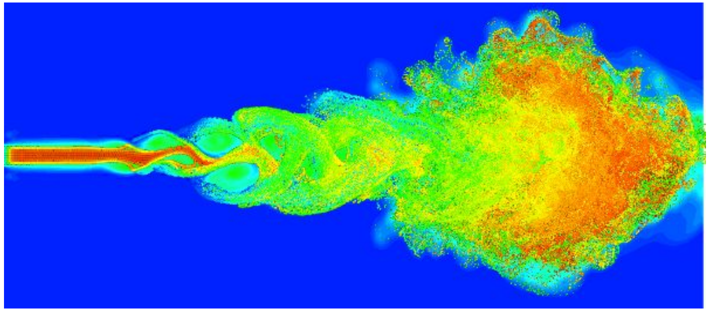
Structure-preserving neural networks: energy conservation for turbulence closure models

Toby van Gastelen, Wouter Edeling, Benjamin Sanderse

CWI

Why?



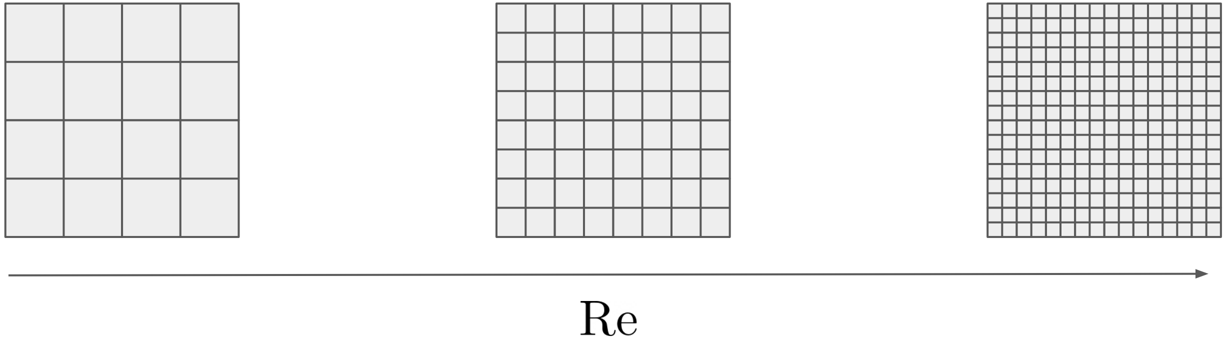
We are dealing with problems of the form

$$\frac{\partial u}{\partial t} = f(u, t),$$
$$u(\mathbf{x}, 0) = u_0(\mathbf{x})$$

Discretizing this system gives

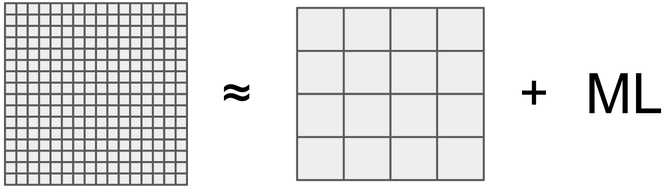
$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = f_h(\mathbf{u}, t),$$
$$\mathbf{u}(0) = \mathbf{u}_0$$

Problem: system can be very high dimensional



Machine learning solution

- Machine learning in LES framework



Name	Loss	Steps	MSE at t_1	MSE at t_2
LR	-	-	2.78e−3	0.057
LES	-	-	2.69e−3	0.051
$\mathrm{NN}_{\mathrm{sup,T}}$	$\mathcal{L}_{\mathrm{T}}$	1	1.52e−3	0.369
$\mathrm{NN}_{1,\mathrm{T}}$	$\mathcal{L}_{\mathrm{T}}$	1	1.65e−3	0.046
NN_{10}	$\mathcal{L}_{2}$	10	4.23e−4	0.018
$\mathrm{NN}_{10,\mathrm{T}}$	$\mathcal{L}_{\mathrm{T}}$	10	4.25e−4	0.022
$\mathrm{NN}_{30,\mathrm{T}}$	$\mathcal{L}_{\mathrm{T}}$	30	4.09e−4	0.021

