On rotational equivariance as an inductive bias in machine learning for fluids

Ryley McConkey (rmcconke@mit.edu), Ali Backour, Julia Balla, Elyssa Hofgard, Jigyasa Nigam, Tess Smidt
Research Laboratory of Electronics, MIT

Data Science and Artificial Intelligence Seminar Series, Chalmers University of Technology

14 April 2025

Outline

1. About our research group
2. **Fluid mechanics and equivariance** - overview and motivation
3. **Early results** - example tasks for ML in fluids
    - Superresolution
    - Subgrid-scale closure modelling
    - RANS anisotropy mapping
4. Future and ongoing work

Atomic Architects Research Group

Massachusetts Institute of Technology

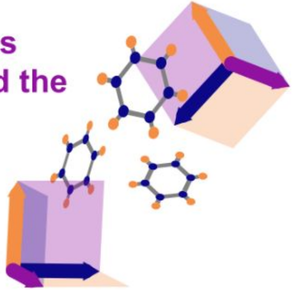
Atomic Architects

Research Group of Prof. Tess Smidt

physics ∩ geometry ∩ machine learning

Spheri-cow Har-moo-nics

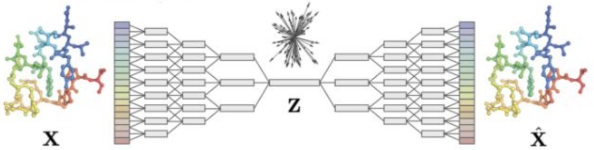
Neural networks that understand the Euclidean symmetry of coordinate systems.



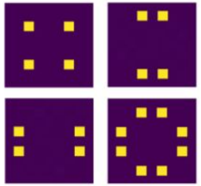
More accurate and better generalizing models with less data. Possess many of same properties as physical systems.

Many applications!

Making big expensive calcs go *zoom* from the atomic to cosmic scale.



Encoding and generating geometry

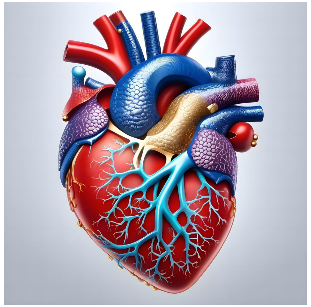


Uncovering symmetry-breaking

Fluid mechanics - overview and motivation








Field goal: understand and predict the behaviour of **liquids**, **gases**, and **plasmas**

Techniques: **experiments** and **numerical simulation**

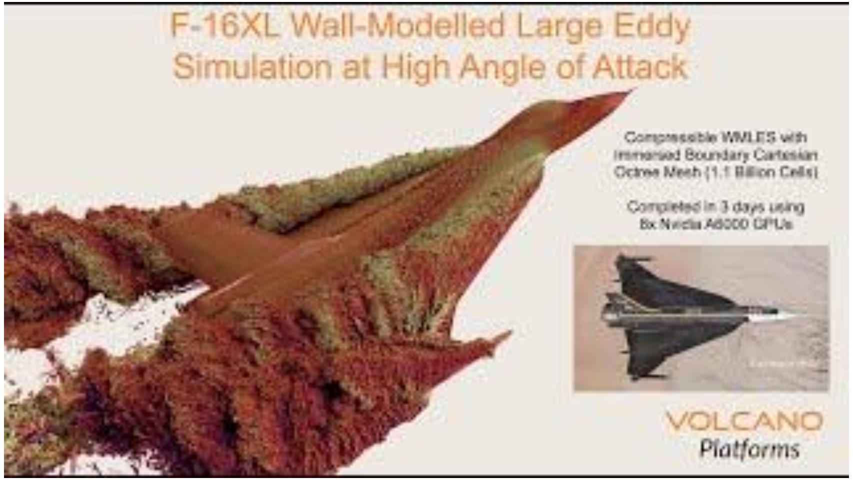
Application domains:

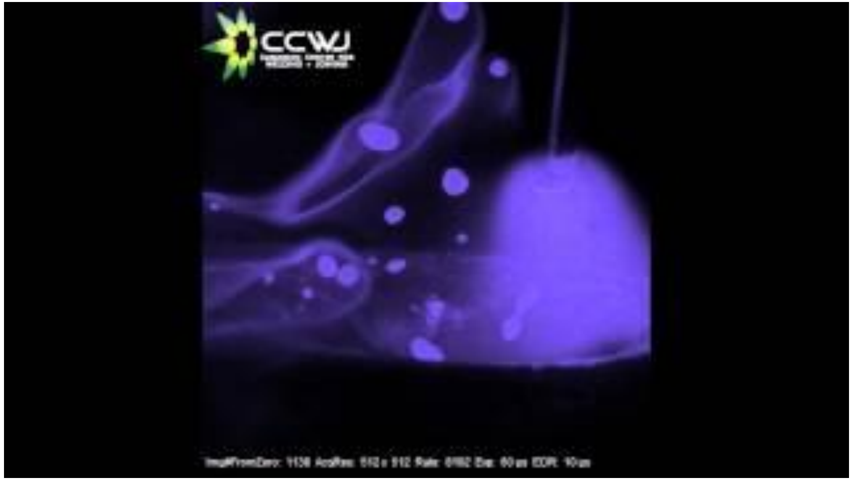
- **Engineering**. Aerospace, automotive, nuclear, chemical, wind
- **Studying nature**. Oceans, weather, astrophysics, flight, swimming
- **Medicine**. Cardiovascular system, respiratory system, drug delivery, spread of contagions

Key problems of practical interest

- Experiments, modelling, and simulation of **turbulence**
- **Multiphysics** - phase changes, multiphase flows, heat transfer, reacting flows

Fluid mechanics - overview and motivation





Fluids: Navier-Stokes

(Incompressible, Newtonian fluid)

ρ = constant $\tau_{ij} = \mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) = 2\mu S_{ij}$

$$\frac{\partial u_i}{\partial x_i} = 0$$

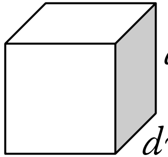
Conservation of mass

(Volume is conserved for a constant density fluid)

$$m a_i = \sum F_i$$

Conservation of momentum

(Newton's second law **per unit volume**)



$$\rho \left(\frac{\partial u_j}{\partial t} + u_i \frac{\partial u_j}{\partial x_i} \right) = -\frac{\partial p}{\partial x_j} + \frac{\partial}{\partial x_i} \left[\mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \right] + \rho g_j$$

Material derivative *Pressure force* *Viscous force* *Body force*

$$\frac{D u_j}{D t} = \frac{\partial u_j}{\partial t} + u_i \frac{\partial u_j}{\partial x_i}$$

3D: 4 equations, 4 fields

Nomenclature

ρ	Mass density (constant)
u_j	Velocity (vector) field
μ	Viscosity (constant)
τ_{ij}	Stress tensor
p	Pressure field
μ	Viscosity (constant)
g_j	Gravity (constant vector)

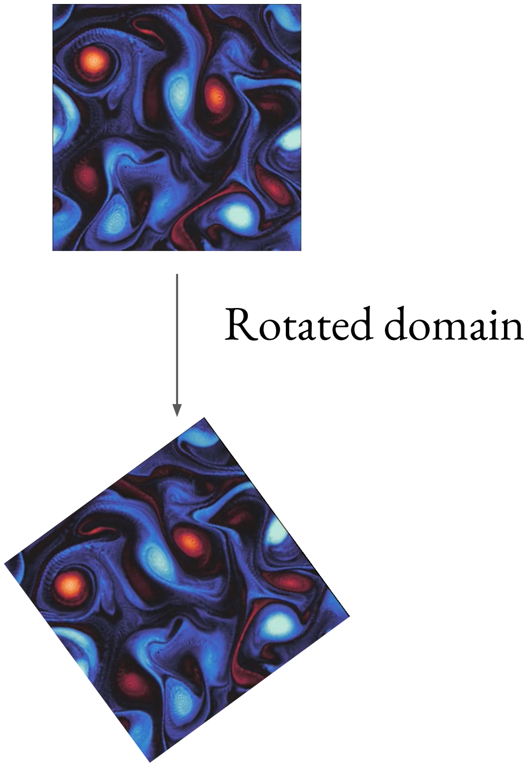
Equivariance

Navier-Stokes equations automatically transform their outputs when the inputs transform (*covariance*)

Should our ML model also generalize to new input orientations/frames?

