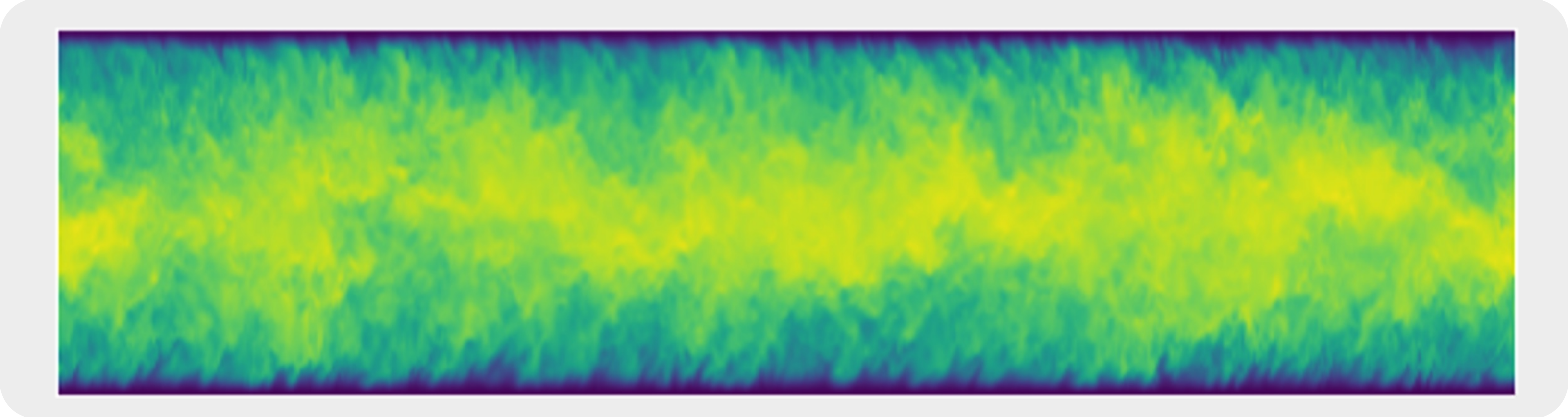
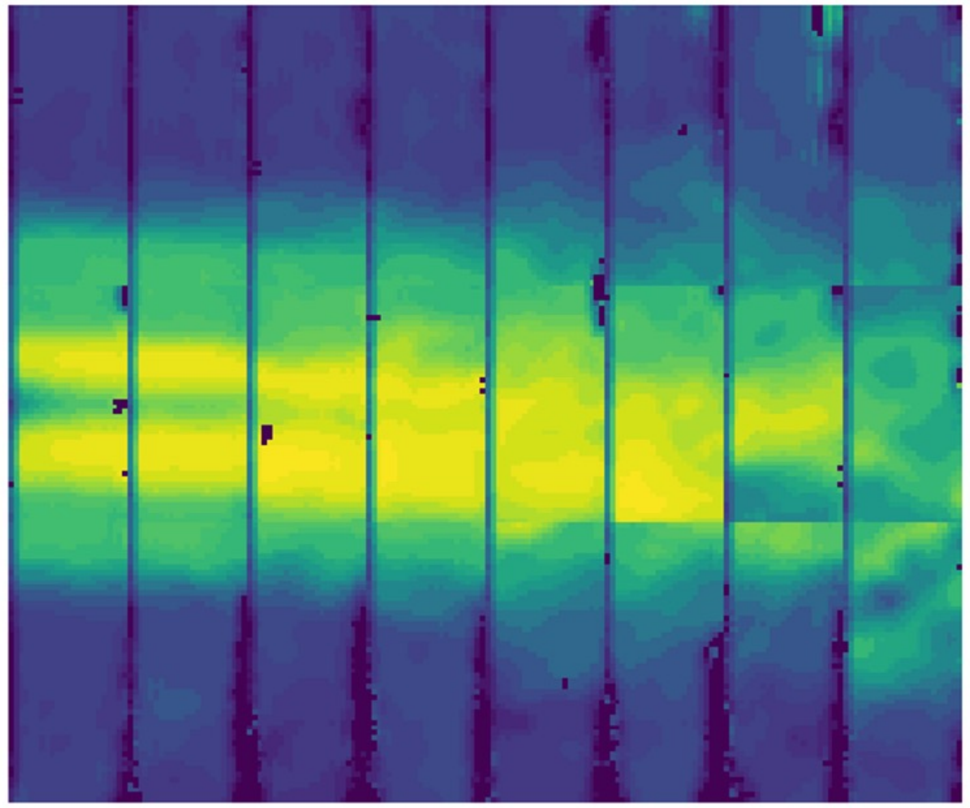
Relaxed Equivariant Networks for Finding Symmetry Breaking in Physical Systems



- **Rui Wang***, Robin Walters*, Rose Yu; Incorporating symmetry into deep dynamics models for improved generalization, ICLR 2021.
- **Rui Wang***, Robin Walters*, Rose Yu; Approximately equivariant networks for imperfectly symmetric dynamics; ICML 2022
- **Rui Wang**, Robin Walters, Tess E Smidt; Relaxed Octahedral Group Convolution for Learning Symmetry Breaking in 3D Physical Systems; arXiv preprint arXiv:2310.02299

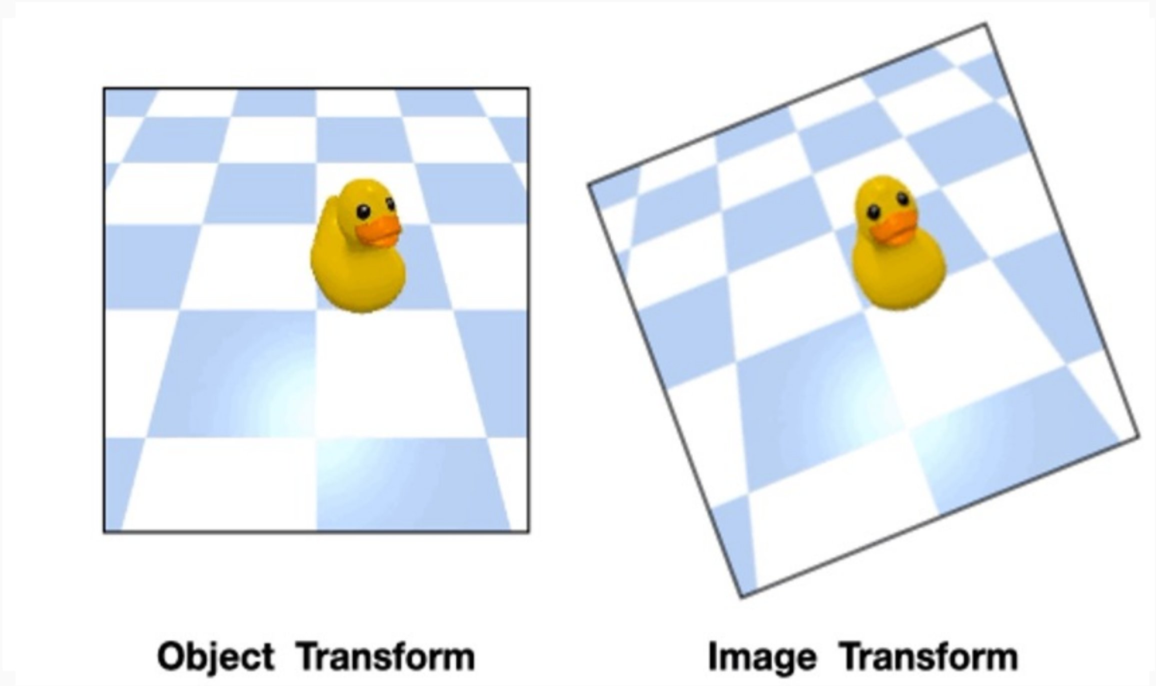
When data is not perfectly symmetric



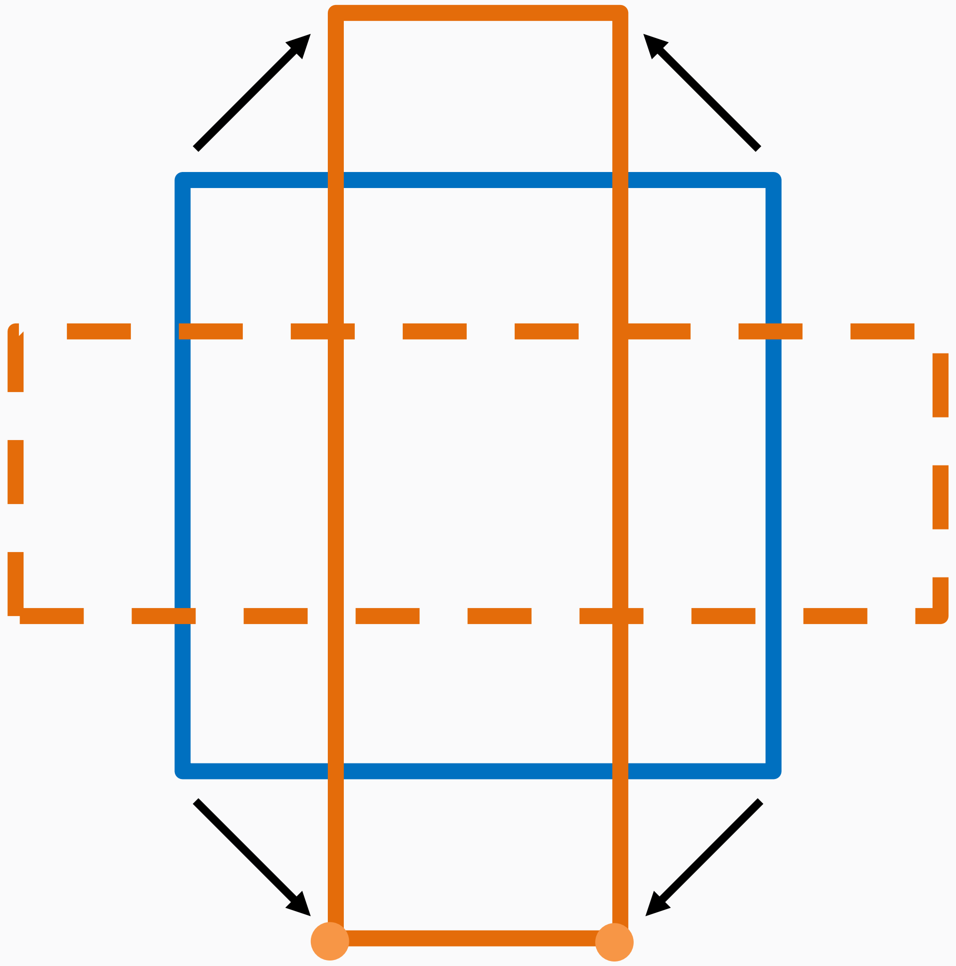
Noisy observations

Unknown external forces

Unknown boundary conditions



Observations not aligned with symmetry or symmetry breaking background

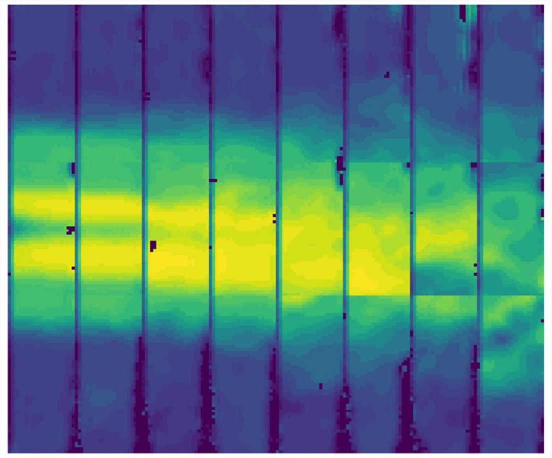


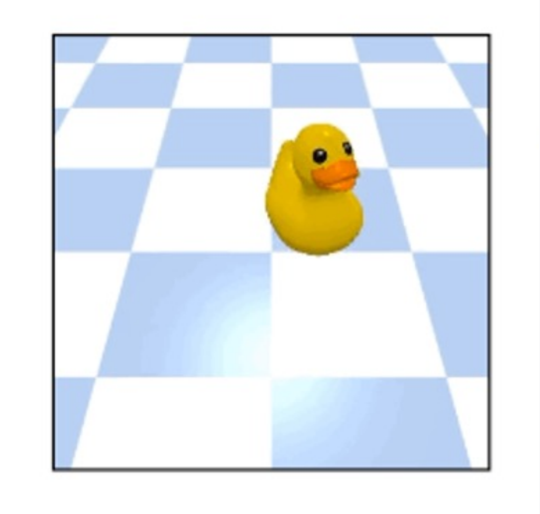
The output has lower symmetry than the input