- **Problem: machine learning representation of closure model can be unstable**
- Current approaches include:
  - Stability training on data with artificial noise *(Marius Kurz, Andrea Beck, 2021)*
  - Minimizing (or eliminating) backscatter *(Jonghwan Park, Haecheon Choi, 2021)*
  - trajectory fitting (*Björn List, Li-Wei Chen, Nils Thuerey, 2022*)

Idea: build physical structure, such as energy conservation, into ML model to boost accuracy and stability

CWI

Korteweg-De Vries equation + filtering

Governing equation

- Korteweg-de Vries (KdV) equation[1]

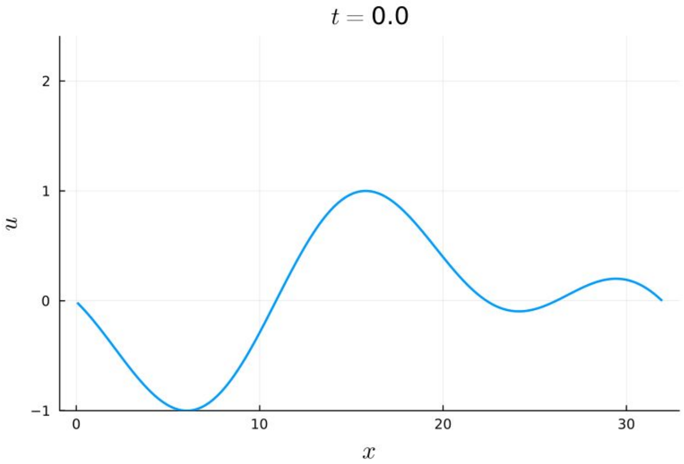
$$\frac{\partial u}{\partial t} + 3\frac{\partial u^2}{\partial x} = -\frac{\partial^3 u}{\partial x^3}$$

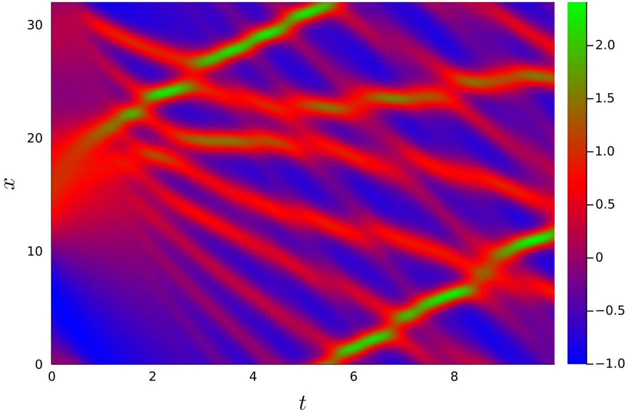
- Exhibits soliton solutions
- Energy conserving (periodic BCs):

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\underbrace{\frac{1}{2}\int_\Omega u^2 d\Omega}_{=:E} = 0$$

- Discretized using skew-symmetric scheme

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = -3\mathbf{G}(\mathbf{u}) - \mathbf{D}_3\mathbf{u}$$

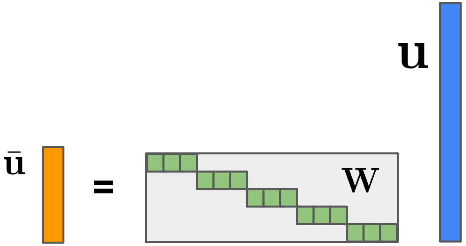


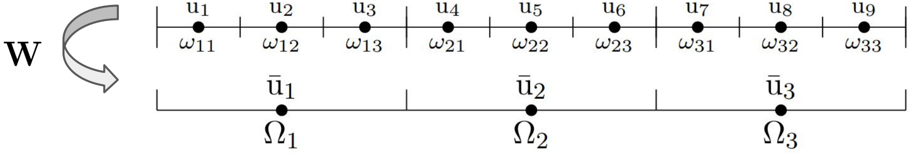


[1] PDE parameters taken from "Learning data driven discretizations for partial differential equations" (*Bar-Sinai 2019*)

Filtering + closure term

- Filtering reduces dimensionality of the problem
- Our filter: discrete spatial average represented by **W**





