EUROPEAN MECHANICS SOCIETY
EFMC14
14th European Fluid Mechanics Conference
13-16 SEPT. 2022
MEGARON ATHENS INTERNATIONAL CONFERENCE CENTRE
ATHENS, GREECE

Data reconstruction of turbulent flows with Gappy POD and Generative Adversarial Networks (GAN)

T. Li[a,b], M. Buzzicotti[a], F. Bonaccorso[a], L. Biferale[a], M. Wan[b] and S. Chen[b]

[a]*Department of Physics and INFN, University of Rome "Tor Vergata", Via della Ricerca Scientifica 1, 00133, Rome, Italy*

and

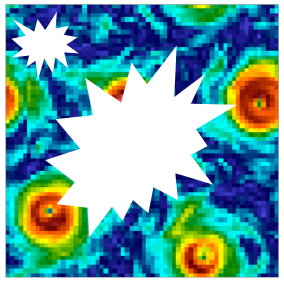
[b]*Department of Mechanics and Aerospace Engineering, Southern University of Science and Technology, Shenzhen, Guangdong 518055, China*

EFMC14, 14th September 2022, Athens

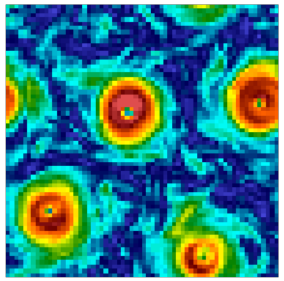
erc European Research Council
TOR VERGATA UNIVERSITÀ DEGLI STUDI DI ROMA
SUSTech Southern University of Science and Technology

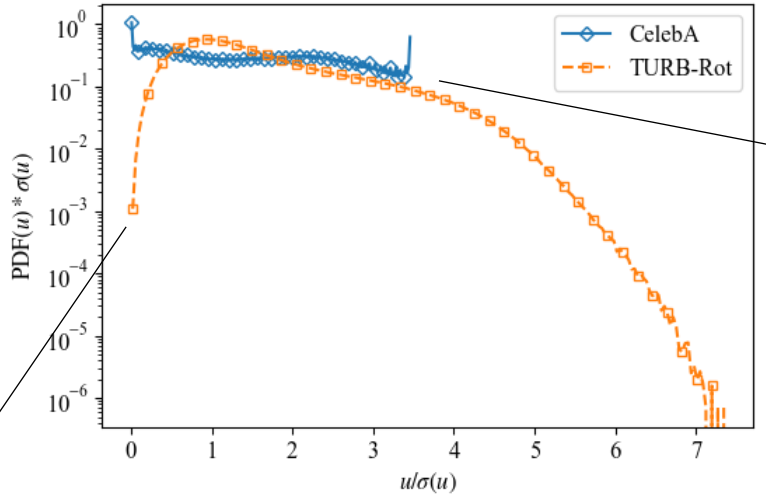
DA of turbulent data vs. image inpainting

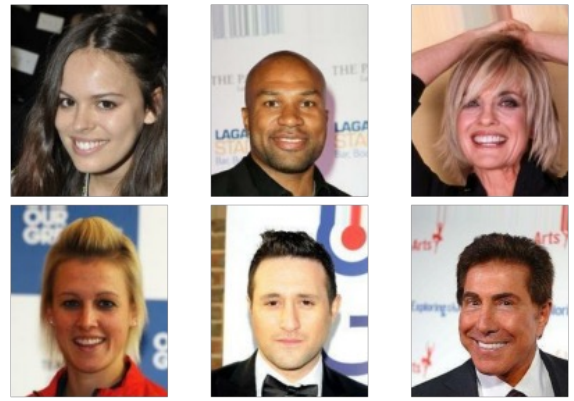
Liu, Ziwei, et al. "Large-scale celebfaces attributes (celeba) dataset." *Retrieved August* 15.2018 (2018): 11.



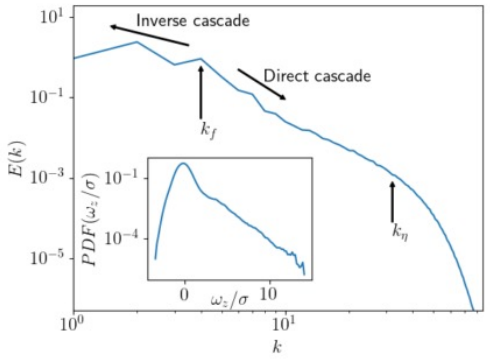
Gappy POD, ML methods, …

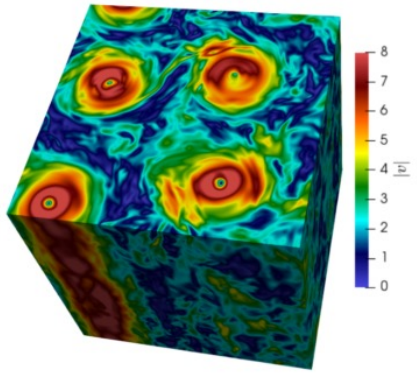






Data complexity





Biferale, Luca, et al. "TURB-Rot. A large database of 3d and 2d snapshots from turbulent rotating flows." *arXiv preprint arXiv:2006.07469* (2020).

