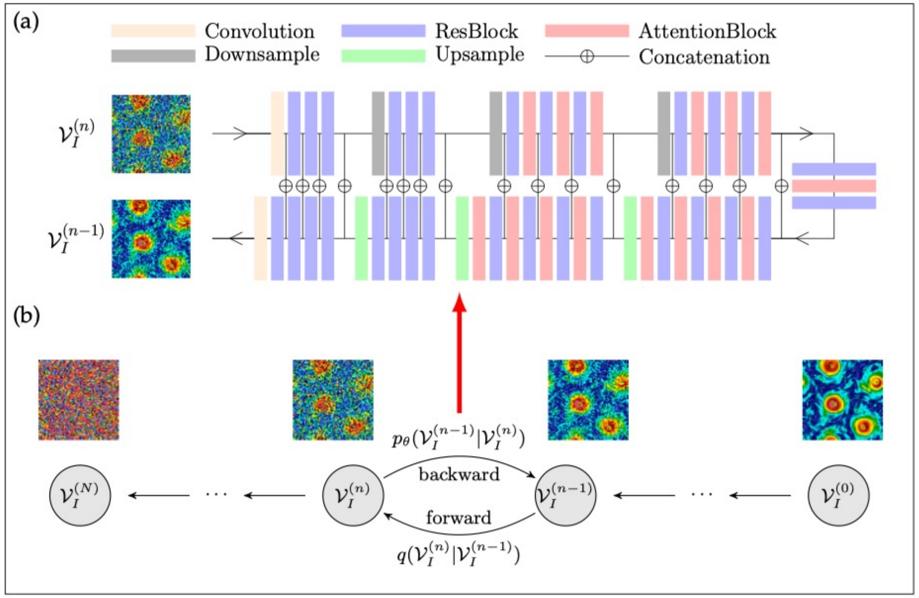
MAXIMIZE:

$$\mathcal{L}_{DIS} = \langle \log(D(\boldsymbol{u}_G^{(t)})) \rangle + \langle \log(1 - D(\boldsymbol{u}_G^{(p)})) \rangle$$

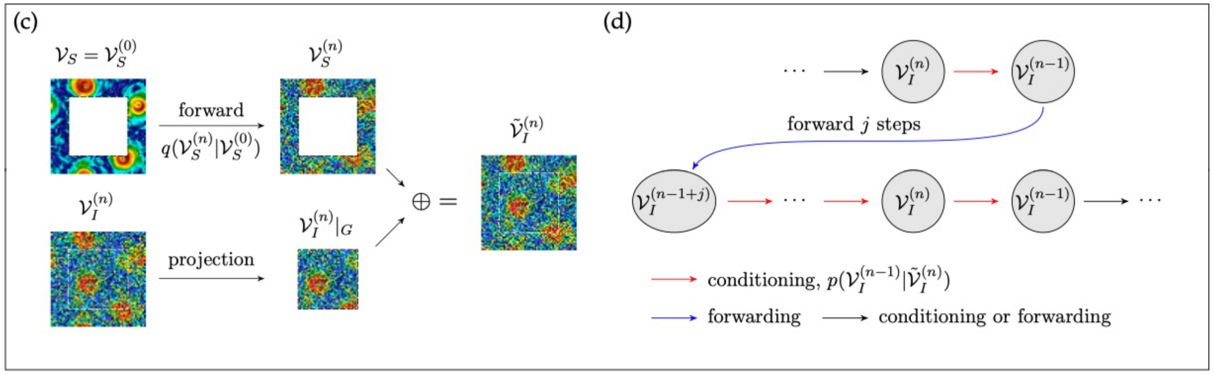
$$= \int [p_t(\mathbf{u}_G) log(D(\mathbf{u}_G) + p_p(\mathbf{u}_G) log(1 - D(\mathbf{u}_G)] \, d\mathbf{u}_G$$

$$KL(p_t \parallel p_p) = \int_{-\infty}^{\infty} p_t(x) \log\left(\frac{p_t(x)}{p_p(x)}\right) dx$$

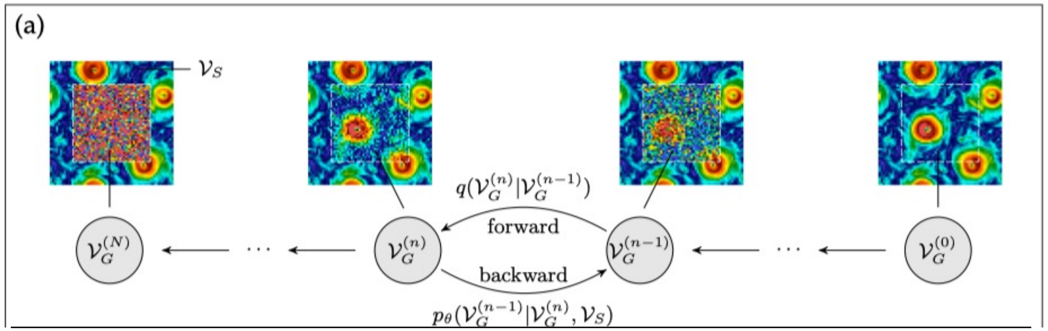
KULLBACK-LEIBLER DISTANCE




## 1. DIFFUSION MODELS
## UNCONDITIONAL
## GENERATION OUT OF I.I.D. NOISE




## RePAINT:
## UNCONDITIONAL
## INPAINTING OUT OF I.I.D. NOISE




## PALETTE:
## CONDITIONAL
## INPAINTING OUT OF I.I.D. NOISE

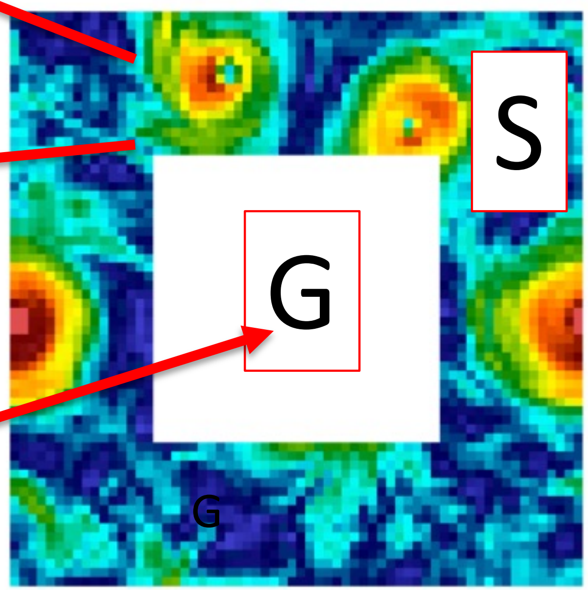
Sohl-Dickstein et al., Deep Unsupervised Learning using Nonequilibrium Thermodynamics, ICML 2015
Ho et al., Denoising Diffusion Probabilistic Models, NeurIPS 2020
Song et al., Score-Based Generative Modeling through Stochastic Differential Equations, ICLR 2021

# LINEAR: EXTENDED-PRINCIPAL ORTHOGONAL DECOMPOSITION (EPOD)

$$R_S(x,y) = \langle u(x)u(y)\rangle$$
$$\int_\Omega R_S(x,y)\phi_S^{(n)}(y)\,dy = \sigma_n \phi_S^{(n)}(x),$$

MSE = $\int_I \|\boldsymbol{u}_G^{(p)}(\boldsymbol{x}) - \boldsymbol{u}_G^{(t)}(\boldsymbol{x})\|^2 \,\mathrm{d}\boldsymbol{x}$

$$u_S(x) = \sum_{n=1}^{N_S} b_S^{(n)} \phi_S^{(n)}(x),$$
$$\phi_S^{(n)}(x) = \langle b_S^{(n)} u_S(x)\rangle/\sigma_n.$$




TRAINING

$$\phi_E^{(n)}(x) = \langle b_S^{(n)} u_G(x)\rangle/\sigma_n.$$

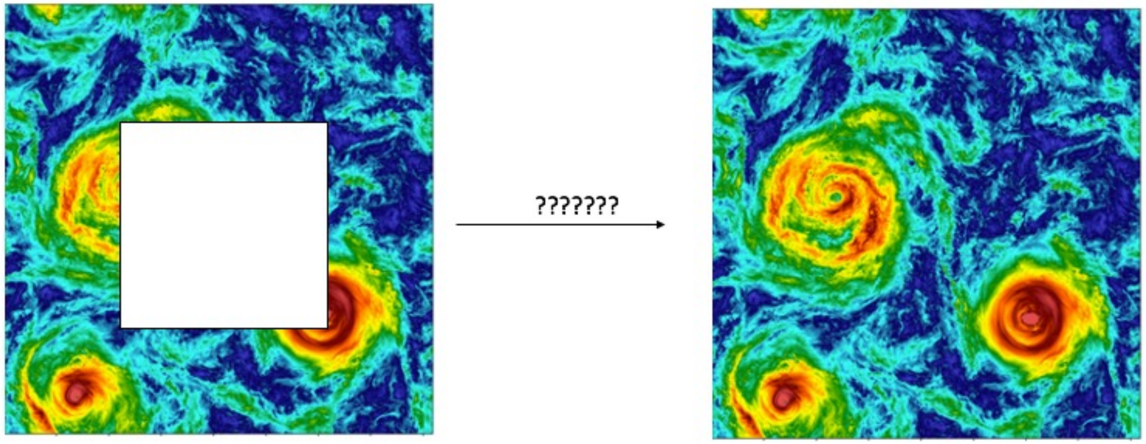
RECONSTRUCTION

$$u_G^{(p)}(x) = \sum_{n=1}^{N_S} b_S^{(n)} \phi_E^{(n)}(x).$$


J. Boree, Extended POD a tool to analyse correlated events in turbulence. Experiments in Fluids 35(2):188 (2003)
T. Li, L. Buzzicotti, F. Bonaccorso, L.B., S. Chen, M. Wan. Data reconstruction of turbulent flows with Gappy-POD and Generative Adversarial Networks. JFM (2023)

