Learned Models for Physical Simulation and Design

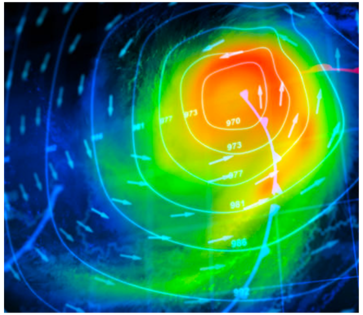
Kimberly Stachenfeld
DeepMind

AMLD EPFL 2022

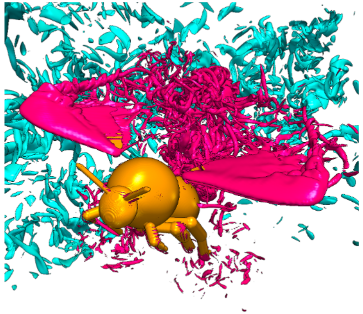
DeepMind

Simulation

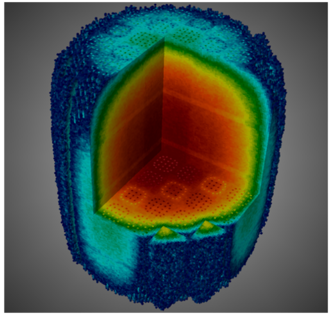
Forecasting



Basic Science



Engineering & Design



How can Machine Learning be useful for simulation?

- Simulators that...
  - Part I → Are more **efficient**
  - Part II → Are **differentiable**; support efficient **engineering & design** optimization problems
  - learn from **real-world data**

Learned Coarse Models for Efficient Turbulence Simulation

FLATIRON INSTITUTE

Stachenfeld, Fielding, Kochkov, Cranmer, Pfaff, Godwin, Cui, Ho, Battaglia, Sanchez-Gonzalez (ICLR 2022)

Learned Coarse Models for Efficient Turbulence Simulation

Classical numerical solvers used for turbulence simulations are powerful but computationally expensive

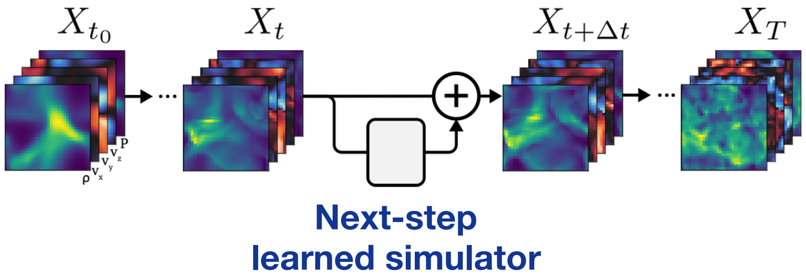
Can fully-learned simulators capture complex, chaotic turbulence accurately at low-resolutions?

FLATIRON INSTITUTE

Stachenfeld, Fielding, Kochkov, Cranmer, Pfaff, Godwin, Cui, Ho, Battaglia, Sanchez-Gonzalez (ICLR 2022)

Our Approach

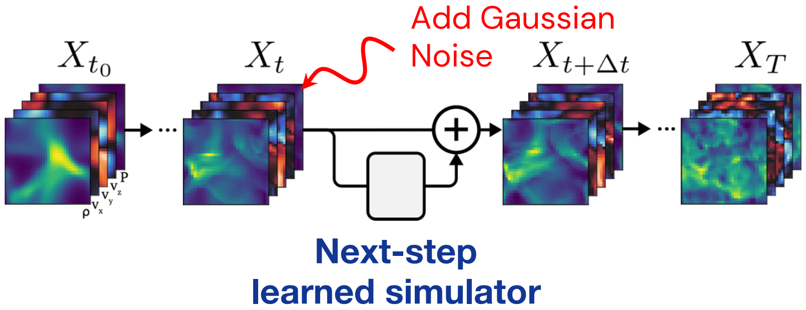
1. Use classical physics solvers to produce **high-resolution trajectories**
2. Downsample these trajectories in **space** and **time** to produce training data
3. Train a neural network to do **next-step prediction** on **low-resolution** frames



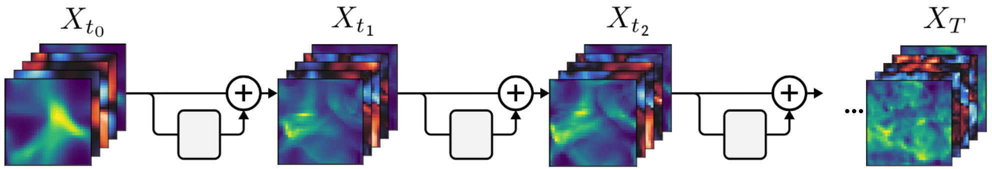
Stachenfeld, Fielding, Kochkov, Cranmer, Pfaff, Godwin, Cui, Ho, Battaglia, Sanchez-Gonzalez (ICLR 2022)

Our Approach

1. Use classical physics solvers to produce **high-resolution trajectories**
2. Downsample these trajectories in **space** and **time** to produce training data
3. Train a neural network to do **next-step prediction** on **low-resolution** frames
4. **Evaluate** on **"rolled out"** trajectories



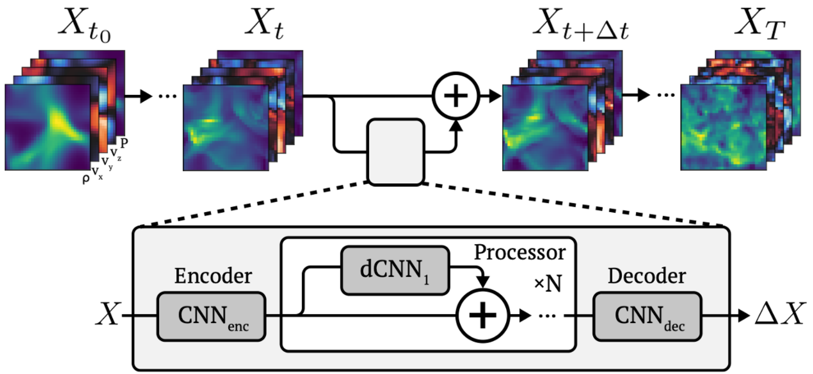
Evaluation:



Stachenfeld, Fielding, Kochkov, Cranmer, Pfaff, Godwin, Cui, Ho, Battaglia, Sanchez-Gonzalez (ICLR 2022)

Our Approach

1. Use classical physics solvers to produce **high-resolution trajectories**
2. Downsample these trajectories in **space** and **time** to produce training data
3. Train a neural network to do **next-step prediction** on **low-resolution** frames
4. **Evaluate** on **“rolled out”** trajectories

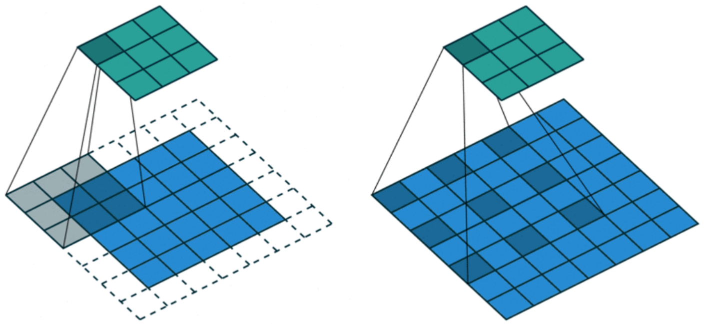


Stachenfeld, Fielding, Kochkov, Cranmer, Pfaff, Godwin, Cui, Ho, Battaglia, Sanchez-Gonzalez (ICLR 2022)

U-shaped Stack of Dilated Convolutions

Dilated Convolution

Yu & Koltun (2015)



Keep local structure

Larger perceptual range

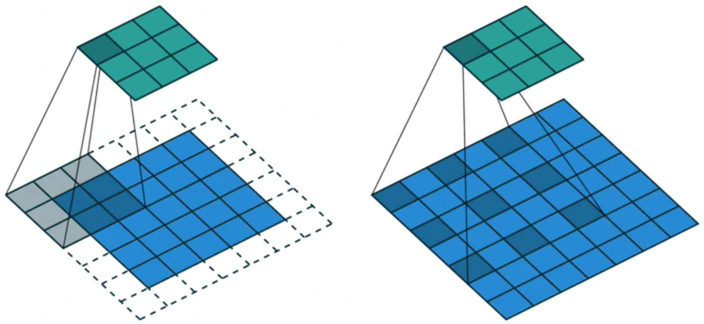
Same # parameters

Stachenfeld, Fielding, Kochkov, Cranmer, Pfaff, Godwin, Cui, Ho, Battaglia, Sanchez-Gonzalez (ICLR 2022)

