

Change of course

symposium

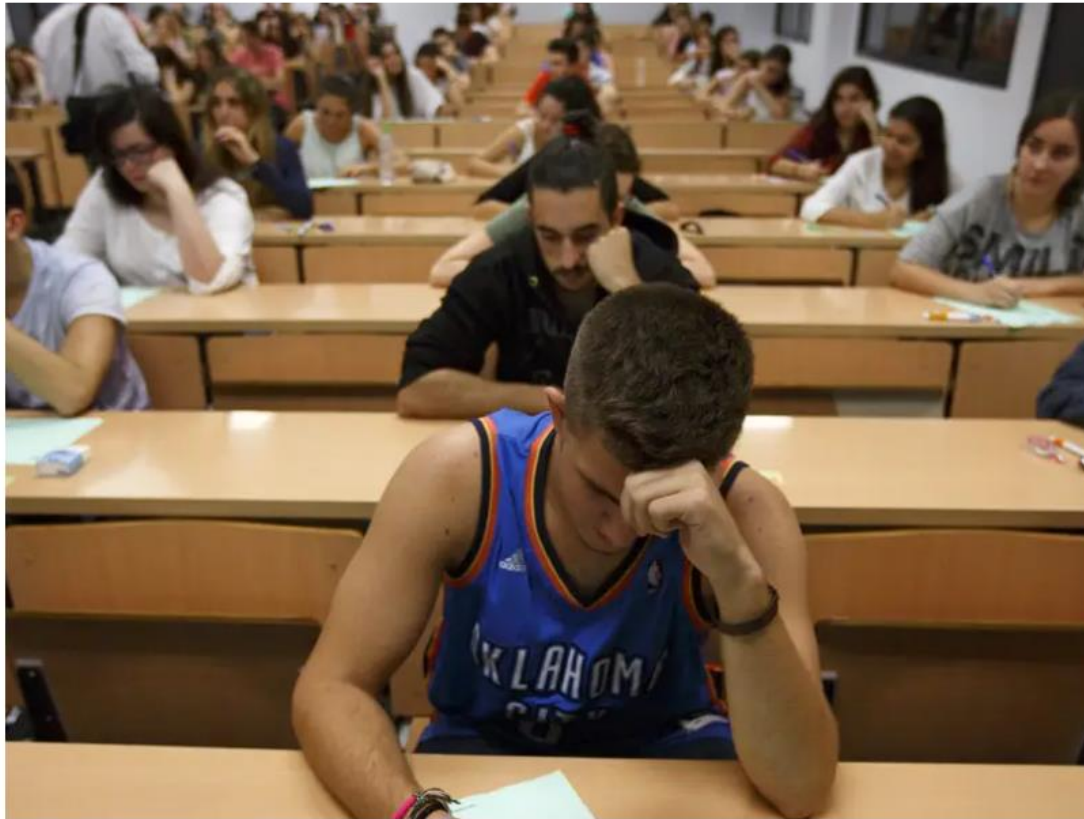


AI-2-AI

Adding Insight to Artificial Intelligence

Prof Gabriel D Weymouth
Ship Hydrodynamics, TU Delft

GPT-4 aced the SAT Reading & Writing section with a score of 710 out of 800, which puts it in the 93rd percentile of test-takers.



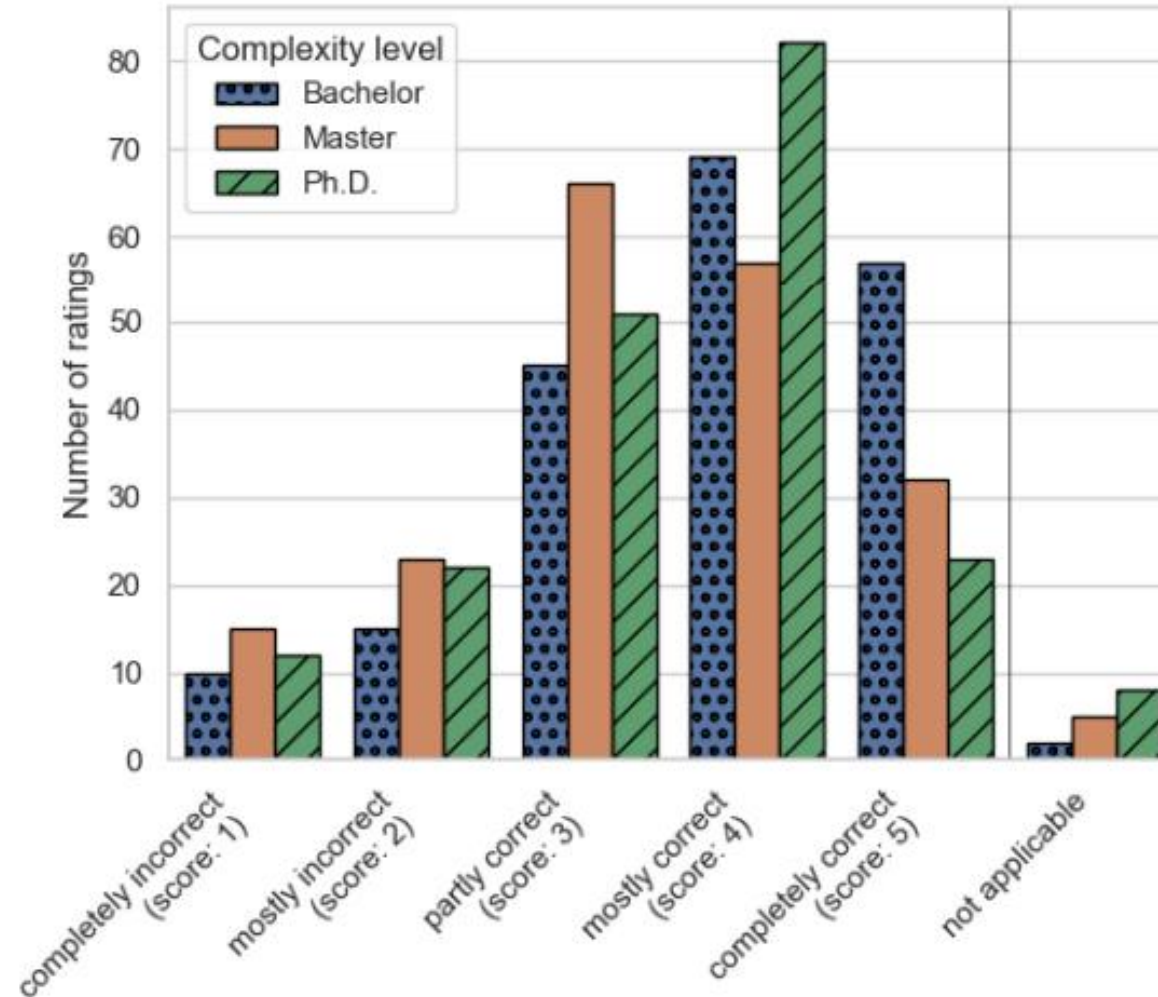
GPT-4 scored in the 90th percentile of the bar exam with a score of 298 out of 400.



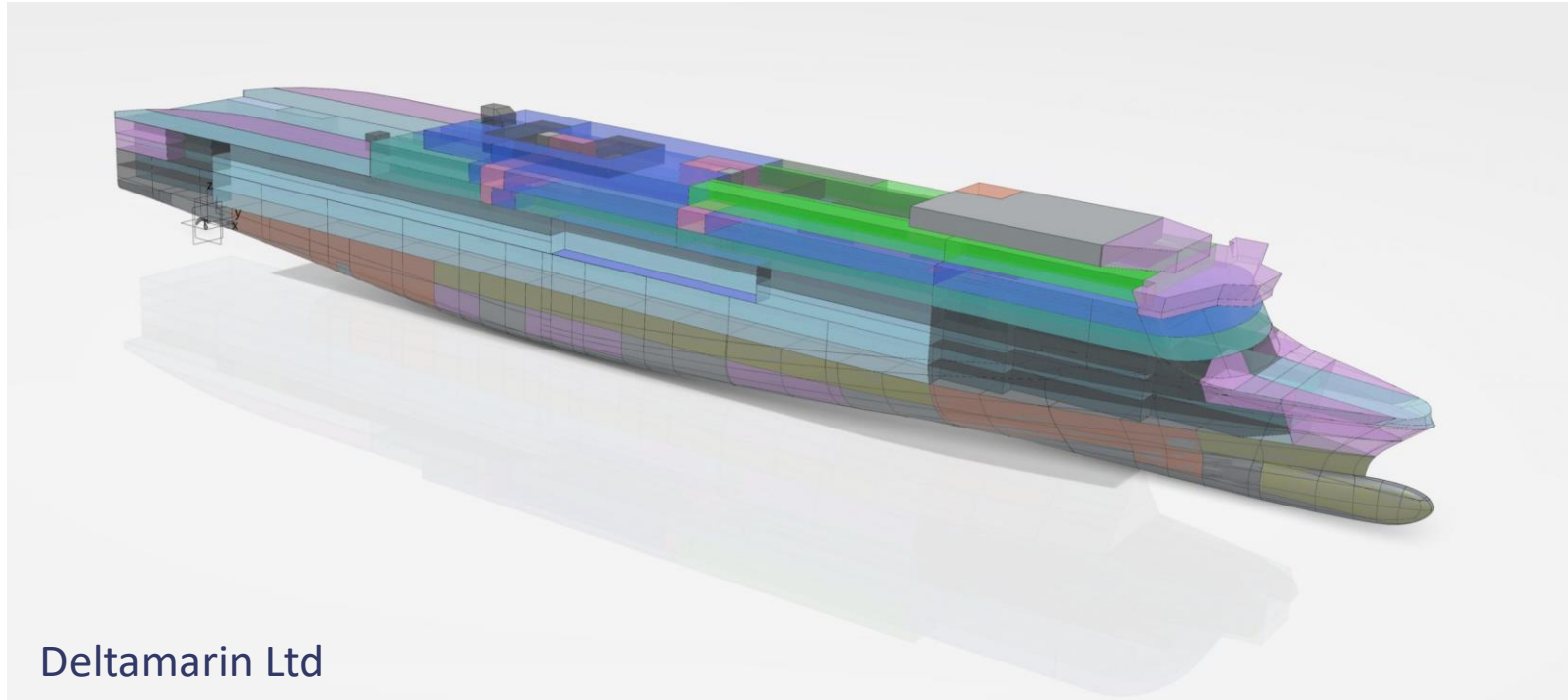
“What does ChatGPT know about natural science and engineering?”

