LOCO: Local Orbit Consistency for Symmetry Robust Neural Operators

Michal Lewandowski
Software Competence Center Hagenberg (SCCH)



Joint work with



Bernhard Heinzl



Roman J. Rainer



Bernhard Nessler



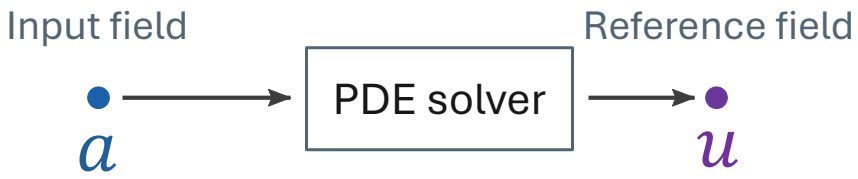
Bernhard A. Moser

Motivation and approach

scch { }

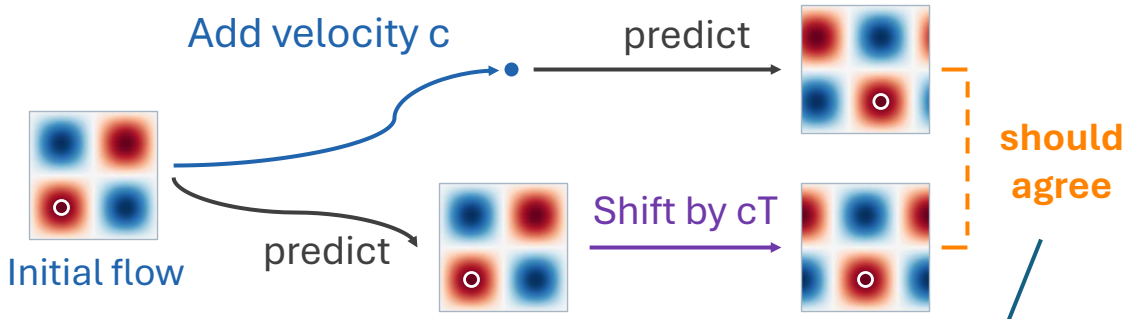
Neural operators learn maps from input fields to PDE solution fields.

Solver-generated labels



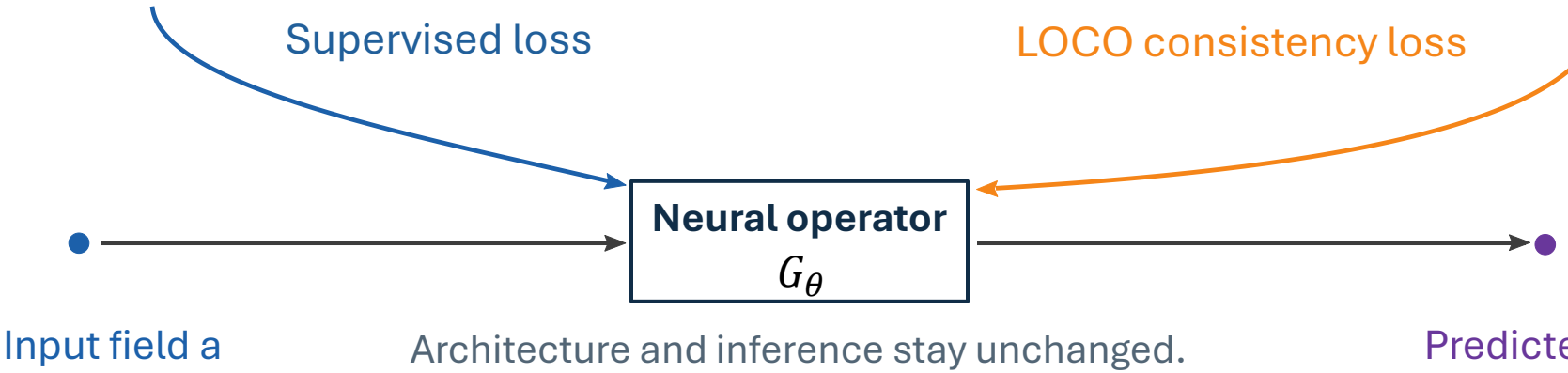
Labels require costly numerical solves.

Example: adding a uniform velocity



Compare predictions from the same input.
No additional reference labels.