Equivariance is cared a lot about in ML for:

- Computational chemistry
- Materials design
- Protein modelling
- Geometry



Equivariance

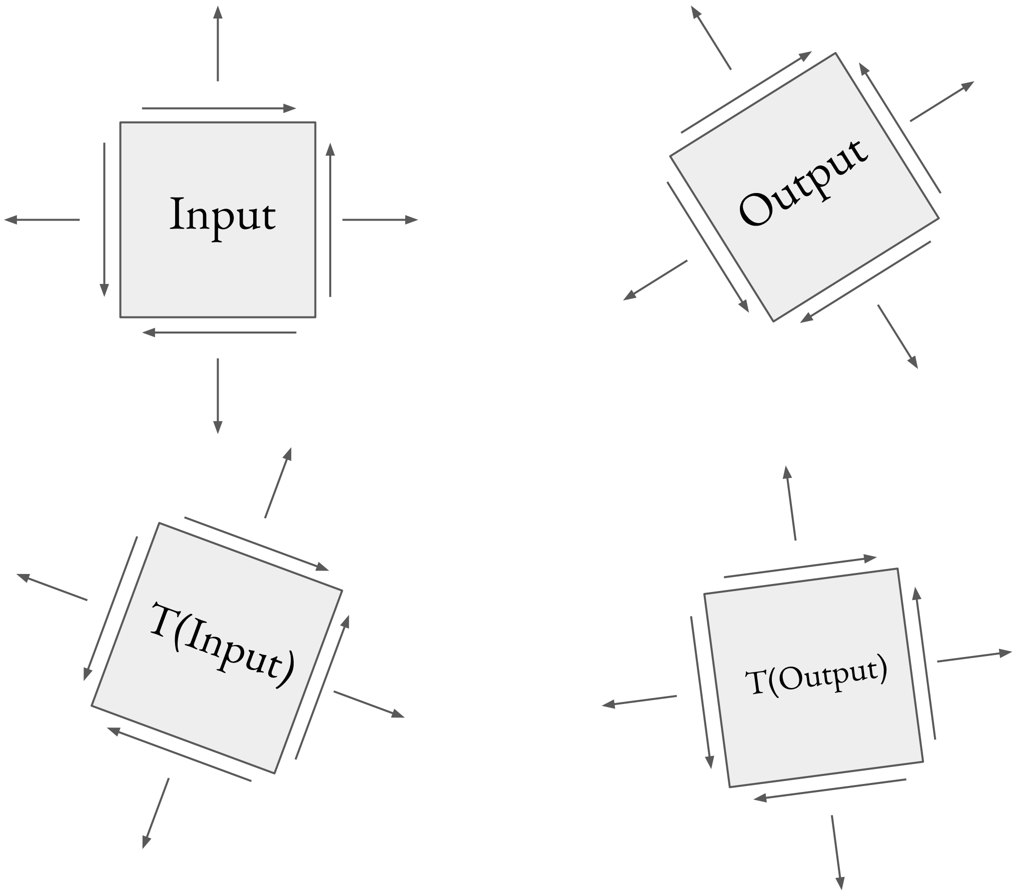
An *equivariant* model automatically transforms its output when the input is transformed.

Relevant transformations (E(3) group):

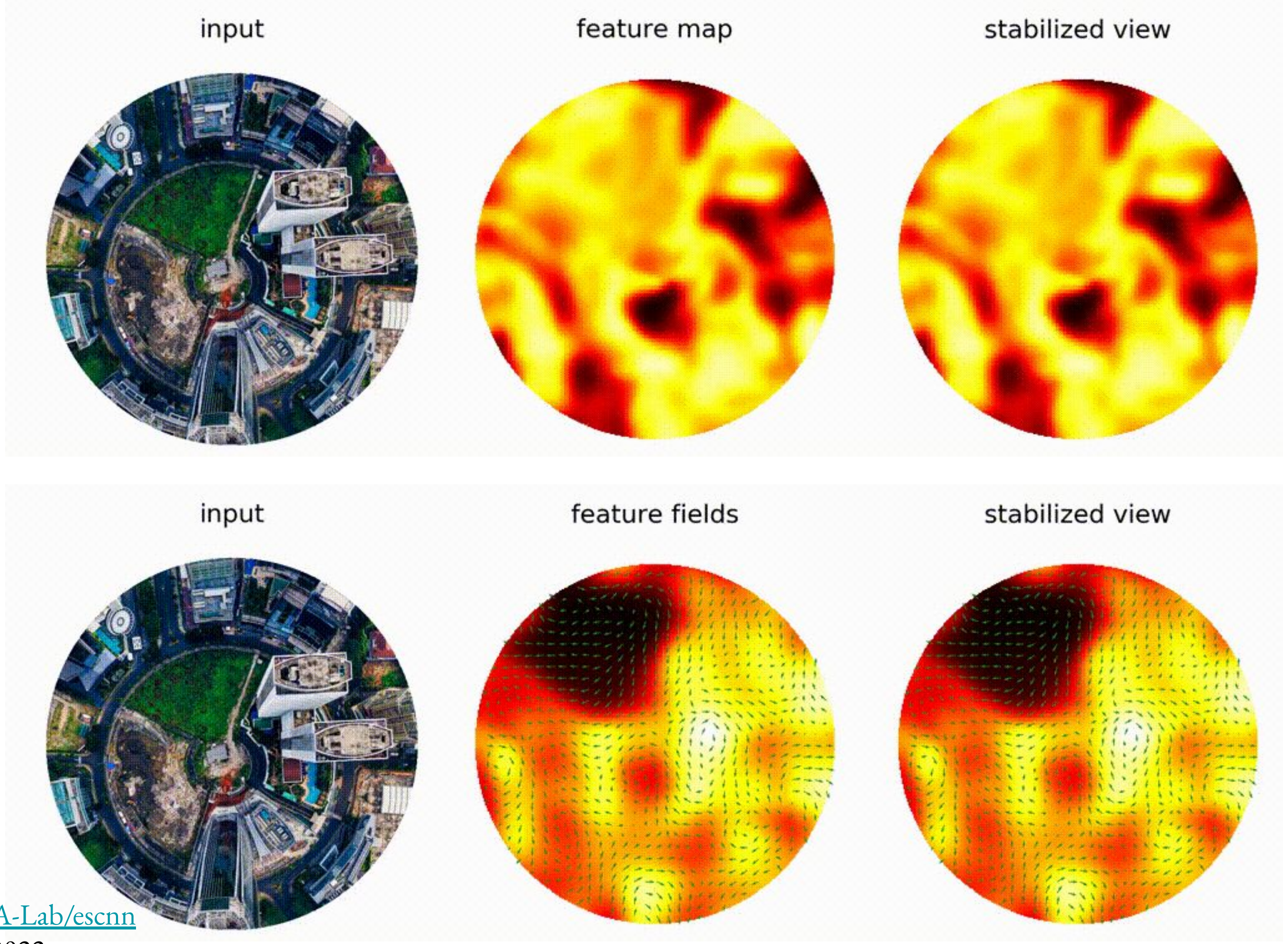
- Translations (automatic with CNNs)
- Rotations
- Reflections
- Inversions

Without equivariance:

- If the input is transformed, the output will not be.