U-shaped Stack of Dilated Convolutions

Dilated Convolution

Yu & Koltun (2015)



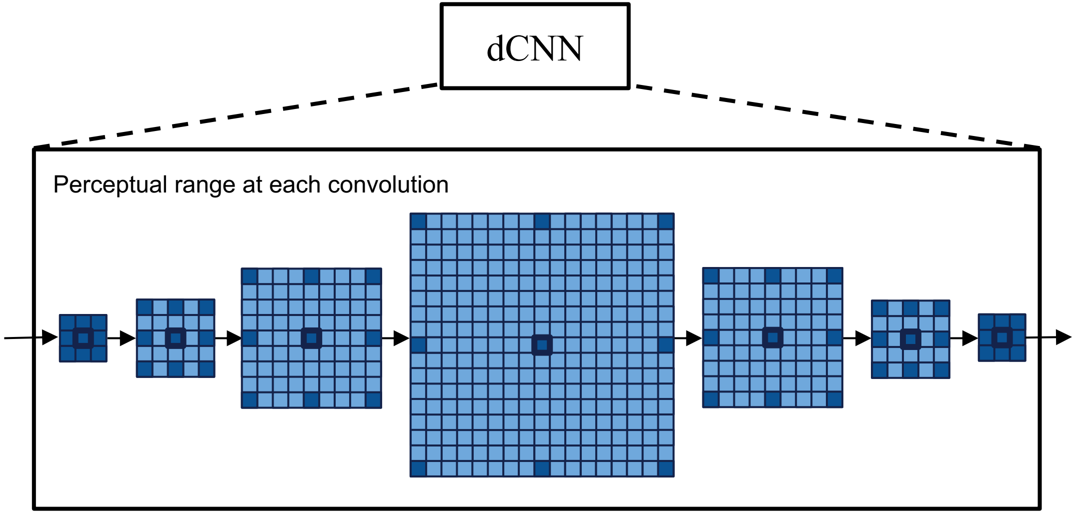
Keep local structure

Larger perceptual range

Same # parameters

dCNN: U-shaped stack

See also: Ronneberger et al. (2015)



7 dilated convolutions in sequence

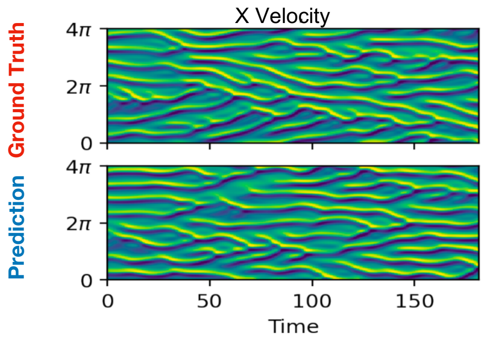
Gradually increase and decrease range of communication

Stachenfeld, Fielding, Kochkov, Cranmer, Pfaff, Godwin, Cui, Ho, Battaglia, Sanchez-Gonzalez (ICLR 2022)

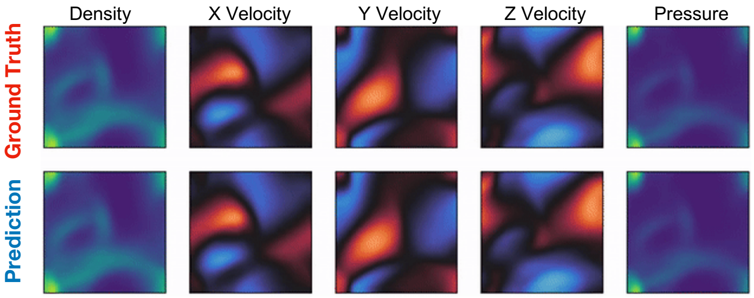
Domain Generality

One model → 4 different domains

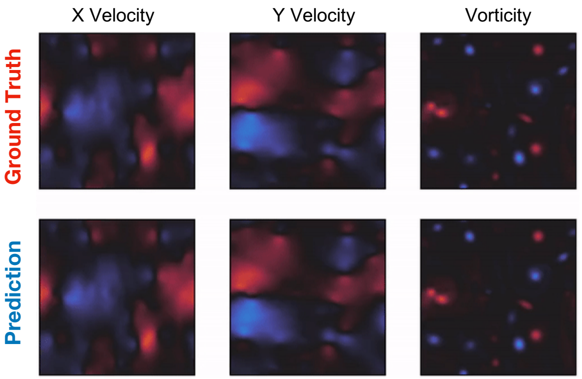
1D Kuramoto–Sivashinsky (KS) Equation



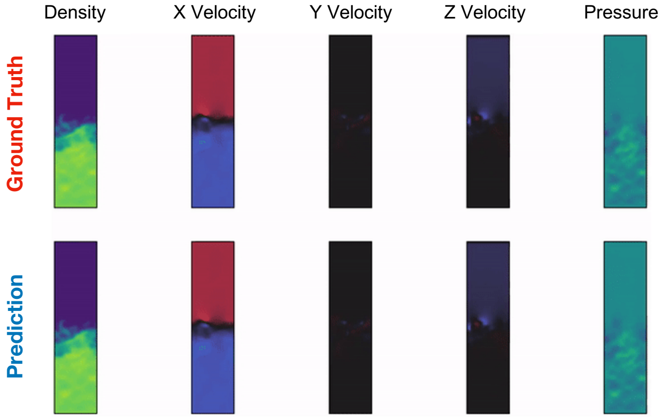
3D Uniform Compressible Decaying Turbulence



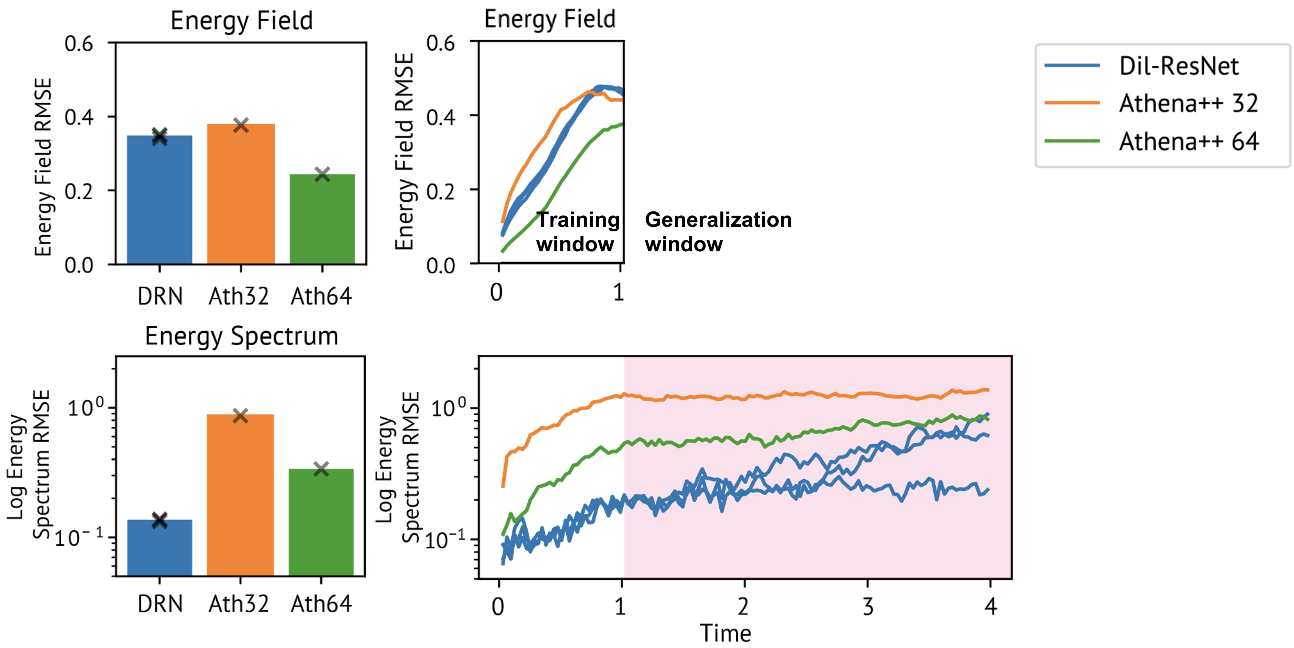
2D Incompressible Turbulence



3D Mixing Layer Turbulence with Radiative Cooling



Comparison to Coarsened Physics Based Simulator (Athena++)

