

Artificial neural network mixed model for large eddy simulation of compressible isotropic turbulence

Cite as: Phys. Fluids **31**, 085112 (2019); <https://doi.org/10.1063/1.5110788>

Submitted: 21 May 2019 . Accepted: 22 July 2019 . Published Online: 20 August 2019

Chenyue Xie (谢晨月),  Jianchun Wang (王建春),  Hui Li (李辉), Minping Wan (万敏平), and Shiyi Chen (陈十一)



[View Online](#)



[Export Citation](#)



[CrossMark](#)

ARTICLES YOU MAY BE INTERESTED IN

Machine learning methods for turbulence modeling in subsonic flows around airfoils
Physics of Fluids **31**, 015105 (2019); <https://doi.org/10.1063/1.5061693>

Investigations of data-driven closure for subgrid-scale stress in large-eddy simulation
Physics of Fluids **30**, 125101 (2018); <https://doi.org/10.1063/1.5054835>

Spatially multi-scale artificial neural network model for large eddy simulation of compressible isotropic turbulence
AIP Advances **10**, 015044 (2020); <https://doi.org/10.1063/1.5138681>

Physics of Fluids
SPECIAL TOPIC: Tribute to
Frank M. White on his 88th Anniversary

SUBMIT TODAY!

Artificial neural network mixed model for large eddy simulation of compressible isotropic turbulence

Cite as: Phys. Fluids 31, 085112 (2019); doi: 10.1063/1.5110788

Submitted: 21 May 2019 • Accepted: 22 July 2019 •

Published Online: 20 August 2019



View Online



Export Citation



CrossMark

Chenyue Xie (谢晨月),^{1,2} Jianchun Wang (王建春),^{1,a)} Hui Li (李辉),² Minping Wan (万敏平),¹ and Shiying Chen (陈十一)^{1,3,a)}

AFFILIATIONS

¹ Shenzhen Key Laboratory of Complex Aerospace Flows, Department of Mechanics and Aerospace Engineering, Southern University of Science and Technology, Shenzhen 518055, People's Republic of China

² School of Power and Mechanical Engineering, Wuhan University, Wuhan 430072, People's Republic of China

³ State Key Laboratory of Turbulence and Complex Systems, Center for Applied Physics and Technology, College of Engineering, Peking University, Beijing 100871, People's Republic of China

^{a)} Authors to whom correspondence should be addressed: wangjc@sustech.edu.cn and chensy@sustech.edu.cn

ABSTRACT

In this work, the subgrid-scale (SGS) stress and the SGS heat flux of compressible isotropic turbulence are modeled by an artificial neural network (ANN) mixed model (ANNMM), which maintains both functional and structural performances. The functional form of the mixed model combining the gradient model and the Smagorinsky's eddy viscosity model is imposed, and the ANN is used to calculate the model coefficients of the SGS anisotropy stress, SGS energy, and SGS heat flux. It is shown that the ANNMM can reconstruct the SGS terms more accurately than the gradient model in the *a priori* test. Specifically, the ANNMM almost recovers the average values of the SGS energy flux and SGS energy flux conditioned on the normalized filtered velocity divergence. In an *a posteriori* analysis, the ANNMM shows advantage over the dynamic Smagorinsky model (DSM) and dynamic mixed model (DMM) in the prediction of the spectra of velocity and temperature, which almost overlap with the filtered direct numerical simulation data, while the DSM and DMM suffer from the problem of the typical tilted spectral distribution. Besides, the ANNMM predicts the probability density functions of SGS energy flux much better than DSM and DMM. ANN with functional model forms can enlighten and deepen our understanding of large eddy simulation modeling.

Published under license by AIP Publishing. <https://doi.org/10.1063/1.5110788>

I. INTRODUCTION

Large eddy simulation (LES) is more efficient than direct numerical simulation (DNS) in computational cost with high accuracy on the prediction of the large scale dynamics of turbulence.^{1–7} LES solves the large scale dynamics of turbulence that dominate the major transport of momentum, energy, and heat in turbulent flows and models the effects of small scale flow structures on the large scale dynamics.^{8–16} As compared to incompressible turbulence, compressible turbulence is more complex, due to the strong couplings of velocity field and thermodynamic field and nonlinear interactions of vortices, acoustic waves, shock waves, and expansionlike waves in compressible turbulence.^{17–25} Compressible LES

contains an extra energy equation.^{1,26–33} Transport equations with different variables are used to close the unclosed terms in compressible turbulence, including the internal energy per unit mass,^{1,34} the thermal pressure,³⁵ the enthalpy per unit mass,^{26,28} or the total energy of the filtered field.^{36,37} Although LES is widely applied for simulating complex turbulent flows in aerospace industry, combustion, astrophysics, and engineering problems, it is still the focus of the LES community to develop more reliable models for the unclosed subgrid-scale (SGS) terms. The main motivation of this paper is to reconstruct the SGS stress and SGS heat flux of compressible turbulence by artificial neural network (ANN) mixed model (ANNMM), which takes into account the functional and structural strategies.³⁸

There are lots of functional and structural closure models to predict the SGS terms, including a Smagorinsky model,^{8,39} a dynamic Smagorinsky model (DSM),^{1,40–42} a similarity model,^{43,44} a dynamic mixed model (DMM),^{26,28,35,43–45} a gradient model,^{44,46} a dynamic Clark model,⁴⁷ an optimal model,^{48–50} the dynamic nonlinear tensorial diffusivity,⁵¹ a variational multiscale formulation,⁵² a modified Leonard's stress,⁵³ a velocity increment model,⁵⁴ a dynamic k-equation model,¹³ a Reynolds-stress-constrained large-eddy simulation model,⁵ a consistent dual-mesh hybrid LES/ Reynolds-Averaged Navier-Stokes (RANS),^{55,56} a grid-independent LES solution by using explicit filtering,⁵⁷ a new eddy viscosity formulation for compressible large eddy simulation, etc.⁵⁸ The kinetic energy cascade of compressible turbulence provides theoretical basis for LES of compressible turbulence.^{59–62} Wang *et al.* studied the interscale energy transfer in solenoidally forced stationary compressible isotropic turbulence and cascades of temperature and entropy fluctuations in compressible turbulence using a filtering method.^{63,101} Buzzicotti *et al.* systematically assessed the impacts of projectors or filters on the multiscale properties of turbulence at high Reynolds numbers and quantified the robustness of scaling as a function of the filtering type in forced homogeneous and isotropic turbulence with up to 2048³ grid points.⁷

Recently, much effort has been devoted to construct turbulence models, including RANS models and LES models by using data-driven methods and machine learning approaches.^{64–91} Ling *et al.* proposed a neural network architecture with embedded invariance properties for the RANS model.⁷⁶ Xiao *et al.* proposed a physics-informed, Bayesian framework for quantifying and reducing model-form uncertainties in RANS simulations.⁷⁷ Machine learning techniques have been used for developing new SGS models in LES recently. ANN has been employed to develop SGS models for a turbulent channel flow from DNS data, which is similar to the gradient model in *a priori* tests and shows no advantage over the Smagorinsky model in *a posteriori* tests.⁷⁸ Vollant *et al.* modeled the subgrid-scale scalar flux by artificial neural network.⁷⁹ Ma *et al.* established a systematic approach for developing reduced models with memory through a natural analogy between recurrent neural networks and the Mori-Zwanzig formalism, which can construct reduced models with good performance on both short-term prediction and long-term statistical properties for the Kuramoto-Sivashinsky equation and the Navier-Stokes equation.⁸² Maulik *et al.* proposed a data-driven strategy to dynamically assess the validity of a turbulence modeling hypothesis in LES.⁹² However, there are few data-driven LES models dealing with the SGS stress and SGS heat flux with functional modeling strategy in compressible turbulence.⁴⁹

In this paper, we develop an ANNMM in the functional form of mixed model combining the gradient model and the Smagorinsky's eddy viscosity model for predicting the SGS anisotropy stress, SGS energy, and SGS heat flux from DNS data of solenoidally forced stationary compressible isotropic turbulence at a grid resolution of 1024³ and turbulent Mach numbers M_t ranging from 0.40 to 1.0. Detailed comparisons between the ANNMM and DNS using both *a priori* and *a posteriori* approaches are made. In Sec. II, we present a brief description about the governing equations and numerical method. In Sec. III, we introduce the DNS database of compressible turbulence. In Sec. IV, we propose the ANNMM for SGS terms (τ_{ij} and Q_j). In Sec. V, we show both *a priori* and *a posteriori* results of

the ANNMM. In Sec. VI, we provide the summary and conclusion of our study.

II. GOVERNING EQUATIONS AND NUMERICAL METHOD

The dimensionless governing equations in a conservation form for compressible isotropic turbulence of ideal gas can be derived by using a set of reference variables to normalize the hydrodynamic and thermodynamic variables,^{93,94} including the reference length L_f , density ρ_f , velocity U_f , temperature T_f , energy per unit volume $\rho_f U_f^2$, viscosity μ_f , thermal conductivity κ_f , and pressure $p_f = \rho_f U_f^2$. The three reference governing parameters are the reference Reynolds number $Re \equiv \rho_f U_f L_f / \mu_f$, the reference Mach number $M = U_f / c_f$, and the reference Prandtl number $Pr \equiv \mu_f C_p / \kappa_f$. Here, the speed of sound is defined by $c_f \equiv \sqrt{\gamma R T_f}$. $\gamma \equiv C_p / C_v$ is the ratio of specific heat at constant pressure C_p to that at constant volume C_v and is assumed to be equal to 1.4. Moreover, R is the specific gas constant, $R \equiv C_p - C_v$. The parameter Pr is assumed to be equal to 0.7. The parameter α is given by $\alpha \equiv Pr Re(\gamma - 1) M^2$. The dimensionless Navier-Stokes equations for compressible flows are

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho u_j)}{\partial x_j} = 0, \quad (1)$$

$$\frac{\partial(\rho u_i)}{\partial t} + \frac{\partial[\rho u_i u_j + p \delta_{ij}]}{\partial x_j} = \frac{1}{Re} \frac{\partial \sigma_{ij}}{\partial x_j} + \mathcal{F}_i, \quad (2)$$

$$\frac{\partial \mathcal{E}}{\partial t} + \frac{\partial[(\mathcal{E} + p) u_j]}{\partial x_j} = \frac{1}{\alpha} \frac{\partial}{\partial x_j} \left(\kappa \frac{\partial T}{\partial x_j} \right) + \frac{1}{Re} \frac{\partial(\sigma_{ij} u_i)}{\partial x_j} - \Lambda + u_j \mathcal{F}_j, \quad (3)$$

$$p = \rho T / (\gamma M^2), \quad (4)$$

where u_i is the velocity component, p is the pressure, ρ is the density, and T is the temperature. Here, \mathcal{F}_i is a large-scale forcing to the fluid momentum and Λ is a large-scale cooling function per unit volume. The viscous stress is defined by $\sigma_{ij} = 2\mu S_{ij} - \frac{2}{3}\mu \delta_{ij} S_{kk}$, where $S_{ij} = \frac{1}{2}(\partial u_i / \partial x_j + \partial u_j / \partial x_i)$ is the strain rate tensor. Molecular viscosity μ is determined from Sutherland's law, and thermal conductivity κ is then calculated from molecular viscosity using a constant Prandtl number assumption.⁹⁵ The total energy per unit volume \mathcal{E} is defined by $\mathcal{E} = \frac{p}{\gamma - 1} + \frac{1}{2}\rho(u_i u_j)$.

Besides, both the Taylor microscale Reynolds number Re_λ and the turbulent Mach number M_t are defined, respectively, by

$$Re_\lambda = Re \frac{\langle \rho \rangle u^{rms} \lambda}{\sqrt{3} \langle \mu \rangle}, \quad M_t = M \frac{u^{rms}}{\langle \sqrt{T} \rangle}, \quad (5)$$

where $\langle \rangle$ stands for a spatial average. Here, the root mean square (rms) value of the velocity magnitude is defined by $u^{rms} = \sqrt{\langle u_i u_i \rangle}$ and the Taylor microscale is defined by

$$\lambda = \sqrt{\frac{\langle u_i u_i \rangle}{((\partial u_1 / \partial x_1)^2 + (\partial u_2 / \partial x_2)^2 + (\partial u_3 / \partial x_3)^2)}}. \quad (6)$$

The Kolmogorov length scale η and the integral length scale L_I are defined, respectively, as

$$\eta = [\langle \mu / (Re \rho) \rangle^3 / \epsilon]^{1/4}, \quad L_I = \frac{3\pi}{2(u^{rms})^2} \int_0^\infty \frac{E(k)}{k} dk, \quad (7)$$

where ϵ is the spatial average of the dissipation rate per unit mass, $\epsilon = \langle \sigma_{ij} S_{ij} / (Re\rho) \rangle$, and $E(k)$ is the spectrum of kinetic energy per unit mass, namely, $\int_0^\infty E(k) dk = (u^{rms})^2/2$.

The physical variables can be separated into the resolved large-scale space and modeled small scales by a filtering operation $\bar{f}(\mathbf{x}) = \bar{\int}_D f(\mathbf{x}') G(\mathbf{x}, \mathbf{x}'; \Delta) d\mathbf{x}'$, where an overbar denotes a filtered variable, G is the filter kernel, D is the overall domain, and Δ is the filter-width. Favre-filtering (mass-weighted filtering: $\bar{f} = \rho \bar{f} / \bar{\rho}$)⁹⁷ is used to avoid additional SGS terms and obtain the large-scale resolved field in the equation of conservation of mass in compressible flows. The dimensionless compressible filtered Navier-Stokes equations for the resolved variables can be given by^{36,98}

$$\frac{\partial \bar{\rho}}{\partial t} + \frac{\partial (\bar{\rho} \bar{u}_j)}{\partial x_j} = 0, \quad (8)$$

$$\begin{aligned} \frac{\partial (\bar{\rho} \bar{u}_i)}{\partial t} + \frac{\partial (\bar{\rho} \bar{u}_i \bar{u}_j + \bar{\rho} \delta_{ij})}{\partial x_j} - \frac{1}{Re} \frac{\partial \tilde{\sigma}_{ij}}{\partial x_j} - \bar{\mathcal{F}}_i \\ = - \frac{\partial \bar{\rho} \tau_{ij}}{\partial x_j} + \frac{1}{Re} \frac{\partial (\tilde{\sigma}_{ij} - \bar{\sigma}_{ij})}{\partial x_j}, \end{aligned} \quad (9)$$

$$\begin{aligned} \frac{\partial \tilde{\mathcal{E}}}{\partial t} + \frac{\partial [(\tilde{\mathcal{E}} + \bar{\rho}) \bar{u}_j]}{\partial x_j} - \frac{1}{Re} \frac{\partial (\tilde{\sigma}_{ij} \bar{u}_i)}{\partial x_j} - \frac{1}{\alpha} \frac{\partial}{\partial x_j} (\tilde{\kappa} \frac{\tilde{T}}{x_j}) \\ + \bar{\Lambda} - \bar{u}_j \bar{\mathcal{F}}_j = R_{\mathcal{E}}, \end{aligned} \quad (10)$$

$$\begin{aligned} R_{\mathcal{E}} = -\bar{u}_i \frac{\partial (\bar{\rho} \tau_{ij})}{\partial x_j} - \frac{1}{(\gamma - 1) M^2} \frac{\partial \bar{\rho} Q_j}{\partial x_j} - \Pi_{dil} + \frac{1}{Re} \varepsilon_v \\ + \frac{\bar{u}_i}{Re} \frac{\partial (\tilde{\sigma}_{ij} - \bar{\sigma}_{ij})}{\partial x_j} + \frac{1}{\alpha} \frac{\partial}{\partial x_j} (\kappa \frac{\partial \tilde{T}}{\partial x_j} - \tilde{\kappa} \frac{\partial \tilde{T}}{\partial x_j}), \end{aligned} \quad (11)$$

$$\bar{p} = \bar{\rho} \tilde{T} / (\gamma M^2), \quad (12)$$

where the resolved total energy $\tilde{\mathcal{E}}$ is defined by $\tilde{\mathcal{E}} = \frac{\bar{p}}{\gamma - 1} + \frac{1}{2} \bar{\rho} (\bar{u}_i \bar{u}_j)$,^{36,98} 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 $\tilde{S}_{ij} = \frac{1}{2} (\partial \bar{u}_i / \partial x_j + \partial \bar{u}_j / \partial x_i)$, and $\tilde{\mu}$ and $\tilde{\kappa}$ are calculated from Sutherland's law.

The SGS terms appearing on the right-hand sides of Eqs. (8)–(12) are defined as

$$\begin{aligned} \tau_{ij} &= \bar{u}_i \bar{u}_j - \bar{u}_i \bar{u}_j, \quad Q_j = \bar{u}_j \tilde{T} - \bar{u}_j \tilde{T}, \\ \Pi_{dil} &= \bar{p} \bar{S}_{kk} - \bar{p} \tilde{S}_{kk}, \quad \varepsilon_v = \bar{\sigma}_{ji} \bar{S}_{ij} - \bar{\sigma}_{ji} \tilde{S}_{ij}, \end{aligned} \quad (13)$$

where τ_{ij} is the SGS stress, Q_j is the SGS heat flux, Π_{dil} is the SGS pressure-dilatation, and ε_v is the SGS viscous dissipation.

The different SGS terms in the energy equation have been compared for compressible turbulent flows.^{16,30,36} The SGS stress τ_{ij} and the SGS heat flux Q_j are the dominant unclosed terms in the current M_t and Re_λ ranges, while other unclosed terms are significantly smaller than the divergence of the SGS heat flux Q_j and can be neglected in LES.¹⁶ In this paper, we model the SGS force M_i (the SGS stress τ_{ij}) and the divergence of SGS heat flux (Q_j) and neglect other unclosed terms. Meanwhile, it is very reasonable to assume that the kinematic viscosity does not change across different length scales such that $\bar{\sigma}_{ij} = 2\rho v (S_{ij} - \frac{1}{3} \delta_{ij} S_{kk}) = 2\bar{\rho} v (\tilde{S}_{ij} - \frac{1}{3} \delta_{ij} \tilde{S}_{kk})$,¹³ where v is the kinematic viscosity, and the term $\frac{1}{Re} \frac{\partial \bar{\sigma}_{ij} - \bar{\sigma}_{ij}}{\partial x_j}$ would not appear in the filtered momentum equation.