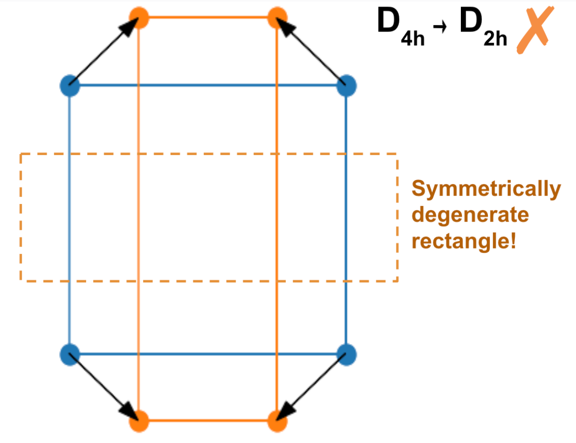
Dian Wang, Jung Yeon Park, Neel Sortur, Lawson L.S. Wong, Robin Walters, Robert Platt; **The Surprising Effectiveness of Equivariant Models in Domains with Latent Symmetry**; *ICLR 2023*

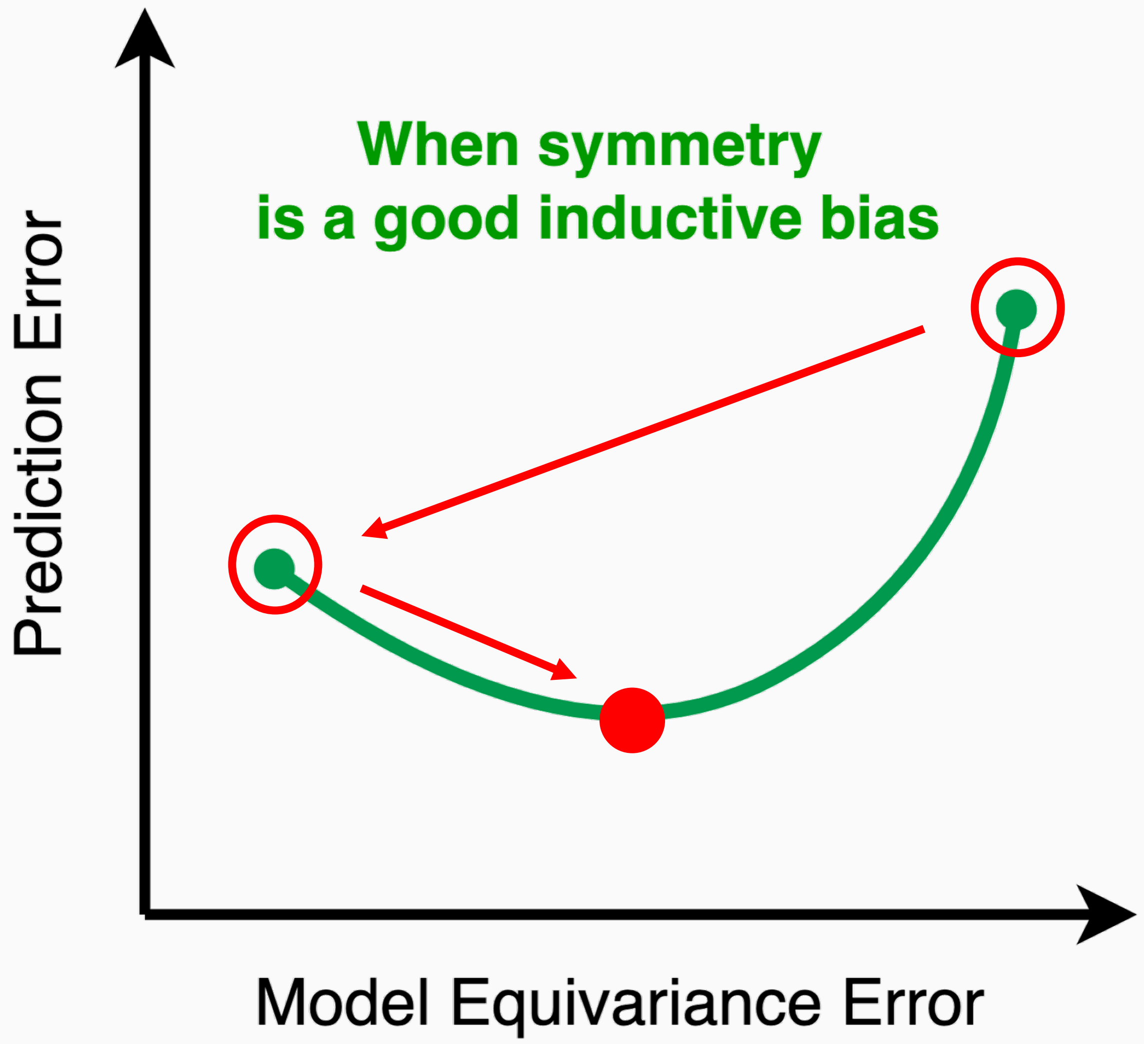
Tess E. Smidt, Mario Geiger, Benjamin Kurt Miller; **Finding symmetry-breaking Order Parameters with Euclidean Neural Networks**; *Physical Review Research 3 (1), L012002*