- Filter commutes with time-derivative

$$\mathbf{W}\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = \frac{\mathrm{d}(\mathbf{W}\mathbf{u})}{\mathrm{d}t} = \frac{\mathrm{d}\bar{\mathbf{u}}}{\mathrm{d}t}$$

$$\frac{\mathrm{d}\bar{\mathbf{u}}}{\mathrm{d}t} = f_H(\bar{\mathbf{u}},t) + \underbrace{(\mathbf{W}f_h(\mathbf{u},t) - f_H(\bar{\mathbf{u}},t))}_{=:\mathbf{c}(\mathbf{u},t)}$$

$$\frac{\mathrm{d}\bar{\mathbf{u}}}{\mathrm{d}t} = \underbrace{-3\mathbf{G}(\bar{\mathbf{u}}) - \mathbf{D}_3\bar{\mathbf{u}}}_{=f_H(\bar{\mathbf{u}})} + \mathbf{c}(\mathbf{u})$$

- Filter does not commute with spatial discretization

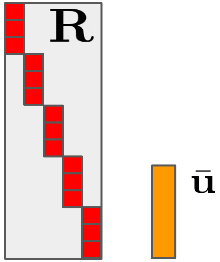
$$\mathbf{W}f_h(\mathbf{u},t) \neq f_H(\mathbf{W}\mathbf{u},t) = f_H(\bar{\mathbf{u}},t)$$

