From Deep to Physics-Informed Learning of Turbulence: Diagnostics

Ryan King [*]   Oliver Hennigh [†]   A. Mohan & M. Chertkov [‡]

Abstract

We describe tests validating progress made toward acceleration and automation of hydrodynamic codes in the regime of developed turbulence by three **Deep Learning** (DL) Neural Network (NN) schemes trained on **Direct Numerical Simulations** of turbulence. Even the bare DL solutions, which do not take into account any physics of turbulence explicitly, are impressively good overall when it comes to qualitative description of important features of turbulence. However, the early tests have also uncovered some caveats of the DL approaches. We observe that the static DL scheme, implementing Convolutional GAN and trained on spatial snapshots of turbulence, fails to reproduce intermittency of turbulent fluctuations at small scales and details of the turbulence geometry at large scales. We show that the dynamic NN schemes, namely LAT-NET and Compressed Convolutional LSTM, trained on a temporal sequence of turbulence snapshots are capable to correct for the caveats of the static NN. We suggest a path forward towards improving reproducibility of the large-scale geometry of turbulence with NN.

1   Introduction

In this manuscript, reporting first results of the approach focusing on developing a Physics-Informed Machine Learning (PIML) framework to improve turbulence model implementations in hydrodynamic codes, see also the Supplementary Material (SM) A, we aim to verify whether various statistical properties, constraints, and relations not enforced explicitly within the DL training on the ground truth, Direct Numerical Simulation (DNS), data hold. To do this, we compare results extracted from the training data and from the generated/synthetic data. Three DL schemes, GAN of (1), LAT-NET of (2), generalized to arbitrary solvers, and newly developed Compressed Convolutional Long Short-Term Memory (CC-LSTM) scheme are juxtaposed within the setting of the homogeneous, isotropic, stationary turbulence. Specifically, we have verified (a) spectrum of energy fluctuations over scales, (b) statistics of velocity gradient (small/viscous scale object), (c) anomalous exponents of higher order velocity increments, and (d) statistics of the coarse-grained velocity-strain alignment represented in the plane of the coarse-grained velocity gradient tensor. Logic behind and details of the diagnostics, suggested in our prior work on the subject (1), are described in the SM B. Static GAN DL scheme is trained on the spatial data (instantaneous snapshots) from the Johns Hopkins turbulence database (3). The dynamic schemes, LAT-NET and CC-LSTM, are trained on the dynamic (sequence of snapshots) data from the SpectralDNS code (4). Description of the DL schemes and results of applying the two step diagnostics to the the DL schemes are described in Section 2. We draw conclusions and suggest a path forward in Section 3.

---

[*]National Renewable Energy Laboratory, Golden, CO, USA (e-mail: Ryan.King@nrel.gov)

[†]Los Alamos National Laboratory, X Computational Physics-4 & CNLS, Los Alamos, NM 87545, USA (e-mail: hennigo@lanl.gov)

[‡]Los Alamos National Laboratory, Theoretical Division T-4 & CNLS, Los Alamos, NM 87545, USA (e-mails: {arvindm,chertkov}@lanl.gov)
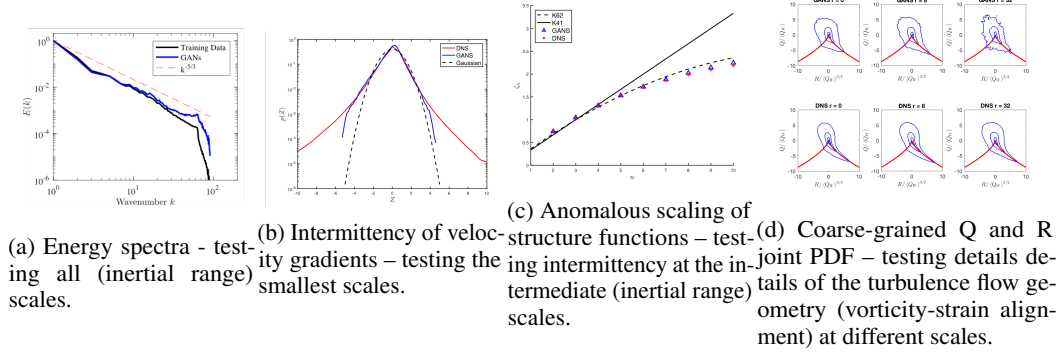
(a) Energy spectra - testing all (inertial range) scales.

