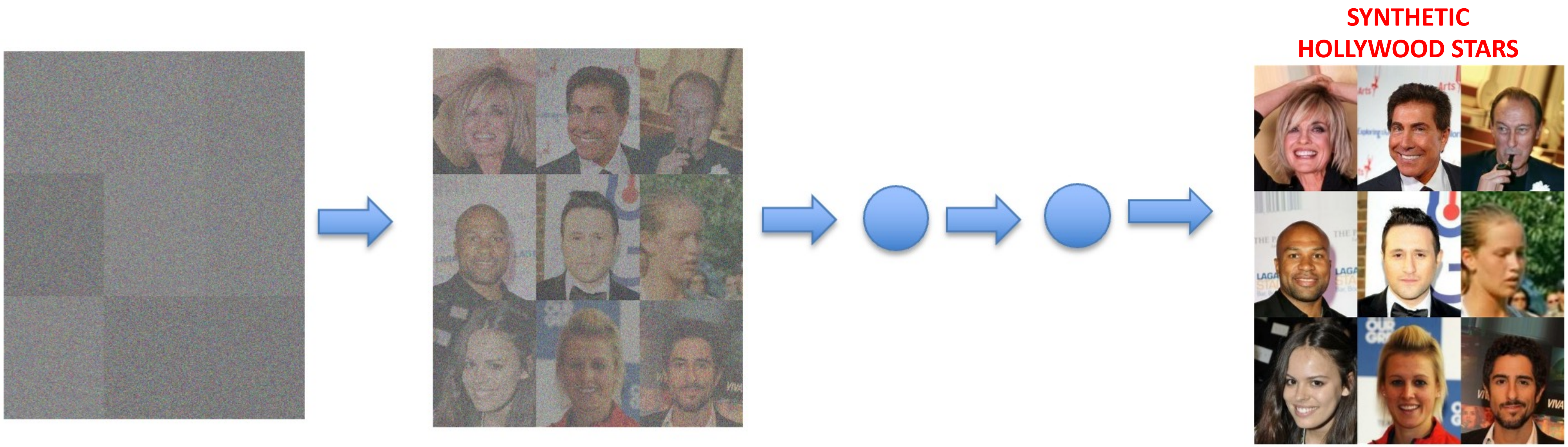
Complex Fluids and Complex Flows Group
Dept. Physics & INFN - University of Rome 'Tor Vergata'
biferale@roma2.infn.it
https://biferale.web.roma2.infn.it/



ERICE – MACHINE LEARNING FOR COMPLEXITY
2024 APRIL
Data driven tools for Eulerian and Lagrangian Turbulence

CREDITS: M. CENCINI, C. CALASCIBETTA, L. PIRO, R. HEINONEN, T. LI, F. BONACCORSO, M. BUZZICOTTI, M. SCARPOLINI

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani Piano Nazionale di Ripresa e Resilienza

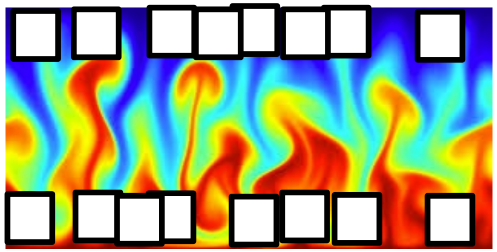
1. Short/personal introduction to Eulerian and Lagrangian Turbulence in 2D, 3D and in between
2. Data-driven and Equation-Informed tools for Eulerian Turbulence
3. Data-driven and Equation-Informed tools for Lagrangian Turbulence

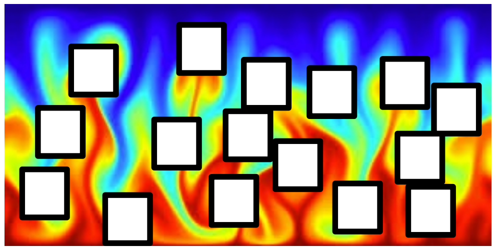
2.1 INPANTING GAPS IN REAL SPACE
2.2 INFERENCE/CLASSIFICATIONS OF FLOW PROPERTIES
2.3 SUB-FILTER MODELLING

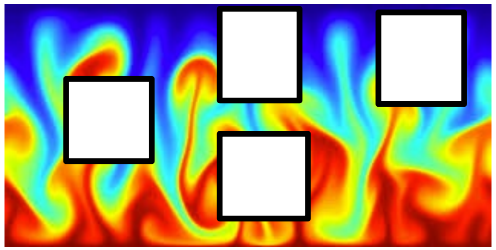
1. RECONSTRUCTION OF MISSING INFORMATION (INPAINTING – SUPER RESOLUTION)
2. FEATURES RANKING: QUALITY AND QUANTITY OF DATA

- IS IT BETTER TO INPUT SPATIAL OR TEMPORAL DATA?
- HOW MANY DATA/VARIABLES YOU NEED TO SUPPLY FOR PERFECT RECONSTRUCTION (SYNCHRONIZATION-TO-DATA)?
- CAN YOU GUESS VELOCITY FIELDS BY MEASURING ONLY TEMPERATURE AND/OR VICEVERSA?
- IS IT BETTER TO PROVIDE INFORMATION FROM BOUNDARIES OR BULK?
- FROM LARGE OR SMALL SCALES?
- DO WE NEED TO KNOW THE EQUATIONS?
- HOW TO COMPARE EQUATIONS-BASED AND EQUATIONS-FREE MODELS?

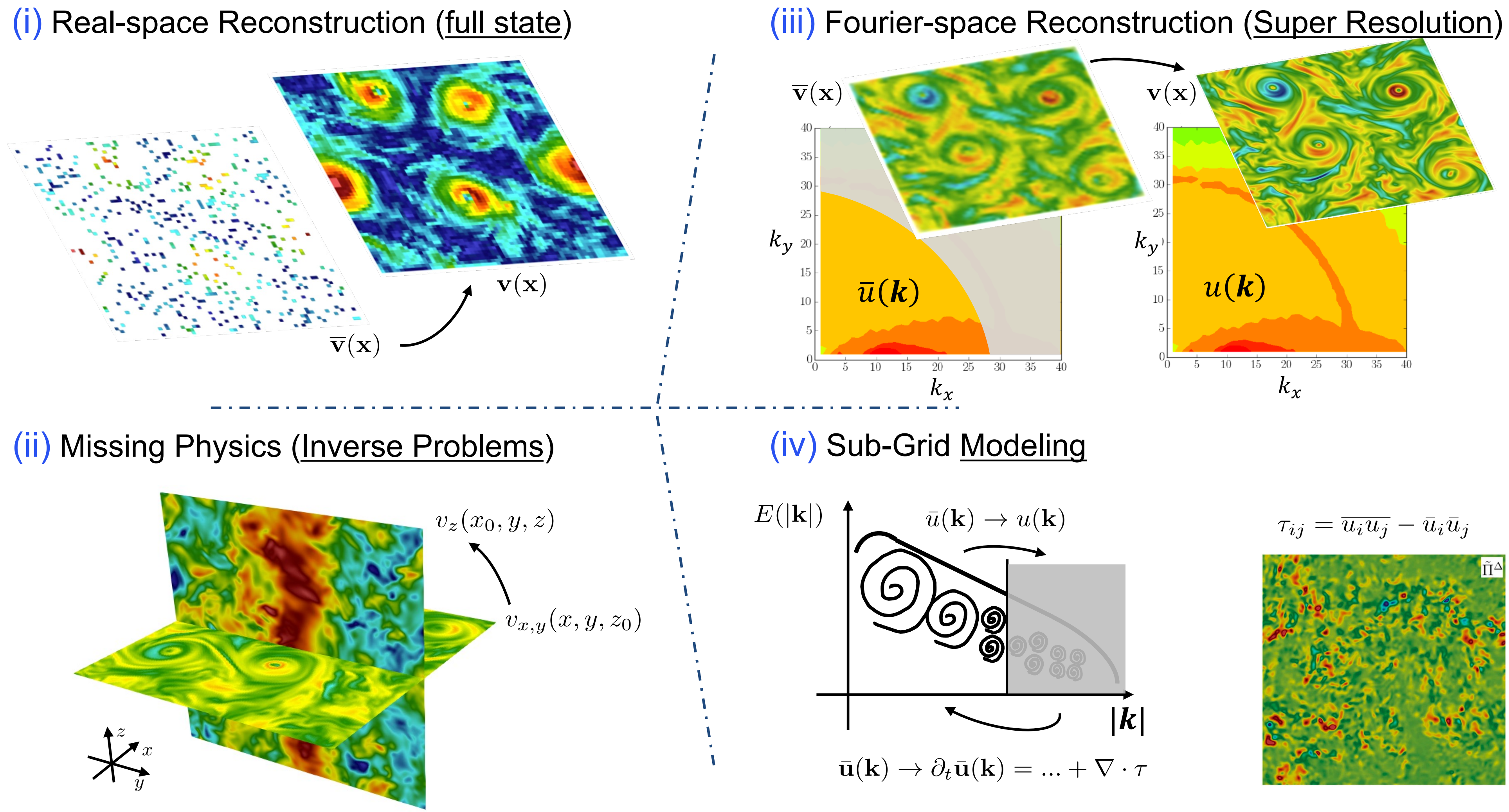
A WAY TO LEARN ABOUT THE UNDERLYING PHYSICS