Closure term

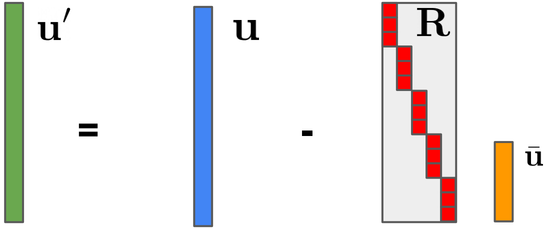
Discretization on coarse grid

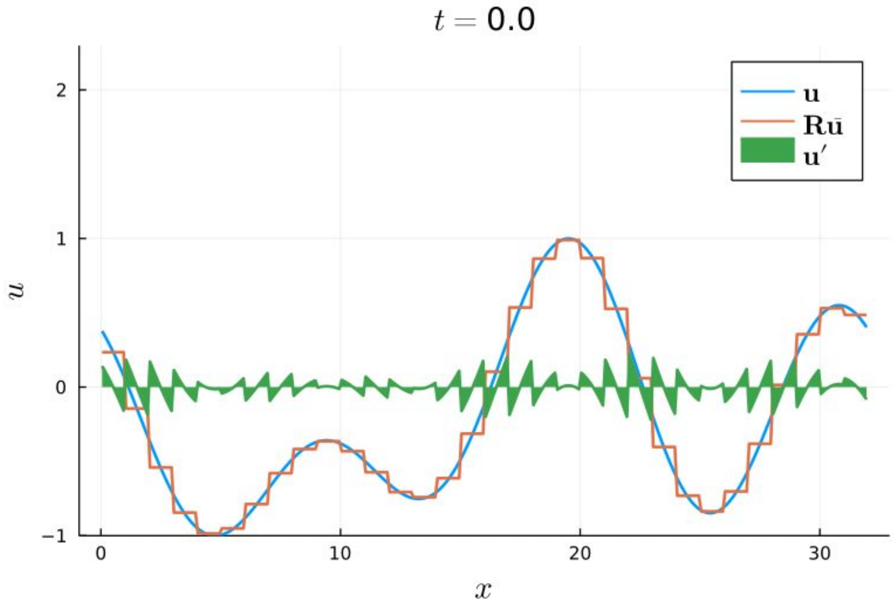
Reconstruction

- Reconstruction operator **R**: projects $\bar{\mathbf{u}}$ onto fine grid



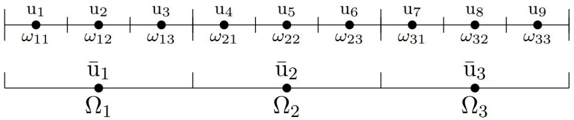
- Definition of subgrid-scale content





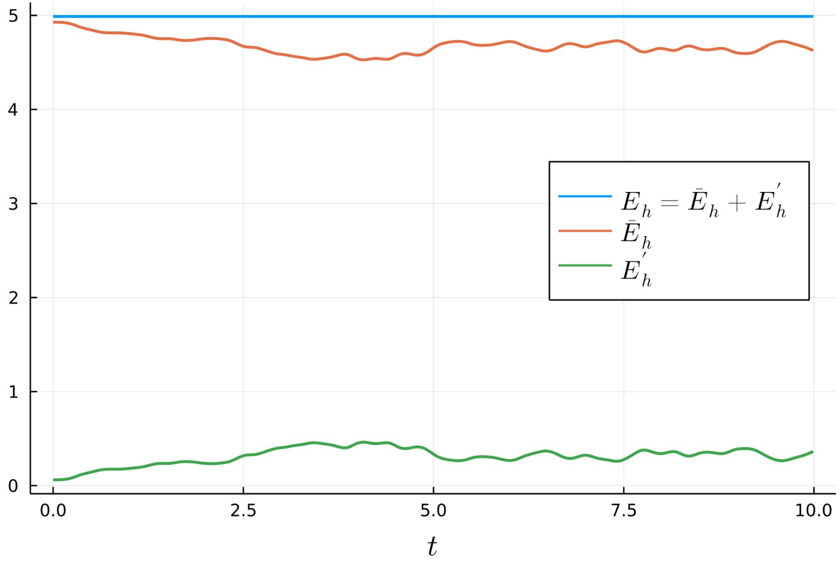
Energy decomposition

- Due to choice of filter we can decompose energy:



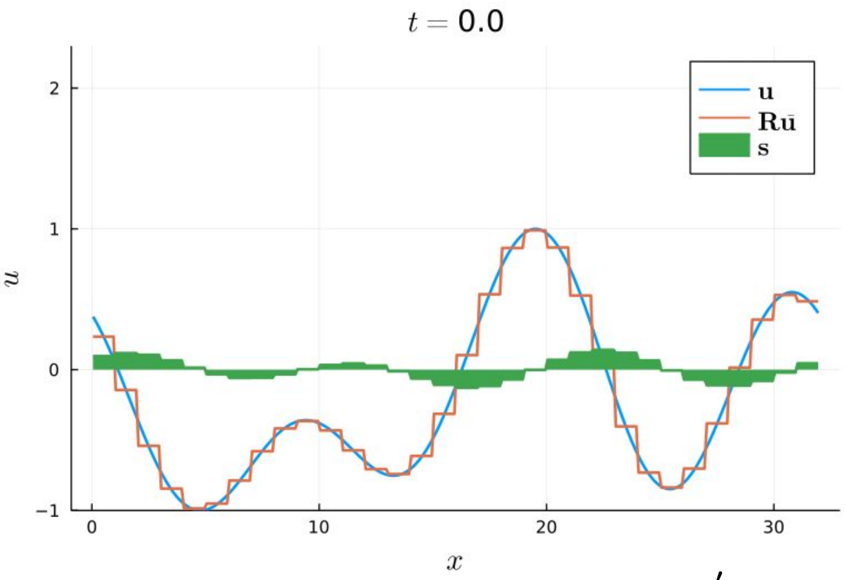
$$E_h = \underbrace{\frac{1}{2}(\bar{\mathbf{u}}, \bar{\mathbf{u}})_\Omega}_{=:\bar{E}_h} + \underbrace{\frac{1}{2}(\mathbf{u}', \mathbf{u}')_\omega}_{=:E'_h}$$

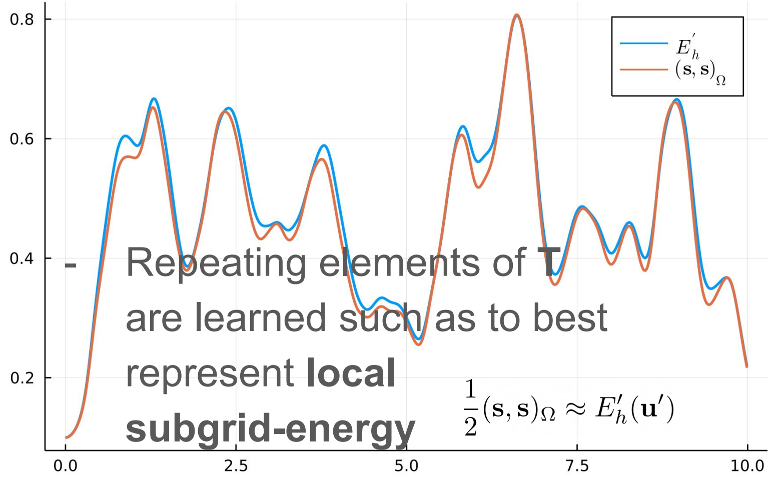
- Leads to the following for the time-evolution



$$\frac{\mathrm{d}E_h}{\mathrm{d}t} = \boxed{\frac{\mathrm{d}\bar{E}_h(\bar{\mathbf{u}})}{\mathrm{d}t}} + \boxed{\frac{\mathrm{d}E'_h(\mathbf{u}')}{\mathrm{d}t}} = 0$$