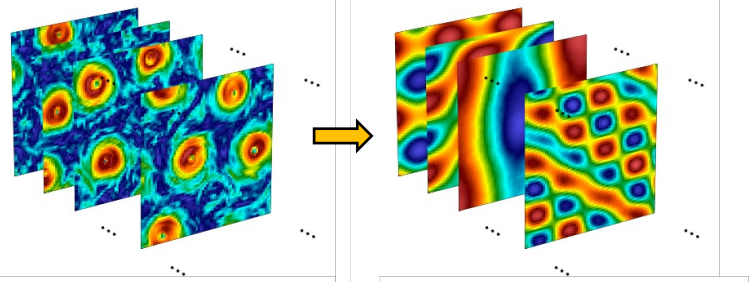
Estimation: Diversity vs. quantitative results

Bordes, Florian, Sina Honari, and Pascal Vincent. "Learning to generate samples from noise through infusion training." *arXiv preprint arXiv:1703.06975* (2017).

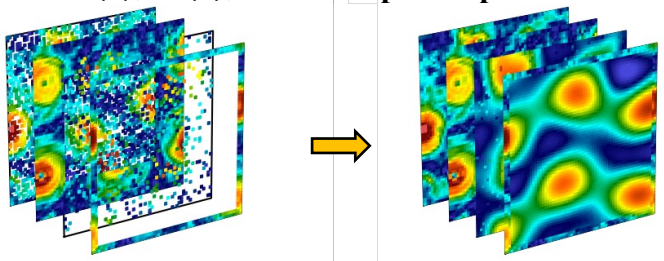
Gappy POD

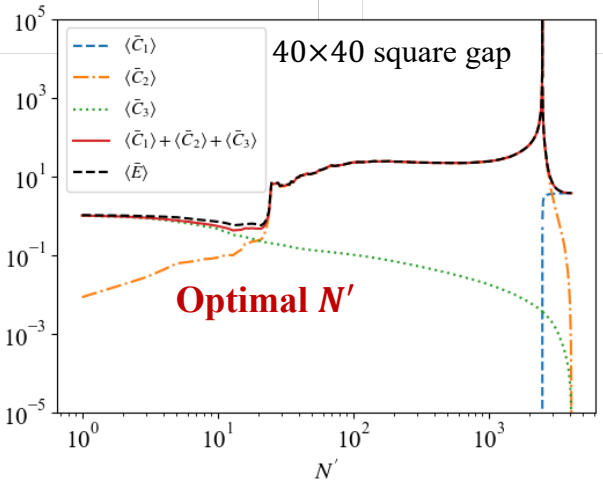
Everson, Richard, and Lawrence Sirovich. "Karhunen–Loeve procedure for gappy data." JOSA A 12.8 (1995): 1657-1664.

Training dataset $u^1(\boldsymbol{x}), u^2(\boldsymbol{x}), \cdots$ **POD modes** $\psi_1(\boldsymbol{x}), \psi_2(\boldsymbol{x}), \cdots$



Gappy data $u^1(\tilde{\boldsymbol{x}}), u^2(\tilde{\boldsymbol{x}}), \cdots$ **Compute expansion coeff.** $a_1, a_2, \cdots$



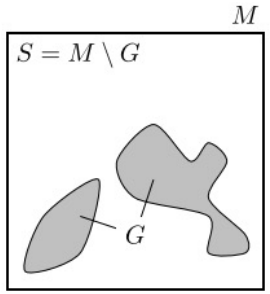


$$u(\boldsymbol{x}) = \sum_{n=1}^{N} a_n \psi_n(\boldsymbol{x}) = \sum_{n=1}^{N'} a_n \psi_n(\boldsymbol{x}) + \sum_{n=N'+1}^{N} a_n \psi_n(\boldsymbol{x})$$

$$\boldsymbol{u} = \boldsymbol{X}\boldsymbol{a} = \boldsymbol{X}'\boldsymbol{a}' + \boldsymbol{r}'$$

$$\boldsymbol{a} = [a_1 \quad a_2 \quad \cdots \quad a_N]^T$$
$$\boldsymbol{a}' = [a_1 \quad a_2 \quad \cdots \quad a_{N'}]^T$$

$$\boldsymbol{X} = [\psi_1(\boldsymbol{x}) \quad \psi_2(\boldsymbol{x}) \quad \cdots \quad \psi_N(\boldsymbol{x})]$$
$$\boldsymbol{X}' = [\psi_1(\boldsymbol{x}) \quad \psi_2(\boldsymbol{x}) \quad \cdots \quad \psi_{N'}(\boldsymbol{x})]$$



$$\boldsymbol{x} = \{x \in M\}$$
$$\bar{\boldsymbol{x}} = \{x \in G\}$$
$$\tilde{\boldsymbol{x}} = \{x \in S\}$$