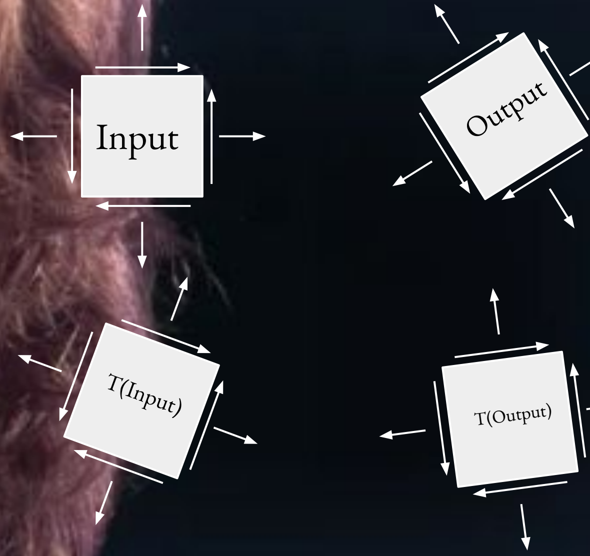


Equivariance



https://github.com/QUVA-Lab/escnn
Cesa, Lang, Weiler ICLR 2022

1. Data augmentation
    a. During training, randomly transform input/output pairs
    b. For fluids - does this happen automatically?
2. Automatically equivariant model (inductive bias)
    a. E.g. e3nn, ESCNN



In fluids, we often don't worry about teaching our models equivariance at all!

Is it possible to learn this power?

Why equivariance?

Active debate between equivariant/non-equivariant models in other domains is ongoing.

Advantages of equivariant models:

- Data efficiency
  - No data augmentation needed
- Automatic encoding/imposition of symmetry
- Model can learn local symmetries

Disadvantages:

- More complicated than your average architecture
- Symmetry is strictly imposed

Selected Tasks

1. Superresolution of a vorticity field
2. Subgrid scale closure modelling
3. Anisotropy mappings for turbulence closure modelling

Goal:

Is equivariance a useful inductive bias in ML for fluids?

Superresolution of vorticity field

A. Backour

J. Balla

Task: Given a coarse resolution image of a flow field, predict a finer resolution

Flow: 2D Kolmogorov Forced Turbulence

Numerics: jax-cfd solver, 256x256 mesh, pseudo spectral solver, Crank Nicholson RK4, first order in time, second order in space, CFL < 0.5

Models: CNN, C_4-Equivariant CNN using escnn

~ 40,000 parameters for each model with 3 convolution blocks + bilinear upsampling

Dataset: Re = [1000, 1500, 2000,... 10000]

Training: Re = [1000, ..., 3500, 7000, ..., 10000]

Test: Re = [5000, 5500]

Super resolution & autoregressive prediction

A. Backour

J. Balla

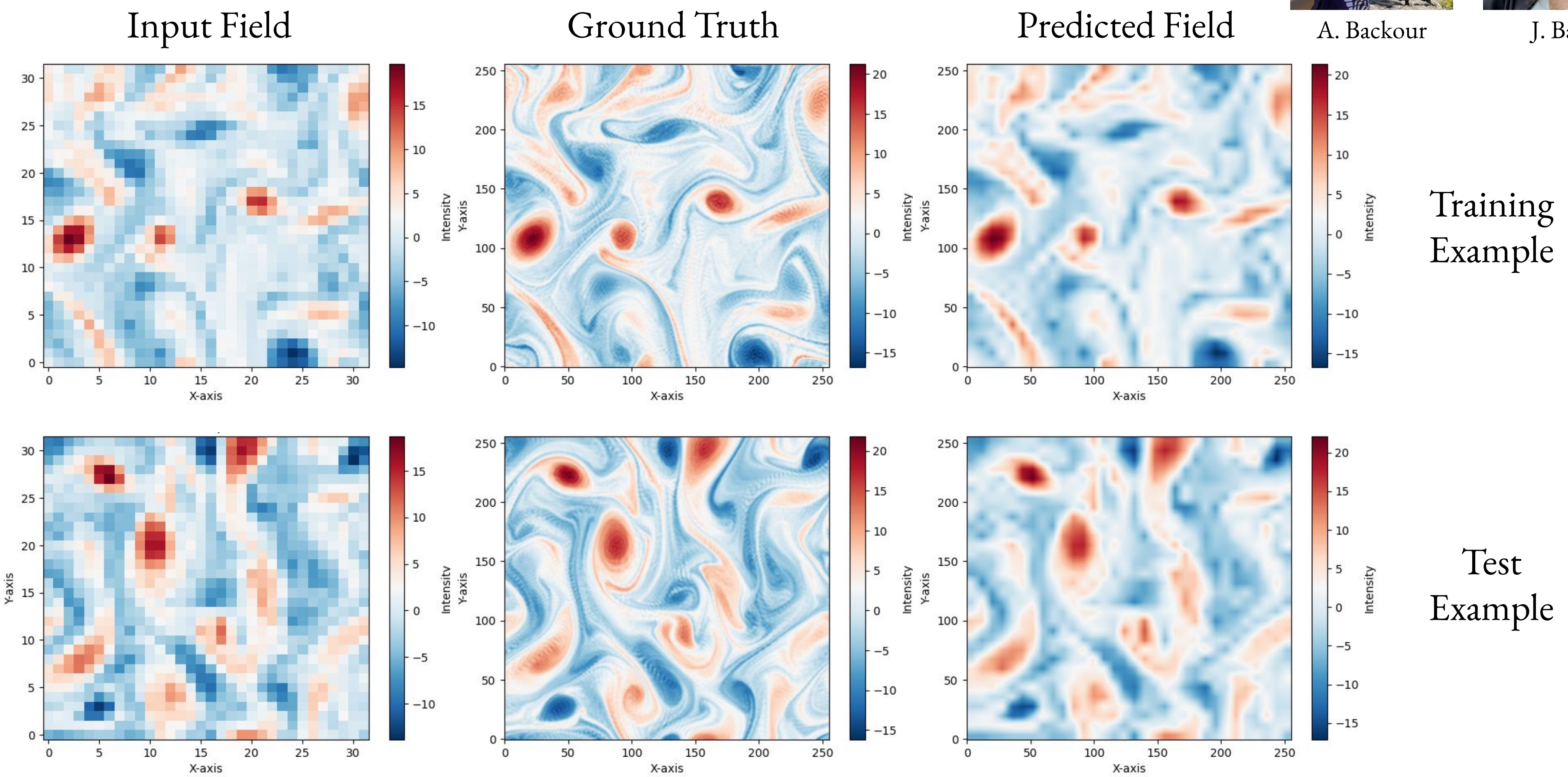
Flow: 2D Kolmogorov turbulence

Tasks:

Super resolution: Predict fine field from coarse field

Autoregressive prediction: Given a time series of images, predict into the future