**NUDGING**: AN **EQUATION-INFORMED UNBIASED** TOOL TO ASSIMILATE AND RECONSTRUCT TURBULENCE DATA/PHYISICS 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)




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

FULLY PHYSICS COMPLIANT

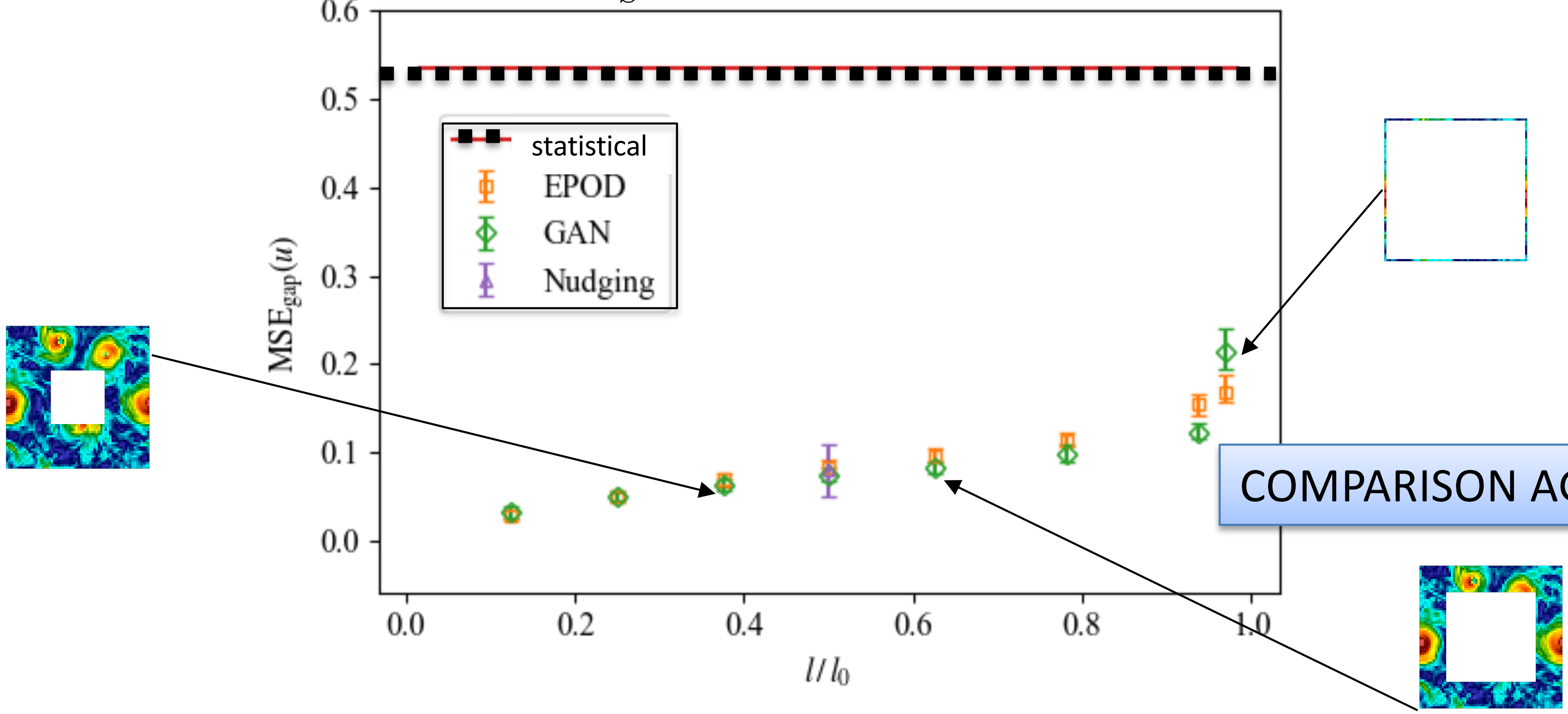
NEED HUGE COMPUTATIONAL RESOURCES

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

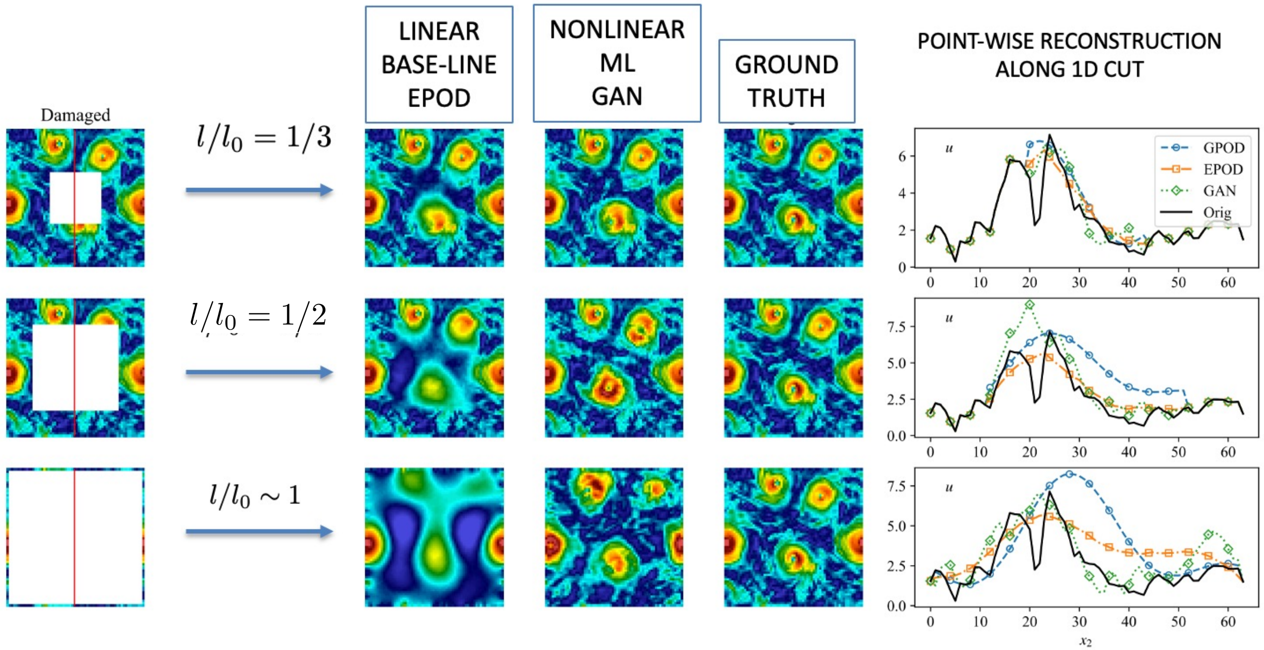
NO NEED TO TRAIN!! NAVIER AND STOKES DID THE JOB FOR YOU: ONE CONF IS ENOUGH

**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)

$$\text{MSE} = \langle \int_S d\mathbf{x}(u_{true}(\mathbf{x}) - u_{pred}(\mathbf{x}, \theta))^2 \rangle$$