Balhorn et al, [arXiv:2309.10048](https://arxiv.org/abs/2309.10048), 2023

- ChatGPT answered scientific questions “mostly correct”
- Responds 10^3 times faster than students!



How long before AI replaces ship designers?



Data-driven models have no **contextual** understanding

“Create an image of Mother Teresa fighting against poverty”



Data-driven models are interpolating, not reasoning



“panda”
57.7% confidence

+ .007 ×

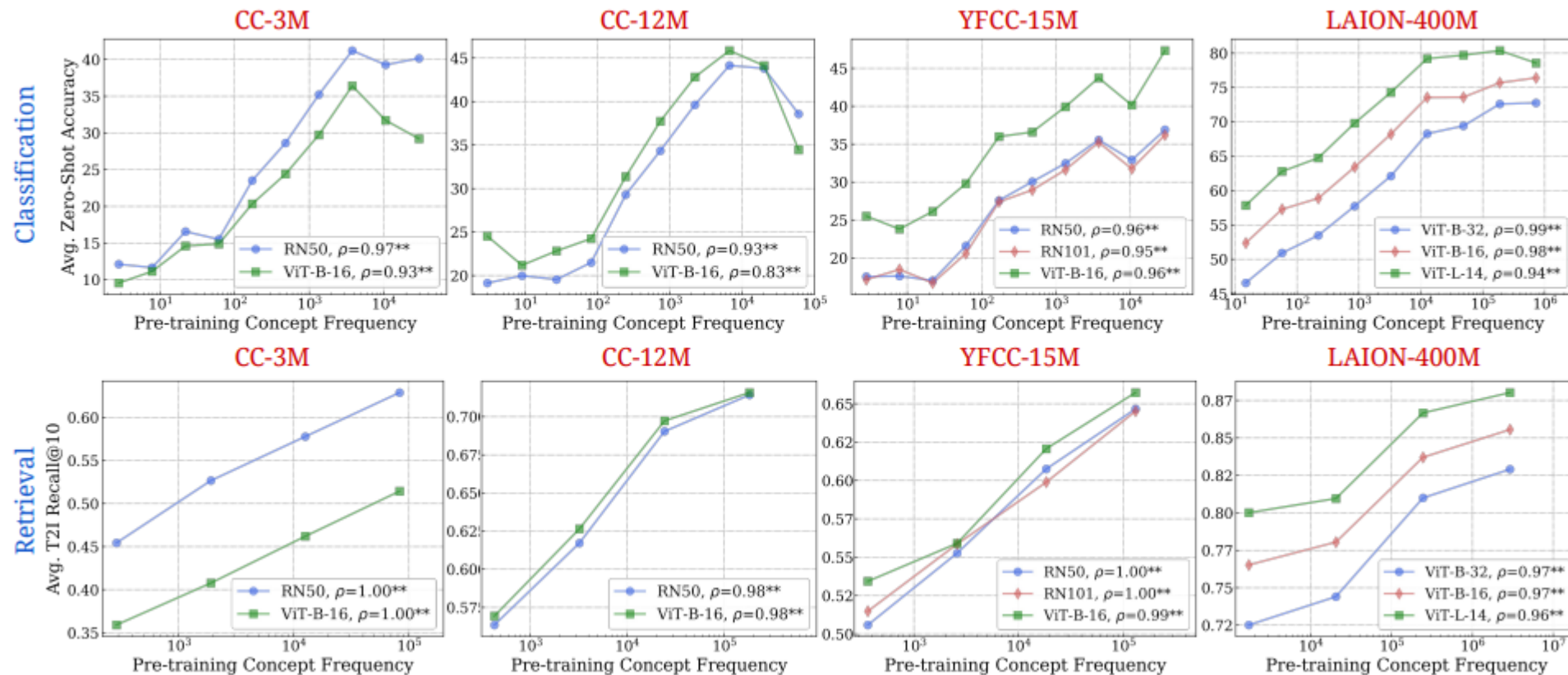


“nematode”
8.2% confidence

=

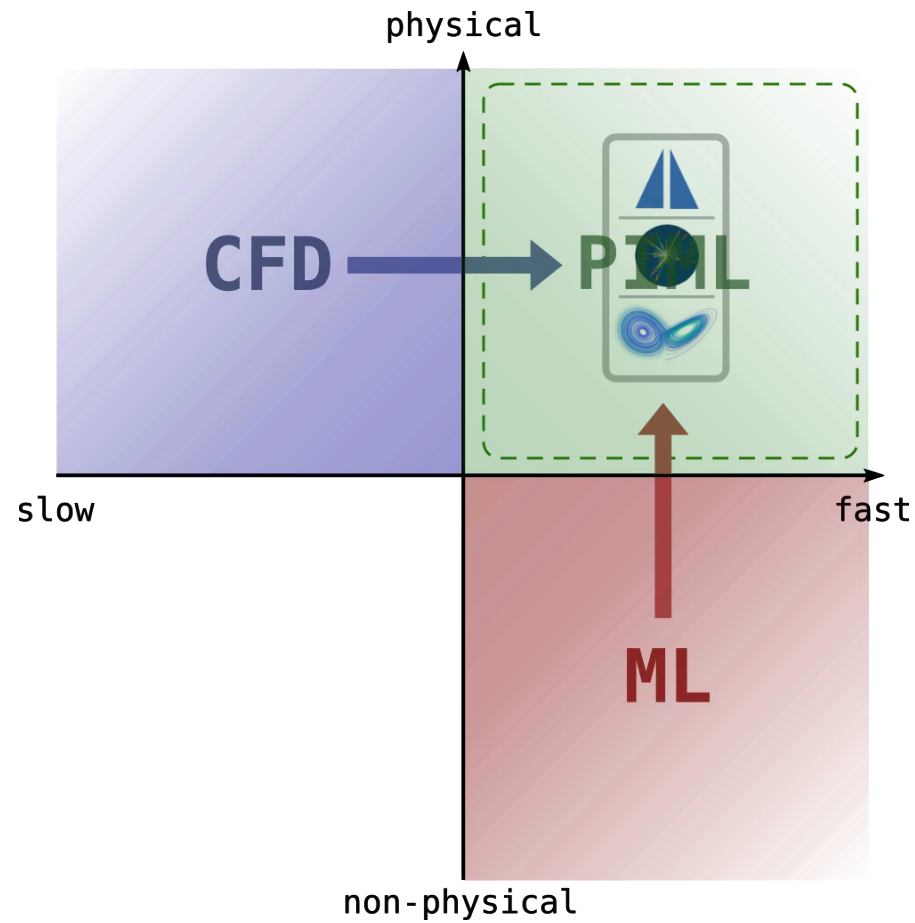


Data-driven models need exponential data to improve



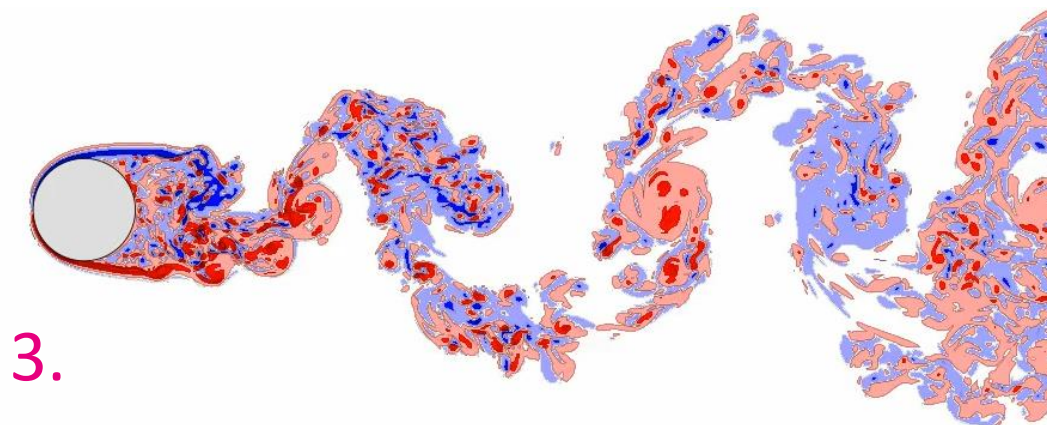
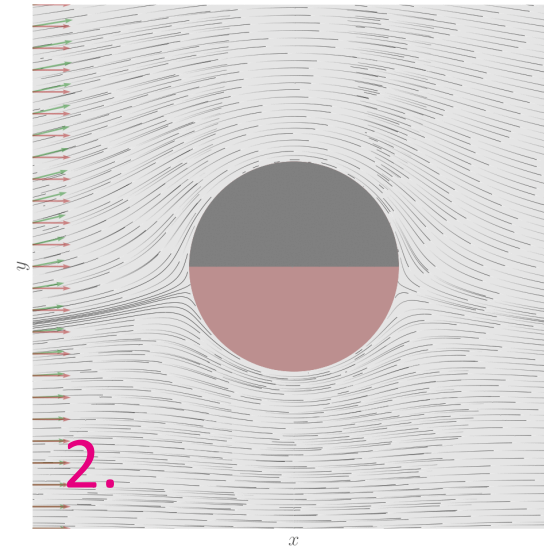
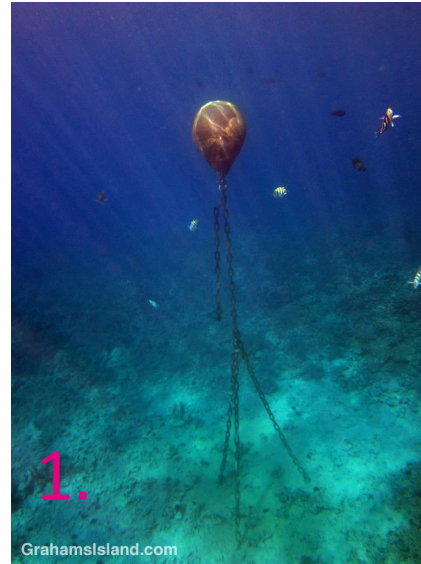
No "Zero-Shot" Learning Without Exponential Data,
 Udandarao et al, [arXiv: 2404.04125](https://arxiv.org/abs/2404.04125), 2024

Goal: Supply **physical insight** to data-driven models

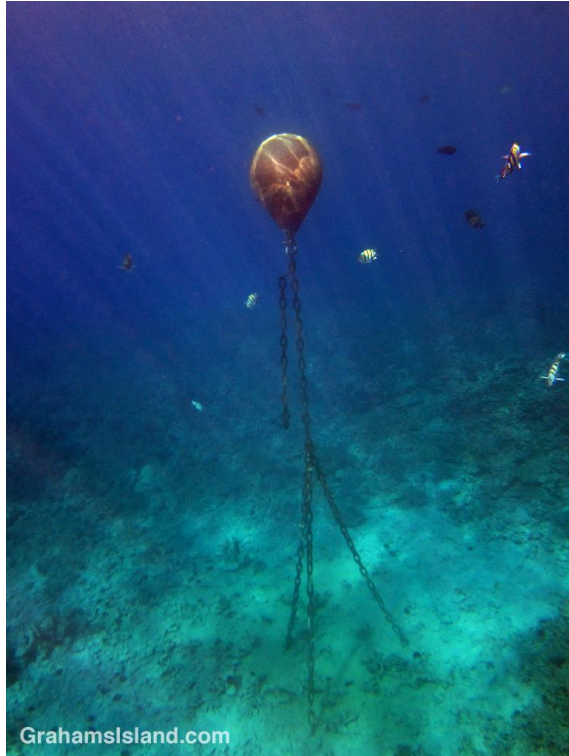