Superresolution - example output



A. Backour

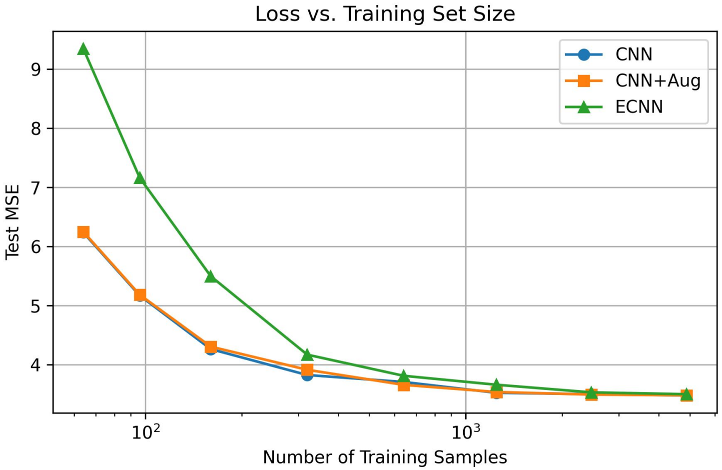
J. Balla

Superresolution - results

A. Backour

J. Balla

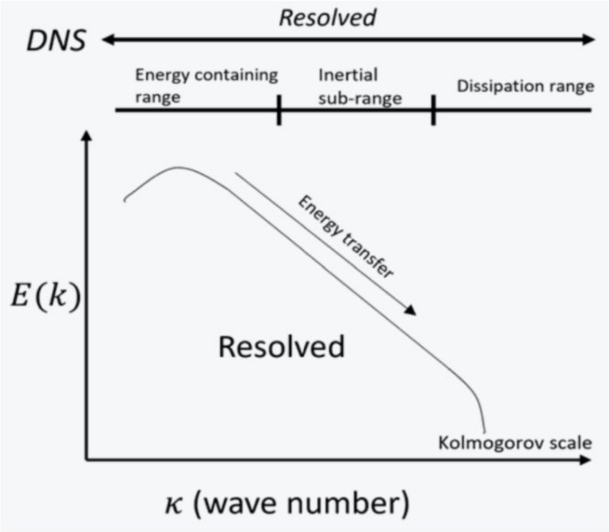
Model	Train MSE	Test MSE	Equivariance Error	Parameters
CNN	3.0760	**3.4738**	0.0447	38624
CNN + Aug	**3.0744**	3.4745	0.0486	38624
ECNN	3.0823	3.4783	$\mathbf{2.113 \cdot 10^{-6}}$	37328



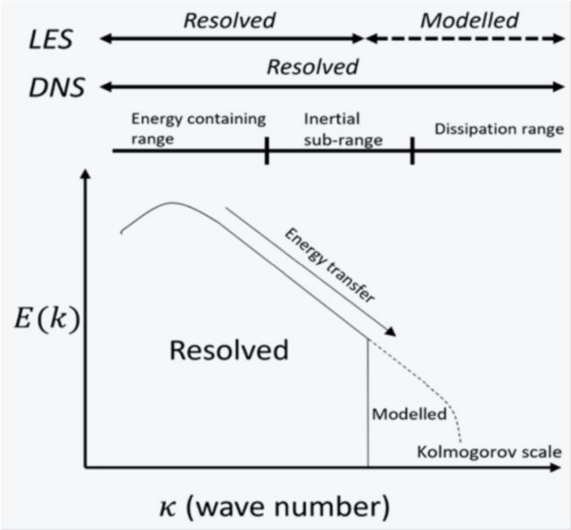
- All models perform similarly on the training and test sets
- "Equivariance error" is not reduced by data augmentation for this task
- We can perform well without completely learning equivariance

Turbulence modelling via machine learning - motivation

DNS – Direct Numerical Simulation



LES – Large Eddy Simulation



RANS – Reynolds Averaged Navier-Stokes

(All modelled)

Subgrid scale turbulence modelling

E. Hofgard

Task (Regression): Predict the subgrid scale stress tensor in terms of resolved tensors

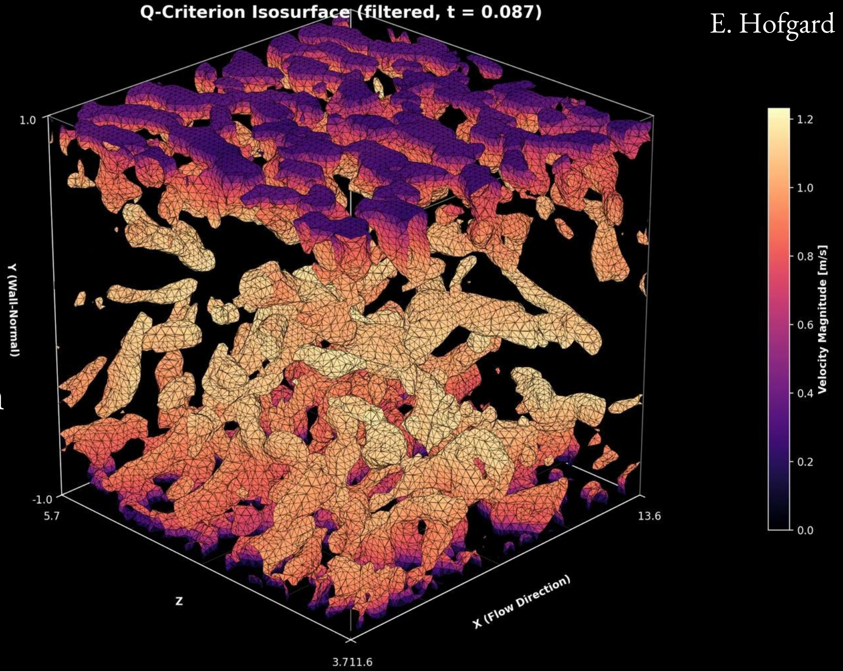
Flow: Turbulent channel flow

Dataset: Johns Hopkins Turbulence Database, $Re_\tau \sim 1000$

Models: 3D CNN, 3D equivariant CNN using [e3nn]

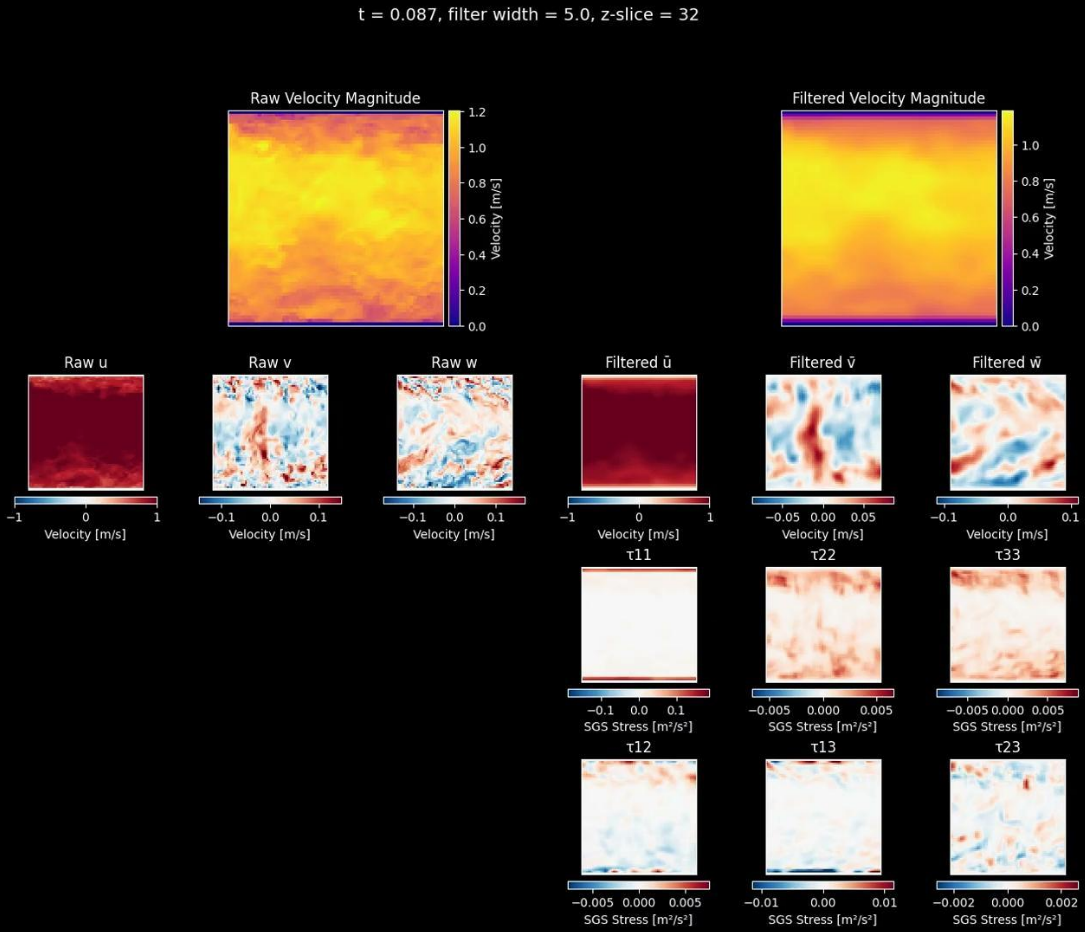
- ~200,000 parameters for each model with 3 convolution blocks

