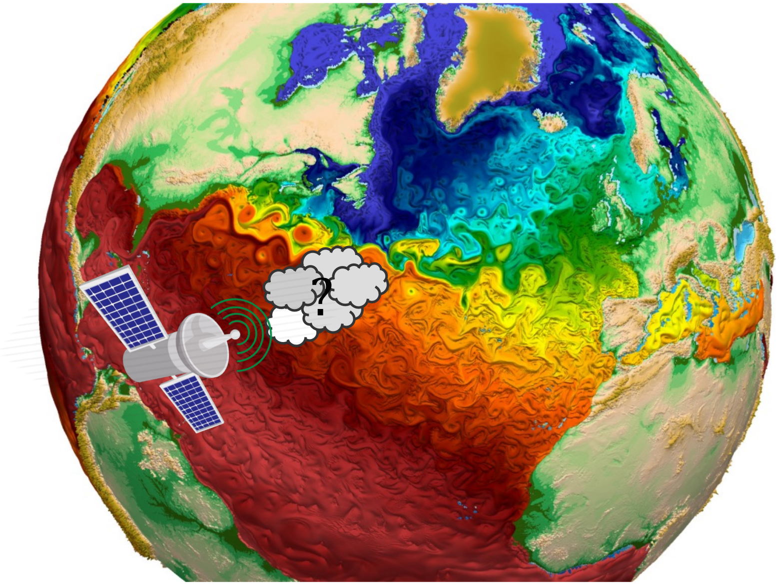
AQTIVATE

WORKSHOP: Data-Driven and Model-Based Tools For Complex Flows and Complex Fluids
Rome - Dept. Physics University of Rome Tor Vergata June/3-7/2024

Stochastic Multi-Scale Reconstruction of Turbulent Flows with Data-Driven Generative Models



Los Alamos National Laboratory ([CC BY-NC-ND 2.0](CC BY-NC-ND 2.0))

Michele Buzzicotti

Dept. Physics & INFN, University of Rome Tor Vergata

https://mbuzzico.github.io/

michele.buzzicotti@roma2.infn.it

Luca Biferale (Uni. Tor Vergata, IT),
Fabio Bonaccorso, **Tianyi Li** (Uni. Tor Vergata, IT),
Hussein Aluie, Ben Storer (Uni. Rochester, NY),
Alessandra Lanotte (CNR-Nanotec, IT)

TOR VERGATA UNIVERSITÀ DEGLI STUDI DI ROMA

INFN

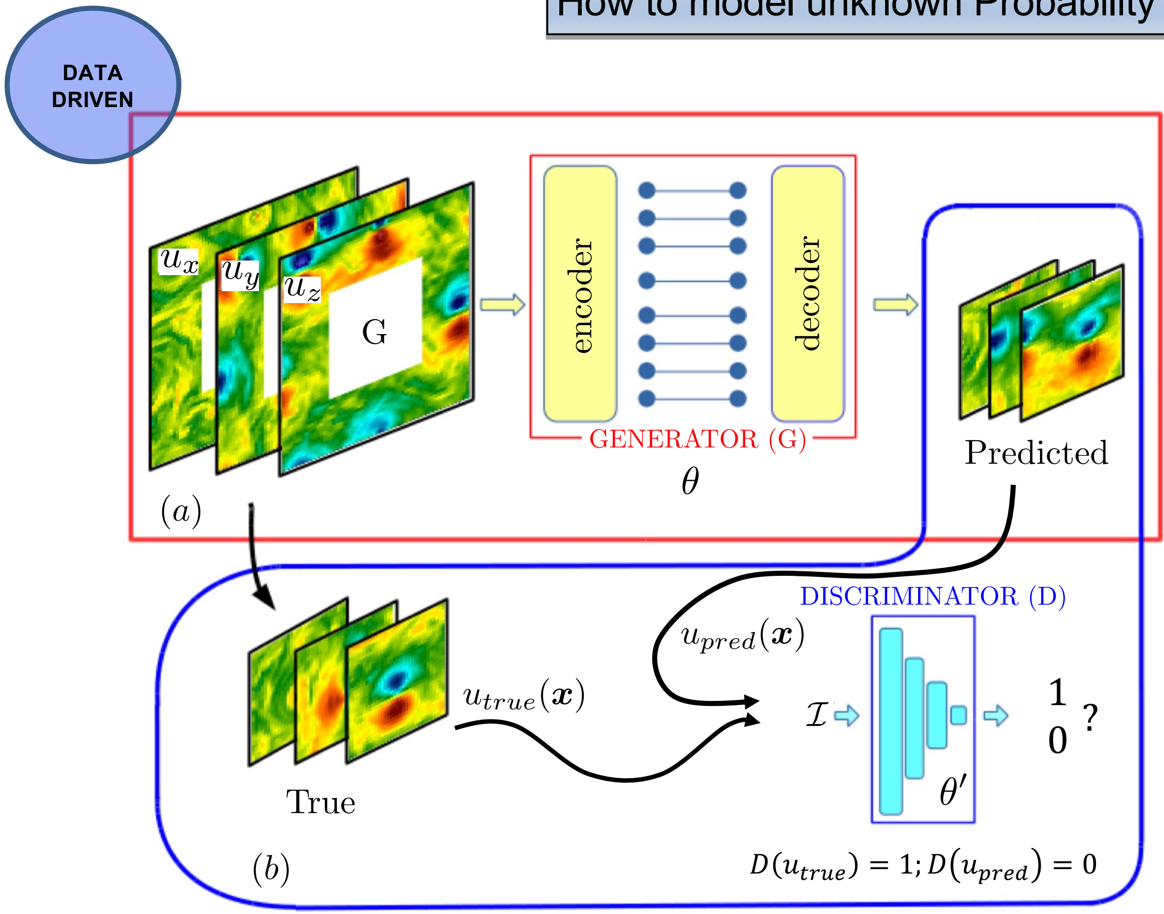
MUR Ministero dell'Università e della Ricerca

FARE RICERCA IN ITALIA FRAMEWORK PER L'ATTRAZIONE E IL RAFFORZAMENTO DELLE ECCELLENZE PER LA RICERCA IN ITALIA

erc European Research Council Established by the European Commission Supporting top researchers from anywhere in the world

Can data-driven solutions be improved with machine learning methods?

Generative Adversarial Networks
How to model unknown Probability Distribution Functions?



The Generator is trained to Minimize:

$$\mathcal{L}_2 = \langle \int_G d\boldsymbol{x} \, (u_{true}(\boldsymbol{x}) - u_{pred}(\boldsymbol{x}, \theta))^2 \rangle$$

$$\mathcal{L}_{tot} = \mathcal{L}_2 + \lambda \mathcal{L}_{adv}$$

The Discriminator is trained to Maximize:

$$\mathcal{L}_{adv} = \log(D(u_{true})) + \log(1 - D(u_{pred}))$$

- The discriminator learns the two probabilities describing true and predicted data:

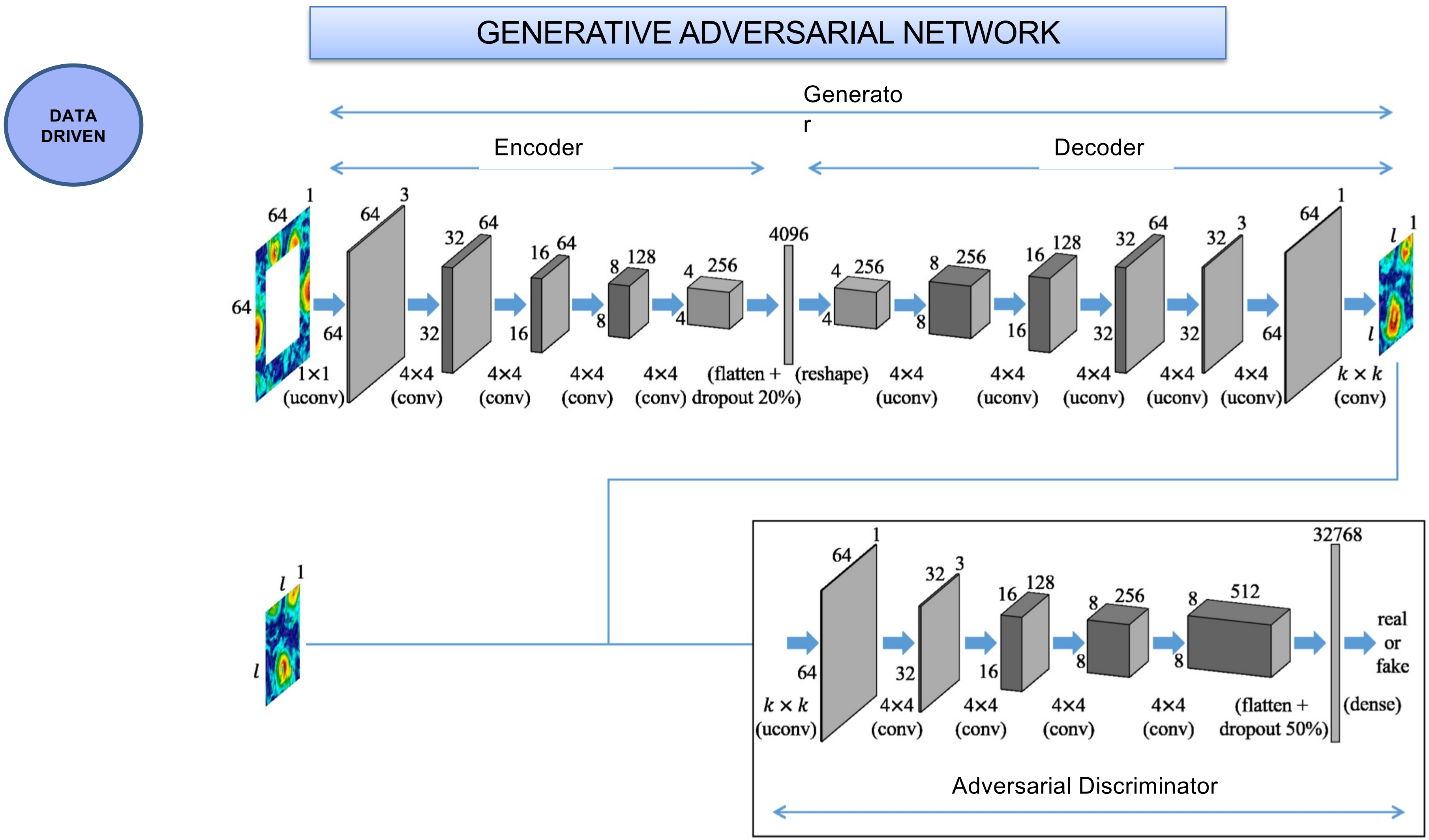
$$P_{true}(\mathbf{u}) \qquad P_{pred}(\mathbf{u})$$

- The Generator minimizes the **Kullback-Leibler Divergence** between predicted and true data

$$\mathrm{KL}(P_{true} || P_{pred}) = \int_{-\infty}^{+\infty} P_{true}(\mathbf{u}) \log \left(\frac{P_{true}(\mathbf{u})}{P_{pred}(\mathbf{u})} \right) \mathrm{d}\mathbf{u}$$

Key point: **Adversarial Training allows to introduce a statistical term in the reconstruction Loss**

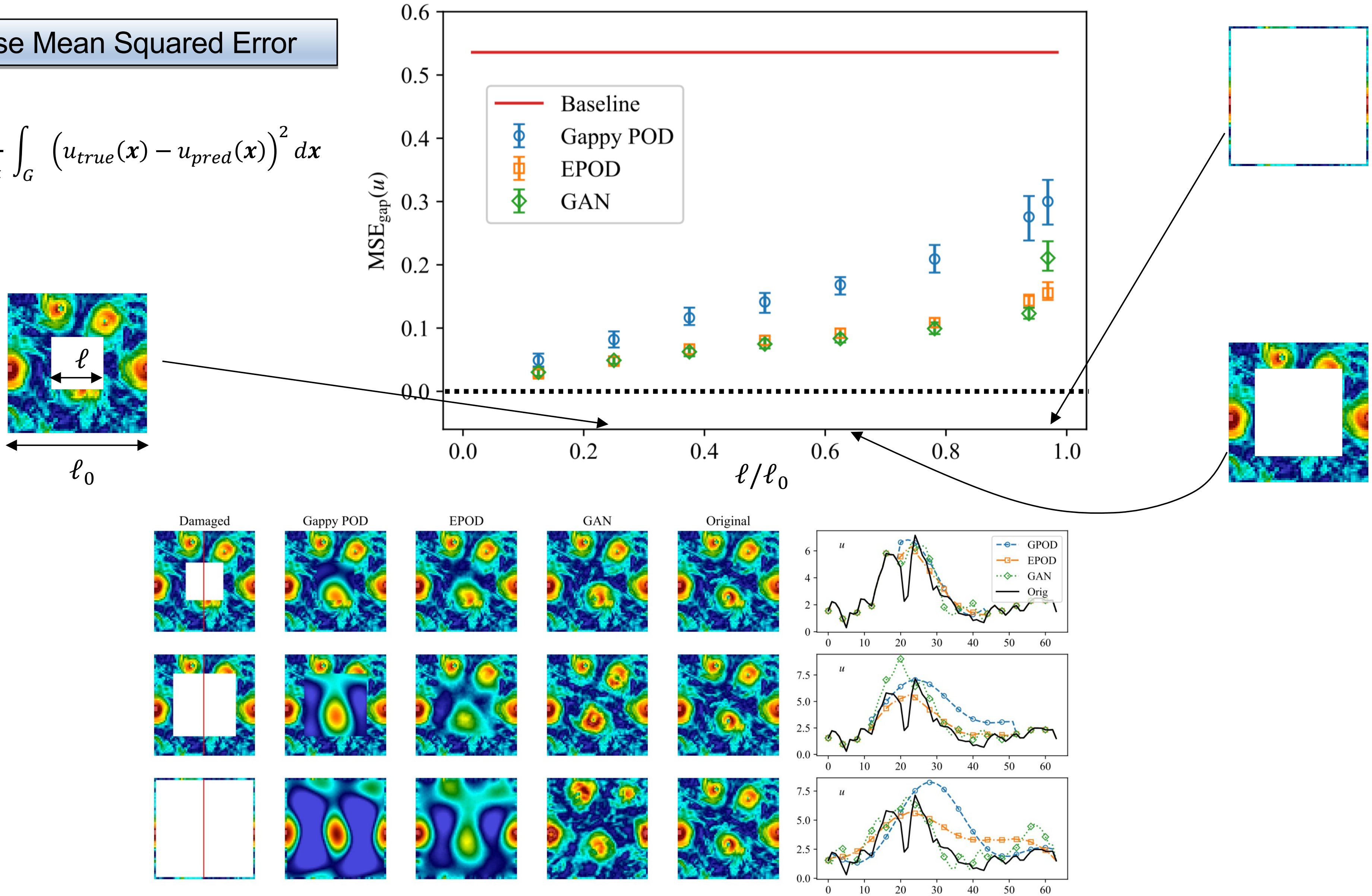
Reconstruction of turbulent data with deep generative models for semantic inpainting from TURB-Rot database
M. Buzzicotti, F. Bonaccorso, P. Clark Di Leoni, and L. Biferale, Phys. Rev. Fluids **6**, 050503, 2021



GENERATIVE ADVERSARIAL NETWORK
DATA DRIVEN
Generato
r
Encoder
Decoder
1×1 (uconv)
4×4 (conv)
4×4 (conv)
4×4 (conv)
4×4 (conv)
(flatten + dropout 20%)
4096
(reshape)
4×4 (uconv)
4×4 (uconv)
4×4 (uconv)
4×4 (uconv)
4×4 (uconv)
k × k (conv)
k × k (uconv)
4×4 (conv)
4×4 (conv)
4×4 (conv)
4×4 (conv)
(flatten + dropout 50%)
32768
(dense)
real or fake
Adversarial Discriminator