## COMPARISON AGAINST GAN

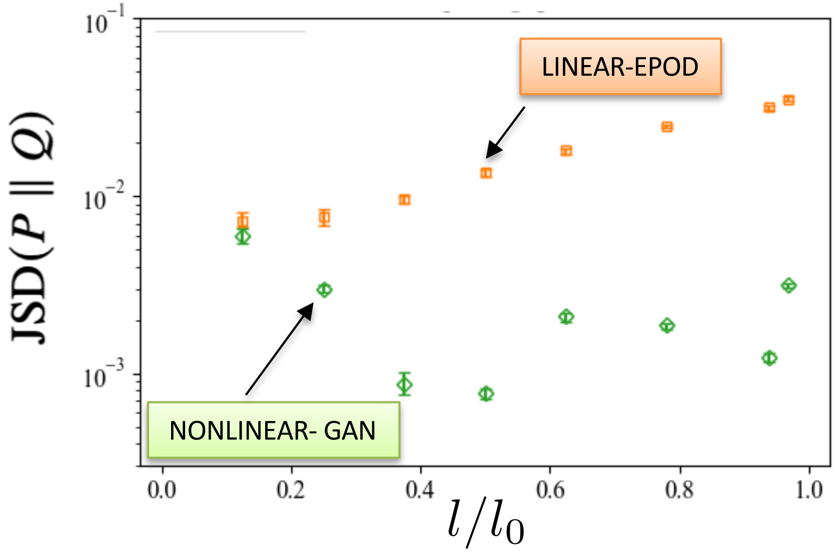


KULLBACK-LEIBLER

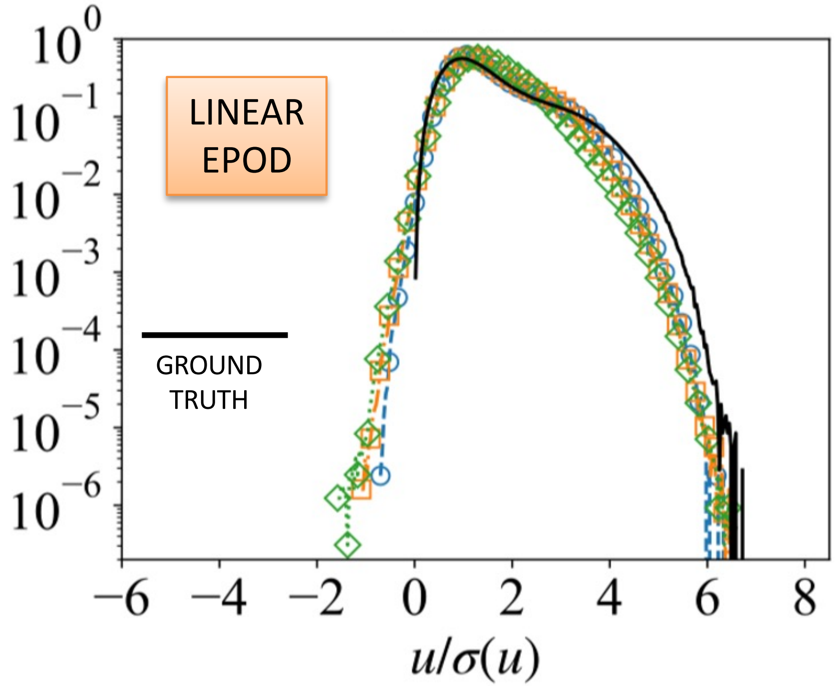
$$D(P \parallel Q) = \int_{-\infty}^{\infty} P(x) \log\left(\frac{P(x)}{Q(x)}\right) \mathrm{d}x$$

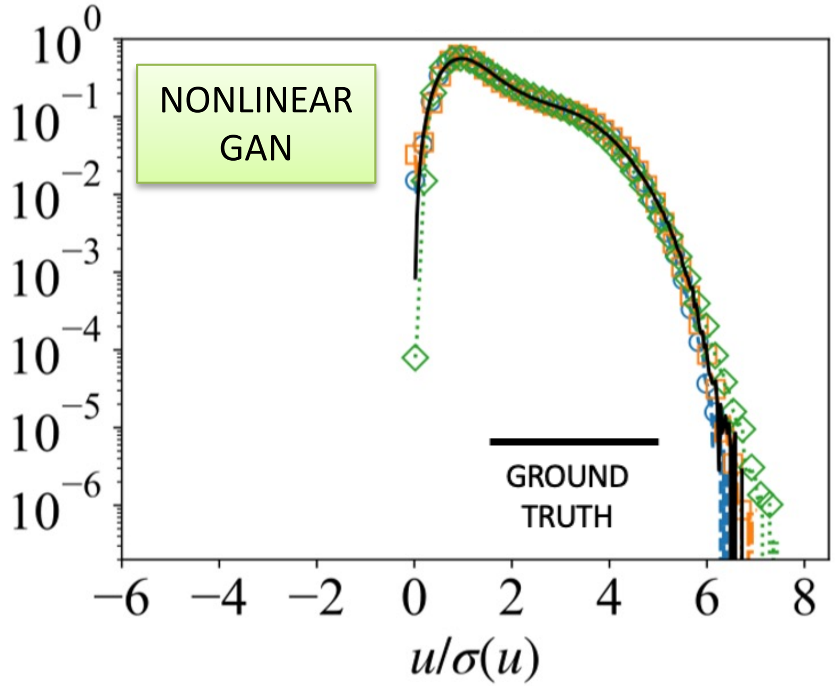
JENSEN-SHANNON

$$\xrightarrow[M = \frac{1}{2}(P+Q)]{} \quad \mathrm{JSD}(P \parallel Q) = \frac{1}{2} D(P \parallel M) + \frac{1}{2} D(Q \parallel M),$$

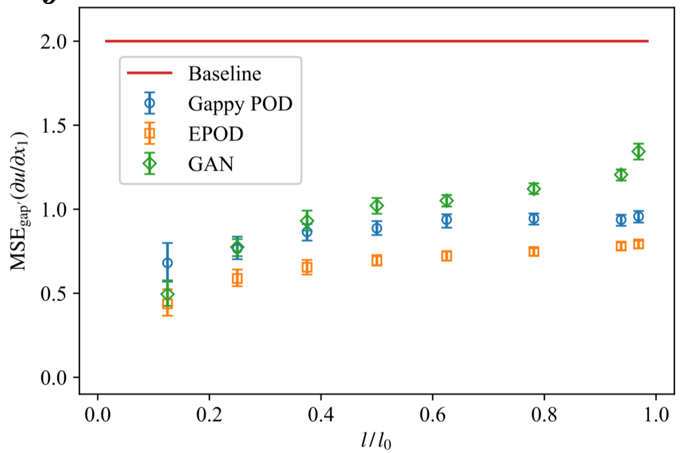



COMPARISON AGAINST GAN

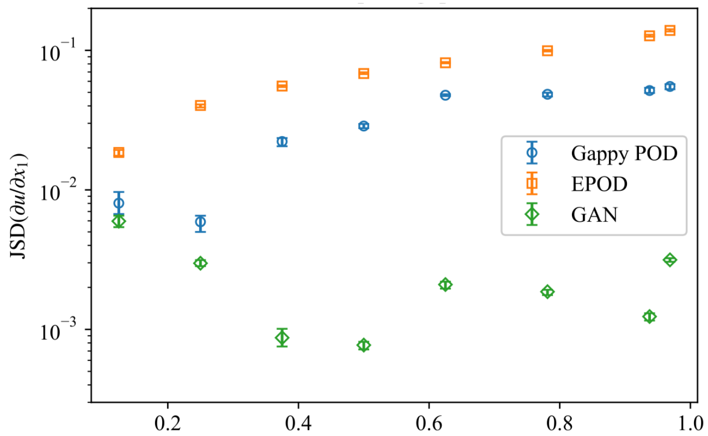


$$\int d\mathbf{x}[\partial_x u^p(\mathbf{x},\theta) - \partial_x u^t(\mathbf{x})]^2$$



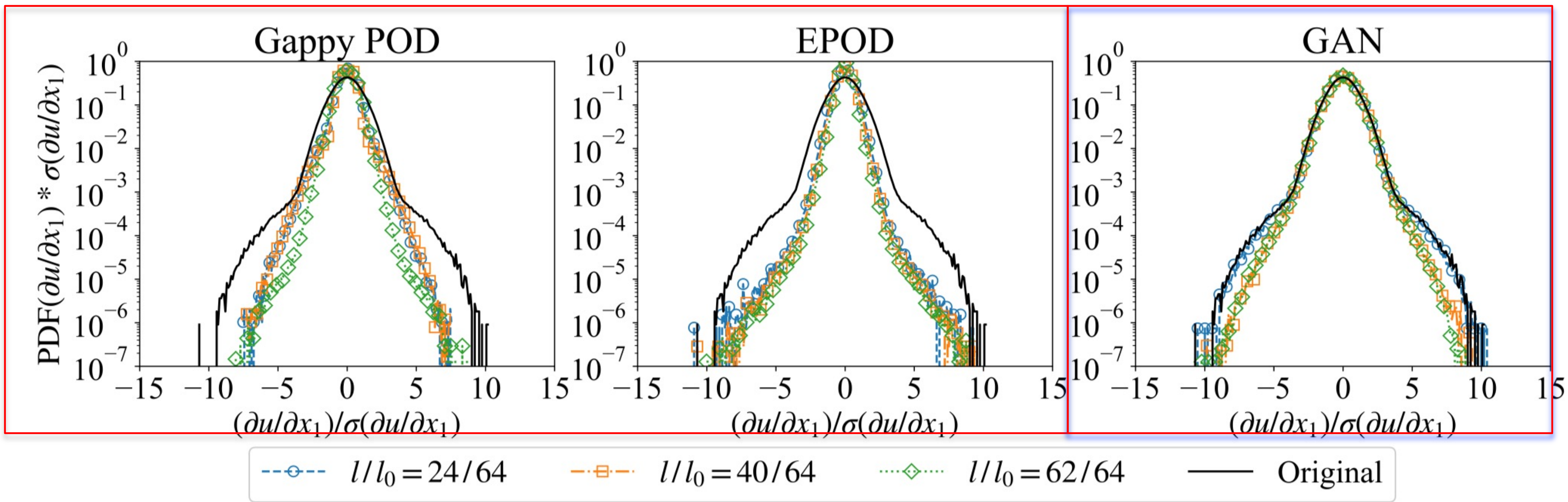

$$\mathrm{JSD}(P \parallel Q) = \frac{1}{2}D(P \parallel M) + \frac{1}{2}D(Q \parallel M),$$