Training/Val/Test: 70/20/10 training/validation/testing split with randomly selected timesteps



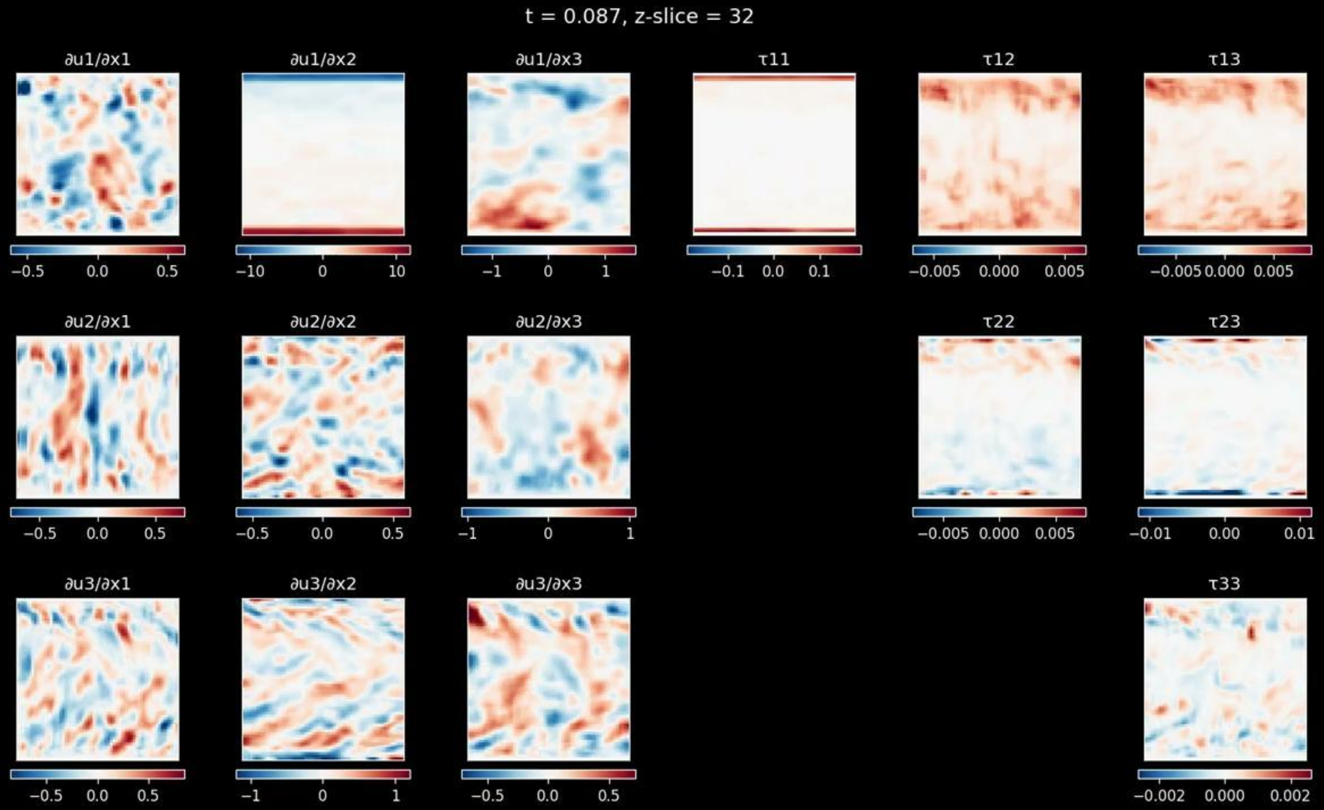
Subgrid scale turbulence modelling

E. Hofgard



Subgrid scale turbulence modelling

E. Hofgard



Input tensor

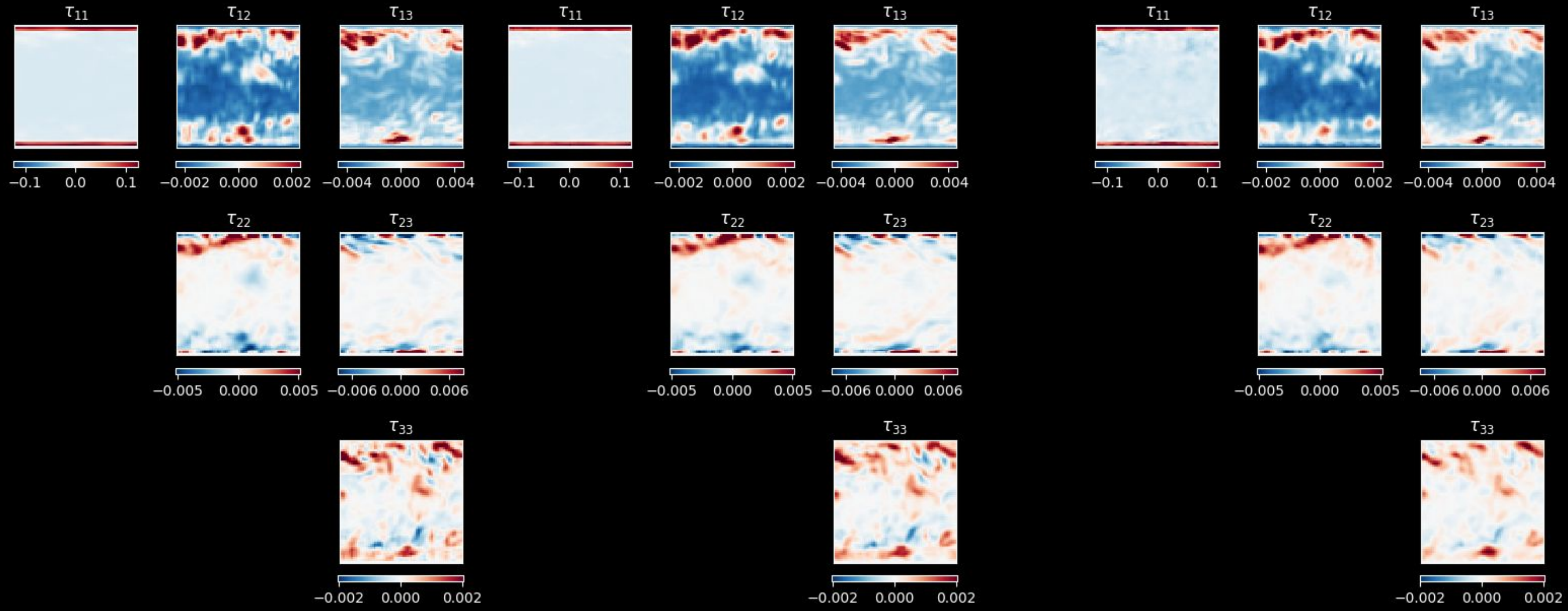
Output tensor

Preliminary Results: subgrid scale turbulence modelling

- Equivariance is not needed to capture large-scale patterns

E. Hofgard

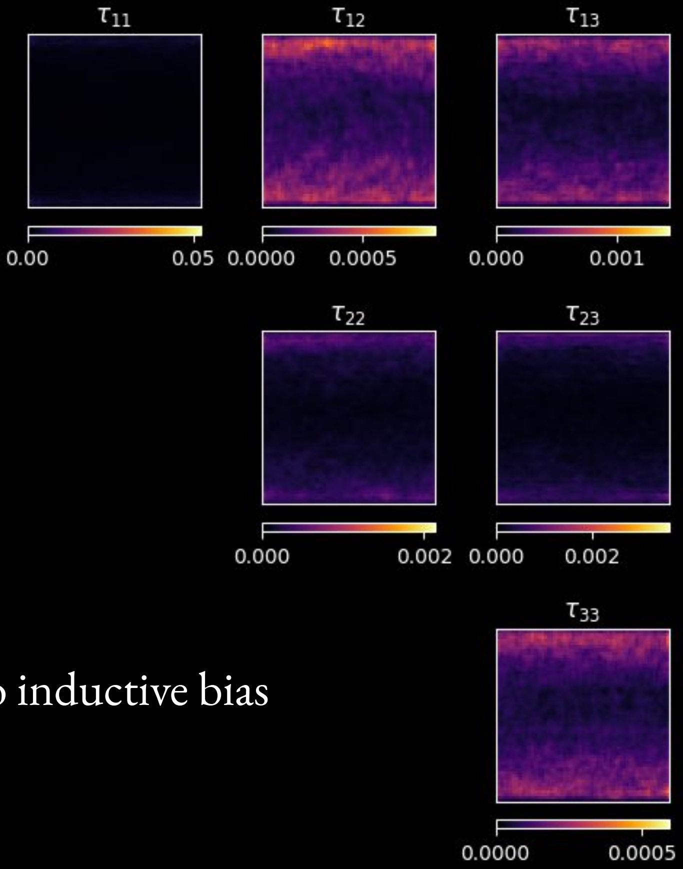
Truth | Prediction (no inductive bias) | Prediction (w/ inductive bias)