(b) Intermittency of velocity gradients – testing the smallest scales.

(c) Anomalous scaling of structure functions – testing intermittency at the intermediate (inertial range) scales.

(d) Coarse-grained Q and R joint PDF – testing details details of the turbulence flow geometry (vorticity-strain alignment) at different scales.

Figure 1: Turbulence diagnostics showing the GANs output captures many characteristics of real turbulence. In panel (a) the GANs output preserves the kinetic energy spectra of the training data except for a bit of extra energy at the smallest length scales. Panel (b) shows a PDF of the normalized velocity gradients that correctly captures the non-Gaussian and negative skewness that is known to characterize intermittency in turbulence. Panel (c) shows the GANs correctly captures anomalous scaling exponents of higher order structure functions. Finally, panel (d) shows the classic teardrop shape of the joint PDF of the Q and R invariants of the velocity gradient tensor, testing details of flow geometry (8).

2 Deep Learning for Turbulence

Generating Turbulence through Adversarial Training. Recent advances in DL have proven remarkably successful on image-based problems such as generation, classification, and denoising. Much of the success of these DL techniques relies on the hierarchical identification and abstraction of features present in images using deep networks. A complex superposition of structures also occurs in turbulent flows, naturally leading us to examine whether techniques developed for images can be used to learn turbulent flow physics. It was demonstrated in (1) that Generative Adversarial Networks (GANs) can be trained on high fidelity DNS to generate high-quality synthetic images of turbulent flows. GAN was originally introduced by Goodfellow et al, (5), and relies on competition between two deep Neural Networks (NNs). The first network, the generator, attempts to create synthetic images that appear to be drawn from the correct high dimensional distribution. The second network, the discriminator, compares the proposed images against a set of training data and estimates the likelihood of the image being genuine. A min-max optimization problem is solved iteratively to train both networks' parameters as they compete against each other. Over time, the generator learns to draw more plausible images while the discriminator's ability to distinguish real and synthetic images also improves. Our GAN implementation makes use of the Convolution (C-GAN) for both the generator and discriminator using the architecture guidelines for stable deep C-GANs recommended by Radford et al (6) and Reed et al (7). We train on 2D slices of 3D homogeneous isotropic turbulence from the Johns Hopkins Turbulence Database (3). The three velocity components map naturally to different red, green, and blue color channels in imaged-based deep learning networks. We calculated a number of common turbulence quantities on the GANs output as part of our survey of DL techniques. The results, summarized in Figure 1, suggest that the generator has successfully learned to sample from the full multi-point Probability Distribution Functions (PDFs) of valid turbulent flow fields.

LAT-NET. Dynamic approach to modeling turbulence is to make temporal predictions given an initial state of the flow. Toward this end, the test was a modified version of LAT-NET (2). LAT-NET works by compressing a snapshot of the flow onto a compact latent space and then mimicking the dynamics of the flow on this latent space. Doing this allows the network to generate the flow simulation for significantly less memory and computation then the flow solver. A convolutional autoencoder is used to create the mapping to and from the latent space while a separate convolutional network is applied iteratively on the latent space to reproduce the dynamics of the underlying flow. Fig. 2 presents the scheme of operations where the mapping on the latent space is referred to as the latent mapping. Each application of the latent mapping is meant to ditto a fixed amount of time in the simulation. In our case, each compression mapping corresponds to 0.06 sec in the simulation thus applying the mapping 10 times propagates the predicted flow forward in time 0.6 seconds. The Lat-Net method was first
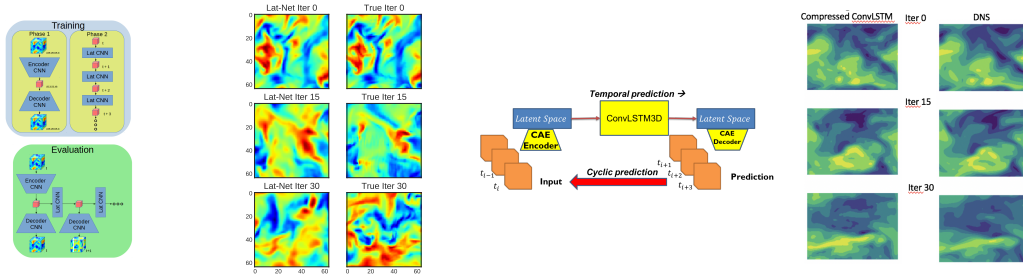
Figure 2: LAT-NET (left) & CC-LSTM (right): (a) training and evaluation & (b) visualization of predicted flow.

