AIAA AVIATION Forum
June 27-July 1, 2022, Chicago, IL & Virtual
AIAA AVIATION 2022 Forum

10.2514/6.2022-3326

Information-Theoretic Approach for Subgrid-Scale Modeling for High-Speed Compressible Wall Turbulence

Emily Williams* and Adrián Lozano-Durán†
Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA 02139

The problem of modeling for turbulent flows is investigated within the framework of information theory. A wall-modeled large-eddy simulation (WMLES) of a compressible turbulent channel flow is conducted using an equilibrium wall model and either the dynamic Smagorinsky (DSM) or information-preserving (IP) subgrid-scale (SGS) model. The IP SGS model is formulated using the Kullback-Leibler (KL) divergence. The model aims at minimizing the information lost between the probability mass distribution of the interscale energy transfer and viscous dissipation at different scales. The statistical quantities of interest are the mean velocity and mean temperature profiles. It is found that the IP SGS model matches or exceeds the accuracy of the DSM SGS model when compared to direct numerical simulation (DNS) data for the compressible channel.

I. Introduction

TURBULENT flows for external aerodynamic applications cannot be simulated by solving all the fluid motions of the Navier-Stokes equations because the range scales involved is large enough such that the computational cost becomes prohibitive. Thus, current state-of-the-art computational fluid dynamics (CFD) algorithms solve a modified set of Navier-Stokes equations for large-eddy simulation (LES). In LES, the large eddies are resolved, and the effect of the small scales on the larger eddies is modeled through a subgrid-scale (SGS) model. Recently, E. Williams and A. Lozano-Durán [1] analyzed the error scaling in predicting statistical quantities of interest with Mach number, Reynolds number, and grid resolution of various SGS models for LES of high-speed channel flows. The errors in the prediction of statistical quantities of interest informed by this analysis were estimated to be at least 10% to 20% in the mean temperature profile and 1% to 3% in the mean velocity profile for real-world applications, such as the Lockheed Martin X-59 QueSST (Quiet SuperSonic Technology). These findings further motivate investigation toward improving the models in order to predict statistical quantities of interest for external aerodynamic applications within the accuracy required by industry and academia.

Significant ongoing efforts have been devoted to capturing the essential flow physics in the form of reduced-order models. Modeling techniques to date have disregarded an essential principle of physics: the conservation of information, which may aid in the development of a new class of models. Information theory is the science about the laws governing information and relies on the notion of information as a fundamental property of physical systems [2]. Reduced-order modeling of chaotic systems can be posed as a problem of conservation of information. Lozano-Durán and Arranz [3] formulated the problem of modeling for high-dimensional, chaotic dynamical systems in information-theoretic terms. Modeled systems contain a smaller number of degrees of freedom than the original system, which in turn entails a loss of information. Therefore, the primary goal of modeling is to preserve the maximum amount of useful information from the original system. To this end, we will explore the potential of information theory in modeling these flows.

Lozano-Durán et al. [4] applied an information-theoretic approach of causality to study the dynamics of energy-containing eddies for wall turbulence and discussed the potential of information transfer in the design of reduced-order models. Particularly noteworthy is the work by Akaike [5], where competing models are selected on the basis of the relative amount of information from observations similar to Bayesian inference. Shavit and Falkovich [6] used information capacity to couple information and modeling to study the turbulence cascade. However, studies coupling the concept of information theory specifically toward advancing modeling techniques for turbulent flows are not abundant. The objective of this work is to assess this formulation of modeling for a high-speed compressible turbulent channel flow.

*Graduate Research Assistant, Department of Aeronautics and Astronautics, ejwllms@mit.edu, and AIAA Student Member.
†Draper Assistant Professor, Department of Aeronautics and Astronautics, adrianld@mit.edu, and AIAA Senior Member.

