



## Modeling subgrid-scale forces by spatial artificial neural networks in large eddy simulation of turbulence

Chenyue Xie <sup>1</sup>, Jianchun Wang <sup>1,\*</sup> and Weinan E <sup>2</sup>

<sup>1</sup>Shenzhen Key Laboratory of Complex Aerospace Flows, Center for Complex Flows and Soft Matter Research, Department of Mechanics and Aerospace Engineering, Southern University of Science and Technology, Shenzhen 518055, People's Republic of China

<sup>2</sup>Department of Mathematics, Program in Applied and Computational Mathematics, Princeton University, Princeton, New Jersey 08544, USA



(Received 20 October 2019; accepted 23 April 2020; published 18 May 2020)

Spatial artificial neural network (ANN) models are developed for subgrid-scale (SGS) forces in the large eddy simulation (LES) of turbulence. The input features are based on the first-order derivatives of the filtered velocity field at different spatial locations. The correlation coefficients of SGS forces predicted by the spatial artificial neural network (SANN) models with reasonable spatial stencil geometry can be made larger than 0.99 in an *a priori* analysis, and the relative error of SGS forces can be made smaller than 15%, much smaller than that of the traditional gradient model. In a *posteriori* analysis, a detailed comparison is made on the results of LES using the SANN model, implicit large eddy simulation (ILES), the dynamic Smagorinsky model (DSM), and the dynamic mixed model (DMM) at grid resolution of  $64^3$ . It is shown that the SANN model performs better than the ILES, DSM, and DMM models in the prediction of the spectrum and other statistical properties of the velocity field, as well as the instantaneous flow structures. These results suggest that artificial neural network with consideration of spatial characteristics is a very effective tool for developing advanced SGS models in LES of turbulence.

DOI: [10.1103/PhysRevFluids.5.054606](https://doi.org/10.1103/PhysRevFluids.5.054606)

### I. INTRODUCTION

Large eddy simulation (LES) is a technique to reduce the effective degrees of freedom of turbulence by modeling the effects of the subgrid-scale flow structures on the dynamics of the resolved scales [1–6]. LES has been widely applied to study many complex turbulent flows in aerospace industry, combustion, astrophysics, and engineering problems for several decades [7–15]. Since the pioneering works of Smagorinsky, Lilly, and Deardorff [1–3], LES models have mainly followed a model-driven approach. These include the dynamic Smagorinsky model [1, 7, 16, 17], the one-equation eddy viscosity LES model [2, 18–22], the dynamic mixed model [23–28], the optimal model [29–31], implicit-LES (ILES) [32–36], etc. [37–43]. Velocity field at the scales near the filter width  $\Delta$  has been widely used to construct the unclosed subgrid-scale (SGS) terms, including a variational multiscale Smagorinsky model [37], the locality of energy cascade in turbulence [44–47], grid-independent LES by explicitly filtering [48, 49], a new length scale for modeling subfilter motions in LES [50].

More recently, various data-driven approaches have been explored to develop more accurate Reynolds averaged Navier-Stokes (RANS) and LES models [51–76]. Neural network architectures

---

\*wangjc@sustech.edu.cn

with embedded invariance properties in RANS simulations have been developed [54,56,59]. A SGS stress model for a turbulent channel flow with artificial neural network (ANN) has been developed by Gamahara *et al.* [60]. The SGS passive scalar flux has been modeled using an ANN-based optimal estimation theory [61]. A systematic approach for developing reduced models with memory through a natural analogy between recurrent neural networks and the Mori-Zwanzig formalism has been proposed by Ma *et al.* [64]. Maulik *et al.* proposed a data-driven closure framework to reconstruct the unclosed SGS terms through localized grid-resolved information in LES [71,77]. Beck *et al.* proposed to learn the SGS force terms using conventional neural network (CNN) architectures for compressible decaying isotropic turbulence [78]. Xie *et al.* proposed an ANN model using the first- and second-order derivatives of the filtered velocity and temperature on local stencil geometry to model the SGS stress, SGS heat flux, and SGS forces of compressible isotropic turbulence [74,75]. We refer to Ref. [72] for a summary of the recent progresses on data-driven turbulence models.

One of fundamental questions for LES we try to answer is in regard to how to reconstruct the SGS terms accurately by the filtered flow fields based on the DNS data. For most traditional SGS models, the filtered flow fields at a single spatial point are used to predict the SGS terms. It was shown that the relative errors of the single-point models are not small in the *a priori* tests [74,75]. One of major reasons is that the single-point models cannot capture the spatial features of subgrid-scale dynamics. According to previous studies on the spatial locality of kinetic energy cascade in turbulence [44–47], the flow structures at the scales between  $\Delta/2$  and  $2\Delta$  give a major contribution to the SGS flux of kinetic energy at the filter scale  $\Delta$ . Thus, it is necessary to use the filtered flow fields at the scales between  $\Delta/2$  and  $2\Delta$  for accurately modeling the SGS terms at the filter width  $\Delta$ . It is worth noting that the flow dynamics at the scale  $\Delta/2$  is critical to achieve a high accuracy of the SGS model, which has rarely been considered in traditional SGS models. Here, we use ANN methods to reconstruct the SGS force from the filtered flow fields at different spatial points, in order to capture the spatial features of SGS dynamics at the scales between  $\Delta/2$  and  $2\Delta$ . We show that much higher accuracy can be achieved by the new SGS model in the *a priori* tests, as compared to the traditional SGS models.

Another fundamental issue for LES we try to address is in regard to how to implement the SGS models properly in the *a posteriori* tests. For most traditional LES simulations, the grid scale is chosen the same as the filter scale, leading to the fact that the numerical errors are comparable to the SGS modeling errors. It is worth noting that the coupled effects of numerical errors and modeling errors can cause the instability of LES. SGS models which are highly accurate in the *a priori* tests probably lead to the instability in the *a posteriori* tests, due to the effect of numerical scheme. Thus, dissipative SGS models such as Smagorinsky model are popular in LES. Here we use the grid scale smaller than the filter width, to ensure that the errors of LES are mainly from the SGS modeling. We point out that the numerical dissipation can be added individually to keep both the stability of LES and the high accuracy of SGS models.

In this paper, we propose a spatial artificial neural network (SANN) framework for reconstructing the SGS forces. We find that the SGS forces predicted by the SANN models exhibit much higher accuracy as compared to traditional gradient-based models in *a priori* analysis. We also study the accuracy of the proposed SANN model in *a posteriori* test by examining the spectrum and other statistical properties of the velocity field. These tests suggest that SANN is a very attractive approach for developing models of SGS forces.

This paper is organized as follows. Section II briefly describes the governing equations and computational method. Section III discusses DNS database of incompressible turbulence. Section IV introduces the SANN models for the reconstruction of SGS forces from the filtered flow fields. Section V presents both *a priori* and *a posteriori* results of the SANN models. Some discussions on the proposed SANN models are presented in Sec. VI. Conclusion are drawn in Sec. VII.

## II. GOVERNING EQUATIONS AND NUMERICAL METHOD

The governing Navier-Stokes equations for incompressible isotropic turbulence are [4,14,15,79]:

$$\frac{\partial u_i}{\partial x_i} = 0, \quad (1)$$

$$\frac{\partial u_i}{\partial t} + \frac{\partial u_i u_j}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \nu \frac{\partial^2 u_i}{\partial x_j \partial x_j} + \mathcal{F}_i, \quad (2)$$

where  $u_i$  is the  $i$ th velocity component,  $p$  is the pressure divided by the constant density,  $\nu$  is the kinematic viscosity, and  $\mathcal{F}$  is a large-scale forcing to the fluid momentum.

The Taylor microscale Reynolds number  $\text{Re}_\lambda$  is defined by [4,79]

$$\text{Re}_\lambda = \frac{u^{\text{rms}} \lambda}{\sqrt{3} \nu}. \quad (3)$$

Here the root-mean-square (rms) value of the velocity magnitude is defined by  $u^{\text{rms}} = \sqrt{\langle u_i u_i \rangle}$ , where  $\langle \rangle$  stands for a spatial average. The dissipation rate  $\epsilon$  is defined by  $\epsilon = 2\nu \langle S_{ij} S_{ij} \rangle$ , where  $S_{ij} = \frac{1}{2}(\partial u_i / \partial x_j + \partial u_j / \partial x_i)$  is the strain rate tensor. The Taylor microscale is defined by [79]

$$\lambda = \sqrt{\frac{5\nu}{\epsilon}} u^{\text{rms}}. \quad (4)$$

The Kolmogorov length scale  $\eta$  and the integral length scale  $L_I$  are defined, respectively, as

$$\eta = \left( \frac{\nu^3}{\epsilon} \right)^{1/4}, \quad L_I = \frac{3\pi}{2(u^{\text{rms}})^2} \int_0^\infty \frac{E(k)}{k} dk, \quad (5)$$

where  $E(k)$  is the spectrum of kinetic energy per unit mass, namely  $\int_0^\infty E(k) dk = (u^{\text{rms}})^2 / 2$ .

The physical variables can be separated into a resolved large-scale component and a subfilter small-scale component by a filtering operation  $\bar{f}(\mathbf{x}) = \int_D f(\mathbf{x}') G(\mathbf{x}, \mathbf{x}'; \Delta) d\mathbf{x}'$ , where an overbar denotes a filtered variable,  $G$  is the filter kernel,  $D$  is the overall domain, and  $\Delta$  is the filter width [5,9,80]. The incompressible filtered Navier-Stokes equations for the resolved variables can be written as [5,6,9,81]:

$$\frac{\partial \bar{u}_i}{\partial x_i} = 0, \quad (6)$$

$$\frac{\partial \bar{u}_i}{\partial t} + \frac{\partial \bar{u}_i \bar{u}_j}{\partial x_j} = -\frac{\partial \bar{p}}{\partial x_i} - \frac{\partial \tau_{ij}}{\partial x_j} + \nu \frac{\partial^2 \bar{u}_i}{\partial x_j \partial x_j} + \bar{\mathcal{F}}_i. \quad (7)$$

The SGS term appearing on the right-hand sides of Eq. (7) is defined by

$$\tau_{ij} = \overline{u_i u_j} - \bar{u}_i \bar{u}_j, \quad M_i = \frac{\partial \tau_{ij}}{\partial x_j}, \quad (8)$$

where  $\tau_{ij}$  is the SGS stress and  $M_i$  is the SGS force.

We numerically simulate the incompressible homogeneous isotropic turbulence by a pseudospectral code in a cubic box of  $(2\pi)^3$  using a uniform grid with periodic boundary conditions [79]. Time marching is performed by a second-order two-step Adams-Bashforth scheme. Full dealiasing is implemented using the two-thirds rule [82]. Force is only applied to the velocity field through fixing the total kinetic energy in the first two wave-number shells [83].

## III. DNS DATABASE OF INCOMPRESSIBLE TURBULENCE

The DNS data are obtained from direct numerical simulation of a forced incompressible isotropic turbulence at the Taylor Reynolds number  $\text{Re}_\lambda$  close to 250. We summarize the detailed numerical parameters and statistics of the turbulent velocity field at a uniform grid of  $1024^3$  in Table I [84].

TABLE I. Numerical simulation parameters and statistical quantities for  $1024^3$  grid resolution.

Reso.	$Re_\lambda$	$\eta/\Delta x$	$L_l/\eta$	$\lambda/\eta$	$\nu$	$u^{\text{rms}}$	$\omega^{\text{rms}}$	$\epsilon$
$1024^3$	259	1.01	233	31.7	0.001	2.29	26.3	0.69

The kinematic viscosity  $\nu$  equals 0.001. The resolution parameter  $\eta/\Delta x$  is close to 1.01 in numerical simulations, where  $\Delta x$  denotes the grid length in each direction. Consequently, the resolution parameter  $k_{\text{max}}\eta \approx 2.11$ . Here the largest wave number  $k_{\text{max}}$  is  $k_{\text{max}} = N/3 = 2\pi/3\Delta x$ , and  $N$  is the number of grids in each direction. The rms values of vorticity magnitude are calculated as  $\omega^{\text{rms}} = \sqrt{\langle \omega_i \omega_i \rangle}$ . A resolution of  $k_{\text{max}}\eta \geq 2.1$  is enough for the convergence of the kinetic energy spectrum at different wave numbers [79,85].

We use a top-hat filter to obtain the filtered flow variables and the corresponding SGS force  $M_i$  in this study, which is calculated in one dimension by [74,80,84,86]

$$\bar{f}_i = \frac{1}{2n} \left( f_{i-n/2} + 2 \sum_{j=i-n/2+1}^{i+n/2-1} f_j + f_{i+n/2} \right), \quad (9)$$

where the filter width is  $\Delta = n\Delta x$ . Filter width  $\Delta = 32\Delta x$  is used in the present study as shown in Fig. 1, which corresponds to the filter size of  $\Delta/\eta \approx 32$  and covers the inertial range; 10% of the turbulent kinetic energy resides in the SGS flow field for  $\Delta/\Delta x = 32$ . LESs are performed on  $64^3$  grid resolution with filter width  $\Delta = 32\Delta x$  in a *posteriori* test. Here the grid size of LES is  $h_{\text{LES}}^\Delta = \Delta/2$ .