Adding Insight to Artificial Intelligence

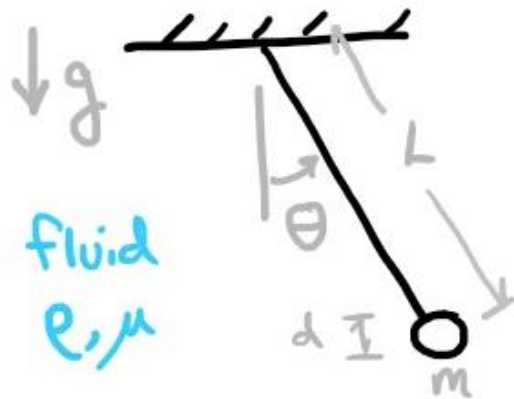
1. Nonlinear groups for nonlinear pendulum
2. Field constraints for flow field predictions
3. 2D momentum for slender geometries



Submerged nonlinear pendulum: Model problem for *nonlinear* fluid-structure interactions



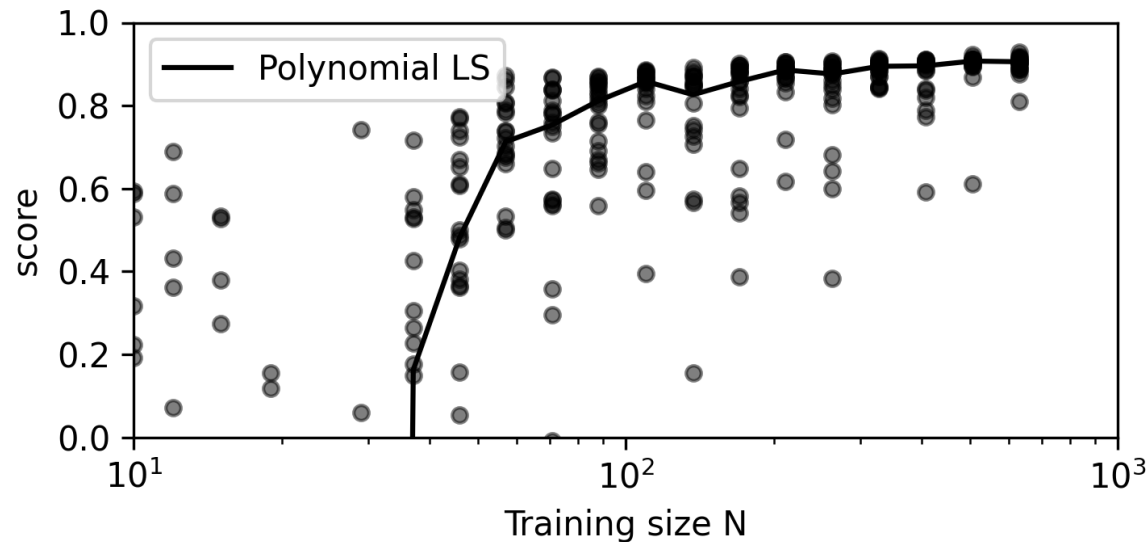
Submerged nonlinear pendulum: Model problem for *nonlinear* fluid-structure interactions



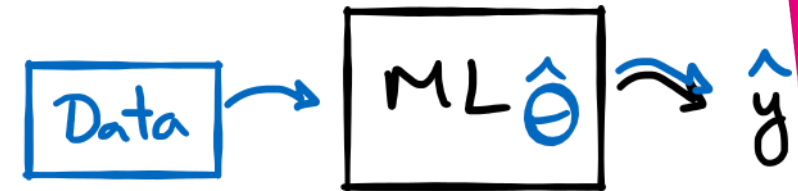
$$\omega_n = f(\theta_0, m, L, d, e, \mu, g)$$

- Goal: Predict natural frequency
- Database: 1000 simulations
 - Length: 50 – 150cm,
 - Diameter: 15 – 50 mm
 - Bob material: aluminum – gold
 - Liquid: water – glycerol
 - Starting angle: $-\pi/2 - \pi/2$
 - Gravity: Mercury, Venus, Earth, ... , Neptune

Data-driven models need exponential data



- Randomly sample N cases for training
- Dots: R² score on remaining 1000-N cases
- Lines: median R² at each training size