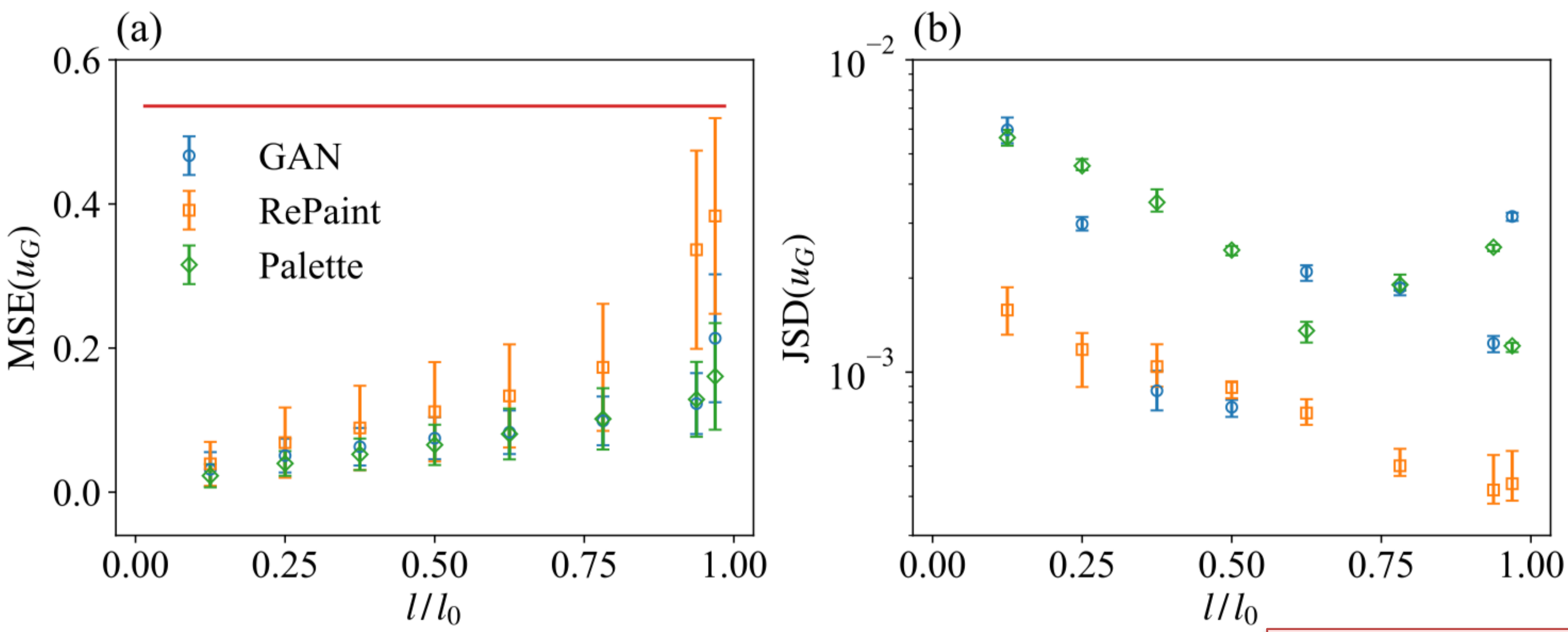



COMPARISON AGAINST GAN

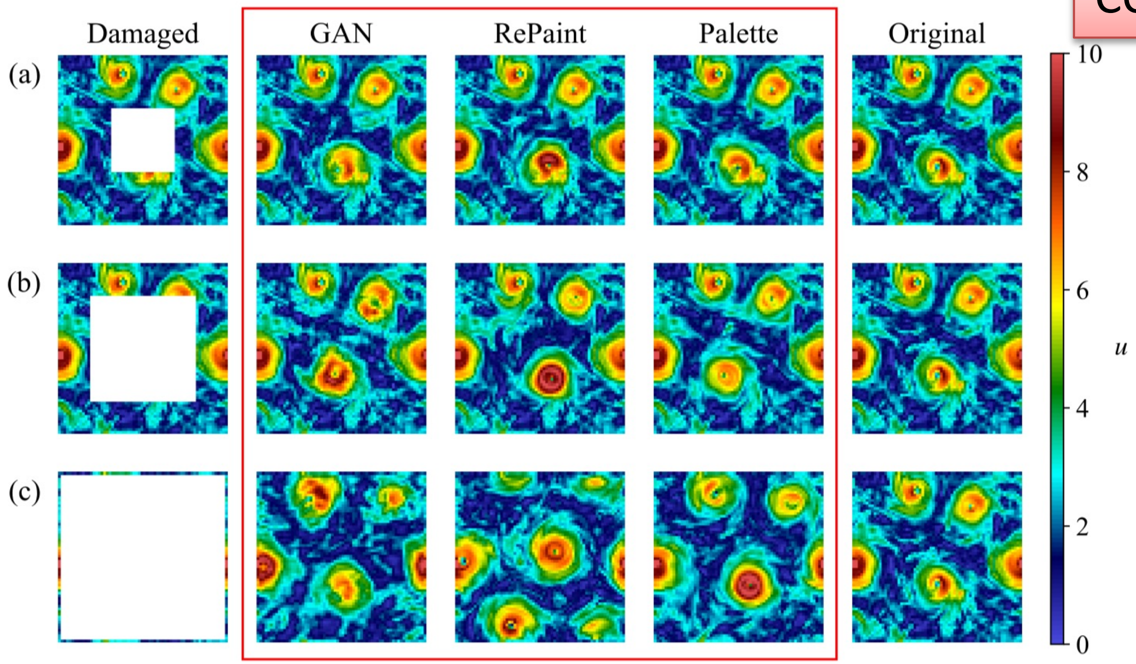
LINEAR GAPPY-POD & EPOD

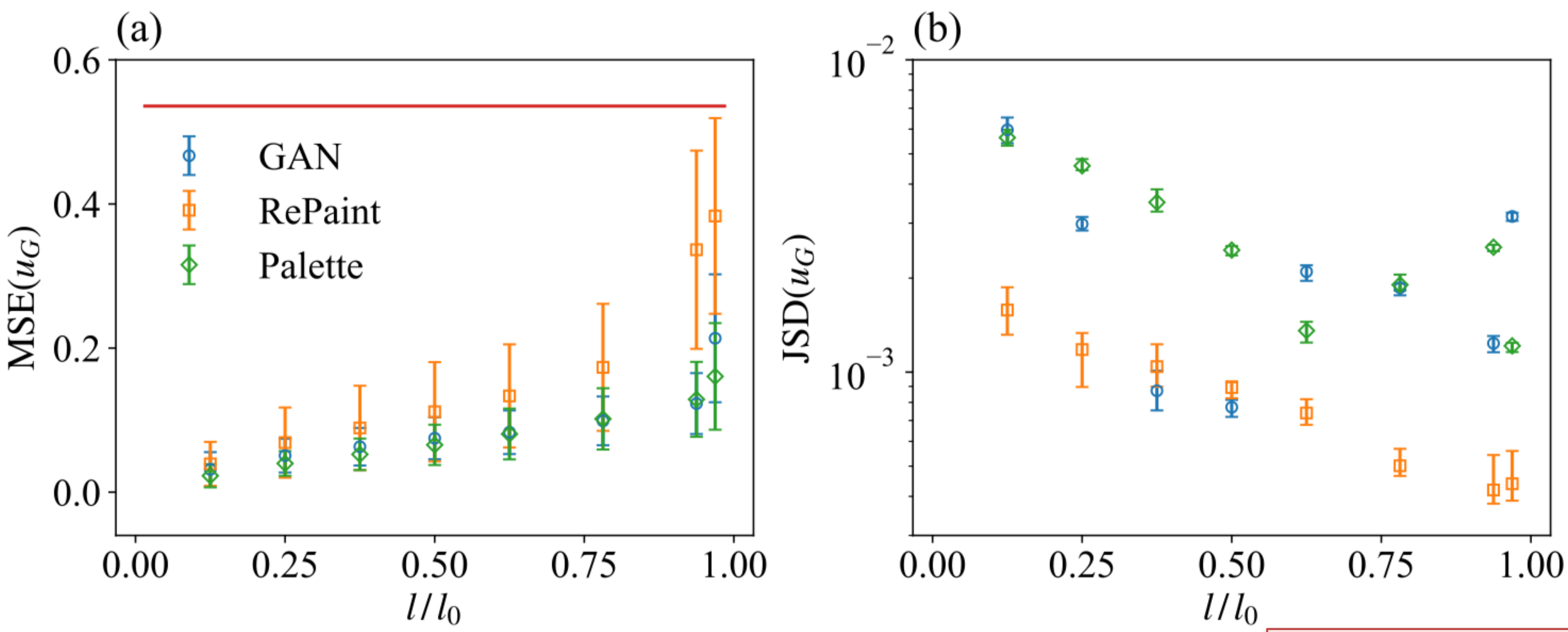
NONLINEAR GAN



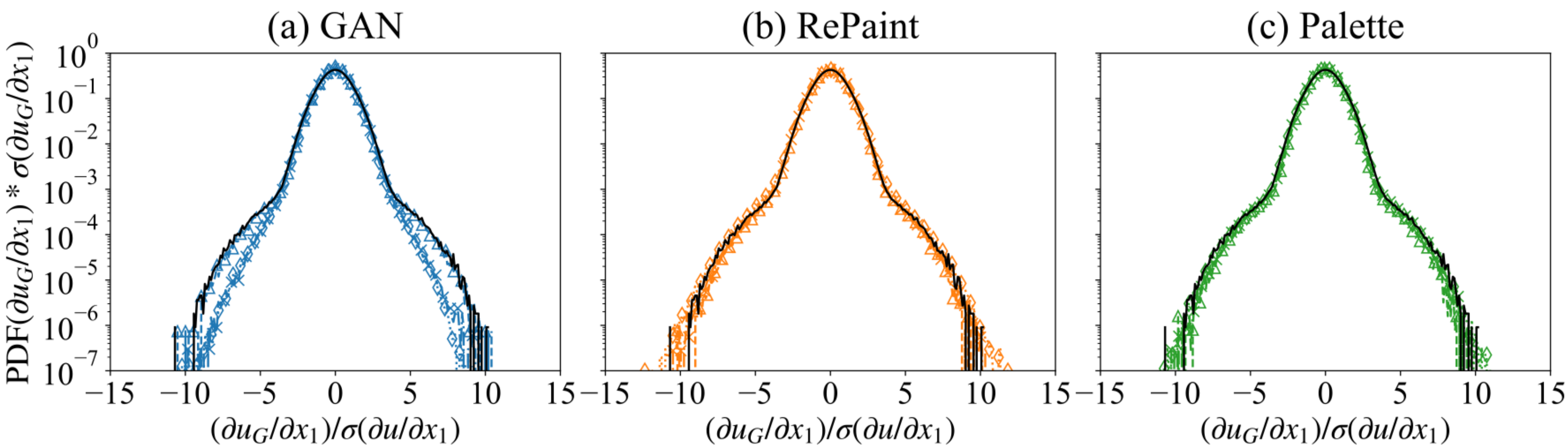



COMPARISON GAN/DM



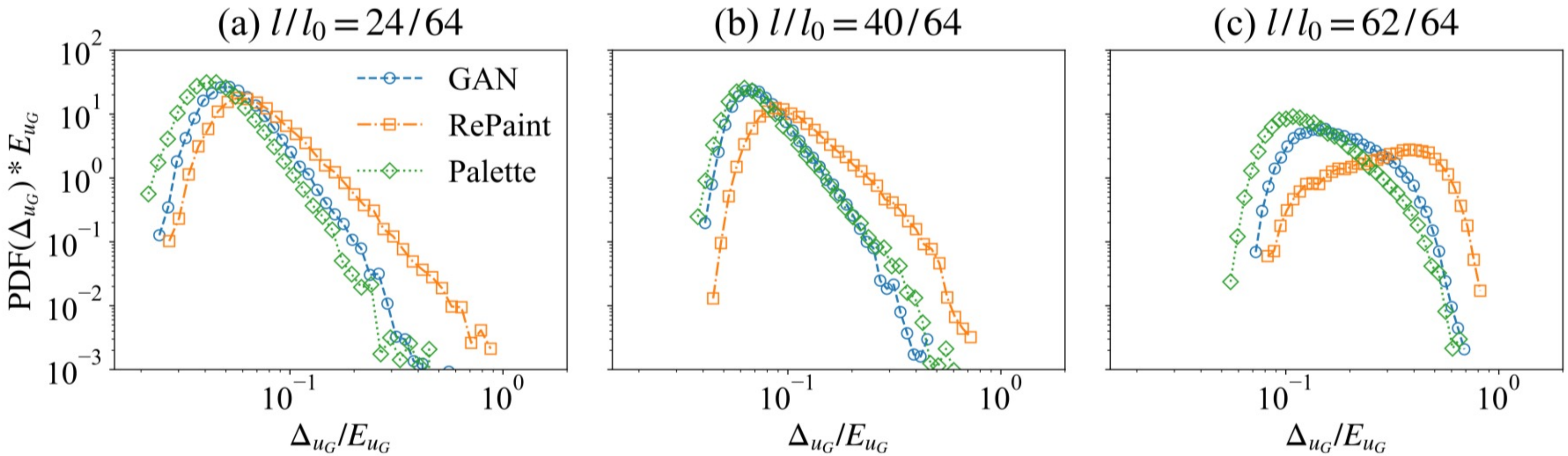


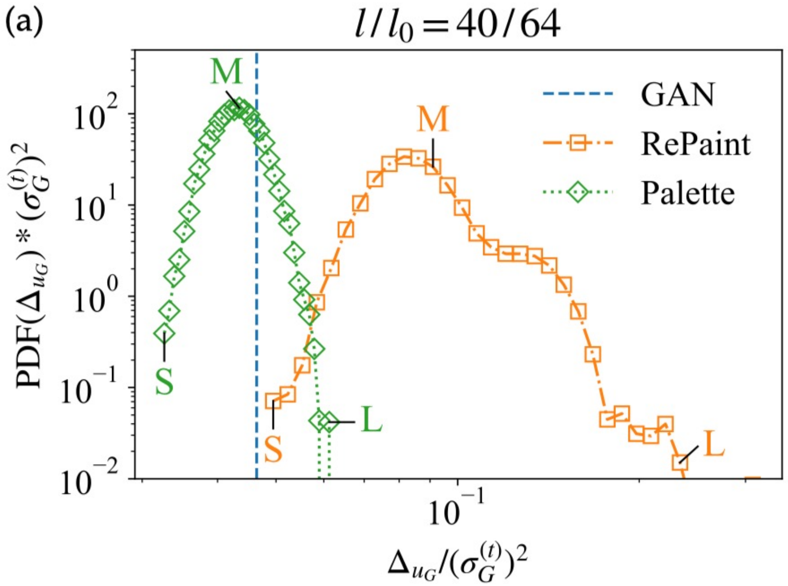

COMPARISON GAN/DM

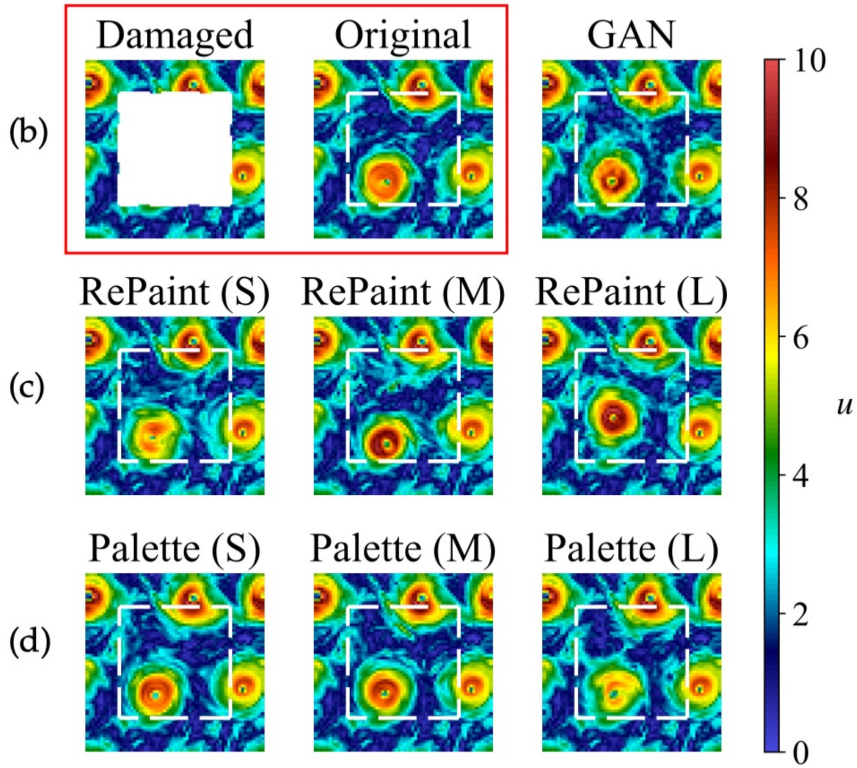