Compute expansion coefficients

$$\tilde{\boldsymbol{u}} = u(\tilde{\boldsymbol{x}})$$
$$\tilde{\boldsymbol{X}}' = [\psi_1(\tilde{\boldsymbol{x}}) \quad \psi_2(\tilde{\boldsymbol{x}}) \quad \cdots \quad \psi_{N'}(\tilde{\boldsymbol{x}})]$$

$$\tilde{E} = \left\|\tilde{\boldsymbol{u}} - \tilde{\boldsymbol{X}}'\boldsymbol{a}'\right\|_2 = \int_S \left[u(\boldsymbol{x}) - \sum_{n=1}^{N'} a_n \psi_n(\boldsymbol{x})\right]^2 \mathrm{d}\boldsymbol{x}$$
$$\boldsymbol{a}'_b = \underset{\boldsymbol{a}'}{\operatorname{argmin}}\, \tilde{E} = \tilde{\boldsymbol{X}}'_+\tilde{\boldsymbol{u}} + \left(\boldsymbol{I}' - \tilde{\boldsymbol{X}}'_+\tilde{\boldsymbol{X}}'\right)\boldsymbol{w}'$$

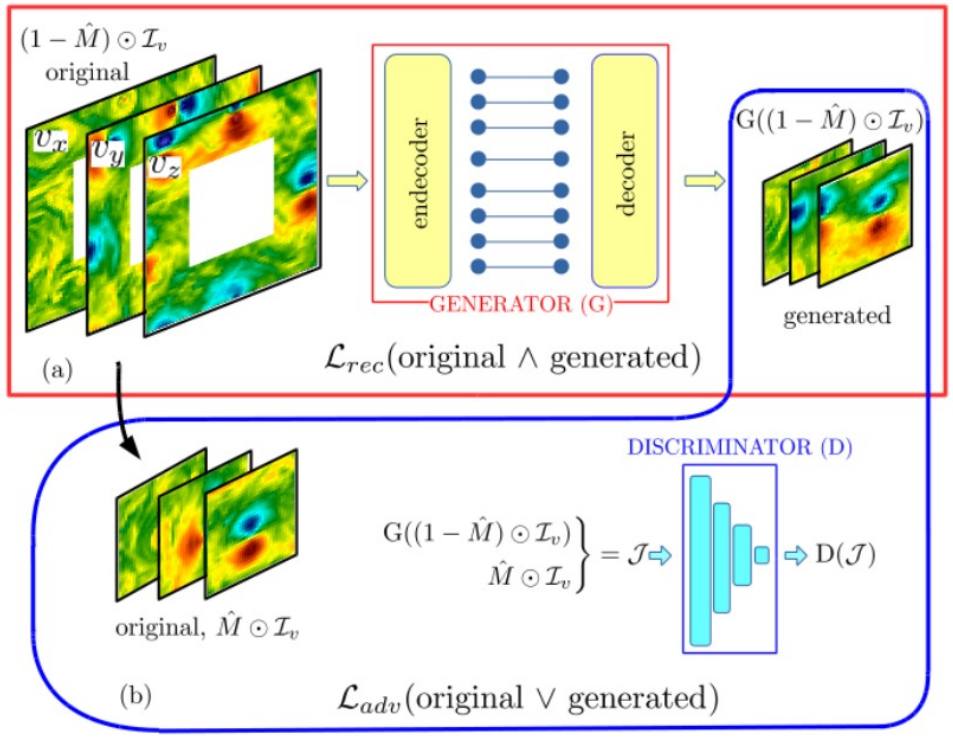
Reconstruction and error analysis

$$\bar{\boldsymbol{u}}_b = \bar{\boldsymbol{X}}'\boldsymbol{a}'_b$$
$$\bar{\boldsymbol{X}}' = [\psi_1(\bar{\boldsymbol{x}}) \quad \psi_2(\bar{\boldsymbol{x}}) \quad \cdots \quad \psi_{N'}(\bar{\boldsymbol{x}})]$$

$$\bar{E} = \|\bar{\boldsymbol{u}} - \bar{\boldsymbol{u}}_b\|_2 = \left\|\bar{\boldsymbol{X}}'\left[\left(\boldsymbol{I}' - \tilde{\boldsymbol{X}}'_+\tilde{\boldsymbol{X}}'\right)(\boldsymbol{a}' - \boldsymbol{w}') - \tilde{\boldsymbol{X}}'_+\boldsymbol{r}'\right] + \boldsymbol{r}'\right\|_2 = \|\bar{\boldsymbol{e}}_1 + \bar{\boldsymbol{e}}_2 + \bar{\boldsymbol{e}}_3\|_2$$
$$\bar{C}_1 = \|\bar{\boldsymbol{e}}_1\|_2, \bar{C}_2 = \|\bar{\boldsymbol{e}}_2\|_2, \bar{C}_3 = \|\bar{\boldsymbol{e}}_3\|_2$$

GAN

Buzzicotti, Michele, et al. "Reconstruction of turbulent data with deep generative models for semantic inpainting from TURB-Rot database." Physical Review Fluids 6.5 (2021): 050503.