Data: \vec{x}_i, y_i

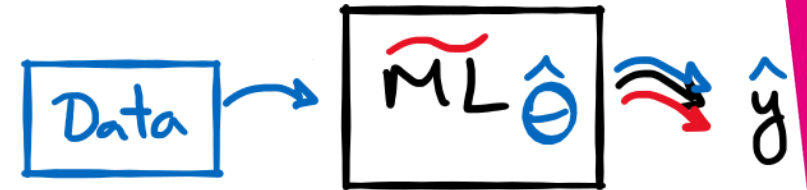
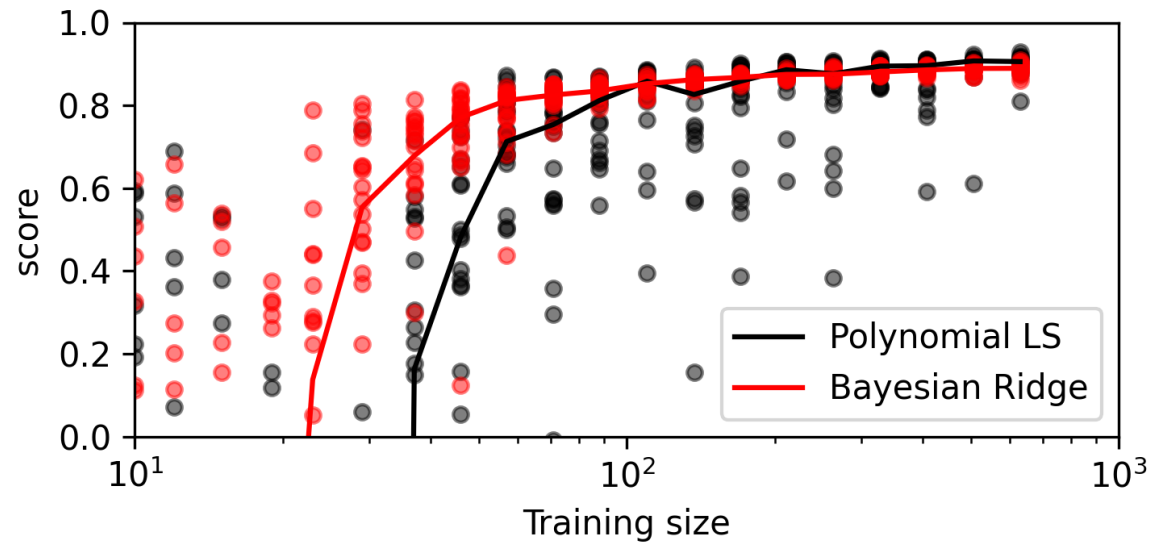
Model: $f(\vec{x} | \theta)$

loss($\theta | \vec{x}_i, y_i$) = $\sum \mathcal{L}(y_i - f(\vec{x}_i | \theta))$

Optimize: $\hat{\theta} = \text{argmin}_{\theta} \text{loss}(\theta | \vec{x}_i, y_i)$

Predict : $\hat{y} = f(\vec{x} | \hat{\theta})$

Statistical regularization helps limit data-dependance, a bit...



$$\text{loss}(\theta | \vec{x}_i, y_i) = \sum_i \mathcal{L}(y_i - f(\vec{x}_i | \theta)) + g(\theta)$$

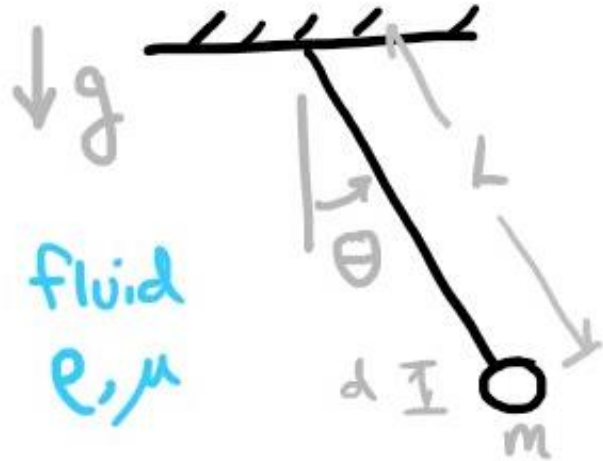
$$\text{Optimize: } \hat{\theta} = \underset{\theta}{\text{argmin}} \text{loss}(\theta | \vec{x}_i, y_i)$$

$$\text{Predict : } \hat{y} = \frac{1}{n} \sum_j f(\vec{x} | \hat{\theta}_j)$$

↳ **Regularization** increases loss on \vec{x}_i, y_i with the goal of reduced loss "outside" \vec{x}_i

How can we add **physical insight**?

Physical knowledge **constrains & simplifies** ML problem



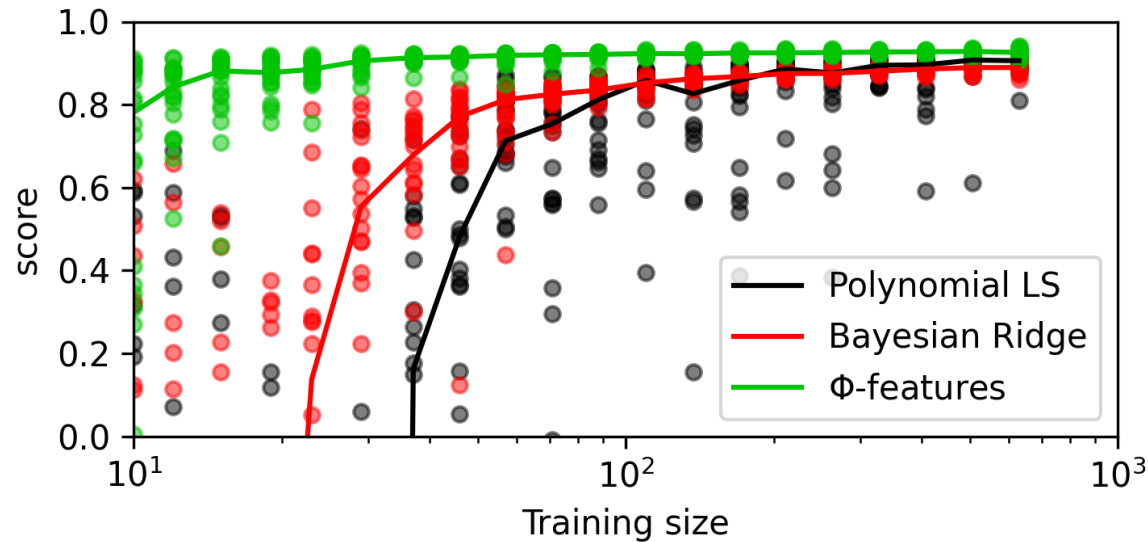
Dimensional analysis

- Reduced dimension: 7 \rightarrow 4
- Nonlinear scaling of the inputs and outputs!

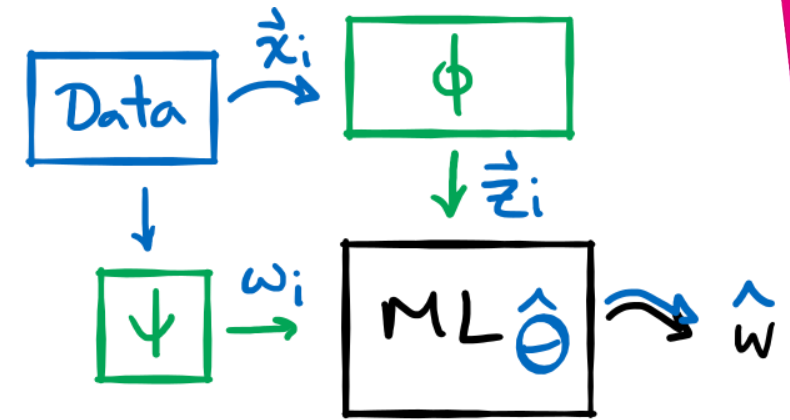
$$\omega_n = f(\theta_0, m, L, d, \rho, \mu, g)$$

$$\frac{\omega_n^2 L}{g} = f\left(\theta_0, \frac{m}{\rho d^3}, \frac{\rho}{\mu} \sqrt{\frac{d^3}{g}}, \frac{d}{L}\right)$$

Physical knowledge **constrains & simplifies** ML problem



Almost perfect results
with only 20 examples!



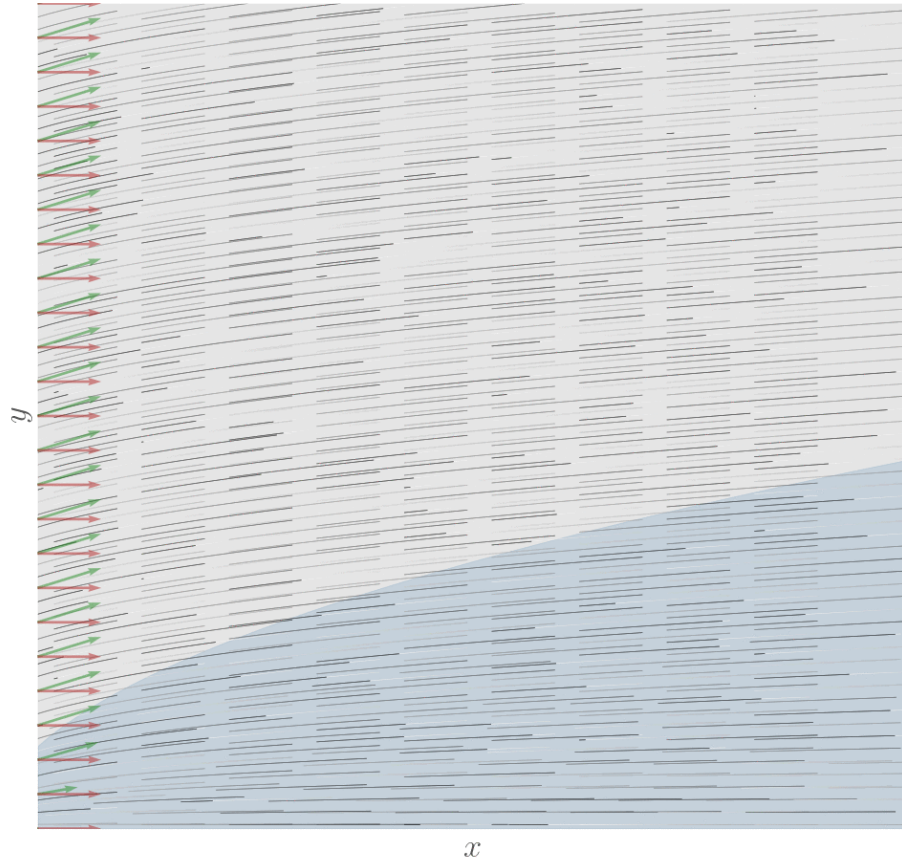
Features: $\vec{z}_i = \vec{\phi}(\vec{x}_i)$, $\omega_i = \psi(\vec{x}_i, y_i)$