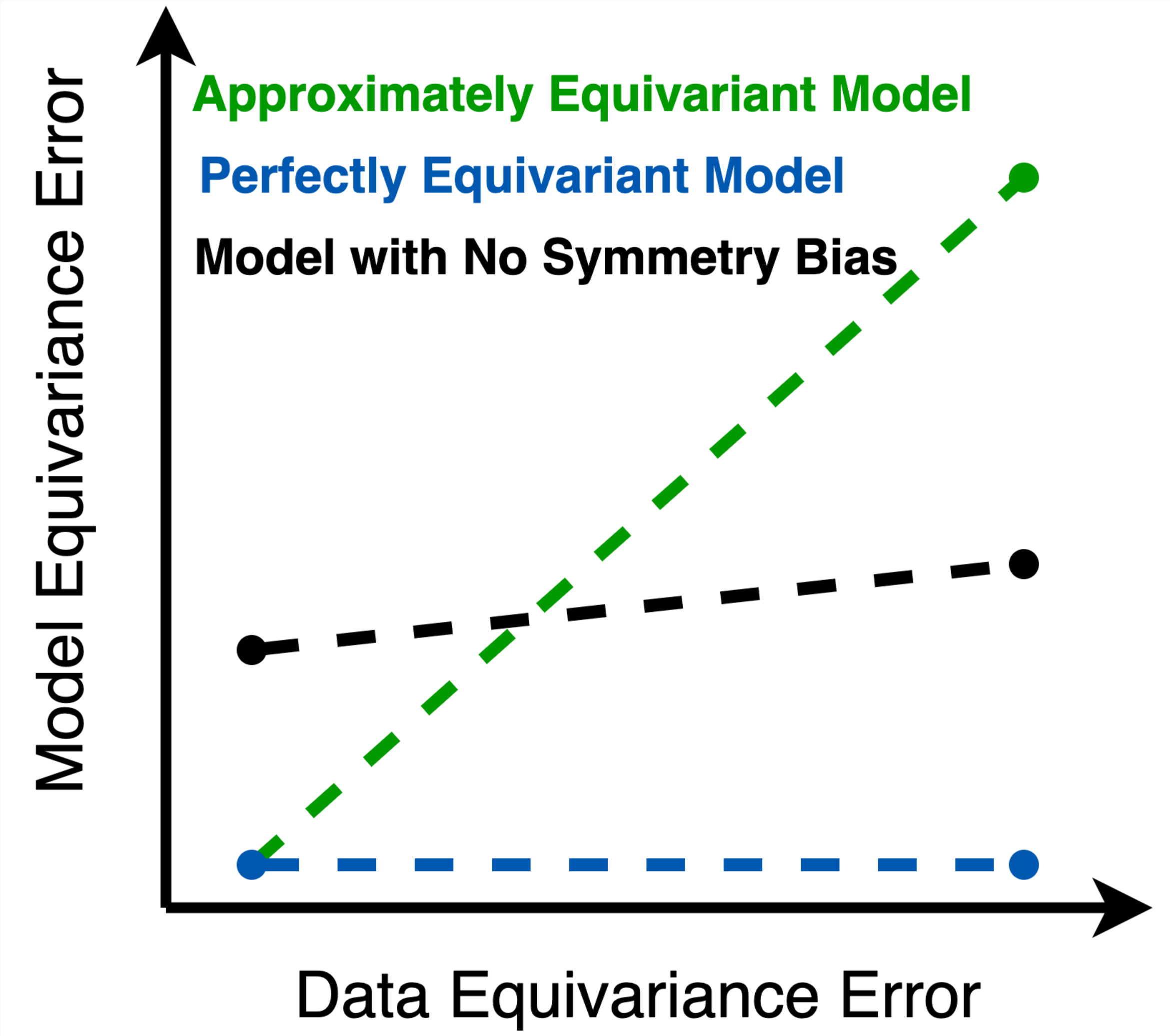
When data is not perfectly symmetric











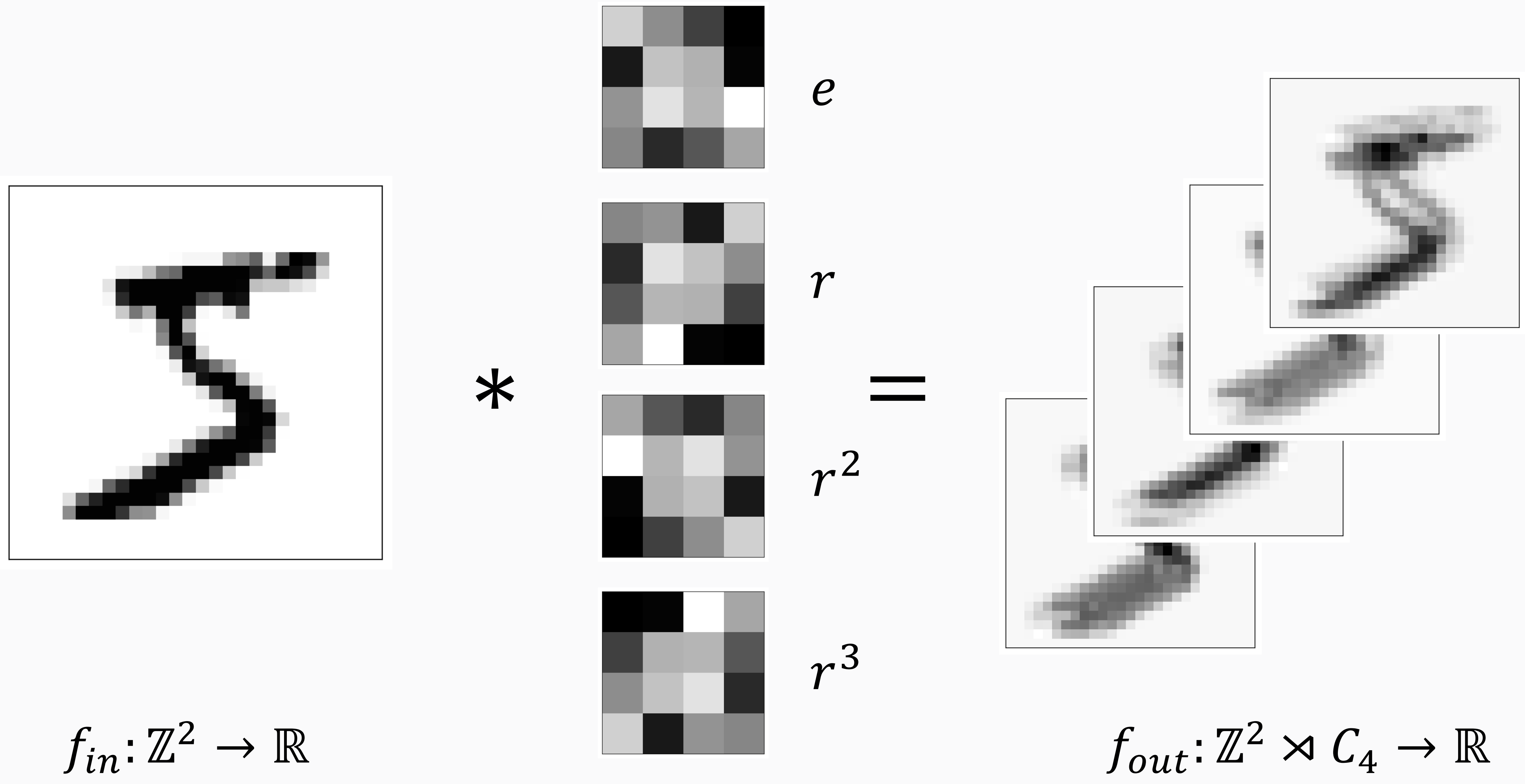
G**-approx-equiv**: $|f(\rho_{in}(g)x) - \rho_{out}(g)f(x)| < \varepsilon$

An ideal model should automatically learn correct amount of symmetry.

Group Convolution Networks

Lift Convolution Layer:

$$f_{out}(\boldsymbol{x}, r) = (f_{in} * \psi)(\boldsymbol{x}) = \sum_{\boldsymbol{y} \in \mathbb{Z}^2} f_{in}(\boldsymbol{y})\, \psi_r(\boldsymbol{y} - \boldsymbol{x}), \quad (\boldsymbol{x}, r) \in \mathbb{Z}^2 \rtimes C_4$$

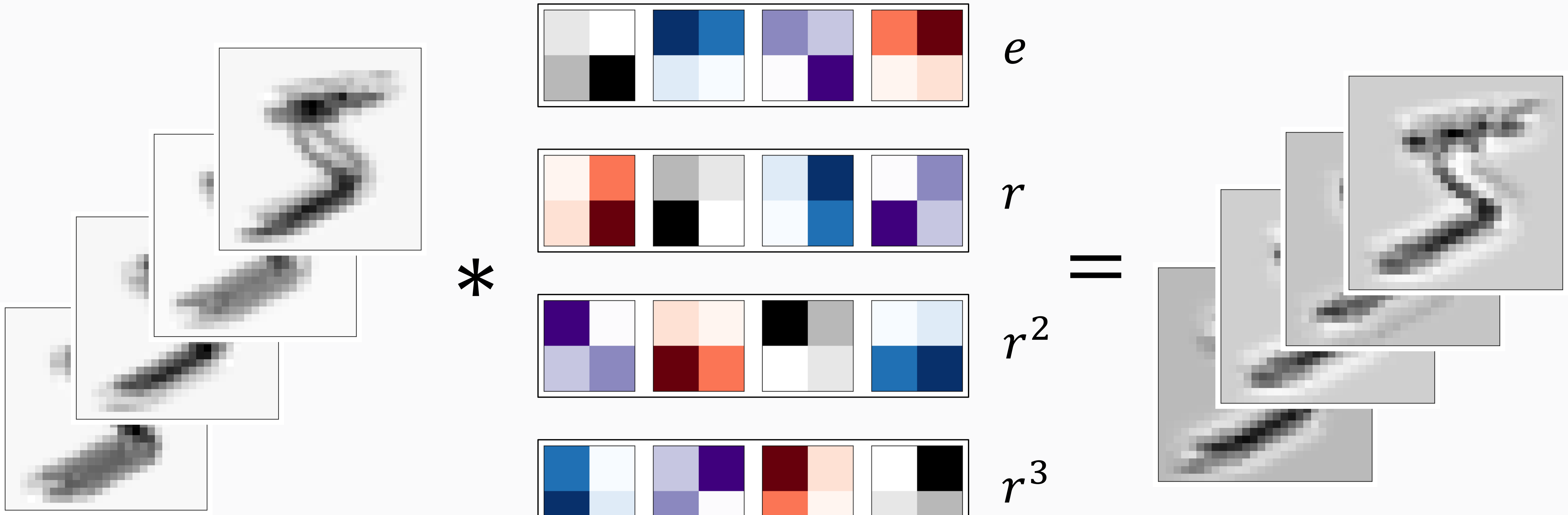


Taco S. Cohen Max Welling ; **Group Equivariant Convolutional Networks**; ICML 2016

Group Convolution Networks

Group Convolution Layer:

$$f_{out}(\boldsymbol{x}, r) = (f_{in} * \psi)(\boldsymbol{x}, r) = \sum_{r' \in C_4} \sum_{\boldsymbol{y} \in \mathbb{Z}^2} f_{in}(\boldsymbol{y}, r)\, \psi\big(\boldsymbol{r}^{-1}(\boldsymbol{y} - \boldsymbol{x}), r^{-1} r'\big), \quad (\boldsymbol{x}, r) \in \mathbb{Z}^2 \rtimes C_4$$



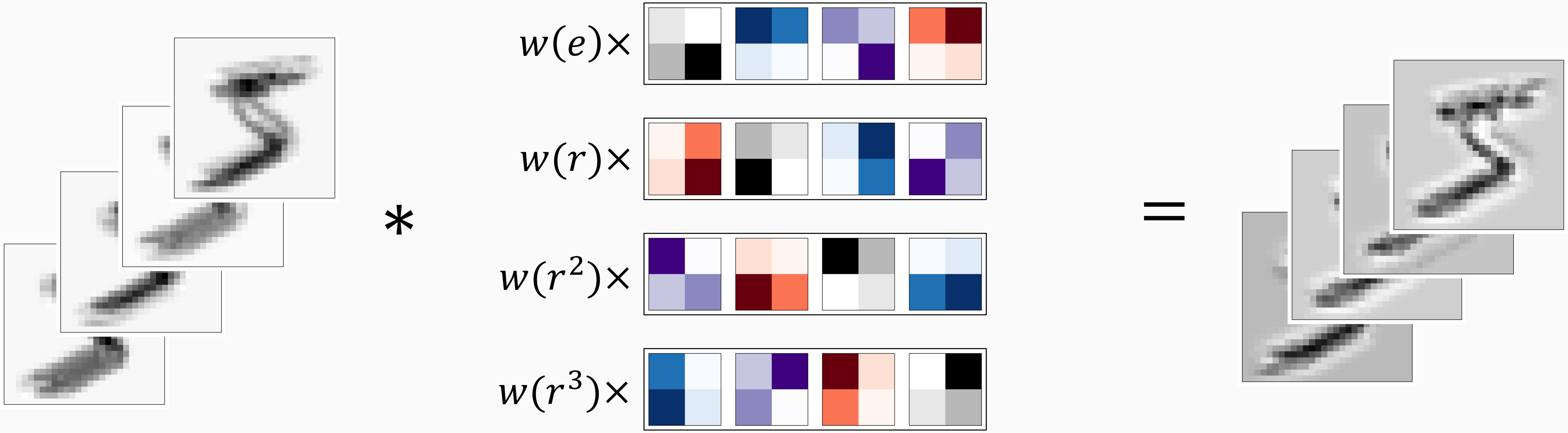
$f_{in}: \mathbb{Z}^2 \rtimes C_4 \to \mathbb{R}$

$f_{out}: \mathbb{Z}^2 \rtimes C_4 \to \mathbb{R}$

Taco S. Cohen Max Welling ; **Group Equivariant Convolutional Networks**; ICML 2016

Relaxed Group Convolution Networks

Relaxing weight-sharing constraints by **introducing group element dependent parameters.**

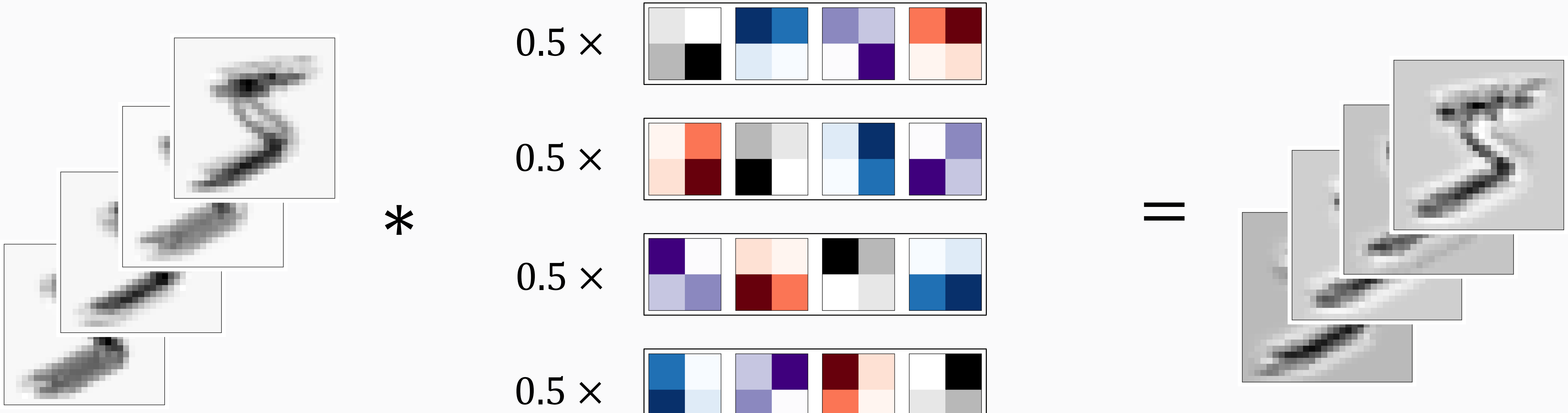


$f_{in}: \mathbb{Z}^2 \rtimes C_4 \rightarrow \mathbb{R}$

$f_{out}: \mathbb{Z}^2 \rtimes C_4 \rightarrow \mathbb{R}$

Relaxed Group Convolution Networks

Relaxing weight-sharing constraints by **introducing group element dependent parameters.**



