A probabilistic approach to subgrid-scale modeling

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

NUS-SJTU PhD Forum
November 14, 2024

Content

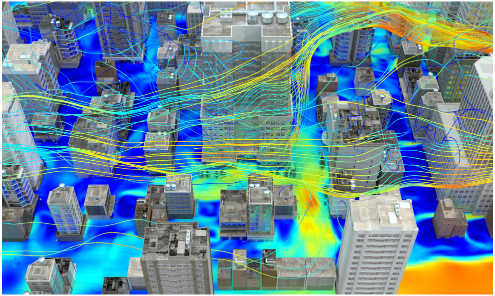
1. Difficulties encountered in the data-driven turbulence modeling.

2. Possible explanation of the problem on 1D Kuramoto–Sivashinsky equation.

3. An effective probabilistic SGS model that improves the regression-based method.

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?

Subgrid-scale stress modeling based on NN

$$
\begin{aligned}
\frac{\partial u_i}{\partial t} + \frac{\partial}{\partial x_j}(u_i u_j) &= -\frac{\partial p}{\partial x_i} + \nu \Delta u_i, \\
\frac{\partial u_i}{\partial x_i} &= 0.
\end{aligned}
$$

Applying a filter G to the equation, i.e. $\overline{u} = G * u$

$$
\begin{aligned}
\frac{\partial \overline{u}_i}{\partial t} + \frac{\partial}{\partial x_j}(\overline{u}_i \overline{u}_j) &= -\frac{\partial \overline{p}}{\partial x_i} + \nu \Delta \overline{u}_i - \frac{\partial \tau_{ij}}{\partial x_j}, \\
\frac{\partial \overline{u}_i}{\partial x_i} &= 0, \\
\tau_{ij} &= \overline{u_i u_j} - \overline{u}_i \overline{u}_j.
\end{aligned}
$$

Classical SGS stress models include the (dynamic) Smagorinsky model, and the Wall-Adapting Local Eddy-Viscosity model, and etc. They all based on physical intuition and empirical data.

Data-driven SGS stress modeling

Machine learning brings powerful approximation toolbox for the development of SGS stress models. Most existing works can be summarized to the following three categories:

1. Rely on classical turbulence models, e.g. Smagorinsky model and use DNS data to fit the model parameters.
2. Choose the input features and fit an end-to-end mapping between them and the QoI.
3. View the SGS stress modeling as a policy and solve it in the framework of reinforcement learning.

Global SGS model is impossible because of the computational cost and data sparsity.

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.

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

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

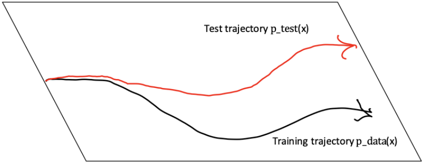
$$\|\widehat{\tau} - \tau\|^2 , \quad \|\overline{u}(T) - \overline{u}(T)\|^2$$

This is very different from classical numerical analysis perspective where a smaller truncation error (higher order) mostly indicates faster convergence via the Lax equivalence theorem.

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

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.



p_data(x) ≠ p_test(x)

The a-priori and a-posteriori performance are not consistent.

TABLE 3. Network and performance details

	Network inputs	Network outputs	*A priori* correlations	*A posteriori* simulations
NN-1	Local $\bar{S}_{ij}$	$\partial_j \tau_{ij}$	0.6	Stable; varying accuracy
NN-2	19-point stencil $\bar{S}_{ij}$	$\partial_j \tau_{ij}$	0.9	Unstable
NN-3	Local $\bar{S}_{ij}$	$L_i - D_i$ (Eq. 2.4)	0.7	Unstable

Kuramoto–Sivashinsky equation

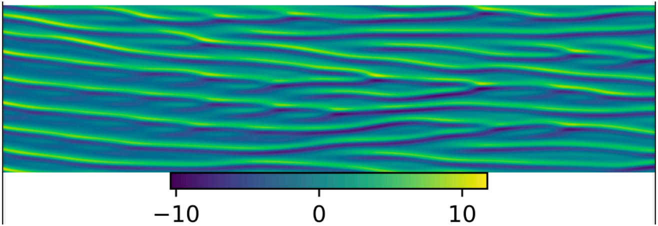
We use the 1D Kuramoto–Sivashinsky equation as a toy model to test the effectiveness of the proposed method.