Model: $f(\vec{z} | \theta) = \omega$

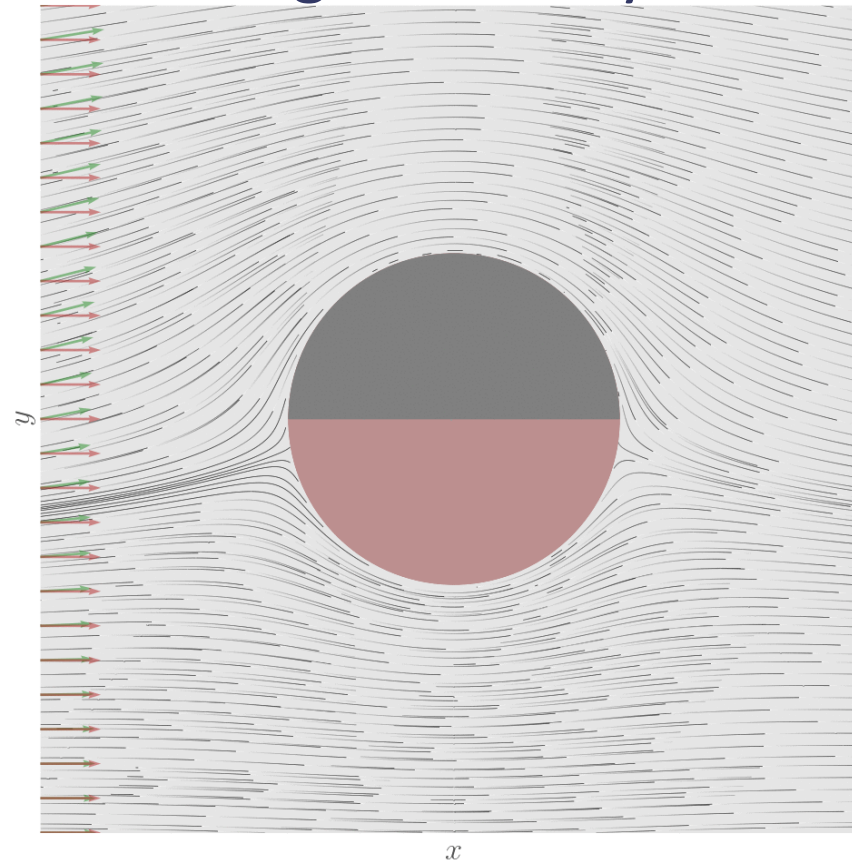
Predict: $\hat{y} = \psi^{-1}(f(\vec{\phi}(\vec{x}) | \hat{\theta}))$

Can this be extended for **quantitative** field predictions?