proposed to predict lattice Boltzmann fluid simulations however we have relaxed this constraint so that the network can be used on flow data regardless of the underlying solver. We have also removed the boundary conditions as our tests are on homogeneous flow with no physical boundaries. Training the network is also modified from the original work by splitting the training process into two phases. First, the autoencoder is trained on snapshots of the data and then frozen. Next, the latent mapping is trained on the compressed data in a sequential fashion in Fig. 2. Breaking up the training process both speeds up learning and reduces the memory usage allowing larger domains to be trained on. Results of the four diagnostic tests, described above and SM B, applied to the dynamic NN LAT-NET are shown in Fig. 3.

Compressed Convolutional LSTM. Here we describe another dynamic scheme based on the Convolutional LSTM by Shi (27), exploiting the strength of the convolutional networks (which captures spatial features) and Long Short Term Memory i.e. LSTM networks (which capture temporal patterns) together in a combined architecture, and also extending it to the case of 3d images. To reduce dimensionality of the data we also compress and decompress input and output images, respectively, by means of two 3d convolutional autoencoders. The resulting Compressed Convolutional LSTM (CC-LSTM) architecture is illustrated in Fig. 2. The results from the CC-LSTM scheme can be seen in Fig. 3. Results for energy spectra and for statistics of the velocity gradient show good-to-reasonable consistency between the synthetic and training data. QR diagnostics show good reproduction by the synthetic data of the training data trends at the larger scales. Respective deviations are significant at the smaller scales – this is a parasitic effect attributed to insufficient spatial resolution of the spatio-temporal (thus more demanding then pure spatial) schemes.

3   Conclusions and Path Forward

In this manuscript we tested three DL schemes in their ability to reproduce turbulent flows. We introduce four tests of increasing complexity. Our first scheme, implementing C-GAN, is static, i.e. it takes as an input only instantaneous snapshots of the turbulence field, completely ignoring any dynamical aspects of turbulence. The scheme paths our first test, testing distribution of energy over scales, with flying colors. The second test, checking statistics of the velocity gradient which is a small scale object, reveals that GAN underestimates tails of the PDF of the velocity gradient, making generated samples less intermittent then the DNS input. Third test, of the intermittency (non-Gaussianity) at the intermediate scales (from inertial range of turbulence) via analysis of velocity increments, was passed by the C-GAN scheme well. Finally, the most elaborate fourth test, consisting in checking statistics of the coarse-grained velocity tensor, reveals a difficulty of GAN to reproduce details of the vorticity-strain alignment at the largest (energy containing) scales of turbulence. Based on the analysis of the static NN scheme we naturally came to the question – if dynamic NNs, trained on temporal sequence of snapshots, is capable of correcting for the static scheme deficiencies in reproducing (a) intermittency of the smallest (close to viscous) scales, and (b) statistics of the vorticity-strain alignment of the largest (close to the energy containing) scales?