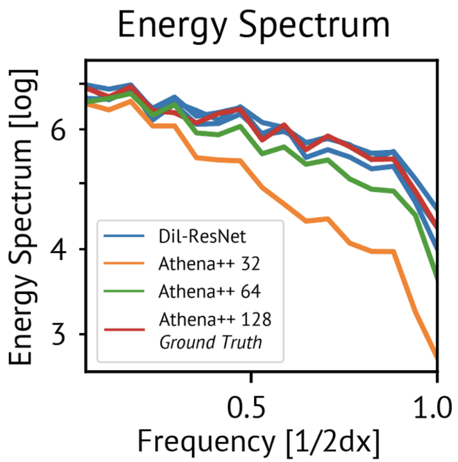


Slightly lower pixel-wise error than classical simulator at same (32^3) resolution during training window

Much better spectral error Athena++ at same (32^3) and higher (64^3) resolution (ground truth: 128^3)

Stachenfeld, Fielding, Kochkov, Cranmer, Pfaff, Godwin, Cui, Ho, Battaglia, Sanchez-Gonzalez (ICLR 2022)

Comparison to Coarsened Physics Based Simulator (Athena++)

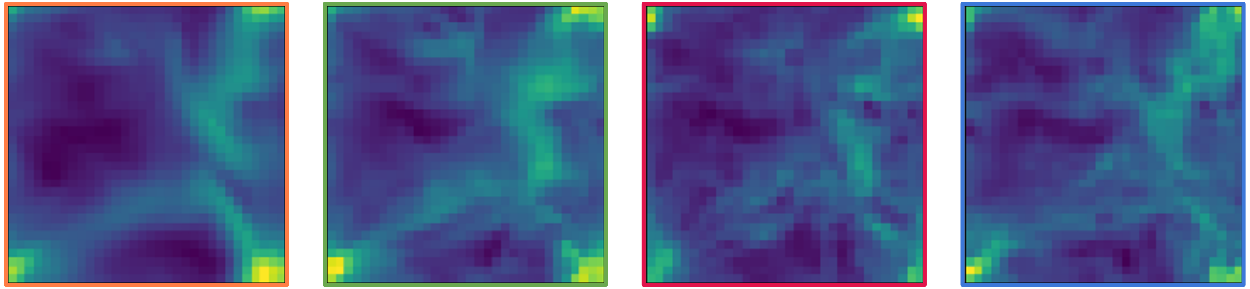


Athena++ 32^3
Low resolution

Athena++ 64^3
Intermediate resolution

Athena++ 128^3
Ground truth
High resolution

Dil-ResNet (32^3)
Low resolution



same resolution

Learned model preserves **high frequency** structure that the classical Athena++ simulator loses at low resolution

Stachenfeld, Fielding, Kochkov, Cranmer, Pfaff, Godwin, Cui, Ho, Battaglia, Sanchez-Gonzalez (ICLR 2022)

Running time

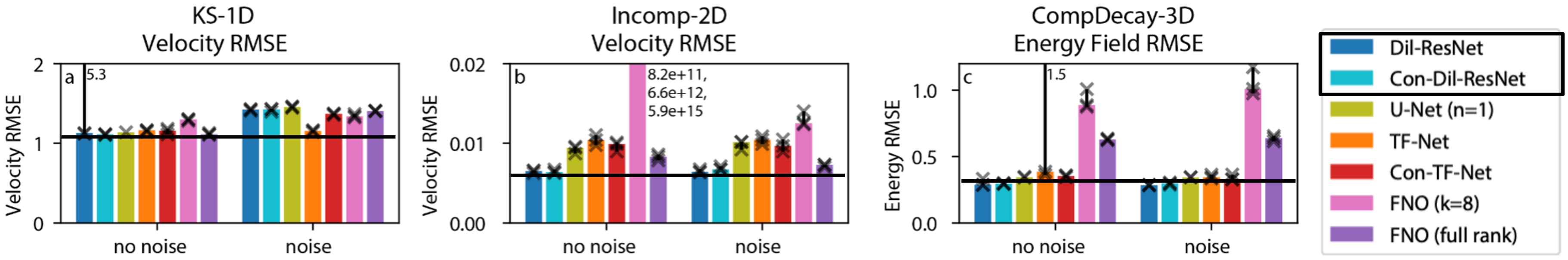
- Athena++
  - Scales O(resolution^4)
  - CPU only

- Learned model:
  - **Up to 1000x faster** than Athena at 128^3

Simulator	Time (s)
Athena++ 32^3	~4
Athena++ 64^3	~60
Athena++ 128^3	~1000
Model $128^3 \rightarrow 32^3$	~20-30
Model $128^3 \rightarrow 32^3$ (GPU)	~1

Stachenfeld, Fielding, Kochkov, Cranmer, Pfaff, Godwin, Cui, Ho, Battaglia, Sanchez-Gonzalez (ICLR 2022)

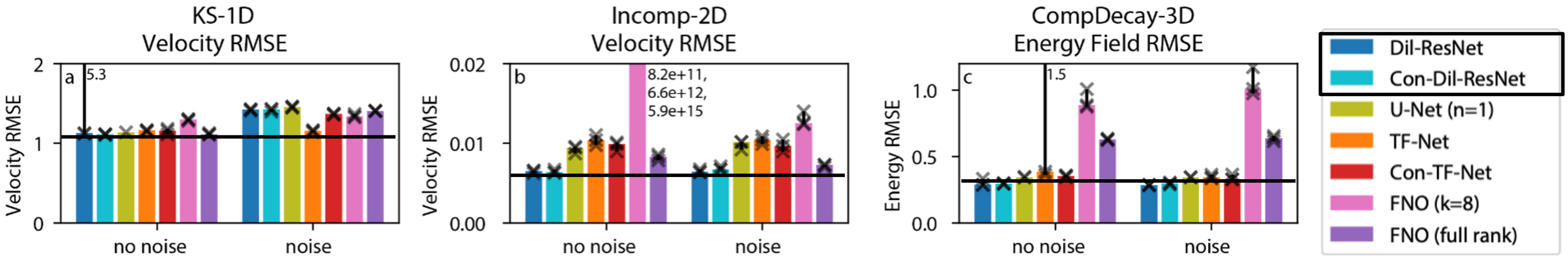
Learned Model Comparison



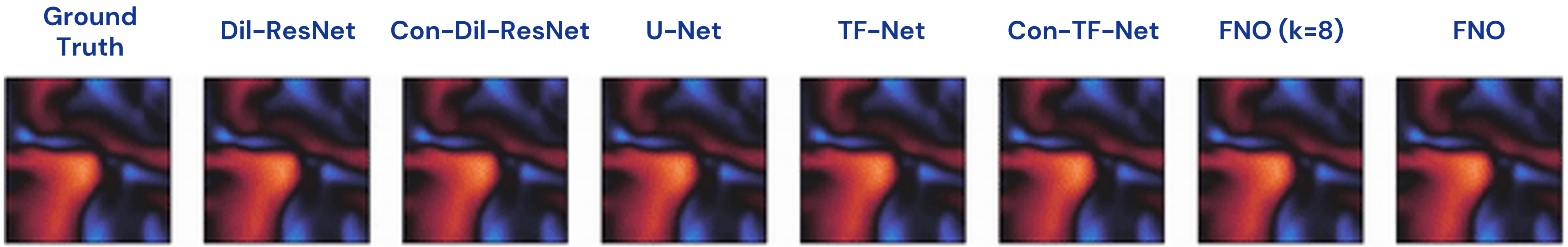
Our models quantitatively outperform other, more specialized, parameterized models

Stachenfeld, Fielding, Kochkov, Cranmer, Pfaff, Godwin, Cui, Ho, Battaglia, Sanchez-Gonzalez (ICLR 2022)

