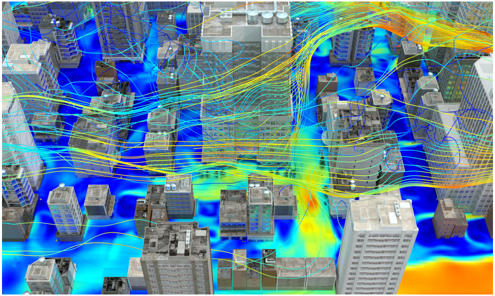
Data-driven subgrid-scale modeling for wall-bounded turbulence

Jiaxi Zhao
joint with Q. Li NUS, N. Thuerey TMU

NUS-SJTU PhD Forum
November 13, 2024

Motivation:

1. Performing DNS is unaffordable, even a LES with 50M grids of length 1000s takes several days.
2. Cheap simulation such as RANS and coarse-grid LES can not obtain accurate quantities such as peak pressure.



Can we design or learn better SGS models based on the LES data so that it can achieve accurate results even on coarse grid LES?

SGS stress modeling

There are three main issues for stress modeling:

1. The mapping from the input features. e.g. filtered velocity to the stress tensor is non-deterministic while most classical turbulence models and data-driven models are deterministic.

$$\overline{\mathbf{U}}, \nabla\overline{\mathbf{U}}, \overline{p} \overset{\text{NOT DETERMINISTIC}}{\Longrightarrow} \tau, \quad \min_{\phi} \left\| \tau - \phi(\overline{\mathbf{U}}) \right\|^2 .$$

2. Discrepancy between a-priori error and a-posteriori error.

$$\|\widehat{\tau} - \tau\|^2 , \quad \left\| \widehat{\overline{\mathbf{U}}}(T) - \overline{\mathbf{U}}(T) \right\|^2$$

3. Difficult to combine the OpenFOAM solver with gradient-based optimization algorithms.

Learning the SGS stress model

We test the following three approaches:

1. Directly predict the stress tensor from the input features.

$$\tau = \mathrm{NN}(\overline{\mathbf{U}}, \nabla\overline{\mathbf{U}}, \overline{p}). \tag{1}$$

2. Learn a correction of the constants to the Smagorinsky model.

$$\widetilde{C} = \mathrm{NN}(\overline{\mathbf{U}}, \nabla\overline{\mathbf{U}}, \overline{p}) + C, \quad \nu_t = \widetilde{C}\Delta^2 \left|\bar{S}\right|. \tag{2}$$

3. Learn a conditional generative model from the input features.

The first approach usually provides a much better a-priori error estimate than the second approach.

A toy case

In a preliminary work, we apply tangent-space regularized method to solve the NS equation using the projection method.

$$\mathbf{u}_{k+1} = \mathbf{u}_k + \Delta t(\nu \Delta \mathbf{u}_k - (\mathbf{u}_k \cdot \nabla)\mathbf{u}_k - \nabla p_k),$$
$$p_k = \phi(\mathbf{u}_k) = \Delta^{-1}(\nabla \cdot (\nu \Delta \mathbf{u}_k - (\mathbf{u}_k \cdot \nabla)\mathbf{u}_k)),$$

Instead of training the data-driven model to solely minimizing the a-priori error, we incorperate dynamical information, i.e. the iteration of the solver into the algorithm that accounts for a-posteriori error.

For fluid people, this part needs more time to explain, I skip most of the details this time.