II. Technical Approach

A. Preliminaries

We introduce preliminary concepts of information theory required to formulate the problem of modeling. Consider the discrete random variable X taking values equal to x with probability mass function $p(x) = \Pr\{X = x\}$ over the finite set of outcomes $\mathcal{X}$ of X. The information of observing the event $X = x$ is defined as [7]

$$\mathcal{I}(x) = -\log_2[p(x)]. \tag{1}$$

The units of $\mathcal{I}(x)$ are bits, as set by the base chosen for this case. If we consider tossing a fair coin such that $p(\text{heads}) = p(\text{tails}) = 0.5$, the information of getting heads after one flip is $\mathcal{I}(\text{heads}) = -\log_2(0.5) = 1$ bit. In other words, observing the outcome of flipping a fair coin provides one bit of information. If the coin is completely biased toward heads, no information is gained since the outcome was already known before flipping the coin. Thus, information is the statistical notion of how unlikely it is to observe an event. The information $\mathcal{I}(x)$ is the number of bits required to unambiguously determine the state x.

The average information in X is given by the expectation $\langle\cdot\rangle$ over all the possible outcomes, defined as

$$H(X) = \langle \mathcal{I}(x) \rangle = \sum_{x\in\mathcal{X}} -p(x)\log_2[p(x)] \geq 0. \tag{2}$$

Equation (2) is referred to as the Shannon entropy. In flipping a fair coin n times, the entropy $H = -\sum 0.5^n \log_2(0.5^n) = n$ bits. Thus, Eq. (2) corresponds to the minimum average number of bits needed to encode a source of n states with probability distribution p [7]. So, H is zero when the process is completely deterministic, with no uncertainty in the outcome.

Interpreting information in the form of bits in terms of uncertainty is used in the formulation of the Kullback-Leibler (KL) divergence. If we consider two probability mass distributions, $p(x)$ and $q(x)$, the KL divergence is defined as

$$\mathrm{KL}(p,q) = \sum_{x\in\mathcal{X}} p(x)\log\frac{p(x)}{q(x)}, \tag{3}$$

which is a measure of the average number of bits required to recover $p(x)$ using the information in $q(x)$. The KL divergence also represents the information lost when $q(x)$ is used to approximate $p(x)$. Note that Eq.(3) is an extension of Eq.(2) and can be referred to as relative entropy [8]. Thus, the KL divergence is equal to zero when the probability mass distributions are exactly the same. These definitions are applicable to scalar random variables, but can be generalized to a vector of random variables with a joint probability.

B. Information Theory for Modeling

The goal of modeling is to preserve the maximum amount of useful information from the original system [3]. Consider a dynamical system governed by

$$\boldsymbol{q}^{n+1} = \boldsymbol{f}(\boldsymbol{q}^n), \tag{4}$$

where we have $\boldsymbol{q}^n = [\boldsymbol{q}^n_1, \dots, \boldsymbol{q}^n_j, \dots, \boldsymbol{q}^n_N]$ as the state vector at time t_n with N being the total number of degrees of freedom in Eq. (4). We define $\boldsymbol{f}$ as the function that advances the state of the system through time. We consider the partition of the phase-space of the full system

$$\boldsymbol{q}^n = [\widetilde{\boldsymbol{q}}^n, \boldsymbol{q}'^n]. \tag{5}$$

We aim to model the subset of the phase-space denoted by $\widetilde{\boldsymbol{q}}^n = [q^n_1, \dots, q^n_{\widetilde{N}}]$ in Eq. (5), where $\widetilde{N} < N$ are the degrees of freedom of the model. Since $\widetilde{\boldsymbol{q}}^n$ is the state to be modeled, $\boldsymbol{q}'^n$ represents the inaccessible degrees of freedom that must be accounted for by the model. The dynamics of the modeled state is governed by

$$\widetilde{\boldsymbol{q}}^{n+1} = \widetilde{\boldsymbol{f}}(\widetilde{\boldsymbol{q}}^n, \boldsymbol{q}'^n), \tag{6}$$

where $\widetilde{\boldsymbol{f}}$ are the components of $\boldsymbol{f}$ corresponding to the states $\widetilde{\boldsymbol{q}}^n$. We now consider a model with access to the information contained in $\widetilde{\boldsymbol{q}}^n$ but not to the information in $\boldsymbol{q}'^n$. The governing equation for the model is denoted by

$$\widetilde{\boldsymbol{q}}^{n+1}_{\text{model}} = \widetilde{\boldsymbol{f}}_{\text{model}}(\widetilde{\boldsymbol{q}}^n), \tag{7}$$