Learned Model Comparison



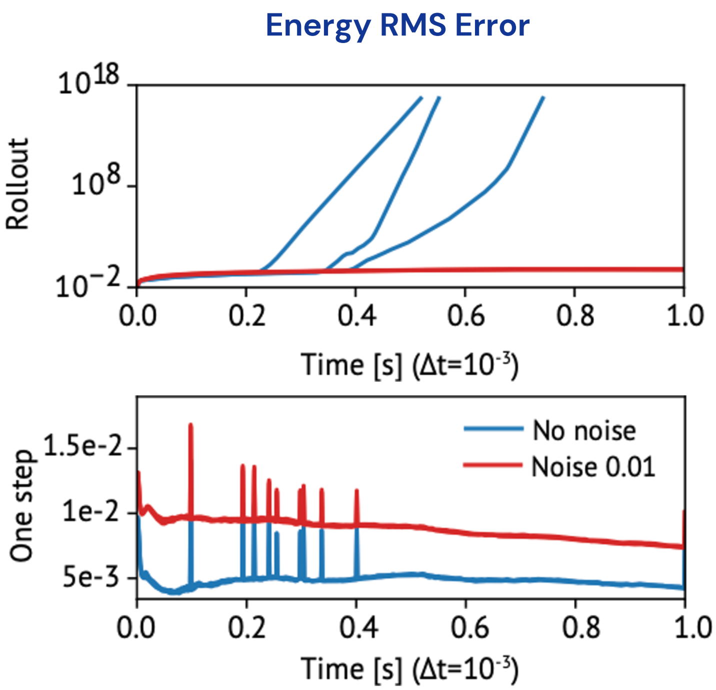
Our models quantitatively outperform other, more specialized, parameterized models



However, most learned models do qualitatively pretty well

Stachenfeld, Fielding, Kochkov, Cranmer, Pfaff, Godwin, Cui, Ho, Battaglia, Sanchez-Gonzalez (ICLR 2022)

Stability

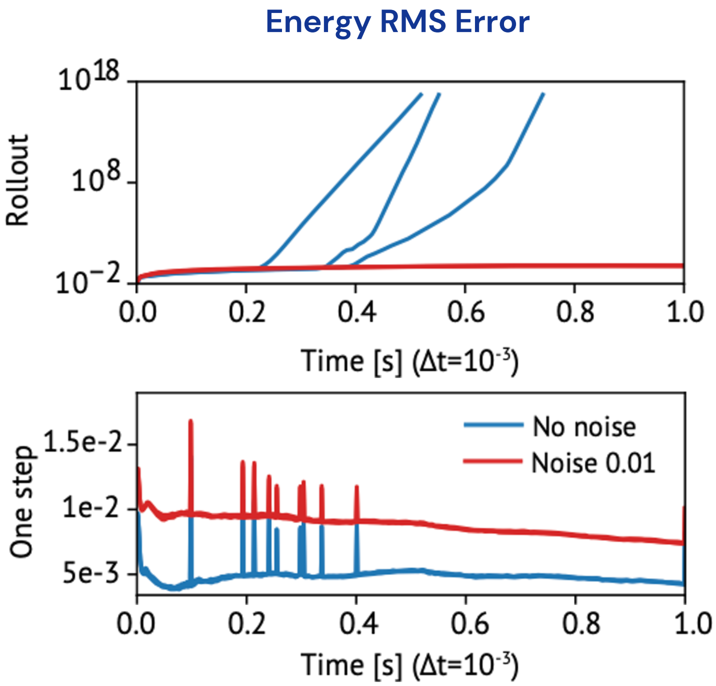


Models **trained without noise** can be unstable

Using **training noise** improves stability but has higher one-step error

Stachenfeld, Fielding, Kochkov, Cranmer, Pfaff, Godwin, Cui, Ho, Battaglia, Sanchez-Gonzalez (ICLR 2022)

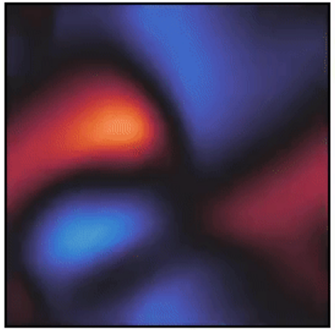
Stability



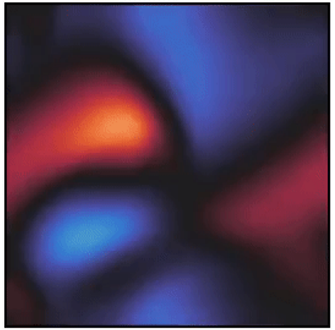
Models **trained without noise** can be unstable

Using **training noise** improves stability but has higher one-step error

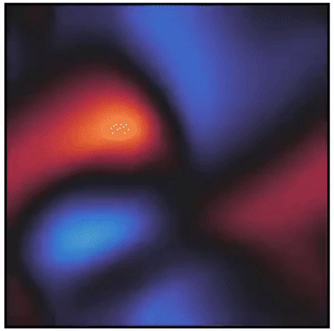
No noise



Noise 0.01



Ground Truth



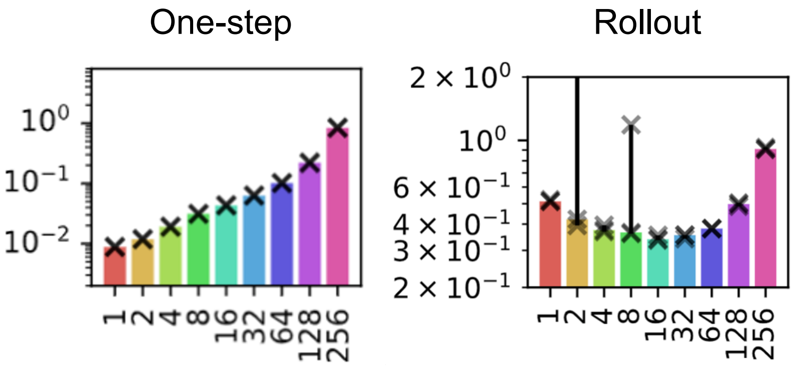
Stachenfeld, Fielding, Kochkov, Cranmer, Pfaff, Godwin, Cui, Ho, Battaglia, Sanchez-Gonzalez (ICLR 2022)

Temporal Coarsening

Learned simulators can be trained on larger timesteps.

Larger timesteps cause larger one-step loss but can lead to greater stability

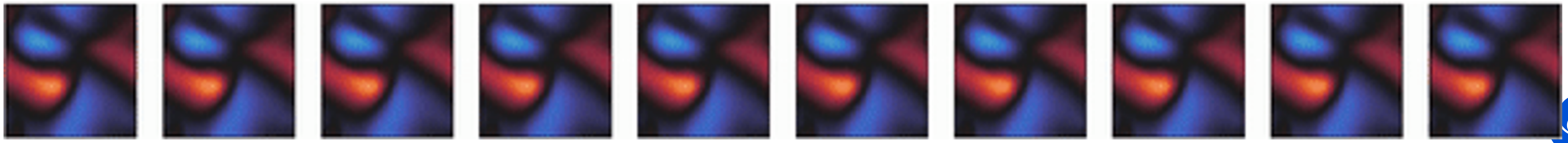
Energy RMS error
(trained with noise)