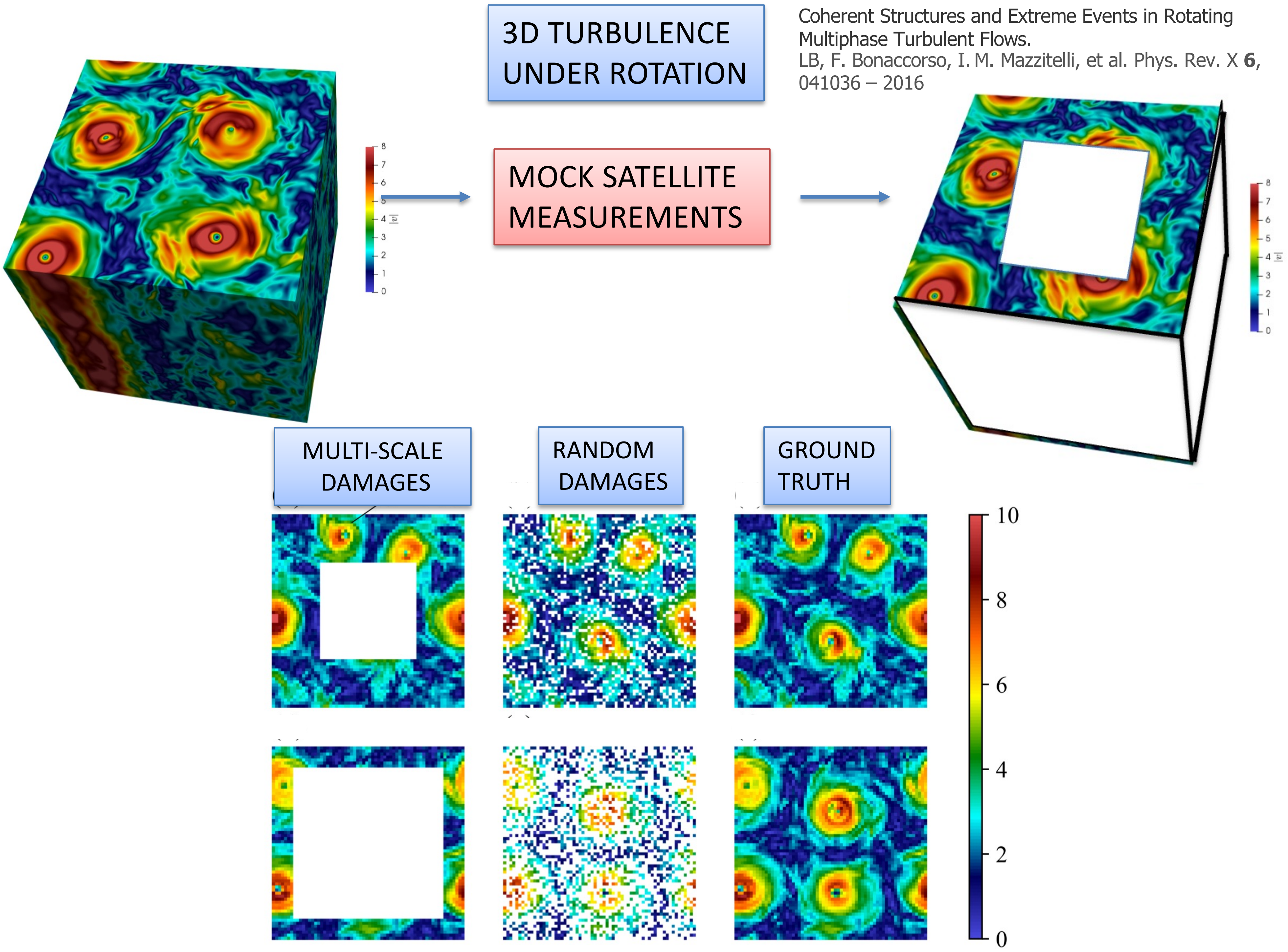




1. RECONSTRUCTION OF MISSING INFORMATION
2. FEATURES RANKING: QUALITY AND QUANTITY OF DATA

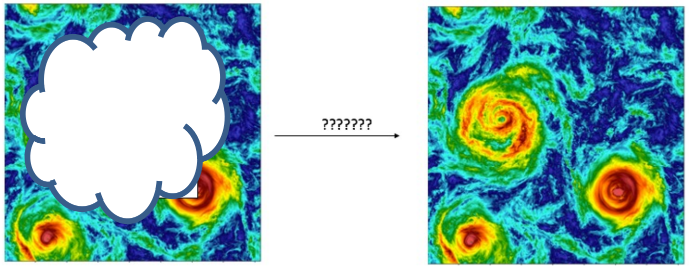


M. Buzzicotti. "Data reconstruction for complex flows using AI: recent progress, obstacles, and perspectives." Europhysics Letters, EPL 142 23001 (2023).



3D TURBULENCE UNDER ROTATION
Coherent Structures and Extreme Events in Rotating Multiphase Turbulent Flows.
LB, F. Bonaccorso, I. M. Mazzitelli, et al. Phys. Rev. X **6**, 041036 – 2016
MOCK SATELLITE MEASUREMENTS
MULTI-SCALE DAMAGES
RANDOM DAMAGES
GROUND TRUTH
10
8
6
4
2
0

DATA ASSIMILATION, FLOW RECONSTRUCTION, INPAINTING, SUPER-RESOLUTION, PHYSICS-INFERRING (CLASSIFICATION)

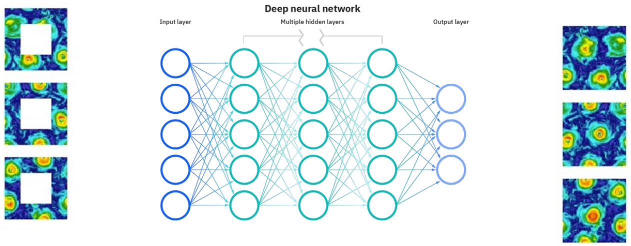


DATA DRIVEN

$$u(\boldsymbol{x}) = \sum_{n=1}^{N} a_n \psi_n(\boldsymbol{x}) = \sum_{n=1}^{N'} a_n \psi_n(\boldsymbol{x}) + \sum_{n=N'+1}^{N} a_n \psi_n(\boldsymbol{x})$$

$$\int_S \left[u(\boldsymbol{x}) - \sum_{n=1}^{N'} a_n \psi_n(\boldsymbol{x}) \right]^2 \mathrm{d}\boldsymbol{x}$$