#### IV. THE STRUCTURE OF THE SANN MODEL

We consider a fully connected ANN as shown in Fig. 2. Different layers are connected by neurons, where a layer  $l$  receives and operates inputs  $X_j^{l-1}$  from layer  $l-1$  and then establishes

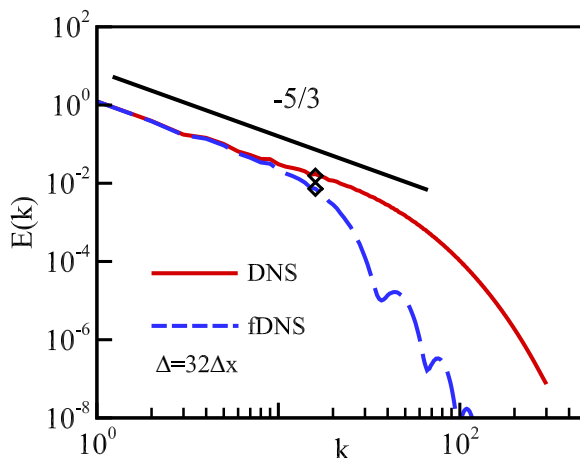


FIG. 1. Velocity spectrum from direct numerical simulation of a forced incompressible isotropic turbulence. Diamond represents filter width  $\Delta/\Delta x = 32$ .

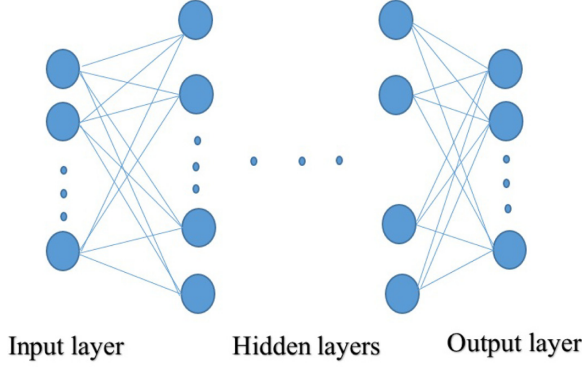


FIG. 2. Schematic diagram of the ANN's network structure.

outputs  $X_i^l$  by the activation function [62,87,88]. The transformation from layer  $l - 1$  to layer  $l$  is calculated as follows:

$$X_i^l = \sigma(s_i^l + b_i^l), \quad (10)$$

$$s_i^l = \sum_j W_{ij}^l X_j^{l-1}, \quad (11)$$

where  $\sigma$  is the activation function,  $W_{ij}^l$  is the weight, and  $b_i^l$  is the bias parameter.  $W_{ij}^l$  are initialized using glorot-uniform, and  $b_i^l$  are initialized to be zero. The ANN consists of four layers (input layer, two hidden layers, and output layer). The neuron numbers of four layers between the set of inputs and final output layer are  $M : 1024 : 512 : 1$ , respectively.  $M$  is the number of neurons of the input layer, while the output layer consists of a single neuron associated with the component of the SGS force  $M_i$ . Meanwhile, two hidden layers are activated by a Leaky-Relu activation function:

$$\sigma(a) = \begin{cases} a, & \text{if } a > 0 \\ 0.2a, & \text{if } a \leq 0 \end{cases}. \quad (12)$$

Linear activation  $\sigma(a) = a$  is applied to the output layer. The reasonable  $W_{ij}^l$  and  $b_i^l$  are determined by the training stage of ANN. The loss function is defined as  $\langle (X^L - M_i)^2 \rangle$ , where  $\langle \rangle$  represents the spatial average over homogeneous directions (i.e., the entire domain for isotropic turbulence) [60]. It measures the difference between the output and the real SGS force  $M_i$ . We use back-propagation to minimize the loss function using the Adam optimizer with learning rate of 0.001 [89].

We show the hyperparameter configuration of the ANN model in Table II. The ANN is trained with four layers of neurons  $M : 1024 : 512 : 1$ , the hidden layers are activated by the Leaky-Relu function, and the output layer is activated by the linear function. The network is trained by the Adam algorithm [89] for 1000 iterations, with the batch size being 1000 for determining the final network until the learning rate is minimal.

TABLE II. Hyperparameter configuration of the ANN model.

Layer structure	Activation (hidden layer)	Activation (output layer)	Loss fuction	Optimizer	Learning rate	Epoch
M:1024:512:1	Leaky-relu	Linear	MSE	Adam	0.001	1000

TABLE III. The number of input variables  $M$  and GPU time for 1000 epoches of ANN training associated with different  $R_s$  and  $R_g$  for the SANN model.

$(R_s, R_g)$	(0,1)	(2,1)	(2,2)	(2,4)	(2,8)	(1,4)	(3,4)	(4,4)
M	9	117	225	441	873	225	657	873
GPU(h)	8.1	11.2	13.4	17.1	27	12.9	21.4	29

We train three ANNs to reconstruct each independent component of the SGS force  $M_i$ . The final output  $X^L$  can be obtained from the input layer  $X_j^1$  of filtered flow fields based on a series of iterations of the bias parameters and the weights. The SGS force  $M_i$  is the derivative of  $\tau_{ij}$  and can be approximated as [81,90,91]:

$$M_i = \frac{\partial \tau_{ij}}{\partial x_j} = \frac{\Delta^2}{12} \left( \frac{\partial^2 \bar{u}_i}{\partial x_j \partial x_k} \frac{\partial \bar{u}_j}{\partial x_k} + \frac{\partial \bar{u}_i}{\partial x_k} \frac{\partial^2 \bar{u}_j}{\partial x_j \partial x_k} \right) + O(\Delta^4). \quad (13)$$

Since higher-order derivatives in Eq. (13) can be approximated by divided differences of the first-order derivative, the first-order derivative of the velocity is used as the input features for the ANN. The velocity derivatives at different spatial points are combined to model the effect of the spatial structure of the flow field. An ANN is then trained to establish a functional relation between input variables and the SGS forces. Since the flow dynamics at the scales near the filter width  $\Delta$  is very important for effective and successful ANN learning due to the multiscale characteristics of the SGS formulation and the locality of energy cascade in turbulence [37,44–47,92–94], different spatial stencils of input features are designed in our ANN architecture in order to improve the accuracy of LES models.

Considering that the set of input features of the SANN model depends on the input spatial stencil width  $\Delta_s$  and the grid length  $\Delta_g$  in each dimension, we use the ratio of the input spatial stencil width  $\Delta_s$  to the filter width  $\Delta$  ( $R_s = \Delta_s/\Delta$ ), and the ratio of the filter width  $\Delta$  to the grid length  $\Delta_g$  ( $R_g = \Delta/\Delta_g$ ) to represent the structure of the input spatial stencil. We consider the input variables of ANN for six spatial directions from the spatial location  $(l, m, n)$  as follows:

$$\left\{ \frac{\partial \bar{u}_i^{l,m,n}}{\partial x_p}, \frac{\partial \bar{u}_i^{l \pm q_1, m, n}}{\partial x_p}, \frac{\partial \bar{u}_i^{l, m \pm q_2, n}}{\partial x_p}, \frac{\partial \bar{u}_i^{l, m, n \pm q_3}}{\partial x_p} \right\},$$

$$\{q_1, q_2, q_3\} \in \{1, 2, \dots, R_s R_g\}, \quad (14)$$

where the indices  $l, m$ , and  $n$  correspond to discrete spatial locations associated with the grid length  $\Delta_g$ ,  $p = 1, 2, 3$  is the subscript of the Cartesian coordinate, and  $i = 1, 2, 3$  is the subscript of the filtered velocity component. The total number of spatial points is  $N = 6R_s R_g + 1$ . Each sample of our training data thus consists of  $M = 9N = 9(6R_s R_g + 1)$  inputs of filtered variables. The SANN model associated with the parameters  $R_s$  and  $R_g$  is abbreviated as SANN( $R_s, R_g$ ). The number of input variables associated with different  $R_s$  and  $R_g$  are shown in Table III for the SANN model. For example, the input set with  $R_s = 2, R_g = 2$  is abbreviated as SANN(2,2), which has 225 input variables:  $M = 225$ .

We consider the input variables for 26 spatial directions from the location  $(l, m, n)$  as follows:

$$\frac{\partial \bar{u}_i^{l+iq_h, m+jq_h, n+kq_h}}{\partial x_p}, \quad q_h = 1, 2, \dots, R_s R_g,$$

$$\{i, j, k\} \in \{-1, 0, 1\}, \quad i^2 + j^2 + k^2 \neq 0. \quad (15)$$

The SANN model associated with the parameters  $R_s$  and  $R_g$  for 26 directions is abbreviated as SANN-D26( $R_s, R_g$ ). The number of input variables associated with different  $R_s$  and  $R_g$  are shown

TABLE IV. The number of input variables  $M$  and GPU time for 1000 epoches of ANN training associated with different  $R_s$  and  $R_g$  for the SANN-D26 model.

$(R_s, R_g)$	(0,1)	(2,1)	(2,2)	(2,4)	(2,8)	(1,4)	(3,4)	(4,4)
M	9	477	945	1881	3753	945	2817	3753
GPU(h)	8.1	17.6	32.2	57.8	184	29.0	87.4	171.7

in Table IV for the SANN-D26 model. For example, the input set with  $R_s = 2, R_g = 2$  can be abbreviated as SANN-D26(2,2), which has 945 input variables:  $M = 945$ .

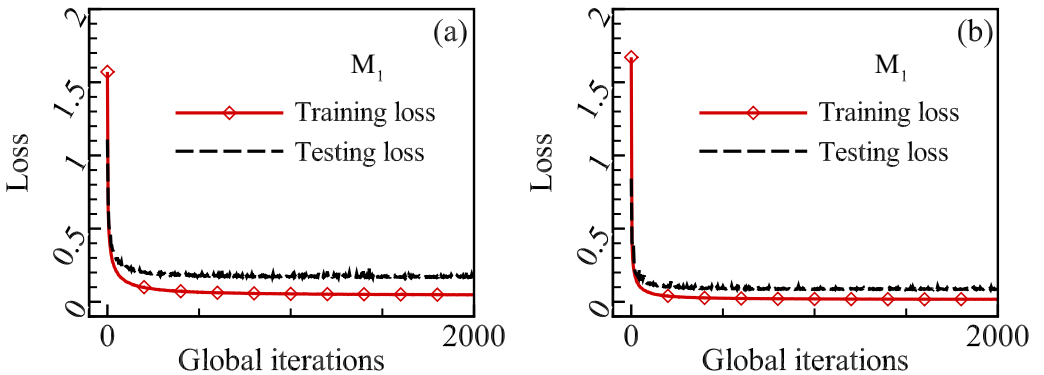
In order to increase the robustness of the ANN training, all input variables  $X_I$  and output variables  $X_O$  are normalized by their rms values  $X_I^{\text{rms}}$  and  $X_O^{\text{rms}}$ , respectively, which is similar to the previous data-driven strategies [52,56,61,62].

$$Z_I = X_I / X_I^{\text{rms}}, \quad (16)$$

$$Z_O = X_O / X_O^{\text{rms}}. \quad (17)$$

Since the real SGS forces are unknown in LES, the rms values of SGS forces reconstructed from the gradient model are used in normalizations of the output variables for the ANN.

We use cross validation to prevent overfitting of the SANN model, using data that have not been trained for. In this research, the LES is performed at grid resolution of  $64^3$  ( $h_{\text{LES}}^\Delta = \Delta/2$ ) with filter width  $\Delta = 32\Delta x$ . Considering that a coarse-graining procedure involves the selection of every 16th grid point in each direction for the DNS data with the grid resolution of  $1024^3$ ,  $16^3$  different data sets of  $64^3$  coarse grid points can be obtained. The total data for ANN training are  $16 \times 64^3$  grid points by randomly choosing from the  $16^3$  different data sets; 70% of data is for training, 30% for testing. Finally, the Adam optimizer with learning rate of 0.001 is used to update  $W_{ij}^l$  and  $b_i^l$  of the ANN. The learning curve of the training and testing loss for the proposed SANN model is shown in Fig. 3, which is trained for a long duration (2000 epochs) with batch size being 1000. The training and testing loss show similar behavior and correlate closely after 500 global iterations, which implies that the learning process is reasonable. The final ANN is selected for 1000 iterations to reconstruct SGS forces accurately. Furthermore, we use four GPU cores (NVIDIA Tesla K80 GPU) for parallel training of neural networks in this research. The GPU time for 1000 epoches of ANN training associated with different  $R_s$  and  $R_g$  for the SANN and SANN-D26 models are listed in Tables III and IV. As the number of input variables  $M$  increases, GPU time increases.