We numerically simulate the compressible homogeneous isotropic turbulence by a hybrid compact-WENO (Weighted Essentially Nonoscillatory) scheme in a cubic box of $(2\pi)^3$ using a uniform grid with periodic boundary conditions.⁹⁵ The hybrid scheme combines an eighth-order compact finite-difference scheme⁹⁹ for smooth regions and a seventh-order WENO scheme¹⁰⁰ for shock regions. We identify the shock front by spatial points with highly negative local dilatation as defined by $\theta < -R_\theta \theta^{rms}$, where R_θ is set to 6.0 for $M_t = 0.4$ and 3.0 for $M_t = 0.6, 0.8, 1.0$. The force is only applied to the solenoidal component of velocity field by fixing the total kinetic energy in the first two wavenumber shells. A uniform thermal cooling Λ is employed to sustain the internal energy in a statistically steady state.⁹⁵

III. DNS DATABASE OF COMPRESSIBLE TURBULENCE

The data used for training and testing are obtained from DNS of a solenoidally forced compressible isotropic turbulence at turbulent Mach numbers M_t ranging from 0.4 to 1.0 and at a uniform grid of 1024^3 . Detailed computational parameters and one-point statistics of four DNS fields are summarized in Table I.¹⁰¹ The Taylor Reynolds number Re_λ is close to 250. The resolution parameter η/dx is in the range $0.98 \leq \eta/dx \leq 1.03$ in numerical simulations, where dx denotes the grid length in each direction. Consequently, the resolution parameter $k_{max}\eta$ is in the range $2.61 \leq k_{max}\eta \leq 3.3$, where the largest wave number k_{max} is half of the number of grids N in each direction: $k_{max} = N/2 = \pi/dx$. In previous works, grid refinement convergence studies were carried out at turbulent Mach number $M_t = 1.0$.^{94,96} The resolutions of $k_{max}\eta \geq 2.61$ are enough for the convergence of small-scale statistics, including the kinetic energy spectrum at different wave numbers and the probability density functions (PDFs) of velocity divergence and vorticity. rms values of velocity divergence and vorticity magnitude are calculated as $\theta^{rms} = \sqrt{\langle \theta^2 \rangle}$ and $\omega^{rms} = \sqrt{\langle \omega_1^2 + \omega_2^2 + \omega_3^2 \rangle}$, respectively.

In the present study, the ANNMM is developed to predict the unclosed SGS stress τ_{ij} and SGS heat flux Q_j using filtered flow fields. Input features, functional input tensors, and the training target of output features (SGS stress and SGS heat flux) of the ANNMM are calculated from DNS data. The functional coefficients mapping input variables to output features are trained using training data and are verified using testing data which are not used in training.

Filtered fields and the corresponding SGS stress and SGS heat flux are obtained using a top-hat filter for its simplicity for both *a priori* and *a posteriori* tests.^{30,101} Filter-width $\Delta = 16\Delta x$ (where Δx is the grid size of DNS) is used in the present study, as shown in Fig. 1, which corresponds to the filter size of $\Delta/\eta \approx 16$ and covers

TABLE I. Simulation parameters and statistical quantities for 1024^3 grid resolutions.

Reso.	M_t	Re_λ	$\eta/\Delta x$	L_1/η	λ/η	θ^{rms}	ω^{rms}	$\theta^{rms}/\omega^{rms}$
1024^3	0.40	259	0.98	235	31.7	0.9	26.1	0.036
1024^3	0.61	259	0.99	232	31.5	2.7	25.9	0.11
1024^3	0.79	261	1.03	229	30.9	4.8	23.6	0.20
1024^3	1.02	248	1.01	226	28.8	8.5	24.0	0.36

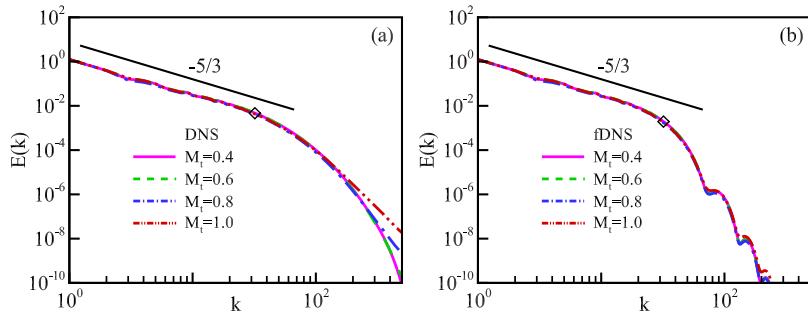


FIG. 1. Velocity spectrum from numerical simulations of solenoidally forced compressible isotropic turbulence, and the diamond represents cutoff $n = 16$: (a) velocity spectrum of DNS and (b) velocity spectrum of filtered-DNS (fDNS).

the inertial range. 5% of the turbulent kinetic energy resides in the SGSs. In *a posteriori* test, LESs are performed on 64^3 ($h_{LES}^\Delta = \Delta$) and 128^3 ($h_{LES}^\Delta = \Delta/2$) grid resolutions with the same filter width Δ for different turbulent Mach numbers. Here, h_{LES}^Δ denotes the grid size of LES. The grid intervals for the filtered fields are in the range $15.24 \leq \Delta/\eta \leq 16.33$ for $Re_\lambda \approx 250$ at different turbulent Mach numbers. Besides, we have performed an *a posteriori* study of the ANNMM applied to LES of compressible turbulence at a larger filter width $\Delta = 32\Delta x$ with a grid resolution of 64^3 ($\Delta = 2h_{LES}^\Delta$).

IV. ARTIFICIAL NEURAL NETWORK

In the context of compressible LES, the functional and structural modeling strategies make use of a Smagorinsky model,^{8,39} a dynamic Smagorinsky model,^{1,40–42} a similarity model,^{43,44} a dynamic mixed model,^{26,28,35,43–45} a gradient model,^{44,46} and a dynamic Clark model.⁴⁷ Verman *et al.* have extended the Clark model⁴⁶ into a dynamic Clark model,^{37,47}

$$\tau_{ij} = C_1 \frac{1}{12} \Delta^2 \partial_k \tilde{u}_i \partial_k \tilde{u}_j + C_2 \Delta^2 |\tilde{S}| \tilde{S}_{ij}, \quad (14)$$

which is very efficient and requires less computational effort than the dynamic mixed model.^{37,47} However, the usual dynamic Clark model needs to use extra filterings to determine the coefficients C_1 and C_2 .

An alternative method to develop a closed mixed model is to learn the coefficients C_1 and C_2 based on the ANN architecture. The ANN's network structure is shown in Fig. 2, which constructs the relation between input features and coefficients C_1 and C_2 .^{76,79} Neurons are the basis of ANN, which are connected by a series of linear

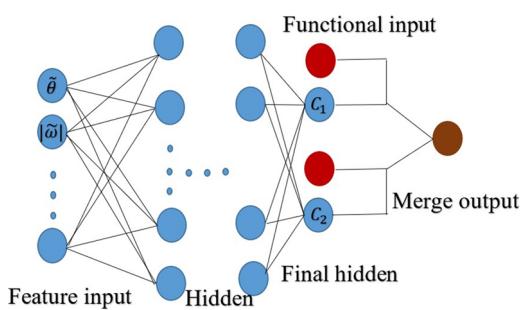


FIG. 2. Schematic diagram of the ANN's network structure.

or nonlinear mathematical operations, receive a signal and process it, and then send the signal to additional neurons connected to them. The data of the first layer X_j^1 are given by the filtered flow fields. Each neuron in a layer l receives and operates inputs X_j^{l-1} from layer $l - 1$ and then establishes outputs X_i^l by the activation function.^{80,102,103} The transfer function is calculated as

$$X_i^l = \sigma(s_i^l + b_i^l), \quad (15)$$

$$s_i^l = \sum_j W_{ij}^l X_j^{l-1}, \quad (16)$$

where σ is the activation function, b_i^l is the bias parameter, and W_{ij}^l is the weight. The bias parameters and the weights are corrected iteratively so that the final hidden layer X^L well approximates the functional coefficients of the real SGS stress and SGS heat flux functional models.

The SGS stress tensor is divided into anisotropic and isotropic parts, respectively, as follows: $\tau_{ij} = \tau_{ij}^A + \tau_{ij}^I$, where $\tau_{ij}^A = \tau_{ij} - \frac{1}{3} \delta_{ij} \tau_{kk}$, $\tau_{ij}^I = -\frac{1}{3} \delta_{ij} \tau_{kk}$. The anisotropic part of the SGS stress tensor τ_{ij}^A , the isotropic part of the SGS stress τ_{ij}^I (SGS kinetic energy τ_{kk}), and the SGS heat flux tensor Q_j can be modeled using the generalization of the Clark model and Smagorinsky model that are separately given by

$$\tau_{ij}^A = \tau_{ij} - 1/3 \delta_{ij} \tau_{kk} = C_1 \tau_{ij}^{A1} + C_2 \tau_{ij}^{A2}, \quad (17)$$

$$\tau_{ij}^I = -\frac{1}{3} \delta_{ij} \tau_{kk}, \quad \tau_{kk} = C_1 \tau_{kk}^1 + C_2 \tau_{kk}^2, \quad (18)$$

$$Q_j = C_1 Q_j^1 + C_2 Q_j^2, \quad (19)$$

where $\tau_{ij}^{A1} = \frac{\Delta^2}{12} \frac{\partial \tilde{u}_i}{\partial x_k} \frac{\partial \tilde{u}_j}{\partial x_k} - \frac{\delta_{ij}}{3} \frac{\Delta^2}{12} \frac{\partial \tilde{u}_k}{\partial x_k} \frac{\partial \tilde{u}_k}{\partial x_k}$, $\tau_{ij}^{A2} = 2\Delta^2 |\tilde{S}| (\tilde{S}_{ij} - \frac{\delta_{ij}}{3} \tilde{S}_{kk})$, $\tau_{kk}^1 = \frac{\Delta^2}{12} \frac{\partial \tilde{u}_k}{\partial x_k} \frac{\partial \tilde{u}_k}{\partial x_k}$, $\tau_{kk}^2 = 2\Delta^2 |\tilde{S}| \tilde{S}_{kk}$, $Q_j^1 = \frac{\Delta^2}{12} \frac{\partial \tilde{u}_j}{\partial x_k} \frac{\partial \tilde{T}}{\partial x_k}$, $Q_j^2 = \Delta^2 |\tilde{S}| \frac{\partial \tilde{T}}{\partial x_j}$.

Three ANNs are trained to predict τ_{ij}^A , τ_{kk} , and Q_j . The choice of proper input variables is important for our ANN architecture. A set of input variables and output variables are shown in Table II for different ANNs. To preserve Galilean invariance, the gradient of filtered velocity instead of filtered velocity is used to model the SGS terms. $|\tilde{\omega}|$ is the filtered vorticity magnitude, $\tilde{\theta}$ is filtered velocity divergence, $\sqrt{\tilde{A}_{ij} \tilde{A}_{ij}}$ is filtered velocity gradient tensor magnitude, $\sqrt{\tilde{S}_{ij} \tilde{S}_{ij}}$ is filtered strain-rate tensor magnitude, and $|\frac{\partial \tilde{T}}{\partial x_i}|$ is filtered temperature gradient magnitude. The velocity gradient tensor is

TABLE II. Set of inputs and outputs for the ANNs.

ANN	Raw inputs	Raw outputs
ANN1	$ \tilde{\omega} , \tilde{\theta}, \sqrt{\tilde{A}_{ij}\tilde{A}_{ij}}, \sqrt{\tilde{S}_{ij}\tilde{S}_{ij}}$	τ_{ij}^A
ANN2	$ \tilde{\omega} , \tilde{\theta}, \sqrt{\tilde{A}_{ij}\tilde{A}_{ij}}, \sqrt{\tilde{S}_{ij}\tilde{S}_{ij}}$	τ_{kk}
ANN3	$ \tilde{\omega} , \tilde{\theta}, \sqrt{\tilde{A}_{ij}\tilde{A}_{ij}}, \sqrt{\tilde{S}_{ij}\tilde{S}_{ij}}, \frac{\partial \tilde{T}}{\partial x_i} $	Q_j

defined by $\tilde{A}_{ij} = \frac{\partial \tilde{u}_i}{\partial x_j}$. There input parameters are used to construct the coefficients C_1 and C_2 .

We have set four layers of neurons ($M:20:20:2$) between the set of inputs and final hidden layer with a tan-sigmoid activation function $\sigma(a) = \frac{1-e^{-2a}}{1+e^{-2a}}$.⁸⁰ M is the number of neurons of the input layer, which depends on the input characters of the SGS unclosed terms as described in Table II, $M = 4$ for ANN1 and ANN2 and $M = 5$ for ANN3. The final hidden layer consists of two neurons, one for each coefficient. The training stage determines the best set of ANN parameters. The back-propagation method is used to minimize the loss function, which is defined as the difference between the merge output and the real SGS anisotropy stress, SGS energy or SGS heat flux ($\langle (\frac{1}{6} \sum_{i=1}^3 \sum_{j=1}^i (\tau_{ij,predicted}^A - \tau_{ij,DNS}^A)^2) \rangle, \langle (\tau_{kk,predicted} - \tau_{kk,DNS})^2 \rangle$, or $\langle (\frac{1}{3} \sum_{j=1}^3 (Q_{j,predicted} - Q_{j,DNS})^2) \rangle$), where $\langle \rangle$ represents the spatial average over homogeneous directions (i.e., the entire domain for isotropic turbulence).⁷⁸

The tan-sigmoid activation function provides outputs between 0 and 1.⁸⁰ Similar to previous data-driven strategies,^{64,77,79,80,91}

all input features $|\tilde{\omega}|, \tilde{\theta}, \sqrt{\tilde{A}_{ij}\tilde{A}_{ij}}, \sqrt{\tilde{S}_{ij}\tilde{S}_{ij}}, |\frac{\partial \tilde{T}}{\partial x_i}|$ are normalized to a range of [0, 1] by min-max scaling, which is defined by $x^* = \frac{x - \min\{x^{DNS}\}}{\max\{x^{DNS}\} - \min\{x^{DNS}\}}$. Meanwhile, τ_{ij}^A , τ_{ij}^{A1} , and τ_{ij}^{A2} are normalized by $x_{ij}^* = \frac{x_{ij} - x_{minmax}}{2x_{minmax}}$, where $x_{minmax} = \frac{1}{6} \sum_{i=1}^3 \sum_{j=1}^i (\max\{x_{ij}^{DNS}\} - \min\{x_{ij}^{DNS}\})/2$, τ_{kk} , τ_{kk}^1 , and τ_{kk}^2 are normalized by $x^* = \frac{x - x_{minmax}}{2x_{minmax}}$, where $x_{minmax} = \max\{x^{DNS}\} - \min\{x^{DNS}\}/2$, Q_j , Q_j^1 , and Q_j^2 are normalized by $x_j^* = \frac{x_j - x_{minmax}}{2x_{minmax}}$, where $x_{minmax} = \frac{1}{3} \sum_{j=1}^3 (\max\{x_j^{DNS}\} - \min\{x_j^{DNS}\})/2$.

Besides, we perform cross-validation to suppress overfitting. The performance of the model is estimated by the data which have not been trained for. In this research, the DNS data are generated using 1024^3 degrees of freedom, while the LES is performed at grid resolutions of 64^3 ($h_{LES}^\Delta = \Delta$) and 128^3 ($h_{LES}^\Delta = \Delta/2$) with the same filter width $\Delta = 16\Delta x$. Therefore, a coarse-graining procedure involves the selection of every sixteenth grid point in each direction in the DNS data, which allow us to devise 16^3 different data sets of 64^3 coarse grid points with the filter width Δ in three dimensions, respectively. For the purpose of training and testing, we randomly choose 6×64^3 grid points for the generation of 70% of training data and 30% of testing data. Finally, the network is trained by the Adam algorithm¹⁰⁴ for 1×10^3 iterations, with the batch size being 1000 for determining the final network until the learning rate is minimal.

In this research, the ANN was trained for a long duration (1000 epochs) until the learning rate is almost minimal when the training and validation loss are correlated closely as shown in Fig. 3. 70% of the total dataset was utilized for training, and the rest was utilized for testing assessment. Figure 3 shows the learning rate of

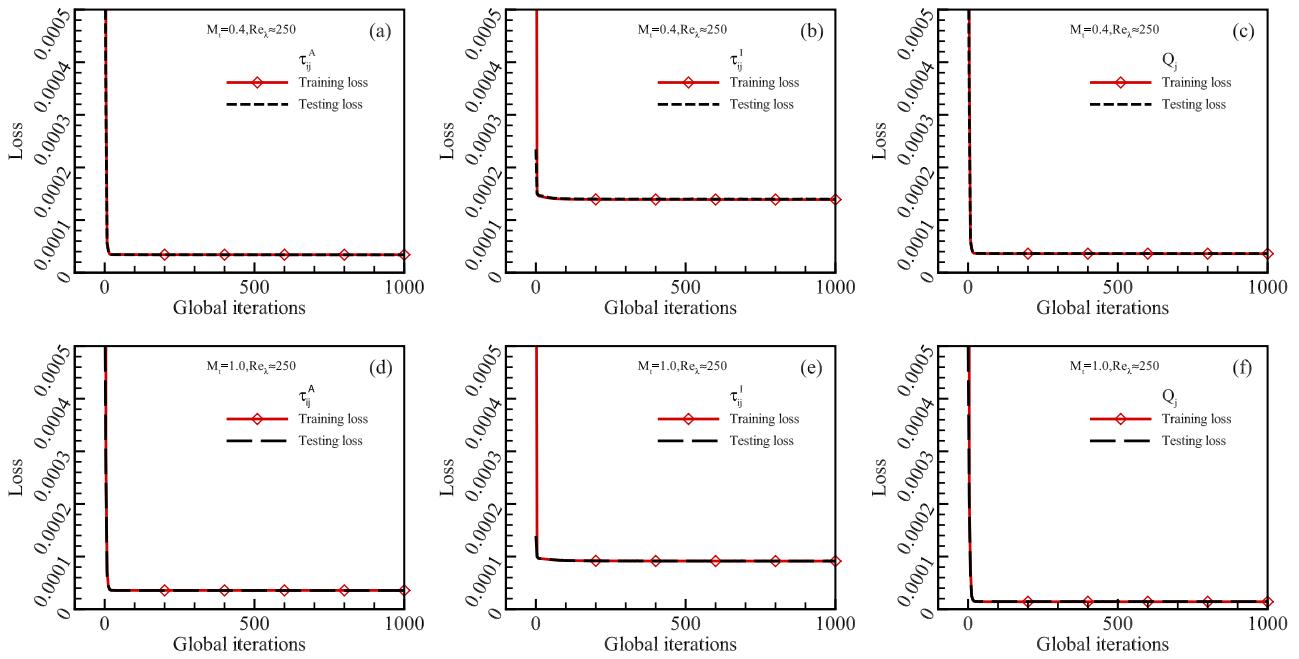


FIG. 3. Learning rate of the proposed ANNM of SGS unclosed terms τ_{ij}^A , τ_{ij}^I , and Q_j , at $M_t = 0.4, 1.0$ and $Re_\lambda \approx 250$: (a) τ_{ij}^A and $M_t = 0.4$, (b) τ_{ij}^I and $M_t = 0.4$, (c) Q_j and $M_t = 0.4$, (d) τ_{ij}^A and $M_t = 1.0$, (e) τ_{ij}^I and $M_t = 1.0$, and (f) Q_j and $M_t = 1.0$.

the proposed framework with very similar behavior between the training and validation loss, implying that the learning process is well-formulated. The final ANN would be selected after a desired number of iterations, which for this study was fixed at 1000.

V. RESULTS

In this section, we outline the results of our proposed framework for the solenoidally forced compressible isotropic turbulence at $M_t = 0.4, 0.6, 0.8$, and 1.0 , which are fully turbulent and various scaling behaviors of spectra of velocity and thermodynamic variables have been observed in the inertial range.^{19–21} The performances of our proposed ANN framework are shown in both *a priori* and *a posteriori* tests. The correlation coefficients, relative errors, and SGS energy flux of the ANNMM are evaluated in *a priori* tests. Furthermore, the large eddy simulations with the ANNMM at grid resolutions of 64^3 ($h_{LES}^\Delta = \Delta$) and 128^3 ($h_{LES}^\Delta = \Delta/2$) with the same filter width $\Delta = 16\Delta x$ are performed and compared with the results from the filtered fields of the DNS database. Moreover, the large-eddy simulation with the ANNMM at a larger filter width $\Delta = 32\Delta x$ and a grid resolution of 64^3 ($h_{LES}^\Delta = \Delta/2$) is performed and compared with the results from the filtered fields of the DNS database. The ANNMM gives reliable predictions on the statistics and structures of filtered DNS data in *a priori* and *a posteriori* tests.