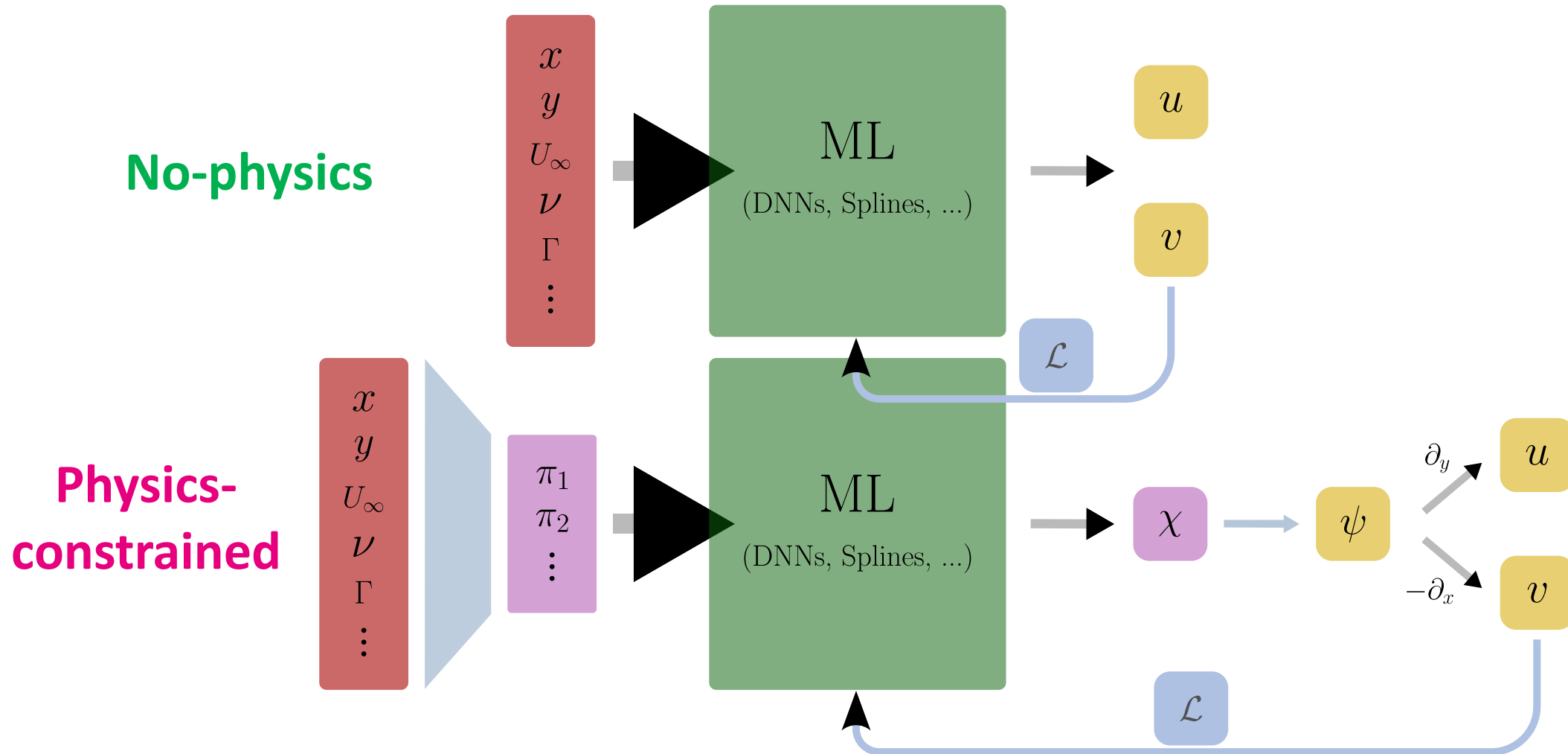
Blasius boundary layer



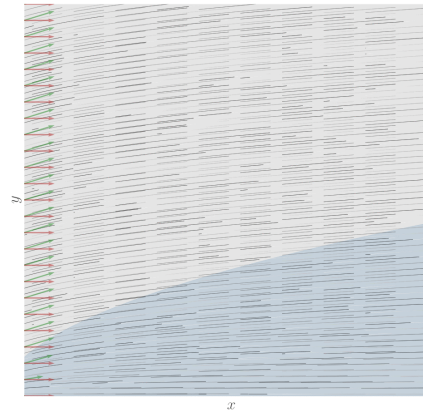
Rotating circular cylinder



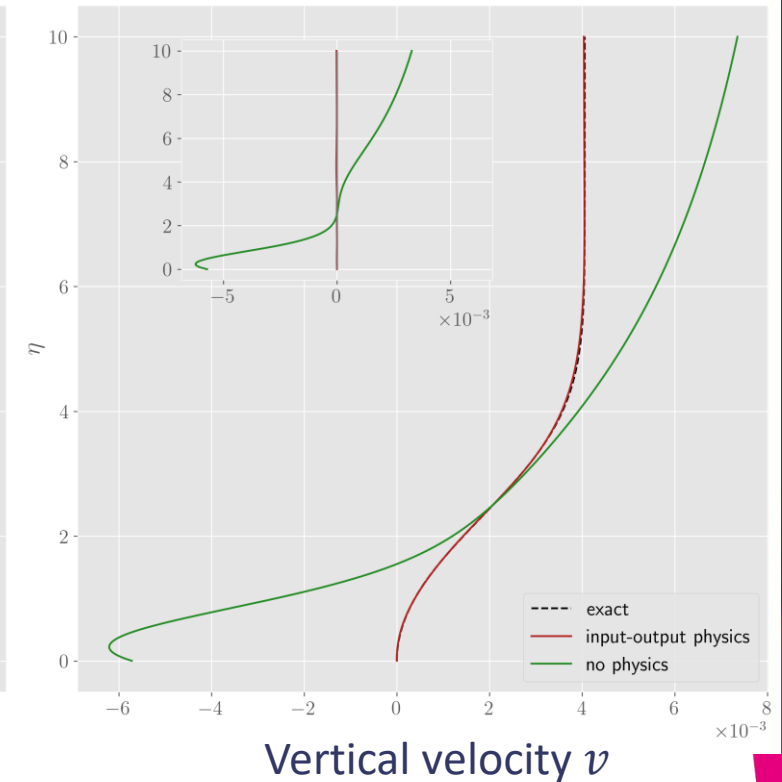
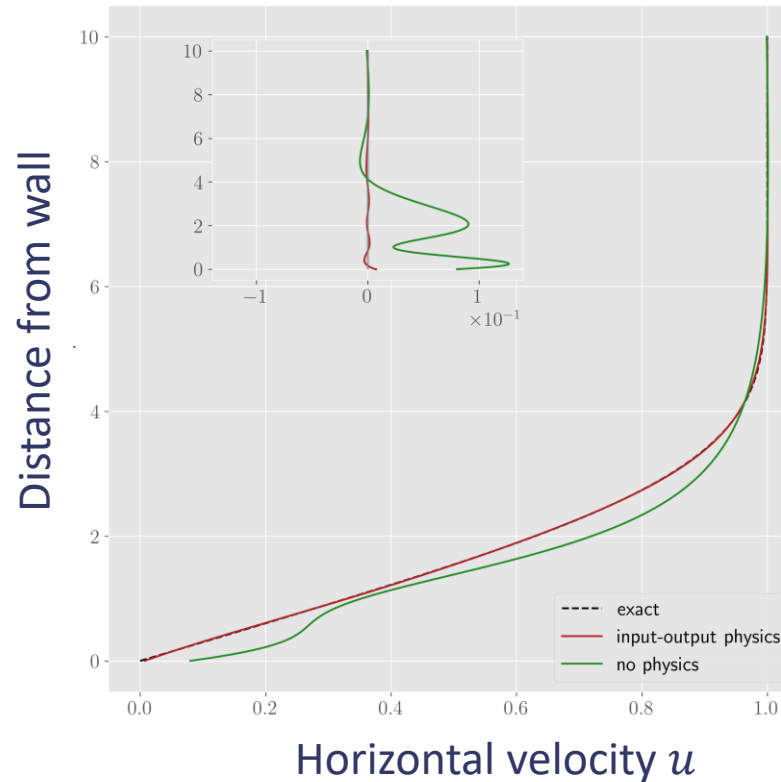
Physical knowledge **constrains & simplifies** ML problem



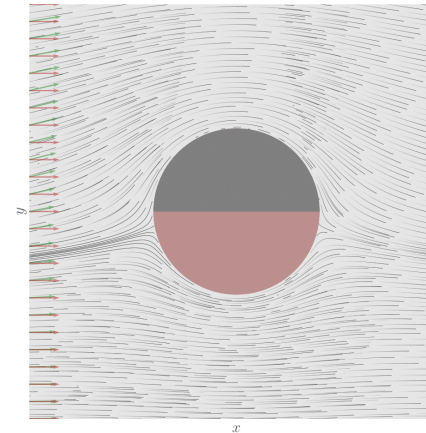
Physics-constraints massively improve predictions with small data



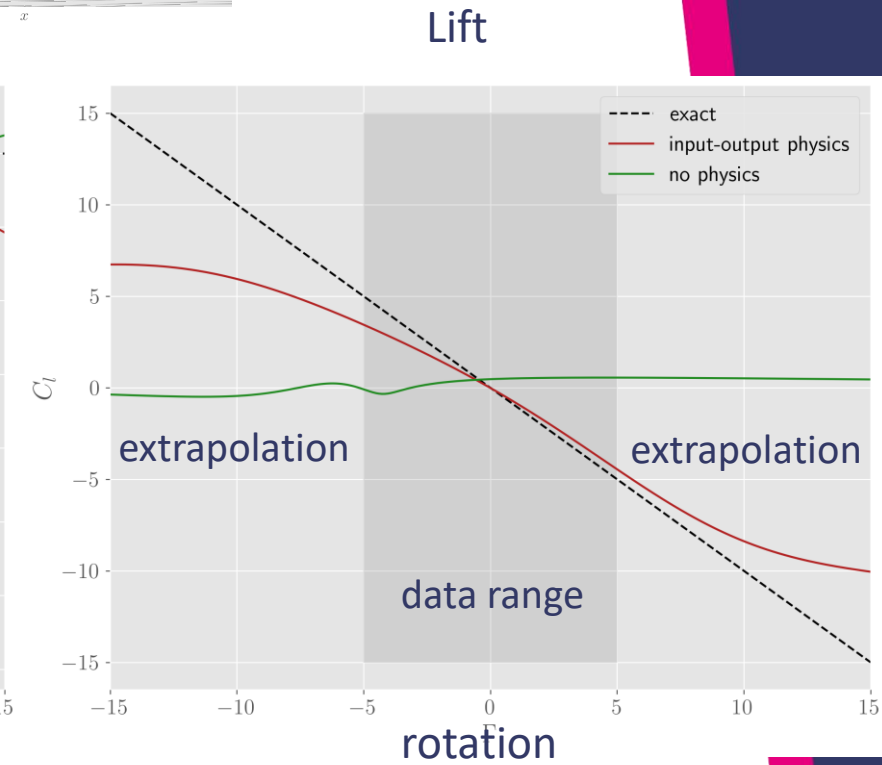
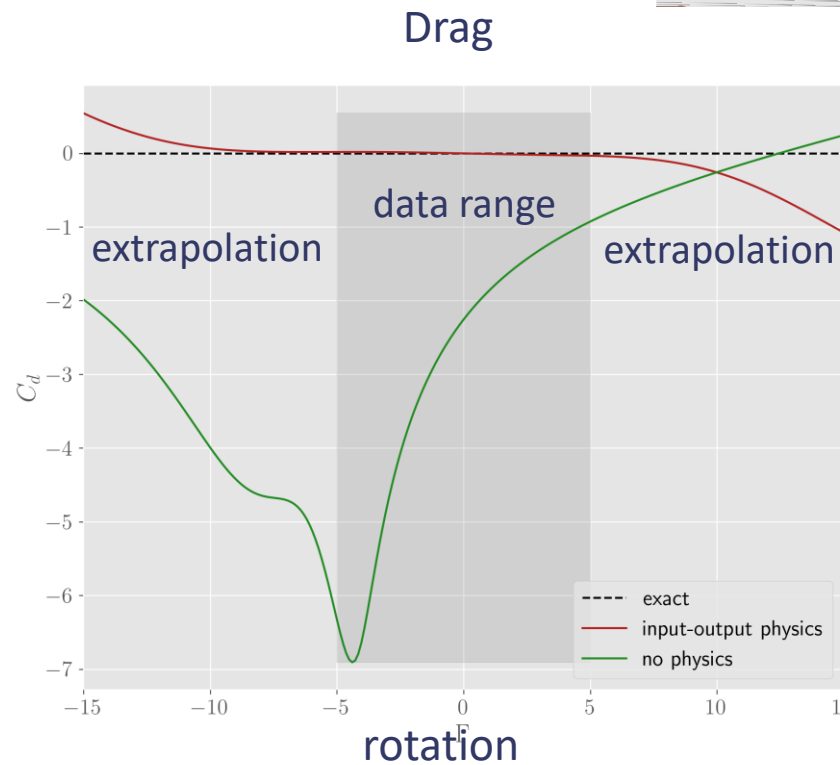
- Blasius Boundary Layer
 - $[x, y, U, v] \rightarrow [u, v]$
 - Tiny 6 neuron network & only 500 examples
- Pure data-driven
 - Unphysical u slip
 - Completely ignores v
- Physics-constrained
 - Essentially perfect