Subgrid-compression: practice



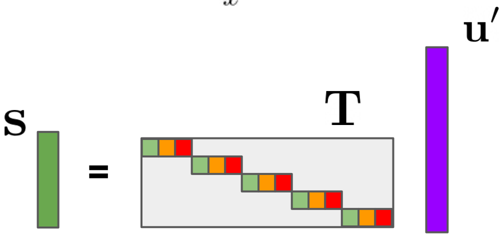


- Repeating elements of **T** are learned such as to best represent **local subgrid-energy**

$$\frac{1}{2}(\mathbf{s},\mathbf{s})_\Omega \approx E_h'(\mathbf{u}')$$

$$= \arg\min \sum_{d=1}^{\mathcal{D}} \sum^{t} \left|\left| \frac{1}{2}\mathbf{s}_d^2 - \frac{1}{2}\mathbf{W}(\mathbf{u}_d')^2 \right|\right|_2^2$$

$$\cancel{\frac{\mathrm{d}E_h(\bar{\mathbf{u}})}{\mathrm{d}t} + \frac{\mathrm{d}E_h'(\mathbf{u}')}{\mathrm{d}t} = 0}$$



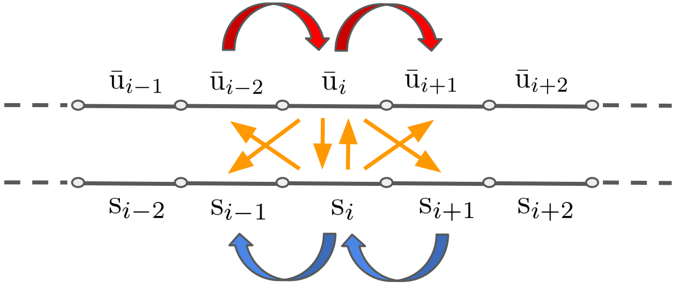
$$\frac{\mathrm{d}\bar{E}_h(\bar{\mathbf{u}})}{\mathrm{d}t} + \frac{1}{2}\frac{\mathrm{d}(\mathbf{s},\mathbf{s})_\omega}{\mathrm{d}t} = 0$$

CWI

Closure modelling

Skew-symmetric framework

- Closure model represented by local **energy** exchange (follows the arrows)



CNN

$$\mathcal{K}(\bar{\mathbf{u}}, \mathbf{s}; \mathbf{\Theta}) = \begin{bmatrix} \mathbf{K}_1 & \mathbf{K}_2 \\ -\mathbf{K}_2^T & \mathbf{K}_3 \end{bmatrix}$$

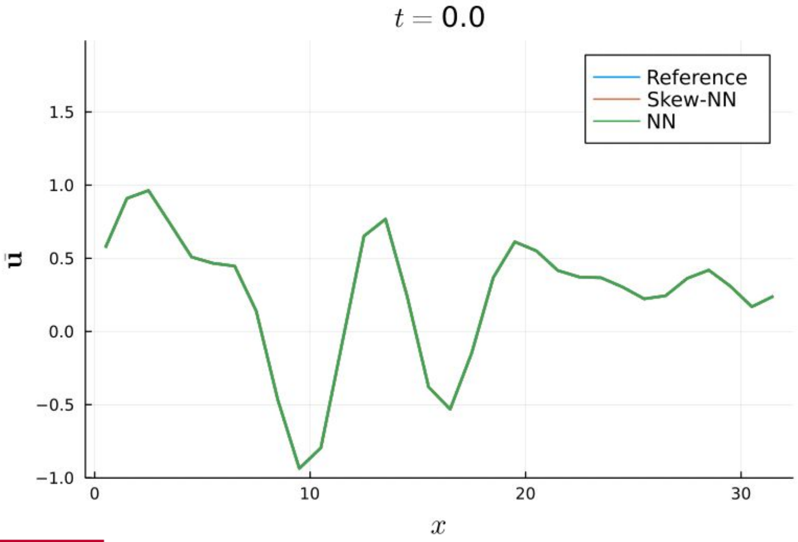
- Represented by skew-symmetric matrix multiplication $\mathcal{K} = \mathbf{K} - \mathbf{K}^T$