Ground Truth

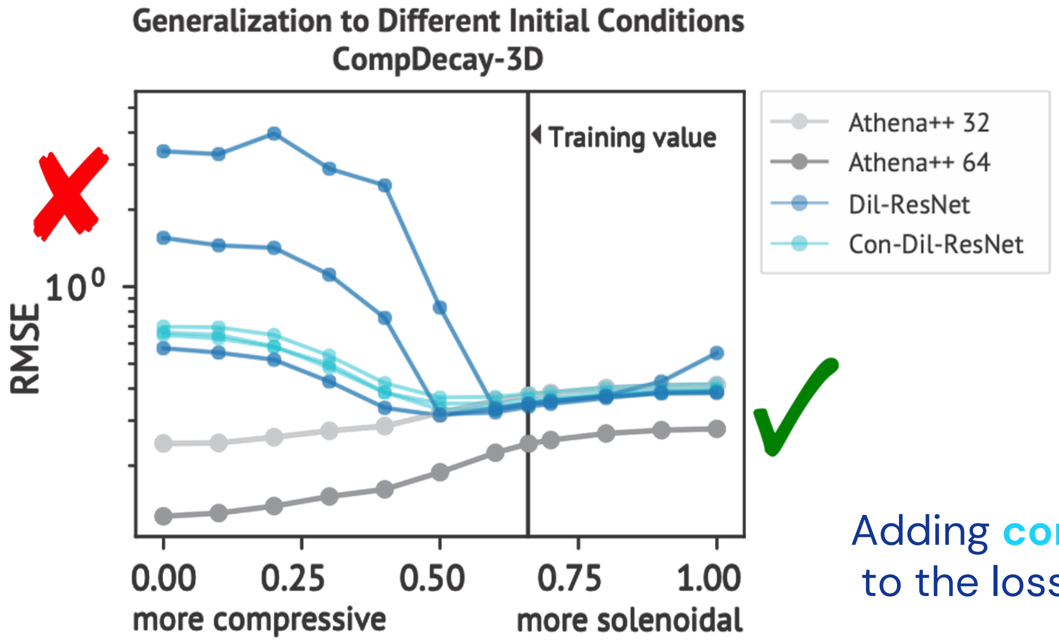
Learned model timestep

1x 2x 4x 8x 16x 32x 64x 128x 256x



Stachenfeld, Fielding, Kochkov, Cranmer, Pfaff, Godwin, Cui, Ho, Battaglia, Sanchez-Gonzalez (ICLR 2022)

Generalization to different initial conditions



Generalization to **more solenoidal** but not **more compressive** components in the initial conditions.

Adding **constraints** (conserving Total Energy) to the loss can help limit generalization error.

Stachenfeld, Fielding, Kochkov, Cranmer, Pfaff, Godwin, Cui, Ho, Battaglia, Sanchez-Gonzalez (ICLR 2022)

Part I: Conclusions

- Domain-general coarse-grained learned simulators
  - Can capture a **variety** of types of challenging turbulence
  - Outperform classical models in terms of **accuracy and speed**
  - Especially for preserving **high frequency information**
- Stability
  - **Training noise** helps
  - **Temporal downsampling** helps
- Generalization
  - **Constraints** help with generalization to different initial conditions
  - **Dataset augmentation** helps with generalization to different box sizes
  - Generalization remains a challenge

Stachenfeld, Fielding, Kochkov, Cranmer, Pfaff, Godwin, Cui, Ho, Battaglia, Sanchez-Gonzalez (ICLR 2022)

Learned Models for Inverse Design


Kelsey Allen


Tatiana Guevara-Lopez

Allen, Guevara-Lopez*, Stachenfeld*, Sanchez-Gonzalez, Battaglia, Hamrick, Pfaff (2022)*

Learned Models for Inverse Design

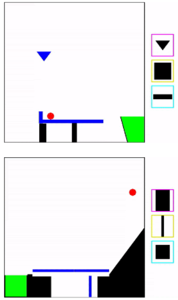
Kelsey Allen

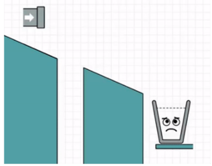
Tatiana Guevara-Lopez

Infrastructure

Engineering

Human Reasoning





"HappyGlass"

Allen*, Smith*, Tenenbaum, PNAS 2020

Allen, Guevara-Lopez*, Stachenfeld*, Sanchez-Gonzalez, Battaglia, Hamrick, Pfaff (2022)*

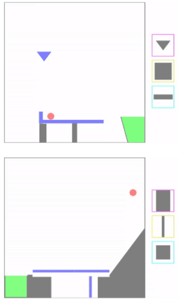
Learned Models for Inverse Design

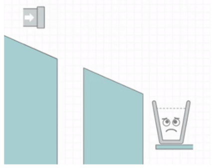

Kelsey Allen Tatiana Guevara-Lopez

Infrastructure

Engineering

Human Reasoning

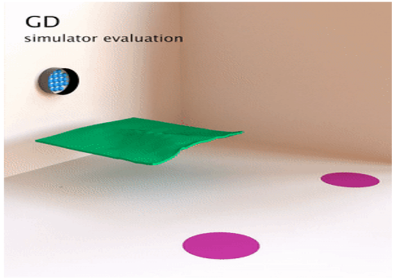




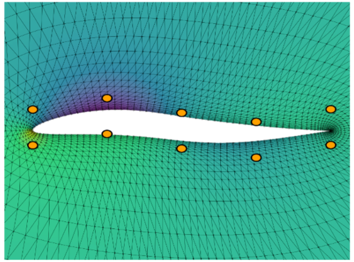
“HappyGlass”

Allen*, Smith*, Tenenbaum, PNAS 2020

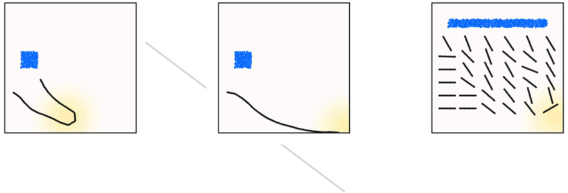
3D WaterCourse



Airfoil Shape Optimization



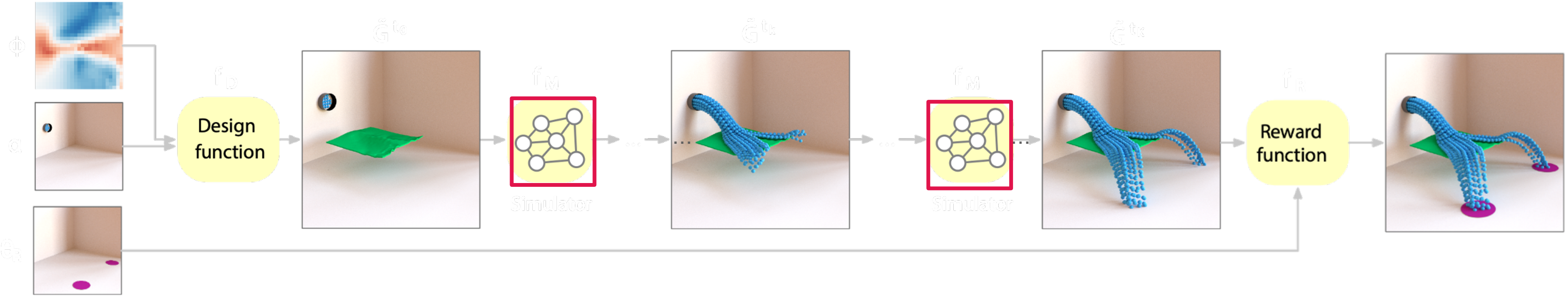
2D Fluid Tools



Can **learned simulators** can be used to solve challenging physical design problems?

Allen, Guevara-Lopez*, Stachenfeld*, Sanchez-Gonzalez, Battaglia, Hamrick, Pfaff (2022)*

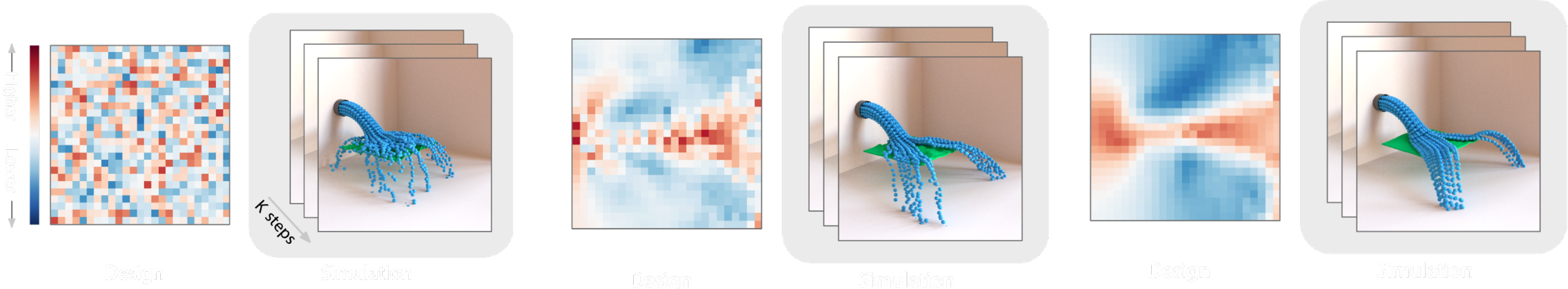
Design Evaluation & Optimization



Parameterize a design space

Learned simulator produces a rollout

Evaluate reward



Across design iterations, **design parameters are optimized** to maximize reward.