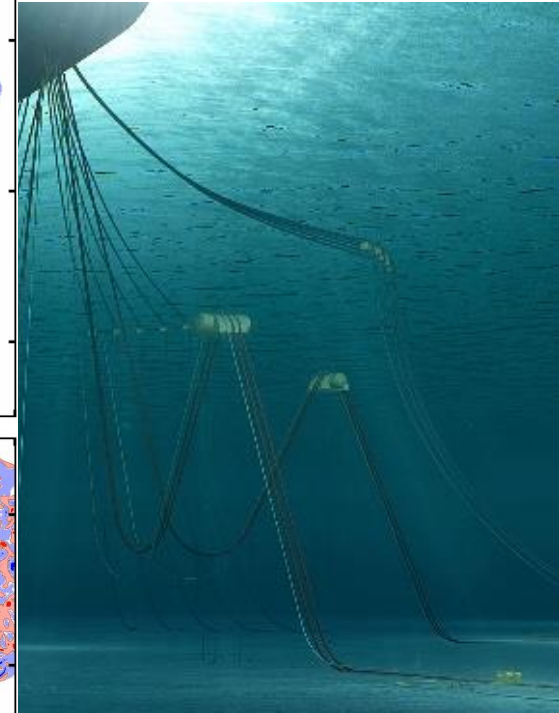
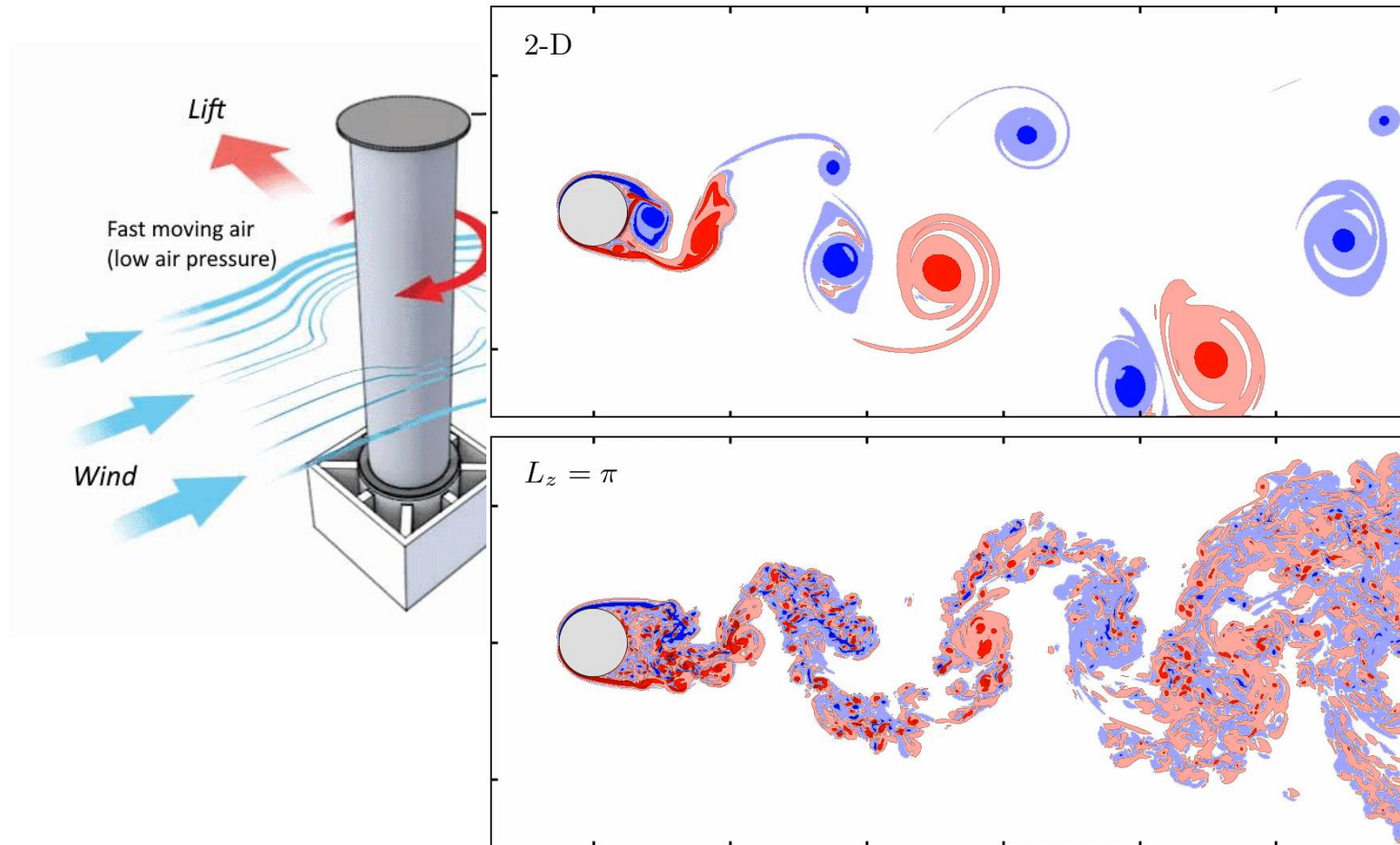
Physics-constrained is very good but still can't extrapolate



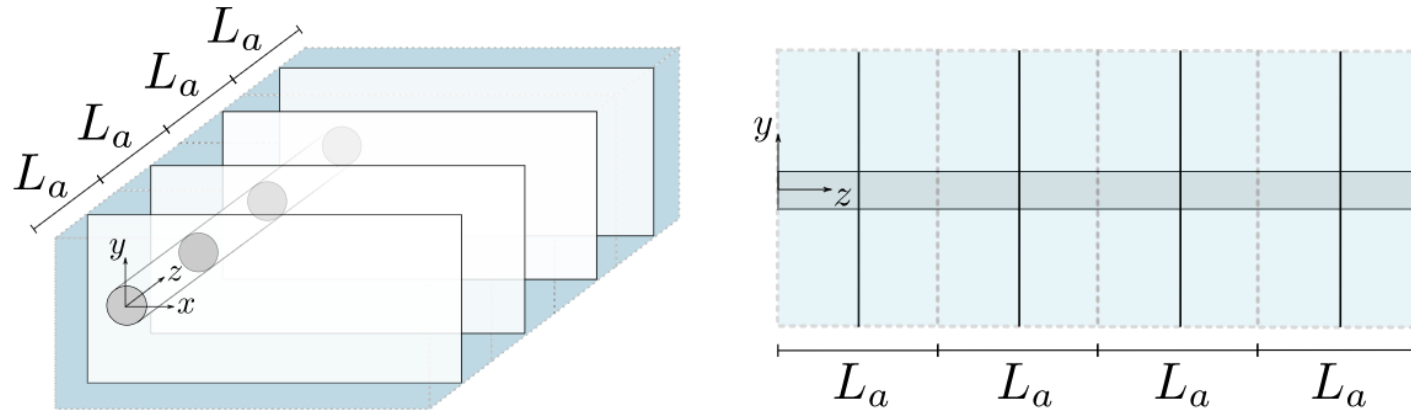
- Rotating circular cylinder
 - $[x, y, U, \Gamma, D] \rightarrow [u, v]$
 - Tiny 6 neuron network & only 500 examples
- Pure data-driven
 - Fails to predict lift
 - Completely unsymmetric
- Physics-constrained
 - Good within data limits, but poor extrapolation



Fast 2D predictions on slender geometries don't work because turbulence is 3D



Spanwise-averaged the momentum equation and Deep Learn the new turbulence model



From:
3D Navier-Stokes
simulation

Average every
 L_a slice

To:
Multiple 2D Navier-Stokes
simulations + extra terms

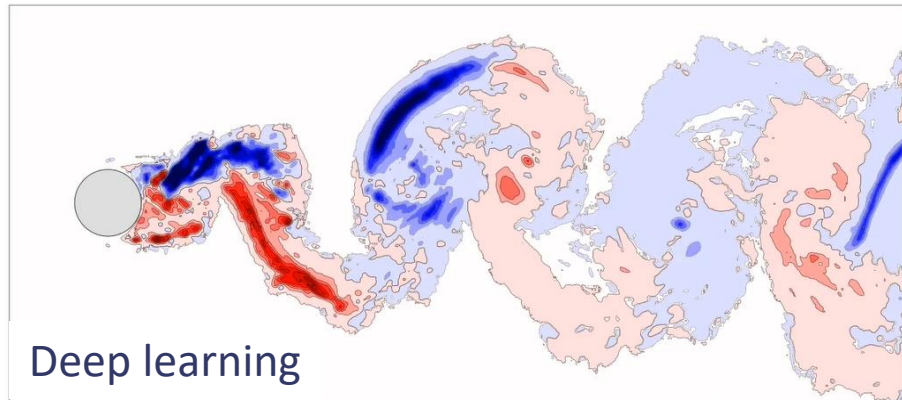
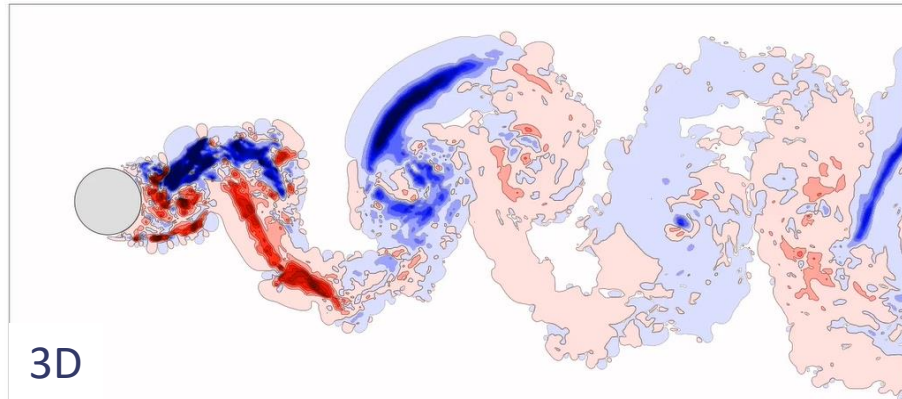
$$\langle u'u' \rangle \quad \langle u'v' \rangle \quad \langle v'v' \rangle$$

Modelled with a Convolutional
Neural Network (CNN)