A. *A priori* tests

The correlation coefficient $C(H)$, the relative error $E_r(H)$, and the root-mean-square value $D(H)$ of the real unclosed term H and the modeled term H^{model} are compared to evaluate the accuracy of the ANNMM. Besides, we consider the gradient model:^{43,44} $\tau_{ij} = \frac{\Delta^2}{12} \alpha_i, \alpha_i = \frac{\partial \bar{u}_i}{\partial x_k} \frac{\partial \bar{u}_j}{\partial x_k}, Q_j = \frac{\Delta^2}{12} \alpha_T, \alpha_T = \frac{\partial \bar{u}_j}{\partial x_k} \frac{\partial \bar{T}}{\partial x_k}$, which is abbreviated as VGM. $C(H)$, $E_r(H)$, and $D(H)$ are calculated, respectively, by

$$C(H) = \frac{\langle (H - \langle H \rangle)(H^{model} - \langle H^{model} \rangle) \rangle}{\langle (H - \langle H \rangle)^2 \rangle \langle (H^{model} - \langle H^{model} \rangle)^2 \rangle^{1/2}}, \quad (20)$$

$$E_r(H) = \frac{\sqrt{\langle (H - H^{model})^2 \rangle}}{\sqrt{\langle H^2 \rangle}}, \quad (21)$$

$$D(H) = \sqrt{\langle (H - \langle H \rangle)^2 \rangle}, \quad (22)$$

where $\langle \cdot \rangle$ denotes averaging over the volume.

Table III shows correlation coefficients, relative errors, and root mean square values of τ_{kk} for both training and testing sets after the training process of ANN. The difference between the results of training and testing sets is always very small, indicating that there is no problem of overfitting in the training process.

Tables IV–VI show correlation coefficients, relative errors, and root mean square values of τ_{11} , τ_{12} , and Q_1 for different models in the testing set at $M_t = 0.4, 0.6, 0.8$, and 1.0 . The correlation coefficients are about 0.94 for VGM and ANNMM. It is seen that the ANNMM is successful since the averaged correlation coefficient C exceeds 0.94 in the *a priori* test.

In order to analyze the impact of hyperparameter of the network architecture, the ANN is trained with $M:40:40:2$ and $M:20:20:20:2$ between the set of input and final hidden layer, which are abbreviated as ANNMM-M1 and ANNMM-M2, respectively.

TABLE III. Correlation coefficient (C), relative error (E_r), and root mean square value (D) of τ_{kk} for training and testing sets with the ANNMM at $M_t = 0.4, 0.6, 0.8$, and 1.0 .

$C \setminus M_t$	0.4	0.6	0.8	1.0
Training	0.945	0.948	0.951	0.949
Testing	0.945	0.947	0.951	0.949
$E_r \setminus M_t$	0.4	0.6	0.8	1.0
Training	0.220	0.215	0.209	0.211
Testing	0.220	0.217	0.210	0.212
$D \setminus M_t$	0.4	0.6	0.8	1.0
DNS	0.256	0.279	0.241	0.233
Training	0.240	0.262	0.229	0.222
Testing	0.241	0.262	0.230	0.221

The *a priori* results of the ANNMM, ANNMM-M1, and ANNMM-M2 are summarized in **Tables IV–IX**. ANNMM, ANNMM-M1 and ANNMM-M2 show similar correlation coefficients, relative errors, and root mean squares of τ_{11} , τ_{12} , Q_1 , τ_{11}^2 , τ_{12}^2 , Q_1^2 at different M_t , indicating that the learning accuracy does not increase with the increase in the number of neurons and hidden layers.

In this work, the coefficient outputs of τ_{ij}^A , τ_{ij}^I , and Q_j are trained separately. It is important to train all 6 coefficient outputs simultaneously with some physical constraints to ensure the physical realizability of τ_{ij}^A , τ_{ij}^I , and Q_j for future ANN models. This is a very preliminary result for training coefficient outputs where no physical constraint is added to the ANN training. We will focus on the ANN models with some physical constraints in the next step.

TABLE IV. Correlation coefficient (C), relative error (E_r), and root mean square value (D) of τ_{11} for different models at $M_t = 0.4, 0.6, 0.8$, and 1.0 .

$C \setminus M_t$	0.4	0.6	0.8	1.0
VGM	0.934	0.938	0.939	0.941
ANNMM	0.937	0.943	0.943	0.944
ANNMM-M1	0.939	0.943	0.943	0.945
ANNMM-M2	0.937	0.943	0.943	0.945
$E_r \setminus M_t$	0.4	0.6	0.8	1.0
VGM	0.434	0.414	0.408	0.399
ANNMM	0.273	0.259	0.260	0.281
ANNMM-M1	0.272	0.259	0.261	0.302
ANNMM-M2	0.288	0.262	0.261	0.287
$D \setminus M_t$	0.4	0.6	0.8	1.0
DNS	0.124	0.117	9.35×10^{-2}	0.109
VGM	8.25×10^{-2}	8.06×10^{-2}	6.54×10^{-2}	7.80×10^{-2}
ANNMM	0.123	0.104	8.35×10^{-2}	9.08×10^{-2}
ANNMM-M1	0.122	0.104	8.33×10^{-2}	8.88×10^{-2}
ANNMM-M2	0.123	0.104	8.31×10^{-2}	8.94×10^{-2}

TABLE V. Correlation coefficient (C), relative error (E_r), and root mean square value (D) of τ_{12} for different models at $M_t = 0.4, 0.6, 0.8$, and 1.0 .

$C \setminus M_t$	0.4	0.6	0.8	1.0
VGM	0.939	0.941	0.944	0.945
ANNMM	0.938	0.942	0.945	0.940
ANNMM-M1	0.939	0.942	0.945	0.944
ANNMM-M2	0.939	0.942	0.945	0.945
$E_r \setminus M_t$	0.4	0.6	0.8	1.0
VGM	0.391	0.382	0.372	0.372
ANNMM	0.389	0.340	0.329	0.370
ANNMM-M1	0.384	0.340	0.329	0.343
ANNMM-M2	0.363	0.337	0.329	0.351
$D \setminus M_t$	0.4	0.6	0.8	1.0
DNS	6.87×10^{-2}	6.59×10^{-2}	5.61×10^{-2}	6.15×10^{-2}
VGM	5.17×10^{-2}	5.02×10^{-2}	4.31×10^{-2}	4.72×10^{-2}
ANNMM	6.82×10^{-2}	6.05×10^{-2}	5.15×10^{-2}	5.23×10^{-2}
ANNMM-M1	6.82×10^{-2}	6.04×10^{-2}	5.15×10^{-2}	5.20×10^{-2}
ANNMM-M2	6.85×10^{-2}	6.02×10^{-2}	5.15×10^{-2}	5.20×10^{-2}

The relative error of the ANNMM is smaller than the relative error of the VGM for τ_{11} . Thus, the ANNMM performs better than the gradient model on the relative error in the *a priori* test. The root mean square values of SGS terms predicted by the VGM reach 70% of those of DNS. The ANNMM predicts the rms value of SGS terms much closer to those of DNS data in the *a priori* test.

TABLE VI. Correlation coefficient (C), relative error (E_r), and root mean square value (D) of Q_1 for different models at $M_t = 0.4, 0.6, 0.8$, and 1.0 .

$C \setminus M_t$	0.4	0.6	0.8	1.0
VGM	0.930	0.940	0.931	0.927
ANNMM	0.931	0.936	0.928	0.926
ANNMM-M1	0.931	0.936	0.929	0.926
ANNMM-M2	0.929	0.937	0.929	0.926
$E_r \setminus M_t$	0.4	0.6	0.8	1.0
VGM	0.395	0.370	0.409	0.443
ANNMM	0.376	0.353	0.400	0.436
ANNMM-M1	0.373	0.353	0.395	0.437
ANNMM-M2	0.369	0.354	0.401	0.453
$D \setminus M_t$	0.4	0.6	0.8	1.0
DNS	7.74×10^{-4}	1.80×10^{-3}	3.04×10^{-3}	8.46×10^{-3}
VGM	6.05×10^{-4}	1.43×10^{-3}	2.26×10^{-3}	5.85×10^{-3}
ANNMM	7.18×10^{-4}	1.65×10^{-3}	2.57×10^{-3}	5.98×10^{-3}
ANNMM-M1	7.18×10^{-4}	1.65×10^{-3}	2.57×10^{-3}	5.99×10^{-3}
ANNMM-M2	7.19×10^{-4}	1.65×10^{-3}	2.56×10^{-3}	5.94×10^{-3}

ANNMM, ANNMM-M1, and ANNMM-M2 show similar correlation coefficients, relative errors, and root mean squares of $\tau_{11}^2, \tau_{12}^2, Q_1^2$ at different M_t in Tables VII–IX, indicating that the ANNMM predicts $\tau_{11}^2, \tau_{12}^2, Q_1^2$ much closer to those of DNS data in the *a priori* test, as compared to the gradient model. It is reasonable to train τ_{ij}^A, τ_{ij}^I , and Q_j with the ANNMM.

Table X shows average adaptive functional model coefficients of τ_{ij}^A, τ_{ij}^I , and Q_j for different models at $M_t = 0.4, 0.6, 0.8$, and 1.0 . We can see that average coefficients C_1^A, C_1^I , and C_1^Q are above 1.0 for the ANNMM.

Besides, it is important to understand the physical mechanism of interscale energy transfer in order to improve the SGS models in LES of turbulence for avoiding numerically unstable problem or excessive SGS dissipation. The interscale transfer of kinetic energy in compressible isotropic turbulence was analyzed previously,¹⁰¹ showing that kinetic energy transferred from large scales to smaller scales, and dissipated finally by viscosity close to the Kolmogorov length scale. Meanwhile, there is backscatter of energy identified as negative values of the subgrid-scale (SGS) flux in the filtering velocity of turbulence. We focus on the performance of the VGM and the ANNMM on SGS flux of kinetic energy $\Pi = -\bar{\rho} \tau_{ij} \bar{S}_{ij}$.

We analyze the effect of compression and expansion motions on the SGS kinetic energy transfer for different models. We plot the joint PDF between the exact and modeled normalized SGS kinetic energy flux for the VGM and the ANNMM at turbulent Mach numbers $M_t = 0.4, 0.6, 0.8$, and 1.0 in Fig. 4. The correlation between the ANNMM and the exact SGS term is better than that between the VGM and the exact SGS term.

The comparison of the PDFs of the normalized SGS flux Π^T/ϵ_T at turbulent Mach numbers $M_t = 0.4, 0.6, 0.8$, and 1.0 is shown in Fig. 5. As M_t increases, the predicted PDF of SGS flux by the VGM and the ANNMM skews toward the positive side, indicating that

TABLE VII. Correlation coefficient (C), relative error (E_r), and root mean square value (D) of τ_{11}^2 for different models at $M_t = 0.4, 0.6, 0.8$, and 1.0 .

$C \setminus M_t$	0.4	0.6	0.8	1.0
VGM	0.908	0.913	0.926	0.916
ANNMM	0.910	0.924	0.931	0.922
ANNMM-M1	0.910	0.923	0.931	0.923
ANNMM-M2	0.910	0.924	0.931	0.925
$E_r \setminus M_t$	0.4	0.6	0.8	1.0
VGM	0.638	0.580	0.566	0.523
ANNMM	0.405	0.392	0.378	0.415
ANNMM-M1	0.405	0.394	0.380	0.430
ANNMM-M2	0.408	0.396	0.381	0.420
$D \setminus M_t$	0.4	0.6	0.8	1.0
DNS	0.127	9.99×10^{-2}	6.20×10^{-2}	9.01×10^{-2}
VGM	5.39×10^{-2}	4.99×10^{-2}	3.13×10^{-2}	5.24×10^{-2}
ANNMM	0.117	8.08×10^{-2}	4.96×10^{-2}	6.76×10^{-2}
ANNMM-M1	0.117	8.04×10^{-2}	4.93×10^{-2}	6.51×10^{-2}
ANNMM-M2	0.121	7.95×10^{-2}	4.91×10^{-2}	6.60×10^{-2}

TABLE VIII. Correlation coefficient (C), relative error (E_r), and root mean square value (D) of τ_{12}^2 for different models at $M_t = 0.4, 0.6, 0.8$, and 1.0 .

$C \setminus M_t$	0.4	0.6	0.8	1.0
VGM	0.923	0.922	0.933	0.926
ANNMM	0.923	0.925	0.932	0.917
ANNMM-M1	0.923	0.925	0.932	0.925
ANNMM-M2	0.924	0.927	0.932	0.924
$E_r \setminus M_t$	0.4	0.6	0.8	1.0
VGM	0.544	0.494	0.487	0.474
ANNMM	0.379	0.377	0.362	0.417
ANNMM-M1	0.378	0.377	0.362	0.404
ANNMM-M2	0.377	0.375	0.362	0.407
$D \setminus M_t$	0.4	0.6	0.8	1.0
DNS	3.19×10^{-2}	2.27×10^{-2}	1.60×10^{-2}	1.86×10^{-2}
VGM	1.71×10^{-2}	1.38×10^{-2}	9.64×10^{-3}	1.18×10^{-2}
ANNMM	3.02×10^{-2}	1.98×10^{-2}	1.38×10^{-2}	1.45×10^{-2}
ANNMM-M1	3.02×10^{-2}	1.99×10^{-2}	1.38×10^{-2}	1.45×10^{-2}
ANNMM-M2	3.04×10^{-2}	1.96×10^{-2}	1.38×10^{-2}	1.44×10^{-2}

kinetic energy transfer from filtered scales to smaller scales. Both models predict that the compressibility can enhance the local SGS energy transfer from large scales to small scales at higher M_t on the positive tail of the SGS flux and recover the SGS backscatter of kinetic energy well on the negative tail of the SGS flux. We plot the average of normalized SGS flux Π/ϵ_T conditioned on the normalized filtered velocity divergence θ/θ^{rms} at turbulent Mach numbers

TABLE IX. Correlation coefficient (C), relative error (E_r), and root mean square value (D) of Q_1^2 for different models at $M_t = 0.4, 0.6, 0.8$, and 1.0 .

$C \setminus M_t$	0.4	0.6	0.8	1.0
VGM	0.873	0.890	0.879	0.833
ANNMM	0.877	0.898	0.884	0.833
ANNMM-M1	0.875	0.898	0.884	0.834
ANNMM-M2	0.878	0.898	0.884	0.835
$E_r \setminus M_t$	0.4	0.6	0.8	1.0
VGM	0.551	0.509	0.685	0.619
ANNMM	0.478	0.436	0.576	0.605
ANNMM-M1	0.480	0.435	0.576	0.604
ANNMM-M2	0.475	0.436	0.575	0.618
$D \setminus M_t$	0.4	0.6	0.8	1.0
DNS	3.22×10^{-6}	1.44×10^{-5}	9.85×10^{-5}	1.11×10^{-3}
VGM	1.95×10^{-6}	9.38×10^{-6}	3.78×10^{-5}	6.13×10^{-4}
ANNMM	2.59×10^{-6}	1.18×10^{-5}	5.35×10^{-5}	6.50×10^{-4}
ANNMM-M1	2.59×10^{-6}	1.18×10^{-5}	5.34×10^{-5}	6.49×10^{-4}
ANNMM-M2	2.63×10^{-6}	1.18×10^{-5}	5.35×10^{-5}	6.12×10^{-4}

TABLE X. Average functional model coefficients of τ_{ij}^A , τ_{ij}^I , and Q_j for the VGM and the ANNMM at $M_t = 0.4, 0.6, 0.8$, and 1.0 .

$C_1^A \setminus M_t$	0.4	0.6	0.8	1.0
VGM	1.0	1.0	1.0	1.0
ANNMM	1.30	1.20	1.19	1.10
$C_2^A \setminus M_t$	0.4	0.6	0.8	1.0
VGM	0.0	0.0	0.0	0.0
ANNMM	2.75×10^{-3}	8.79×10^{-4}	3.81×10^{-4}	2.47×10^{-4}
$C_1^I \setminus M_t$	0.4	0.6	0.8	1.0
VGM	1.0	1.0	1.0	1.0
ANNMM	1.36	1.06	1.10	1.01
$C_2^I \setminus M_t$	0.4	0.6	0.8	1.0
VGM	0.0	0.0	0.0	0.0
ANNMM	2.69×10^{-2}	2.53×10^{-2}	2.07×10^{-2}	1.24×10^{-2}
$C_1^Q \setminus M_t$	0.4	0.6	0.8	1.0
VGM	1.0	1.0	1.0	1.0
ANNMM	1.19	1.16	1.12	1.02
$C_2^Q \setminus M_t$	0.4	0.6	0.8	1.0
VGM	0.0	0.0	0.0	0.0
ANNMM	1.08×10^{-4}	4.53×10^{-3}	4.85×10^{-3}	-3.75×10^{-5}

$M_t = 0.4, 0.6, 0.8$, and 1.0 in Fig. 6. As M_t increases, strong compression motions induce strong direct SGS flux of kinetic energy, while strong expansion motions induce strong reverse SGS flux of kinetic energy. The conditional average of SGS flux of kinetic energy can be recovered by the ANNMM accurately.

Overall, in *a priori* analysis, the ANNMM is a successful model, which gives the high correlation coefficient, and recovers the PDFs of the SGS kinetic energy flux accurately. Thus, the proposed ANNMM includes more useful information of filtered fields and performs substantially better than the gradient model in *a priori* test.

B. A posteriori tests

We investigate and validate the performances of the trained ANNMM for $M_t = 0.4, 0.6, 0.8$, and 1.0 at grid resolutions of 64^3 ($h_{LES} = \Delta$) and 128^3 ($h_{LES} = \Delta/2$) with the same filter width $\Delta = 16\Delta x$ in LES in *a posteriori* test,^{105–107} which is the most proper way to evaluate the performances of LES models. Furthermore, we present an *a posteriori* study of ANNMM applied to LES of compressible turbulence at a larger filter width $\Delta = 32\Delta x$ with a grid resolution of 64^3 ($\Delta = 2h_{LES}^A$). Simulations without any LES model (no-model) and using DSM and DMM are also carried out for comparison.