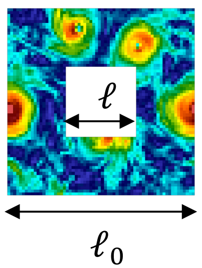
Point-wise Mean Squared Error

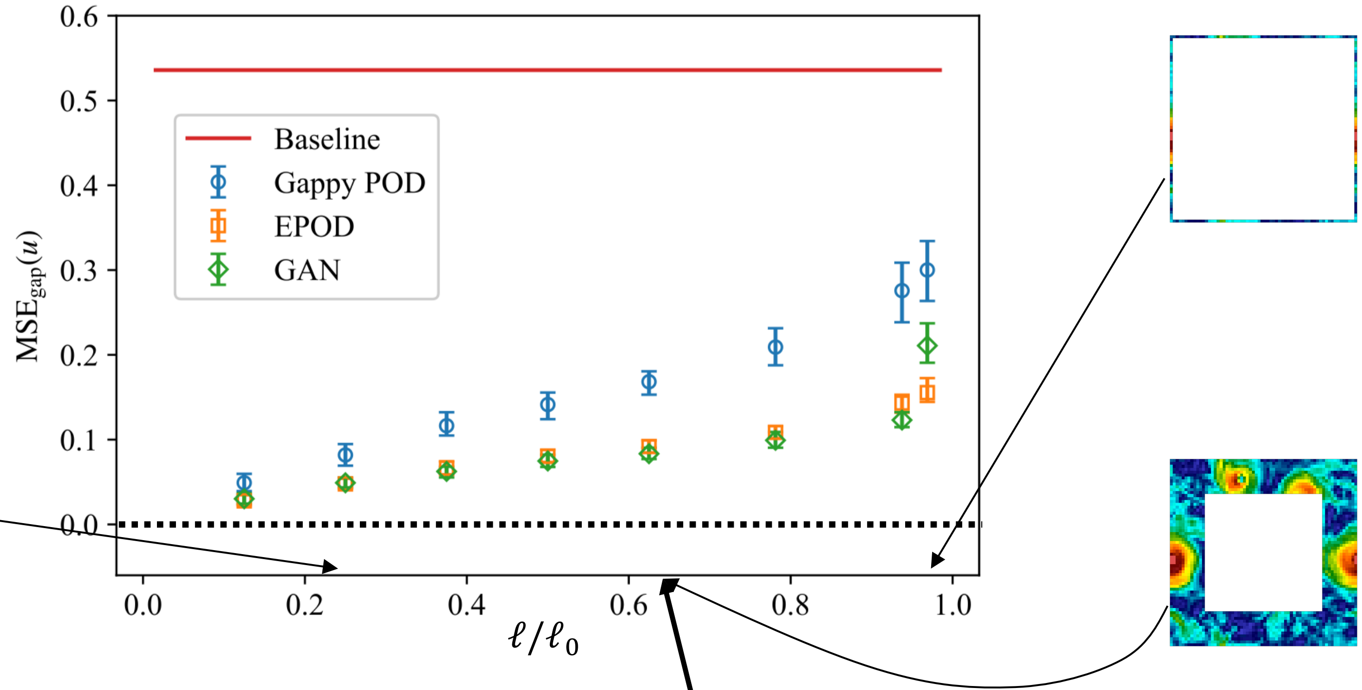
$$MSE_{gap} = \frac{1}{\mathrm{E}_k} \int_G \left(u_{true}(\boldsymbol{x}) - u_{pred}(\boldsymbol{x})\right)^2 d\boldsymbol{x}$$

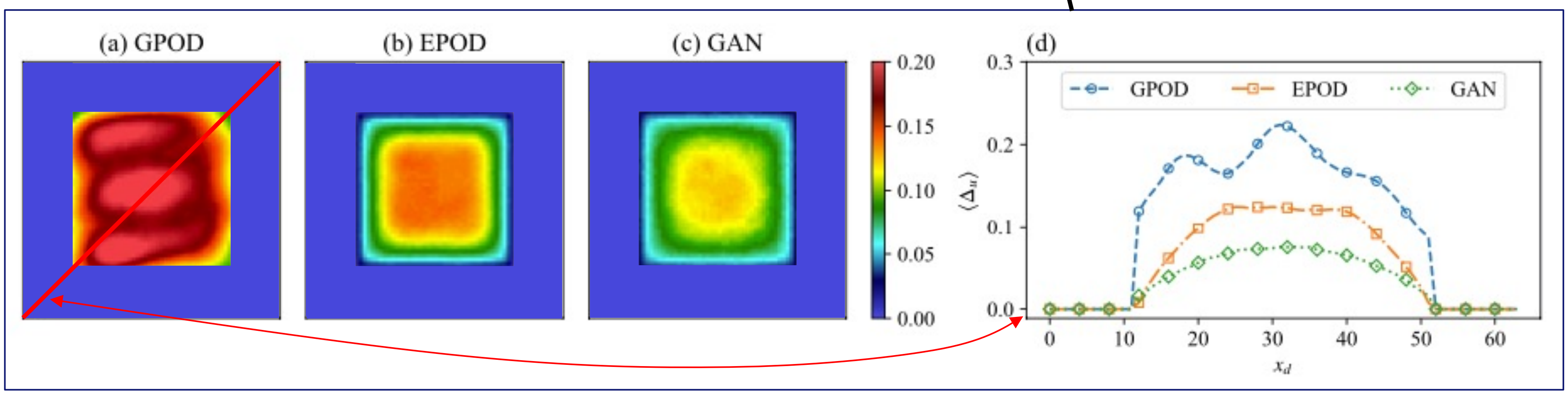


Point-wise Mean Squared Error

$$MSE_{gap} = \frac{1}{\mathrm{E}_k} \int_G \left(u_{true}(\boldsymbol{x}) - u_{pred}(\boldsymbol{x})\right)^2 d\boldsymbol{x}$$