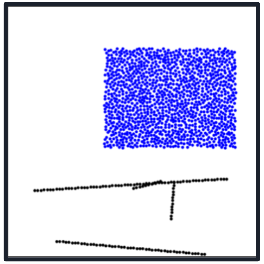
Allen, Guevara-Lopez*, Stachenfeld*, Sanchez-Gonzalez, Battaglia, Hamrick, Pfaff (2022)*

Task Agnostic Model Training with GNNs

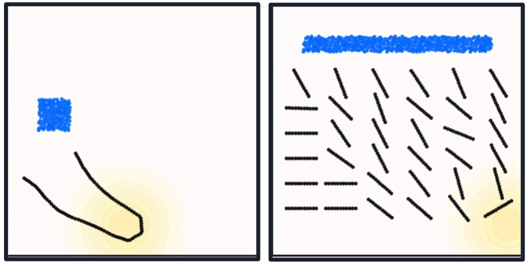
GNN -based learned simulators
- work for many types of physics
- efficient, accurate, stable
- differentiable, permitting gradient-based design optimization
- **generalize**

Simulator is **pre-trained** on **next-step prediction** with data **qualitatively different** from scenes encountered during design.

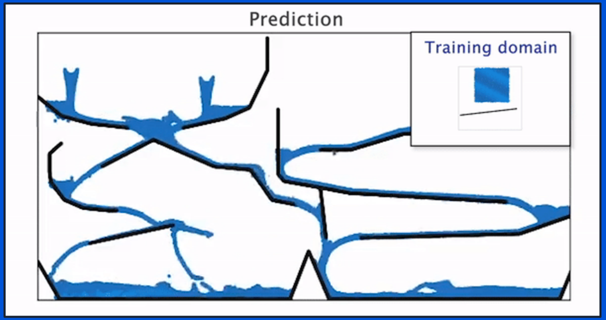
2D Fluid Tools
Training data



2D Fluid Tools
Example design tasks



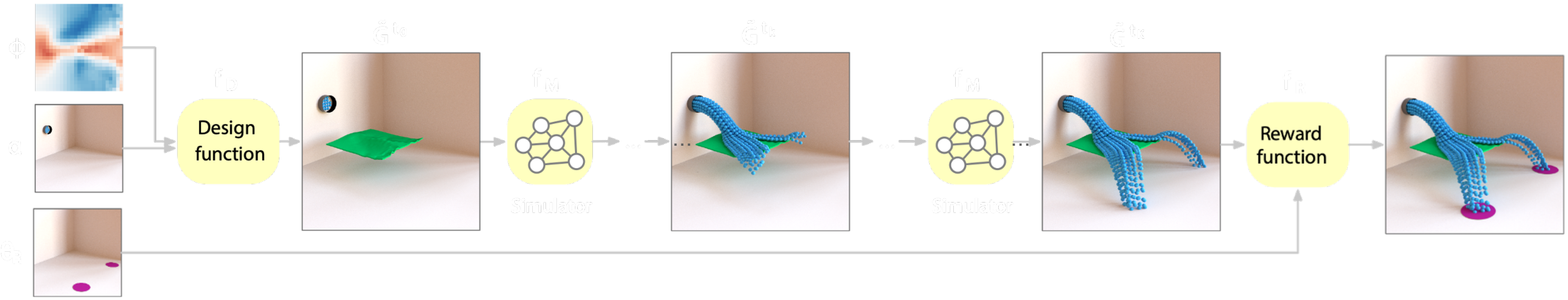
Generalization with GNN simulators



Pfaff*, Fortunato*, *et al*, ICLR 2021;
Sanchez-Gonzalez*, Godwin*, Pfaff*, Ying*, *et al*, ICML 2020

Allen, Guevara-Lopez*, Stachenfeld*, Sanchez-Gonzalez, Battaglia, Hamrick, Pfaff (2022)*

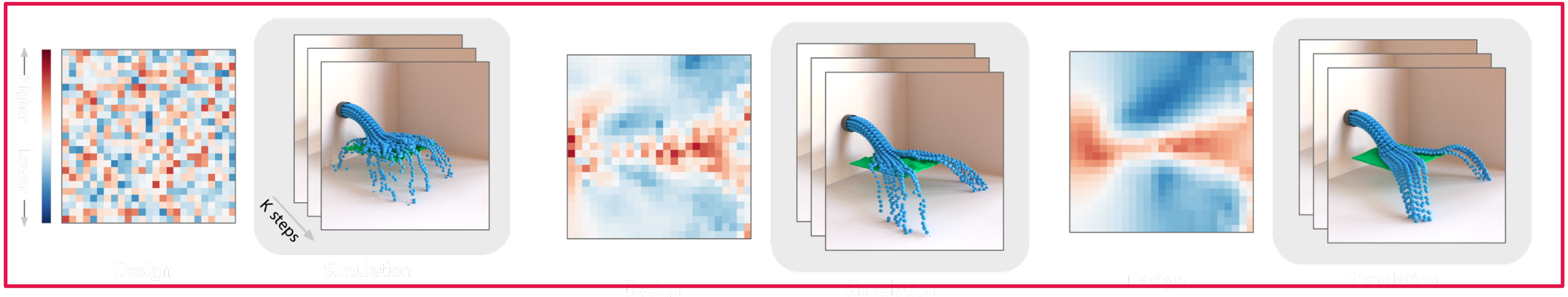
Design Evaluation & Optimization



Parameterize a design space

Learned simulator produces a rollout

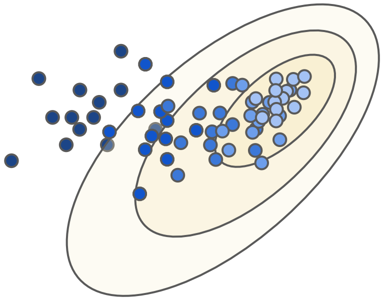
Evaluate reward

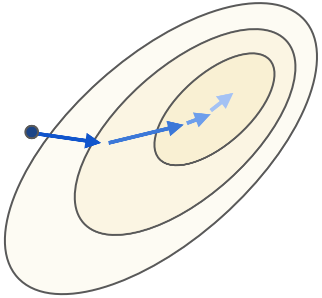


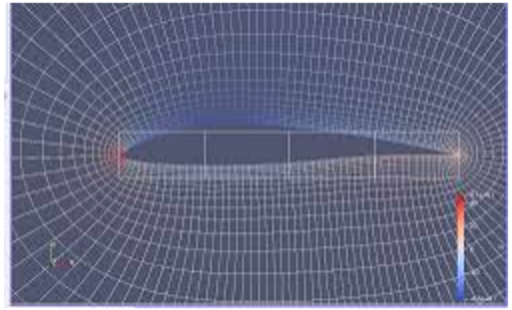
Across design iterations, **design parameters are optimized** to maximize reward.

Allen, Guevara-Lopez*, Stachenfeld*, Sanchez-Gonzalez, Battaglia, Hamrick, Pfaff (2022)*

Approaches to Design Optimization







Sampling-based with black-box forward models

Cross-entropy method
(De Boer et al, 2005)

Gradient-based with learned models

Gradient descent with ADAM
(Kingma & Ba, 2014)

Gradient-based with hand-crafted models

DAFoam, specialied for aerodynamics
(He et al, 2020)

Allen, Guevara-Lopez*, Stachenfeld*, Sanchez-Gonzalez, Battaglia, Hamrick, Pfaff (2022)*