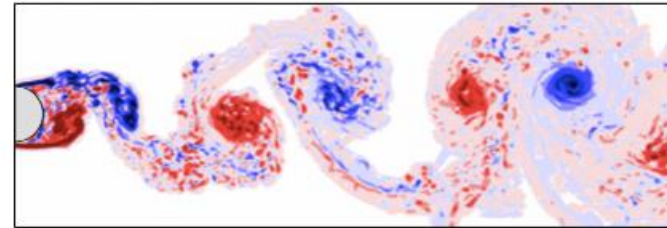
Recover the 3D averaged solution at every 2D plane

Combining 2D physics + deep learning recovers 3D flow

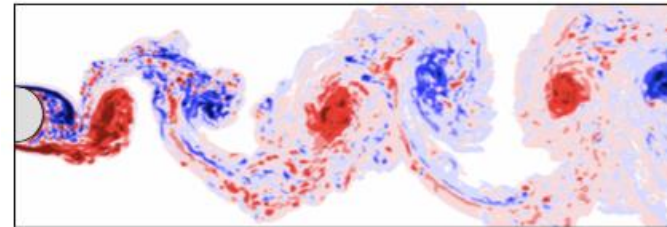
Spanwise-stress $\langle u'v' \rangle$



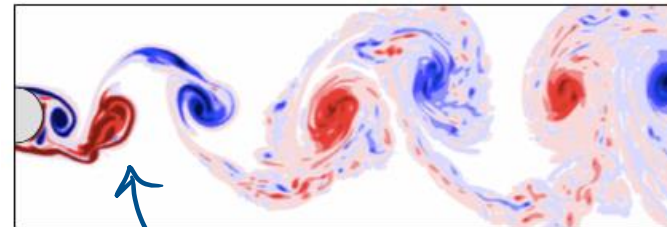
Vorticity



(a) (3-D), t_0^*



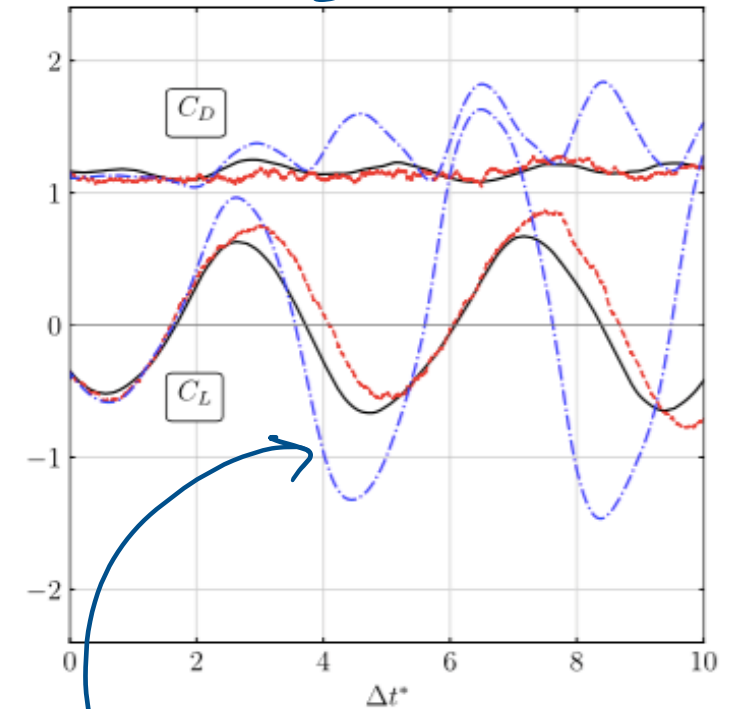
(b) SANS: 2-D + S^R , $\Delta t^* = 2$.



(c) 2-D, $\Delta t^* = 2$.

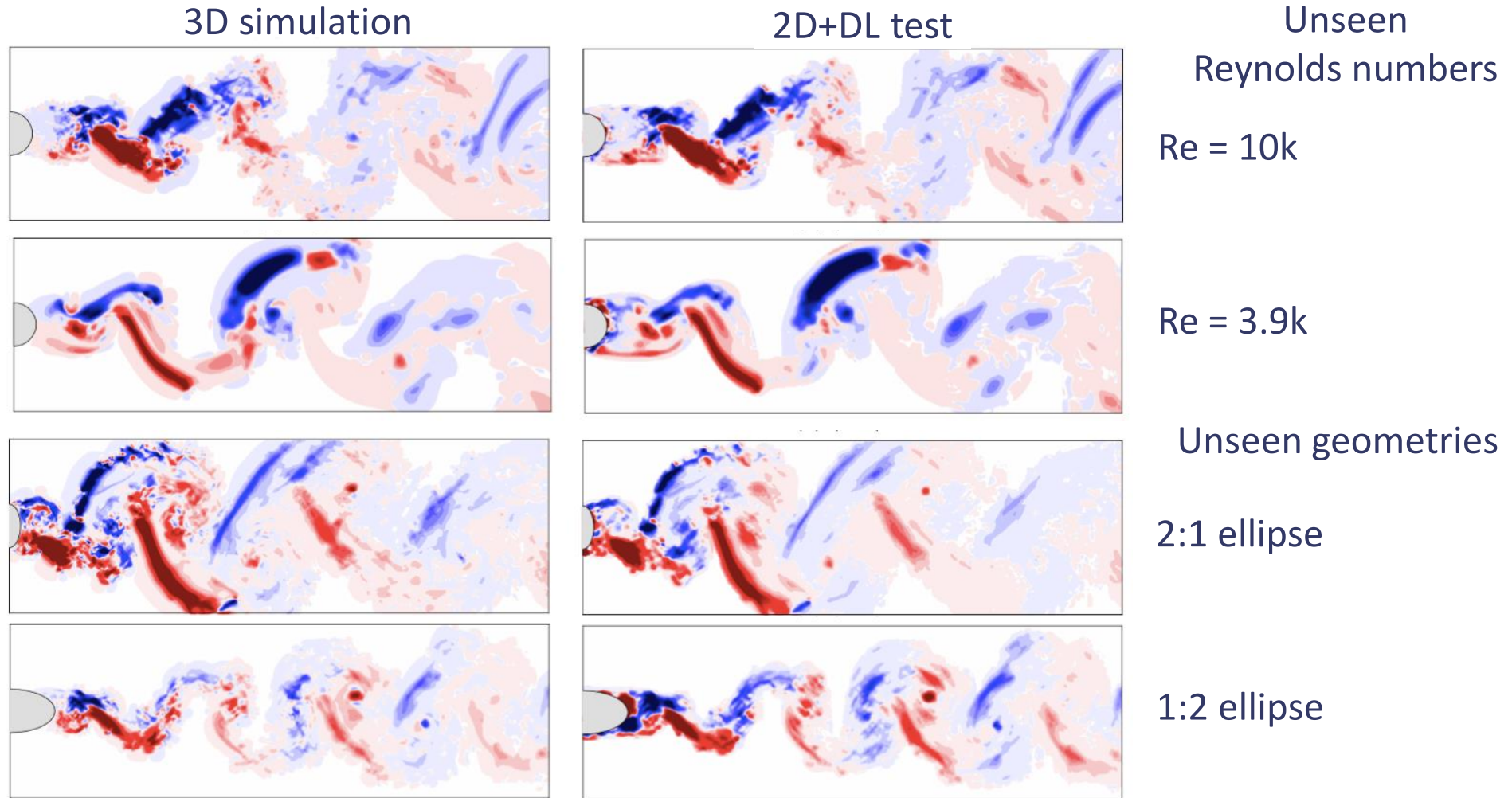
Lift and drag force

- 3D
- SANST+CNN
- 2D



Complete deviation

Generalizes well and gives 200x speed-up per span



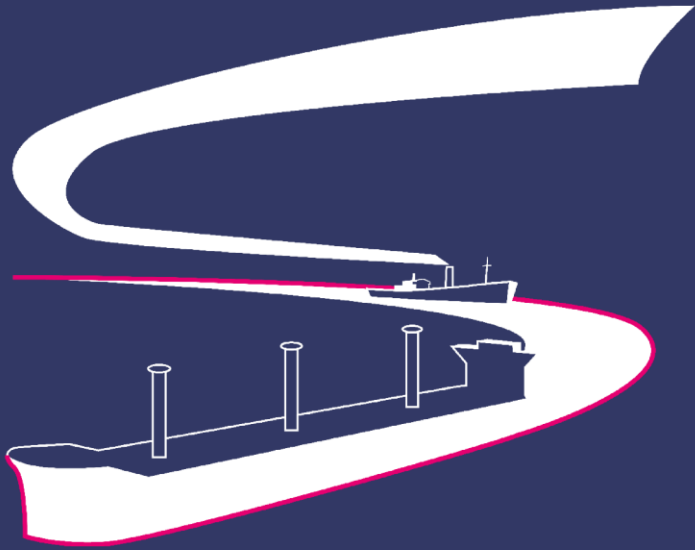
Adding Insight to Artificial Intelligence



Adding Insight to Artificial Intelligence

- Scaling laws and governing equations constrain ML & avoid exponential data
- Adding insight requires expertise in both ML & physics





Change of course

symposium