The filter width Δ is crucial for modeling the SGS unclosed terms, which usually is proportional to the grid size h_{LES}^A ,^{108,109} i.e.,

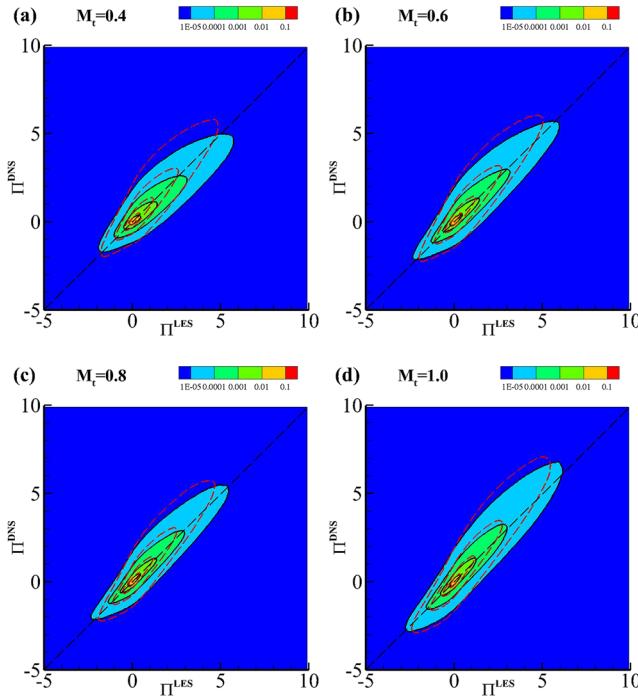


FIG. 4. Joint probability density function (J-PDF) between the exact and modeled normalized SGS transfer terms, for the VGM (dashed red lines) and the ANNMM (solid black lines) in *a priori* analysis. The isocontours are in the range 10^{-5} – 10^{-1} with a logarithm scale: (a) $M_t = 0.4$, (b) $M_t = 0.6$, (c) $M_t = 0.8$, and (d) $M_t = 1.0$.

the filter grid ratio (FGR) ($FGR = \Delta/h_{LES}^\Delta$) is typically chosen to be equal to 1 or 2.^{105–108,110} In *a posteriori* tests of LES with 64^3 and 128^3 points, we implement the ANNMM with the same weight W_{ij}^l and bias parameter b_i^l in the four layers of neurons $M:20:20:2$ between

the set of input and final hidden layers. The numerical errors are relatively small at a grid resolution of 128^3 , giving rise to the fact that the differences between LES results and filtered DNS results are mainly due to the LES models at a grid resolution of 128^3 .

The time steps in LES and DNS are different: $\Delta t_{LES}^\Delta = 5\Delta t_{DNS}$. Here, Δt_{LES}^Δ and Δt_{DNS} are the time steps for LES and DNS, respectively. We neglect the temporal closure term and focus on the spatial closure only for the time step of a practical LES is small, by assuming that the time discretization is sufficiently accurate and the filter commutes with the time derivative.

The DSM reproduces the global exchange of energy between the resolved and unresolved stress by mimicking the drain of energy associated with the turbulence energy cascade. The SGS stress can be written as follows:^{1,8,30,42}

$$\tau_{ij} - \frac{\delta_{ij}}{3}\tau_{kk} = -2C_s^2\Delta^2|\tilde{S}| \left(\tilde{S}_{ij} - \frac{\delta_{ij}}{3}\tilde{S}_{kk} \right), \quad (23)$$

$$\tau_{kk} = 2C_l\Delta^2|\tilde{S}|^2, \quad (24)$$

$$Q_j = -\frac{v_T}{Pr_T} \frac{\partial \tilde{T}}{\partial x_j} = -C \frac{\Delta^2 |\tilde{S}|}{Pr_T} \frac{\partial \tilde{T}}{\partial x_j}, \quad (25)$$

where Δ is the characteristic SGS length scale, which is equal to the LES grid interval. C_s^2 , C_l , and C/Pr_T can be calculated dynamically,

$$C_s^2 = \frac{\langle \mathcal{L}_{ij}\mathcal{M}_{ij} \rangle}{\langle \mathcal{M}_{kl}\mathcal{M}_{kl} \rangle}, \quad C_l = \frac{\langle \mathcal{L}_{kk} \rangle}{\langle \beta - \alpha \rangle}, \quad \frac{C}{Pr_T} = \frac{\langle \mathcal{K}_j \mathcal{T}_j \rangle}{\langle \mathcal{T}_k \mathcal{T}_k \rangle}, \quad (26)$$

where $\mathcal{L}_{ij} = (\widehat{\rho u_i} \widehat{\rho u_j} / \widehat{\rho}) - \widehat{\rho u_i} \widehat{\rho u_j} / \widehat{\rho}$, $\alpha_{ij} = -2\Delta^2 \widehat{\rho} |\tilde{S}| (\tilde{S}_{ij} - \frac{\delta_{ij}}{3} \tilde{S}_{kk})$, $\beta_{ij} = -2\widehat{\Delta}^2 \widehat{\rho} |\tilde{S}| (\widehat{\tilde{S}_{ij}} - \delta_{ij} \widehat{\tilde{S}_{kk}}/3)$, $\mathcal{M}_{ij} = \beta_{ij} - \alpha_{ij}$, $\alpha = 2\widehat{\rho} \Delta^2 |\tilde{S}|^2$, and $\beta = 2\widehat{\Delta}^2 \widehat{\rho} |\tilde{S}|^2$. $\mathcal{T}_j = -\widehat{\Delta}^2 \widehat{\rho} |\tilde{S}| \frac{\partial \tilde{T}}{\partial x_j} + \Delta^2 \widehat{\rho} |\tilde{S}| \frac{\partial \tilde{T}}{\partial x_j}$, $\mathcal{K}_j = (\widehat{\frac{\rho u_j \rho T}{\widehat{\rho}}} - \widehat{\frac{\rho u_j \rho T}{\widehat{\rho}}})$. An overbar denotes the filter at grid scale dx , a hat represents a test filter coarser than the grid filter, and a hat over the overbar denotes a filter at scale $\widehat{\Delta} = 2\Delta$.

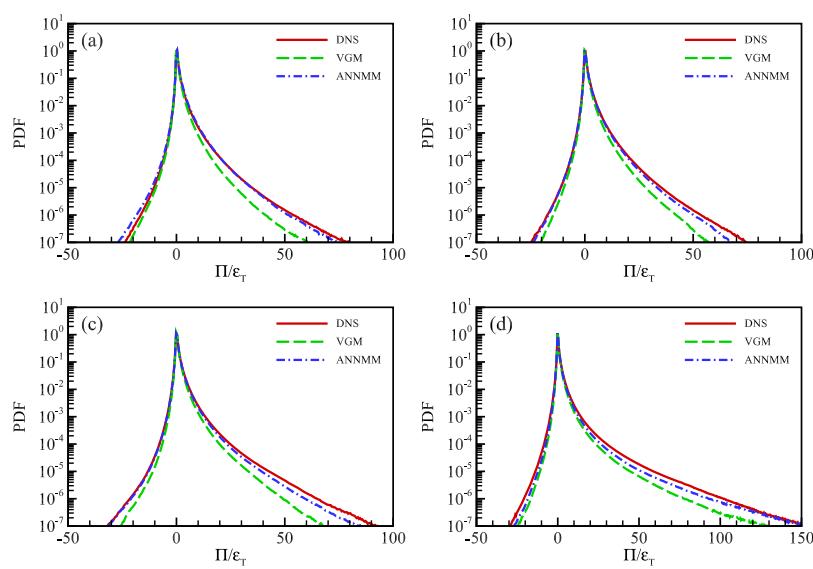


FIG. 5. PDF of the normalized SGS flux Π/ϵ_T for the VGM and the ANNMM in *a priori* analysis: (a) $M_t = 0.4$, (b) $M_t = 0.6$, (c) $M_t = 0.8$, and (d) $M_t = 1.0$.

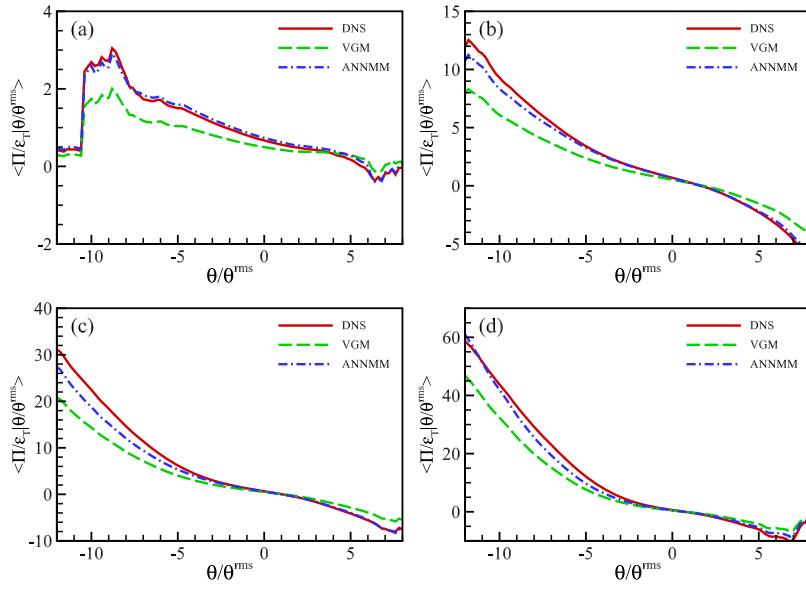


FIG. 6. Average of the normalized SGS flux Π/ε_T conditioned on the normalized filtered velocity divergence θ/θ' for the VGM and the ANNMM in *a priori* analysis: (a) $M_t = 0.4$, (b) $M_t = 0.6$, (c) $M_t = 0.8$, and (d) $M_t = 1.0$.

DMM is a combination of a scale-similarity part and an eddy-viscosity part.^{26,43} We consider the scale similarity part as $(\widehat{\tilde{u}_i \tilde{u}_j} - \widehat{\tilde{u}_i \tilde{u}_j})$.^{44,45} A specific form is prescribed with several undetermined coefficients that should be obtained following a dynamic procedure based on the Germano identity,⁴¹

$$L_{ij} = T_{ij} - \widehat{\rho \tau_{ij}}, \quad (27)$$

where $T_{ij} = \widehat{\rho}(\widehat{\tilde{u}_i \tilde{u}_j} - \widehat{\tilde{u}_i \tilde{u}_j})$ is the stress at scale Δ and $L_{ij} = \widehat{\rho}(\widehat{\tilde{u}_i \tilde{u}_j} - \widehat{\tilde{u}_i \tilde{u}_j})$ is the resolved stress tensor. Replacing τ_{ij} and T_{ij} with the model stresses τ_{ij}^{mod} and T_{ij}^{mod} induces a square error given by

$$E_{mod} = \langle (L_{ij} - L_{ij}^{mod})^2 \rangle, \quad (28)$$

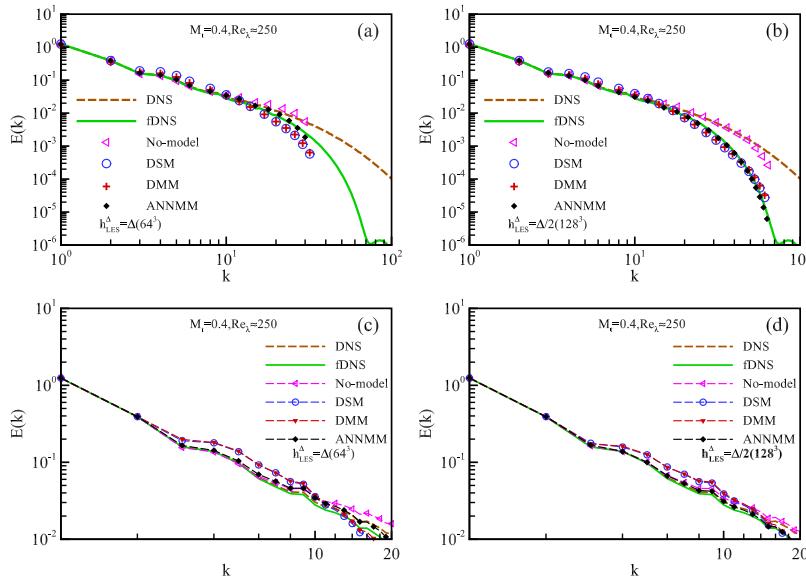


FIG. 7. Spectrum of velocity at $M_t = 0.4$ for LES at grid resolutions of 64^3 ($h_L^\Delta = \Delta$) and 128^3 ($h_L^\Delta = \Delta/2$) with the same filter width $\Delta = 16\Delta_x$: (a) spectrum of velocity at 64^3 ($h_L^\Delta = \Delta$), (b) spectrum of velocity at 128^3 ($h_L^\Delta = \Delta/2$), (c) spectrum of velocity at $k \leq 20$ and 64^3 ($h_L^\Delta = \Delta$), and (d) spectrum of velocity at $k \leq 20$ and 128^3 ($h_L^\Delta = \Delta/2$).

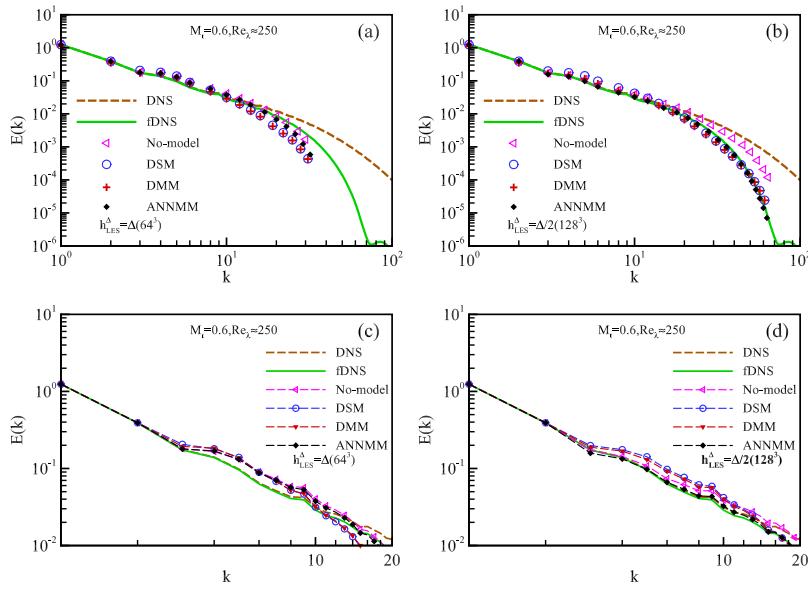


FIG. 8. Spectrum of velocity at $M_t = 0.6$ for LES at grid resolutions of 64^3 ($h_{LES}^\Delta = \Delta$) and 128^3 ($h_{LES}^\Delta = \Delta/2$) with the same filter width $\Delta = 16\Delta x$: (a) spectrum of velocity at 64^3 ($h_{LES}^\Delta = \Delta$), (b) spectrum of velocity at 128^3 ($h_{LES}^\Delta = \Delta/2$), (c) spectrum of velocity at $k \leq 20$ and 64^3 ($h_{LES}^\Delta = \Delta$), and (d) spectrum of velocity at $k \leq 20$ and 128^3 ($h_{LES}^\Delta = \Delta/2$).

model coefficients to be determined by minimizing Eq. (28),

$$C_1 = \frac{\langle M_{ij}^2 \rangle \langle L_{ij} N_{ij} \rangle - \langle M_{ij} N_{ij} \rangle \langle L_{ij} M_{ij} \rangle}{\langle N_{ij}^2 \rangle \langle M_{ij}^2 \rangle - \langle M_{ij} N_{ij} \rangle^2}, \quad (31)$$

$$C_2 = \frac{\langle N_{ij}^2 \rangle \langle L_{ij} M_{ij} \rangle - \langle M_{ij} N_{ij} \rangle \langle L_{ij} N_{ij} \rangle}{\langle N_{ij}^2 \rangle \langle M_{ij}^2 \rangle - \langle M_{ij} N_{ij} \rangle^2}, \quad (32)$$

where $L_{ij} = \hat{\rho}(\tilde{u}_i \tilde{u}_j - \hat{u}_i \hat{u}_j)$, $M_{ij} = H_{1,ij} - \hat{\rho}h_{1,ij}$, and $N_{ij} = H_{2,ij} - \hat{\rho}h_{2,ij}$.

Meanwhile, the SGS heat flux Q_j is modeled as^{30,45}

$$Q_j = C_{q1} h_{q1,j} + C_{q2} h_{q2,j}, \quad (33)$$

$$Q_j = C_{q1} H_{q1,j} + C_{q2} H_{q2,j}, \quad (34)$$

where $Q_j = \hat{\rho}(\widehat{u_j T} - \hat{u}_j \hat{T})$ is the heat flux at scale $\hat{\Delta}$. The coefficients are computed as

$$C_{q1} = \frac{\langle T_j^2 \rangle \langle L_j V_j \rangle - \langle T_j V_j \rangle \langle L_j T_j \rangle}{\langle V_j^2 \rangle \langle T_j^2 \rangle - \langle T_j V_j \rangle^2}, \quad (35)$$

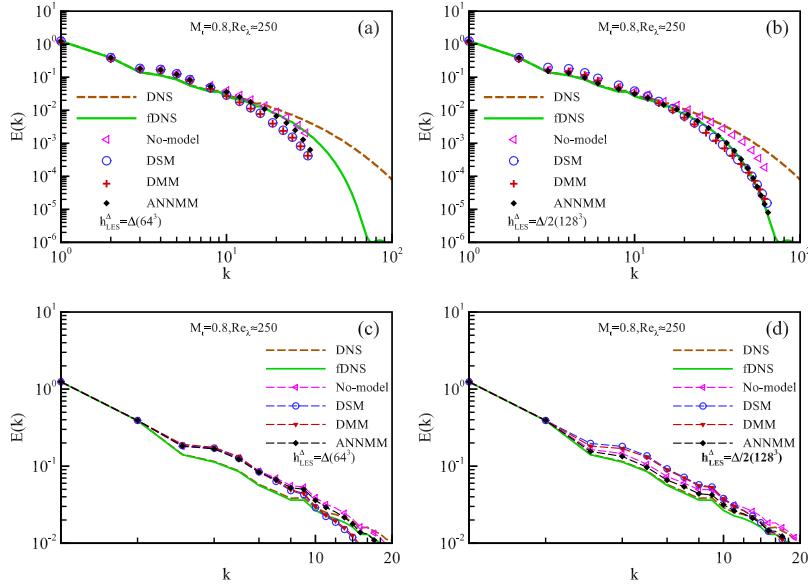


FIG. 9. Spectrum of velocity at $M_t = 0.8$ for LES at grid resolutions of 64^3 ($h_{LES}^\Delta = \Delta$) and 128^3 ($h_{LES}^\Delta = \Delta/2$) with the same filter width $\Delta = 16\Delta x$: (a) spectrum of velocity at 64^3 ($h_{LES}^\Delta = \Delta$), (b) spectrum of velocity at 128^3 ($h_{LES}^\Delta = \Delta/2$), (c) spectrum of velocity at $k \leq 20$ and 64^3 ($h_{LES}^\Delta = \Delta$), and (d) spectrum of velocity at $k \leq 20$ and 128^3 ($h_{LES}^\Delta = \Delta/2$).