2D Fluid Tools

100 - 1000 particles, 16-36 design dimensions

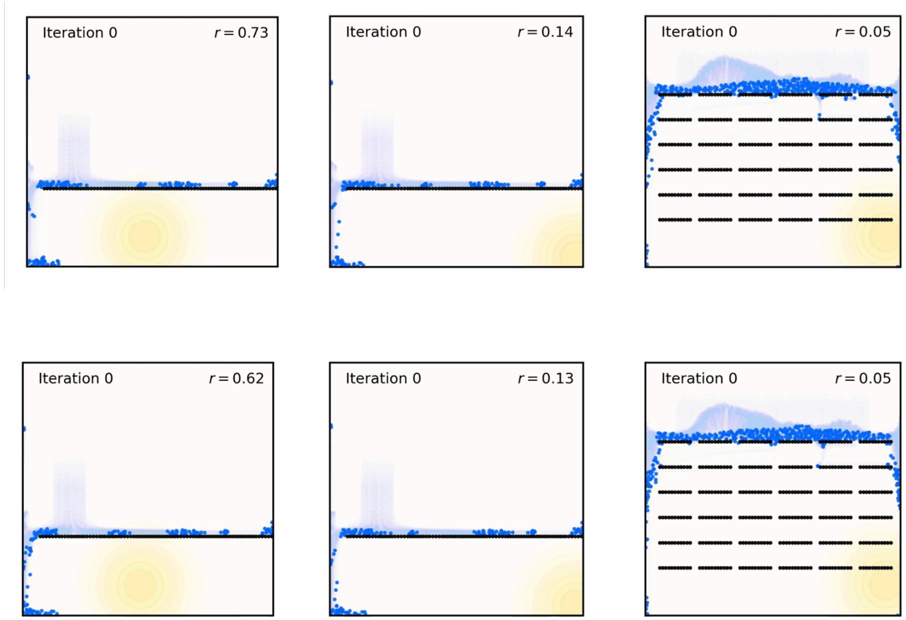
Allen, Guevara-Lopez*, Stachenfeld*, Sanchez-Gonzalez, Battaglia, Hamrick, Pfaff (2022)*

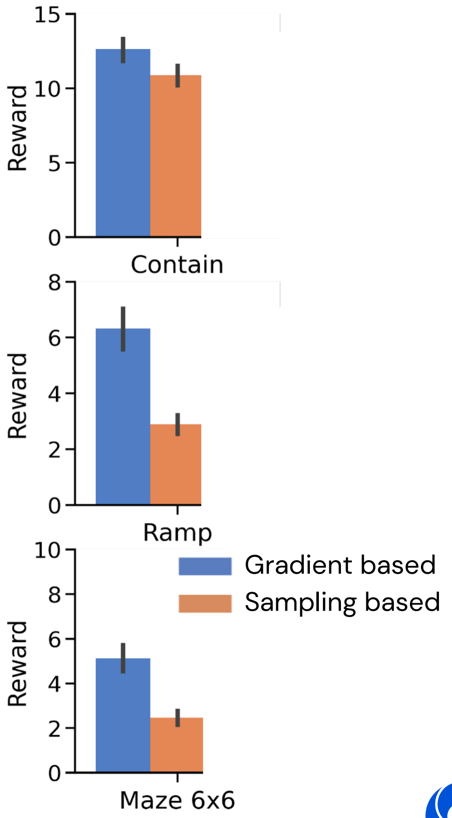
2D Fluid Tools

100 – 1000 particles, 16–36 design dimensions

Gradient-based optimization
GD

Sampling-based optimization
CEM



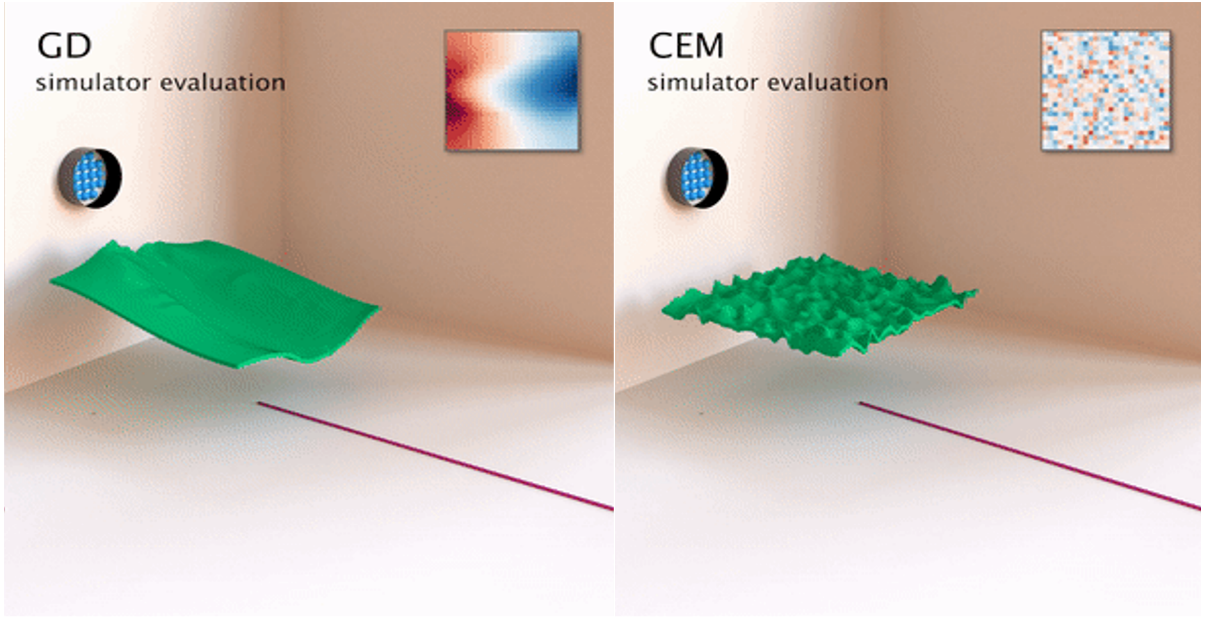


Gradient-based optimization finds **smoother, more accurate** designs than the sampling-based approach (CEM).

Allen, Guevara-Lopez*, Stachenfeld*, Sanchez-Gonzalez, Battaglia, Hamrick, Pfaff (2022)*

3D WaterCourse

2k - 4k particles, 625 design dimensions



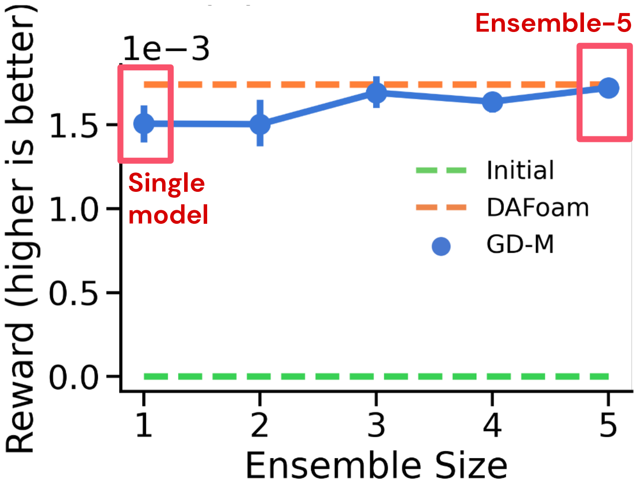
Gradient-based optimization (GD) with the learned simulator can solve **high-dimensional design tasks** where a sampling-based approach (CEM) is intractable.

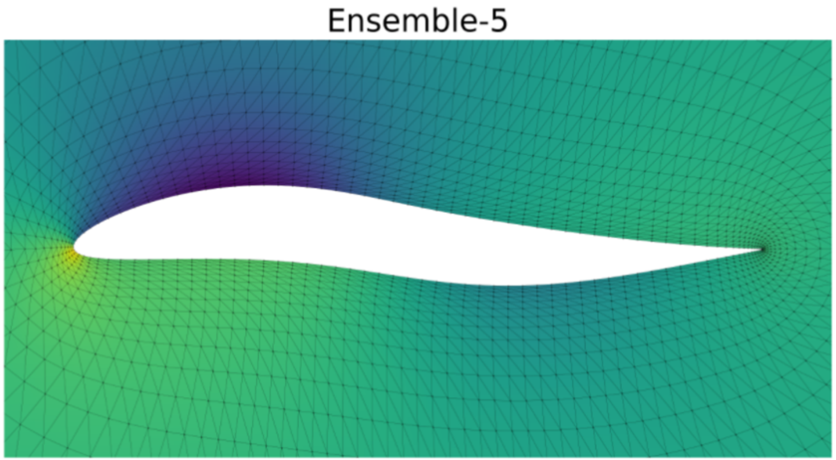
Allen, Guevara-Lopez*, Stachenfeld*, Sanchez-Gonzalez, Battaglia, Hamrick, Pfaff (2022)*

Airfoil Shape Optimization

~4k particles, 12 design dimensions

Comparable designs to specialized DAFoam* Solver
(16-48x faster on a single A100 GPU than DAFoam on an 8-core workstation)





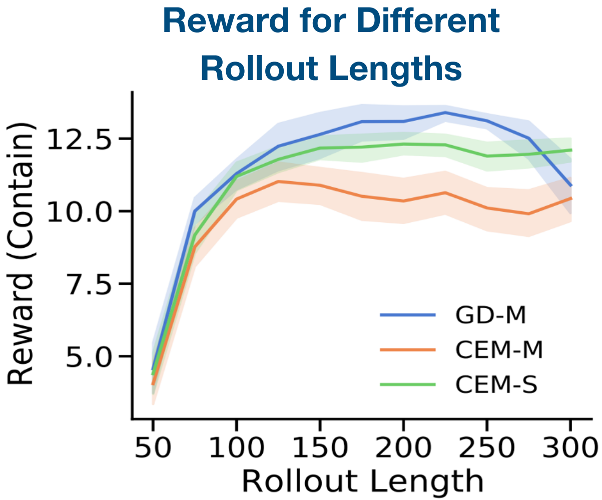
Simple tricks like **model ensembles** can yield an extra level of accuracy.