 FIG. 3. Learning curves of the proposed SANN model of unclosed SGS force  $M_1$ : (a) SANN(2,2) and (b) SANN-D26(2,2).



## V. TEST RESULTS OF THE SANN MODEL

In this section, we test the ability of our proposed SANN models to produce reliable, repeatable predictions by learning from the training data for the forced isotropic turbulence at  $Re_\lambda \approx 250$ . Both *a priori* and *a posteriori* tests are conducted to study the performance of the SANN models. In the *a priori* test, the correlation coefficients and relative errors of the predicted SGS forces are compared for  $\Delta/\Delta x = 32$  in the inertial range. The multiscale and local characteristics of the SGS models  $SANN(R_s, R_g)$  are analyzed by changing two parameters  $(R_s, R_g)$ . In the *a posteriori* test, results of the large eddy simulations with the SGS model  $SANN(2,2)$  at grid resolution of  $64^3$  ( $h_{LES} = \Delta/2$ ,  $\Delta = 32\Delta x$ ) are compared with the ILES, DSM, and DMM models and filtered DNS database. It is shown that the proposed SANN model can predict the spectra and statistics of the filtered DNS data with high accuracy in LES.

### A. A priori tests

We evaluate the performance of the SGS models  $SANN(R_s, R_g)$  by calculating the correlation coefficient  $C(R)$ , the relative error  $E_r(R)$ , and the root-mean-square value  $D(R)$  of SGS forces.  $C(R)$ ,  $E_r(R)$ , and  $D(R)$  are defined, respectively, by

$$C(R) = \frac{\langle (R - \langle R \rangle)(R^{\text{model}} - \langle R^{\text{model}} \rangle) \rangle}{(\langle (R - \langle R \rangle)^2 \rangle \langle (R^{\text{model}} - \langle R^{\text{model}} \rangle)^2 \rangle)^{1/2}}, \quad (18)$$

$$E_r(R) = \frac{\sqrt{\langle (R - R^{\text{model}})^2 \rangle}}{\sqrt{\langle R^2 \rangle}}, \quad (19)$$

$$D(R) = \sqrt{\langle (R - \langle R \rangle)^2 \rangle}, \quad (20)$$

where  $\langle \cdot \rangle$  denotes averaging over the volume. We consider two types of gradient models to compare with the SGS models  $SANN(R_s, R_g)$  [25,27]:

$$(1) : M_i = \frac{\Delta^2}{12} \alpha_i, \quad \alpha_i = \frac{\partial}{\partial x_j} \left( \frac{\partial \bar{u}_i}{\partial x_k} \frac{\partial \bar{u}_j}{\partial x_k} \right), \quad (21)$$

$$(2) : M_i = C_i \alpha_i, \quad C_i = \frac{\langle M_i \alpha_i \rangle}{\langle \alpha_i \alpha_i \rangle}. \quad (22)$$

The first model is the gradient model, which is abbreviated as VG1. The second model is the gradient model with coefficient  $C_i$  determined from DNS data by least square method, which is abbreviated as VG1m.

In this research, we consider a fully connected ANN as shown in Fig. 2. There is no additional constraint to the ANN architecture. The back propagation method is used to minimize the loss function, which is defined as the difference between the output and the real SGS force [ $\langle (X^L - M_i)^2 \rangle$ ], where  $\langle \cdot \rangle$  represents the spatial average over homogeneous directions (i.e., the entire domain for isotropic turbulence) [60]. The network is trained by the Adam algorithm [89] for 1000 iterations, with batch size being 1000 for determining the final network until the learning rate is minimal. All the weights corresponding to the added inputs are activated. The weights in different hidden layers are different from the previous levels. More details about the characteristics of the weights of different hidden layers will be further studied in the next step.

In an *a priori* analysis, the total dataset for different  $SANN(R_s, R_g)$  and  $SANN\text{-D26}(R_s, R_g)$  models are  $16 \times 64^3$  grid points; 70% of dataset is for training, 30% for testing. The  $SANN(R_s, R_g)$  and  $SANN\text{-D26}(R_s, R_g)$  models are trained by the Adam optimizer with learning rate of 0.001 for 1000 epochs with batch size being 1000. The correlation coefficient ( $C$ ), relative error ( $E_r$ ), and root-mean-square value ( $D$ ) of  $M_1$  for the SANN and SANN-D26 models with different  $R_s$  and  $R_g$  of the training and testing datasets are shown in Tables V–X.

In order to analyze the impact of the network architecture of neurons, the ANN is trained with four layers of neurons  $M : 1536 : 768 : 1$  and five layers of neurons  $M : 1024 : 512 :: 256 : 1$ , which



TABLE V. Correlation coefficient ( $C$ ), relative error ( $E_r$ ), and root-mean-square value ( $D$ ) of  $M_1$  for different SANN models at the filter width  $\Delta = 32\Delta x$  with  $R_s = 2$  and  $R_g = 2$ .

$C$	VG1	VG1m	SANN(2,2)	SANN(2,2)-D	SANN(2,2)-L	SANN(2,2)-R	SANN(2,2)-SP	
Train	0.738	0.738	0.989	0.994	0.991	0.990	0.980	
Test	0.738	0.738	0.978	0.986	0.980	0.979	0.970	
$E_r$	VG1	VG1m	SANN(2,2)	SANN(2,2)-D	SANN(2,2)-L	SANN(2,2)-R	SANN(2,2)-SP	
Train	0.713	0.674	0.163	0.112	0.140	0.144	0.200	
Test	0.713	0.674	0.220	0.171	0.203	0.206	0.245	
$D$	DNS	VG1	VG1m	SANN(2,2)	SANN(2,2)-D	SANN(2,2)-L	SANN(2,2)-R	SANN(2,2)-SP
Train	1.41	0.718	1.04	1.40	1.40	1.40	1.41	1.40
Test	1.41	0.718	1.04	1.39	1.39	1.39	1.40	1.40

are abbreviated as SANN-D and SANN-L (SANN-D26-D and SANN-D26-L), respectively. The *a priori* results of the SANN, SANN-D and SANN-L (SANN-D26, SANN-D26-D and SANN-D26-L) models are summarized in Tables V and VI. The SANN-D, SANN-L and SANN models show similar correlation coefficients and relative errors, indicating that the learning accuracy does not increase with the increase of the number of neurons and hidden layers, and the nearly convergent results have been achieved with the enough numbers of neurons.

In order to illustrate the rationality of the SANN model with Leaky-ReLU activation function used in this research, we consider several different SANN models with hidden layers, which are activated by two other activation functions without normalization: (i) rectified linear (ReLU) function

 TABLE VI. Correlation coefficient ( $C$ ), relative error ( $E_r$ ), and root-mean-square value ( $D$ ) of  $M_1$  for different SANN-D26 models at the filter width  $\Delta = 32\Delta x$  with  $R_s = 2$  and  $R_g = 2$ .

$C$	VG1	VG1m	SANN-D26(2,2)	
Train	0.738	0.738	0.995	
Test	0.738	0.738	0.989	
SANN-D26(2,2)-D	SANN-D26(2,2)-L	SANN-D26(2,2)-R	SANN-D26(2,2)-SP	
0.996	0.995	0.995	0.986	
0.991	0.989	0.989	0.981	
$E_r$	VG1	VG1m	SANN-D26(2,2)	
Train	0.713	0.674	0.102	
Test	0.713	0.674	0.151	
SANN-D26(2,2)-D	SANN-D26(2,2)-L	SANN-D26(2,2)-R	SANN-D26(2,2)-SP	
0.887	0.096	0.126	0.171	
0.132	0.146	0.169	0.200	
$D$	DNS	VG1	VG1m	SANN-D26(2,2)
Train	1.41	0.718	1.04	1.41
Test	1.41	0.718	1.04	1.40
SANN-D26(2,2)-D	SANN-D26(2,2)-L	SANN-D26(2,2)-R	SANN-D26(2,2)-SP	
1.41	1.40	1.39	1.40	
1.39	1.39	1.39	1.39	

TABLE VII. Correlation coefficient ( $C$ ), relative error ( $E_r$ ), and root-mean-square value ( $D$ ) of  $M_1$  for different models (VG1, VG1m, and SANN models) at the filter width  $\Delta = 32\Delta x$  with same  $R_s$  ( $R_s = 2$ ).

$C$	VG1	VG1m	SANN(0,1)	SANN(2,1)	SANN(2,2)	SANN(2,4)	SANN(2,8)	
Train	0.738	0.738	0.544	0.981	0.989	0.992	0.992	
Test	0.738	0.738	0.517	0.964	0.978	0.985	0.985	
$E_r$	VG1	VG1m	SANN(0,1)	SANN(2,1)	SANN(2,2)	SANN(2,4)	SANN(2,8)	
Train	0.713	0.674	0.839	0.195	0.163	0.124	0.125	
Test	0.713	0.674	0.859	0.266	0.220	0.174	0.173	
$D$	DNS	VG1	VG1m	SANN(0,1)	SANN(2,1)	SANN(2,2)	SANN(2,4)	SANN(2,8)
Train	1.41	0.718	1.04	0.80	1.39	1.40	1.41	1.41
Test	1.41	0.718	1.04	0.78	1.38	1.39	1.40	1.40

$\sigma(a) = \begin{cases} a, & \text{if } a > 0 \\ 0, & \text{if } a \leq 0 \end{cases}$  (SANN-R) and (ii) softplus activation function  $\sigma(a) = \log(1 + e^a)$  (SANN-SP). The SANN model with the two different activation functions are abbreviated as SANN-R and SANN-SP, respectively. The *a priori* tests of the SANN-R, SANN-SP models are summarized in Tables V and VI. All the models with different activation functions show similar results, suggesting that the SANN models are not very sensitive to the choice of the activation function.

Tables V–X show correlation coefficients, relative errors and root-mean-square values of  $M_1$  for different models in both training and testing sets after the training process of ANN with filter width  $\Delta/\Delta x = 32$ . The small difference of the results of training and testing sets show that: the ANN training process is not overfitting. Each component of the SGS force  $M_i$  is trained separately without physical constraint.

Tables VII and VIII show the  $C$ ,  $E_r$ ,  $D$ , of  $M_1$  for the SGS models SANN( $R_s, R_g$ ). The relative errors of the VG1 and VG1m models are more than 60%. The correlation coefficient is close to 55%, and the relative error is close to 0.80 for the local ANN model with  $R_s R_g = 0$ , which means that single point input features cannot accurately reconstruct the  $M_i$ . The correlation coefficients increase and the relative errors decrease for the SANN( $2, R_g$ ) models with the increase of  $R_g$  as shown in Table VII. Moreover, the performance of the SGS models SANN( $R_s, R_g$ ) with the same  $R_g = 4$  for different  $R_s$  are compared in Table VIII. When the stencil width increases, the SANN models perform much better than the VG1 and VG1m models. The correlation coefficients can be made larger than 0.99 and the relative errors less than 0.15.

 TABLE VIII. Correlation coefficient ( $C$ ), relative error ( $E_r$ ), and root-mean-square value ( $D$ ) of  $M_1$  for different models (VG1, VG1m, and SANN models) at the filter width  $\Delta = 32\Delta x$  with same  $R_g$  ( $R_g = 4$ ).

$C$	VG1	VG1m	SANN(0,1)	SANN(1,4)	SANN(2,4)	SANN(3,4)	SANN(4,4)	
Train	0.738	0.738	0.544	0.988	0.992	0.994	0.995	
Test	0.738	0.738	0.517	0.977	0.985	0.988	0.989	
$E_r$	VG1	VG1m	SANN(0,1)	SANN(1,4)	SANN(2,4)	SANN(3,4)	SANN(4,4)	
Train	0.713	0.674	0.839	0.155	0.124	0.109	0.109	
Test	0.713	0.674	0.859	0.214	0.174	0.156	0.152	
$D$	DNS	VG1	VG1m	SANN(0,1)	SANN(1,4)	SANN(2,4)	SANN(3,4)	SANN(4,4)
Train	1.41	0.718	1.04	0.80	1.40	1.41	1.41	1.41
Test	1.41	0.718	1.04	0.78	1.39	1.40	1.40	1.40

The results for  $C$ ,  $E_r$ , and  $D$  of  $M_1$  for the SGS models SANN-D26( $R_s, R_g$ ) with 26 directions are shown in Tables IX and X. With the same  $R_s, R_g$ , the SANN-D26 models with 26 directions show significant advantage over the SANN models with 6 directions. The relative errors of the SGS models SANN-D26( $R_s, R_g$ ) can be made less than 12% with reasonable  $R_s$  and  $R_g$ , which are much smaller than those of VG1 and VG1m models. The root-mean-square value of SGS force  $M_1$  predicted by VG1 reaches 70% of that of DNS, while the SANN-D26 models predict rms value of SGS force  $M_1$  very close to that of DNS data in the *a priori* test. The spatial stencil of the filtered velocity gradient with reasonable  $R_s, R_g$  can provide a good approximation of the SGS forces. By considering more information near the filter width, the ANN model can build a better relation between filtered flow variables and the unclosed SGS terms.