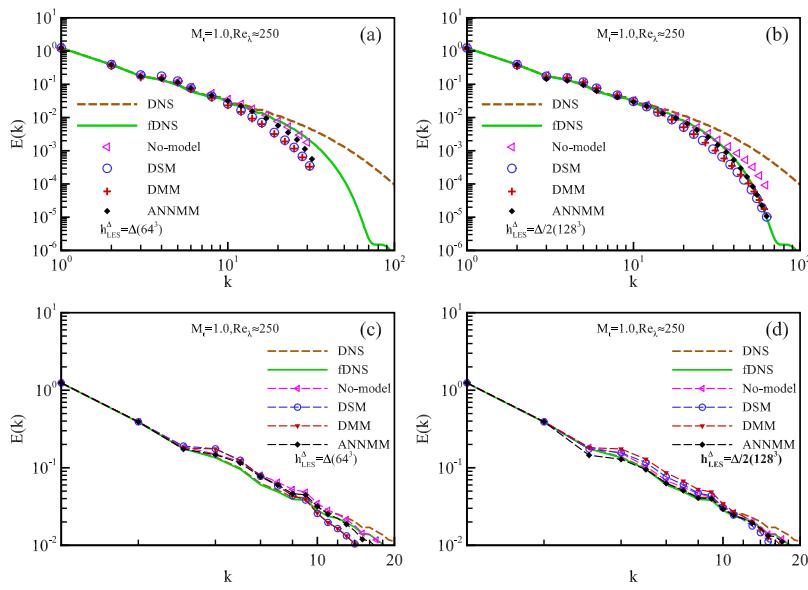


FIG. 10. Spectrum of velocity at $M_t = 1.0$ for LES at grid resolutions of 64^3 ($h_{LES}^\Delta = \Delta$) and 128^3 ($h_{LES}^\Delta = \Delta/2$) with the same filter width $\Delta = 16\Delta x$: (a) spectrum of velocity at 64^3 ($h_{LES}^\Delta = \Delta$), (b) spectrum of velocity at 128^3 ($h_{LES}^\Delta = \Delta/2$), (c) spectrum of velocity at $k \leq 20$ and 64^3 ($h_{LES}^\Delta = \Delta/2$), and (d) spectrum of velocity at $k \leq 20$ and 128^3 ($h_{LES}^\Delta = \Delta/2$).

$$C_{q2} = \frac{\langle V_j^2 \rangle \langle L_j T_j \rangle - \langle T_j V_j \rangle \langle L_j V_j \rangle}{\langle V_j^2 \rangle \langle T_j^2 \rangle - \langle T_j v_j \rangle^2}, \quad (36)$$

where $L_j = \hat{\rho}(\widehat{u_j \tilde{T}} - \widehat{\tilde{u}_j \tilde{T}})$, $h_{q1,j} = -\Delta^2 |\widehat{S}_{ij}| \frac{\partial \widehat{T}}{\partial x_j}$, $h_{q2,j} = \widehat{u_j \tilde{T}} - \widehat{\tilde{u}_j \tilde{T}}$, $H_{q1,j} = -\hat{\Delta}^2 \hat{\rho} |\widehat{S}_{ij}| \frac{\partial \widehat{T}}{\partial x_j}$, and $H_{q2,j} = \hat{\rho}(\widehat{\tilde{u}_j \tilde{T}} - \widehat{\tilde{\tilde{u}}_j \tilde{T}})$, $T_j = H_{q1,j} - \widehat{\rho h_{q1,j}}$, and $V_j = H_{q2,j} - \widehat{\rho h_{q2,j}}$.

The performances of the ANNMM are assessed by comparing the spectra and statistical properties of velocity and temperature. It is important for the LES model to reproduce the spectrum and transfer of kinetic energy accurately. Based on Helmholtz decomposition, the

spectrum of velocity can be decomposed into solenoidal and compressible components.^{18,38,93,94} In Figs. 7–12, the spectra of velocity and its compressible component for different LES models are compared to the filtered DNS results. The inertial region with a $k^{-5/3}$ scaling of velocity spectrum is clearly visible. Noting that the small scales near the cutoff wavenumbers are hard to predict, the spectra prediction errors increase as k increases. The ANNMM predicts the velocity spectrum similar to the DSM and DMM at a grid resolution of 64^3 , especially for compressible velocity component. The implicit numerical dissipation of “no-model” decreases as the FGR (FGR = Δ/h_{LES}^Δ) increases. The spectrum of the velocity field predicted by the “no-model” is higher than that of the filtered DNS data

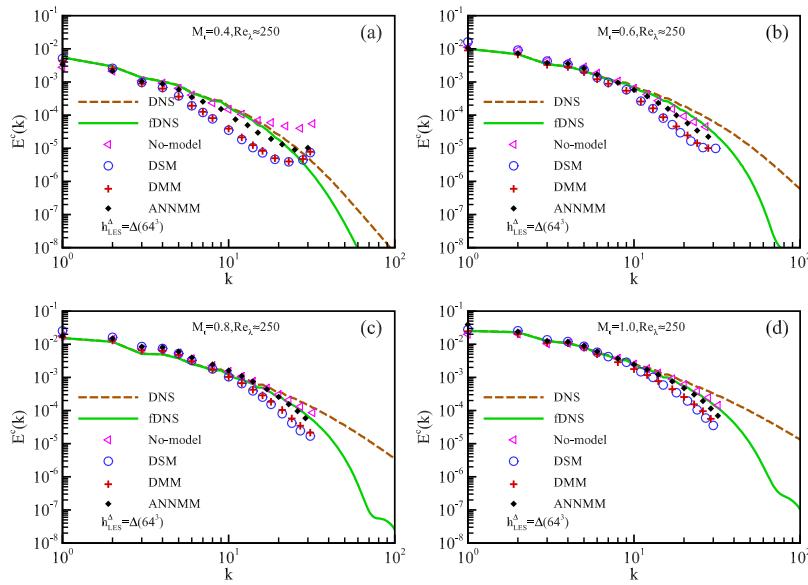


FIG. 11. Spectrum of compressible component of velocity for LES at grid resolutions of 64^3 ($h_{LES}^\Delta = \Delta$) with the filter width $\Delta = 16\Delta x$: (a) $M_t = 0.4$, (b) $M_t = 0.6$, (c) $M_t = 0.8$, and (d) $M_t = 1.0$.

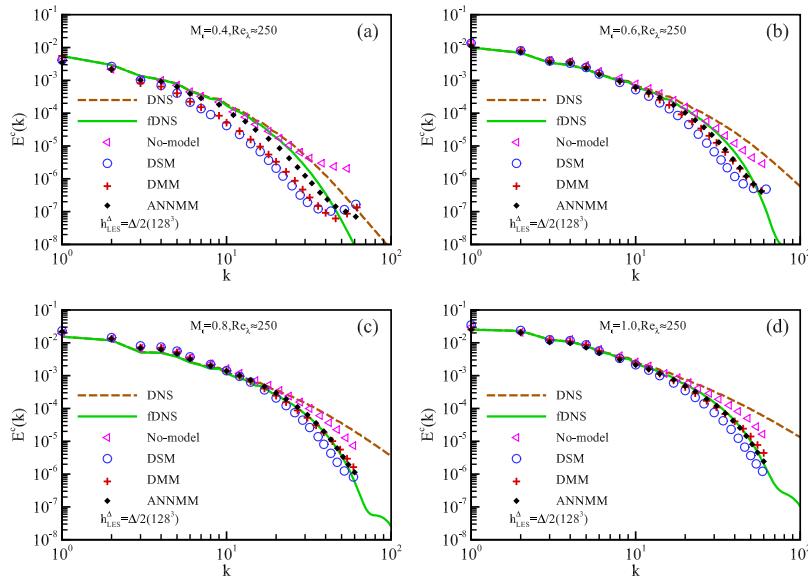


FIG. 12. Spectrum of compressible component of velocity for LES at grid resolutions of 128^3 ($h_{LES}^A = \Delta/2$) with the filter width $\Delta = 16\Delta x$: (a) $M_t = 0.4$, (b) $M_t = 0.6$, (c) $M_t = 0.8$, and (d) $M_t = 1.0$.

at large wavenumbers at 128^3 grid resolution. As FGR increases, the DSM, DMM, and ANNMM perform better due to the decrease in the numerical errors. It is worth noting that for LES at a grid resolution of 128^3 , the spectrum of the overall velocity field predicted by the ANNMM nearly overlap with the filtered DNS data. The DSM and DMM lead to the bump spectral distribution, where the low wavenumber is too energy-rich, while those near the cutoff are damped too strongly for the spectra of velocity.

We show the temperature spectra of different LES models in Figs. 13 and 14. The temperature spectra of DSM and DMM are in reasonable agreement with filtered-DNS (fDNS) for $k \leq 10$. As M_t increase, the spatial fluctuations of thermodynamic quantities

become stronger. The DSM predicts the spectra quite low because it is very dissipative near truncated wavenumbers. Meanwhile, the ANNMM predicts the temperature spectrum exactly even for wavenumber $k \leq 30$, which catches more smaller thermal scale structures as compared to other models at a grid resolution of 128^3 .

The spectra of velocity and its compressible component, and temperature for the filter width $\Delta = 32\Delta x$ are shown in Figs. 15 and 16. The spectrum of the overall velocity field predicted by the ANNMM nearly overlaps with the filtered DNS data.

The statistical properties and spatial structures of LES flow fields can be further analyzed by the longitudinal structure

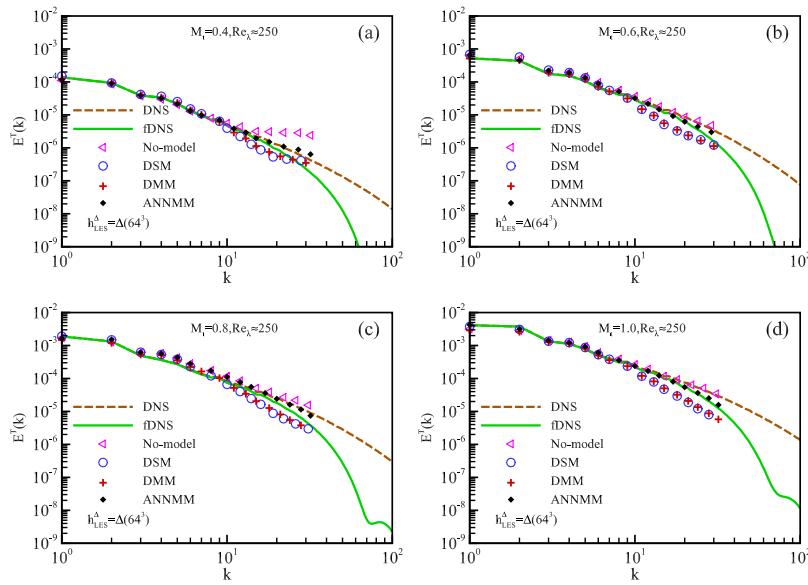


FIG. 13. Spectrum of temperature T for LES at grid resolutions of 64^3 ($h_{LES}^A = \Delta$) with the filter width $\Delta = 16\Delta x$: (a) $M_t = 0.4$, (b) $M_t = 0.6$, (c) $M_t = 0.8$, and (d) $M_t = 1.0$.

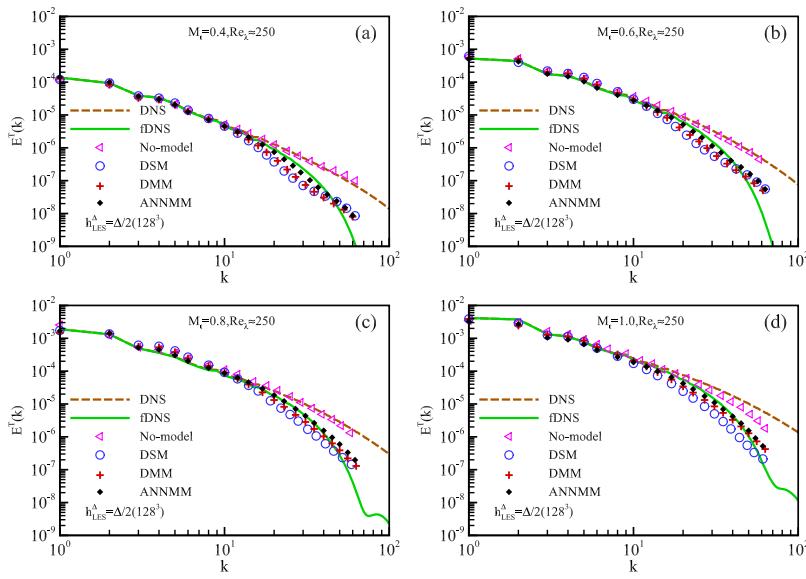


FIG. 14. Spectrum of temperature T for LES at grid resolutions of 128^3 ($h_{LES}^A = \Delta/2$) with the filter width $\Delta = 16\Delta x$: (a) $M_t = 0.4$, (b) $M_t = 0.6$, (c) $M_t = 0.8$, and (d) $M_t = 1.0$.

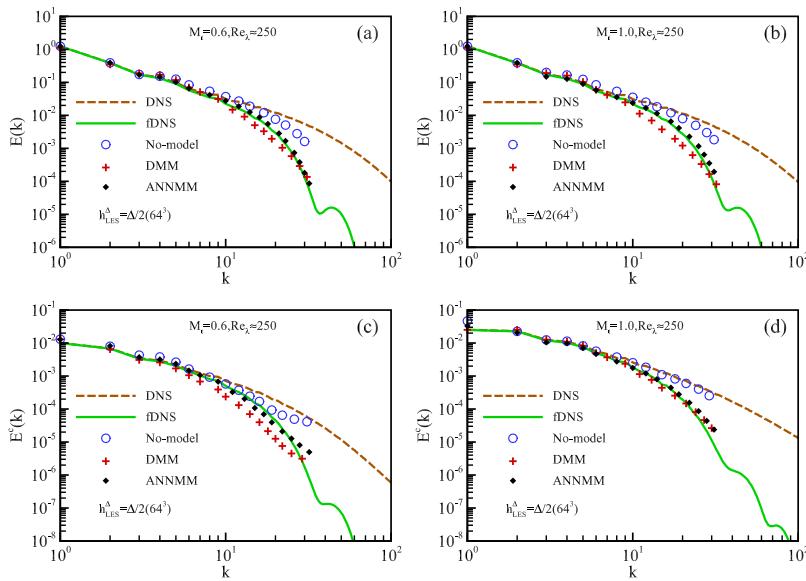


FIG. 15. Spectra of velocity and its compressible component for LES at grid resolutions of 64^3 ($h_{LES}^A = \Delta/2$) with the filter width $\Delta = 32\Delta x$: (a) spectrum of velocity at $M_t = 0.6$, (b) spectrum of velocity at $M_t = 1.0$, (c) spectrum of compressible component of velocity at $M_t = 0.6$, and (d) spectrum of compressible component of velocity at $M_t = 1.0$.

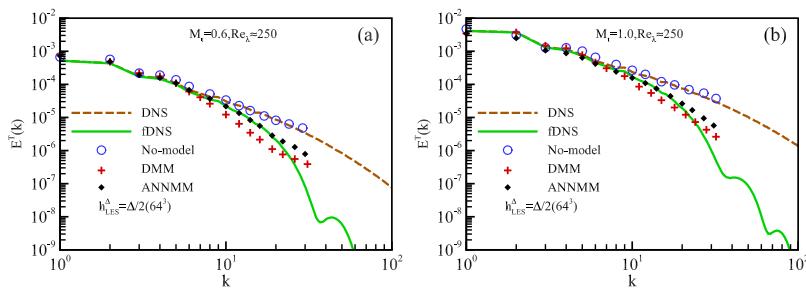


FIG. 16. Spectrum of temperature T for LES at grid resolutions of 64^3 ($h_{LES}^A = \Delta/2$) with the filter width $\Delta = 32\Delta x$: (a) $M_t = 0.6$ and (b) $M_t = 1.0$.

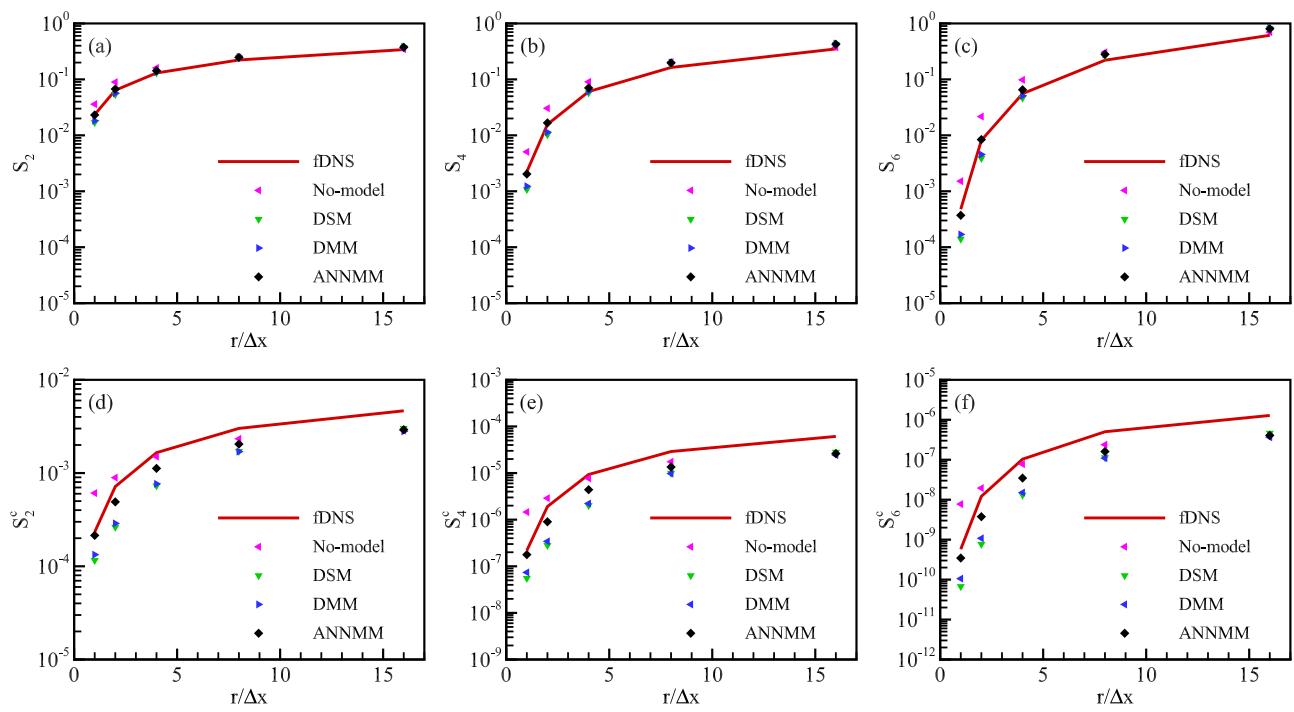


FIG. 17. Structure functions of the velocity and its compressible component at $M_t = 0.4$ for LES at grid resolution of 64^3 ($h_{LES}^\Delta = \Delta$) with the filter width $\Delta = 16\Delta_x$, where the first row shows the structure functions of the velocity and the second row shows the structure functions of the compressible velocity component: (a) S_2 , (b) S_4 , (c) S_6 , (d) S_2^c , (e) S_4^c , and (f) S_6^c .