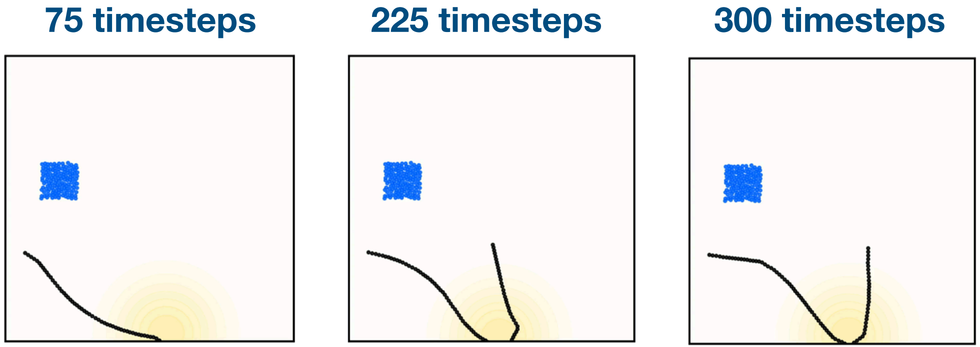
Note the sharper wing tip, stronger S-shape

*DAFoam (He et al, AIAA 2020)

Allen, Guevara-Lopez*, Stachenfeld*, Sanchez-Gonzalez, Battaglia, Hamrick, Pfaff (2022)*

Stable Gradients over long rollouts





Solutions from gradient-based optimization (GD) continue to improve up to 225 timesteps, and outperform sampling-based CEM up to 275 timesteps.

Allen, Guevara-Lopez*, Stachenfeld*, Sanchez-Gonzalez, Battaglia, Hamrick, Pfaff (2022)*

Part II: Conclusions

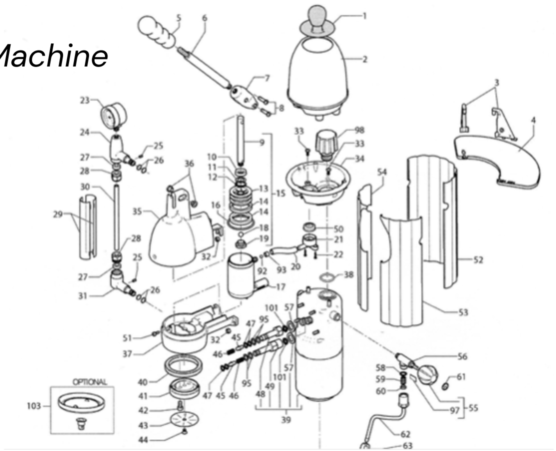
- GNN-based learned simulators can support **general-purpose design** over a **variety of challenging physical domains**
  - Problems feature **high dimensional, complex** state spaces and design spaces
  - Useful gradients over **100s of timesteps**
- Task-agnostic training on data still permits **out-of-distribution** design
- Match the **accuracy of specialized solvers** on airfoil shape optimization
  - Solutions obtained **efficiently**
  - Model **ensembles** can achieve an extra level of high accuracy

Allen, Guevara-Lopez*, Stachenfeld*, Sanchez-Gonzalez, Battaglia, Hamrick, Pfaff (2022)*

Next directions

- New domains with new, challenging types of physics
- Exploring more robust optimization procedures
    - Gradient descent suffers from zero or noisy gradients, local optima
- New models of dynamics and design spaces
    - Learned forward models that are better optimized for design tasks
    - Rich models of the design space that support hierarchical, compositional design

Espresso Machine



Acknowledgements

Contact:
stachenfeld@deepmind.com
@neuro_kim

Kelsey Allen

Tatiana Guevara-Lopez

Tobias Pfaff

Alvaro Sanchez-Gonzalez

Jessica Hamrick

Peter Battaglia

DeepMind

FLATIRON INSTITUTE

Google Research

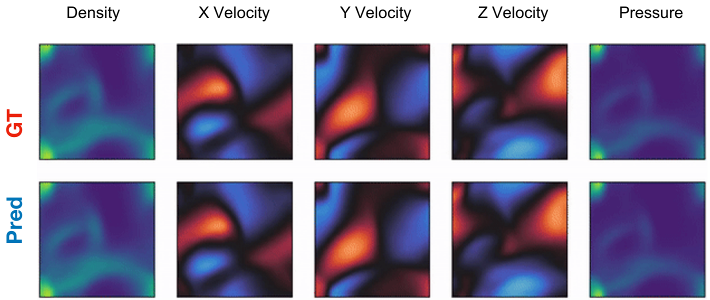
Check out the papers:

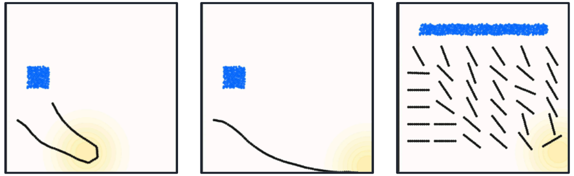
Learned Coarse Models for Efficient Turbulence Simulation. Stachenfeld, Fielding, Kochkov, Cranmer, Pfaff, Godwin, Cui, Ho, Battaglia, Sanchez-Gonzalez (ICLR 2022). arxiv.org/abs/2112.15275

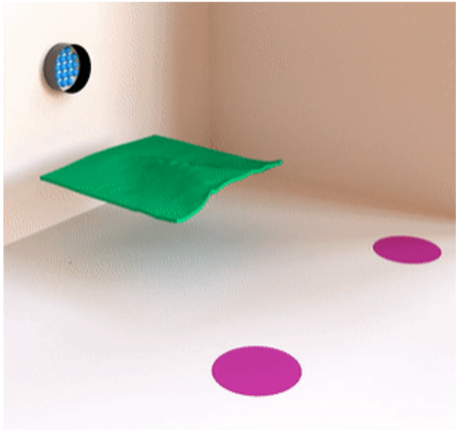
Physical Design using Differentiable Learned Simulators. Allen*, Guevara-Lopez*, Stachenfeld*, Sanchez-Gonzalez, Battaglia, Hamrick, Pfaff (2022). arxiv.org/abs/2202.00728

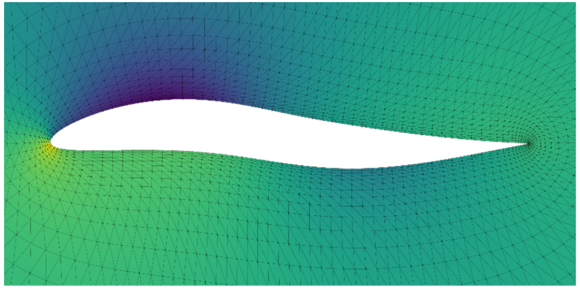
Questions

Contact:
stachenfeld@deepmind.com
@neuro_kim









DeepMind

FLATIRON INSTITUTE

Google Research

Check out the papers:

Learned Coarse Models for Efficient Turbulence Simulation. Stachenfeld, Fielding, Kochkov, Cranmer, Pfaff, Godwin, Cui, Ho, Battaglia, Sanchez-Gonzalez (ICLR 2022).
arxiv.org/abs/2112.15275

Physical Design using Differentiable Learned Simulators. Allen*, Guevara-Lopez*, Stachenfeld*, Sanchez-Gonzalez, Battaglia, Hamrick, Pfaff (2022).
arxiv.org/abs/2202.00728