COMPARISON GAN/DM

PDF OF L2 ERROR – SINGLE IMAGE



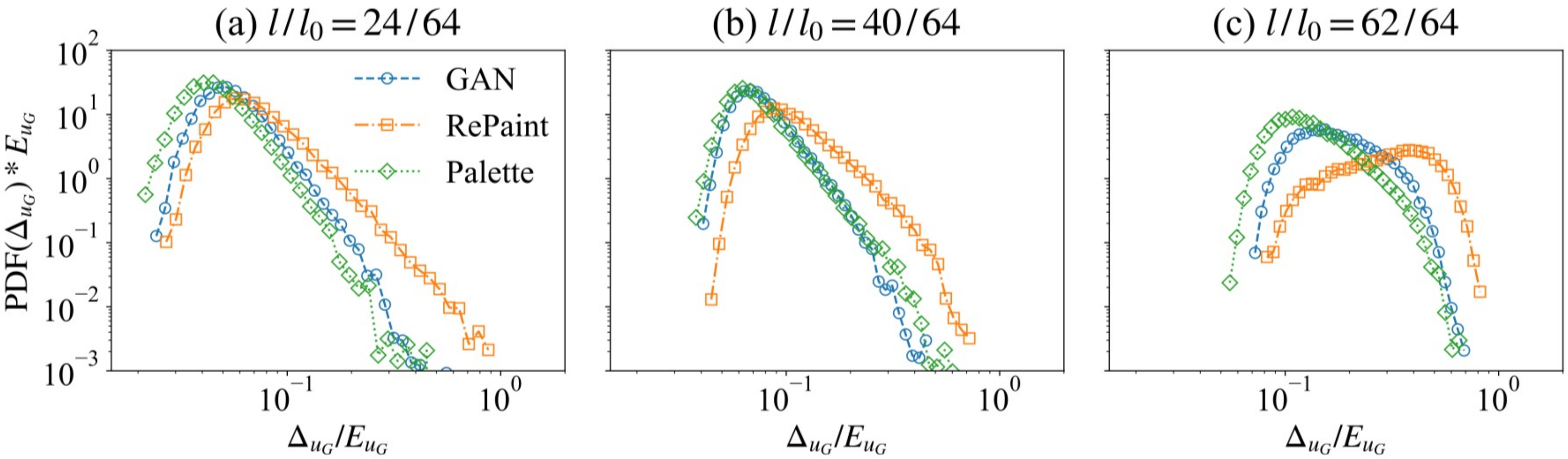



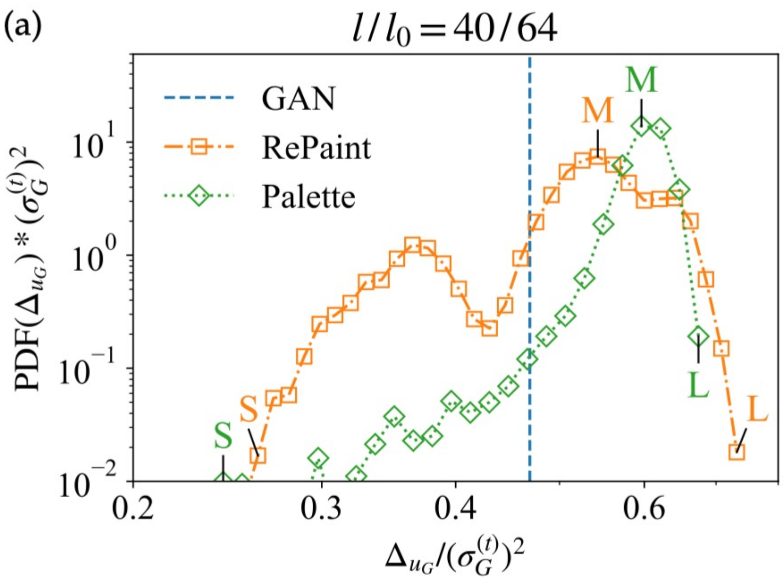



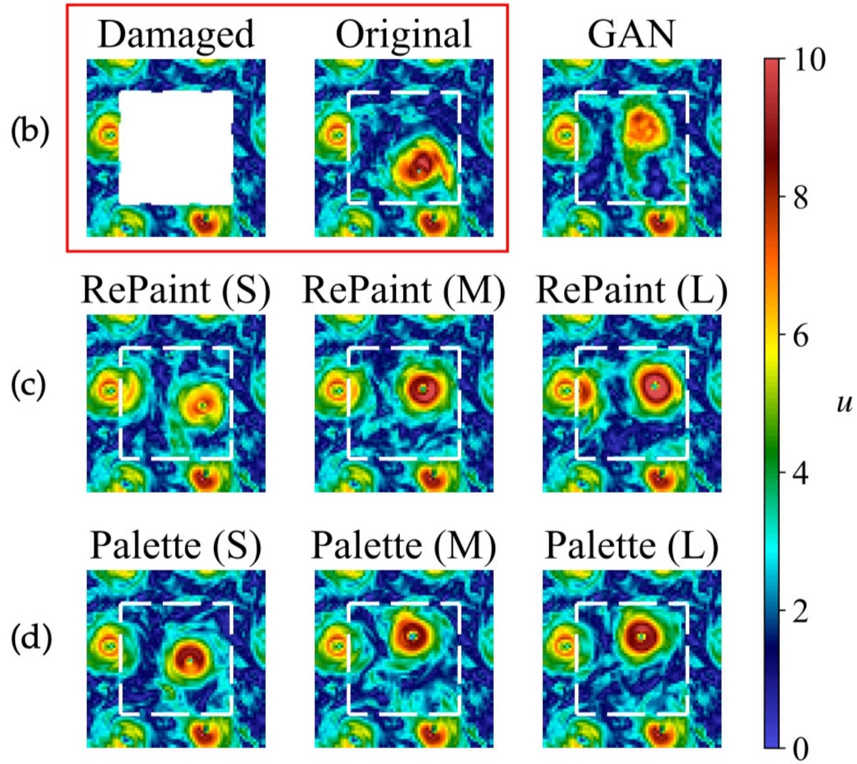

UNCERTAINTY QUANTIFICATION

# COMPARISON GAN/DM

## PDF OF L2 ERROR – SINGLE IMAGE

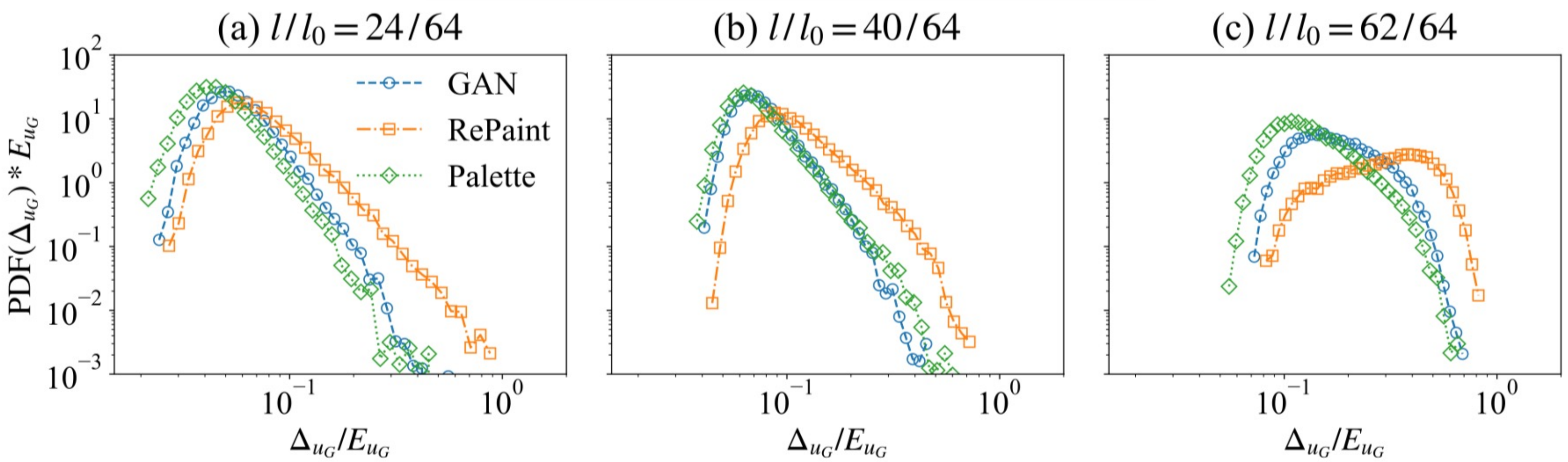


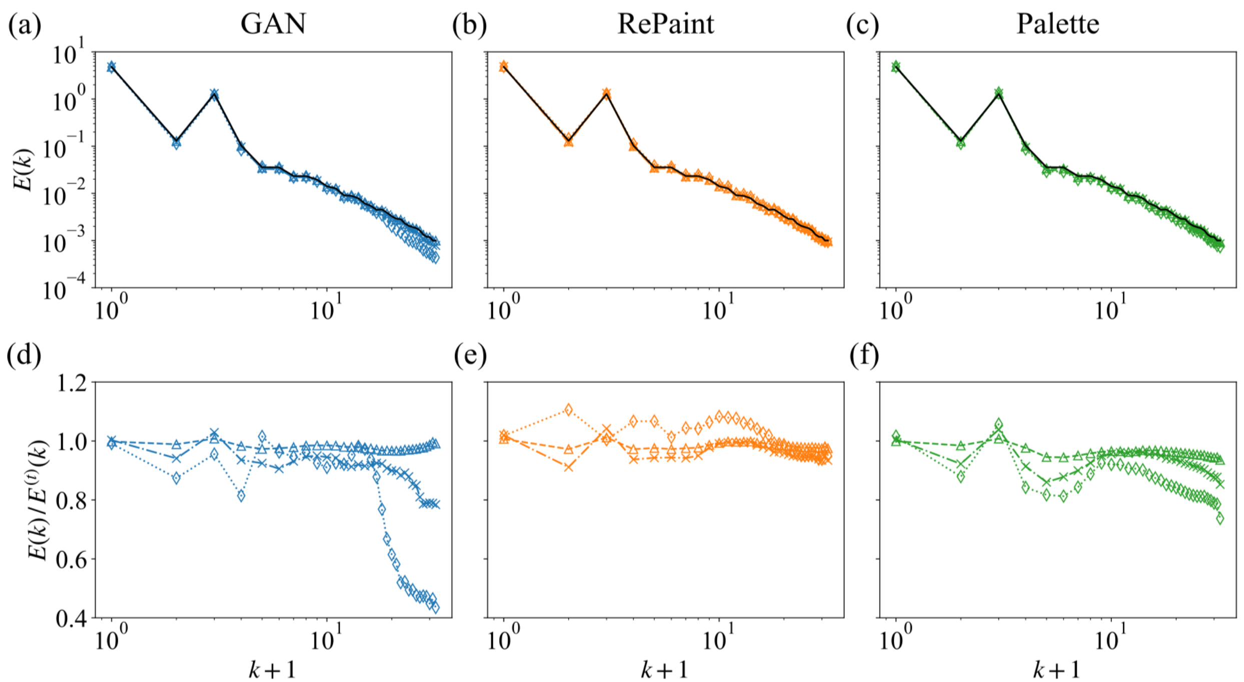




## UNCERTAINTY QUANTIFICATION

# PDF OF L2 ERROR – SINGLE IMAGE



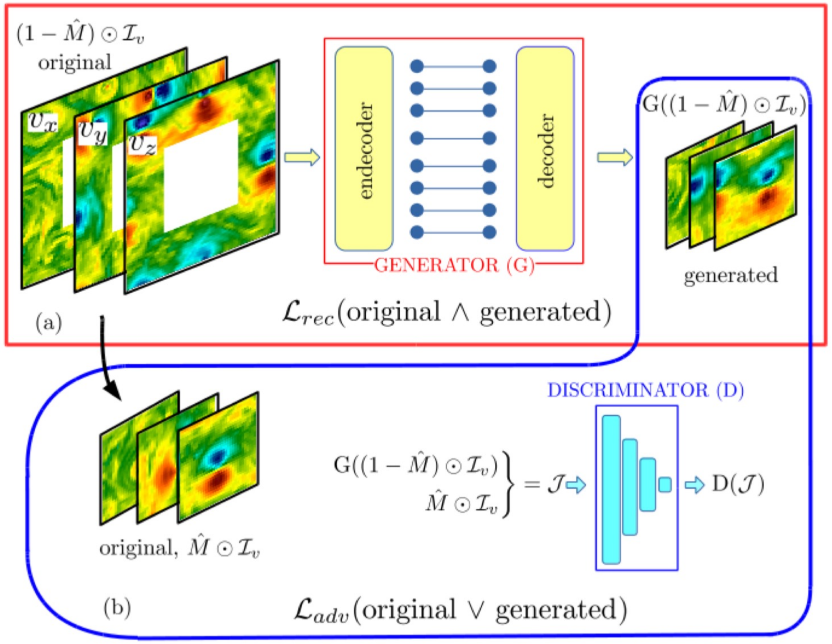