1. EQUATION FREE
PRINCIPAL ORTHOGONAL DECOMPOSITION
GAPPY-POD & EXTENDED POD



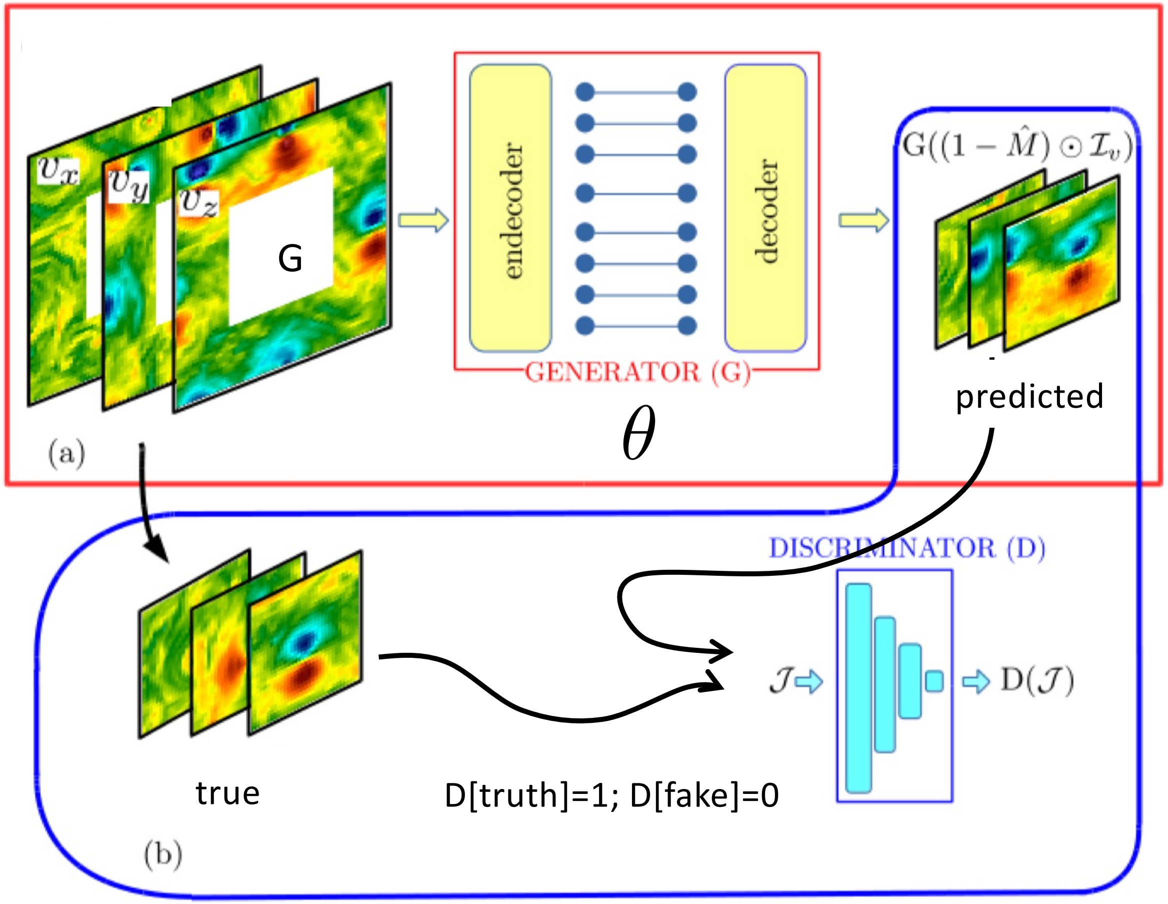
2. EQUATION FREE
GENERATIVE-ADVERSARIAL-NETWORK

EQ. BASED

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

3. NUDGING

GENERATIVE ADVERSARIAL NETWORK: CONTEXT ENCODER (ACTOR-CRITIC)



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:

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

$$\mathcal{L}_{adv} = \log\left(1 - D\left(u_{\text{pred}}\right)\right).$$

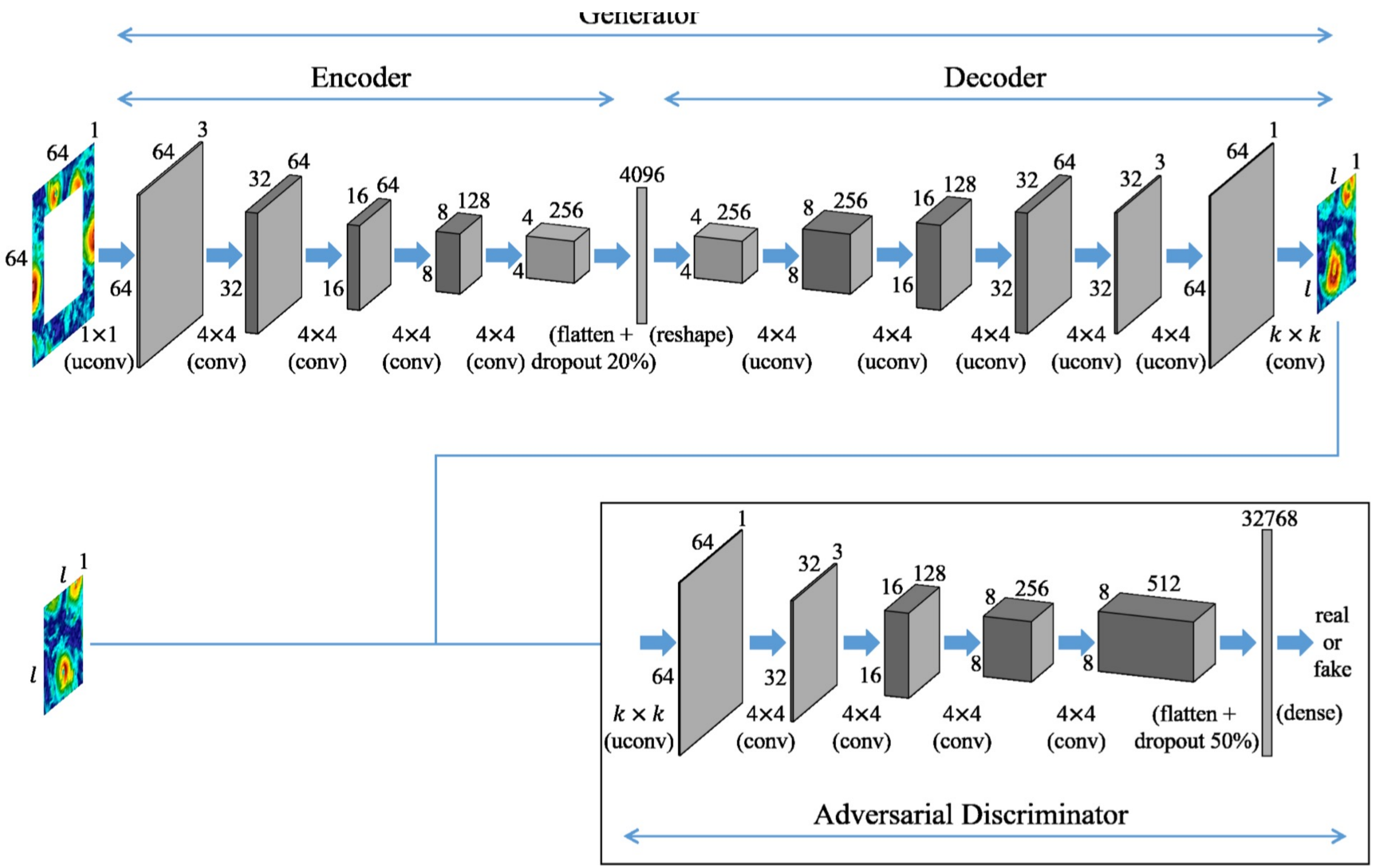
$$\mathcal{L}_{TOT} = \mathcal{L}_G + \lambda \mathcal{L}_{ADV}$$

MAXIMIZE:

$$\mathcal{L}_{DIS} = \log\left(D\left(u_{\text{true}}\right)\right) + \log\left(1 - D\left(u_{\text{pred}}\right)\right).$$