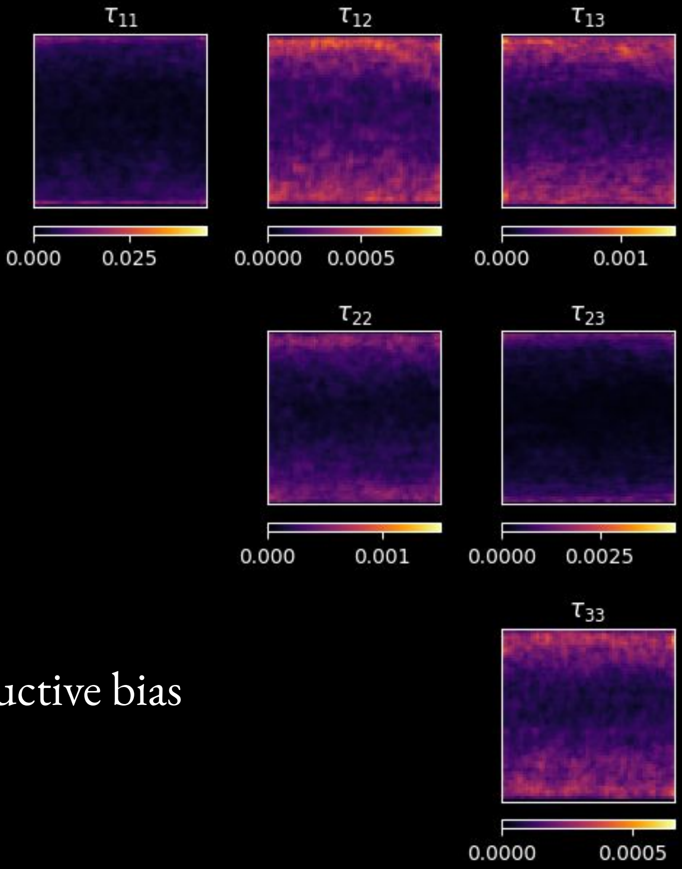
Preliminary Results: subgrid scale turbulence modelling

E. Hofgard

(Test set MAE per pixel, horizontal slice through centre of domain)



No inductive bias



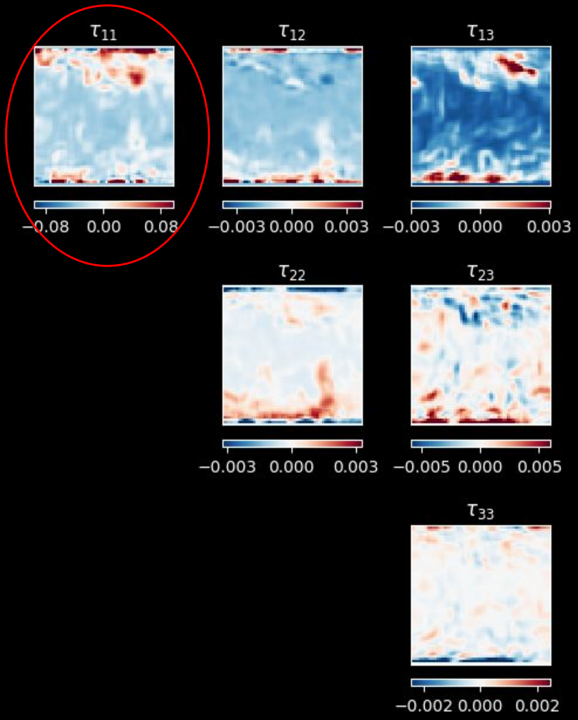
Inductive bias

Preliminary Results: subgrid scale turbulence modelling

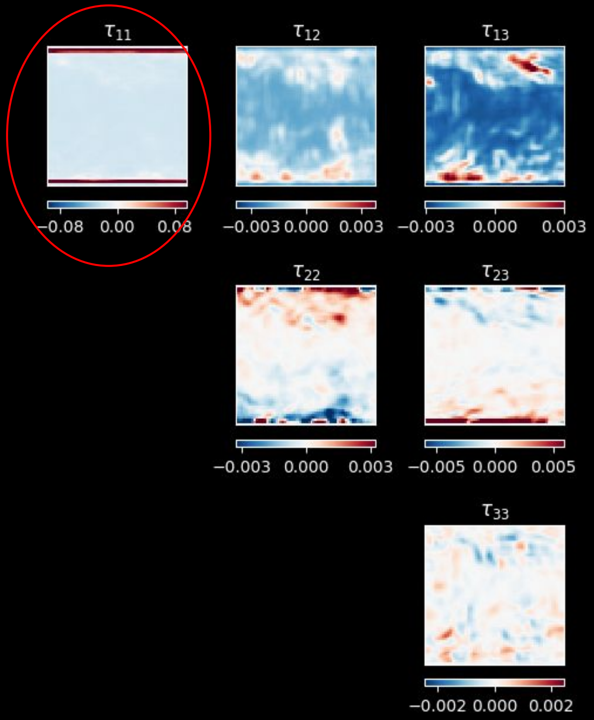
E. Hofgard

- Generalization (covariance) test after rotation of the input tensor

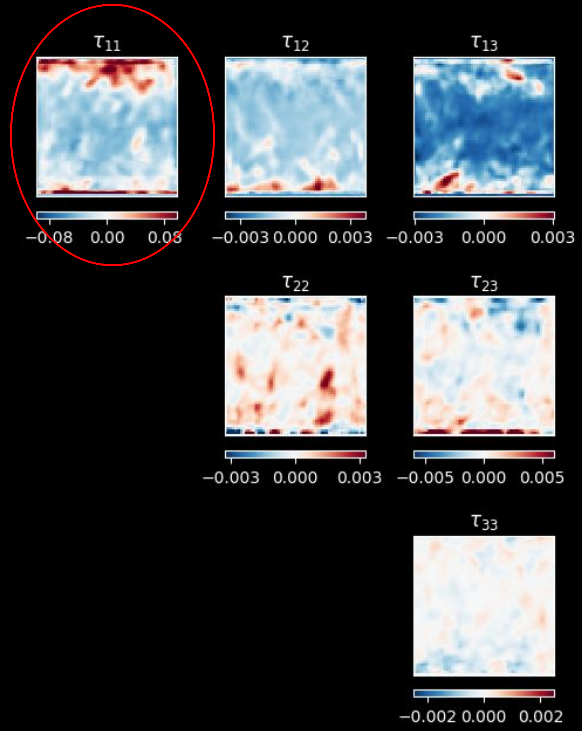
Truth



Prediction (no inductive bias)