Equivariance: transforming the input produces the corresponding transformation of the prediction

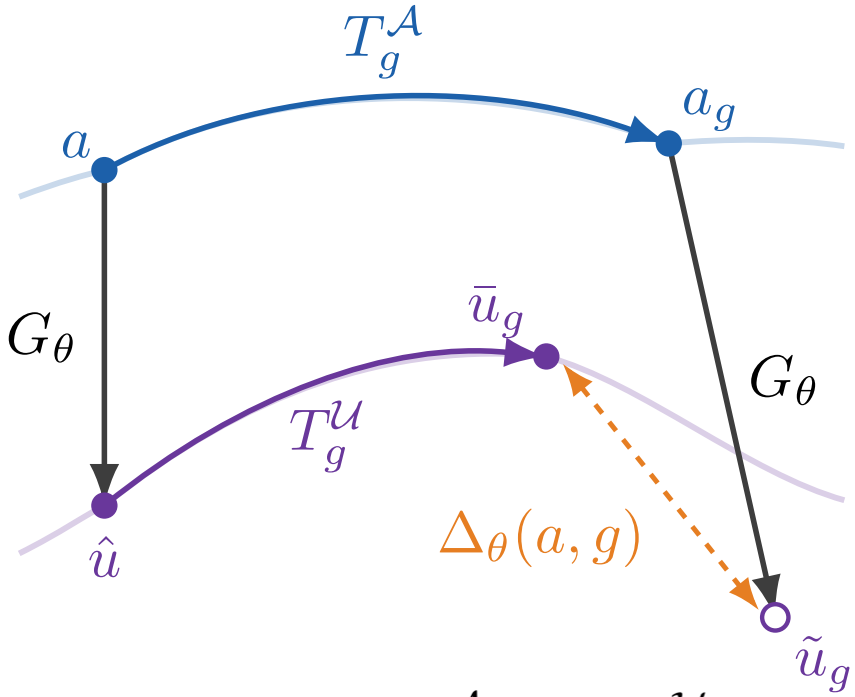


Can known symmetries improve learning from limited labels?

N. Kovachki et al. *Neural Operator: Learning Maps Between Function Spaces With Applications to PDEs*. JMLR, 2023

LOCO: Local Orbit Consistency

scch { }



- LOCO samples a small symmetry transformation g and penalizes disagreement between the two prediction paths.
- "predicting after transforming" ≈ "transforming the prediction"

$$\mathcal{L}_{\text{orb}}(\theta) = \mathbb{E}_{a,\,(X,\epsilon)\sim\rho_{\text{orb}}}\left[\frac{\|\Delta_\theta(a,g)\|_{\mathcal{U}}^2}{\epsilon^2+\eta}\right],\quad g=\exp(\epsilon X)$$

The orbit term compares predictions; no reference solution is needed

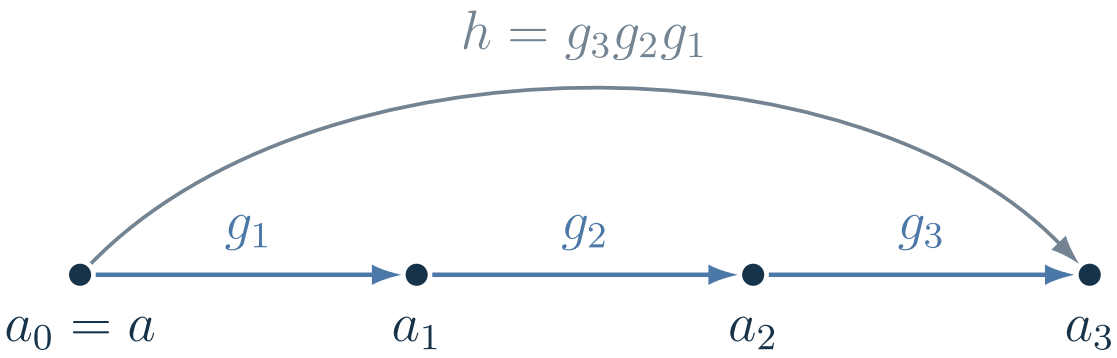
The small-step expansion connects LOCO to tangent propagation [2]

$$\Delta_\theta(a,g) := \underbrace{G_\theta(T_g^{\mathcal{A}}a)}_{\text{transform, then predict}} - \underbrace{T_g^{\mathcal{U}}G_\theta(a)}_{\text{predict, then transform}}$$

Method	**Compares the prediction $G_\theta(T_g^{\mathcal{A}}a)$ with…**	**Exact training objective**
Augmentation [1]	The transformed reference solution $T_g^{\mathcal{U}}u$	$\min_\theta \ \mathcal{L}_{\text{sup}}(\theta) + \lambda_{\text{aug}}\mathcal{L}_{\text{aug}}(\theta)$
LOCO (ours)	The transformed model prediction $T_g^{\mathcal{U}}G_\theta(a)$	$\min_\theta \ \mathcal{L}_{\text{sup}}(\theta) + \lambda_{\text{aug}}\mathcal{L}_{\text{aug}}(\theta) + \lambda_{\text{orb}}\mathcal{L}_{\text{orb}}(\theta)$

[1] J. Brandstetter et al. *Lie Point Symmetry Data Augmentation for Neural PDE Solvers*. ICML, 2022.
[2] P. Simard et al. *Tangent Prop - A formalism for specifying selected invariances in an adaptive network*. NIPS, 1991.