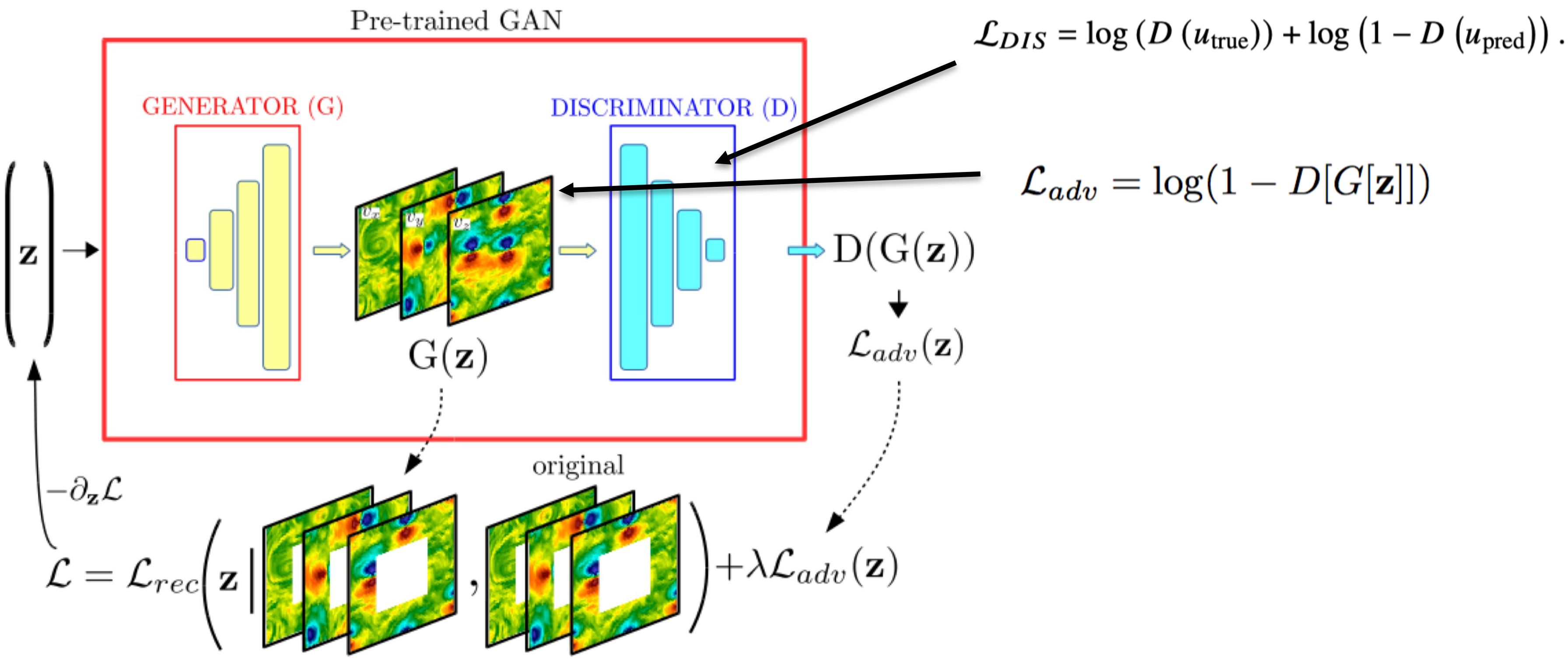
DATA DRIVEN NO-EQUATIONS

[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.



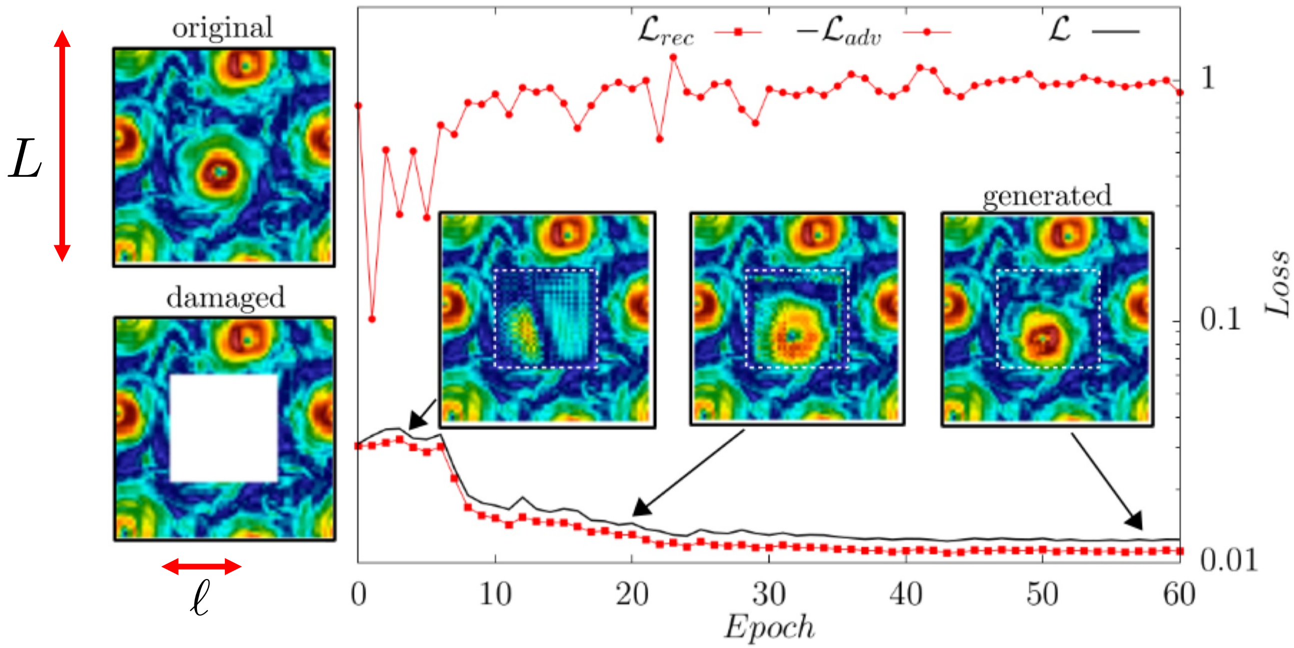
Generator
Encoder
Decoder
64
64
1
1×1
(uconv)
64
64
3
4×4
(conv)
32
32
64
4×4
(conv)
16
16
64
4×4
(conv)
8
8
128
4×4
(conv)
4
4
256
(flatten +
dropout 20%)
4096
(reshape)
4
4
256
4×4
(uconv)
8
8
256
4×4
(uconv)
16
16
128
4×4
(uconv)
32
32
64
4×4
(uconv)
32
32
3
4×4
(uconv)
64
64
1
k × k
(conv)
l
l
1
l
l
1
64
64
1
k × k
(uconv)
32
32
3
4×4
(conv)
16
16
128
4×4
(conv)
8
8
256
4×4
(conv)
8
8
512
4×4
(conv)
(flatten +
dropout 50%)
32768
(dense)
real
or
fake
Adversarial Discriminator

CONTEXT ENCODER 2 (CE2)
CONSTRAINED IMAGE GENERATION



Raymond A Yeh, Chen Chen, Teck Yian Lim, Alexander G Schwing, Mark Hasegawa-Johnson, and Minh N Do. Semantic image inpainting with deep generative models. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5485–5493, 2017.

DURING TRAINING
80K 64x64 images of velocity amplitude for training
20K 64x64 images of velocity amplitude for validation



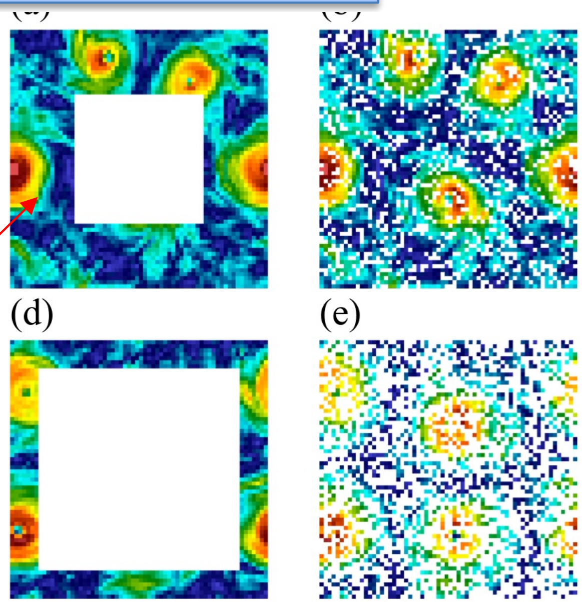
ℓ :much larger than differentiable scale, i.e. velocity fields are rough (no linear interpolation here)

GAPPY-POD (PRINCIPAL ORTHOGONAL DECOMPOSITION)

$$u_{pred}(\mathbf{x}) = \sum_{n=1}^{N} a_n \psi_n(\boldsymbol{x}) = \sum_{n=1}^{N'} a_n \psi_n(\boldsymbol{x}) + \sum_{n=N'+1}^{N} a_n \psi_n(\boldsymbol{x}),$$

$$K(\mathbf{x}, \mathbf{y}) = \langle u_{true}(\mathbf{x}) u_{true}(\mathbf{y}) \rangle$$

$$\int K(\mathbf{x}, \mathbf{y}) \psi_n(\mathbf{y}) d\mathbf{y} = \lambda_n \psi_n(\mathbf{x})$$