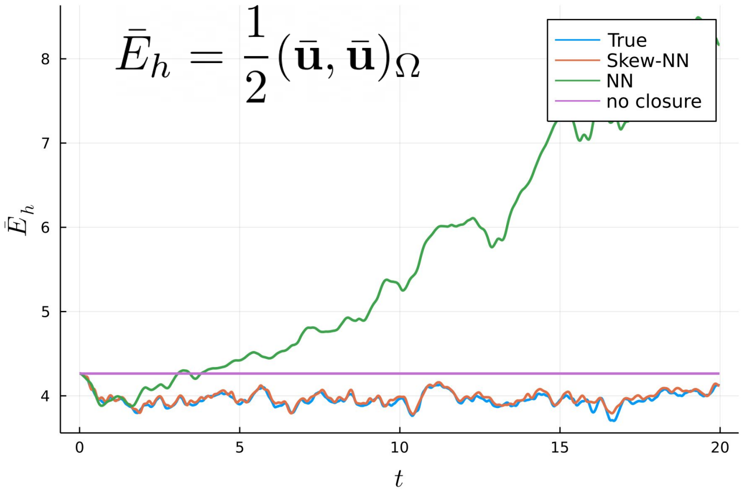
$$\begin{bmatrix} \frac{\mathrm{d}\bar{\mathbf{u}}}{\mathrm{d}t} \\ \frac{\mathrm{d}\mathbf{s}}{\mathrm{d}t} \end{bmatrix} = \begin{bmatrix} f_H(\bar{\mathbf{u}}) \\ \mathbf{0} \end{bmatrix} + \boxed{\mathcal{K}(\bar{\mathbf{u}}, \mathbf{s}; \mathbf{\Theta})} \begin{bmatrix} \bar{\mathbf{u}} \\ \mathbf{s} \end{bmatrix} \quad \Rightarrow \quad \frac{\mathrm{d}\bar{E}_h(\bar{\mathbf{u}})}{\mathrm{d}t} + \frac{1}{2}\frac{\mathrm{d}(\mathbf{s},\mathbf{s})_\omega}{\mathrm{d}t} = 0$$

Results

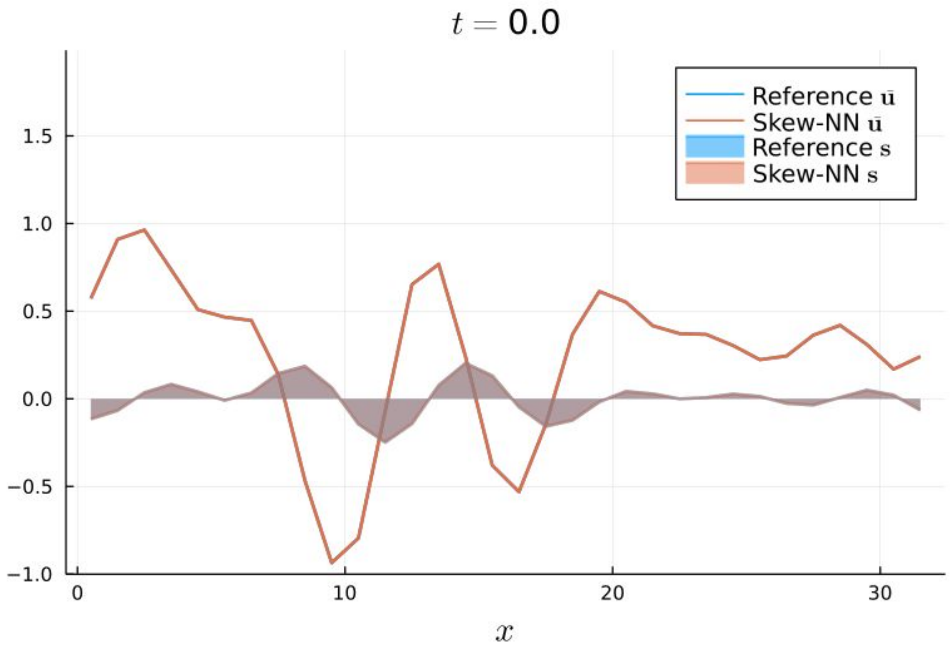
Presented framework improves quality+stability of solution