LOCO: Theoretical Results



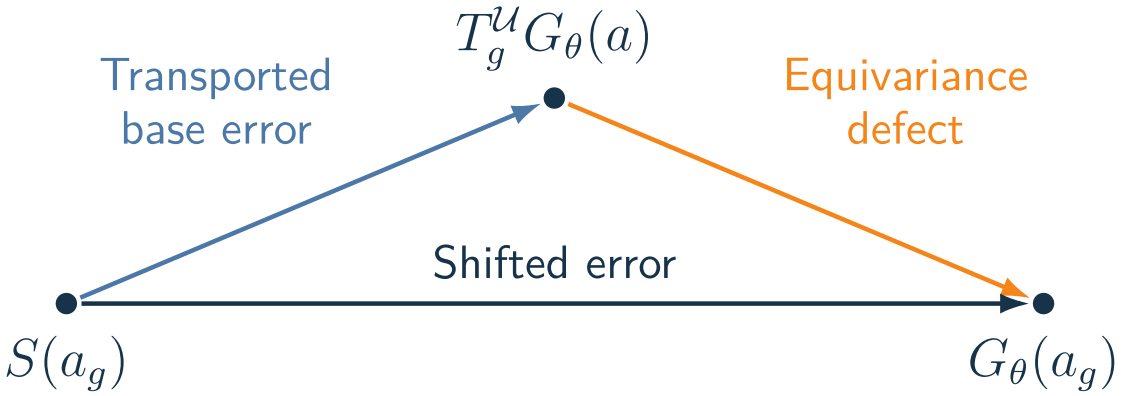
$$\|\Delta_\theta(a, h)\| \le \sum_{j=1}^{3} \|\Delta_\theta(a_{j-1}, g_j)\|$$

$a_j = T_{g_j}^{\mathcal{A}} a_{j-1}$

Output actions are affine and isometric on differences.

- We can bound the defect of a composed transformation by accumulated local defects along the orbit.
- This connects the local training constraint to behavior under larger transformations.

scch { }



$$\|e_g\| \le \|e_0\| + \|\Delta_\theta(a, g)\|$$

$a_g = T_g^{\mathcal{A}} a, \quad e_0 = G_\theta(a) - S(a), \quad e_g = G_\theta(a_g) - S(a_g)$

Exact symmetry; output actions affine and isometric on differences.

- We can decompose prediction error under a symmetry shift into transported base error and equivariance defect.