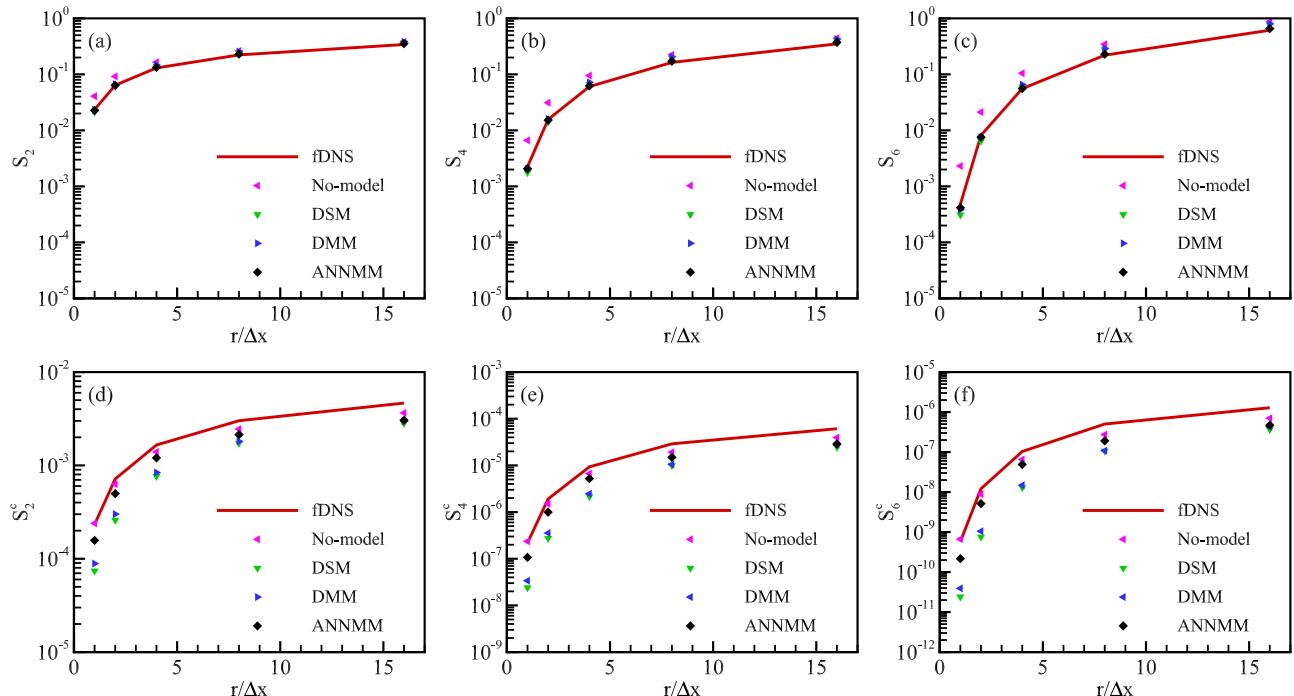


FIG. 18. Structure functions of the velocity and its compressible component at $M_t = 0.4$ for LES at grid resolution of 128^3 ($h_{LES}^\Delta = \Delta/2$) with the filter width $\Delta = 16\Delta_x$, where the first row shows the structure functions of the velocity and the second row shows the structure functions of the compressible velocity component: (a) S_2 , (b) S_4 , (c) S_6 , (d) S_2^c , (e) S_4^c , and (f) S_6^c .

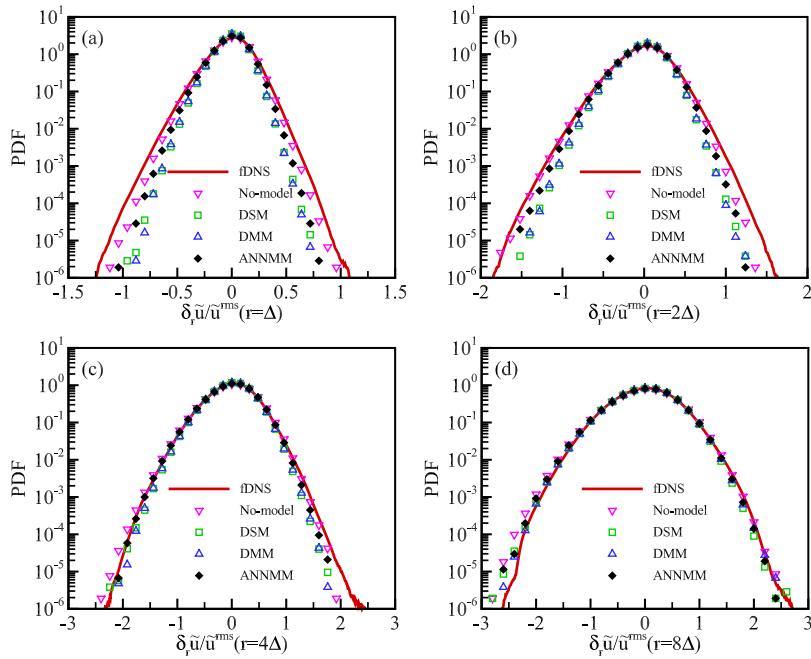


FIG. 19. PDF of the normalized increments of the velocity at $M_t = 0.6$ for LES at grid resolutions of 64^3 ($h_{LES}^\Delta = \Delta$) with the filter width $\Delta = 16\Delta x$: (a) $\delta_r \tilde{u}/\tilde{u}^{rms}(r = \Delta)$, (b) $\delta_r \tilde{u}/\tilde{u}^{rms}(r = 2\Delta)$, (c) $\delta_r \tilde{u}/\tilde{u}^{rms}(r = 4\Delta)$, and (d) $\delta_r \tilde{u}/\tilde{u}^{rms}(r = 8\Delta)$.

functions for the velocity and its two components, which can be defined by

$$S_n^L(r) \equiv \left\langle \left| \frac{\delta_r \tilde{u}}{\tilde{u}^{rms}} \right|^n \right\rangle, \quad (37)$$

$$S_n^{L,s}(r) \equiv \left\langle \left| \frac{\delta_r \tilde{u}^s}{\tilde{u}^{rms}} \right|^n \right\rangle, \quad (38)$$

where $\delta_r \tilde{u} = [\tilde{\mathbf{u}}(\mathbf{x} + \mathbf{r}) - \tilde{\mathbf{u}}(\mathbf{x})] \cdot \hat{\mathbf{r}}$, $\delta_r \tilde{u}^s = [\tilde{\mathbf{u}}^s(\mathbf{x} + \mathbf{r}) - \tilde{\mathbf{u}}^s(\mathbf{x})] \cdot \hat{\mathbf{r}}$, and $\delta_r \tilde{u}^c = [\tilde{\mathbf{u}}^c(\mathbf{x} + \mathbf{r}) - \tilde{\mathbf{u}}^c(\mathbf{x})] \cdot \hat{\mathbf{r}}$ are, respectively, the longitudinal increments of the velocity and its two components for the separation \mathbf{r} . Here, $\hat{\mathbf{r}} = \mathbf{r}/|\mathbf{r}|$. We have performed normalization by using the rms velocity \tilde{u}^{rms} .

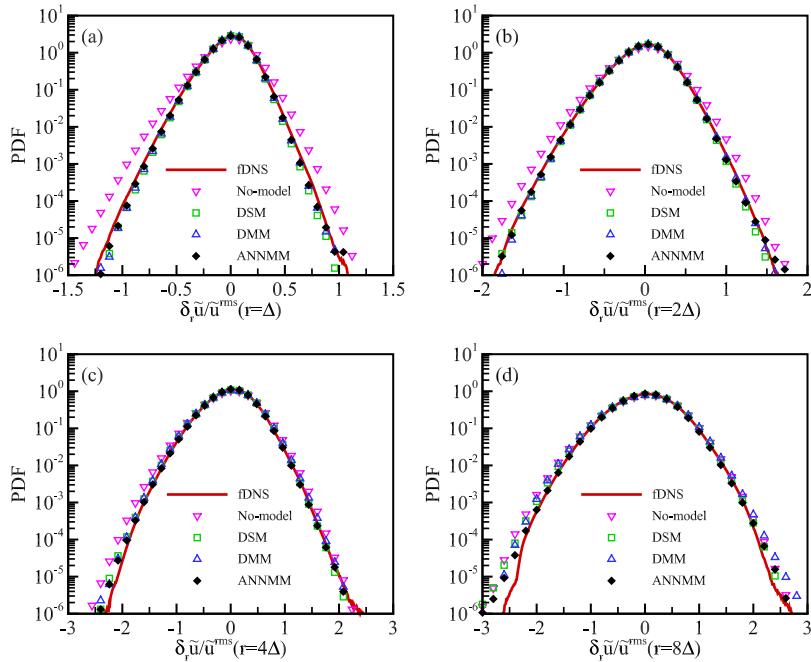


FIG. 20. PDF of the normalized increments of the velocity at $M_t = 0.6$ for LES at grid resolutions of 128^3 ($h_{LES}^\Delta = \Delta/2$) with the filter width $\Delta = 16\Delta x$: (a) $\delta_r \tilde{u}/\tilde{u}^{rms}(r = \Delta)$, (b) $\delta_r \tilde{u}/\tilde{u}^{rms}(r = 2\Delta)$, (c) $\delta_r \tilde{u}/\tilde{u}^{rms}(r = 4\Delta)$, and (d) $\delta_r \tilde{u}/\tilde{u}^{rms}(r = 8\Delta)$.

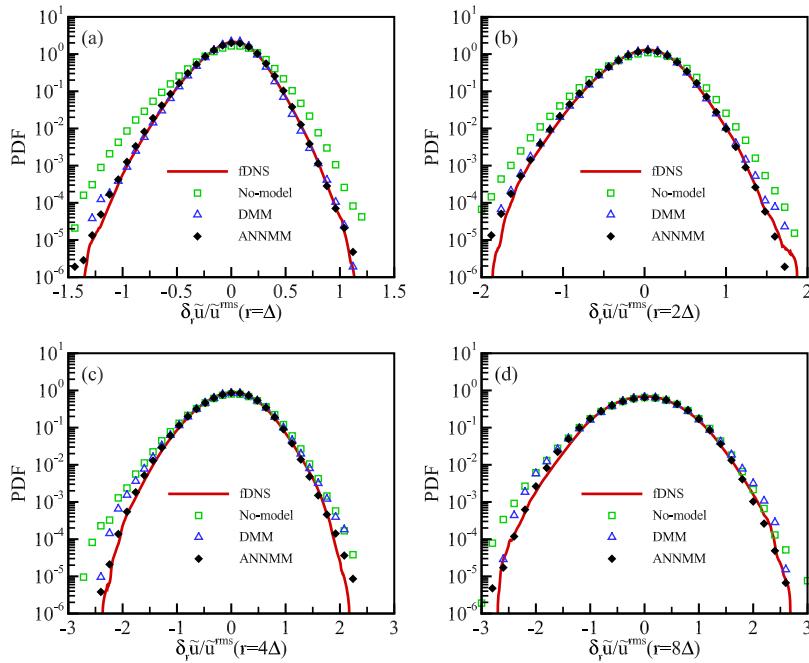


FIG. 21. PDF of the normalized increments of the velocity at $M_t = 0.6$ for LES at grid resolutions of 64^3 ($h_{LES}^3 = \Delta/2$) with the filter width $\Delta = 32\Delta x$: (a) $\delta_r \tilde{u}/\tilde{u}^{rms}(r = \Delta)$, (b) $\delta_r \tilde{u}/\tilde{u}^{rms}(r = 2\Delta)$, (c) $\delta_r \tilde{u}/\tilde{u}^{rms}(r = 4\Delta)$, and (d) $\delta_r \tilde{u}/\tilde{u}^{rms}(r = 8\Delta)$.

Figures 17 and 18 shows the normalized structure functions of the velocity and its compressible component. All the models can predict the structure functions well at large separations. As the separation decreases, the prediction error of all models increases, which is consistent with the results of the spectra at high wavenumbers. The filtered small scales are hard to model for they are strongly coupling with the unfiltered scales. The structure functions of

velocity and its compressible component at all scales predicted by the ANNMM are in good agreement with the filtered DNS data at a grid resolution of 128^3 , which indicates that the ANNMM can predict small scale fluctuations of turbulence accurately at a grid resolution of 128^3 .

Figures 19–24 show the normalized PDFs of the velocity increment and temperature increment \tilde{u}^{rms} and $\tilde{T}^{rms} = \sqrt{\langle (\tilde{T} - \tilde{T}_0)^2 \rangle}$

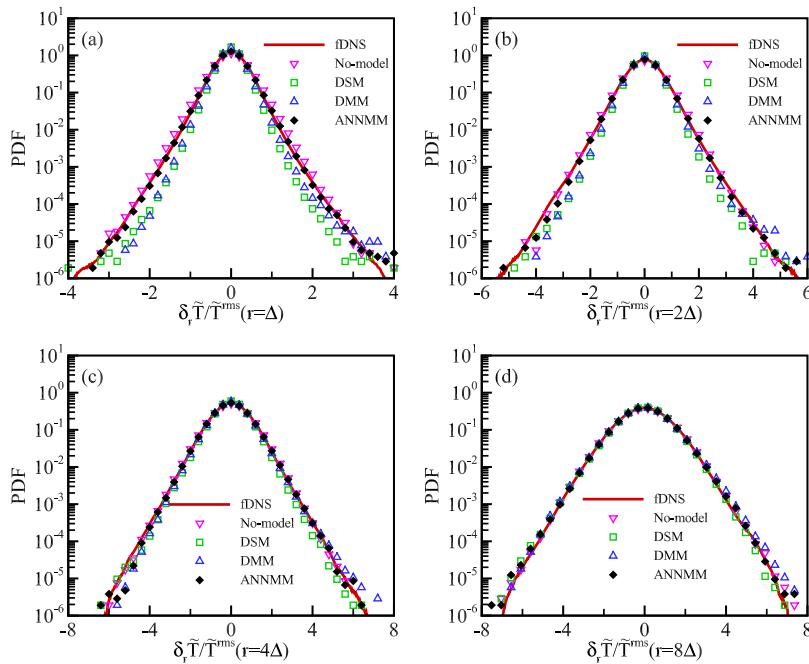


FIG. 22. PDF of the normalized increments of the temperature T at $M_t = 0.6$ for LES at grid resolutions of 64^3 ($h_{LES}^3 = \Delta$) with the filter width $\Delta = 16\Delta x$: (a) $\delta_r \tilde{T}/\tilde{T}^{rms}(r = \Delta)$, (b) $\delta_r \tilde{T}/\tilde{T}^{rms}(r = 2\Delta)$, (c) $\delta_r \tilde{T}/\tilde{T}^{rms}(r = 4\Delta)$, and (d) $\delta_r \tilde{T}/\tilde{T}^{rms}(r = 8\Delta)$.

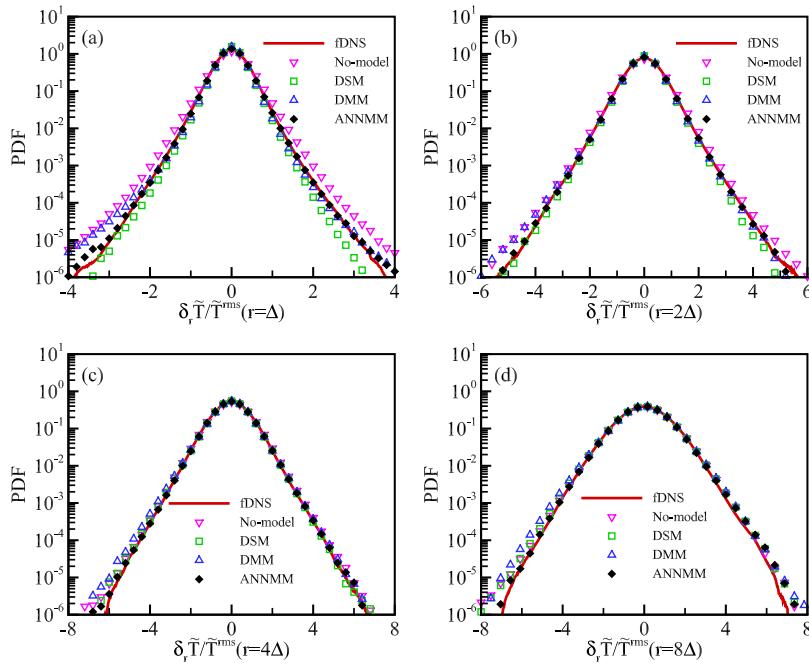


FIG. 23. PDF of the normalized increments of the temperature T at $M_t = 0.6$ for LES at grid resolutions of 128^3 ($h_{LES}^\Delta = \Delta/2$) with the filter width $\Delta = 16\Delta x$: (a) $\delta_r \tilde{T}/\tilde{T}^{rms}$ ($r = \Delta$), (b) $\delta_r \tilde{T}/\tilde{T}^{rms}$ ($r = 2\Delta$), (c) $\delta_r \tilde{T}/\tilde{T}^{rms}$ ($r = 4\Delta$), and (d) $\delta_r \tilde{T}/\tilde{T}^{rms}$ ($r = 8\Delta$).

are used to normalize the PDFs. All the PDFs exhibit a strong symmetry. As r increases, the scope increases. The ANNMM performs similar to DSM and DMM in the prediction of normalized PDF of temperature increment at small r . All the LES models predict the normalized PDF of velocity increment narrower than the result of fDNS at small r with a grid resolution of 64^3 ($h_{LES}^\Delta = \Delta$). As the grid

resolution increases to 128^3 with $h_{LES}^\Delta = \Delta/2$. All the models accurately reproduce the normalized PDF of velocity increment at all r , especially at small r . Besides, the prediction error of DSM and DMM increases near the tails of PDF for temperature increment, and the ANNMM accurately reproduces the PDF of temperature increment at all r , especially at small r , which are consistent with the results of

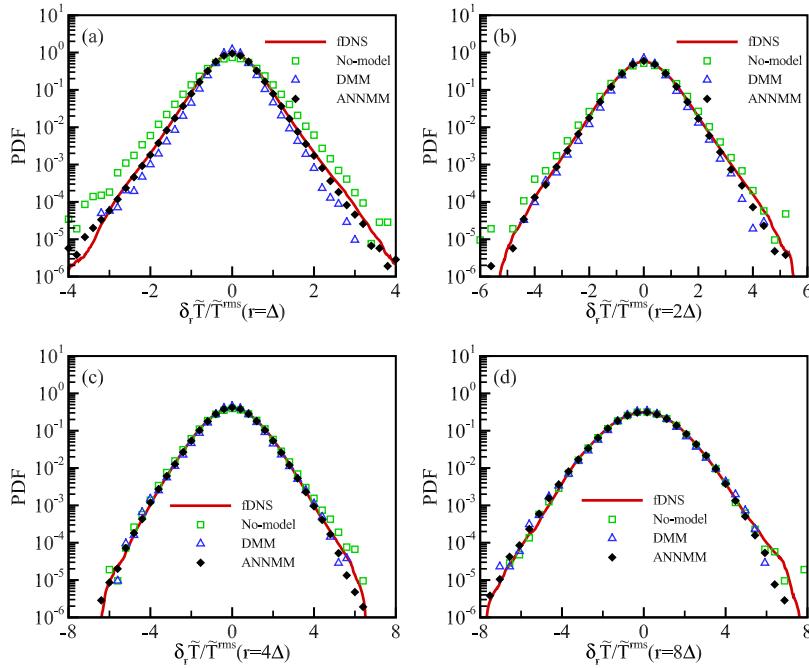


FIG. 24. PDF of the normalized increments of the temperature T at $M_t = 0.6$ for LES at grid resolutions of 64^3 ($h_{LES}^\Delta = \Delta/2$) with the filter width $\Delta = 32\Delta x$: (a) $\delta_r \tilde{T}/\tilde{T}^{rms}$ ($r = \Delta$), (b) $\delta_r \tilde{T}/\tilde{T}^{rms}$ ($r = 2\Delta$), (c) $\delta_r \tilde{T}/\tilde{T}^{rms}$ ($r = 4\Delta$), and (d) $\delta_r \tilde{T}/\tilde{T}^{rms}$ ($r = 8\Delta$).