$f_{in}: \mathbb{Z}^2 \rtimes C_4 \to \mathbb{R}$

$f_{out}: \mathbb{Z}^2 \rtimes C_4 \to \mathbb{R}$

Relaxed Group Convolution Networks

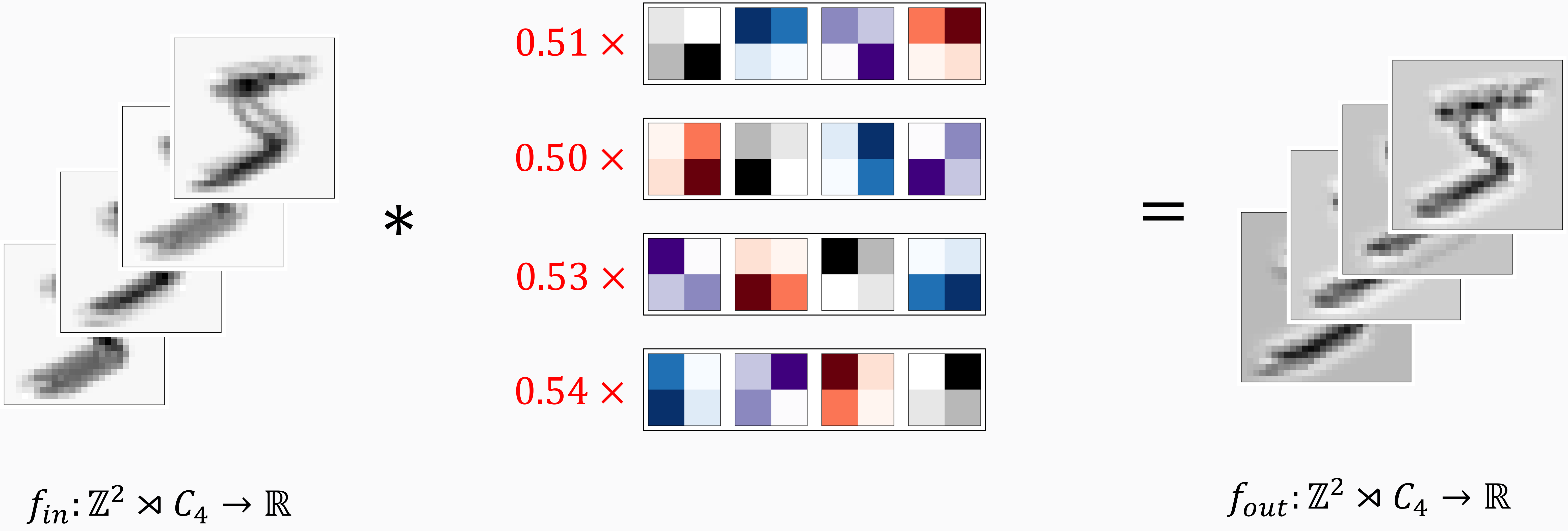
Relaxing weight-sharing constraints by **introducing group element dependent parameters.**

$f_{in}: \mathbb{Z}^2 \rtimes C_4 \rightarrow \mathbb{R}$

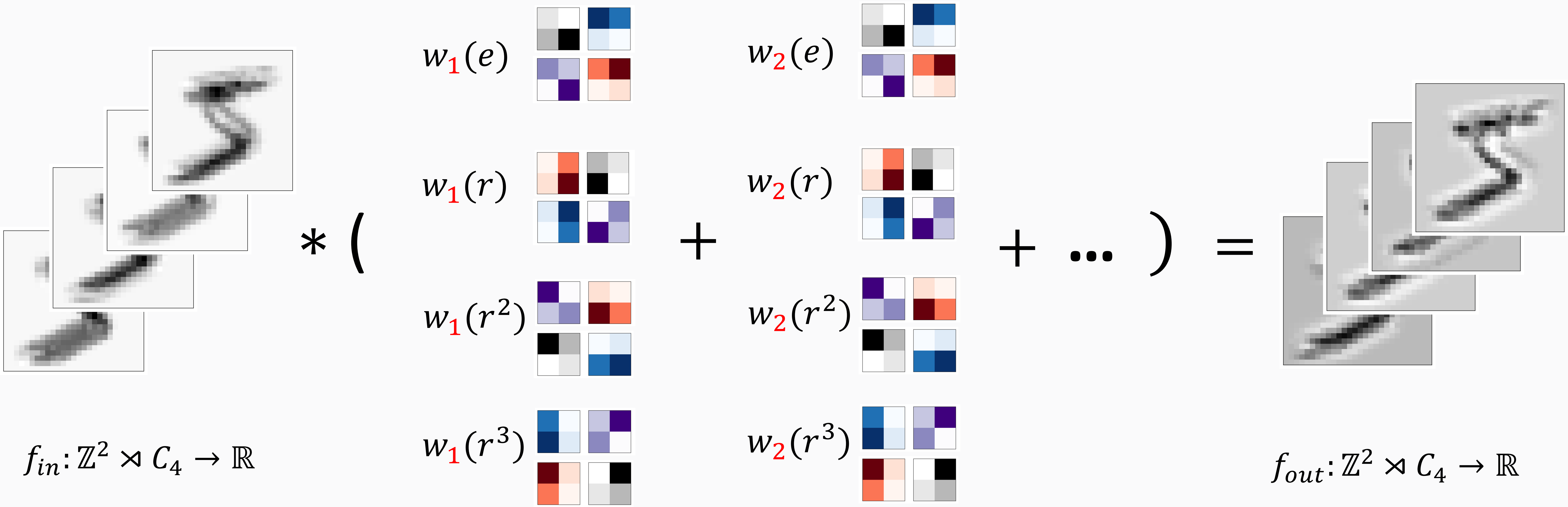
$f_{out}: \mathbb{Z}^2 \rtimes C_4 \rightarrow \mathbb{R}$

Relaxed Group Convolution Networks

Relaxing weight-sharing constraints by **introducing group element dependent parameters.**

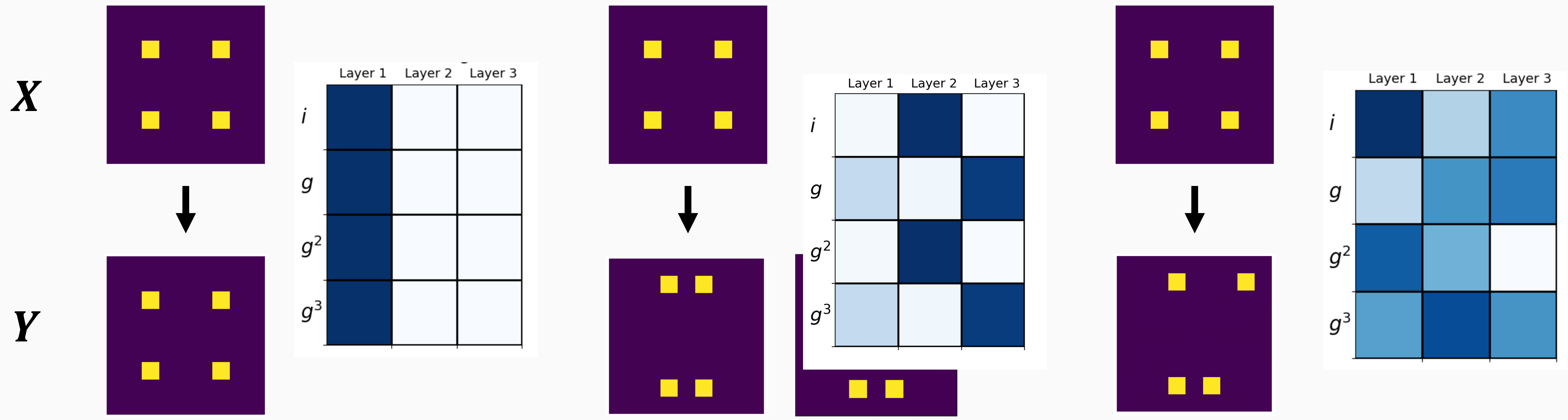


Relaxed Group Convolution Networks



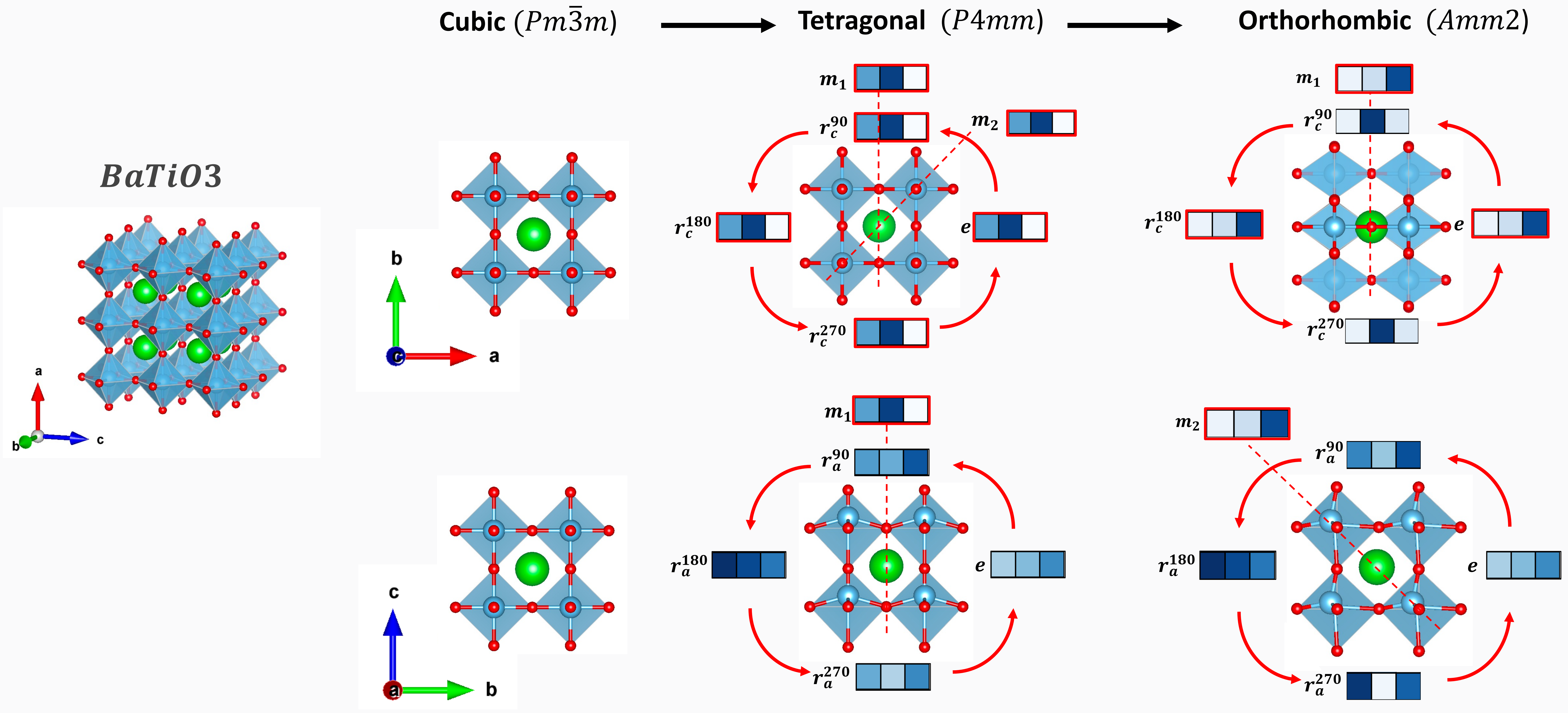
$$[f \star \psi](g) = \sum_{h \in G} f(h)\psi(g,h) = \sum_{h \in G} \sum_{l=1}^{L} f(h) w_l(h) \psi_l(g^{-1}h),$$