### B. *A posteriori* tests

We evaluate the performance of the SGS model SANN(2,2) with  $R_s = 2$  and  $R_g = 2$  at grid resolution of  $64^3$  ( $h_{\text{LES}} = \Delta/2$ ,  $\Delta = 32\Delta x$ ) [92–94]. The filter width  $\Delta$  is usually proportional to the grid size  $h_{\text{LES}}^\Delta$ : The filter grid ratio (FGR =  $\Delta/h_{\text{LES}}^\Delta$ ) is typically chosen to be equal to 1 or 2 [95,96]. The value FGR = 2 used in this work is a reasonable value to ensure that the differences between LES results and filtered DNS results are mainly due to the SGS models [48–50,92–94]. Furthermore, the *a posteriori* tests of the SANN(4,2) model with more input features are shown in Appendix A. In order to show that the SANN model can be applied to other conditions, the *a posteriori* studies of the SANN models applied to homogenous shear turbulence are shown in Appendix B.

In *a posteriori* tests, we implement the SGS model of SANN(2,2) with the weight  $W_{ij}^l$  and bias parameter  $b_i^l$  in the ANN of neurons 225 : 1024 : 512 : 1. A physical constraint is added to the output of the SANN(2,2) model by  $M_i^{\text{updated}} = M_i - \langle M_i \rangle$  since the average of  $M_i$  in the whole domain equals zero in the conservation equations. This is reasonable since the ratio  $|\langle M_i \rangle / M_i^{\text{rms}}|$  is less than 5%, where  $\langle M_i \rangle$  is the spatial average of  $M_i$ , and  $M_i^{\text{rms}}$  is the rms value of  $M_i$ .

Simulation of ILES based on a numerical viscosity without explicit modeling has been carried out as comparison [32–36]. We introduce a dissipative numerical method with higher numerical dissipation for larger wave numbers:

$$\bar{\mathbf{u}} = \mathbf{u} - \Delta t \cdot C_0^l \left( \frac{\mathbf{k}}{k_c} \right)^n \mathbf{u}, \quad k_c = \frac{4\pi}{3\Delta}, \quad (23)$$

where  $n = 4$  and  $C_0^l = 3$ . Further more, simulations using the DSM, DMM, and SANN(2,2) models based on ILES are also carried out. The ratio of the computational cost of different SGS models with the same grid resolution is found to be: No-model : DSM : DMM : SANN(2, 2) = 0.4 : 1 : 1.5 : 256. The SANN(2,2) model requires most computational cost because of matrix operations in neural network. The ratio of the time steps in LES and DNS is  $\Delta t_{\text{LES}} / \Delta t_{\text{DNS}} = 8$ , where  $\Delta t_{\text{LES}}$  and  $\Delta t_{\text{DNS}}$  are the time steps for LES and DNS, respectively. We mainly focus on the spatial closure since the time discretization error of a practical LES is small and can be neglected.

Based on the equilibrium hypothesis, the DSM model reproduces the global exchange of energy between the resolved and unresolved scales by mimicking the drain of energy associated with the turbulence energy cascade. In the DSM model, the SGS stress can be written as follows [1,5,7,9,17]:

$$\tau_{ij} - \frac{\delta_{ij}}{3} \tau_{kk} = -2C_s^2 \Delta^2 |\bar{S}| \left( \bar{S}_{ij} - \frac{\delta_{ij}}{3} \bar{S}_{kk} \right), \quad (24)$$

where  $\Delta$  is the characteristic SGS length scale, which is equal to the filter width.  $C_s^2$  can be calculated dynamically as follows:

$$C_s^2 = \frac{\langle \mathcal{L}_{ij} \mathcal{M}_{ij} \rangle}{\langle \mathcal{M}_{kl} \mathcal{M}_{kl} \rangle}, \quad (25)$$

TABLE IX. Correlation coefficient ( $C$ ), relative error ( $E_r$ ), and root-mean-square value ( $D$ ) of  $M_1$  for different models (VG1, VG1m, and SANN-D26 models) at the filter width  $\Delta = 32\Delta x$  with same  $R_s$  ( $R_s = 2$ ).

$C$	VG1	VG1m	SANN-D26(0,1)	SANN-D26(2,1)	SANN-D26(2,2)	SANN-D26(2,4)	SANN-D26(2,8)	
Train	0.738	0.738	0.544	0.992	0.995	0.996	0.996	
Test	0.738	0.738	0.517	0.982	0.989	0.992	0.992	
$E_r$	VG1	VG1m	SANN-D26(0,1)	SANN-D26(2,1)	SANN-D26(2,2)	SANN-D26(2,4)	SANN-D26(2,8)	
Train	0.713	0.674	0.839	0.130	0.102	0.089	0.094	
Test	0.713	0.674	0.859	0.191	0.151	0.130	0.132	
$D$	DNS	VG1	VG1m	SANN-D26(0,1)	SANN-D26(2,1)	SANN-D26(2,2)	SANN-D26(2,4)	SANN-D26(2,8)
Train	1.41	0.718	1.04	0.80	1.41	1.41	1.41	1.40
Test	1.41	0.718	1.04	0.78	1.40	1.40	1.40	1.39

TABLE X. Correlation coefficient ( $C$ ), relative error ( $E_r$ ), and root-mean-square value ( $D$ ) of  $M_1$  for different models (VG1, VG1m, and SANN-D26 models) at the filter width  $\Delta = 32\Delta x$  with same  $R_g$  ( $R_g = 4$ ).

$C$	VG1	VG1m	SANN-D26(0,1)	SANN-D26(1,4)	SANN-D26(2,4)	SANN-D26(3,4)	SANN-D26(4,4)	
Train	0.738	0.738	0.544	0.994	0.996	0.997	0.997	
Test	0.738	0.738	0.517	0.987	0.992	0.993	0.994	
$E_r$	VG1	VG1m	SANN-D26(0,1)	SANN-D26(1,4)	SANN-D26(2,4)	SANN-D26(3,4)	SANN-D26(4,4)	
Train	0.713	0.674	0.839	0.112	0.089	0.081	0.079	
Test	0.713	0.674	0.859	0.163	0.130	0.118	0.114	
$D$	DNS	VG1	VG1m	SANN-D26(0,1)	SANN-D26(1,4)	SANN-D26(2,4)	SANN-D26(3,4)	SANN-D26(4,4)
Train	1.41	0.718	1.04	0.80	1.41	1.41	1.41	1.40
Test	1.41	0.718	1.04	0.78	1.40	1.40	1.40	1.39

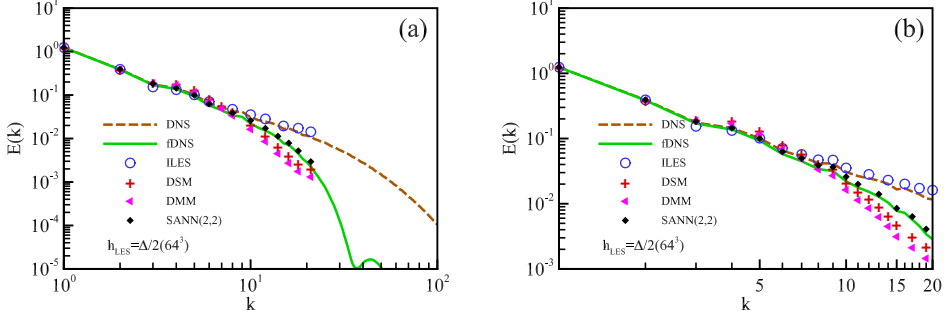


FIG. 4. Spectrum of velocity for LES at grid resolutions of  $64^3$  ( $h_{\text{LES}} = \Delta/2$ ) with the filter width  $\Delta = 32\Delta x$ .

where  $\mathcal{L}_{ij} = \widehat{u_i u_j} - \hat{u}_i \hat{u}_j$ ,  $\alpha_{ij} = -2\Delta^2 |\bar{S}| (\bar{S}_{ij} - \frac{\delta_{ij}}{3} \bar{S}_{kk})$ ,  $\beta_{ij} = -2\hat{\Delta}^2 |\hat{S}| (\hat{S}_{ij} - \delta_{ij} \hat{S}_{kk}/3)$ , and  $\mathcal{M}_{ij} = \beta_{ij} - \hat{\alpha}_{ij}$ . An overbar denotes the filter at grid scale  $\Delta x$ , a hat represents a test filter coarser than the grid filter, and a hat over the overbar denotes a filter at scale  $\hat{\Delta} = 2\Delta$ .

DMM model has been developed based on a combination of a scale-similarity part and an eddy-viscosity part [23,25], which not only has the high correlation coefficient of scale similarity part, but also guarantees sufficient dissipation with eddy-viscosity part. The coefficients in the DMM model can be determined by a dynamic procedure based on the Germano identity [8],

$$L_{ij} = T_{ij} - \widehat{\rho \tau_{ij}}, \quad (26)$$

where an overbar denotes the filter at the scale  $\Delta$ , a hat represents a test filter coarser than the grid filter, and a hat over the overbar denotes a filter at scale  $\hat{\Delta} = 2\Delta$ ,  $T_{ij} = \widehat{u_i u_j} - \hat{u}_i \hat{u}_j$  is the stress at scale  $\hat{\Delta}$ , and  $L_{ij} = \widehat{u_i u_j} - \hat{u}_i \hat{u}_j$  is the resolved stress tensor. We can obtain a square error by replacing  $\tau_{ij}$  and  $T_{ij}$  with the model stresses  $\tau_{ij}^{\text{mod}}$  and  $T_{ij}^{\text{mod}}$  for  $L_{ij}$ :

$$E_{\text{mod}} = \langle (L_{ij} - L_{ij}^{\text{mod}})^2 \rangle, \quad (27)$$

where  $L_{ij}^{\text{mod}} = T_{ij}^{\text{mod}} - \widehat{\tau_{ij}^{\text{mod}}}$ . By assuming that the coefficients are scale invariant and minimizing the right-hand side of (27), we get a set of equations for the coefficients, including the SGS terms at scale  $\Delta$  and  $\hat{\Delta}$  [27,28]

$$\tau_{ij}^{\text{mod}} = C_1 h_{1,ij} + C_2 h_{2,ij}, \quad (28)$$

$$T_{ij}^{\text{mod}} = C_1 H_{1,ij} + C_2 H_{2,ij}, \quad (29)$$

where  $h_{1,ij} = -2\Delta^2 |\bar{S}_{ij}| (\bar{S}_{ij} - \frac{1}{3} \delta_{ij} \bar{S}_{kk})$ ,  $h_{2,ij} = \widehat{u_i u_j} - \hat{u}_i \hat{u}_j$ ,  $H_{1,ij} = -2\hat{\Delta}^2 |\hat{S}_{ij}| (\hat{S}_{ij} - \frac{1}{3} \delta_{ij} \hat{S}_{kk})$ , and  $H_{2,ij} = \overrightarrow{(\hat{u}_i \hat{u}_j - \vec{\hat{u}}_i \vec{\hat{u}}_j)}$ , where an overrightarrow represents filtering at scale  $4\Delta$ . These give

$$C_1 = \frac{\langle N_{ij}^2 \rangle \langle L_{ij} M_{ij} \rangle - \langle M_{ij} N_{ij} \rangle \langle L_{ij} N_{ij} \rangle}{\langle N_{ij}^2 \rangle \langle M_{ij}^2 \rangle - \langle M_{ij} N_{ij} \rangle^2}, \quad (30)$$

$$C_2 = \frac{\langle M_{ij}^2 \rangle \langle L_{ij} N_{ij} \rangle - \langle M_{ij} N_{ij} \rangle \langle L_{ij} M_{ij} \rangle}{\langle N_{ij}^2 \rangle \langle M_{ij}^2 \rangle - \langle M_{ij} N_{ij} \rangle^2}, \quad (31)$$

where  $L_{ij} = \widehat{u_i u_j} - \hat{u}_i \hat{u}_j$ ,  $M_{ij} = H_{1,ij} - \widehat{h_{1,ij}}$ , and  $N_{ij} = H_{2,ij} - \widehat{h_{2,ij}}$ .