where $\widetilde{\boldsymbol{q}}^{n+1}_{\mathrm{model}}$ is the model prediction, which does not need to match the exact solution $\widetilde{\boldsymbol{q}}^{n+1}$ from Eq. (6). We want to find $\widetilde{\boldsymbol{f}}_{\mathrm{model}}$ that predicts the future state to within some error ε defined as

$$||\widetilde{\boldsymbol{q}}^{n+1}_{\mathrm{model}} - \widetilde{\boldsymbol{q}}^{n+1}|| \leq \varepsilon, \tag{8}$$

where $||\cdot||$ is the L_1 norm. This error constraint can be relaxed to

$$||p(\widetilde{\boldsymbol{q}}^{n+1}_{\mathrm{model}}) - p(\widetilde{\boldsymbol{q}}^{n+1})|| \leq \varepsilon, \tag{9}$$

where $p(\widetilde{\boldsymbol{q}}^{n+1})$ is the true probability distribution of the system state and $p(\widetilde{\boldsymbol{q}}^{n+1}_{\mathrm{model}})$ is the probability distribution of the model state. Note that this constraint is weaker in that it is possible that a model can replicate the statistics of the actual state, yet the dynamics of the model may not coincide with the true state. Nevertheless, the error defined by Eq. (9) yields an estimated upper bound for the expectation of the modeling error of probabilities.

We reintroduce the KL divergence between $p(\widetilde{\boldsymbol{q}}^{n+1})$ and $p(\widetilde{\boldsymbol{q}}^{n+1}_{\mathrm{model}})$ as

$$\mathrm{KL}(\widetilde{\boldsymbol{q}}^{n+1}, \widetilde{\boldsymbol{q}}^{n+1}_{\mathrm{model}}) = \sum p(\widetilde{\boldsymbol{q}}^{n+1}) \log \frac{p(\widetilde{\boldsymbol{q}}^{n+1})}{p(\widetilde{\boldsymbol{q}}^{n+1}_{\mathrm{model}})}, \tag{10}$$

with $\mathrm{KL}(\widetilde{\boldsymbol{q}}^{n+1}, \widetilde{\boldsymbol{q}}^{n+1}_{\mathrm{model}}) = 0$ if and only if the model predictions are statistically identical to those from the original system. It can be shown via the Pinsker's inequality [9] that

$$\mathrm{KL}(\widetilde{\boldsymbol{q}}^{n+1}, \widetilde{\boldsymbol{q}}^{n+1}_{\mathrm{model}}) \geq \frac{1}{2\ln 2} ||p(\widetilde{\boldsymbol{q}}^{n+1}_{\mathrm{model}}) - p(\widetilde{\boldsymbol{q}}^{n+1})||^2. \tag{11}$$

Substituting to reintroduce the prescribed error for the upper bound gives

$$\mathrm{KL}(\widetilde{\boldsymbol{q}}^{n+1}, \widetilde{\boldsymbol{q}}^{n+1}_{\mathrm{model}}) \geq \frac{\varepsilon^2}{2\ln 2}. \tag{12}$$

Equation (12) therefore yields a connection between information loss and probabilistic model performance. The model $\widetilde{\boldsymbol{f}}_{\mathrm{model}}$ that we seek will minimize Eq. (10), thus containing the coherent information in the data, while disregarding the incoherent noise. This formulation can be extended to the problem of modeling for turbulent flows.

C. Information-Preserving Subgrid-Scale Model for Large-Eddy Simulation

In LES, only the large eddies are resolved by the grid, whereas the information from the small-scale eddies is lost. Instead, the effect of the small scales on the larger eddies is modeled through an SGS model. This approach reduces computational cost while still capturing statistical quantities of interest. The previous formulation can therefore be applied in devising an SGS model for LES. The governing equations for LES for compressible flows are formally derived by applying a spatial filter to the Navier-Stokes equations for mass and momentum