Relaxed Group Convolution Networks



✓ **Proposition** (informal): The relaxed weights will learn to be distinct across group elements during training in a way such that the model is equivariant to $Stab(X) \cap Stab(Y)$

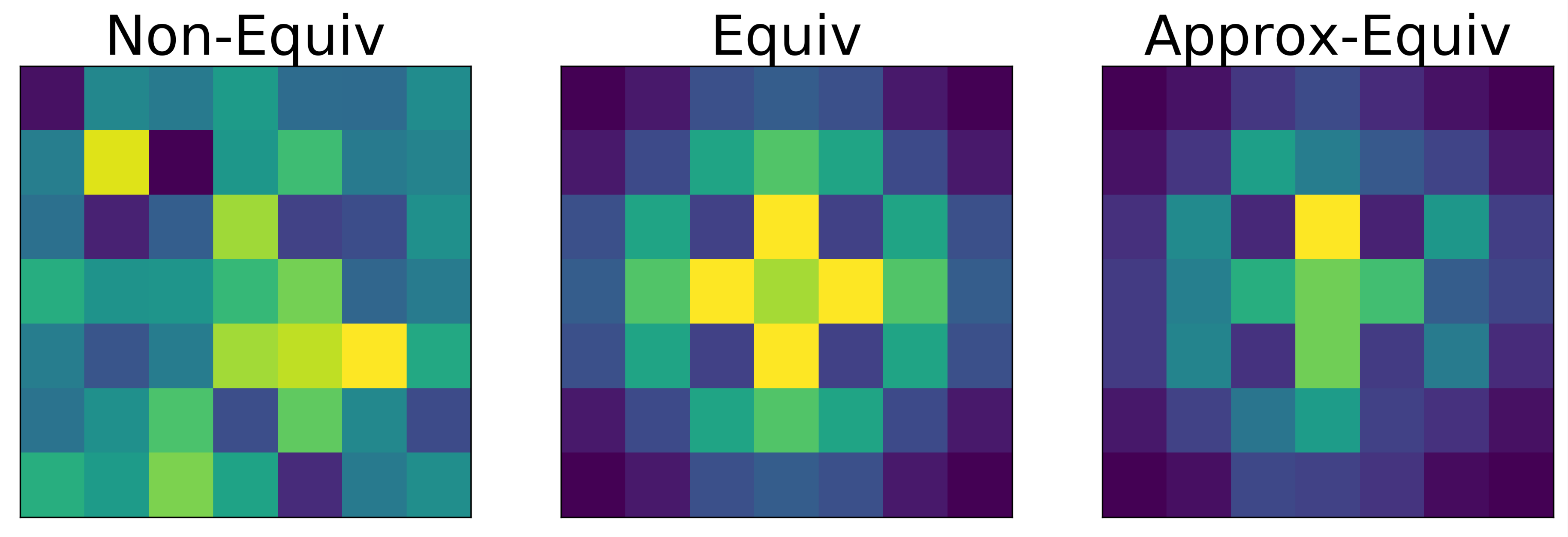
Relaxed Group Convolution Networks



Relaxed Steerable Convolution Network

Steerable Kernels: $\phi(gx) = \rho_{out}(g)\phi(x)\rho_{in}(g^{-1}),\ \forall g \in G$

$$f_{out}(\boldsymbol{x}) = \sum_{\boldsymbol{y}} \sum_{i=1}^{N} (w_i \odot \phi_i(\boldsymbol{y})) f_{in}(\boldsymbol{x}+\boldsymbol{y}) \longrightarrow \sum_{\boldsymbol{y}} \sum_{i=1}^{N} (w_i(\boldsymbol{y}) \odot \phi_i(\boldsymbol{y})) f_{in}(\boldsymbol{x}+\boldsymbol{y})$$



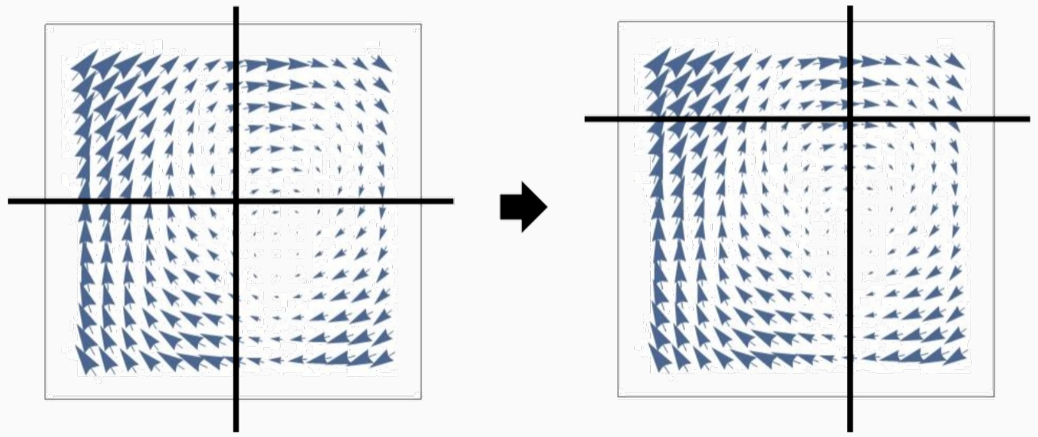
Maurice Weiler , Gabriele Cesa; **General E(2) - Equivariant Steerable CNNs**; NeurIPS 2019

Symmetries of Navier-Stoke Equation

$$\partial \boldsymbol{w}/\partial \boldsymbol{t} + (\boldsymbol{w} \cdot \boldsymbol{\nabla})\boldsymbol{w} = -\mathbf{1}/\boldsymbol{\rho}_0 \boldsymbol{\nabla} \boldsymbol{p} + \boldsymbol{\nu} \Delta \boldsymbol{w} + \boldsymbol{f}$$

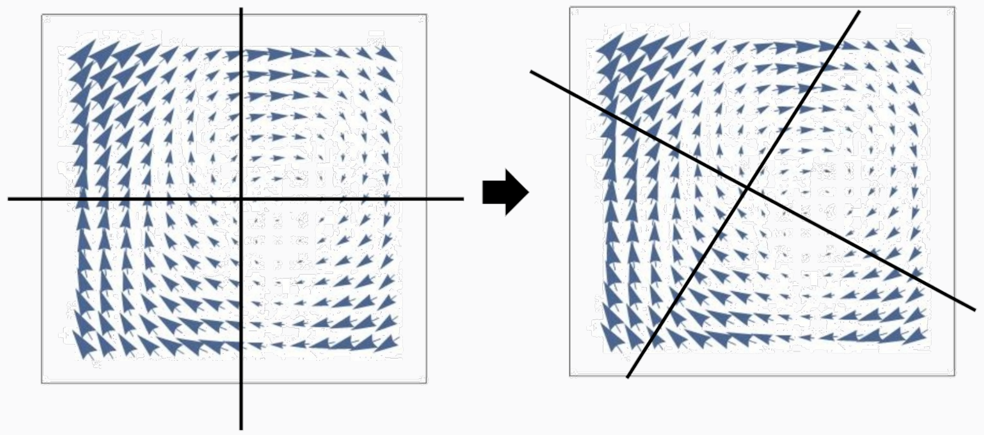
$$T_{\boldsymbol{v}}^{sp}\boldsymbol{w}(\boldsymbol{x}, t) = \boldsymbol{w}(\boldsymbol{x} - \boldsymbol{v}, t), v \in \mathbb{R}^2$$

Translation



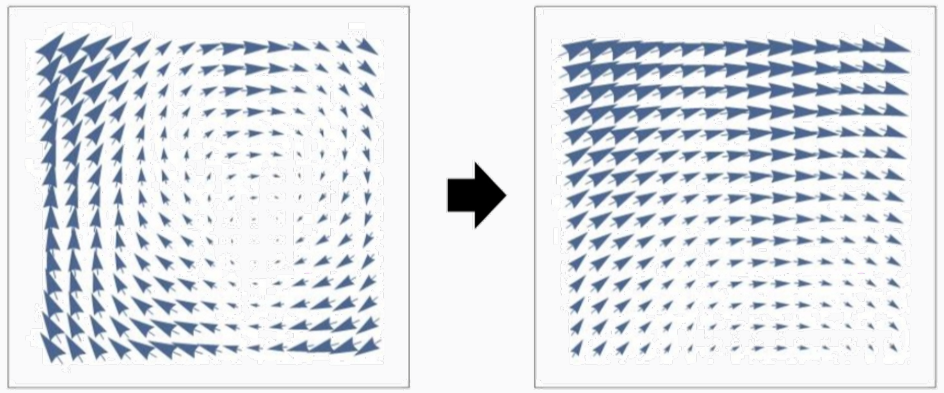
$$T_R^{Rot}\boldsymbol{w}(\boldsymbol{x}, t) = R\boldsymbol{w}(R^{-1}\boldsymbol{x}, t), R \in SO(2)$$

Rotation



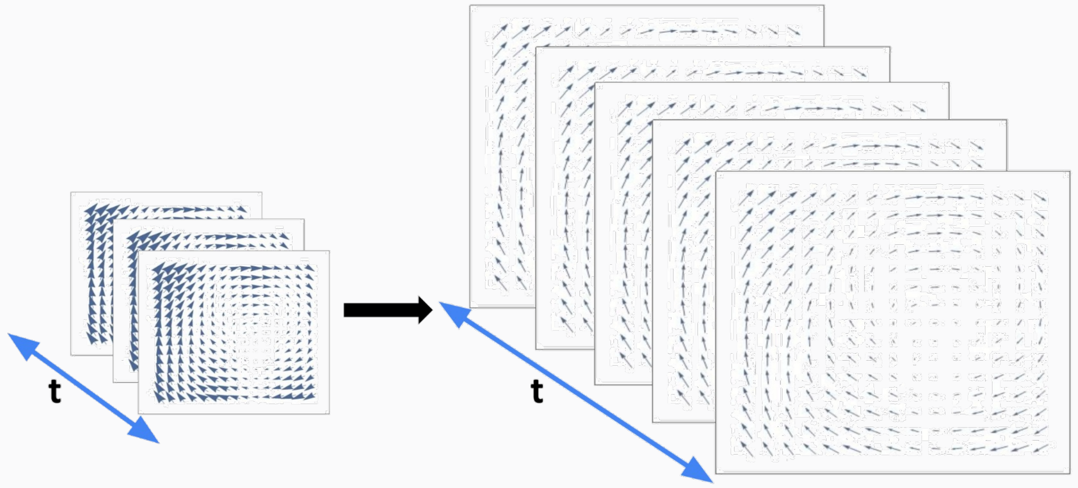
$$T_{\boldsymbol{c}}^{Gal}\boldsymbol{w}(\boldsymbol{x}, t) = \boldsymbol{w}(\boldsymbol{x} - \boldsymbol{c}t, t) + \boldsymbol{c}, \boldsymbol{c} \in \mathbb{R}^2$$