The performances of the SGS model SANN(2,2) are evaluated by calculating the spectrum and statistical properties of the velocity. The spectrum of the velocity field is displayed in Fig. 4 for the DNS, filtered DNS (fdNS) data and LES computations with the ILES, DSM, DMM, and

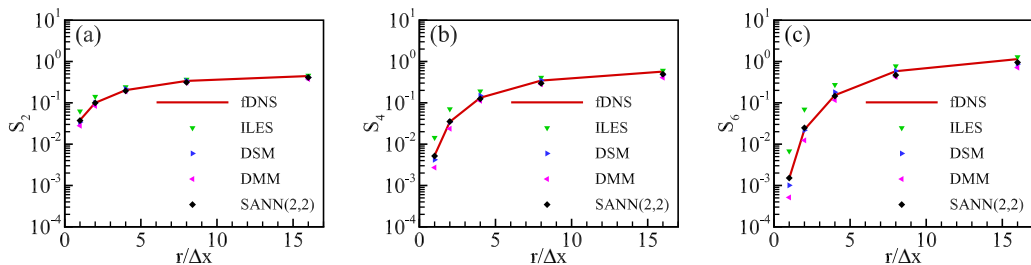


FIG. 5. Structure functions of the velocity for LES at grid resolutions of  $64^3$  ( $h_{\text{LES}}^{\Delta} = \Delta/2$ ) with the filter width  $\Delta = 32\Delta x$ : (a)  $S_2$ , (b)  $S_4$ , (c)  $S_6$ .

SANN(2,2) models. There is a long inertial region with a  $k^{-5/3}$  scaling of velocity spectrum for DNS flow field. LES results deviate slightly from the results of fDNS. In particular, the errors of the spectrum predicted by LES models increase as  $k$  increases since it is hard to model the dynamics at scales near the largest resolved wave number. The spectrum of the velocity field predicted by the “ILES” is obviously higher than that of the filtered DNS data at large wave numbers. The spectra of velocity predicted by the DSM and DMM models show a bump, where the spectra at low wave numbers with  $k \leq 10$  are too energy-rich, while near the cut-off wave number the spectra are damped too strongly. In contrast, the velocity spectrum predicted by the SGS model SANN(2,2) nearly overlaps with that of the filtered DNS data.

We also evaluate the performance of the LES models by comparing many statistical properties [14,15,97,98]. The longitudinal structure functions of velocity is defined by

$$S_n^L(r) \equiv \left\langle \left| \frac{\delta_r \bar{u}}{\bar{u}^{\text{rms}}} \right|^n \right\rangle, \quad (32)$$

where  $\delta_r \bar{u} = [\bar{\mathbf{u}}(\mathbf{x} + \mathbf{r}) - \bar{\mathbf{u}}(\mathbf{x})] \cdot \hat{\mathbf{r}}$ , is the longitudinal increment of the velocity for the separation  $\mathbf{r}$ . Here  $\hat{\mathbf{r}} = \mathbf{r}/|\mathbf{r}|$ . The rms velocity  $\bar{u}^{\text{rms}}$  has been used to normalize the longitudinal structure functions.

We compare the normalized structure functions of the velocity in Fig. 5. All models [“ILES,” DSM, DMM, and SANN(2,2) models] accurately predict the structure functions at large separations. At small separations, traditional LES models including ILES, DSM, and DMM do a poor job. In contrast, the structure functions predicted by the SGS model SANN(2,2) are in good agreement with the filtered DNS data at grid resolution of  $64^3$  ( $h_{\text{LES}}^{\Delta} = \Delta/2$ ), suggesting that the SANN model can predict small scale fluctuations of turbulence accurately. This is consistent with the results of the velocity spectrum at high wave numbers near the filter width.

The probability density functions (PDFs) of the normalized velocity increment are shown in Fig. 6. We normalize the velocity increment by the rms value of velocity  $\bar{u}^{\text{rms}}$ . All the PDFs are nearly symmetrical and their tails become longer as  $\mathbf{r}$  increases. The PDFs of all the models are in reasonable agreement with fDNS at large separations. The PDFs predicted by ILES are much wider than those of fDNS at small separations. The PDFs predicted by SANN(2,2) are much closer to those of fDNS, compared to other LES models.

Figure 7 shows the comparison of the PDFs of the normalized vorticity  $\bar{\omega}/\bar{\omega}_{\text{fDNS}}^{\text{rms}}$  ( $\bar{\omega} = \sqrt{\bar{\omega}_i \bar{\omega}_i}$ ) for LES with different SGS models. The PDF of  $\bar{\omega}/\bar{\omega}_{\text{fDNS}}^{\text{rms}}$  is skewed toward the positive side. The PDF predicted by ILES is substantially higher than that of fDNS for large vorticity. Meanwhile, the PDF tail predicted by the DMM is significantly shorter than that of fDNS. The DSM and SANN(2,2) models are in reasonable agreement with fDNS for the tail of PDF. The SANN(2,2) model performs better than other models on capturing the peak of the PDF.

The temporal evolution of the instantaneous coherent structures are examined by plotting the instantaneous normalized vorticity contours of  $\bar{\omega}/\bar{\omega}_{\text{fDNS}}^{\text{rms}}$ . LES with different SGS models are



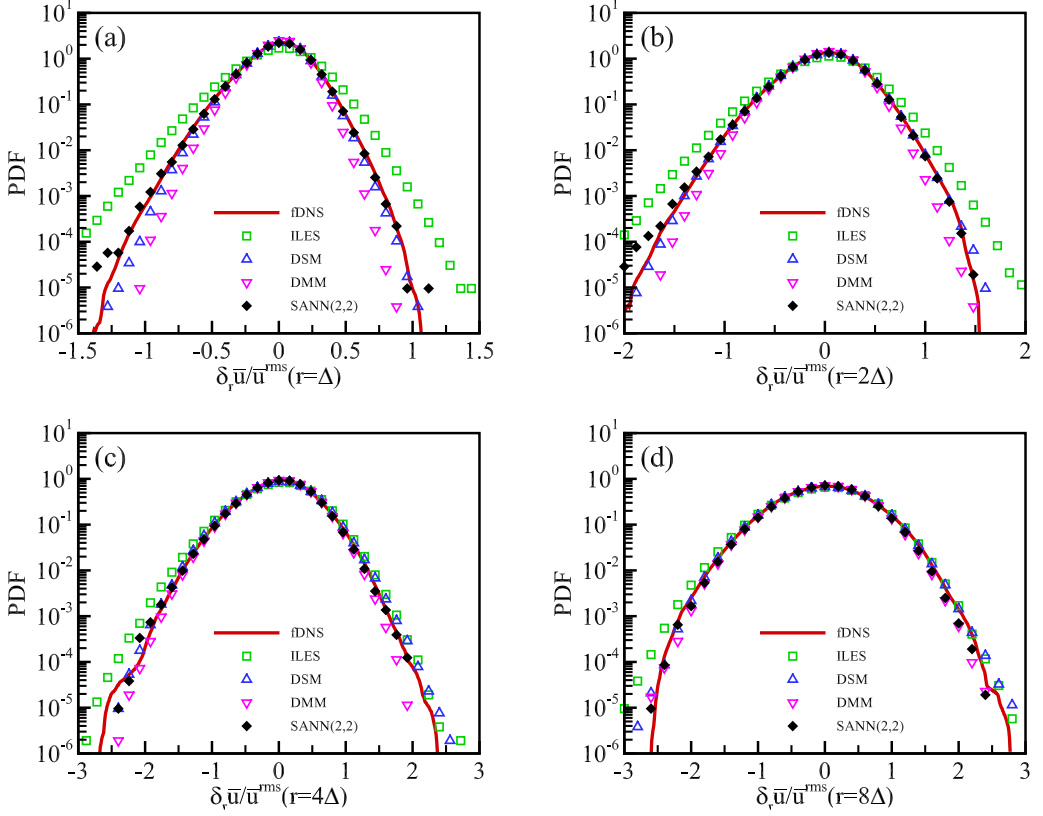


FIG. 6. PDFs of the normalized increments of the velocity for LES at grid resolutions of  $64^3$  ( $h_{LES}^\Delta = \Delta/2$ ) with the filter width  $\Delta = 32\Delta x$ : (a)  $\delta_r \bar{u}/\bar{u}^{rms}(r = \Delta)$ , (b)  $\delta_r \bar{u}/\bar{u}^{rms}(r = 2\Delta)$ , (c)  $\delta_r \bar{u}/\bar{u}^{rms}(r = 4\Delta)$ , (d)  $\delta_r \bar{u}/\bar{u}^{rms}(r = 8\Delta)$ .

simulated with the same initial condition by using an instantaneous flow field of the filtered DNS data. We display the normalized vorticity contours of fDNS and LES with different SGS models [DSM, DMM, SANN(2,2)] at the same instant  $t/\tau = 6.17$  in Fig. 8 where  $\tau \equiv L_I/u^{rms}$  is the large

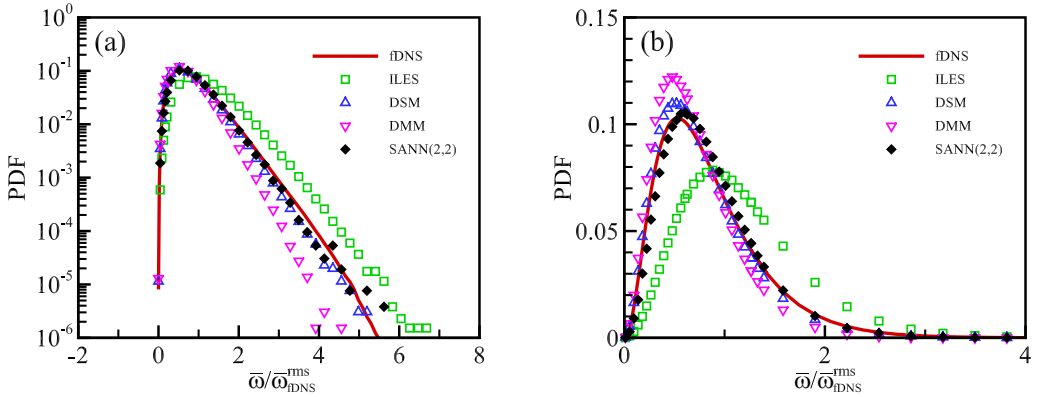


FIG. 7. PDFs of the normalized vorticity  $\bar{\omega}/\bar{\omega}_{fDNS}^{rms}$  for LES at grid resolutions of  $64^3$  ( $h_{LES}^\Delta = \Delta/2$ ) with the filter width  $\Delta = 32\Delta x$ .

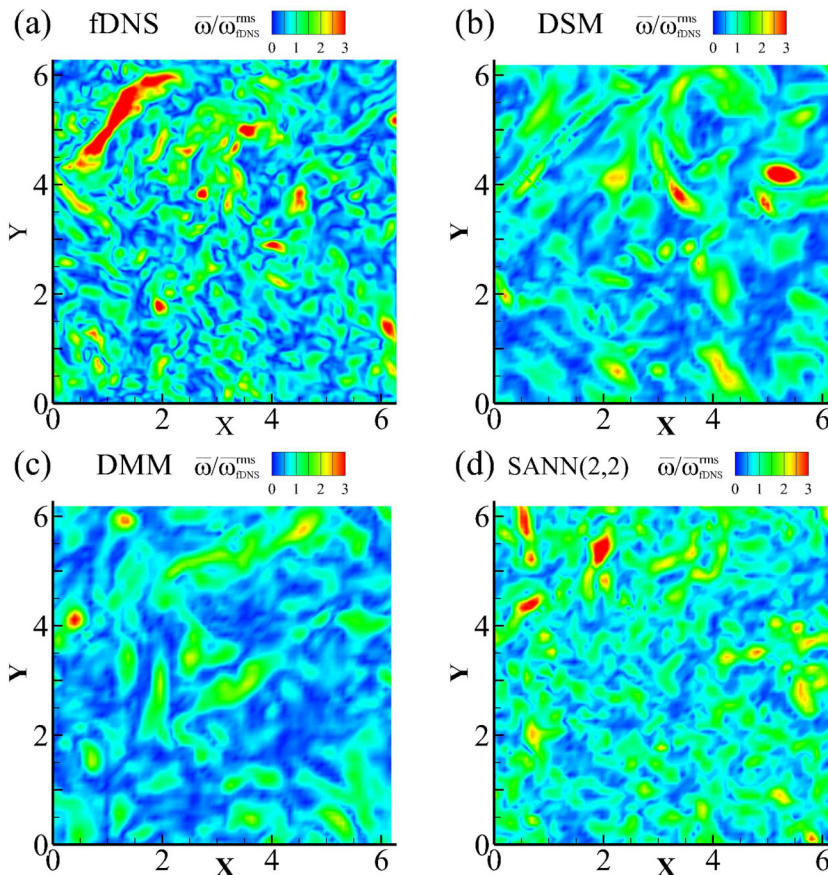


FIG. 8. Contours of the normalized vorticity  $\bar{\omega}/\bar{\omega}_{\text{fDNS}}^{\text{rms}}$  on an arbitrarily selected  $x$ - $y$  slice, at  $t/\tau = 6.17$  for LES at grid resolutions of  $64^3$  ( $h_{\text{LES}}^{\Delta} = \Delta/2$ ) with the filter width  $\Delta = 32\Delta x$ : (a) fDNS, (b) DSM, (c) DMM, and (d) SANN(2,2).

eddy turnover time. Localized vortical structures at high vortex intensity are clearly seen in fDNS. One can see that some vortical structures are missing in the predicted flow fields by the DSM and DMM models. In contrast, the results of the SANN(2,2) model are very close to those of the fDNS.