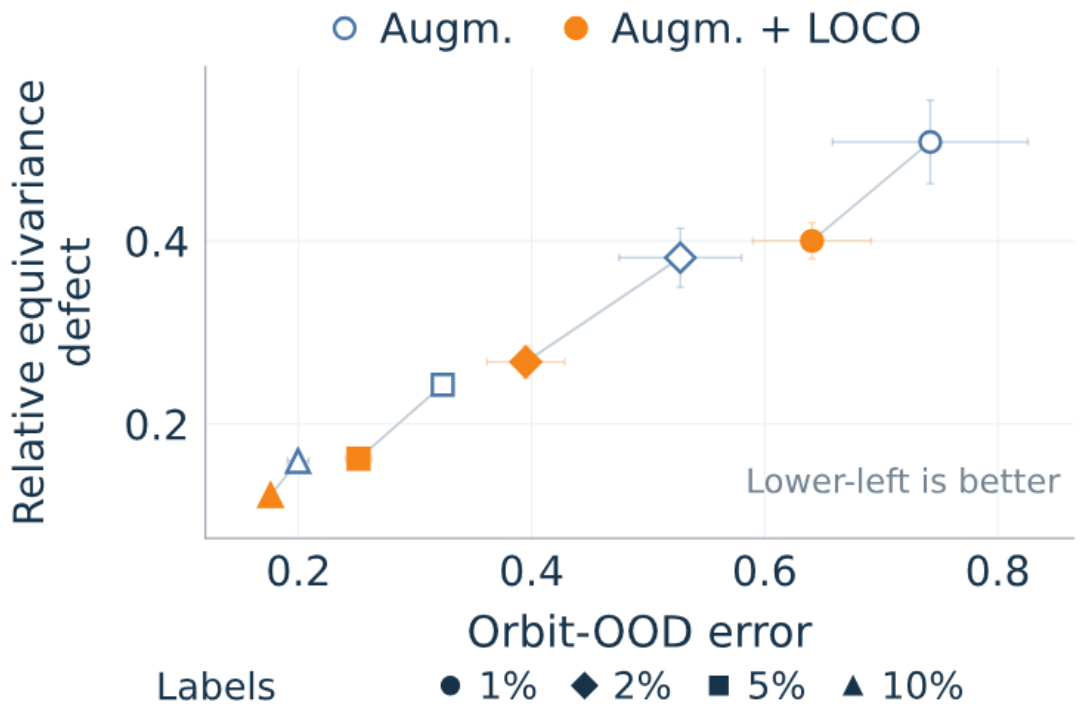
Lower error with fewer labels

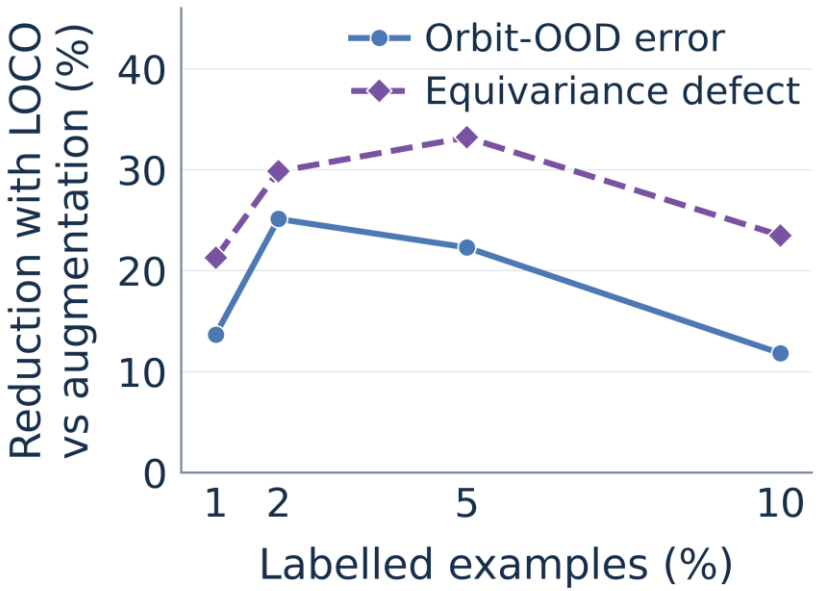
scch { }

Periodic Navier–Stokes experiment:

- FNO backbone [1]
- Compared against augmentations

One labelled example pairs an initial vorticity field and ambient velocity with the future vorticity field computed by a numerical solver.





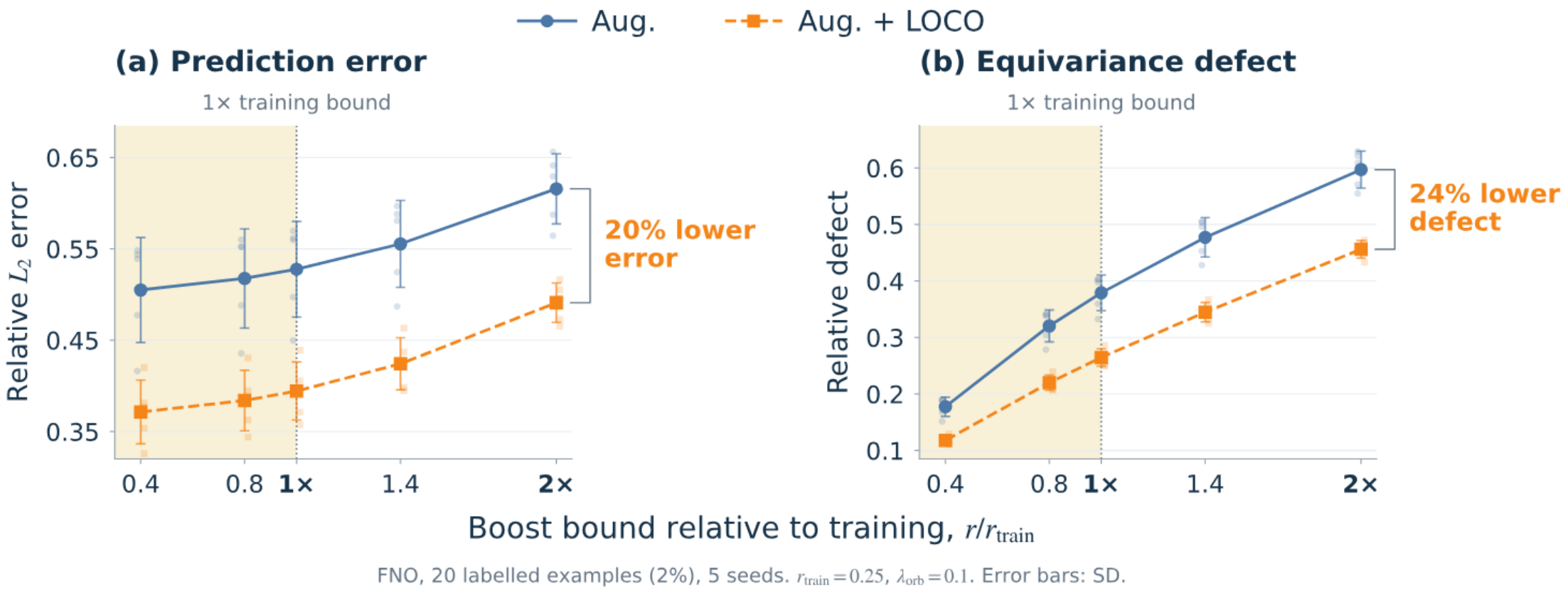
For $\lambda_{orb} = 0.1$ at 2% labels (aver. 5 seeds):