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho u_j}{\partial x_j} = 0, \tag{13a}$$

$$\frac{\partial \rho u_i}{\partial t} + \frac{\partial \rho u_i u_j}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \frac{\partial \sigma_{ij}}{\partial x_j}, \tag{13b}$$

where ρ is the density, u_i is the velocity component in the i^{th} direction, p is the pressure, and σ_{ij} is the stress-tensor. We then employ the filter operator on a variable ϕ for scale separation, defined as

$$\overline{\phi}(\boldsymbol{x}, t) \equiv \int_V G(\boldsymbol{x} - \boldsymbol{x'}; \overline{\Delta}) \phi(\boldsymbol{x'}, t) \mathrm{d}\boldsymbol{x'}, \tag{14}$$

where G is the filter kernel with filter size $\bar{\Delta}$ and V is the domain of integration. Applying this spatial filter to Eq. (13) yields the governing equations for LES for compressible flows as

$$\frac{\partial \overline{\rho}}{\partial t} + \frac{\partial \bar{\rho} \bar{u}_j}{\partial x_j} = 0, \tag{15a}$$

$$\frac{\partial \bar{\rho} \bar{u}_i}{\partial t} + \frac{\partial \bar{\rho} \bar{u}_i \bar{u}_j}{\partial x_j} = -\frac{\partial \overline{p}}{\partial x_i} + \frac{1}{Re} \frac{\partial \overline{\sigma}_{ij}}{\partial x_j} - \frac{\partial \tau^{\mathrm{SGS}}_{ij}}{\partial x_j}, \tag{15b}$$

where $\bar{(\cdot)}$ denotes spatially filtered quantity and τ_{ij}^{SGS} is the effect of the subgrid scales on the resolved eddies, which has to be modeled. The system in Eq. (15) is assumed to have a severely truncated number of degrees of freedom with respect to Eq. (13). The objective of LES is to model the SGS tensor as a function of known filtered quantities defined by

$$\tau_{ij}^{\mathrm{SGS}} = \tau_{ij}^{\mathrm{SGS}}(\overline{S}_{ij}, \overline{\Omega}_{ij}, \overline{\Delta}; \boldsymbol{\theta}), \tag{16}$$

where $\overline{S}_{ij}$ and $\overline{\Omega}_{ij}$ are the filtered rate-of-strain and rate-of-rotation tensors, respectively, and $\boldsymbol{\theta}$ are model parameters.

The map $\boldsymbol{f}$ in Eq. (4) corresponds to a discrete version of Eq. (13), resolving all the space and time scales. The state vector $\boldsymbol{q}^n$ is given by the discretization of the velocities and pressure in a grid fine enough to capture all the relevant scales of motion. The map $\widetilde{\boldsymbol{f}}_{\mathrm{model}}$ for the model is derived from the discretization of Eq. (4). The model state vector $\widetilde{\boldsymbol{q}}^n_{\mathrm{model}}$ therefore corresponds to the filtered velocities and pressure.

We start from the general expansion of the SGS tensor in terms of $\overline{S}_{ij}$ and $\overline{\Omega}_{ij}$ proposed by Lund and Novikov [10]. Retaining the two leading terms for the functional form considered for the SGS stress tensor results in

$$\tau_{ij}^{\mathrm{SGS}} - \frac{1}{3}\tau_{kk}^{\mathrm{SGS}}\delta_{ij} = \theta_1 \overline{\Delta}^2 \overline{S}_{ij}\sqrt{\overline{S}_{nm}\overline{S}_{nm}} + \theta_2 \overline{\Delta}^2 (\overline{S}_{ik}\overline{\Omega}_{kj} - \overline{\Omega}_{ik}\overline{S}_{kj}), \tag{17}$$

where δ_{ij} is the Kronecker delta, and θ_1 and θ_2 are modeling parameters to be determined. We also introduce the interscale energy transfer and viscous dissipation in the case of LES at the filter cut-off $\overline{\Delta}$ given by

$$\overline{\Gamma} = (\overline{u_i u_j} - \bar{u}_i \bar{u}_j)\overline{S}_{ij} - 2\nu \overline{S}_{ij}\overline{S}_{ij} + \tau_{ij}^{\mathrm{SGS}}\overline{S}_{ij}. \tag{18}$$