# COMPARISON GAN/DM

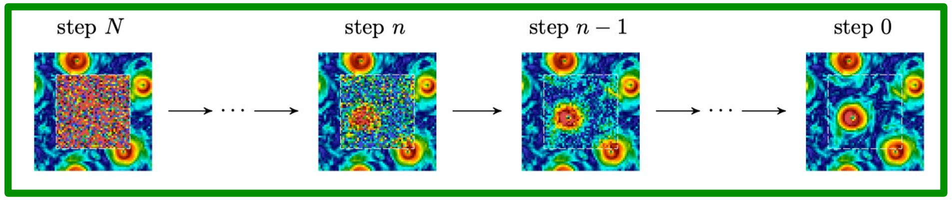



# STATISTICAL RECONSTRUCTION ENERGY SPECTRA

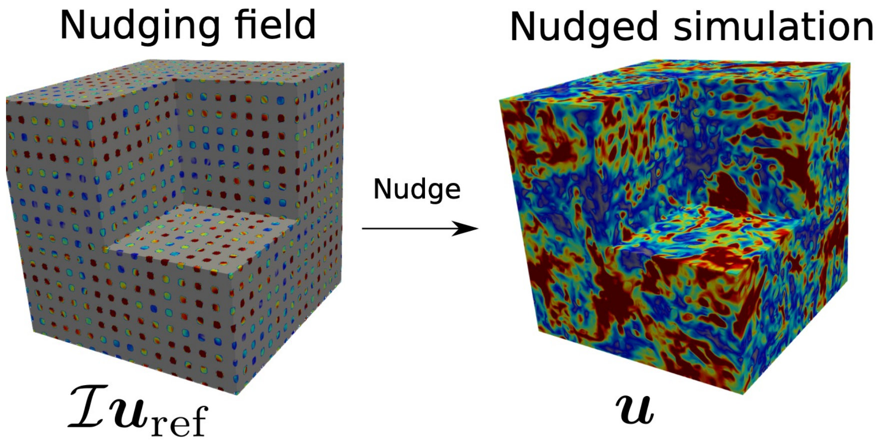
# GAN/DM

# NUDGING




$$\begin{cases}\partial_t \mathbf{v} + \mathbf{v}\cdot\partial_{\mathbf{x}}\mathbf{v} + \partial_{\mathbf{x}}P - \nu\Delta\mathbf{v} = 2\mathbf{v}\times\mathbf{\Omega} + \mathcal{S}\mathbf{v} + \alpha