Galilean



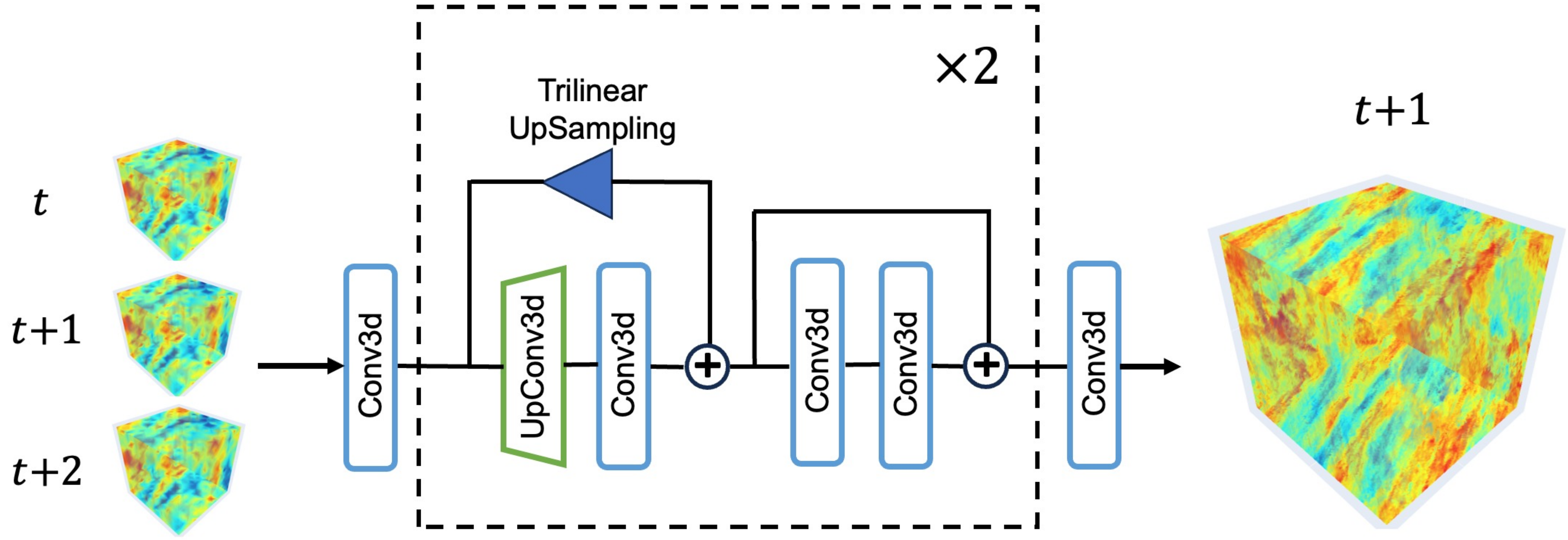
$$T_\lambda^{Scale}\boldsymbol{w}(\boldsymbol{x}, t) = \lambda\boldsymbol{w}(\lambda\boldsymbol{x}, \lambda^2 t), \lambda \in \mathbb{R}_{>0}$$

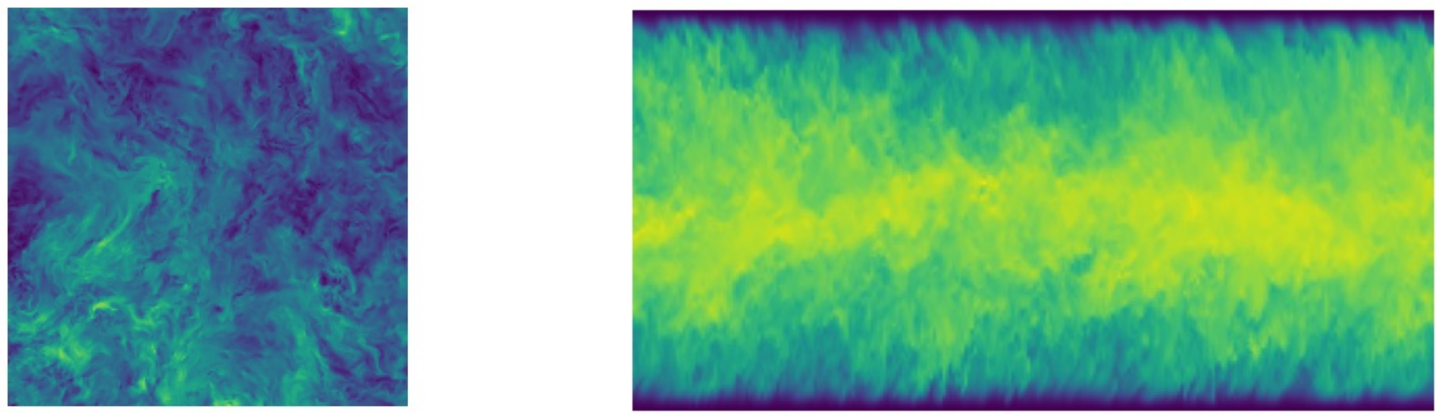
Scaling



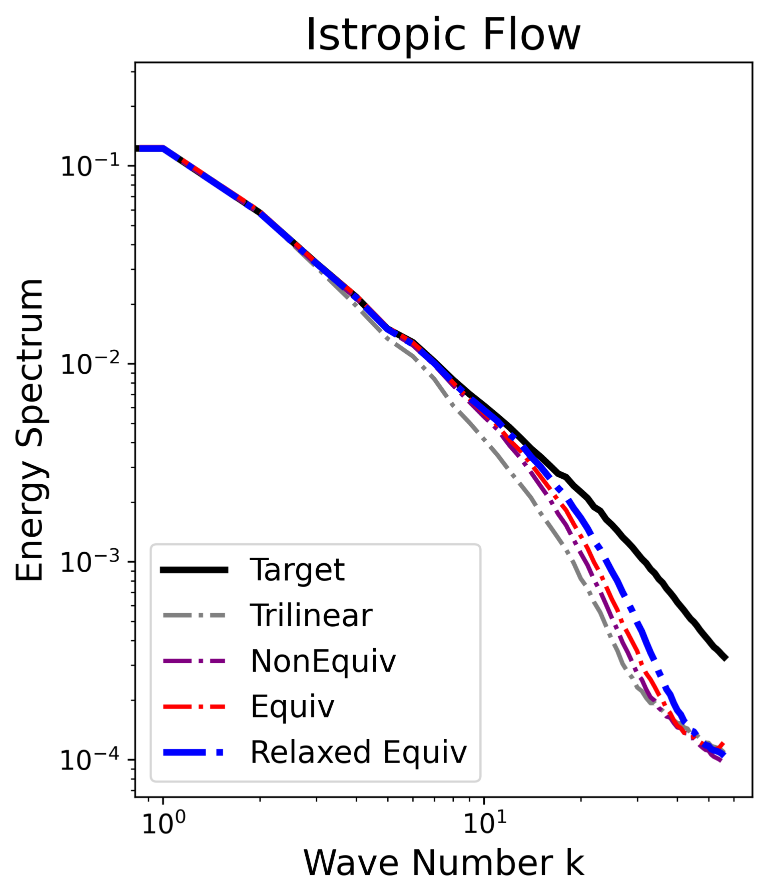
Rui Wang*, Robin Walters*, Rose Yu; Incorporating symmetry into deep dynamics models for improved generalization, ICLR 2021.