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

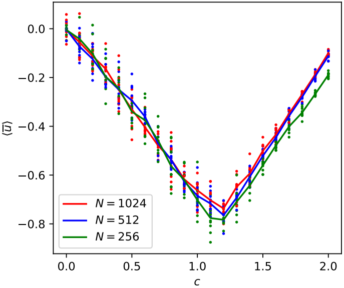
This equation is known to exhibit chaotic behavior and we use the statistics of its first two moments to evaluate the performance of the SGS model.

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

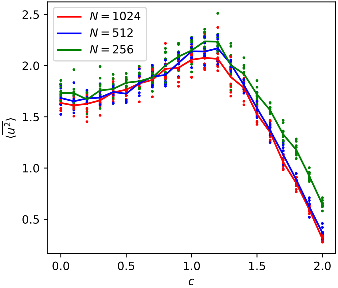
Statistics



−10
0
10



$\langle \overline{u} \rangle$
c
$N = 1024$
$N = 512$
$N = 256$
0.0
−0.2
−0.4
−0.6
−0.8
0.0
0.5
1.0
1.5
2.0



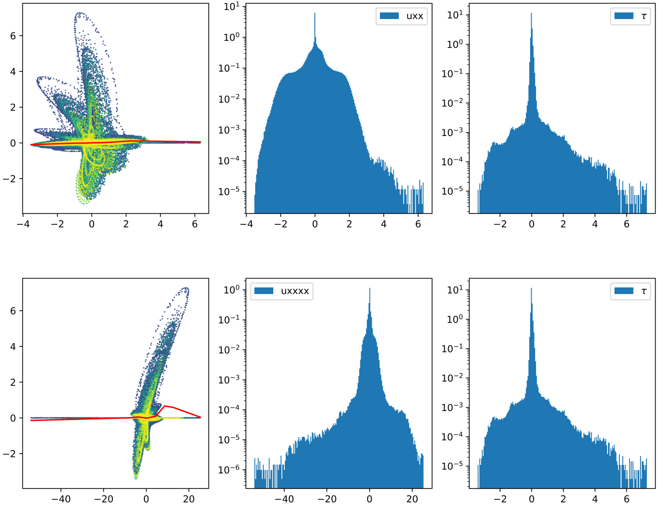
$\langle \overline{u^2} \rangle$
c
$N = 1024$
$N = 512$
$N = 256$
2.5
2.0
1.5
1.0
0.5
0.0
0.5
1.0
1.5
2.0

SGS modeling of the KS equation

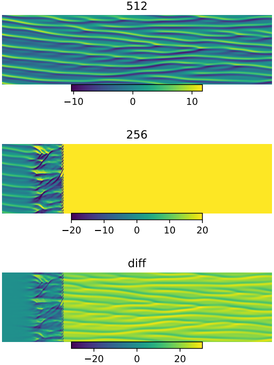
We use the simulatiosn of KS equation on two grid resolutions (256, 1024). Denote a projection operator from the fine grid to the coarse grid as P and numerical scheme as f_{1024}, f_{256}, the SGS model we want to handle is given by

$$\tau_{256} = \mathrm{P}(f_{1024}(u_{1024})) - f_{256}(\mathrm{P}\, u_{1024}) \sim p(\cdot | u_{256}). \tag{1}$$

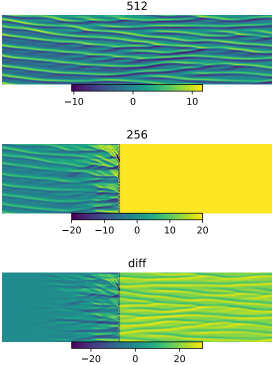
Multivalue issue



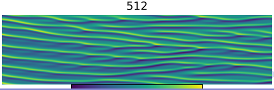
A priori and a posteriori



(a) validation loss: 3.5000e-05



(b) validation loss: 1.3050e-05



Proposed model

We model the relation between the input features and the SGS stress by a conditional probability $p_\theta(\tau|\overline{u})$ Intuitively speaking, this model can not only capture the "average" behavior of the SGS stress but also the uncertainty of the stress depends on how complex the probability family we choose. For example, by using Gaussian family, we also capture the variation of the stress.