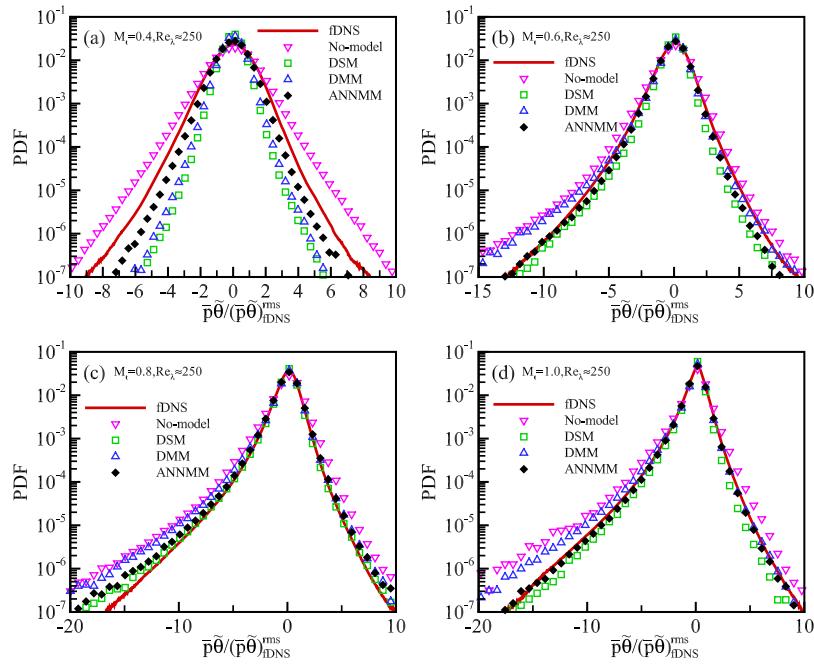


FIG. 25. PDF of the normalized pressure dilatation $\bar{p}\tilde{\theta}/(\bar{p}\tilde{\theta})_{IDNS}^{rms}$ for LES at grid resolutions of 128^3 ($h_{LES}^\Delta = \Delta/2$) with the filter width $\Delta = 16\Delta x$: (a) $M_t = 0.4$, (b) $M_t = 0.6$, (c) $M_t = 0.8$, and (d) $M_t = 1.0$.

velocity and temperature spectra. The PDFs of the normalized increments of velocity and temperature for the filter width $\Delta = 32\Delta x$ are shown in Figs. 21 and 24. The ANNMM is in agreement with the filtered DNS data.

The PDFs of the normalized pressure-dilatation $\bar{p}\tilde{\theta}/(\bar{p}\tilde{\theta})_{IDNS}^{rms}$ and velocity divergence $\tilde{\theta}/\tilde{\theta}_{IDNS}^{rms}$ for LES at a grid resolution of 128^3

are shown in Figs. 25 and 26 for all four M_t . The PDFs are almost symmetric at $M_t = 0.4$. As M_t increases, the PDFs skew toward the negative side and the tails become longer. The right tails are very short, while the left tails fall very slowly. When the turbulent Mach number becomes higher, the rms value of velocity divergence increases and its PDF becomes more skewed negatively due

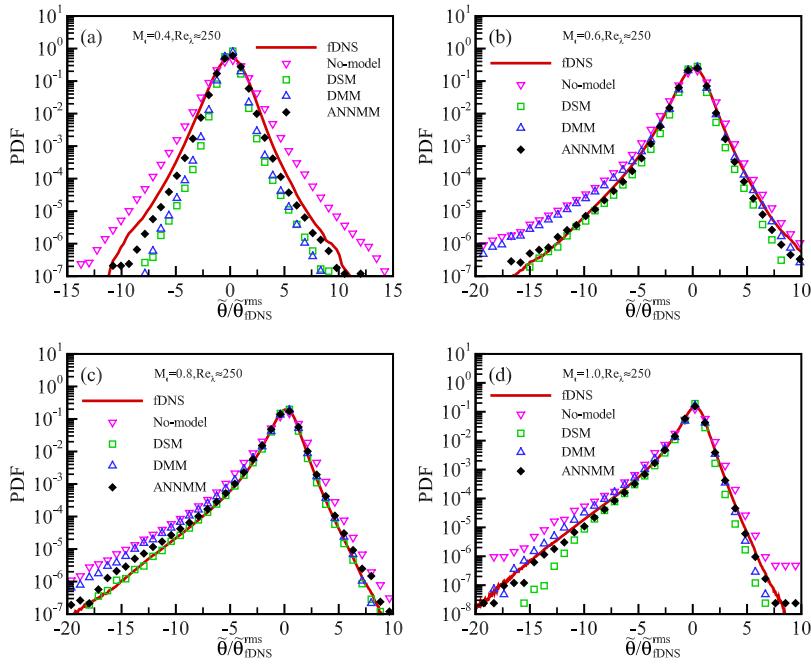


FIG. 26. PDF of the normalized velocity divergence $\tilde{\theta}/\tilde{\theta}_{IDNS}^{rms}$ for LES at grid resolutions of 128^3 ($h_{LES}^\Delta = \Delta/2$) with the filter width $\Delta = 16\Delta x$: (a) $M_t = 0.4$, (b) $M_t = 0.6$, (c) $M_t = 0.8$, and (d) $M_t = 1.0$.

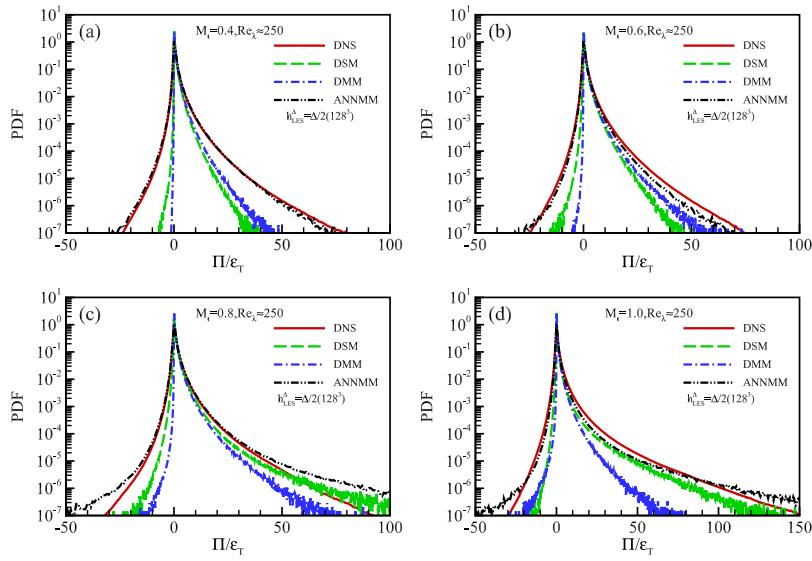


FIG. 27. PDF of the normalized SGS flux Π/ε_T for LES at 128^3 grid resolution, where $h_{LES}^3 = \Delta/2$ with the filter width $\Delta = 16\Delta x$: (a) $M_t = 0.4$, (b) $M_t = 0.6$, (c) $M_t = 0.8$, and (d) $M_t = 1.0$.

to more shocklets with stronger strength in compressible turbulence.²¹ The ANNMM shows high values at the tails of PDF at lower M_t , which are closer to the values of fDNS. The peaks of fDNS are recovered most accurately by the ANNMM. The results of the

PDFs for “no-model” LES are not in agreement with the filtered DNS data. The DMM predicts higher at the negative tail of PDFs of $\tilde{p}\tilde{\theta}/(\tilde{p}\tilde{\theta})_{fDNS}^{rms}$. Meanwhile, the DSM and ANNMM reproduce the PDFs more accurately for higher M_t .

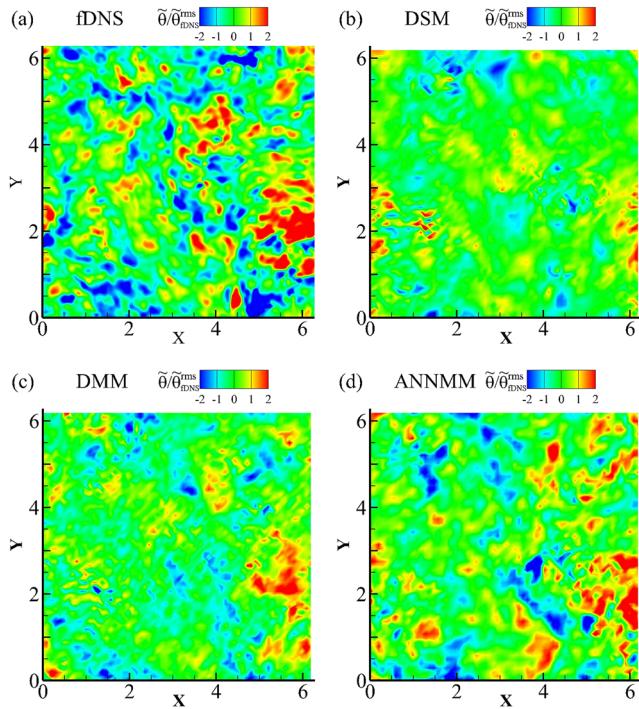


FIG. 28. Contours of the normalized velocity divergence $\tilde{\theta}/\tilde{\theta}_{fDNS}^{rms}$ on an arbitrarily selected x-y slice, at $M_t = 0.4$, and $t/\tau = 3.37$ (here $\tau \equiv L/u^{rms}$ is the large-eddy turnover time): (a) fDNS, (b) DSM, (c) DMM, and (d) ANNMM for LES at grid resolutions of 64^3 ($h_{LES}^3 = \Delta$) with the filter width $\Delta = 16\Delta x$.

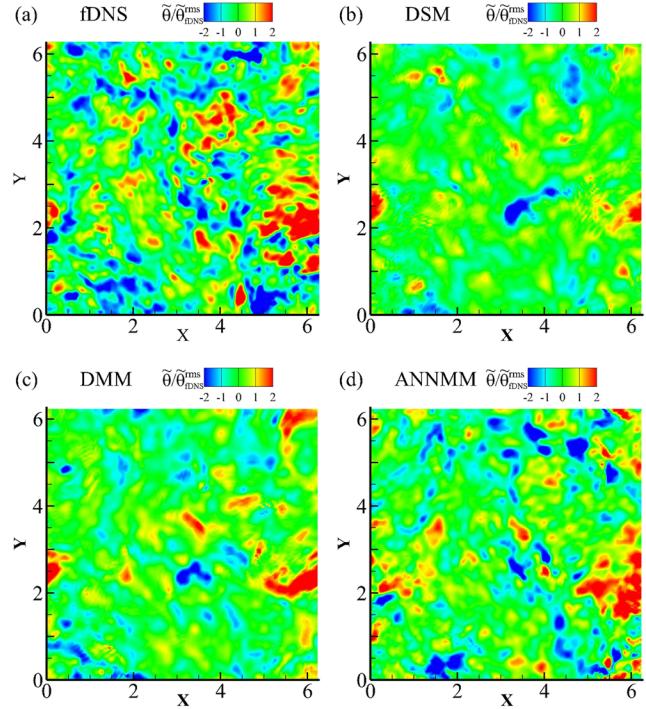


FIG. 29. Contours of the normalized velocity divergence $\tilde{\theta}/\tilde{\theta}_{fDNS}^{rms}$ on an arbitrarily selected x-y slice, at $M_t = 0.4$, and $t/\tau = 3.37$ (here $\tau \equiv L/u^{rms}$ is the large-eddy turnover time): (a) fDNS, (b) DSM, (c) DMM, and (d) ANNMM for LES at grid resolutions of 128^3 ($h_{LES}^3 = \Delta/2$) with the filter width $\Delta = 16\Delta x$.

We show the normalized PDFs of Π/ε_T , where $\varepsilon_T = -\langle p\theta \rangle + \langle \sigma_{ij}S_{ij}/Re \rangle$ is the total dissipation rate of kinetic energy in Fig. 27. The SGS energy flux Π is crucial for reproducing the correct properties of the SGS energy transfer. As M_t increases, the PDFs of Π/ε_T skew toward the positive side. The negative values of the SGS flux in the filtering velocity of turbulence predicted by the ANNMM nearly overlap with the DNS data, which represent the local reverse transfer of energy from small scales to large scales. The positive tail of PDF of the SGS flux of kinetic energy can be recovered more accurately by the ANNMM. The ANNMM performs better than the DSM and the DMM, which captures the overall shape of the PDF, especially for the tails of PDF.

We also explore the different characters of LES models in predicting the temporal evolution of the velocity fields and instantaneous coherent structures, which can be used to evaluate the performance of SGS models to reproduce both the phase and magnitude of flow structures predicted by DNS when the initial field is same. LES with different SGS models with an instantaneous flow field from filtered DNS as the initial condition is tested. In Figs. 28–31, the structures of velocity divergence for fDNS and LES are compared at the same instant in order to examine whether or not LES captures coherent structures correctly. There are more large-scale bandlike structures at high compression levels in fDNS as M_t increases, which indicate the occurrence of shocklets.^{19–21,96,101} All the models can catch the bandlike structures of velocity divergence, which are

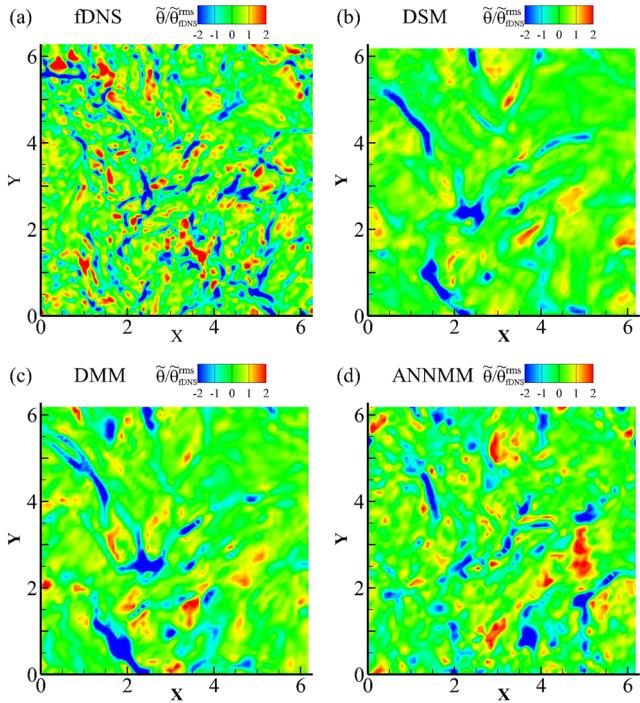


FIG. 30. Contours of the normalized velocity divergence $\tilde{\theta}/\tilde{\theta}_{fDNS}^{rms}$ on an arbitrarily selected x-y slice, at $M_t = 1.0$, and $t/\tau = 3.21$: (a) fDNS, (b) DSM, (c) DMM, and (d) ANNMM for LES at grid resolutions of 64^3 ($h_{LES}^\Delta = \Delta$) with the filter width $\Delta = 16\Delta x$.

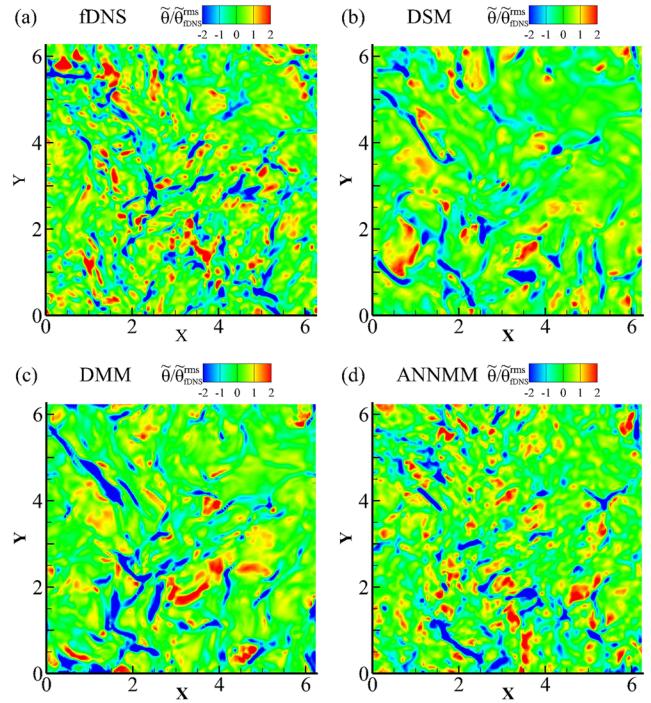


FIG. 31. Contours of the normalized velocity divergence $\tilde{\theta}/\tilde{\theta}_{fDNS}^{rms}$ on an arbitrarily selected x-y slice, at $M_t = 1.0$, and $t/\tau = 3.21$: (a) fDNS, (b) DSM, (c) DMM, and (d) ANNMM for LES at grid resolutions of 128^3 ($h_{LES}^\Delta = \Delta/2$) with the filter width $\Delta = 16\Delta x$.

nearly identical to those from fDNS. Meanwhile, some small structures are missing in the flow fields from the DSM and the DMM. The ANNMM can recover more small-scale structures, which are similar to those from fDNS.

VI. CONCLUSION

To summarize, the SGS stress and SGS heat flux of compressible turbulence are reconstructed based on the ANN framework, using DNS data for solenoidally forced compressible isotropic turbulence with M_t ranging from 0.4 to 1.0. We use the functional form of the mixed model combining the gradient model and the Smagorinsky's eddy viscosity model to predict the SGS stress and heat flux. The ANN is used to construct the model coefficients of the functional form of the mixed model for the SGS anisotropy stress, the SGS isotropy stress, and the SGS heat flux. The systematic analysis of the ANNMM shows that in an *a priori* test, the ANNMM gives the high correlation coefficient and recovers the PDFs of the SGS flux of kinetic energy accurately. In an *a posteriori* analysis, the ANNMM performs better than DSM and DMM in the prediction of spectra and statistical properties of velocity and temperature and instantaneous flow structures at a grid resolution of 128^3 ($h_{LES}^\Delta = \Delta/2$). The DSM and DMM lead to the bump spectral distribution where low wavenumbers are too energy-rich, while those near the cutoff are damped too strongly. The spectrum of the overall velocity field predicted by the ANNMM nearly overlaps with the filtered DNS data at

a grid resolution of 128^3 ($h_{LES}^\Delta = \Delta/2$). Besides, the ANNMM reconstructs the PDFs of SGS flux of kinetic energy accurately at a grid resolution of 128^3 ($h_{LES}^\Delta = \Delta/2$). Meanwhile, the ANNMM improves the prediction of the turbulence structures and statistics near the smallest resolved scales. Thus, the ANNMM with reasonable functional model forms and input features based on physical nature of turbulence can reconstruct the SGS stress τ_{ij} and SGS heat flux Q_j well.

More physical functional and structural characteristics of turbulence should be considered in ANN models in order to achieve the efficiency of traditional models and provide better predictions of flow statistics and structures in LES of compressible turbulence, including the different types of coherent structures, vortex dynamics, the nonlinear couplings of the solenoidal mode, compressible model and thermodynamic mode, and the nonlocal temporal and spatial effects of the dynamics of SGS scales.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (NSFC Grant Nos. 11702127, 91752201, and 11672123) and by the Technology and Innovation Commission of Shenzhen Municipality (Grant Nos. JCYJ20170412151759222 and ZDSYS201802081843517). This work was also supported by the Center for Computational Science and Engineering of Southern University of Science and Technology. J.W. acknowledges the support from the Young Elite Scientist Sponsorship Program by CAST (Grant No. 2016QNRC001).