- ~25% lower prediction error on transformed inputs than augmentation
- ~30% lower equivariance defect

[1] Z. Li et al. *Fourier Neural Operator for Parametric Partial Differential Equations*. ICLR, 2021.

Beyond the training boost range

scch { }

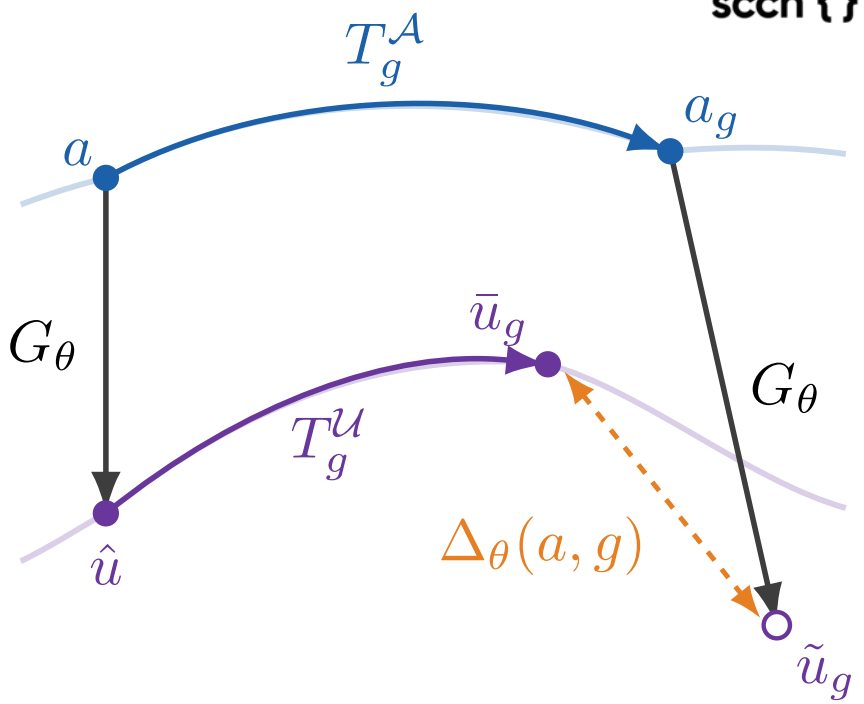


FNO, 20 labelled examples (2%), 5 seeds. $r_{\text{train}} = 0.25$, $\lambda_{\text{orb}} = 0.1$. Error bars: SD.

- At twice the training boost bound, adding LOCO reduces prediction error by 20% relative to augmentation.
- Methods matched in model forward evaluations. Error bars show standard deviations.

Summary

- LOCO adds a label-free symmetry loss during training, unchanged architecture and inference.
- Theory links local defects to composed transformations and prediction error under symmetry shifts.
- Lower prediction error and equivariance defect on FNO and DeepONet
- Requires known, valid symmetry actions.



Periodic Navier-Stokes with FNO

20 labelled examples 2%, averaged over 5 seeds, $\lambda_{\text{orb}} = 0.1$

$\approx 25\%$ lower prediction error under symmetry shifts

$\approx 30\%$ lower equivariance defect

(relative to augmentation with matched model forward evaluations)

Paper (Openreview)