We invoke the modeling assumption that the information content content of $p(\overline{\Gamma}_1)$ must be equal to the information content of $p(\overline{\Gamma}_2\gamma)$, where $\overline{\Gamma}_1$ and $\overline{\Gamma}_2$ are $\overline{\Gamma}$ at scale $\overline{\Delta}_1$ and $\overline{\Delta}_2$, respectively, and $\gamma = (\overline{\Delta}_1/\overline{\Delta}_2)^{2/3}$ is a scaling factor. The model aims at minimizing the information lost when $p(\overline{\Gamma}_1)$ is used to approximate $p(\overline{\Gamma}_2\gamma)$ in the LES solution. Thus, it follows that the model is formulated using the KL divergence, ensuring that the average information required for reconstructing $p(\overline{\Gamma}_2\gamma)$ is minimum given the information in $p(\overline{\Gamma}_1)$ given by

$$\boldsymbol{\theta} = \arg\min_{\boldsymbol{\theta}'} \mathrm{KL}(\overline{\Gamma}_2\gamma, \overline{\Gamma}_1), \tag{19}$$

where $\boldsymbol{\theta} = (\theta_1, \theta_2)$ from Eq. (17). This model will be referred to as the information-preserving (IP) SGS model. For more details on the technical approach outlined above, the reader is referred to Lozano-Durán and Arranz [3] for an in-depth exposition of the information-theoretic formulation of modeling for high-dimensional, chaotic dynamical systems.

III. Computational Setup

A. Compressible Channel

The simulation is performed using wall-modeled large-eddy simulation (WMLES) with the code charLES from Cascade Tech., Inc. The validation of the algorithm can be found in Fu et al. [11]. The solver integrates the filtered Navier-Stokes equations using a second-order accurate finite volume formulation. The numerical discretization relies on a flux formulation that is approximately entropy preserving in the inviscid limit, thereby limiting the amount of numerical dissipation added into the calculation. The time integration is performed with a third-order Runge-Kutta explicit method. The mesh generator is based on a Voronoi hexagonal close packed point-seeding method which automatically builds high-quality meshes for arbitrarily complex geometries with minimal user input.

An equilibrium wall model (EQWM) is used to overcome the restrictive grid-resolution requirements to resolve the small-scale flow motions in the vicinity of the walls. The wall stress is obtained from an algebraic wall model derived from the integration of the one-dimensional equilibrium stress model along the wall-normal direction. The EQWM is an ODE-based wall-stress-model with simplified momentum and total energy equations given by [12]

$$\frac{\partial}{\partial y}\left[(\mu + \mu_t)\frac{\partial u}{\partial y}\right] = 0, \tag{20a}$$

$$\frac{\partial}{\partial y}\left[(\mu + \mu_t)u\frac{\partial u}{\partial y} + c_p\left(\frac{\mu}{Pr} + \frac{\mu_t}{Pr_t}\right)\frac{\partial T}{\partial y}\right] = 0, \tag{20b}$$

where μ is the dynamic viscosity, Pr is the Prandtl number, and the subscript "t" denotes a turbulent quantity. The eddy-viscosity is taken from a mixing-length model as

$$\mu_t = \kappa \rho y \sqrt{\frac{\tau_w}{\rho}} \left[1 - \exp\left(\frac{y^+}{A^+} \right) \right]^2, \tag{21}$$

where τ_w is the instantaneous wall stress and the superscript "+" denotes "plus" units (i.e., normalization by viscous wall quantities). The model parameters κ, Pr, and A^+ are constants. More details on this wall model can be found in the work by Larsson and Kawai [12]. Effectively, given an instantaneous velocity at some height above the wall, the model estimates the instantaneous wall stress and heat flux. The no-slip boundary condition at the walls is replaced instead by a wall-stress boundary condition. The wall is assumed to be isothermal.

The simulation is performed in a canonical compressible channel flow at the bulk Reynolds number $Re_b = 24$K and bulk Mach number $M_b = 3.0$ such that results can be compared to direct numerical simulation (DNS) data from Trettel and Larsson [13]. The simulation has an isotropic grid resolution of approximately $2\delta/24$, where δ is the channel half-height, with 24^3 total points. Figure 1 shows a schematic for channel flow, where $\langle u \rangle$ is the mean velocity along the x-direction.