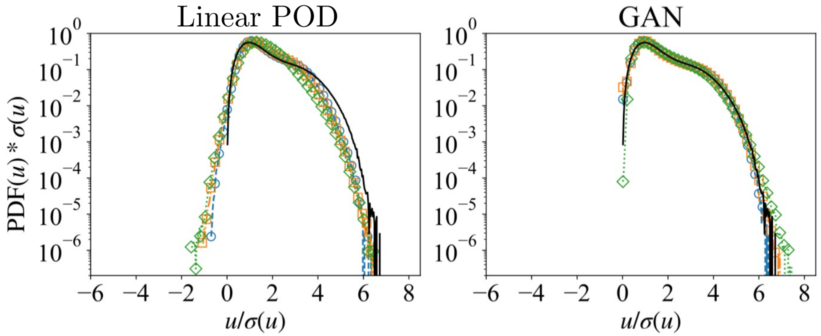




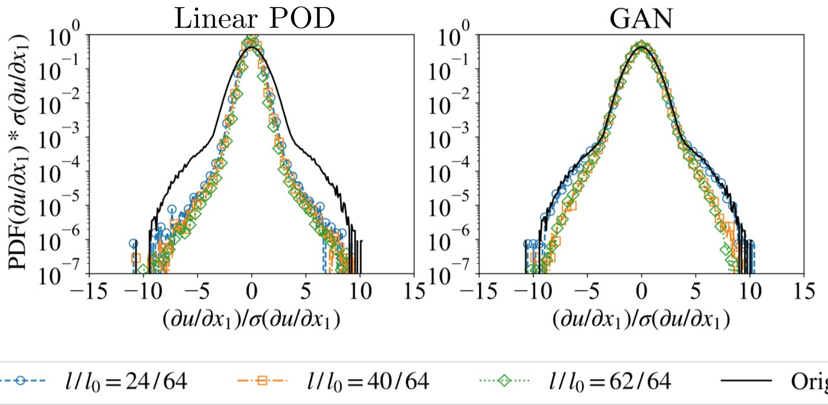


Statistical evaluation of reconstructed data; from mean value to extreme-events

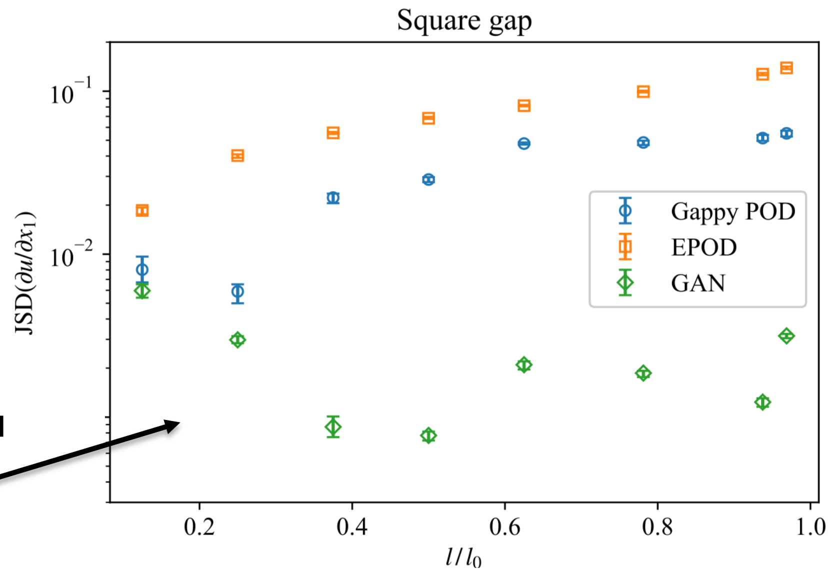
PDF of velocity module



PDF of velocity gradients



JENSEN-SHANNON Divergence



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

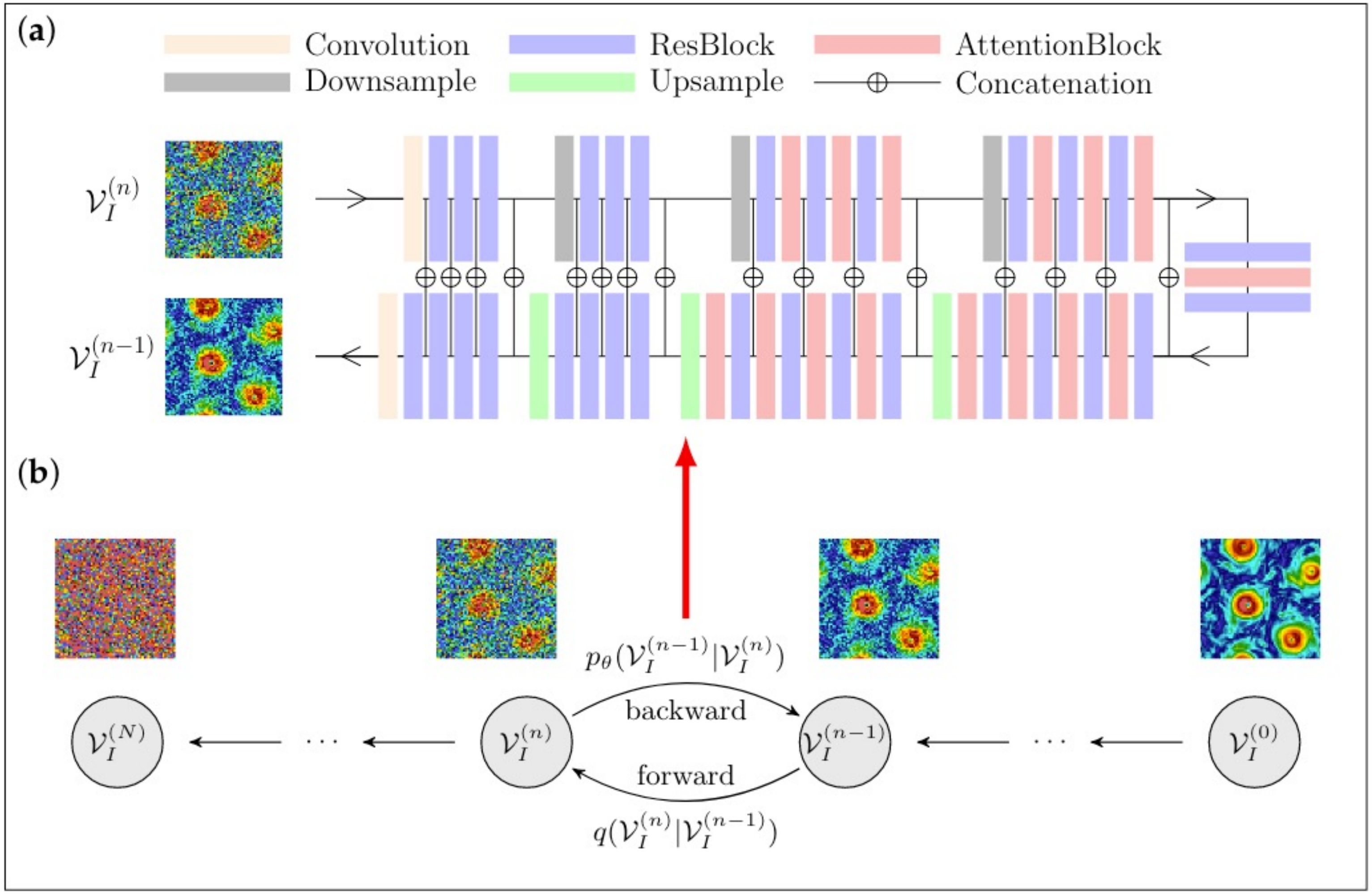
$$D(P \parallel Q) = \int_{-\infty}^{\infty} P(x) \log\left(\frac{P(x)}{Q(x)}\right) \mathrm{d}x$$

$$M = \tfrac{1}{2}(P + Q)$$

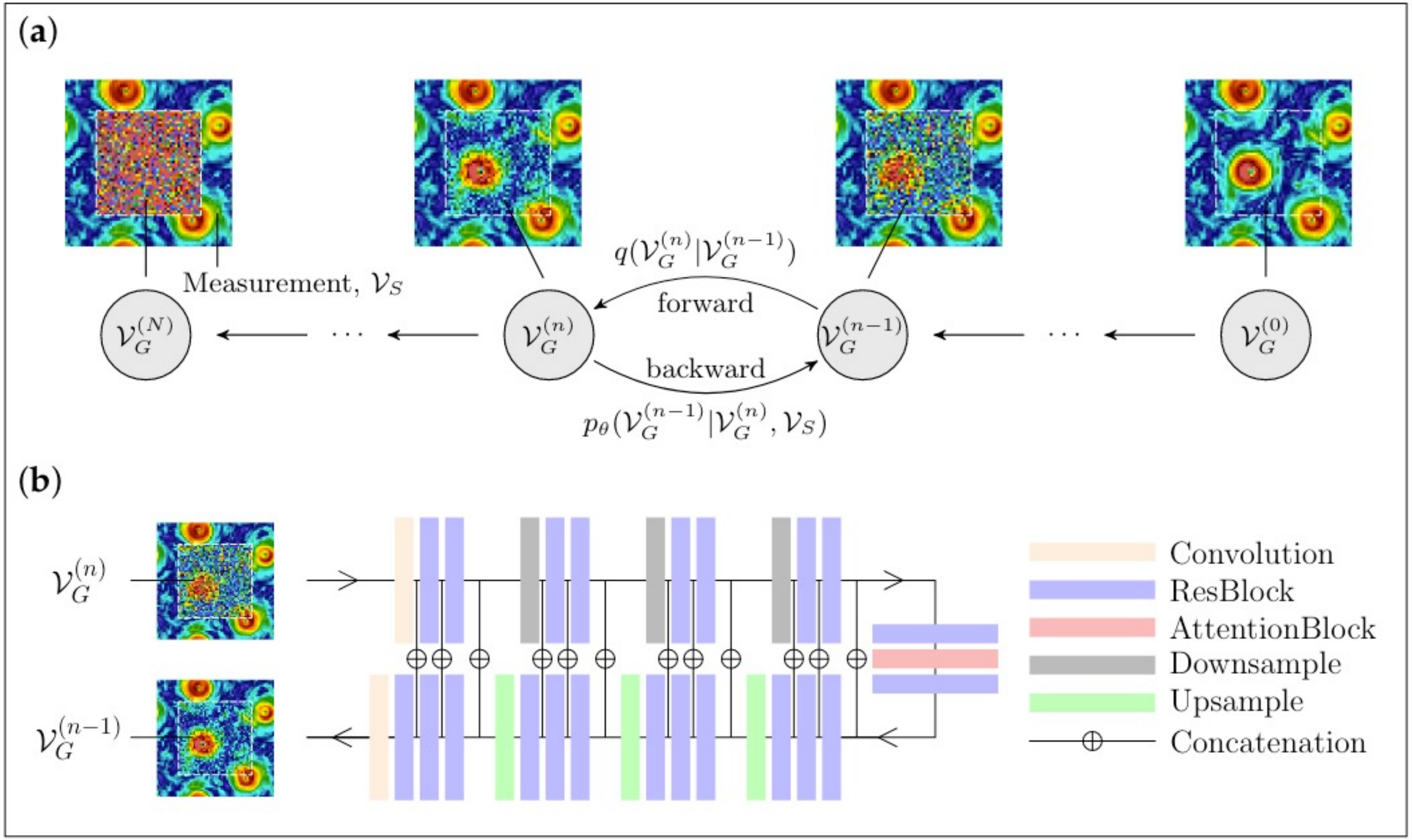
KULLBACK-LEIBLER Divergence

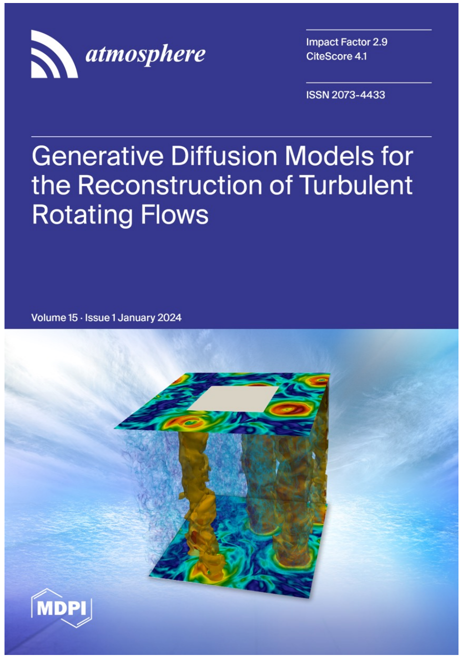
Reconstruction of turbulent data with deep generative models for semantic inpainting from TURB-Rot database
M. Buzzicotti, F. Bonaccorso, P. Clark Di Leoni, and L. Biferale, Phys. Rev. Fluids **6**, 050503, 2021