g\hat{\mathbf{z}}T + \mathcal{F} - N(\mathbf{v}_N - \mathbf{v}) \\ \partial_t T + \mathbf{v}\cdot\partial_{\mathbf{x}}T - \chi\Delta T = \mathcal{G}v_z + \mathcal{L} - N_T(T_N - T)\end{cases}$$

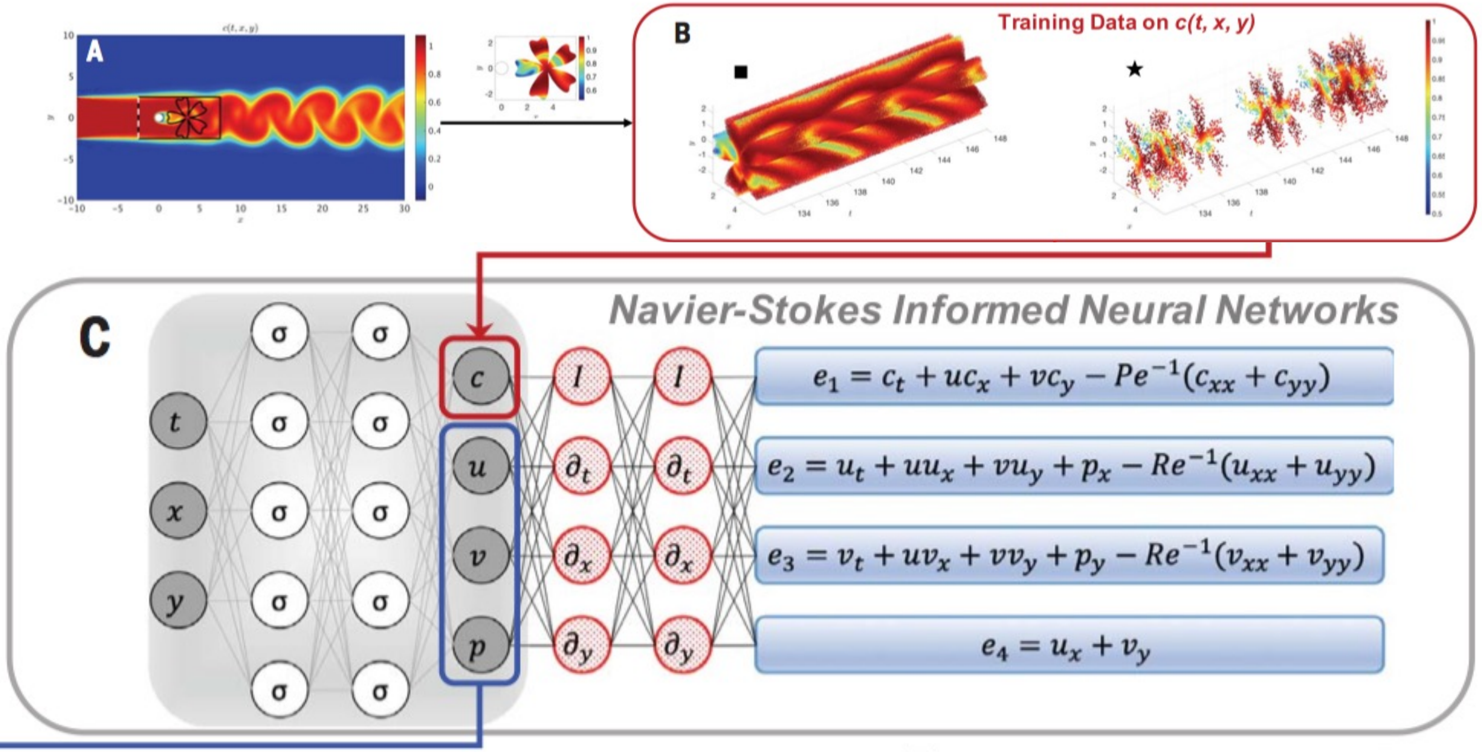



**GAN/DM**

- -EQUATION-FREE
- -DATA HUNGRY
- -MISSING PHYSICS (SEE LATER)
- +ONCE TRAINED -> INSTANTANEOUS
- +MIXED INPUT FEATURES
- +OK STATISTICAL UNCERTAINTIES