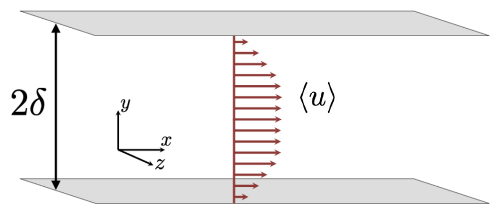

Fig. 1 Canonical channel flow.

The statistical quantities of interest are the mean velocity and temperature profiles. The performance of the IP SGS model is evaluated in the prediction of these statistical quantities of interest of DNS results and compared against the predictions provided by the dynamic Smagorinsky (DSM) SGS model.

IV. Results

The mean velocity and temperature profiles predicted by the WMLES are given in Fig. 2a and Fig. 2b, respectively. DNS data is plotted as a black dashed line for a reference. The profiles are predicted using an equilibrium wall model and either IP or DSM SGS model. The velocity is normalized using the centerline velocity value u_c, while the temperature is normalized using the temperature at the wall T_w.

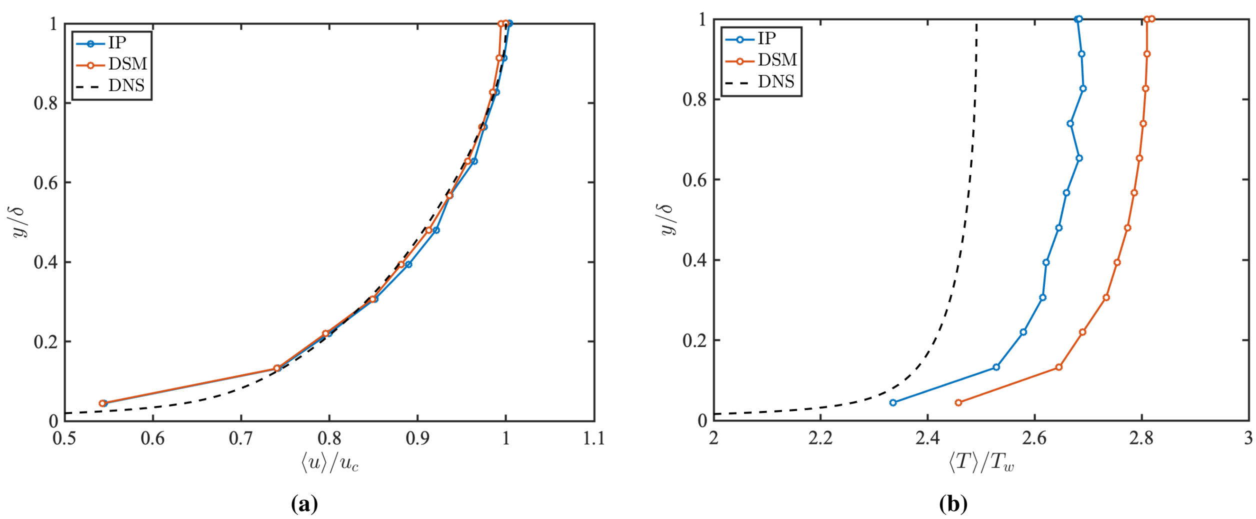

Fig. 2 (a) Mean velocity profiles. (b) Mean temperature profiles.

The results show that the IP SGS model offers an accuracy comparable to the DSM SGS model in the prediction of the mean velocity profile. For the mean temperature profile, the prediction given by the IP SGS model is improved compared with the DSM SGS model. The error for each statistical quantity of interest can be systematically quantified as the average difference between the LES and DNS solutions defined as [1]

$$\varepsilon_q = \left[\frac{\int_{0.2\delta}^{\delta} \left(\langle q_{\mathrm{LES}} \rangle - \langle q_{\mathrm{DNS}} \rangle\right)^2 \,\mathrm{d}y}{\int_{0.2\delta}^{\delta} \langle q_{\mathrm{DNS}} \rangle^2 \mathrm{d}y} \right]^{1/2}. \tag{22}$$

Note that Eq. (22) excludes the region close to the wall below 0.2δ [14]. Table 1 shows the percent error in the prediction of the mean velocity and temperature profiles as calculated using Eq. (22). Indeed, it is found that the IP and DSM SGS models both yield predictions of the mean velocity profile within 1% of DNS data. Further, there is 5% improvement in accuracy in the prediction of the temperature profile using the IP SGS model compared to the DSM SGS model.