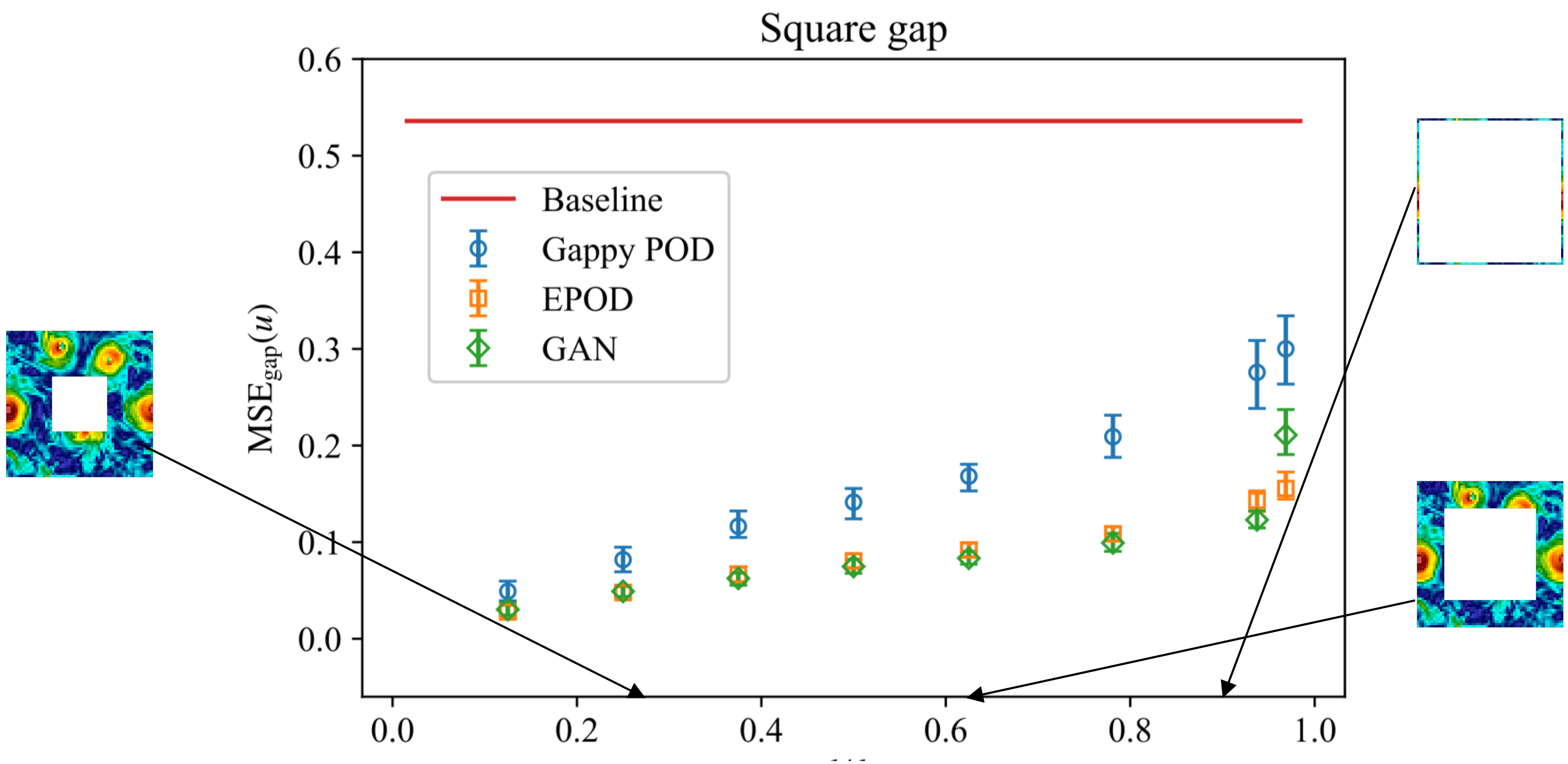


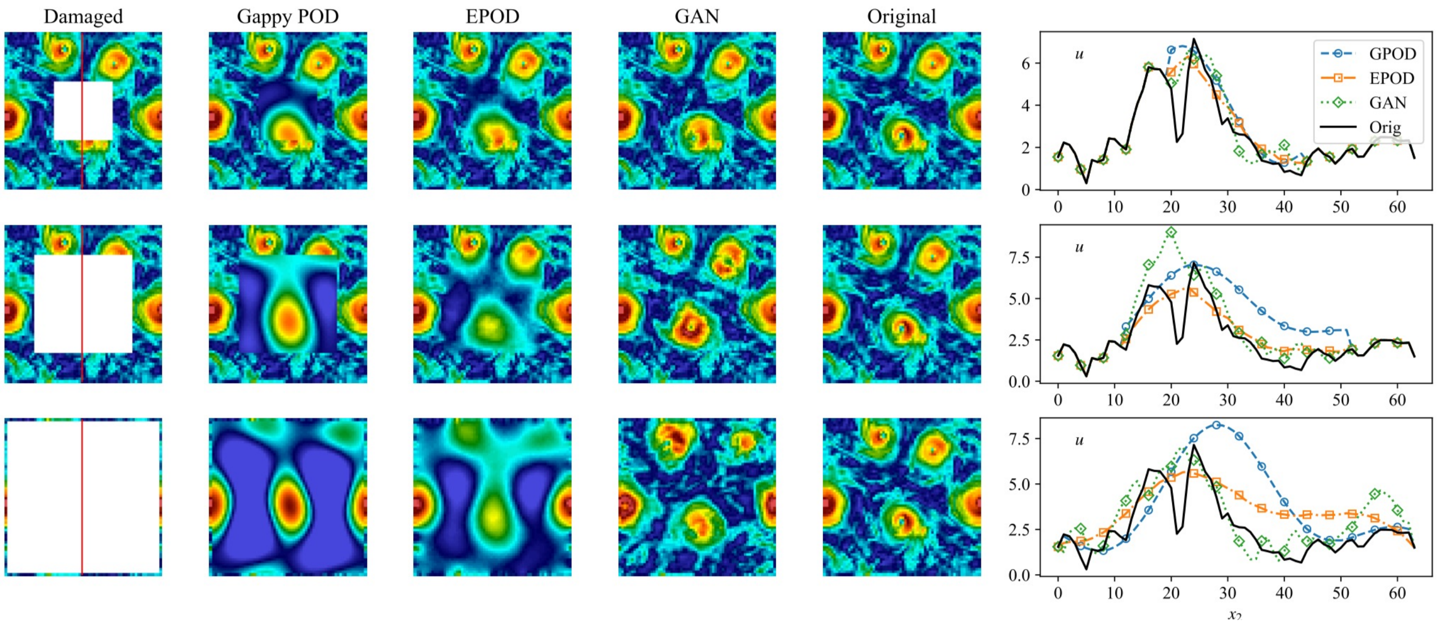
$$X' = \psi \oplus \psi$$

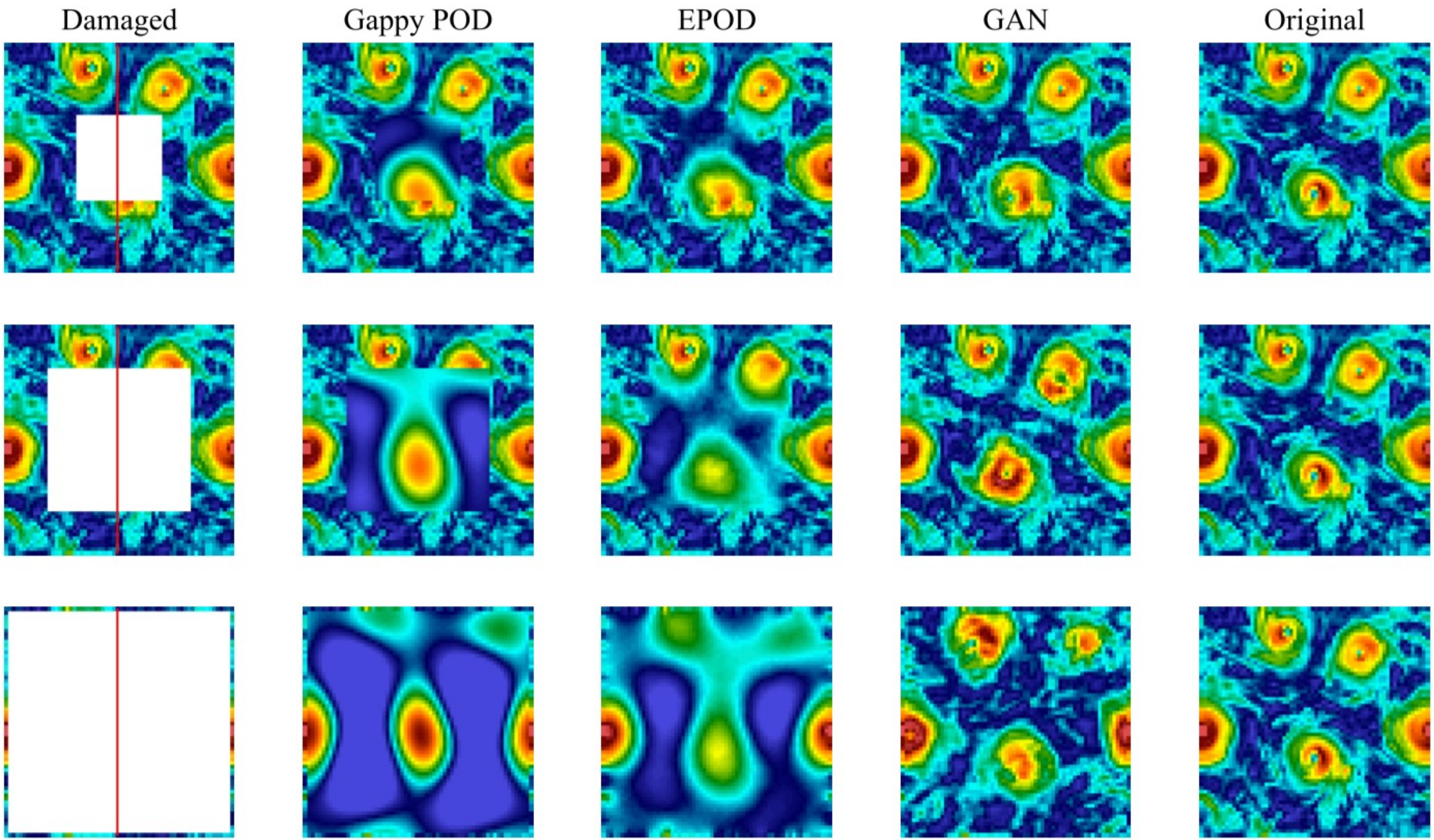
$$\tilde{\boldsymbol{E}} = \|\tilde{\boldsymbol{u}} - \tilde{\boldsymbol{X}}' \boldsymbol{a}'\| = \int_S d\mathbf{x} \left(u_{true}(\mathbf{x}) - \sum_{n=1}^{N'} a_n \psi_n(\mathbf{x})\right)^2$$