Diffusion Models to Generate turbulent flows

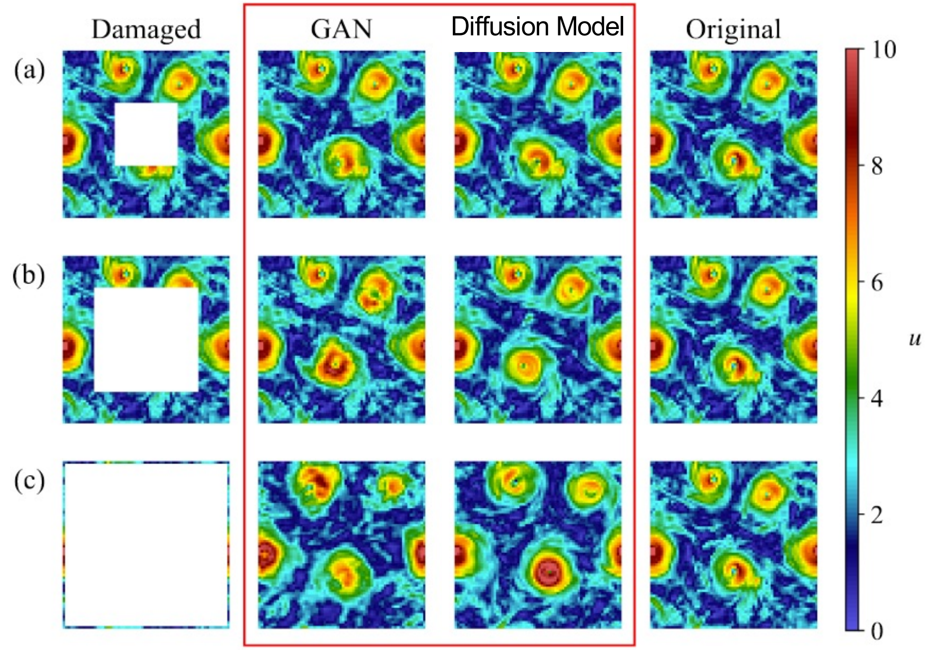


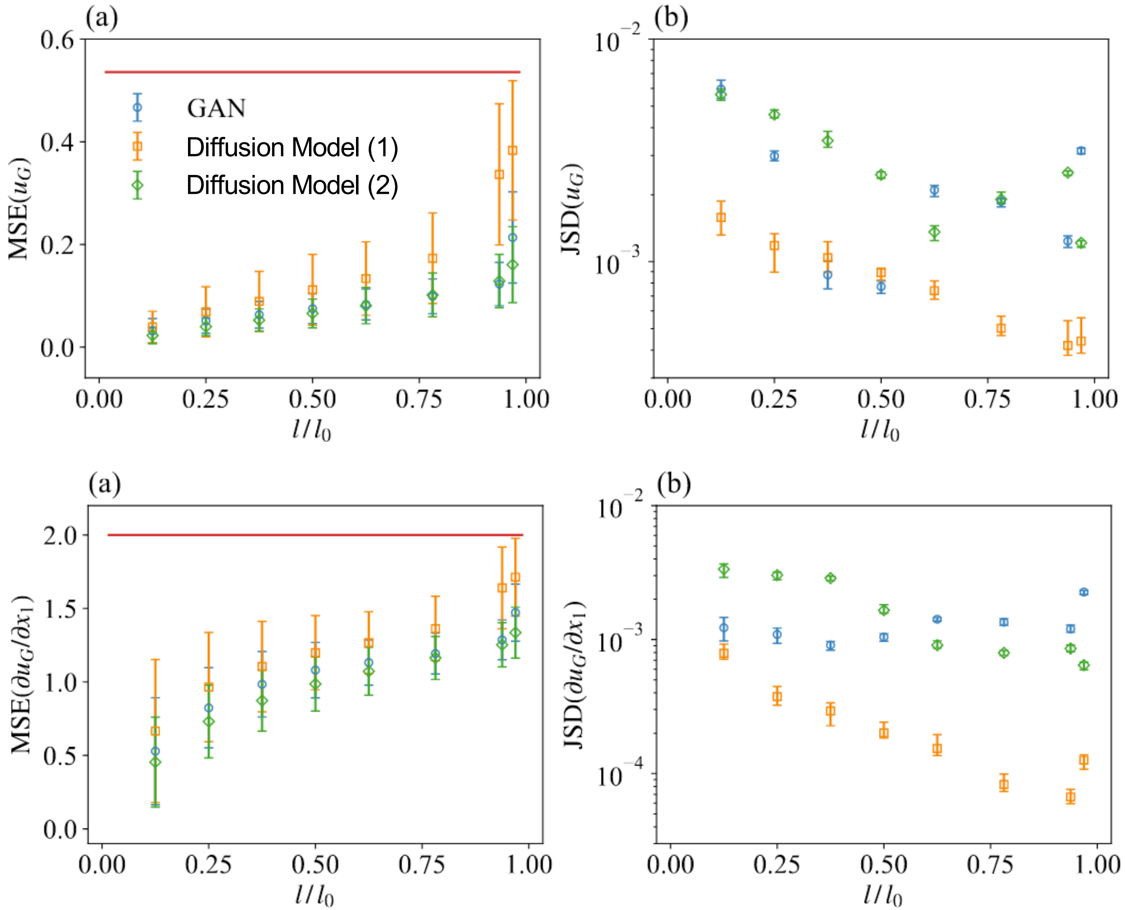
Diffusion Models to Reconstruct [Conditioned Generation]