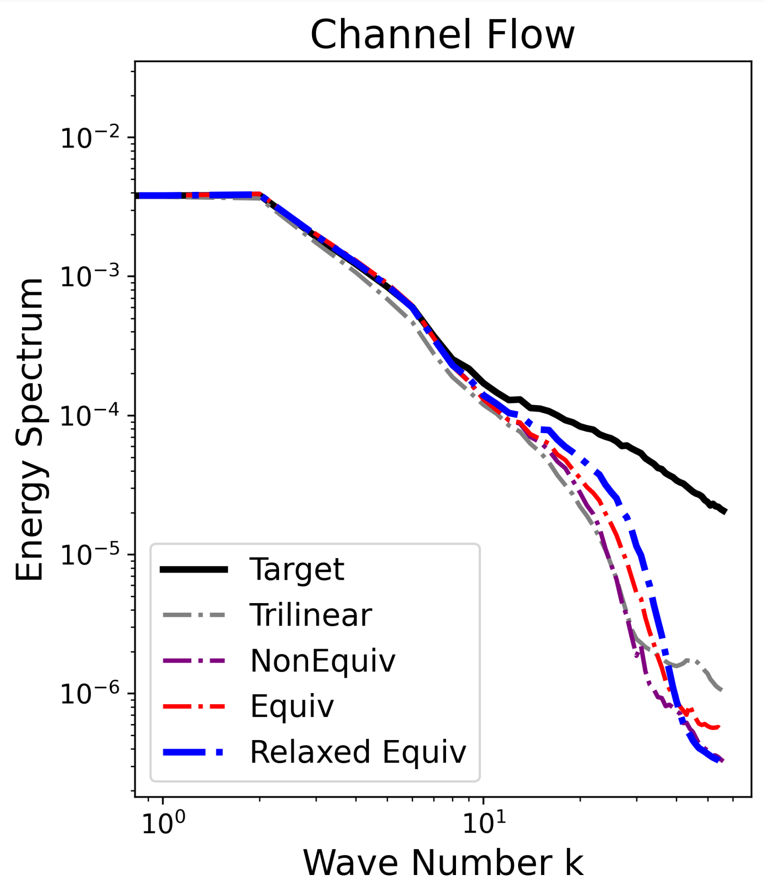
3D Turbulence Super-Resolution

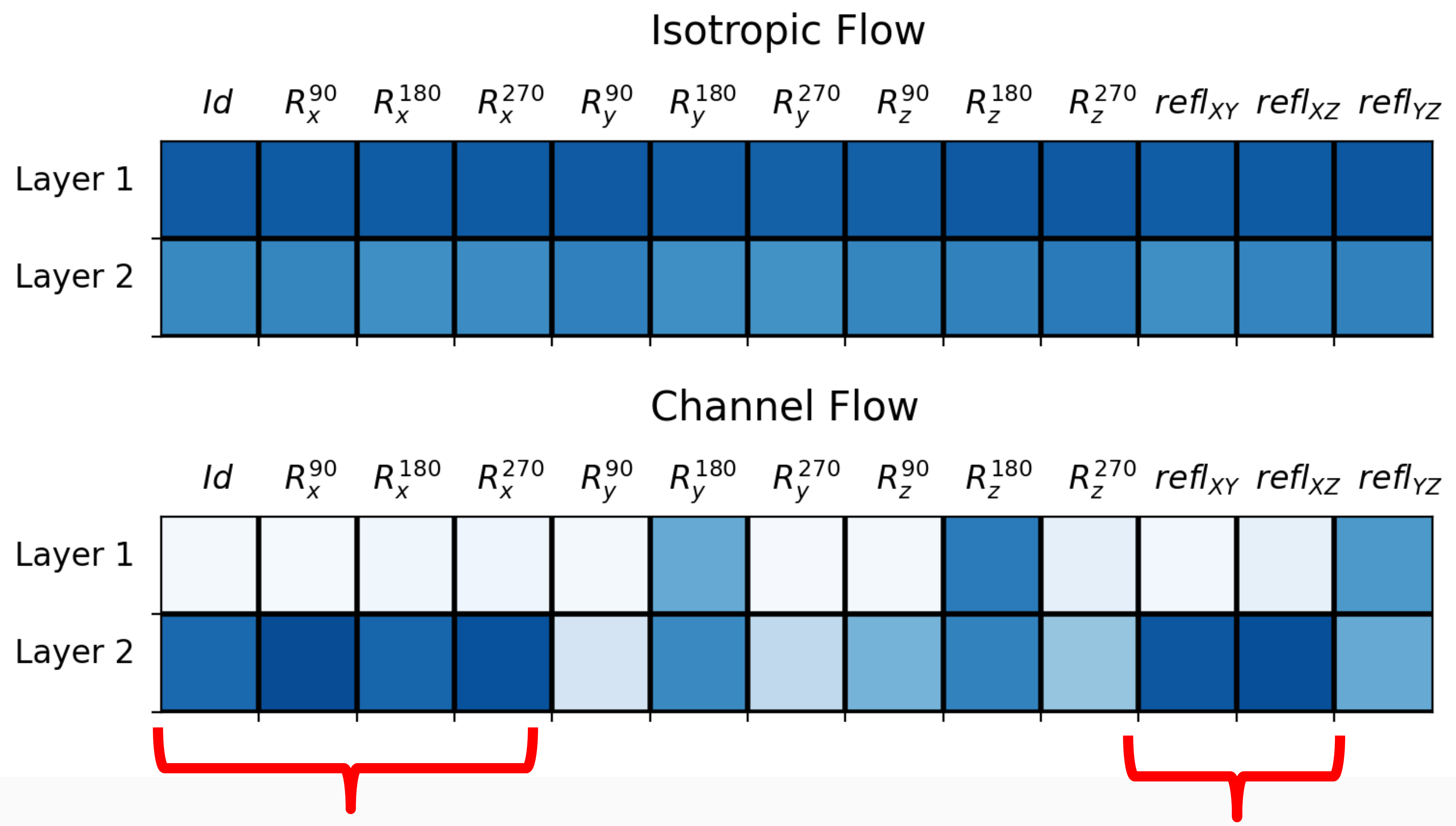




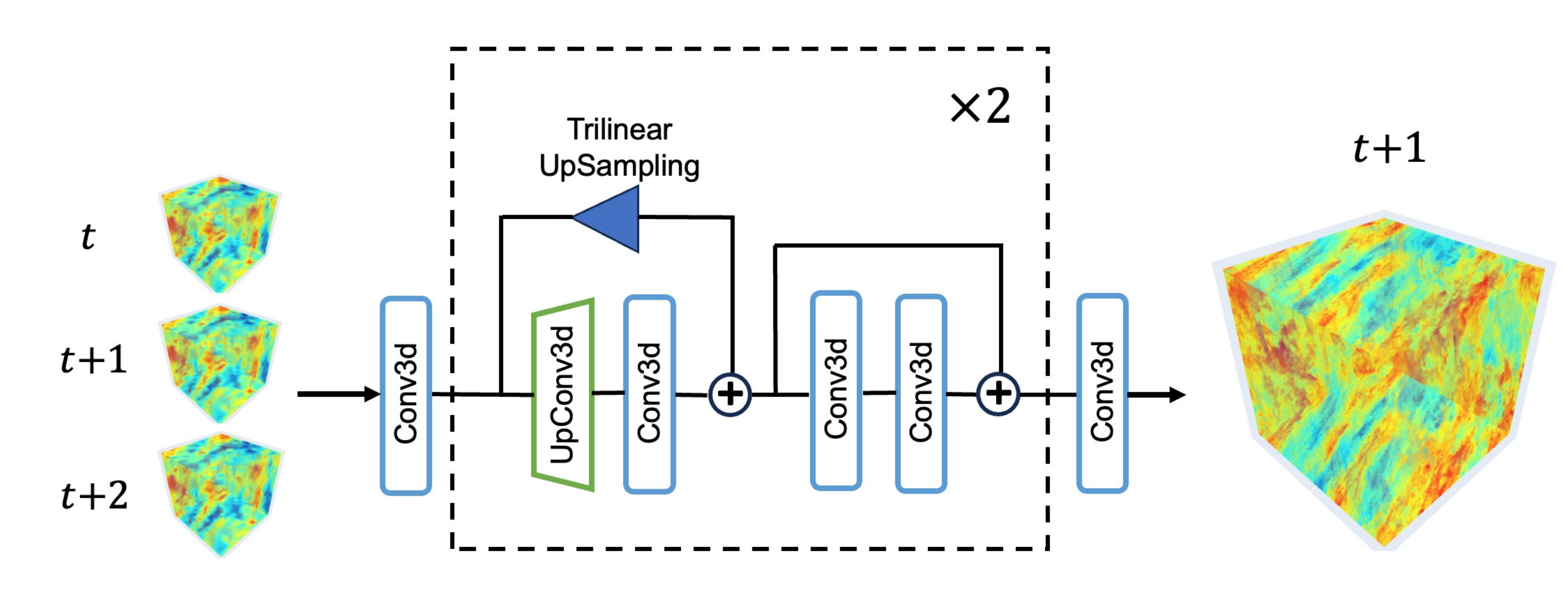
Model	Channel Flow (10^{-2})				Isotropic Flow (10^{-1})			
	TrilinearConv		Equiv	R-Equiv	TrilinearConv		Equiv	R-Equiv
MAE	5.241	2.602	2.540	**2.441**	5.248	1.215	1.119	**1.000**