LINEAR OPTIMAL REGERESSION

T. Li, L. Buzzicotti, F. Bonaccorso, L.B., S. Chen, M. Wan. Data reconstruction of turbulent flows with Gappy-POD and Generative Adversarial Networks. In preparation 2022.


Square gap
Baseline
Gappy POD
EPOD
GAN
MSE_gap(u)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.0
0.2
0.4
0.6
0.8
1.0


Damaged
Gappy POD
EPOD
GAN
Original
u
GPOD
EPOD
GAN
Orig
0
10
20
30
40
50
60


Damaged
Gappy POD
EPOD
GAN
Original

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

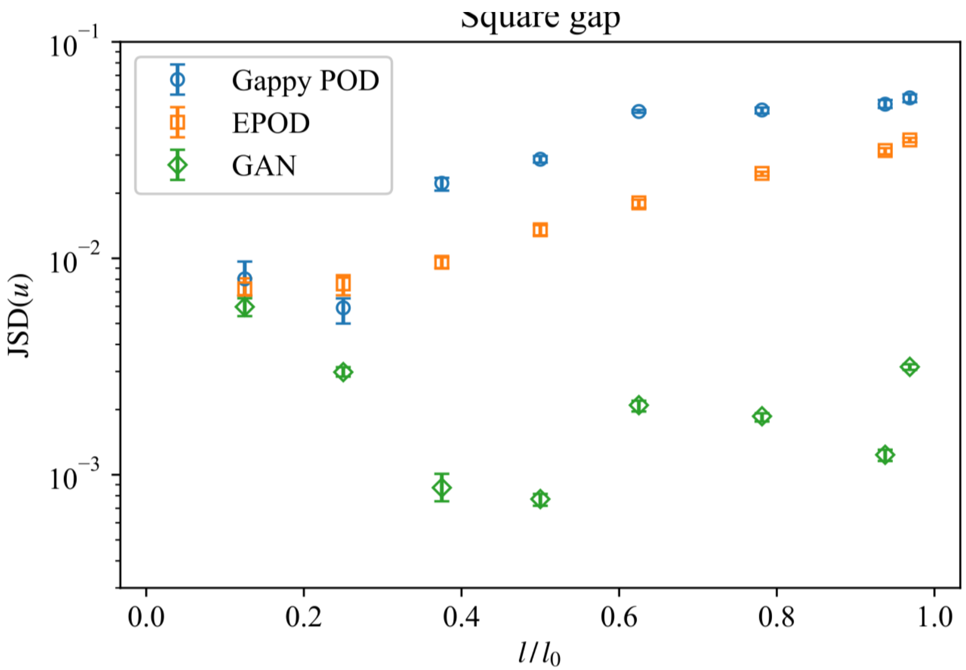
$$M = \tfrac{1}{2}(P+Q)$$

KULLBLACK-LEIBLER

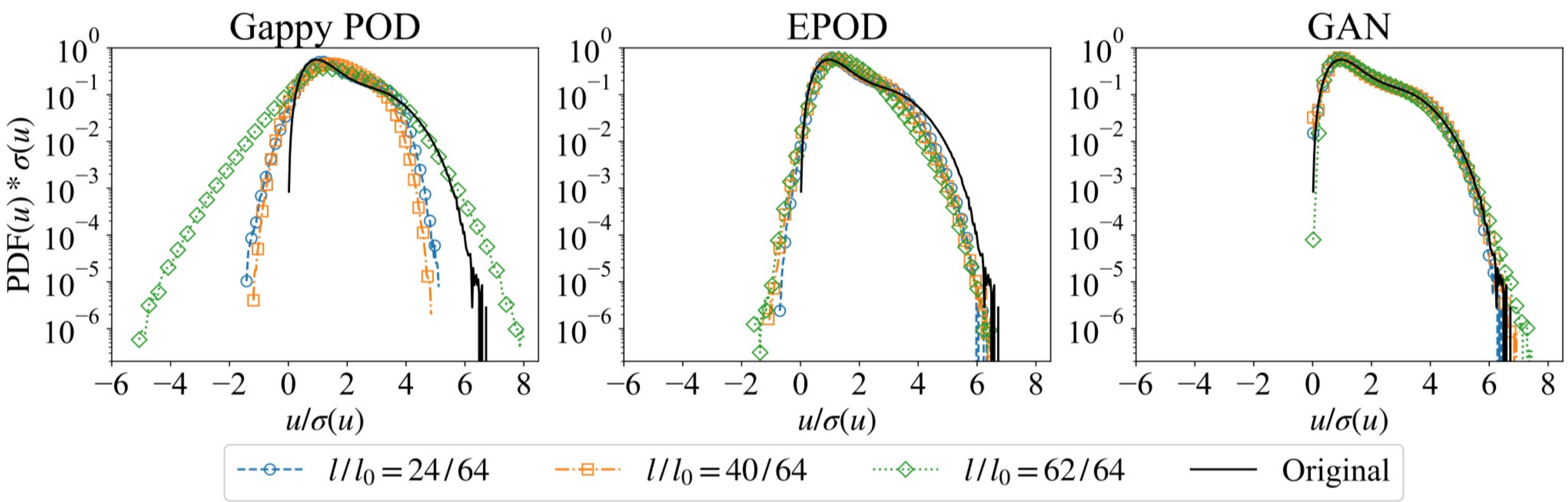
→

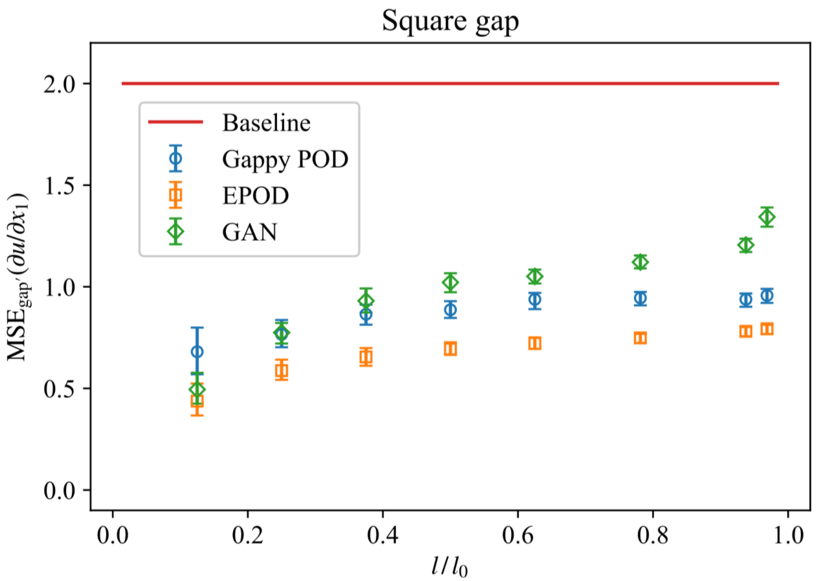
JENSEN-SHANNON

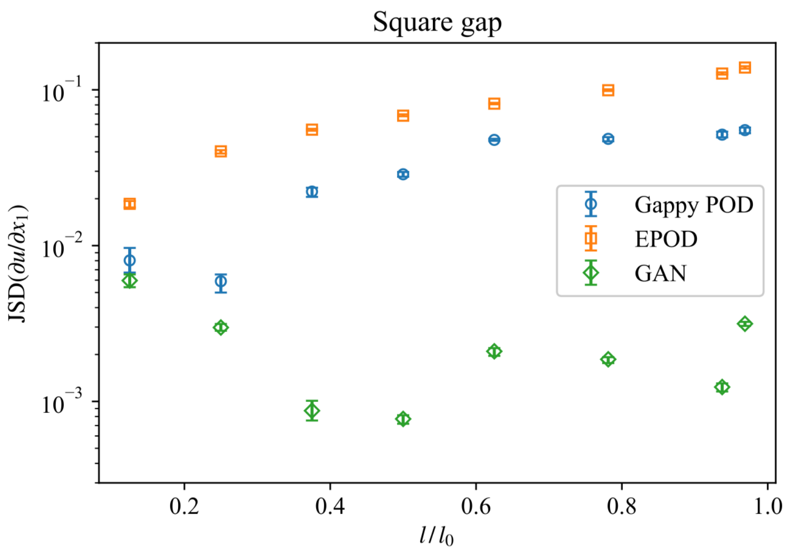
$$\mathrm{JSD}(P \parallel Q) = \frac{1}{2}D(P \parallel M) + \frac{1}{2}D(Q \parallel M),$$

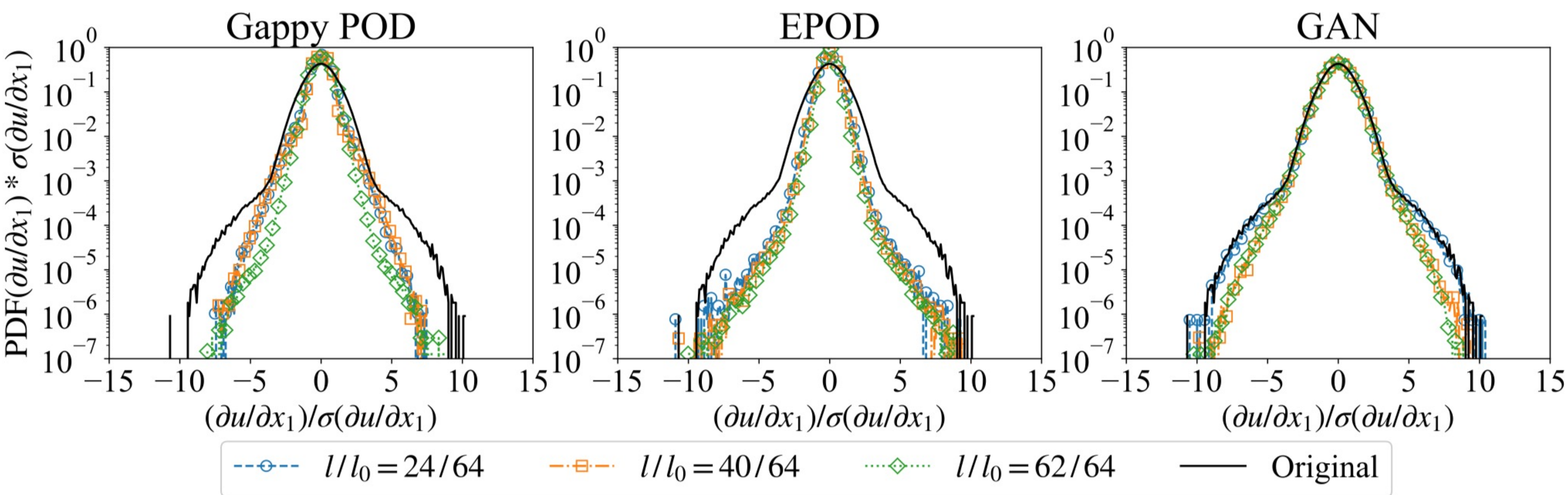