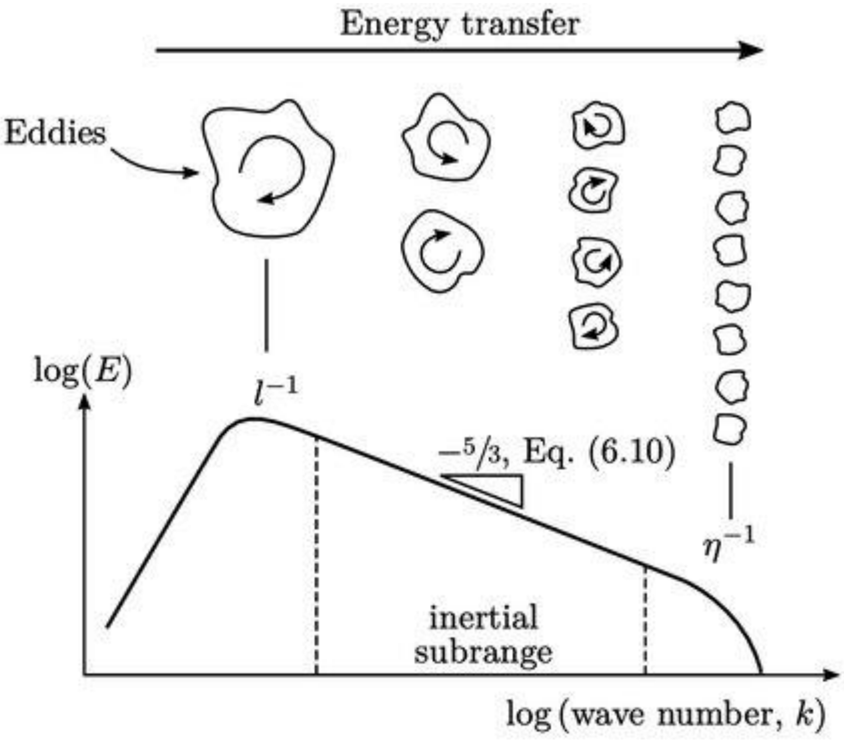


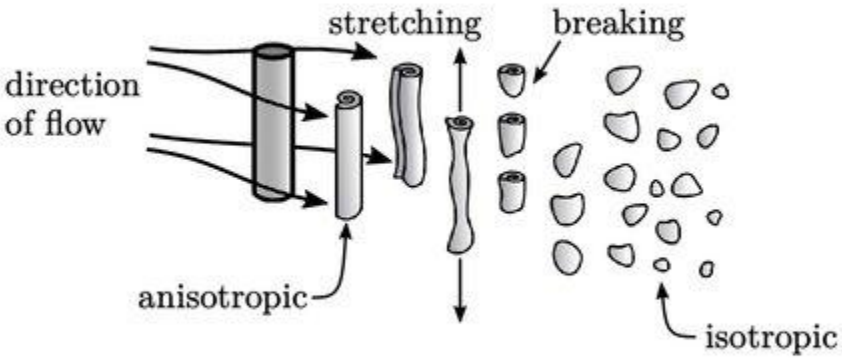
Prediction (w/ inductive bias)



Energy cascade and isotropy

Notes on Computational Fluid Dynamics: General Principles, Greenshields & Weiler





Is your model currently equivariant?

Distributional symmetry

Statistical homogeneity/isotropy of the dataset

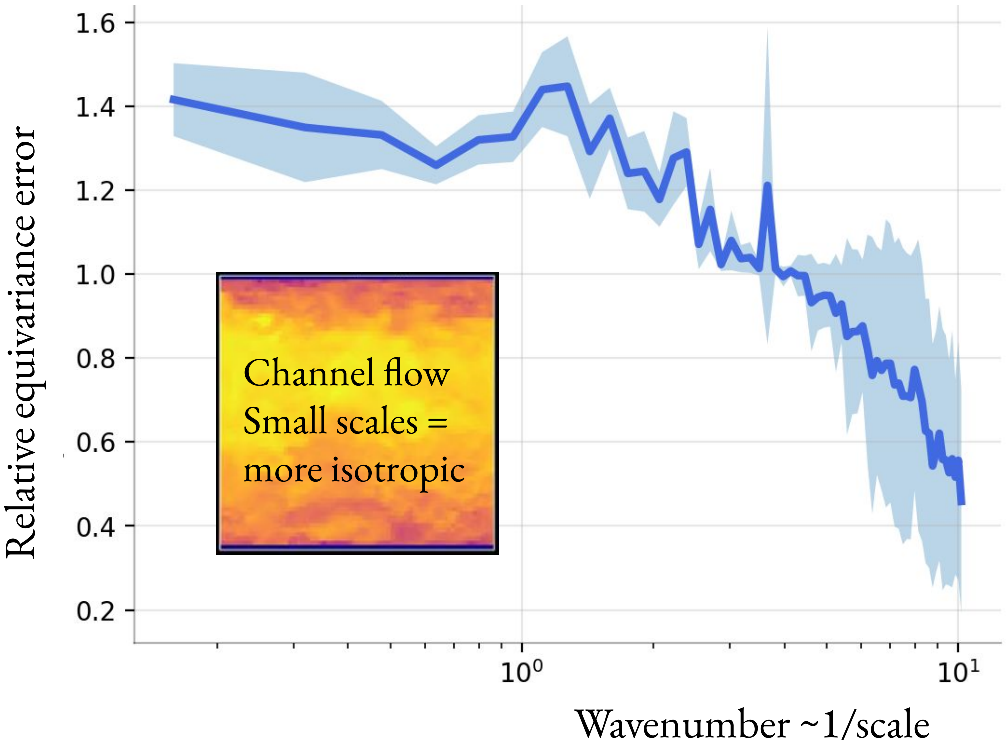
- In certain directions
- At smaller scales (Kolmogorov hypothesis)

(individual frames are inhomogeneous/anisotropic)

Result: *scale-dependent rotational data augmentation*

- Our models might be learning equivariance, but *only at the small scales* due to this Kolmogorov hypothesis-based data augmentation.