REFERENCES

- ¹P. 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).
- ²C. Meneveau and J. Katz, "Scale-invariance and turbulence models for large-eddy simulation," *Annu. Rev. Fluid Mech.* **32**, 1 (2000).
- ³D. I. Pullin, "A vortex-based model for the subgrid flux of a passive scalar," *Phys. Fluids* **12**, 2311 (2000).
- ⁴C. Meneveau, "Lagrangian dynamics and models of the velocity gradient tensor in turbulent flows," *Annu. Rev. Fluid Mech.* **43**, 219 (2011).
- ⁵S. 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).
- ⁶P. A. Durbin, "Some recent developments in turbulence closure modeling," *Annu. Rev. Fluid Mech.* **50**, 77 (2018).
- ⁷M. Buzzicotti, M. Linkmann, H. Aluie, L. Biferale, J. Brasseur, and C. Meneveau, "Effect of filter type on the statistics of energy transfer between resolved and subfilter scales from *a priori* analysis of direct numerical simulations of isotropic turbulence," *J. Turbul.* **19**, 167 (2018).
- ⁸J. Smagorinsky, "General circulation experiments with the primitive equations, I. The basic experiment," *Mon. Weather Rev.* **91**, 99 (1963).
- ⁹D. K. Lilly, "The representation of small-scale turbulence in numerical simulation experiments," in *Proceedings of IBM Scientific Computing Symposium on Environmental Sciences* (IBM, 1967), pp. 195–210.
- ¹⁰J. W. Deardorff, "A numerical study of three-dimensional turbulent channel flow at large Reynolds numbers," *J. Fluid Mech.* **41**, 453 (1970).
- ¹¹W. Schmidt, J. C. Niemeyer, and W. Hillebrandt, "A localised subgrid scale model for fluid dynamical simulations in astrophysics—I. Theory and numerical tests," *Astron. Astrophys.* **450**, 265 (2006).
- ¹²W. Schmidt and C. Federrath, "A fluid-dynamical subgrid scale model for highly compressible astrophysical turbulence," *Astron. Astrophys.* **528**, A106 (2011).
- ¹³X. C. Chai and K. Mahesh, "Dynamic k -equation model for large-eddy simulation of compressible flows," *J. Fluid Mech.* **699**, 385 (2012).
- ¹⁴R. Teyssier, "Grid-based hydrodynamics in astrophysical fluid flows," *Annu. Rev. Astron. Astrophys.* **53**, 325 (2015).
- ¹⁵C. P. Yu, Z. L. Xiao, and X. L. Li, "Scale-adaptive subgrid-scale modelling for large-eddy simulation of turbulent flows," *Phys. Fluids* **29**, 035101 (2017).
- ¹⁶C. 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).
- ¹⁷M. R. Petersen and D. Livescu, "Forcing for statistically stationary compressible isotropic turbulence," *Phys. Fluids* **22**, 116101 (2010).
- ¹⁸S. Jagannathan and D. A. Donzis, "Reynolds and Mach number scaling in solenoidally-forced compressible turbulence using high-resolution direct numerical simulations," *J. Fluid Mech.* **789**, 669 (2016).
- ¹⁹J. Wang, T. Gotoh, and T. Watanabe, "Spectra and statistics in compressible isotropic turbulence," *Phys. Rev. Fluids* **2**, 013403 (2017).
- ²⁰J. Wang, T. Gotoh, and T. Watanabe, "Scaling and intermittency in compressible isotropic turbulence," *Phys. Rev. Fluids* **2**, 053401 (2017).
- ²¹J. Wang, T. Gotoh, and T. Watanabe, "Shocklet statistics in compressible isotropic turbulence," *Phys. Rev. Fluids* **2**, 023401 (2017).
- ²²J. C. Wang, M. P. Wan, S. Chen, C. Y. Xie, and S. Y. Chen, "Effect of shock waves on the statistics and scaling in compressible isotropic turbulence," *Phys. Rev. E* **97**, 043108 (2018).
- ²³S. 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).
- ²⁴S. 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).
- ²⁵D. D. Zhang, J. G. Tan, and X. Yao, "Direct numerical simulation of spatially developing highly compressible mixing layer: Structural evolution and turbulent statistics," *Phys. Fluids* **31**, 036102 (2019).
- ²⁶C. G. Speziale, G. Erlebacher, T. A. Zang, and M. Y. Hussaini, "The subgrid-scale modeling of compressible turbulence," *Phys. Fluids* **31**, 940 (1988).
- ²⁷G. Erlebacher, M. Y. Hussaini, C. G. Speziale, and T. A. Zang, "On the large-eddy simulation of compressible isotropic turbulence," in *Lecture Notes in Physics* (Springer-Verlag, 1990), Vol. 371, p. 121.
- ²⁸G. Erlebacher, M. Hussaini, C. Speziale, and T. Zang, "Toward the large-eddy simulation of compressible turbulent flows," *J. Fluid Mech.* **238**, 155 (1992).
- ²⁹D. Knight, G. Zhou, N. Okong's, and V. Shukla, "Compressible large eddy simulation using unstructured grids," AIAA Paper No. 98-0535, 1998.
- ³⁰M. P. Martín, U. Piomelli, and G. V. Candler, "Subgrid-scale models for compressible large-eddy simulations," *Theor. Comput. Fluid Dyn.* **13**, 361 (2000).
- ³¹S. Stolz, N. A. Adams, and L. Kleiser, "The approximate deconvolution model for large-eddy simulations of compressible flows and its application to shock-turbulent-boundary-layer interaction," *Phys. Fluids* **13**, 2985 (2001).
- ³²T. Dubois, J. A. Domaradzki, and A. Honein, "The subgrid-scale estimation model applied to large eddy simulations of compressible turbulence," *Phys. Fluids* **14**, 1781 (2002).
- ³³S. Boquet, P. Sagaut, and J. Jouhaut, "A compressible wall model for large-eddy simulation with application to prediction of aerothermal quantities," *Phys. Fluids* **24**, 065103 (2012).
- ³⁴N. M. El-Hady, T. A. Zang, and U. Piomelli, "Application of the dynamic subgrid-scale model to axisymmetric transitional boundary layer at high speed," *Phys. Fluids* **6**, 1299 (1994).
- ³⁵T. 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).
- ³⁶B. Vreman, B. Geurts, and H. Kuerten, "Subgrid-modeling in LES of compressible flow," *Appl. Sci. Res.* **54**, 191 (1995).
- ³⁷B. Vreman, B. Geurts, and H. Kuerten, "Large-eddy simulation of the turbulent mixing layer," *J. Fluid Mech.* **339**, 357 (1997).
- ³⁸P. Sagaut and C. Cambon, *Homogeneous Turbulence Dynamics* (Cambridge University Press, 2008).

- ³⁹R. Rogallo and P. Moin, "Numerical simulation of turbulent flows," *Annu. Rev. Fluid Mech.* **16**, 99–137 (1984).
- ⁴⁰M. Germano, U. Piomelli, P. Moin, and W. Cabot, "A dynamic subgrid-scale eddy-viscosity model," *Phys. Fluids A* **3**, 1760–1765 (1991).
- ⁴¹M. Germano, "Turbulence: The filtering approach," *J. Fluid Mech.* **238**, 325 (1992).
- ⁴²D. Lilly, "A proposed modification of the Germano subgrid scale closure method," *Phys. Fluids A* **4**, 633 (1992).
- ⁴³J. Bardina, J. H. Ferziger, and W. C. Reynolds, "Improved subgrid scale models for large eddy simulation," AIAA Paper No. 80-1357, 1980, p. 10.
- ⁴⁴S. Liu, C. Meneveau, and J. Katz, "On the properties of similarity subgrid-scale models as deduced from measurements in a turbulent jet," *J. Fluid Mech.* **275**, 83 (1994).
- ⁴⁵Y. P. Shi, Z. L. Xiao, and S. Y. Chen, "Constrained subgrid-scale stress model for large eddy simulation," *Phys. Fluids* **20**, 011701 (2008).
- ⁴⁶R. A. Clark, "Evaluation of sub-grid scalar models using an accurately simulated turbulent flow," *J. Fluid Mech.* **91**, 1–16 (1979).
- ⁴⁷B. Vreman, B. Geurts, and H. Kuerten, "Large eddy simulation of the temporal mixing layer using the Clark model," *Theor. Comput. Fluid Dyn.* **8**, 309–324 (1996).
- ⁴⁸J. A. Langford and R. D. Moser, "Optimal LES formulations for isotropic turbulence," *J. Fluid Mech.* **398**, 321–346 (1999).
- ⁴⁹J. A. Langford and R. D. Moser, "Optimal large-eddy simulation results for isotropic turbulence," *J. Fluid Mech.* **521**, 273–294 (2004).
- ⁵⁰R. 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).
- ⁵¹B. C. Wang, J. Yin, E. Yee, and D. J. Bergstrom, "A complete and irreducible dynamic SGS heat-flux modelling based on the strain rate tensor for large-eddy simulation of thermal convection," *Int. J. Heat Fluid Flow* **28**, 1227–1243 (2007).
- ⁵²T. J. R. Hughes, L. Mazzei, and K. E. Jansen, "Large eddy simulation and the variational multiscale method," *Comput. Visualization Sci.* **3**, 47–59 (2000).
- ⁵³C. Brun and R. Friedrich, "Modeling the test SGS tensor T_{ij} : An issue in the dynamic approach," *Phys. Fluids* **13**, 2373–2385 (2001).
- ⁵⁴C. Brun, R. Friedrich, and C. B. da Silva, "A non-linear SGS model based on the spatial velocity increment," *Theor. Comput. Fluid Dyn.* **20**, 1–21 (2006).
- ⁵⁵H. Xiao and P. Jenny, "A consistent dual-mesh framework for hybrid LES/RANS modeling," *J. Comput. Phys.* **231**, 1848–1865 (2012).
- ⁵⁶H. Xiao, J. X. Wang, and P. Jenny, "An implicitly consistent formulation of a dual-mesh hybrid LES/RANS method," *Commun. Comput. Phys.* **21**, 570–599 (2017).
- ⁵⁷D. You, S. Bose, and P. Moin, "Grid-independent large-eddy simulation of compressible turbulent flows using explicit filtering," in *Proceedings of the Summer Program* (Center for Turbulence Research, 2010), pp. 203–210.
- ⁵⁸M. Germano, A. Abba, R. Arina, and L. Bonaventura, "On the extension of the eddy viscosity model to compressible flows," *Phys. Fluids* **26**, 041702 (2014).
- ⁵⁹H. Aluie, "Compressible turbulence: The cascade and its locality," *Phys. Rev. Lett.* **106**, 174502 (2011).
- ⁶⁰R. Wagner, G. Falkovich, and A. G. Kritsuk, "Flux correlations in supersonic isothermal turbulence," *J. Fluid Mech.* **713**, 482–490 (2012).
- ⁶¹H. Aluie, "Scale decomposition in compressible turbulence," *Physica D* **247**, 54–65 (2013).
- ⁶²J. C. Wang, Y. T. Yang, Y. P. Shi, Z. L. Xiao, X. T. He, and S. Y. Chen, "Cascade of kinetic energy in three-dimensional compressible turbulence," *Phys. Rev. Lett.* **110**, 214505 (2013).
- ⁶³J. C. Wang, M. P. Wan, S. Chen, C. Y. Xie, L. P. Wang, and S. Y. Chen, "Cascades of temperature and entropy fluctuations in compressible turbulence," *J. Fluid Mech.* **867**, 195–215 (2019).
- ⁶⁴J. 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).
- ⁶⁵Z. J. Zhang and K. Duraisamy, *Machine Learning Methods for Data-Driven Turbulence Modeling* (AIAA, 2015).
- ⁶⁶J. Ling, R. Jones, and J. Templeton, "Machine learning strategies for systems with invariance properties," *J. Comput. Phys.* **318**, 22 (2016).
- ⁶⁷B. D. Tracey, K. Duraisamy, and J. J. Alonso, *A Machine Learning Strategy to Assist Turbulence Model Development* (AIAA, 2015).
- ⁶⁸S. 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. U. S. A.* **113**, 3932 (2016).
- ⁶⁹A. T. Mohan, and D. V. Gaitonde, "A deep learning based approach to reduced order modeling for turbulent flow control using LSTM neural networks," e-print arXiv:1804.09269 (2018).
- ⁷⁰M. Raissi and G. E. Karniadakis, "Hidden physics models: Machine learning of nonlinear partial differential equations," *J. Comput. Phys.* **357**, 125 (2018).
- ⁷¹Z. 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).
- ⁷²J.-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).
- ⁷³S. W. Pan and K. Duraisamy, "Data driven discovery of closure models," *SIAM J. Appl. Dyn. Syst.* **17**(4), 2381 (2018).
- ⁷⁴S. W. Pan and K. Duraisamy, "Long time predictive modeling of nonlinear dynamical system using neural networks," *Complexity* **2018**, 480102.
- ⁷⁵F. Sarghini, G. de Felice, and S. Santini, "Neural networks based subgrid scale modeling in large eddy simulations," *Comput. Fluids* **32**, 97 (2003).
- ⁷⁶J. Ling, A. Kurzawski, and J. Templeton, "Reynolds averaged turbulence modelling using deep neural networks with embedded invariance," *J. Fluid Mech.* **807**, 155 (2016).
- ⁷⁷H. 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).
- ⁷⁸M. Gamahara and Y. Hattori, "Searching for turbulence models by artificial neural network," *Phys. Rev. Fluids* **2**(5), 054604 (2017).
- ⁷⁹A. Volland, G. Balarac, and C. Corre, "Subgrid-scale scalar flux modelling based on optimal estimation theory and machine-learning procedures," *J. Turbul.* **18**(9), 854 (2017).
- ⁸⁰R. Maulik and O. San, "A neural network approach for the blind deconvolution of turbulent flows," *J. Fluid Mech.* **831**, 151 (2017).
- ⁸¹J. N. Kutz, "Deep learning in fluid dynamics," *J. Fluid Mech.* **814**, 1 (2017).
- ⁸²C. Ma, J. Wang, and W. E, "Model reduction with memory and the machine learning of dynamical systems," *Commun. Comput. Phys.* **25**(4), 947–962 (2019).
- ⁸³L. Jofre, S. P. Domino, and G. Iaccarino, "A framework for characterizing structural uncertainty in large-eddy simulation closures," *Flow, Turbul. Combust.* **100**(2), 341 (2018).
- ⁸⁴Z. 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).
- ⁸⁵R. Maulik, O. San, A. Rasheed, and P. Vedula, "Data-driven deconvolution for large eddy simulations of Kraichnan turbulence," *Phys. Fluids* **30**, 125109 (2018).
- ⁸⁶A. A. Mishra, K. Duraisamy, and G. Iaccarino, "Estimating uncertainty in homogeneous turbulence evolution due to coarse-graining," *Phys. Fluids* **31**, 025106 (2019).
- ⁸⁷L. Y. Zhu, W. W. Zhang, J. Q. Kou, and Y. L. Liu, "Machine learning methods for turbulence modeling in subsonic flows around airfoils," *Phys. Fluids* **31**, 015105 (2019).
- ⁸⁸R. Maulik, O. San, A. Rasheed, and P. Vedula, "Subgrid modelling for two-dimensional turbulence using neural networks," *J. Fluid Mech.* **858**, 122 (2019).
- ⁸⁹K. Duraisamy, G. Iaccarino, and H. Xiao, "Turbulence modeling in the age of data," *Annu. Rev. Fluid Mech.* **51**, 357 (2019).
- ⁹⁰C. 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).

- ⁹¹J. X. Wang, J. L. Wu, J. L. Ling, G. Iaccarino, and H. Xiao, "A comprehensive physics-informed machine learning framework for predictive turbulence modeling," e-print [arXiv:1701.07102v1](https://arxiv.org/abs/1701.07102v1) (2017).
- ⁹²R. 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).
- ⁹³R. Samtaney, D. I. Pullin, and B. Kosović, "Direct numerical simulation of decaying compressible turbulence and shocklet statistics," *Phys. Fluids* **13**, 1415 (2001).
- ⁹⁴J. Wang, Y. Shi, L.-P. Wang, Z. Xiao, X. T. He, and S. Chen, "Effect of compressibility on the small scale structures in isotropic turbulence," *J. Fluid Mech.* **713**, 588 (2012).
- ⁹⁵J. Wang, L.-P. Wang, Z. Xiao, Y. Shi, and S. Chen, "A hybrid numerical simulation of isotropic compressible turbulence," *J. Comput. Phys.* **229**, 5257 (2010).
- ⁹⁶J. Wang, Y. Shi, L.-P. Wang, Z. Xiao, X. T. He, and S. Chen, "Effect of shocklets on the velocity gradients in highly compressible isotropic turbulence," *Phys. Fluids* **23**, 125103 (2011).
- ⁹⁷A. Favre, "Equations des gaz turbulents compressibles - I - Formes générales," *J. Méc.* **4**, 361–390 (1965).
- ⁹⁸E. Garnier, N. Adams, and P. Sagaut, *Large Eddy Simulation for Compressible Flows* (Springer Verlag, 2009).
- ⁹⁹S. K. Lele, "Compact finite difference schemes with spectral-like resolution," *J. Comput. Phys.* **103**, 16 (1992).
- ¹⁰⁰D. S. Balsara and C. W. Shu, "Monotonicity preserving weighted essentially nonoscillatory schemes with increasingly high order of accuracy," *J. Comput. Phys.* **160**, 405 (2000).
- ¹⁰¹J. C. Wang, M. P. Wan, S. Chen, and S. Y. Chen, "Kinetic energy transfer in compressible isotropic turbulence," *J. Fluid Mech.* **841**, 581 (2018).
- ¹⁰²G. Zhang, B. E. Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks: The state of the art," *Int. J. Forecasting* **14**, 35 (1998).
- ¹⁰³H. B. Demuth, M. H. Beale, O. D. Jess, and M. T. Hagan, *Neural Network Design* (Stillwater, OK, 2014).
- ¹⁰⁴D. P. Kingma and J. B. Adam, "A method for stochastic optimization," e-print [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) (2014).
- ¹⁰⁵S. Ghosal, "An analysis of numerical errors in large-eddy simulations of turbulence," *J. Comput. Phys.* **125**, 187–206 (1996).
- ¹⁰⁶J. Meyers, B. J. Geurts, and M. Baelsmans, "Database analysis of errors in large-eddy simulation," *Phys. Fluids* **15**, 2740 (2003).
- ¹⁰⁷F. K. Chow and P. Moin, "A further study of numerical errors in large-eddy simulations," *J. Comput. Phys.* **184**, 366–380 (2003).
- ¹⁰⁸O. J. Mcmillan and J. H. Ferziger, "Direct testing of subgrid-scale models," *AIAA J.* **17**, 1340–1346 (1979).
- ¹⁰⁹S. B. Pope, "Ten questions concerning the large-eddy simulation of turbulent flows," *New J. Phys.* **6**, 35–59 (2004).
- ¹¹⁰U. Piomelli, A. Rouhi, and B. J. Geurts, "A grid-independent length scale for large-eddy simulations," *J. Fluid Mech.* **766**, 499–527 (2015).