11




Square gap
Baseline
Gappy POD
EPOD
GAN
$\mathrm{MSE}_{\mathrm{gap}}(\partial u/\partial x_1)$
l/l_0


Square gap
Gappy POD
EPOD
GAN
$\mathrm{JSD}(\partial u/\partial x_1)$
l/l_0


Gappy POD
EPOD
GAN
$\mathrm{PDF}(\partial u/\partial x_1) * \sigma(\partial u/\partial x_1)$
$(\partial u/\partial x_1)/\sigma(\partial u/\partial x_1)$
$l/l_0 = 24/64$
$l/l_0 = 40/64$
$l/l_0 = 62/64$
Original

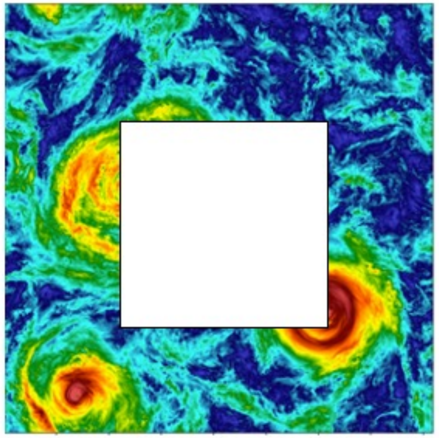
NUDGING: AN **EQUATION-INFORMED UNBIASED** TOOL TO ASSIMILATE AND RECONSTRUCT TURBULENCE DATA/PHYISICS BY ADDING A DRAG TERM AGAINST PARTIAL FIELD MEASUREMENTS

EQUATIONS BASED

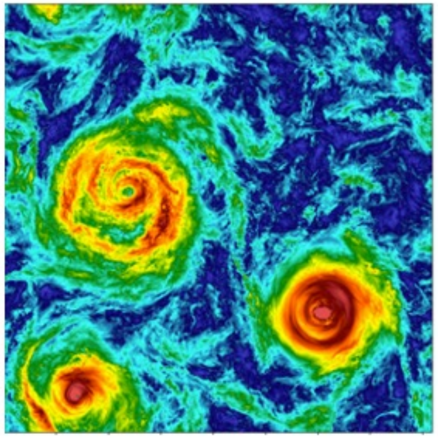
C.C. Lalescu, C. Meneveau and G.L. Eyink. Synchronization of Chaos in Fully Developed Turbulence. Phys. Rev. Lett. 110, 084102 (2013)

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)



???????