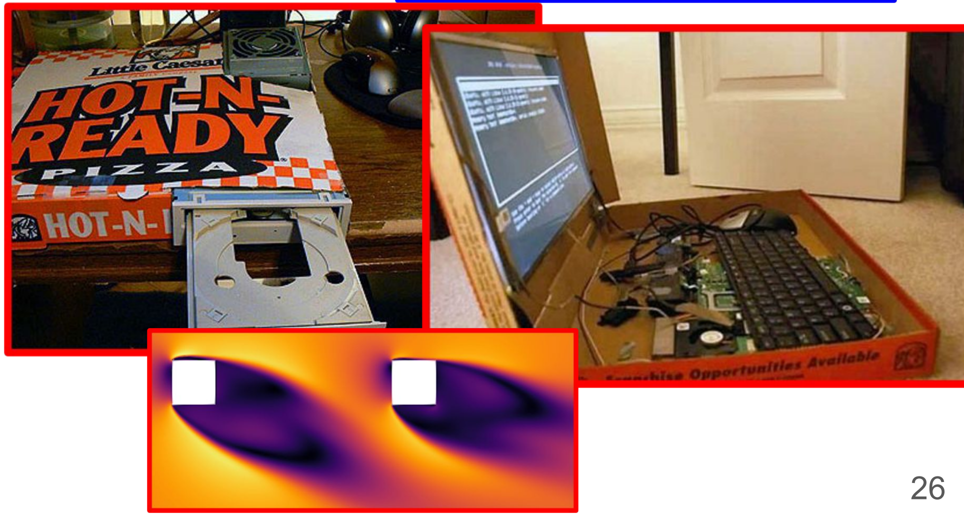
E. Hofgard



RANS Anisotropy Mappings

$$\frac{D\vec{u}}{Dt} = -\nabla p + \nu \nabla^2 \vec{u}$$

$$\frac{D\vec{U}}{Dt} = -\nabla P + \nu \nabla^2 \vec{U} - \nabla \cdot \tau$$



RANS Closure Problem

NS $$\nabla \cdot (\vec{u}\vec{u}) = -\nabla P + \nu \nabla^2 \vec{u}$$

RANS $$\nabla \cdot (\vec{U}\vec{U}) = -\nabla p + \nu \nabla^2 \vec{U} - \nabla \cdot \tau$$

Closure problem: finding a relationship for τ as a function of $\vec{U}$, p

RANS Anisotropy Mappings

- Recent work by others shows that equivariance is a useful inductive bias for predicting closure tensors
- We are working on reproducing and extending these results

RESEARCH ARTICLE | FEBRUARY 07 2025

Implicit modeling of equivariant tensor basis with Euclidean turbulence closure neural network

Grzegorz Kaszuba; Tomasz Krakowski; Bartosz Ziegler; Andrzej Jaszkiewicz; Piotr Sankowski

Check for updates

+ Author & Article Information

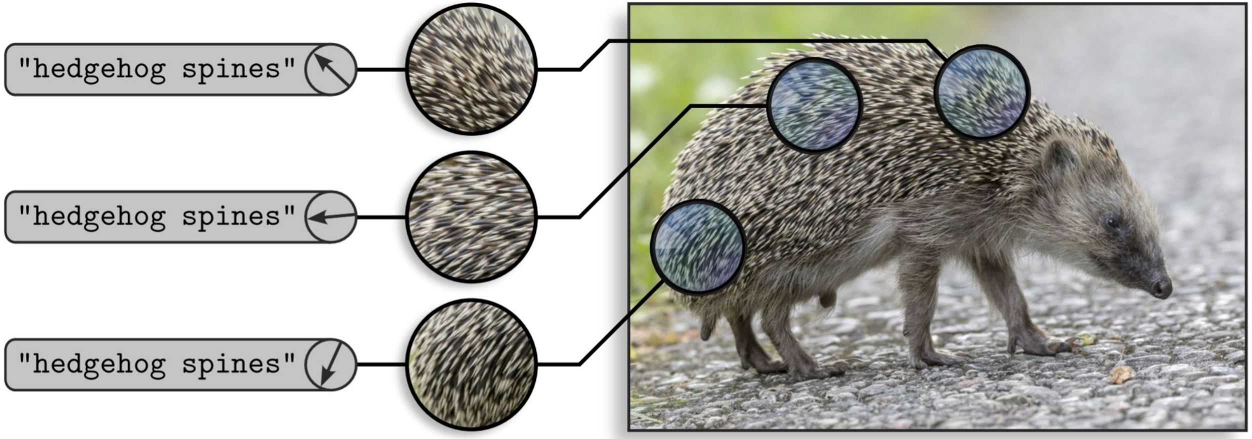
Physics of Fluids 37, 025137 (2025)
https://doi.org/10.1063/5.0249490 **Article history**

TABLE I. Mean squared error of the proposed e3nn-based turbulence model (e3TM) in the prediction of optimal eddy viscosity and nonlinear part of Reynolds stress as proposed by McConkey, compared to a dense network performance reported in his paper[17].

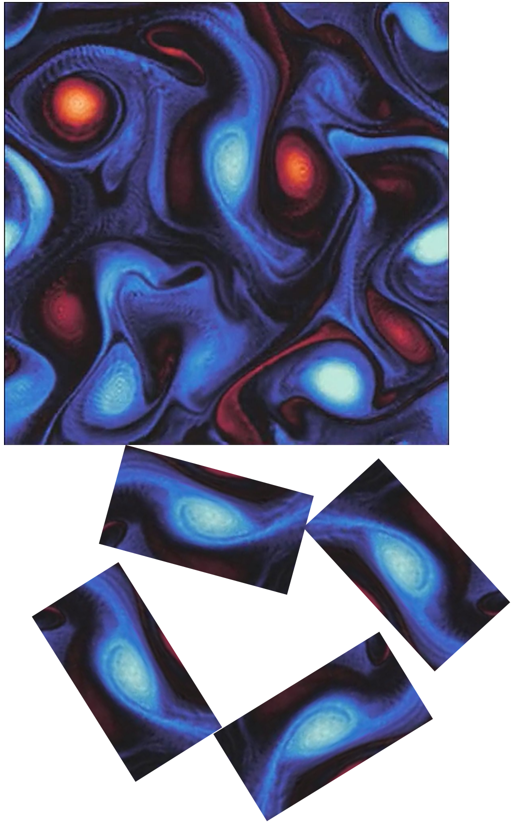
Quantity	e3TM MSE	Dense[17] MSE	Improvement
$a^{\perp}_{xx}$	4.96×10^{-13}	2.31×10^{-12}	78.3%
$a^{\perp}_{xy}$	2.86×10^{-13}	9.40×10^{-13}	69.6%
$a^{\perp}_{yy}$	2.84×10^{-13}	9.56×10^{-13}	70.3%
$\nu^{\dagger}$	1.34×10^{-7}	1.89×10^{-7}	29.1%

Local Symmetry

Weiler, M., Forré, P., Verlinde, E., & Welling, M. (2024). Equivariant and Coordinate Independent Convolutional Networks. WORLD SCIENTIFIC. https://doi.org/10.1142/14143



(Hedgehog adapted under the free license by courtesy of Freepik.)



Summary

Equivariance: widely used in other scientific ML domains - has pros and cons, but the debate is ongoing

Goal: determine whether equivariance is a useful inductive bias for fluids

Limitations:

- Simple (easy) tasks considered
- Turbulent flows with a limited range of scales

Preliminary conclusions

- We *don't* need equivariance to predict large-scale flow structures
- We *do* need it to generalize to new coordinate frames
- The more anisotropic the flow, the more equivariance will help
  - For more isotropic flows - implicit data augmentation (less dependent on the coordinate frame)

Future work

- Investigate local symmetries and patterns in anisotropic flows
- Harder generalization tests - can equivariance help turbulence models generalize better?

Acknowledgements

NSERC
CRSNG

DOE
CSGF

NSF

DEPARTMENT OF DEFENSE
UNITED STATES OF AMERICA
NDSEG

MIT UROP
Undergraduate Research
Opportunities Program

Appendix

Anisotropy test - results

A. Backour

J. Balla

Model	Anisotropic MSE	Equivariance Error
CNN	5.300	0.0654
CNN + Aug	5.325	0.0657
ECNN	45.625	$4.0128 \cdot 10^{-13}$