Besides, the *a posteriori* tests of the SANN(4,2) model with more input features in Appendix A show that the SANN(4,2) model produces best agreement with the velocity spectrum of the filtered DNS data, similar to the results of the SANN(2,2) model. The *a posteriori* studies of the SANN models applied to homogenous shear turbulence are shown in Appendix B. It is found that the spectrum of the velocity recovered by the SANN(2,2) model is closest to the filtered DNS data of the homogenous shear turbulence.

## VI. DISCUSSION

Forced canonical isotropic turbulence is one of standard examples for developing new SGS models of LES [14,15,99,100], where the Kolmogorov  $k^{-5/3}$  scaling of velocity spectrum in inertial range of turbulence can be observed clearly. One of important assumptions for LES modeling is that the statistical properties and flow structures of fully developed turbulence share some universal features in the inertial range where the effect of large-scale anisotropy is weak [4]. This assumption makes it easier to develop SGS models for the fully developed turbulent regions. We will try

to model the effects of boundary conditions for LES of wall-bounded turbulence in the ANN framework in a follow-up study. Another important characteristics of LES is that the SGS force and SGS stress exhibit complex spatial structures that are very irregular and diverse in the inertial range of turbulence. This fact is one of major reasons for that the relative errors of the traditional SGS models are not small in the *a priori* tests. One advantage of the rich spatial structures of SGS terms is that the possibility of data overfitting in the ANN framework is suppressed.

It is worth noting that the spatial features of filtered flow fields are important for developing an accurate SGS model. The spatial locality of kinetic energy cascade was identified for both incompressible turbulence and compressible turbulence [44–47], indicating that the flow structures at the scales between  $\Delta/2$  and  $2\Delta$  dominate the kinetic energy flux at the scale  $\Delta$ . The relative errors of single-point SGS models are not small in the *a priori* tests, partly due to the fact that these models cannot accurately capture the spatial characteristics of energy cascades. The filtered flow fields at different spatial points should be incorporated to accurately reconstruct the SGS terms. Due to the complex and irregular nature of turbulence, it is difficult to explicitly specify the dependence of SGS flux of kinetic energy on the flow dynamics at different scales by using simple forms of analytical functions. We have shown that the ANN method is a powerful tool which can efficiently learn the high-dimensional and nonlinear relations between the SGS terms and the filtered flow fields at different spatial scales. It has been demonstrated that the relative errors of proposed SANN models are much smaller than traditional SGS models in the *a priori* tests.

We have observed that the SGS models which are highly accurate in the *a priori* tests can lead to the instability in the *a posteriori* tests if the grid scale is equal to the filter width, while the same SGS models can give convergent and accurate results in the *a posteriori* tests using the fine grids with the same resolution as the DNS. In order to suppress the effect of numerical errors and achieve highly accurate results in the *a posteriori* tests, we use the grid scale which is smaller than the filter width. Moreover, we add the numerical dissipation individually to keep the stability of LES, rather than using the dissipative SGS models such as the Smagorinsky model. The numerical dissipation is necessary to mimic the flux of kinetic energy to smaller scales which cannot be resolved at the LES grids. It is expected that as the grid scale becomes smaller and the filter width is fixed, the effects of numerical scheme and numerical dissipation decrease, and the flow fields of LES using accurate SGS models can be convergent. We have shown that the statistical properties of the LES using SANN models are very close to those of the filtered DNS data, while the statistics of LES using traditional SGS models deviate substantially from the filtered DNS data at the scales close to the filter width. If no SGS model is used in coarse-grid numerical simulations with artificial dissipations (implicit LES), then the results cannot be easily convergent as the grid scale becomes smaller: The velocity spectrum will become longer and finally convergent to that of the DNS field. It is worth noting that the degree of freedom for the filtered DNS field is much smaller than that of the DNS field. Thus, we can achieve convergent results of LES using the explicit SGS models more easily, as compared to the implicit LES.

## VII. CONCLUSIONS

In this work, we propose a framework of SANN for the SGS forces in large eddy simulation of turbulence. The proposed SANN models depend on two parameters  $R_s$  and  $R_g$ . As  $R_s$  and  $R_g$  increase, the accuracy of the SANN models increase. The correlation coefficients can be made larger than 0.99 and the relative errors smaller than 15% for the SANN models with reasonable  $R_s$  and  $R_g$  in the *a priori* test, which are more accurate than traditional gradient models. In an *a posteriori* analysis, the performances of the SANN(2,2) model are compared with that of the ILES, DSM, and DMM models in the prediction of the spectra and statistical properties of velocity field, as well as the instantaneous flow structures. The velocity spectra predicted by the DSM and DMM models are typical tilted: They are damped too strongly near the cut-off wave number and are too energy rich at low wave numbers. In contrast, SANN(2,2) predicts the spectrum accurately, and it

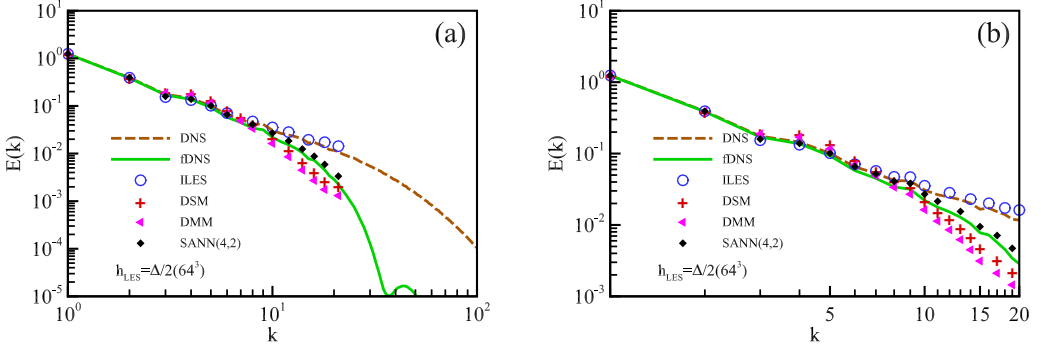


FIG. 9. Spectrum of velocity for LES at grid resolutions of  $64^3$  ( $h_{LES} = \Delta/2$ ) with the filter width  $\Delta = 32\Delta x$ .

also accurately reconstructs the PDFs of velocity increment and the vorticity. In addition, turbulence structures predicted by the SGS model SANN(2,2) are also close to those of the fDNS data.

Several issues require further investigation. These include the physical mechanism for the complex spatial correlation between the filtered flow fields and the SGS terms, the interpretability, efficiency and universality of the neural network models, and the spatial-temporal characteristics of the SGS dynamics.

#### ACKNOWLEDGMENTS

This work was supported in part by the National Natural Science Foundation of China (NSFC Grants No. 91952104, No. 11702127, and No. 91752201), and by the Technology and Innovation Commission of Shenzhen Municipality (Grants No. KQTD20180411143441009, No. JCYJ20170412151759222, and No. ZDSYS201802081843517). This work was also supported by the Center for Computational Science and Engineering of the Southern University of Science and Technology. J.W. acknowledges the support from the Young Elite Scientist Sponsorship Program by CAST (Grant No. 2016QNRC001).

#### APPENDIX A: THE *A POSTERIORI* TEST OF THE SANN(4,2) MODEL

In order to analyze the effect of the input features on the SANN models, an *a posteriori* study of the SANN(4,2) model applied to LES of turbulence with grid resolution  $64^3$  ( $\Delta = 2h^{LES}$ ) has been compared in this section. The SANN(4,2) model contains more input features than the SANN(2,2) model. Figure 9 shows the spectrum of velocity for different SGS models. The SANN(4,2) model performs better than other models.

#### APPENDIX B: THE *A POSTERIORI* TEST OF SHEAR FLOW

In order to provide more promising results that the ANN model can be applied to more complex conditions, we apply the SANN(2,2) model to LES of weakly compressible homogeneous shear turbulence (HST) at Taylor microscale Reynolds number  $Re_\lambda \approx 108$  and  $M_t = 0.20$  [101–103].

Based on the dimensionless governing equations for the compressible HST [101–103], the filtered dimensionless Navier-Stokes equations for the resolved variables can be expressed as follows [6,80,81,86]:

$$\frac{\partial \bar{\rho}}{\partial t} + \frac{\partial(\bar{\rho}\tilde{u}_j)}{\partial x_j} + Sx_2 \frac{\partial \bar{\rho}}{\partial x_1} = 0, \quad (\text{B1})$$

TABLE XI. Numerical simulation parameters and statistical quantities for compressible homogeneous shear turbulence [102].

Resol.	Re	$M$	$\text{Re}_\lambda$	$M_T$	$\eta/\Delta x$
$512 \times 256 \times 256$	220	0.10	108	0.20	0.79

$$\frac{\partial(\bar{\rho}\tilde{u}_i)}{\partial t} + \frac{\partial(\bar{\rho}\tilde{u}_i\tilde{u}_j + \bar{p}\delta_{ij})}{\partial x_j} - \frac{1}{\text{Re}} \frac{\partial\tilde{\sigma}_{ij}}{\partial x_j} - \tilde{\mathcal{F}}_i + Sx_2 \frac{\partial(\bar{\rho}\tilde{u}_i)}{\partial x_1} + S\bar{\rho}\tilde{u}_2\delta_{i1} = -\frac{\partial\bar{\rho}\tau_{ij}}{\partial x_j} + \frac{1}{\text{Re}} \frac{\partial(\tilde{\sigma}_{ij} - \bar{\sigma}_{ij})}{\partial x_j}, \quad (\text{B2})$$

$$\frac{\partial\tilde{\mathcal{E}}}{\partial t} + \frac{\partial[(\tilde{\mathcal{E}} + \bar{p})\tilde{u}_j]}{\partial x_j} - \frac{1}{\text{Re}} \frac{\partial(\tilde{\sigma}_{ij}\tilde{u}_i)}{\partial x_j} - \frac{1}{\alpha} \frac{\partial}{\partial x_j} \left( \tilde{\kappa} \frac{\tilde{T}}{x_j} \right) + \bar{\Lambda} - \tilde{u}_j\tilde{\mathcal{F}}_j + Sx_2 \frac{\partial\tilde{\mathcal{E}}}{\partial x_1} + \bar{\rho}S\tilde{u}_1\tilde{u}_2 = R_\mathcal{E}, \quad (\text{B3})$$

$$R_\mathcal{E} = -\tilde{u}_i \frac{\partial(\bar{\rho}\tau_{ij})}{\partial x_j} - \frac{1}{(\gamma-1)\gamma M^2} \frac{\partial\bar{\rho}Q_j}{\partial x_j} - \Pi_{\text{dil}} + \frac{1}{\text{Re}} \varepsilon_v + \frac{\tilde{u}_i}{\text{Re}} \frac{\partial(\tilde{\sigma}_{ij} - \bar{\sigma}_{ij})}{\partial x_j} + \frac{1}{\alpha} \frac{\partial}{\partial x_j} \left( \tilde{\kappa} \frac{\partial\tilde{T}}{\partial x_j} - \tilde{\kappa} \frac{\partial\tilde{T}}{\partial x_j} \right), \quad (\text{B4})$$

$$\bar{p} = \bar{\rho}\tilde{T}/(\gamma M^2), \quad (\text{B5})$$

where the Favre filtering is applied [104]:  $\tilde{f} = \overline{\rho f}/\bar{\rho}$ , where  $\rho$  denotes the fluid density, and  $f$  represents the velocity component or temperature. The resolved total energy  $\tilde{\mathcal{E}}$  is defined by  $\tilde{\mathcal{E}} = \frac{\bar{p}}{\gamma-1} + \frac{1}{2}\bar{\rho}(\tilde{u}_j\tilde{u}_j)$  [6,81], the filtered viscous stress is  $\tilde{\sigma}_{ij} = 2\tilde{\mu}\tilde{S}_{ij} - \frac{2}{3}\tilde{\mu}\delta_{ij}\tilde{S}_{kk}$ , where  $\tilde{S}_{ij} = \frac{1}{2}(\partial\tilde{u}_i/\partial x_j + \partial\tilde{u}_j/\partial x_i)$ , and  $\tilde{\mu}$  and  $\tilde{\kappa}$  are calculated from Sutherland's law.