**NUDGING**

- -NEEDS EQUATIONS
- -CPU HUNGRY
- -RESTRICTED INPUT FEATURES
- +PHYSICS COMPLIANT
- +NO TRAINING
- +OK STATISTICAL UNCERTAINTIES




**PHYSICS INFORMED**

$$MSE = \frac{1}{N}\sum_{n=1}^{N}\left|c(t^n, x^n, y^n, z^n) - c^n\right|^2 + \sum_{i=1}^{5} 4\,\frac{1}{M}\sum_{m=1}^{M}\left|e_i(t^m, x^m, y^m, z^m)\right|^2$$

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


**Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. M. Raissi A. Yazdani , G. E. Karniadakis , Science 367, 1026–1030 (2020)**

**Physics-Informed Neural Network for Ultrasound Nondestructive Quantification of Surface Breaking Cracks** K. Shukla, P. Clark Di Leoni, J. Blackshire, D. Sparkman & G. E. Karniadakis *Journal of Nondestructive Evaluation* 39 (2020)

## WHAT WE HAVE:

+ "QUICK" STOCHASTIC TOOL TO GENERATE/AUGMENT REALISTIC 3D EULERIAN CONFIGURATIONS AND 3D LAGRANGIAN TRAJECTORIES IN TURBULENCE, EASY TO GENERALISE FOR DIFFERENT APPLICATIONS

+ HIGH QUANTITATIVE AGREEMENT WITH MULTI-SCALE STATISTICAL PROPERTIES

## WHAT WE MISS:

- UNDERSTADING OF ROBUSTNESS IN GENERALISING OUT-OF-SAMPLE: EXTREME EVENTS, DIFFERENT REYNOLDS NUMBERS & DIFFERENT FLUIDS PROPERTIES

- UNDERSTANDING SCALING PROPERTIES FOR TIME-TO-SOLUTION AT CHANGING IN-SAMPLE PROPERTIES, I.E. AT CHANGING DIMENSION OF THE TRAINING DATASET, SETS OF HYPER-PARAMETERS, CNN ARCHITECTURES: GAN, DM, TRANSFORMERS

-WHAT-IF QUESTIONS: EXPLICABILITY OF THE GENERATED DATA, FEATURES RANKINGS, PHYSICS DISCOVERY

**Wavelet Score-Based Generative Modeling**

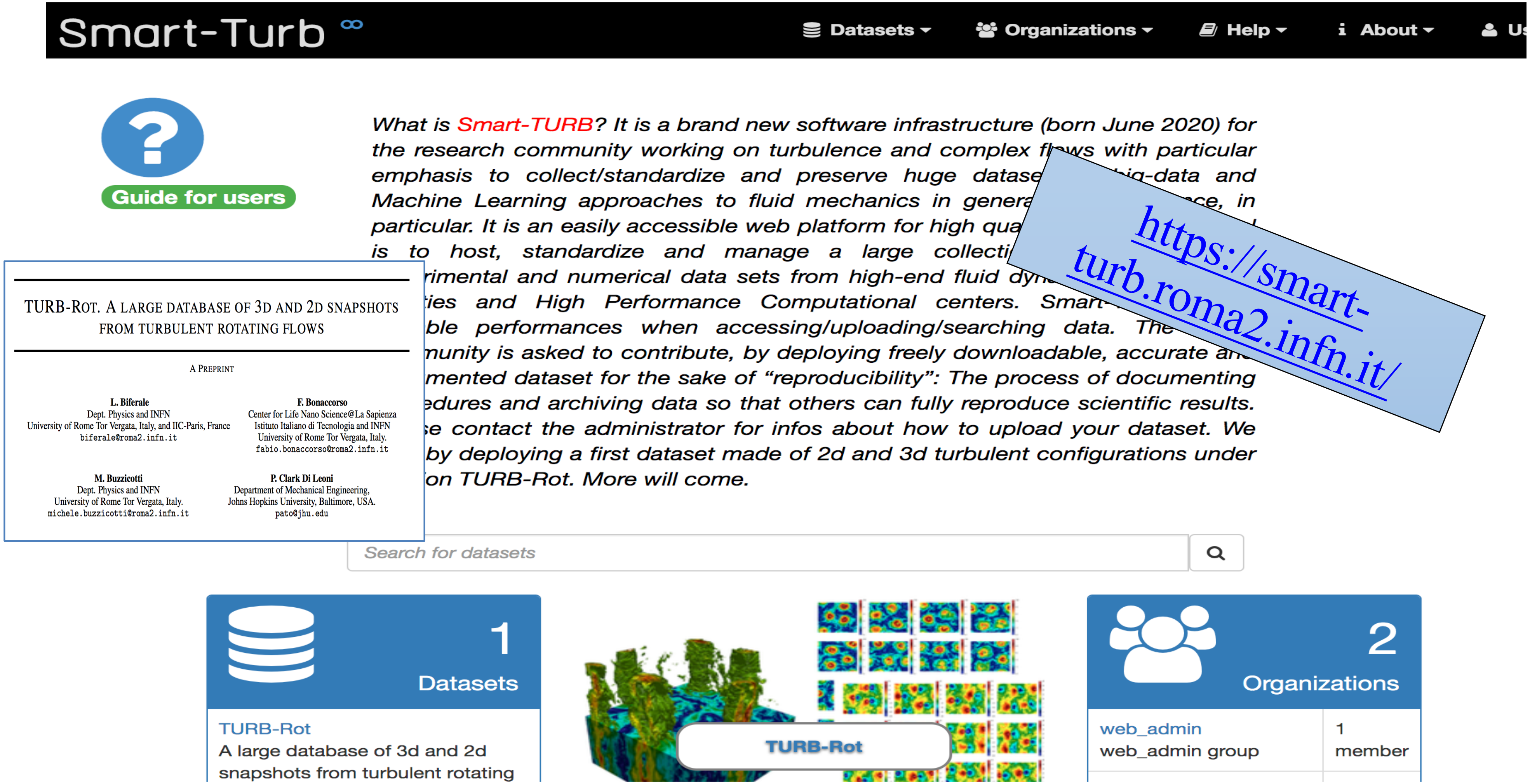
**Florentin Guth**
Computer Science Department,
ENS, CNRS, PSL University

**Simon Coste**
Computer Science Department,
ENS, CNRS, PSL University

**Valentin De Bortoli**
Computer Science Department,
ENS, CNRS, PSL University

**Stéphane Mallat**
Collège de France, Paris, France
Flatiron Institute, New York, USA

arxive 2208.05003



Smart-Turb
Datasets
Organizations
Help
About
Guide for users
What is Smart-TURB? It is a brand new software infrastructure (born June 2020) for the research community working on turbulence and complex flows with particular emphasis to collect/standardize and preserve huge datase
Machine Learning approaches to fluid mechanics in genera
particular. It is an easily accessible web platform for high qua
is to host, standardize and manage a large collecti
rimental and numerical data sets from high-end fluid dyn
ies and High Performance Computational centers. Smart-
ble performances when accessing/uploading/searching data. The
munity is asked to contribute, by deploying freely downloadable, accurate an
mented dataset for the sake of "reproducibility": The process of documenting
edures and archiving data so that others can fully reproduce scientific results.
e contact the administrator for infos about how to upload your dataset. We
by deploying a first dataset made of 2d and 3d turbulent configurations under
on TURB-Rot. More will come.
https://smart-turb.roma2.infn.it/
TURB-Rot. A large database of 3D and 2D snapshots from turbulent rotating flows
A Preprint
L. Biferale
Dept. Physics and INFN
University of Rome Tor Vergata, Italy, and IIC-Paris, France
biferale@roma2.infn.it
F. Bonaccorso
Center for Life Nano Science@La Sapienza
Istituto Italiano di Tecnologia and INFN
University of Rome Tor Vergata, Italy.
fabio.bonaccorso@roma2.infn.it
M. Buzzicotti
Dept. Physics and INFN
University of Rome Tor Vergata, Italy.
michele.buzzicotti@roma2.infn.it
P. Clark Di Leoni
Department of Mechanical Engineering,
Johns Hopkins University, Baltimore, USA.
pato@jhu.edu
Search for datasets
1
Datasets
TURB-Rot
A large database of 3d and 2d snapshots from turbulent rotating
TURB-Rot
2
Organizations
web_admin
web_admin group
1
member