FULLY PHYSICS COMPLIANT

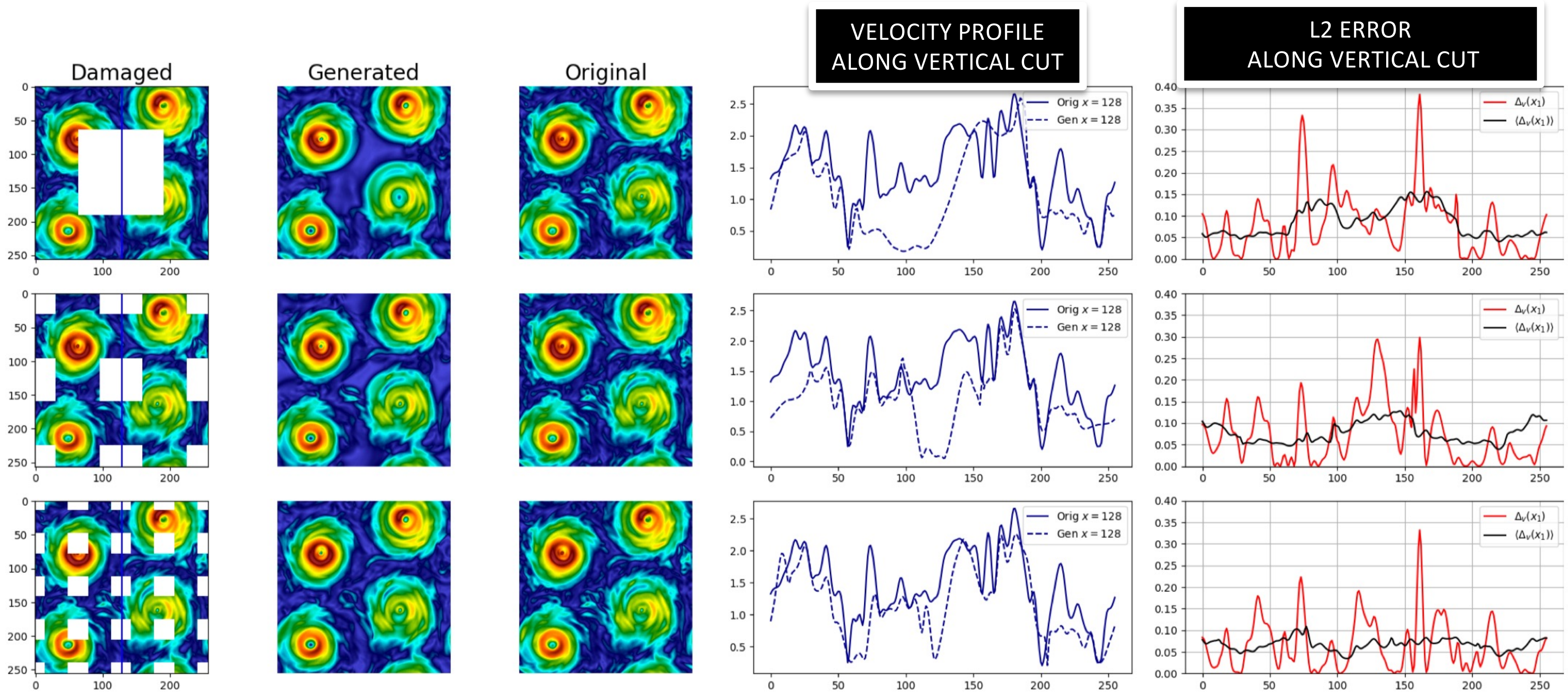
NEED HUGE COMPUTATIONAL RESOURCES

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

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

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

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

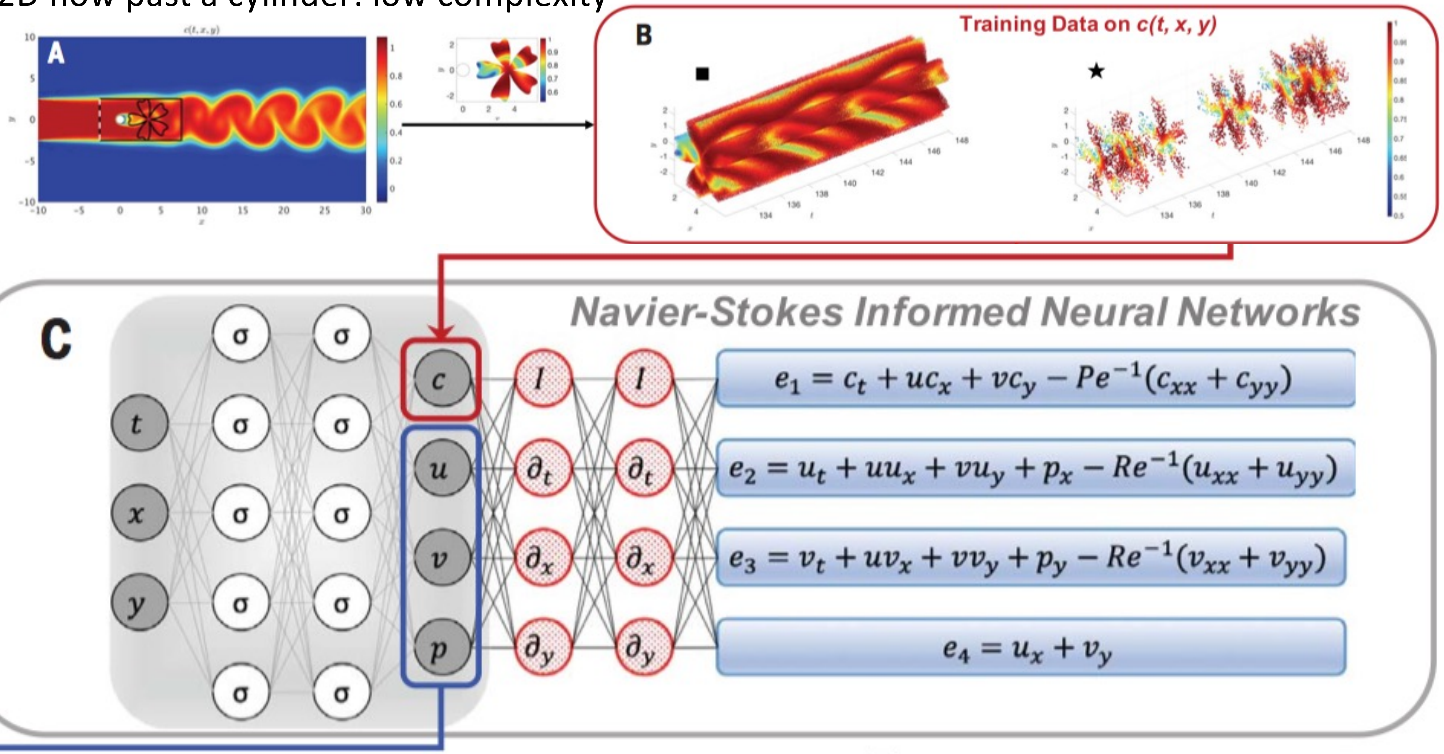


<u>Synchronization to Big Data: Nudging the Navier-Stokes Equations for Data Assimilation of Turbulent Flows</u>
Patricio Clark Di Leoni, Andrea Mazzino, and L. B. Phys. Rev. X **10**, 011023 (2020)

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

2D flow past a cylinder: low complexity

PHYSICS INFORMED

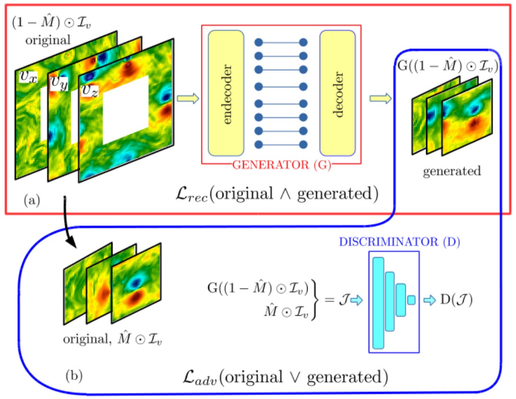


$$MSE = \frac{1}{N}\sum_{n=1}^{N}\left|c(t^n, x^n, y^n, z^n) - c^n\right|^2 + \sum_{i=1}^{5}\frac{4}{}\frac{1}{M}\sum_{m=1}^{M}\left|e_i(t^m, x^m, y^m, z^m)\right|^2$$

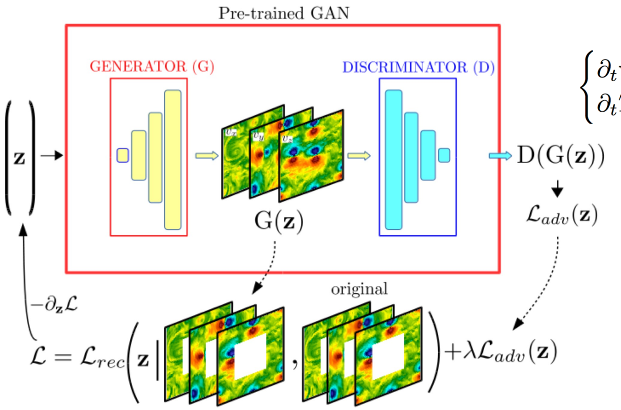
ML-TRAINED ON A SPARSE SPATIO+TEMPORAL DATASET FOR CONCENTRATION -> INFER VELOCITY + PRESSURE -> BACK PROPAGATE FOR GRADIENTS **(AUTOMATIC DIFFERENTIATION)**-> NAVIER-STOKES

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)

CONTEXT ENCODER 1



CONTEXT ENCODER 2



NUDGING

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

CNN-GAN
-EQUATION-FREE
GENERATION OF MISSING DATA ONLY
+ONCE TRAINED -> INSTANTANEOUS
+MIXED INPUT FEATURES

LINEAR CASE: EXTENDED-POD

PRETRAINED CNN-GAN
-EQUATION-FREE
GENERATION OF FRAME & MISSING DATA
-NEW MINIMIZATION FOR EACH DA
+MIXED INPUT FEATURES

LINEAR CASE: GAPPY-POD

+EQUATION-INFORMED
GENERATION OF FRAME & MISSING DATA
-NEW 3D DNS FOR EACH DA
-RESTRICTED INPUT FEATURES