The SGS terms appearing on the right-hand sides of Eqs. (B1)–(B5) are defined as

$$\tau_{ij} = \widetilde{u_i u_j} - \tilde{u}_i \tilde{u}_j, \quad Q_j = \widetilde{u_j T} - \tilde{u}_j \tilde{T}, \quad \Pi_{\text{dil}} = \overline{p S_{kk}} - \bar{p} \tilde{S}_{kk}, \quad \varepsilon_v = \overline{\sigma_{ji} S_{ij}} - \bar{\sigma}_{ji} \tilde{S}_{ij}, \quad (\text{B6})$$

$$M_i = \frac{\partial\bar{\rho}\tau_{ij}}{\partial x_j}, \quad M_T = \frac{\partial\bar{\rho}Q_j}{\partial x_j}, \quad (\text{B7})$$

where  $\tau_{ij}$  is the SGS stress,  $Q_j$  is SGS heat flux,  $\Pi_{\text{dil}}$  is SGS pressure-dilatation,  $\varepsilon_v$  is SGS viscous dissipation,  $M_i$  is the SGS force, and  $M_T$  is the divergence of SGS heat flux. We model the SGS force  $M_i$  and the divergence of SGS heat flux  $M_T$ , and neglect other unclosed terms [22,74,81,86,106].

The compressible homogeneous shear turbulence is solved by a hybrid scheme [83], which combines an eighth-order compact finite-difference scheme [105] for smooth regions and a seventh-order weighted essentially nonoscillatory scheme [107] for shock regions. The hybrid scheme can be combined with Rogallo's method to simulate the compressible HST [101–103]. The governing equations of compressible HST are solved with periodic boundary conditions in a rectangular domain with side lengths  $L_x = 4\pi$ ,  $L_y = L_z = 2\pi$  [102,103]. A uniform thermal cooling  $\Lambda$  is employed to sustain the internal energy in a statistically steady state [83]. The simulation parameters for the HST are summarized in Table XI.

In the *a posteriori* tests of the HST, the SANN(2,2) model trained in the incompressible homogenous isotropic turbulence is used to reconstruct  $M_i (i = 1, 2, 3)$  with grid resolution of  $128 \times 64 \times 64$  ( $h_{\text{LES}} = \Delta/2$ ,  $\Delta = 8\Delta x$ ) [92–94], meanwhile,  $M_T$  is modeled by the DSM model [74,86,106]. Further more, simulations using the DSM and DMM models are also carried out. The ratio of the time steps in LES and DNS is  $\Delta t_{\text{LES}}/\Delta t_{\text{DNS}} = 8$ . Figure 10 shows the spectrum of

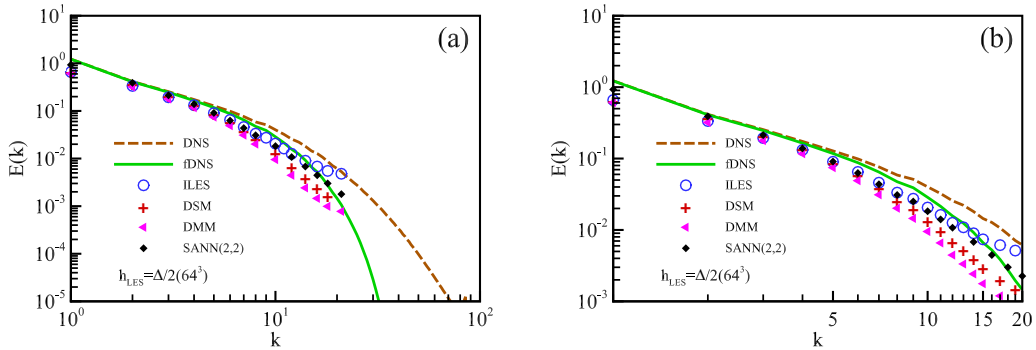


FIG. 10. Spectrum of velocity for LES at grid resolutions of  $64^3$  ( $h_{LES} = \Delta/2$ ) with the filter width  $\Delta = 32\Delta x$ .

velocity for different SGS models. The spectrum of the velocity predicted by the SANN(2,2) model is closest to the filtered DNS data of HST. Thus, the proposed SANN(2,2) model can recover the spectrum more accurately than the other models for homogeneous shear turbulence.

- [1] J. Smagorinsky, General circulation experiments with the primitive equations, i. The basic experiment, *Mon. Weath. Rev.* **91**, 99 (1963).
- [2] D. K. Lilly, The representation of small-scale turbulence in numerical simulation experiments, in *Proceedings of the IBM Scientific Computing Symposium on Environmental Sciences*, edited by H. H. Goldstone (IBM, Yorktown Heights, NY, 1967), p. 195.
- [3] J. W. Deardorff, A numerical study of three-dimensional turbulent channel flow at large Reynolds numbers, *J. Fluid Mech.* **41**, 453 (1970).
- [4] S. B. Pope, *Turbulent Flows* (Cambridge University Press, Cambridge, UK, 2000).
- [5] P. Sagaut, *Large Eddy Simulation for Incompressible Flows* (Springer, Berlin, 2006).
- [6] E. Garnier, N. Adams, and P. Sagaut, *Large Eddy Simulation for Compressible Flows* (Springer-Verlag, Berlin, 2009).
- [7] P. Moin, K. Squires, W. Cabot, and S. Lee, A dynamic subgrid-scale model for compressible turbulence and scalar transport, *Phys. Fluids A* **3**, 2746 (1991).
- [8] M. Germano, Turbulence: The filtering approach, *J. Fluid Mech.* **238**, 325 (1992).
- [9] C. Meneveau and J. Katz, Scale-invariance and turbulence models for large-eddy simulation, *Annu. Rev. Fluid Mech.* **32**, 1 (2000).
- [10] D. I. Pullin, A vortex-based model for the subgrid flux of a passive scalar, *Phys. Fluids* **12**, 2311 (2000).
- [11] C. Meneveau, Lagrangian dynamics and models of the velocity gradient tensor in turbulent flows, *Annu. Rev. Fluid Mech.* **43**, 219 (2011).
- [12] S. Y. Chen, Z. H. Xia, S. Y. Pei, J. C. Wang, Y. T. Yang, Z. L. Xiao, and Y. P. Shi, Reynolds-stress-constrained large-eddy simulation of wall-bounded turbulent flows, *J. Fluid Mech.* **703**, 1 (2012).
- [13] P. A. Durbin, Some recent developments in turbulence closure modeling, *Annu. Rev. Fluid Mech.* **50**, 77 (2018).
- [14] M. Buzicotti, M. Linkmann, H. Aluie, L. Biferale, J. Brasseur, and C. Meneveau, Effect of filter type on the statistics of energy transfer between resolved and sub-filter scales from a-priori analysis of direct numerical simulations of isotropic turbulence, *J. Turbul.* **19**, 167 (2018).
- [15] M. Linkmann, M. Buzicotti, and L. Biferale, Multi-scale properties of large eddy simulations: correlations between resolved-scale velocity-field increments and subgrid-scale quantities, *J. Turbul.* **19**, 493 (2018).



- [16] M. Germano, U. Piomelli, P. Moin, and W. Cabot, A dynamic subgrid-scale eddy-viscosity model, *Phys. Fluids A* **3**, 1760 (1991).
- [17] D. Lilly, A proposed modification of the Germano subgrid scale closure method, *Phys. Fluids A* **4**, 633 (1992).
- [18] J. W. Deardorff, Three-dimensional numerical modeling of the planetary boundary layer, in *Workshop on Micrometeorology*, edited by D. A. Haugen (American Meteorological Society, Boston, 1973), p. 271.
- [19] U. Schumann, Subgrid scale model for finite difference simulations of turbulent flows in plane channels and annuli, *J. Comput. Phys.* **18**, 376 (1975).
- [20] S. Menon and W. W. Kim, High Reynolds number flow simulations using the localized dynamic subgrid-scale model, in *Proceedings of the 34th AIAA Aerospace Sciences Meeting and Exhibit*, AIAA Paper No. 96-0425 (AIAA, Reno, NV, 1996).
- [21] F. Génin and S. Menon, Dynamics of sonic jet injection into supersonic cross-flow, *J. Turbul.* **11**, N4 (2010).
- [22] X. C. Chai and K. Mahesh, Dynamic  $k$ -equation model for large-eddy simulation of compressible flows, *J. Fluid Mech.* **699**, 385 (2012).
- [23] C. G. Speziale, G. Erlebacher, T. A. Zang, and M. Y. Hussaini, The subgrid-scale modeling of compressible turbulence, *Phys. Fluids* **31**, 940 (1988).
- [24] G. Erlebacher, M. Hussaini, C. Speziale, and T. Zang, Toward the large-eddy simulation of compressible turbulent flows, *J. Fluid Mech.* **238**, 155 (1992).
- [25] J. Bardina, J. H. Ferziger, and W. C. Reynolds, Improved subgrid scale models for large eddy simulation, in *Proceedings of the 13th Fluid and Plasma Dynamics Conference*, AIAA Paper No. 80-1357 (AIAA, Reston, VA, 1980).
- [26] T. A. Zang, R. B. Dahlburg, and J. P. Dahlburg, Direct and large-eddy simulations of three-dimensional compressible Navier-Stokes turbulence, *Phys. Fluids A* **4**, 127 (1992).
- [27] S. Liu, C. Meneveau, and J. Katz, On the properties of similarity subgridscale models as deduced from measurements in a turbulent jet, *J. Fluid Mech.* **275**, 83 (1994).
- [28] Y. P. Shi, Z. L. Xiao, and S. Y. Chen, Constrained subgrid-scale stress model for large eddy simulation, *Phys. Fluids* **20**, 011701 (2008).
- [29] J. A. Langford and R. D. Moser, Optimal LES formulations for isotropic turbulence, *J. Fluid. Mech.* **398**, 321 (1999).
- [30] J. A. Langford and R. D. Moser, Optimal large-eddy simulation results for isotropic turbulence, *J. Fluid. Mech.* **521**, 273 (2004).
- [31] R. D. Moser, N. P. Malaya, H. Chang, P. S. Zandonade, P. Vedula, A. Bhattacharya, and A. Haselbacher, Theoretically based optimal large-eddy simulation, *Phys. Fluids*. **21**, 105104 (2009).
- [32] J. P. Boris, F. F. Grinstein, E. S. Oran, and R. L. Kolbe, New insights into large Eddy simulation, *Fluid Dyn. Res.* **10**, 199 (1992).
- [33] E. Garnier, M. Mossi, P. Sagaut, P. Comte, and M. Deville, On the use of shockcapturing schemes for large-eddy simulation, *J. Comput. Phys.* **153**, 273 (1999).
- [34] M. R. Visbal, P. E. Morgan, and D. P. Rizzetta, An implicit LES approach based on high-order compact differencing and filtering schemes, in *Proceedings of the 16th AIAA Computational Fluid Dynamics Conference*, AIAA Paper No. 2003-4098 (AIAA, Orlando, Florida, 2003).
- [35] N. A. Adams, S. Hickel, and S. Franz, Implicit subgrid-scale modeling by adaptive deconvolution, *J. Comput. Phys.* **200**, 412 (2004).
- [36] F. F. Grinstein, L. G. Margolin, and W. J. Rider, *Implicit Large Eddy Simulation, Computing Turbulent Fluid Dynamics* (Cambridge University Press, Cambridge, UK, 2007).
- [37] T. J. R. Hughes, L. Mazzei, and K. E. Jansen, Large eddy simulation and the variational multiscale method, *Comput. Vis. Sci.* **3**, 47 (2000).
- [38] C. Brun and R. Friedrich, Modeling the test SGS tensor  $T_{ij}$ : An issue in the dynamic approach, *Phys. Fluids* **13**, 2373 (2001).
- [39] B. Kosović, D. I. Pullin, and R. Samtaney, Subgrid-scale modeling for large-eddy simulations of compressible turbulence, *Phys. Fluids* **14**, 1511 (2002).