This methodology is also recognized by the molecular dynamics community, e.g. molecular docking.

Optimizing the model

Instead of using the MSE as the loss function, we optimize our model by maximize the likelihood over data:

$$
\begin{aligned}
&\min_{\theta} \sum_{i=1}^{N} \frac{(y_i - \mu_\theta(x_i))^2}{2(\sigma_\theta(x_i))^2} + \log \sigma_\theta(x_i), \\
&\min_{\theta} \sum_{i=1}^{N} -\log\left(\sum_{j=1}^{M} \frac{\mathrm{softmax}(c_\theta^j(x_i))}{\sigma_\theta^j(x_i)} \exp\left[-\frac{(y_i - \mu_\theta^j(x_i))^2}{2(\sigma_\theta^j(x_i))^2}\right]\right).
\end{aligned}
\tag{2}
$$

Currently, we naively use the Adam solver to optimize this non-linear optimization problem.

Integrating in the simulation

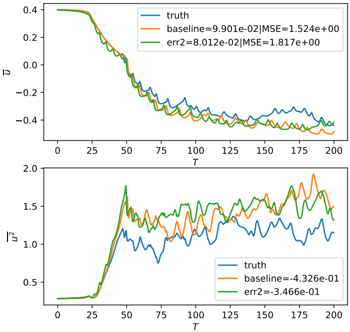
How can we deploy this model in the simulation?

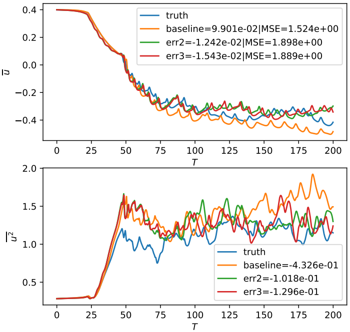
$$
\begin{aligned}
\text{Gaussian: } & u_i \Longrightarrow \mu_\theta(u_i), \sigma_\theta(u_i), z \sim N(0,1), \\
& \tau_{ij} = \mu_\theta(u_i) + \sigma_\theta(u_i) z, \\
\text{Gaussian mixture: } & u_i \Longrightarrow \mu_\theta^j(u_i), \sigma_\theta^j(u_i), z \sim N(0,1), j \sim [M], \\
& \tau_i = \mu_\theta^j(u_i) + \sigma_\theta^j(u_i) z,
\end{aligned} \tag{3}
$$

There is temporal and spatial consistency issue. Should we use the same latent variable z for all the grid points at all the time step or we should use different z_i for different u_i?

Comparison with the regression-based method

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





More experiments

	baseline	regression	gaussian, fix	gaussian sample
A priori error	NA	0.976	-2.173	-2.173
$\int \|\mathbf{u} - \mathbf{u}_0\|_2^2 \, dxdt$	1.524	2.036	1.720	1.597
$\langle \overline{u} \rangle - \langle \overline{u_0} \rangle$	9.901E-02	1.011E-01	3.870E-02	3.214E-02
$\left\langle \overline{u^2} \right\rangle - \left\langle \overline{u_0^2} \right\rangle$	-4.326E-01	-4.241E-01	-1.895E-01	-6.577E-02
A priori error	NA	0.987	-2.583	-2.583
$\int \|\mathbf{u} - \mathbf{u}_0\|_2^2 \, dxdt$	1.524	1.817	1.898	1.889
$\langle \overline{u} \rangle - \langle \overline{u_0} \rangle$	9.901E-02	8.012E-02	-1.242E-02	-1.543E-02
$\left\langle \overline{u^2} \right\rangle - \left\langle \overline{u_0^2} \right\rangle$	-4.326E-01	-3.466E-01	-1.018E-01	-1.296E-01

Future work

1. Implement more expressive generative models for SGS modeling and investigate their performance.
2. Deploy to practical problems: Subgrid-scale modeling in large eddy simulation of the urban environment.
3. Theoretical understanding of the difference between regression-based and generative-based SGS modeling, especially in different application scenarios such as CFD and MD.