Reconstruction Loss (L2 norm)

$$\mathcal{L}_{rec} = \mathbb{E}_{x\in G}\left[\left\|u_{\text{pred}}(\boldsymbol{x}) - u_{\text{true}}(\boldsymbol{x})\right\|_2\right]$$

Adversarial Loss (Binary Cross-Entropy)

A loss function measuring the distance between PDFs of the original and generated data

$$\mathcal{L}_{adv} = \mathbb{E}_{u_{\text{pred}}\in G}\left[\log\left(1 - D\left(u_{\text{pred}}\right)\right)\right]$$

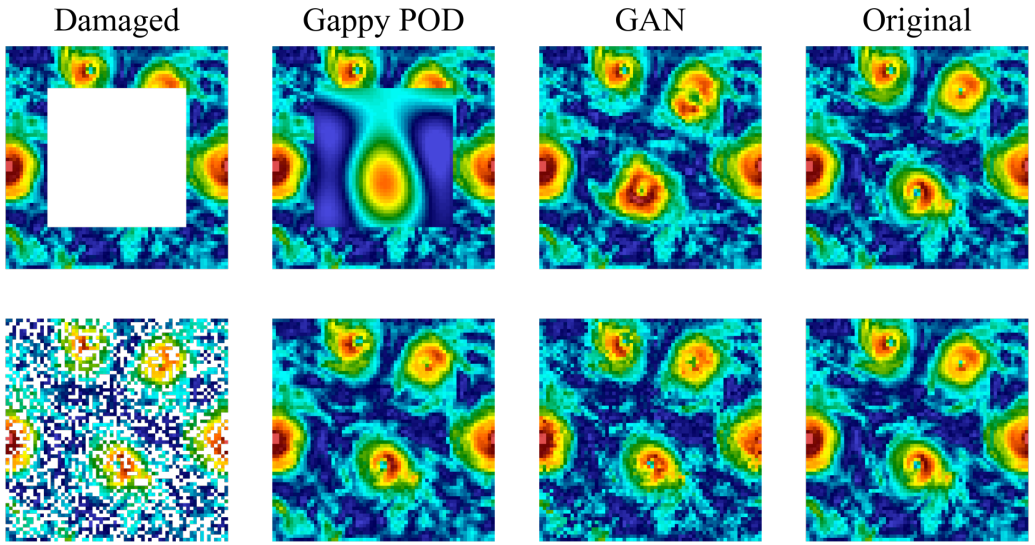
Discrimanator, wants to MAXimize the Adversarial Loss

Generator, wants to MINimize the Total Loss

$$\mathcal{L}_{tot} = (1-\lambda_{adv})\mathcal{L}_{rec} + \lambda_{adv}\,\mathcal{L}_{adv}$$

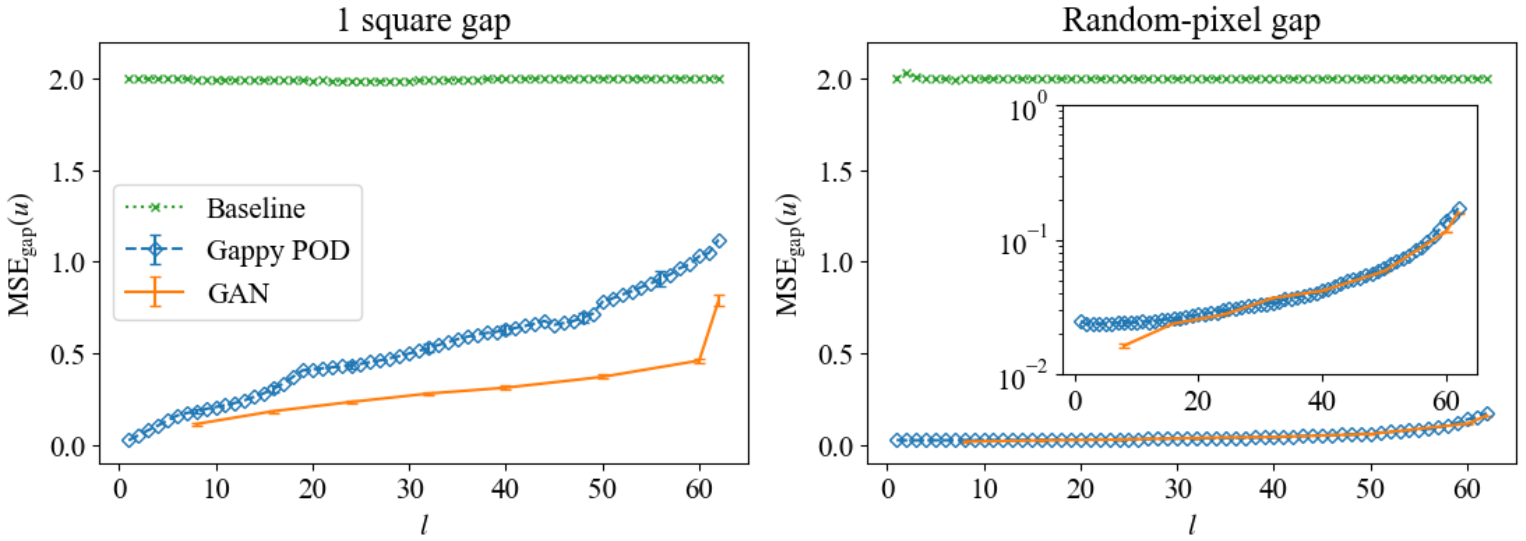
Gappy POD vs GAN

- **Different gap geometries**
  - One square gap
  - random-pixel gap
- **Different gap sizes**