Table 1 Percent error in mean velocity and temperature profiles

	DSM	IP
ε_u (%)	0.4	0.8
ε_T (%)	12.4	7.3

V. Conclusion

The information-theoretic formulation presented in this work is shown to be an effective framework for reduced-order modeling of highly chaotic systems with a large number of degrees of freedom. In the case of LES, an SGS model was developed based on minimizing the KL divergence between the probabilities of the model state and the true state, enabling the accurate prediction of the statistical quantities of interest. The model aims at minimizing the information lost between the probability mass distribution of the interscale energy transfer and viscous dissipation at different scales. We have shown that the IP SGS model offers comparable accuracy as the DSM SGS model in the prediction of the mean velocity profile given by DNS data. We have also shown that the IP SGS model yields an improved prediction of the mean temperature profile given by DNS data compared to the DSM SGS model.

Acknowledgments

The authors acknowledge the MIT SuperCloud and Lincoln Laboratory Supercomputing Center for providing HPC resources that have contributed to the research results reported within this work. This work was supported by the National Science Foundation under Grant No. 032707-00001.

References

[1] Williams, E., and Lozano-Durán, A., "Error scaling of wall-modeled large-eddy simulation of compressible wall turbulence," Bulletin of the American Physical Society 66, 2021.

[2] Landauer, R., "The physical nature of information," *Phys. Lett. A.*, Vol. 217, 1996, pp. 188–193. https://doi.org/10.1016/0375-9601(96)00453-7.

[3] Lozano-Durán, A., and Arranz, G., "Information-theoretic formulation of dynamical systems: causality, modeling, and control," 2022. https://doi.org/10.48550/arXiv.2111.09484.

[4] Lozano-Durán, A., Bae, H., and Encinar, M., "Causality of energy-containing eddies in wall turbulence," *J. Fluid Mech.*, Vol. 882, 2019, p. A2. https://doi.org/10.1017/jfm.2019.801.

[5] Akaike, H., "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, Vol. 19, 1974, pp. 716–723. https://doi.org/10.1109/TAC.1974.1100705.

[6] Shavit, M., and Falkovich, G., "Singular Measures and Information Capacity of Turbulent Cascades," *Phys. Rev. Lett.*, Vol. 125, 2020. https://doi.org/10.1103/PhysRevLett.125.104501.

[7] Shannon, C., "A Mathematical Theory of Communication," *Bell Syst. Tech. J.*, Vol. 27, 1948, pp. 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.

[8] Hobson, A., and Cheng, B., "A Comparison of the Shannon and Kullback Information Measures," *J. Stat. Phys.*, Vol. 7, 1973, pp. 301–310. https://doi.org/10.1007/BF01014906.

[9] Weissman, T., Ordentlich, E., Seroussi, G., Verdu, S., and Weinberger, M., "Inequalities for the L_1 Deviation of the Empirical Distribution," *Hewlett-Packard Labs, Tech. Rep*, 1973.

[10] Lund, T. S., and Novikov, E. A., "Parameterization of subgrid-scale stress by the velocity gradient tensor," *Center for Turbulence Research, Annual Research Briefs*, 1992.

[11] Fu, L., Karp, M., Bose, S. T., Moin, P., and Urzay, J., "Shock-induced heating and transition to turbulence in a hypersonic boundary layer," *J. Fluid Mech.*, Vol. 909, 2020, p. A8. https://doi.org/10.1017/jfm.2020.935.

[12] Larsson, J., and Kawai, S., "Wall-modeling in large eddy simulation: length scales, grid resolution and accuracy," *Phys. Fluids*, Vol. 24, 2012. https://doi.org/10.1063/1.3678331.

[13] Trettel, A., and Larsson, J., "Mean velocity scaling for compressible wall turbulence with heat transfer," *Phys. Fluids*, Vol. 28, 2016. https://doi.org/10.1063/1.4942022.

[14] Lozano-Durán, A., and Bae, H. J., "Error scaling of large-eddy simulation in the outer region of wall-bounded turbulence," *Journal of Computational Physics*, Vol. 392, 2019, pp. 532–555. https://doi.org/10.1016/j.jcp.2019.04.063.