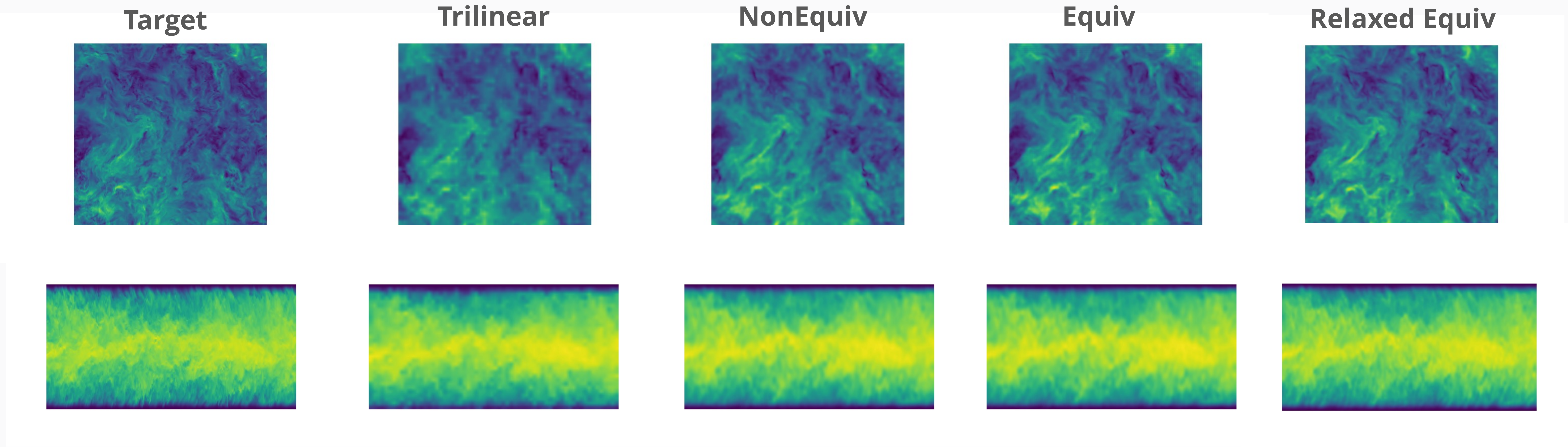






3D Turbulence Super-Resolution

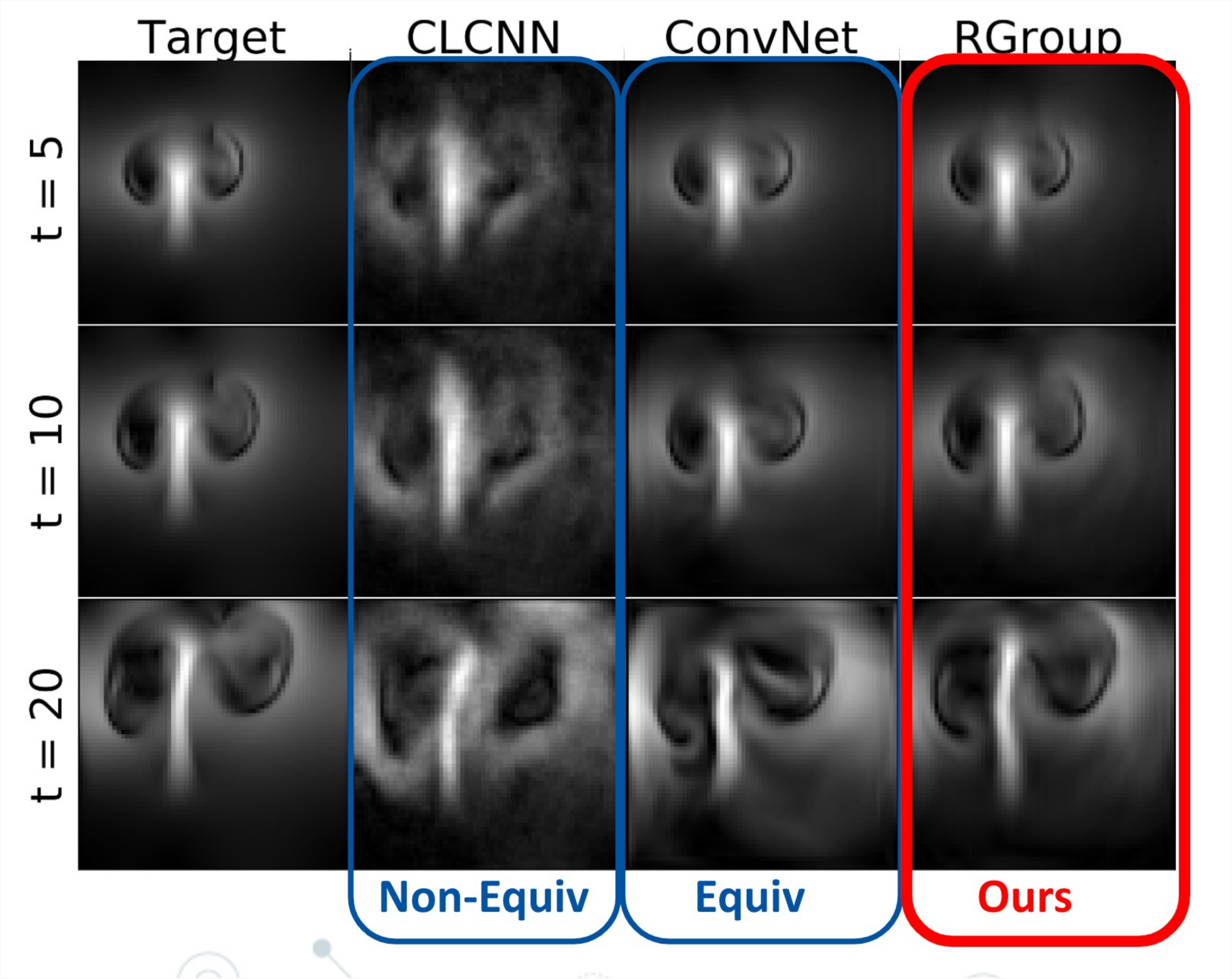




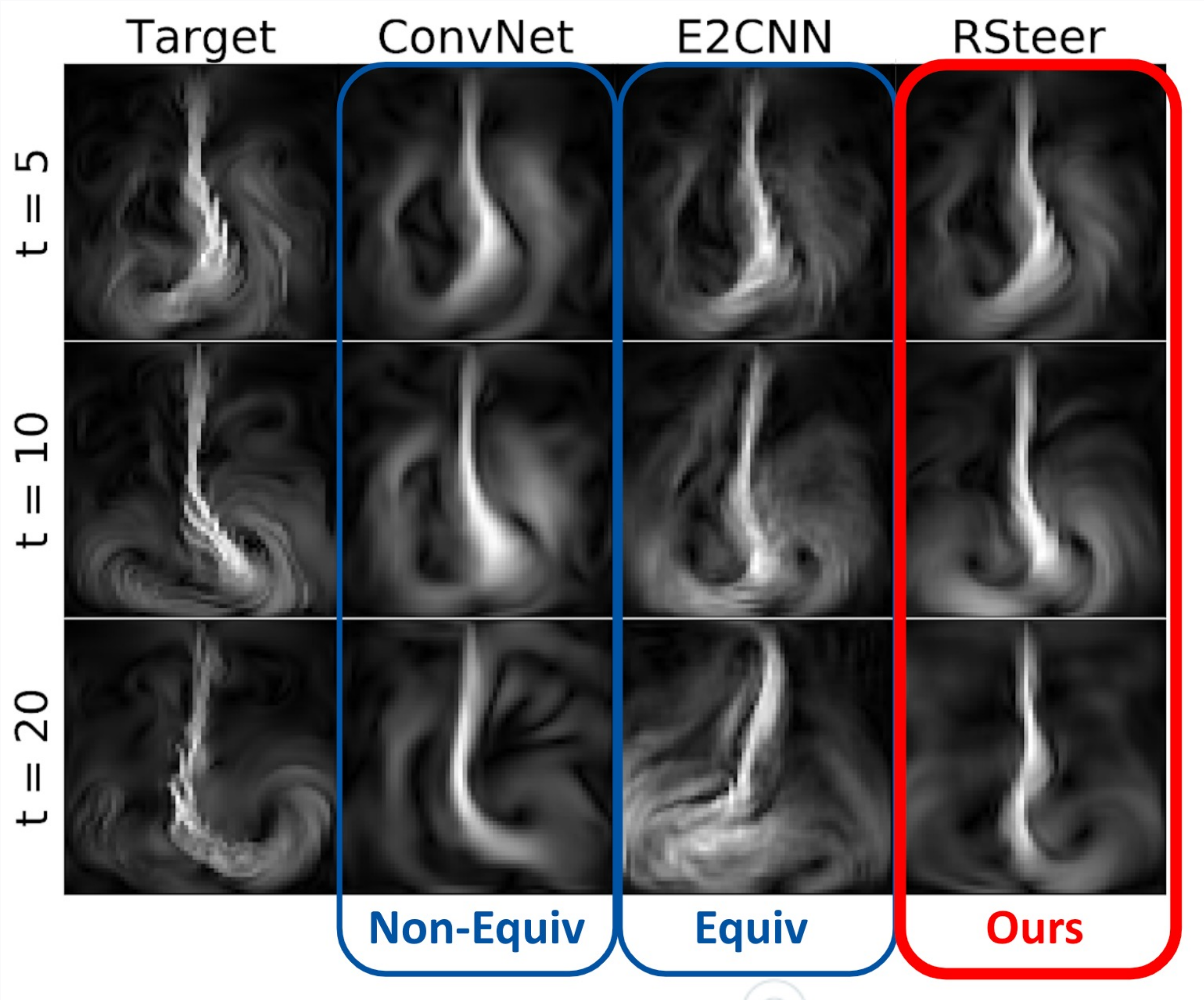
Smoke Plume Simulation

Dynamics Forecasting: $f_\theta(u_{t-q}, \dots, u_t) = \hat{u}_{t+1}, \dots, \hat{u}_{t+h}$

The buoyant forces are different at different subdomains



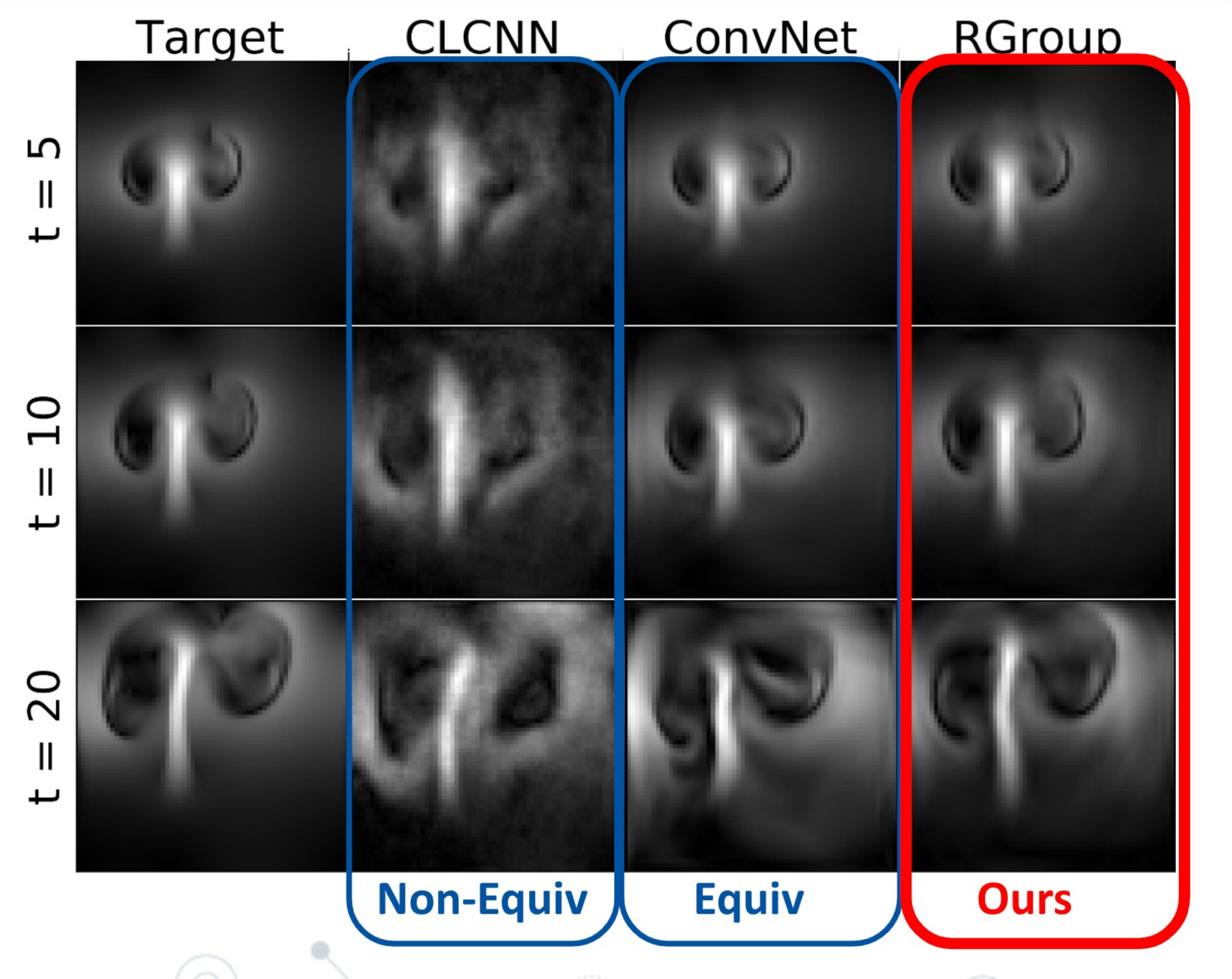
The initial velocities varies with the inflow positions to break the **rotation** symmetry



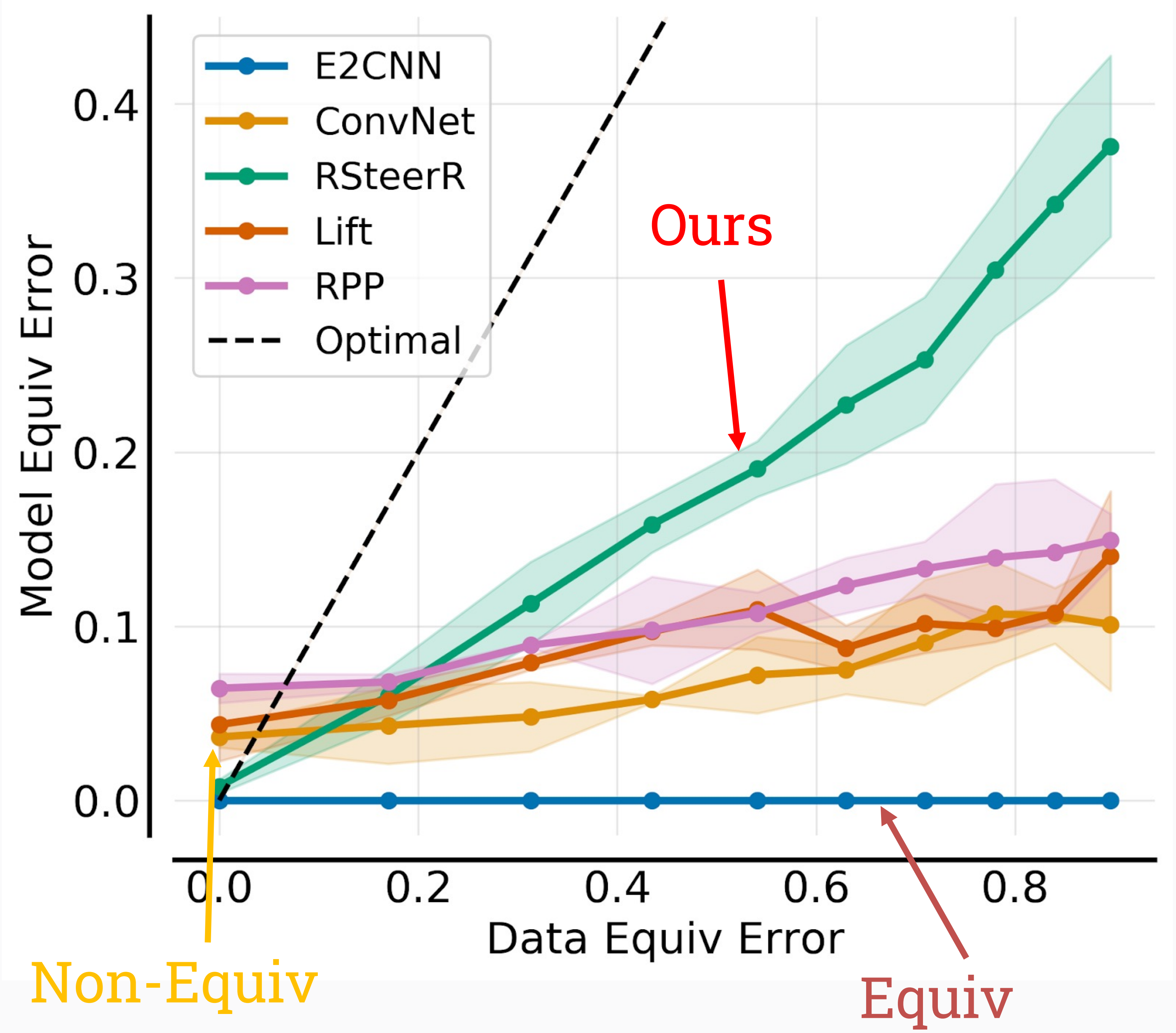
Smoke Plume Simulation

Dynamics Forecasting: $f_\theta(u_{t-q}, \ldots, u_t) = \hat{u}_{t+1}, \ldots, \hat{u}_{t+h}$

The buoyant forces are different at different subdomains



❖ Learn different levels of equivariance



Summary

- ✓ Relaxed group convolution networks always maintain the highest level of equivariance that is consistent with data.
- ✓ The relaxed weights can be used to discover the symmetry and symmetry-breaking factors in the data.
- ✓ Superior performance on turbulence super-resolution and predictions.
- ✓ Future works including investigating the benefits of relaxed weights in optimization and finding more potential in material science.

Acknowledgement

Tess E.Smidt
MIT

Robin Walters
Northeastern University

Rose Yu
UC San Diego

Thank you for your attention!