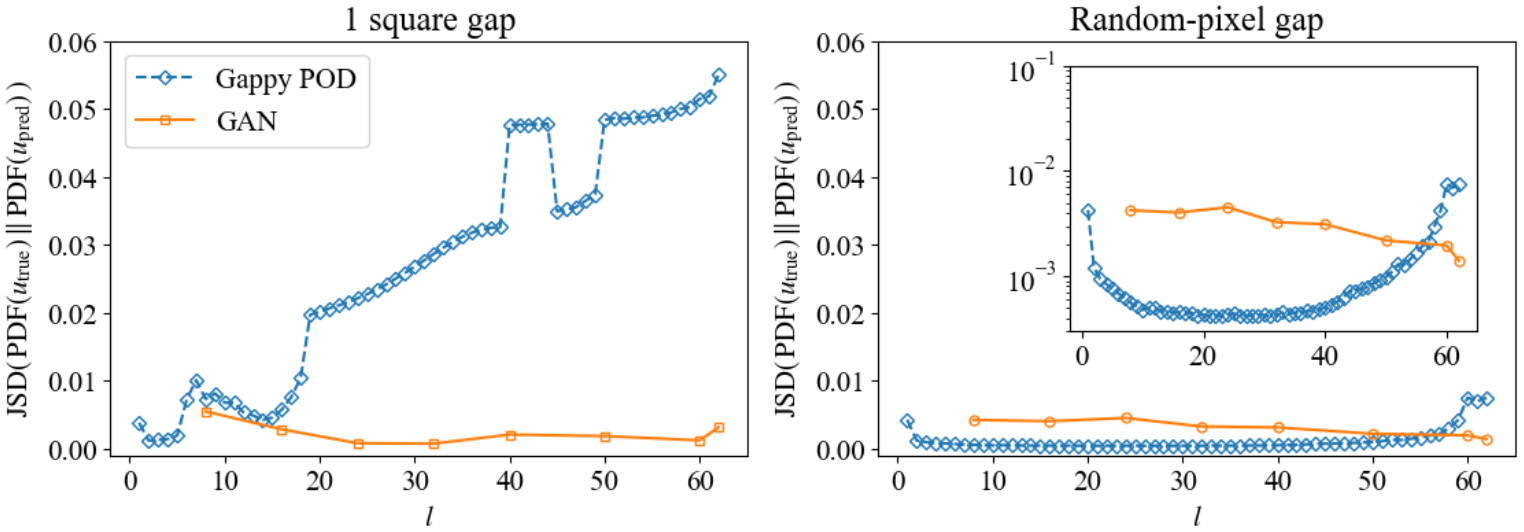


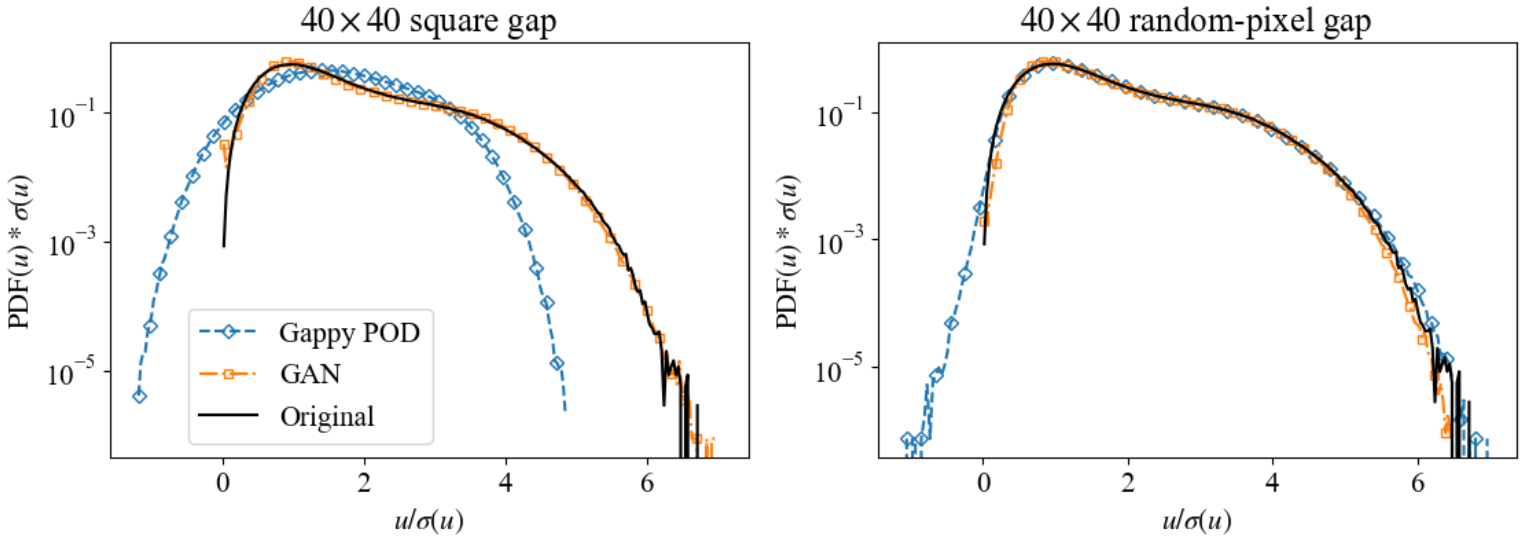
- **Estimation**
  - Reconstruction error
  - Statistical quantities