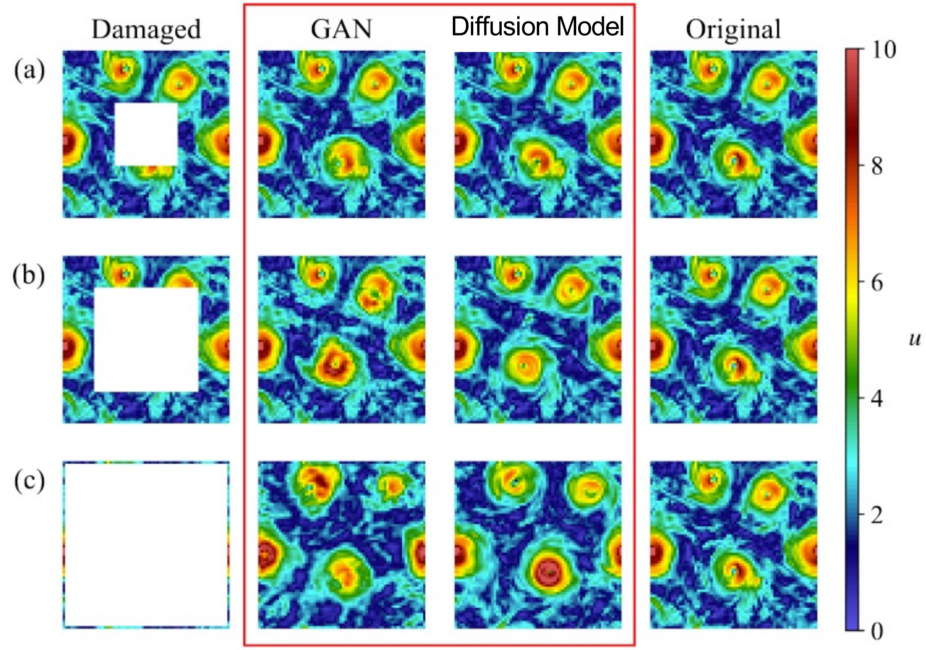
Diffusion Models vs GANs

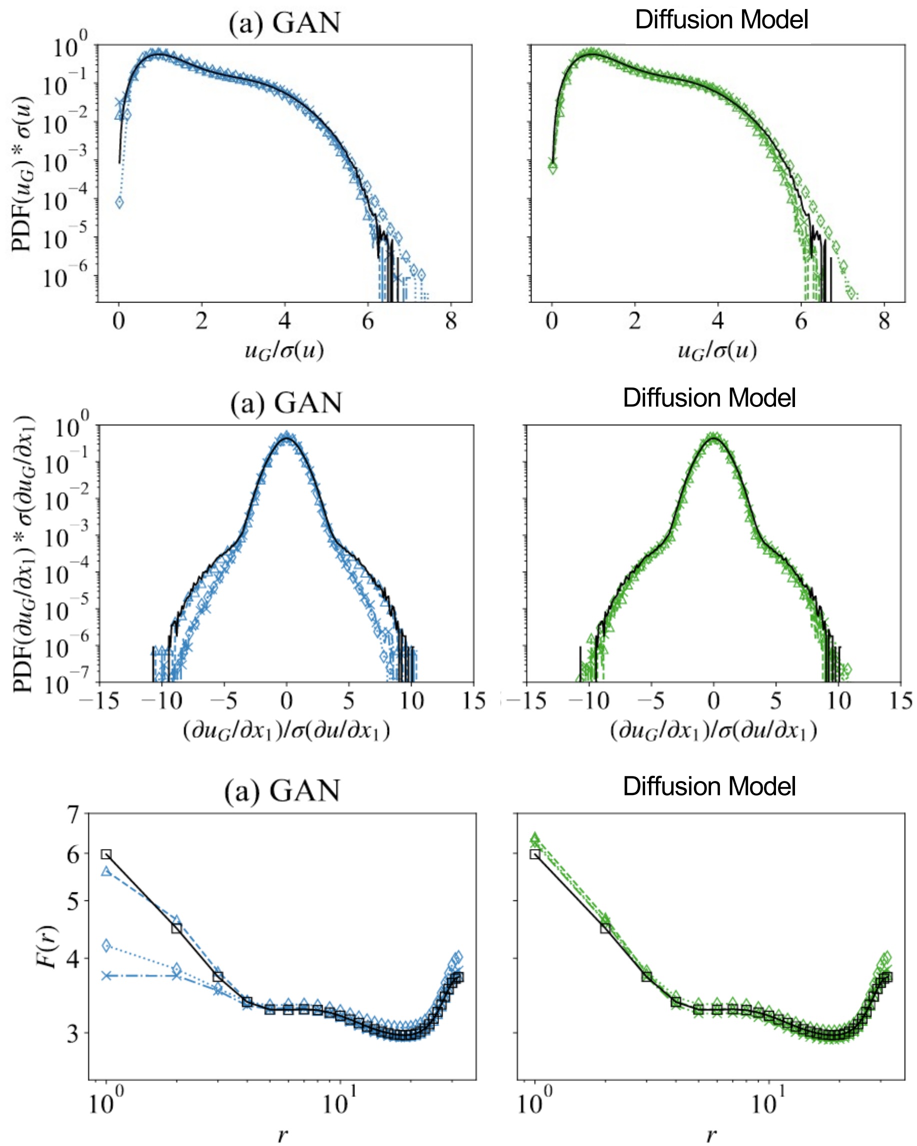




Li, T., Lanotte, A. S., Buzzicotti, M., Bonaccorso, F., & Biferale, L. (2023). **Multi-Scale Reconstruction of Turbulent Rotating Flows with Generative Diffusion Models**. *Atmosphere*, *15*(1), 60.

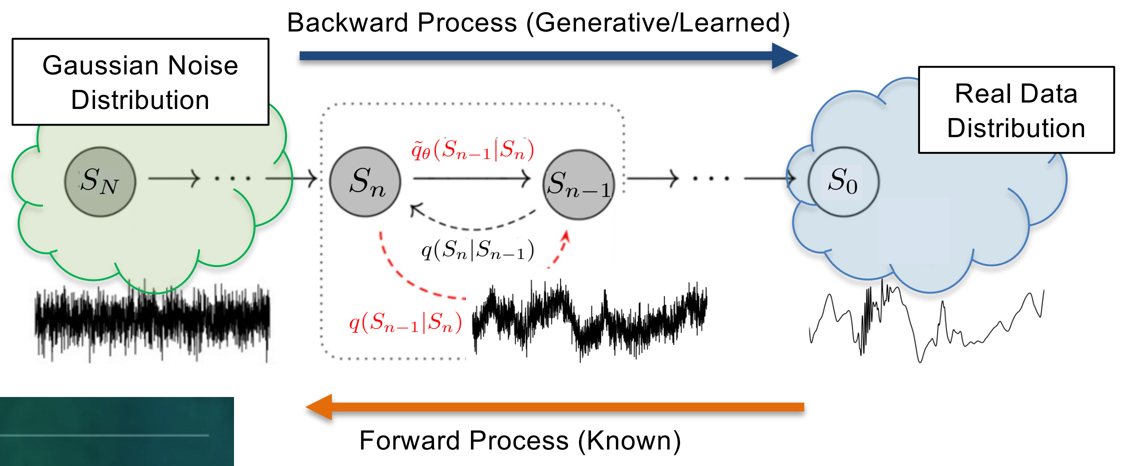
Diffusion Models vs GANs





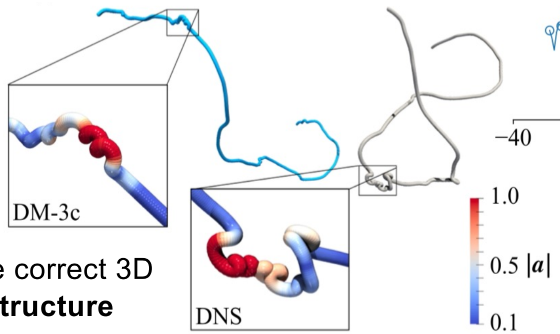
Li, T., Lanotte, A. S., Buzzicotti, M., Bonaccorso, F., & Biferale, L. (2023). **Multi-Scale Reconstruction of Turbulent Rotating Flows with Generative Diffusion Models**. *Atmosphere*, *15*(1), 60.

Synthetic Lagrangian Turbulence by Generative Diffusion Models

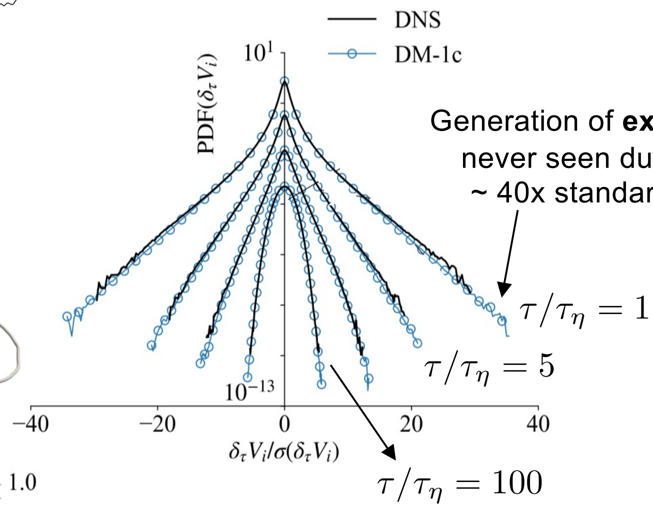


First Application of DM for the generation of turbulent data

Lagrangian particle trajectories advected by 3D fully developed Turbulent flows



Generation of the correct 3D **topological structure**



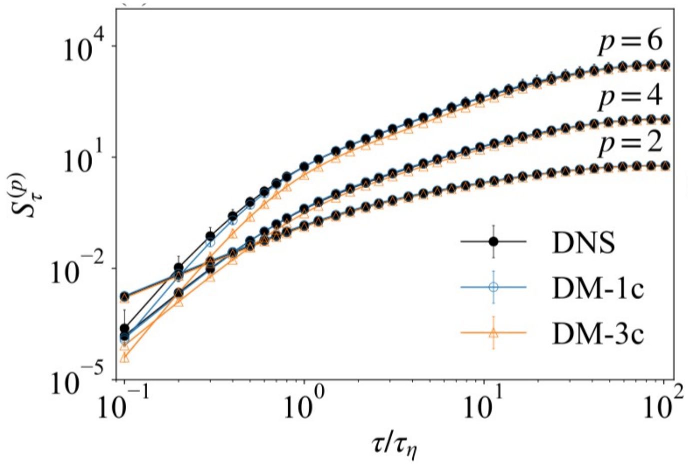
Generation of **extreme events** never seen during training ~ 40x standard deviation