- [40] C. Brun, R. Friedrich, and C. B. da Silva, A non-linear SGS model based on the spatial velocity increment, *Theor. Comput. Fluid Dyn.* **20**, 1 (2006).
- [41] D. You, S. Bose, and P. Moin, Grid-independent large-eddy simulation of compressible turbulent flows using explicit filtering, in *Proceedings of the 2010 Center for Turbulence Research Summer Program* (Stanford University, 2010), pp. 203–210.
- [42] H. Xiao and P. Jenny, A consistent dual-mesh framework for hybrid LES/RANS modeling, *J. Comput. Phys.* **231**, 1848 (2012).
- [43] H. Xiao, J. X. Wang, and P. Jenny, An implicitly consistent formulation of a dual-mesh hybrid LES/RANS method, *Commun. Comput. Phys.* **21**, 570 (2017).
- [44] G. L. Eyink, Locality of turbulent cascades, *Physica D* **207**, 91 (2005).
- [45] G. L. Eyink and H. Aluie, Localness of energy cascade in hydrodynamic turbulence. I. Smooth coarse graining, *Phys. Fluids* **21**, 115107 (2009).
- [46] H. Aluie and G. L. Eyink, Localness of energy cascade in hydrodynamic turbulence. II. Sharp spectral filter, *Phys. Fluids* **21**, 115108 (2009).
- [47] H. Aluie, Compressible Turbulence: The Cascade and its Locality, *Phys. Rev. Lett.* **106**, 174502 (2011).
- [48] S. T. Bose, P. Moin, and D. You, Grid-independent large-eddy simulation using explicit filtering, *Phys. Fluids* **22**, 105103 (2010).
- [49] S. Radhakrishnan and J. Bellan, Explicit filtering to obtain grid-spacing-independent and discretization-order-independent large-eddy simulation of compressible single-phase flow, *J. Fluid Mech.* **697**, 399 (2012).
- [50] U. Piomelli, A. Rouhi, and B. J. Geurts, A grid-independent length scale for large-eddy simulations, *J. Fluid Mech.* **766**, 499 (2015).
- [51] F. Sarghini, G. de Felice, and S. Santini, Neural networks based subgrid scale modeling in large eddy simulations, *Comput. Fluids* **32**, 97 (2003).
- [52] J. Ling and J. Templeton, Evaluation of machine learning algorithms for prediction of regions of high Reynolds averaged Navier Stokes uncertainty, *Phys. Fluids* **27**, 085103 (2015).
- [53] Z. J. Zhang and K. Duraisamy, Machine learning methods for data-driven turbulence modeling, in *Proceedings of the 22nd AIAA Computational Fluid Dynamics Conference*, AIAA Paper No. 2015-2460 (AIAA, Reston, VA, 2015).
- [54] J. Ling, A. Kurzawski, and J. Templeton, Reynolds averaged turbulence modeling using deep neural networks with embedded invariance, *J. Fluid Mech.* **807**, 155 (2016).
- [55] J. Ling, R. Jones, and J. Templeton, Machine learning strategies for systems with invariance properties, *J. Comput. Phys.* **318**, 22 (2016).
- [56] H. Xiao, J. L. Wu, J. X. Wang, J. X. Sun, and C. J. R. Roy, Quantifying and reducing model-form uncertainties in Reynolds-averaged Navier-Stokes simulations: A data-driven, physics-informed Bayesian approach, *J. Comput. Phys.* **324**, 115 (2016).
- [57] B. D. Tracey, K. Duraisamy, and J. J. Alonso, A machine learning strategy to assist turbulence model development, in *Proceedings of the 53rd AIAA Aerospace Sciences Meeting, Kissimmee, FL*, AIAA Paper No. 2015-1287 (AIAA, Reston, VA, 2015).
- [58] S. L. Brunton, J. L. Proctor, and J. N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, *Proc. Natl Acad. Sci. USA* **113**, 3932 (2016).
- [59] J. X. Wang, J. L. Wu, and H. Xiao, Physics-informed machine learning approach for reconstructing reynolds stress modeling discrepancies based on dns data, *Phys. Rev. Fluids* **2**, 034603 (2017).
- [60] M. Gamahara and Y. Hattori, Searching for turbulence models by artificial neural network, *Phys. Rev. Fluids* **2**, 054604 (2017).
- [61] A. Vollant, G. Balarac, and C. Corre, Subgrid-scale scalar flux modeling based on optimal estimation theory and machine-learning procedures, *J. Turbulence* **18**, 854 (2017).
- [62] R. Maulik and O. San, A neural network approach for the blind deconvolution of turbulent flows, *J. Fluid Mech.* **831**, 151 (2017).
- [63] J. N. Kutz, Deep learning in fluid dynamics, *J. Fluid Mech.* **814**, 1 (2017).
- [64] C. Ma, J. Wang, and W. E, Model reduction with memory and the machine learning of dynamical systems, *Commun. Comput. Phys.* **25**, 947 (2019).

- [65] A. T. Mohan and D. V. Gaitonde, A deep learning based approach to reduced order modeling for turbulent flow control using LSTM neural networks, [arXiv:1804.09269](#).
- [66] M. Raissi and G. E. Karniadakis, Hidden physics models: Machine learning of nonlinear partial differential equations, *J. Comput. Phys.* **357**, 125 (2018).
- [67] Z. Y. Wan, P. Vlachas, P. Koumoutsakos, and T. Sapsis, Data-assisted reduced-order modeling of extreme events in complex dynamical systems, *PloS One* **13**, e0197704 (2018).
- [68] J.-L. Wu, H. Xiao, and E. Paterson, Physics-informed machine learning approach for augmenting turbulence models: A comprehensive framework, *Phys. Rev. Fluids* **3**, 074602 (2018).
- [69] L. Jofre, S. P. Domino, and G. Iaccarino, A framework for characterizing structural uncertainty in large-eddy simulation closures, *Fluid, Turbul. Combust.* **100**, 341 (2018).
- [70] Z. Wang, K. Luo, D. Li, J. H. Tan, and J. R. Fan, Investigations of data-driven closure for subgrid-scale stress in large-eddy simulation, *Phys. Fluids* **30**, 125101 (2018).
- [71] R. Maulik, O. San, A. Rasheed, and P. Vedula, Subgrid modeling for two-dimensional turbulence using neural networks, *J. Fluid Mech.* **858**, 122 (2019).
- [72] K. Duraisamy, G. Iaccarino, and H. Xiao, Turbulence modeling in the age of data, *Annu. Rev. Fluid Mech.* **51**, 357 (2019).
- [73] J. L. Wu, H. Xiao, R. Sun, and Q. Q. Wang, Reynolds-averaged Navier-Stokes equations with explicit data-driven Reynolds stress closure can be ill-conditioned, *J. Fluid Mech.* **869**, 553 (2019).
- [74] C. Y. Xie, J. C. Wang, K. Li, and C. Ma, Artificial neural network approach to large-eddy simulation of compressible isotropic turbulence, *Phys. Rev. E* **99**, 053113 (2019).
- [75] C. Y. Xie, K. Li, C. Ma, and J. C. Wang, Modeling subgrid-scale force and divergence of heat flux of compressible isotropic turbulence by artificial neural network, *Phys. Rev. Fluids* **4**, 104605 (2019).
- [76] C. Y. Xie, J. C. Wang, H. Li, M. P. Wan, and S. Y. Chen, Artificial neural network mixed model for large eddy simulation of compressible isotropic turbulence, *Phys. Fluids* **31**, 085112 (2019).
- [77] R. Maulik, O. San, J. D. Jacob, and C. Crick, Sub-grid scale model classification and blending through deep learning, *J. Fluid Mech.* **870**, 784 (2019).
- [78] A. Beck, D. Flad, and C.-D. Munz, Deep neural networks for data-driven LES closure models, *J. Comput. Phys.* **398**, 108910 (2019).
- [79] T. Ishihara, T. Gotoh, and Y. Kaneda, Study of high-Reynolds number isotropic turbulence by direct numerical simulation, *Annu. Rev. Fluid Mech.* **41**, 165 (2009).
- [80] A. Leonard, Energy cascade in large-eddy simulations of turbulent fluid flows, *Adv. Geophys.* **18**, 237 (1975).
- [81] B. Vreman, B. Geurts, and H. Kuerten, Subgrid-modeling in LES of compressible flow, *Appl. Sci. Res.* **54**, 191 (1995).
- [82] G. S. Patterson and S. A. Orszag, Spectral calculations of isotropic turbulence: Efficient removal of aliasing interactions, *Phys. Fluids* **14**, 2538 (1971).
- [83] J. Wang, L.-P. Wang, Z. Xiao, Y. Shi, and S. Chen, A hybrid numerical simulation of isotropic compressible turbulence, *J. Comp. Phys.* **229**, 5257 (2010).
- [84] J. C. Wang, M. P. Wan, S. Chen, and S. Y. Chen, Kinetic energy transfer in compressible isotropic turbulence, *J. Fluid Mech.* **841**, 581 (2018).
- [85] T. Ishihara, Y. Kaneda, M. Yokokawa, K. Itakura, and A. Uno, Small-scale statistics in high-resolution direct numerical simulation of turbulence: Reynolds number dependence of one-point velocity gradient statistics, *J. Fluid Mech.* **592**, 335 (2007).
- [86] P. Martín, U. Piomelli, and G. Candler, Subgrid-scale models for compressible large-eddy simulations, *Theor. Comput. Fluid Dyn.* **13**, 361 (2000).
- [87] G. Zhang, B. E. Patuwo, and M. Y. Hu, Forecasting with artificial neural networks: The state of the art, *Int. J. Forecast.* **14**, 35 (1998).
- [88] H. B. Demuth, M. H. Beale, O. D. Jess, and M. T. Hagan, *Neural Network Design* (Martin Hagan, Oklahoma State University, Stillwater, United States, 2014).
- [89] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, [arXiv:1412.6980](#).
- [90] K. W. Bedford and W. K. Yeo, *Large Eddy Simulation of Complex Engineering and Geophysical Flows* (Cambridge University Press, Cambridge, UK, 1993).

- [91] D. G. Vlaykov, P. Grete, W. Schmidt, and D. R. G. Schleicher, A nonlinear structural subgrid-scale closure for compressible MHD. I. Derivation and energy dissipation properties, *Phys. Plasmas* **23**, 062316 (2016).
- [92] S. Ghosal, An analysis of numerical errors in large-eddy simulations of turbulence, *J. Comput. Phys.* **125**, 187 (1996).
- [93] J. Meyers, B. J. Geurts, and M. Baelmans, Database analysis of errors in large-eddy simulation, *Phys. Fluids* **15**, 2740 (2003).
- [94] F. K. Chow and P. Moin, A further study of numerical errors in large-eddy simulations, *J. Comput. Phys.* **184**, 366 (2003).
- [95] O. J. Mcmillan and J. H. Ferziger, Direct testing of subgrid-scale models, *AIAA J.* **17**, 1340 (1979).
- [96] S. B. Pope, Ten questions concerning the large-eddy simulation of turbulent flows, *New J. Phys.* **6**, 35 (2004).
- [97] C. Meneveau, Statistics of turbulence subgrid-scale stresses: Necessary conditions and experimental tests, *Phys. Fluids* **6**, 815 (2004).
- [98] R. J. A. M. Stevens, M. Wilczek, and C. Meneveau, Large-eddy simulation study of the logarithmic law for second- and higher-order moments in turbulent wall-bounded flow, *J. Fluid Mech.* **757**, 888 (2014).
- [99] M. Kessar, G. Balarac, and F. Plunian, The effect of subgrid-scale models on grid-scale/subgrid-scale energy transfers in large-eddy simulation of incompressible magnetohydrodynamic turbulence, *Phys. Plasmas* **23**, 102305 (2016).
- [100] G. P. Offermans, L. Biferale, M. Buzzicotti, and M. Linkman, *A priori* study of the subgrid energy transfers for small-scale dynamo in kinematic and saturation regimes, *Phys. Plasmas* **25**, 122307 (2018).
- [101] Z. Q. Ma and Z. L. Xiao, Turbulent kinetic energy production and flow structures in compressible homogeneous shear flow, *Phys. Fluids* **28**, 096102 (2016).
- [102] S. Chen, J. C. Wang, H. Li, M. P. Wan, and S. Y. Chen, Spectra and Mach number scaling in compressible homogeneous shear turbulence, *Phys. Fluids* **30**, 065109 (2018).
- [103] S. Chen, J. C. Wang, H. Li, M. P. Wan, and S. Y. Chen, Effect of compressibility on small scale statistics in homogeneous shear turbulence, *Phys. Fluids* **31**, 025107 (2019).
- [104] A. Favre, Equations des gaz turbulents compressible. I. Formes generales, *J. Mec.* **4**, 361 (1965).
- [105] S. K. Lele, Compact finite difference schemes with spectral-like resolution, *J. Comput. Phys.* **103**, 16 (1992).
- [106] C. Y. Xie, J. C. Wang, H. Li, M. P. Wan, and S. Y. Chen, A modified optimal LES model for highly compressible isotropic turbulence, *Phys. Fluids* **30**, 065108 (2018).
- [107] D. S. Balsara and C. W. Shu, Monotonicity preserving weighted essentially nonoscillatory schemes with increasingly high order of accuracy, *J. Comput. Phys.* **160**, 405 (2000).