$$\mathrm{MSE}_{\mathrm{gap}}(u) = \frac{1}{G}\int_G \left[u_{\mathrm{pred}}(\boldsymbol{x}) - u_{\mathrm{true}}(\boldsymbol{x})\right]^2 \mathrm{d}\boldsymbol{x}$$



Gappy POD vs GAN





Measure of the similarity between two probability distributions

Kullback-Leibler divergence

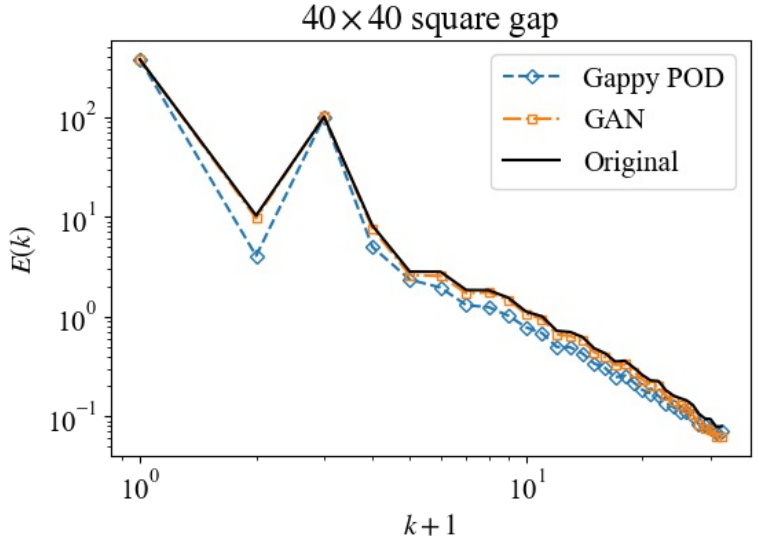
$D(P \parallel Q) = \sum_{x \in \chi} P(x) \log(P(x)/Q(x))$