$$\boxed{\delta_\tau V_i(t) = V_i(t+\tau) - V_i(t)}$$

T. Li, L. Biferale, F. Bonaccorso, M. A. Scarpolini, and M. Buzzicotti, *Nature Machine Intelligence* (2024): 1-11.

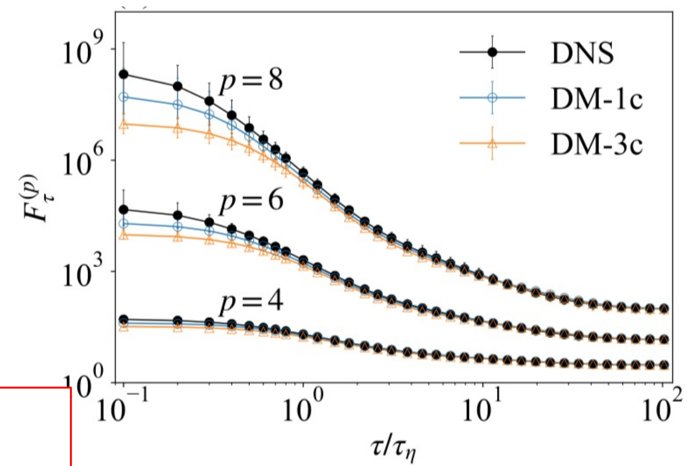
LAGRANGIAN STRUCTURE FUNCTIONS

$$S_\tau^{(p)} = \langle (\delta_\tau V_i)^p \rangle$$

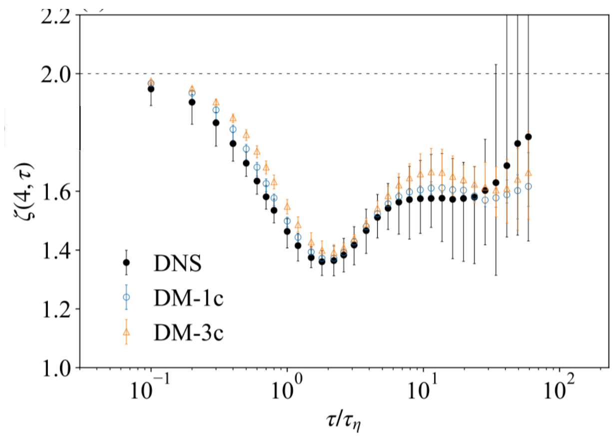


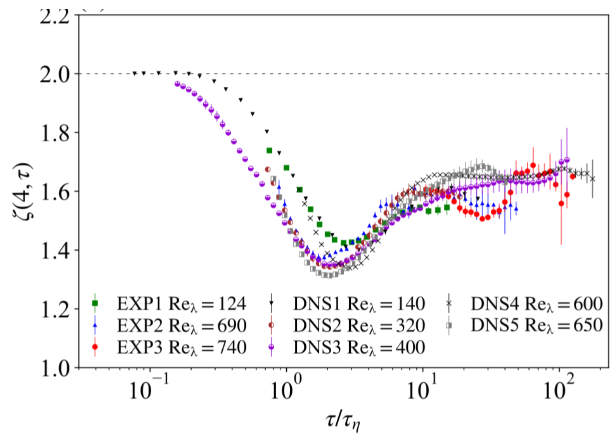
GENERALIZED FLATNESS

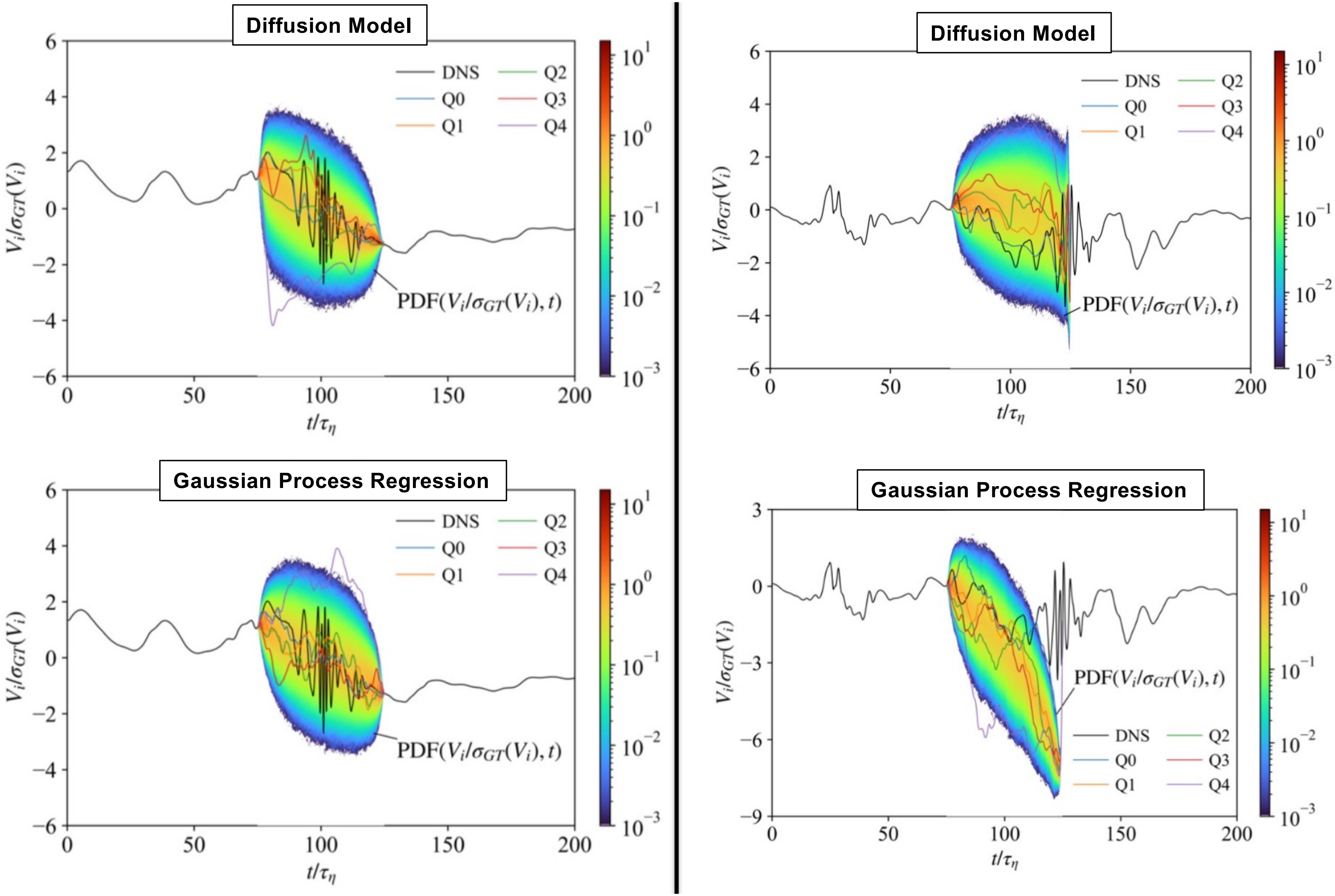
$$F_\tau^{(p)} = S_\tau^{(p)} / [S_\tau^{(2)}]^{p/2}.$$



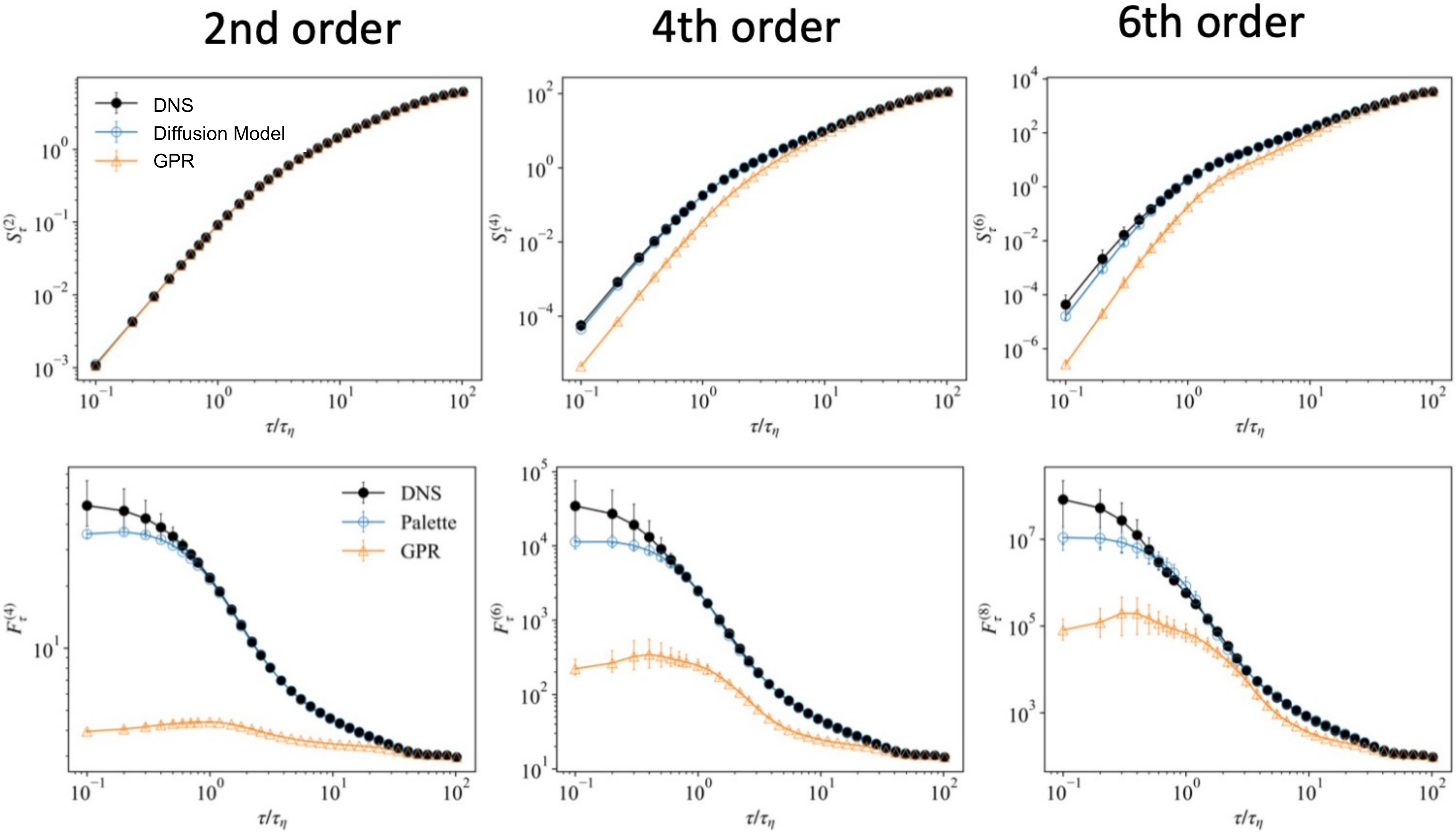
$$\boxed{\zeta(p,\tau) = \frac{d \log S_\tau^{(p)}}{d \log S_\tau^{(2)}}.}$$