Kuramoto–Sivashinsky equation

We use the 1D Kuramoto–Sivashinsky equation as a toy model to test the effectiveness of the proposed method.

$$
\begin{aligned}
u_t &= -(c+u)u_x - uu_x - u_{xx} - \nu u_{xxxx}, \\
u(0,t) &= u(L,t) = 0, \\
u_x(0,t) &= u_x(L,t) = 0, \forall t.
\end{aligned}
$$

This equation is known to exhibit chaotic behavior and we use the statistics of its first two moments to evaluate the performance of the SGS model.

$$
\begin{aligned}
\langle \overline{u} \rangle &= \frac{1}{LT}\int_{[0,L]}\int_t^{t+T} u(x,t)\,dtdx, \\
\left\langle \overline{u^2} \right\rangle &= \frac{1}{LT}\int_{[0,L]}\int_t^{t+T} u^2(x,t)\,dtdx,
\end{aligned}
$$

Statistics

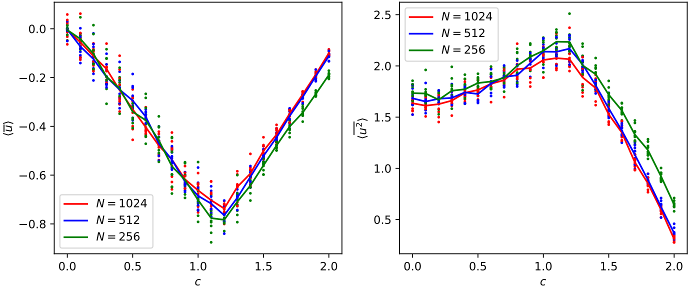
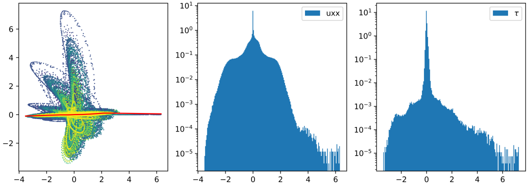


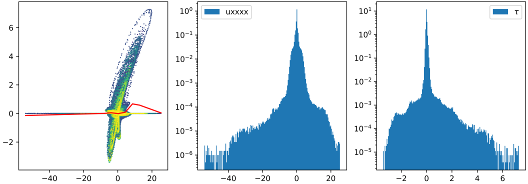
Figure: Simulation statistics under different parameters c.

A-priori and a-posteriori discrepancy

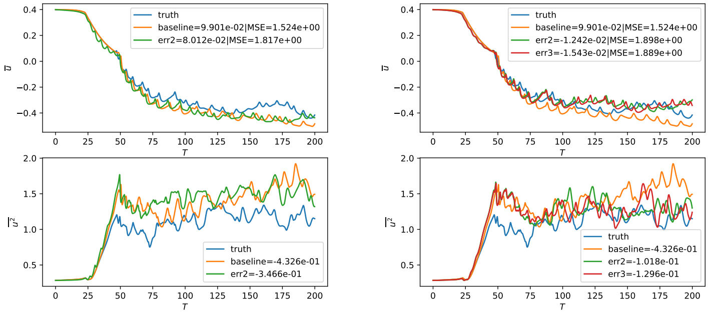
The inconsistency between the a priori error and a posteriori error arises because the training algorithm does not take the solver dynamics into account.



(a) Relation between u_{xx}, τ.



Comparison with the regression-based method



(a) (b)

Figure: Prediction τ using u_{xx}.

Future work

1. Generate a systematical dataset of the flow around bluffs; train SGS models within the dataset and design algorithms which improves the a-posteriori performance.
2. Deploy to practical problems: Subgrid-scale modeling in large eddy simulation of the urban environment.
3. Investigate the statistical and numerical properties of the data-driven turbulence modeling, with special focus on the dataset characteristic and the relation between data with different resolution.

References

M. Benjamin, S. Domino, and G. Iaccarino
Neural Networks for Large Eddy Simulations of Wall-bounded Turbulence: Numerical Experiments and Challenges
The European Physical Journal E

J. Zhao and Q. Li (2024)
Mitigating Distribution Shift in Machine Learning-augmented Hybrid Simulation
Arxiv preprint https://arxiv.org/pdf/2401.09259

S. Arisaka and Q. Li (2024)
Accelerating Legacy Numerical Solvers by Non-intrusive Gradient-based Meta-solving
International Conference on Machine Learning 2024