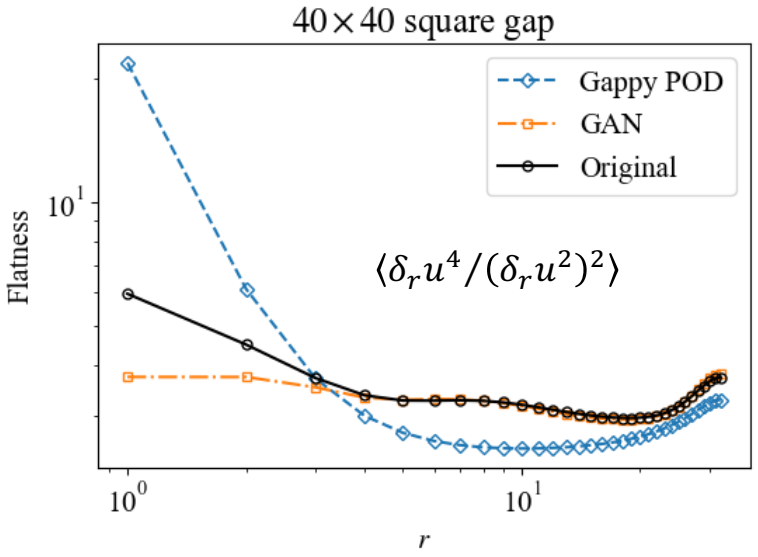
Jensen-Shannon divergence

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

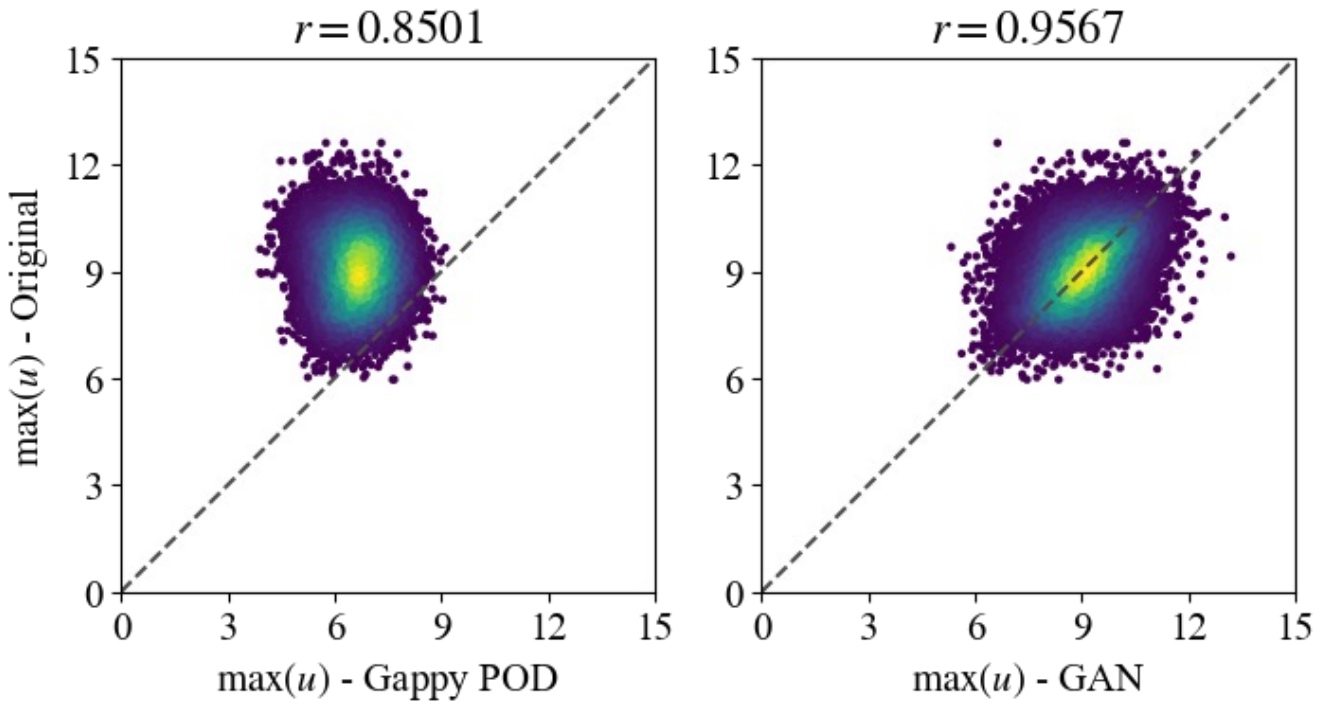
where $M = \frac{1}{2}(P + Q)$

Gappy POD vs GAN





Scatter plot of the maximum values inside the reconstructed region

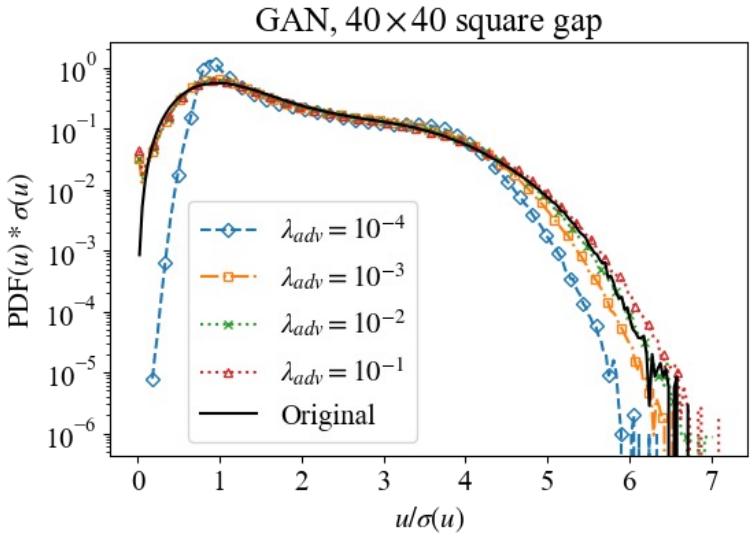


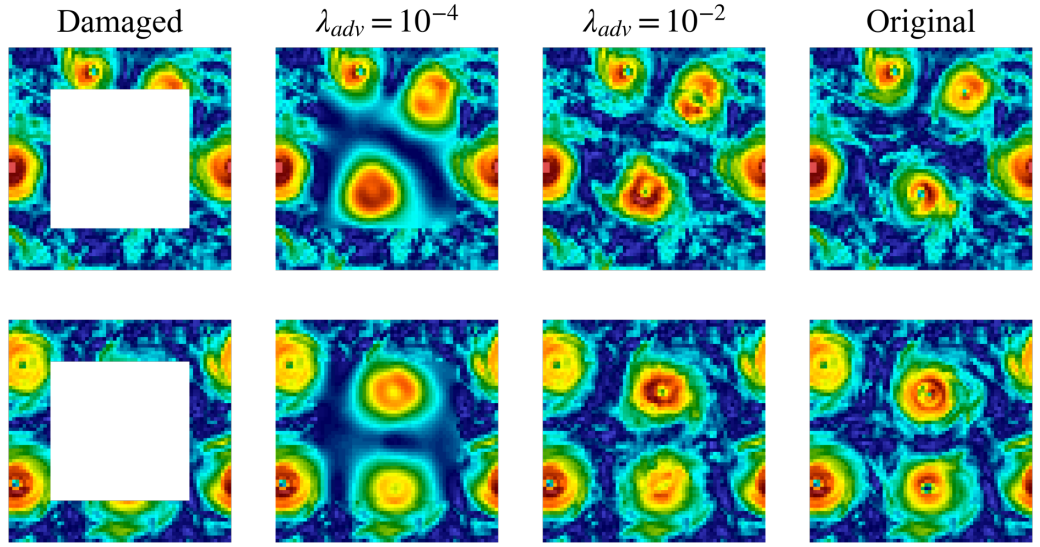
Correlation coefficient

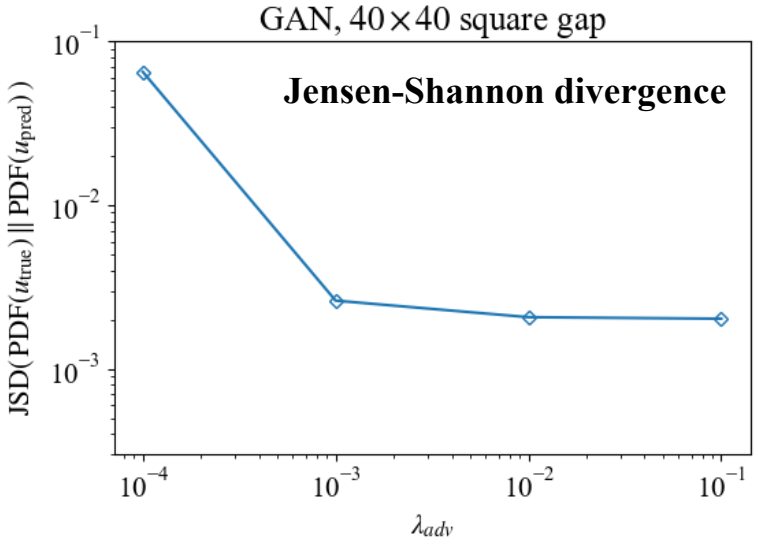
$$r = \left\langle 1 - \left|\max(u^c_{\text{true}}) - \max(u^c_{\text{pred}})\right| \Big/ \sqrt{2\left(\max(u^c_{\text{true}})^2 + \max(u^c_{\text{pred}})^2\right)} \right\rangle$$