- Trained on many different initial conditions and tested on unseen initial conditions
- Reduction from N = 600 -> N = 30
- Compare to standard CNN





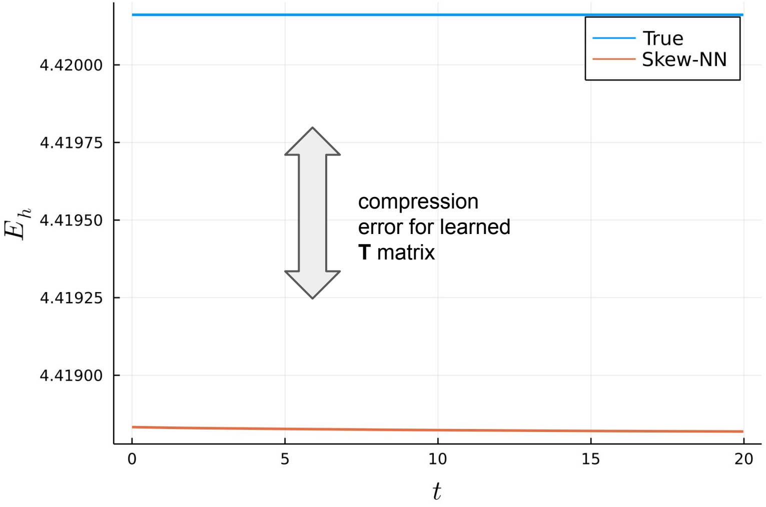
Evolution of the subgrid-content matches nicely

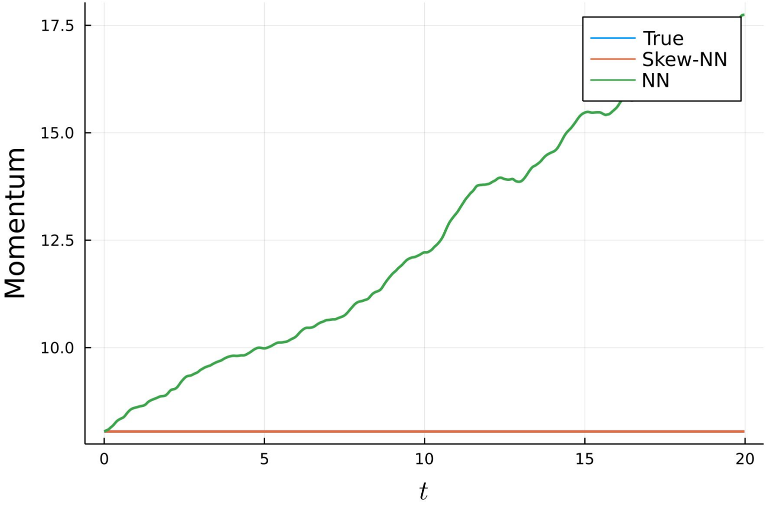


$$\begin{bmatrix} \frac{\mathrm{d}\bar{\mathbf{u}}}{\mathrm{d}t} \\ \frac{\mathrm{d}\mathbf{s}}{\mathrm{d}t} \end{bmatrix} = \begin{bmatrix} f_H(\bar{\mathbf{u}}) \\ \mathbf{0} \end{bmatrix} + \mathcal{K}(\bar{\mathbf{u}}, \mathbf{s}; \boldsymbol{\Theta}) \begin{bmatrix} \bar{\mathbf{u}} \\ \mathbf{s} \end{bmatrix}$$