Diffusion Model
Gaussian Process Regression
DNS
Q0
Q1
Q2
Q3
Q4
PDF($V_i/\sigma_{GT}(V_i), t$)
$V_i/\sigma_{GT}(V_i)$
t/τ_η
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}


2nd order
4th order
6th order
DNS
Diffusion Model
GPR
DNS
Palette
GPR
$S_\tau^{(2)}$
$S_\tau^{(4)}$
$S_\tau^{(6)}$
$F_\tau^{(4)}$
$F_\tau^{(6)}$
$F_\tau^{(8)}$
τ/τ_η

Hands on Session

Implementation and training of a VAE and a GAN for the generation of turbulent flows on a rotating reference frame.

Smart-Turb ∞

Datasets ▾ Organizations ▾ Help ▾ About ▾ U

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 dataset for big-d... Machine Learning approaches to fluid mechanics in general and turb... easily accessible web platform for high quality data. It... dardize and manage a large collection of heterog... numerical data sets from high-end fluid dynamics experimental... gh Performance Computational centers. Smart-TURB offers ...ances when accessing/uploading/searching data. The whole ...d to contribute, by deploying freely downloadable, accurate and ...set for the sake of "reproducibility": The process of documenting ...chiving data so that others can fully reproduce scientific results. ...e administrator for infos about how to upload your dataset. We ... a first dataset made of 2d and 3d turbulent configurations under ... 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

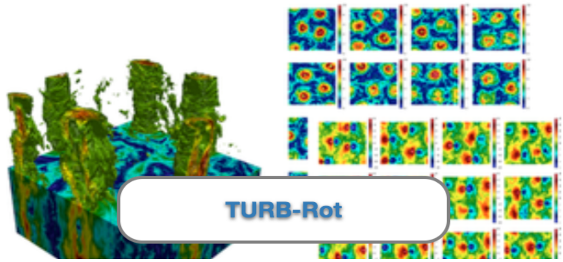
M. Buzzicotti
Dept. Physics and INFN
University of Rome Tor Vergata, Italy.
michele.buzzicotti@roma2.infn.it

P. Clark Di Leoni
Department of Mechanical Engineering,
Johns Hopkins University, Baltimore, USA.
pato@jhu.edu

Search for datasets

1 Datasets

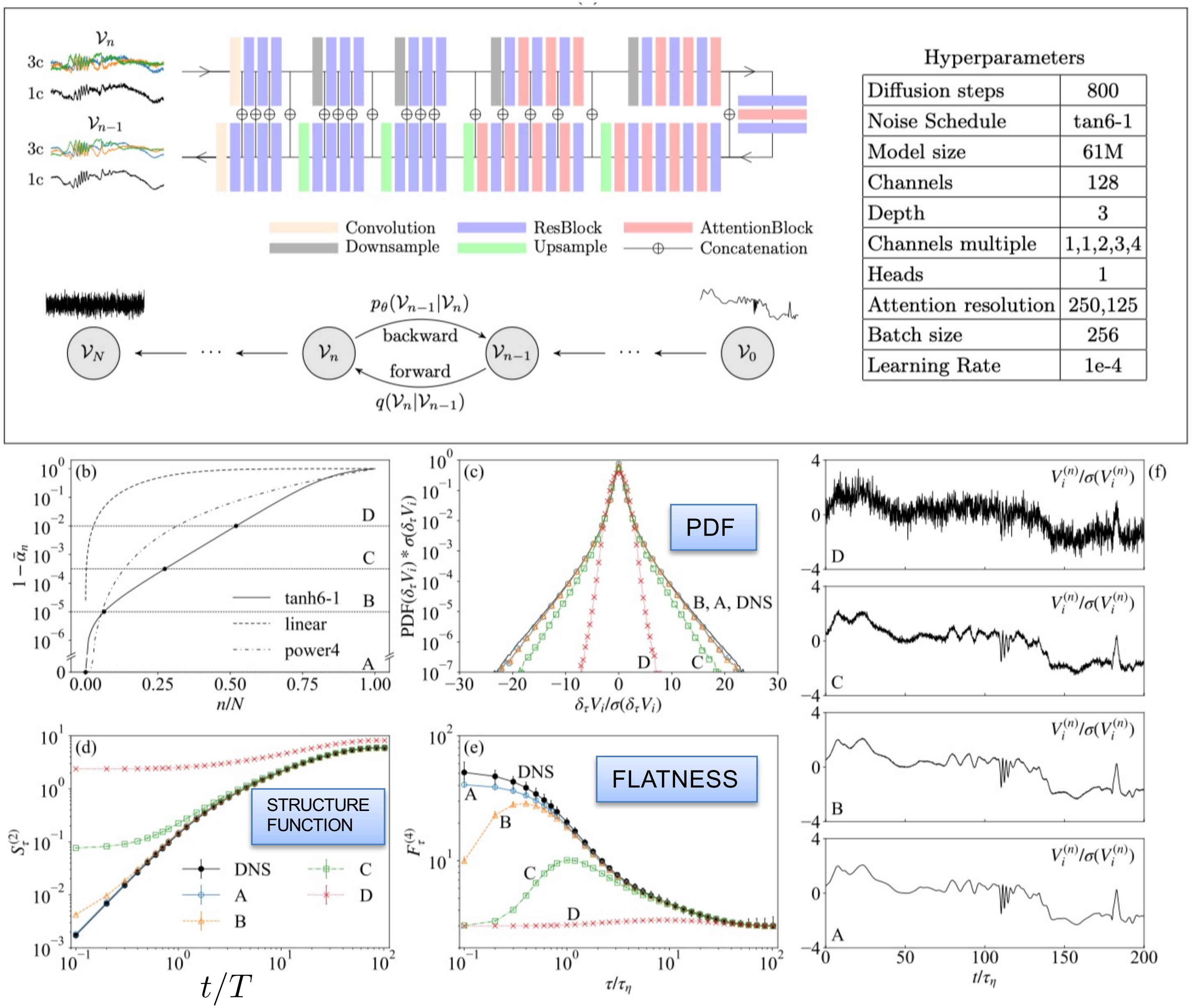
TURB-Rot
A large database of 3d and 2d snapshots from turbulent rotating



2 Organizations

web_admin web_admin group	1 member

michele.buzzicotti@roma2.infn.it



$\mathcal{V}_n$
3c
1c
$\mathcal{V}_{n-1}$
3c
1c
Convolution
Downsample
ResBlock
Upsample
AttentionBlock
Concatenation
Hyperparameters
Diffusion steps 800
Noise Schedule tan6-1
Model size 61M
Channels 128
Depth 3
Channels multiple 1,1,2,3,4
Heads 1
Attention resolution 250,125
Batch size 256
Learning Rate 1e-4
$p_\theta(\mathcal{V}_{n-1}|\mathcal{V}_n)$
backward
forward
$q(\mathcal{V}_n|\mathcal{V}_{n-1})$
$\mathcal{V}_N$
$\mathcal{V}_n$
$\mathcal{V}_{n-1}$
$\mathcal{V}_0$
(b)
$1-\bar{\alpha}_n$
tanh6-1
linear
power4
n/N
(c)
PDF
$\mathrm{PDF}(\delta_\tau V_i) * \sigma(\delta_\tau V_i)$
B, A, DNS
$\delta_\tau V_i/\sigma(\delta_\tau V_i)$
(f)
$V_i^{(n)}/\sigma(V_i^{(n)})$
t/τ_η
(d)
STRUCTURE FUNCTION
$S_\tau^{(2)}$
DNS
t/T
(e)
FLATNESS
$F_\tau^{(4)}$
τ/τ_η

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

EQ. BASED

Nudging Simulation:

DNS with the addition of a drag term against partial field measurements

$$\partial_t \boldsymbol{v} + \boldsymbol{v} \cdot \partial_x \boldsymbol{v} + \partial_x P - \nu \Delta \boldsymbol{v} = 2\, \boldsymbol{v} \times \boldsymbol{\Omega} + \mathbf{f} \; - N(\boldsymbol{v}_{\mathbf{N}} - \boldsymbol{v})$$