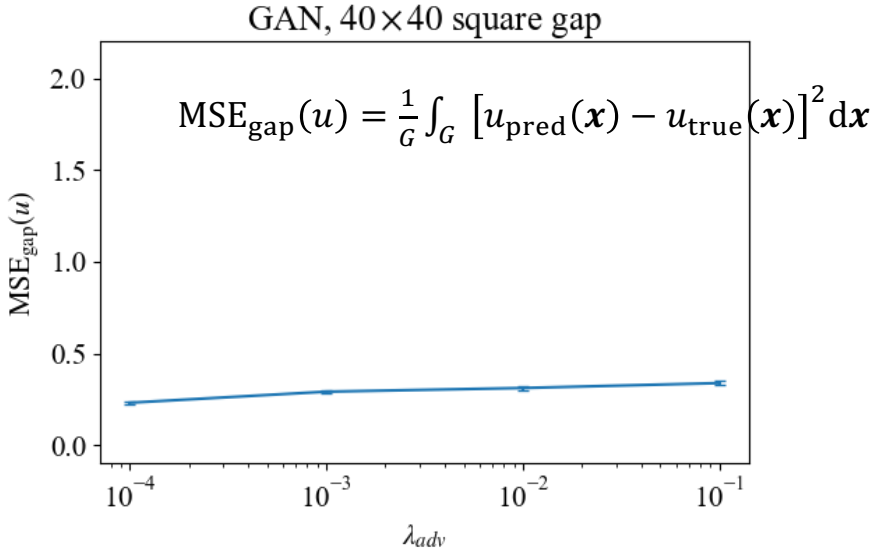
GAN: Effects of adversarial ratio

$$\mathcal{L}_{tot} = (1 - \lambda_{adv})\mathcal{L}_{rec} + \lambda_{adv}\,\mathcal{L}_{adv}$$









Summary

☐ Gappy POD

- **Error can be optimized, by tuning N' to reach a balance between different components**

☐ GAN

- **Adversarial ratio controls the trade-off between MSE and statistical quantities**

☐ Gappy POD vs. GAN

- **Gappy POD: less computation, medium random-pixel mask**
- **GAN: small MSE and good turbulent statistical properties**

Thank you !