CWI

Skew-symmetric NN indeed preserves relevant structure



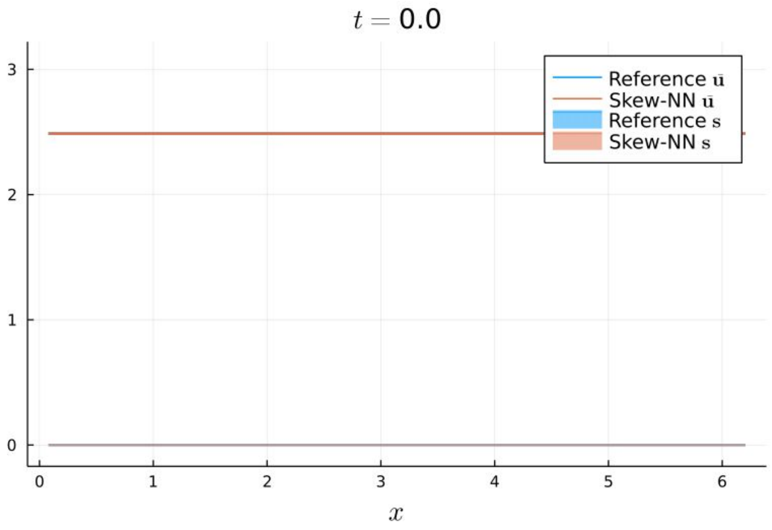


$$\frac{\mathrm{d}\bar{E}_h(\bar{\mathbf{u}})}{\mathrm{d}t} + \frac{1}{2}\frac{\mathrm{d}(\mathbf{s},\mathbf{s})_\omega}{\mathrm{d}t} = 0$$

$$(\mathbf{1}, \frac{\mathrm{d}\bar{\mathbf{u}}}{\mathrm{d}t})_\Omega = 0$$

Model can be enhanced with diffusive terms to work on Burgers’ equation with viscosity

- Reduction from N = 1000 -> N = 40



$$\mathcal{P} = -\mathcal{Q}^T\mathcal{Q}$$

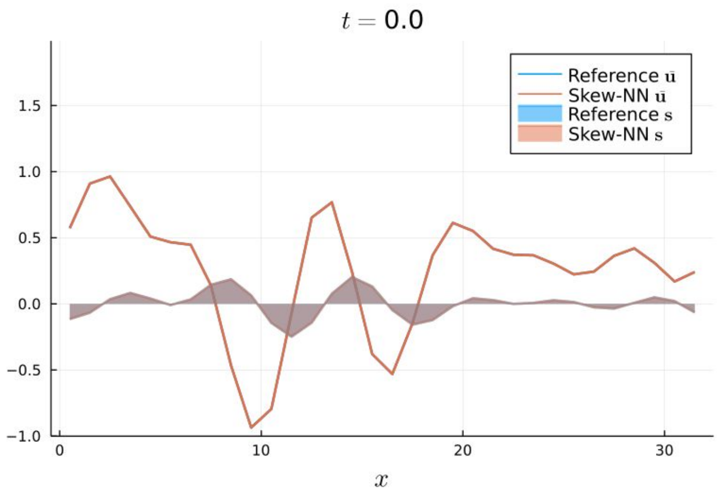
Strictly dissipative form based on Cholesky-decomposition

$$\begin{bmatrix} \frac{\mathrm{d}\bar{\mathbf{u}}}{\mathrm{d}t} \\ \frac{\mathrm{d}\mathbf{s}}{\mathrm{d}t} \end{bmatrix} = \begin{bmatrix} f_H(\bar{\mathbf{u}}) \\ \mathbf{0} \end{bmatrix} + \mathcal{K}(\bar{\mathbf{u}}, \mathbf{s}; \boldsymbol{\Theta}) \begin{bmatrix} \bar{\mathbf{u}} \\ \mathbf{s} \end{bmatrix} + \boxed{\mathcal{P}(\bar{\mathbf{u}}, \mathbf{s}; \boldsymbol{\Theta}) \begin{bmatrix} \bar{\mathbf{u}} \\ \mathbf{s} \end{bmatrix}}$$

Conclusions

Conclusions

- Introduced subgrid-scale compression, representing the local unresolved energy content
- Introduced skew-symmetric CNN framework for closure models
- Framework conserves energy & momentum and is stable by design
- Successfully applied to KdV equation & Burgers' equation (BCs)
- Downsides:
    - Additional unknowns $\mathbf{s}$
    - Compression is not perfect



What next?

- Write article
- Navier-Stokes (2D/3D)