To resolve the question we analyzed the dynamic LAT-NET and CC-LSTM NN schemes trained on temporal sequence of turbulence snapshots. We observed that the dynamic schemes perform on the three-test diagnostics (excluding the anomalous scaling test) at least as well as the static
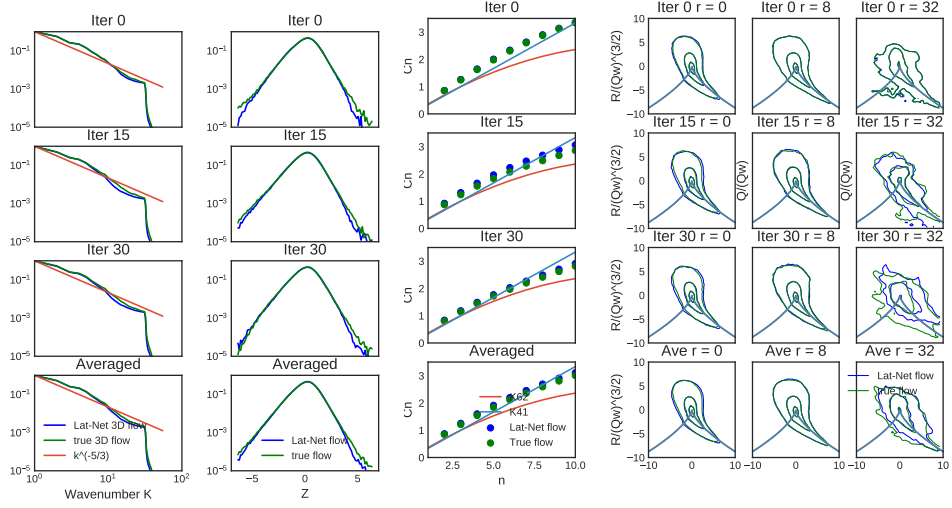
(a) Energy spectra. (b) Intermittency. (c) Anomalous Scaling. (d) Coarsed-grained Q and R joint PDF's.
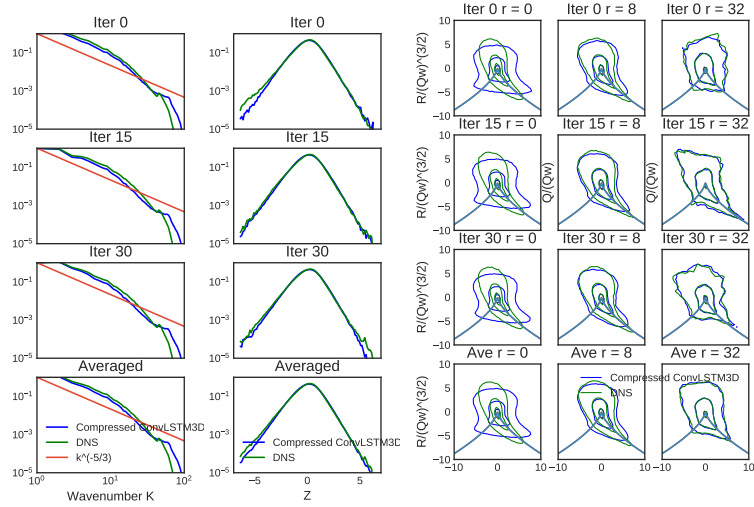
(e) Energy spectra. (f) Intermittency. (g) Coarsed-grained Q and R joint PDF's.

Figure 3: Results of the turbulence diagnostics tests for the dynamic NN LAT-NET (first row) and CC-LSTM (second raw) schemes. Diagnostics is static (applied to 3d snapshots). Notations are equivalent to ones used in Fig. 1. See also SM B for description of the details.

NN scheme. Moreover, we discovered that the small scale intermittency (feature (a) above) is now reproduced well, therefore correcting for the caveat of the static NN. We attribute this success of the dynamic schemes to the direct cascade nature of turbulence – turbulent dynamics makes small scale statistics universal, i.e. minimally sensitive to larger scale details of the vorticity-strain alignment. We also note that CC-LSTM is reproducing geometry of the flow better at the larger scales than at the smaller scales, while LAT-NET show the opposite tendency. Sub-optimal performance of the two dynamic schemes in the anomalous scaling test is attributed to problems with insufficient spatial resolution. This analysis suggests that dynamic networks bring in additional information which improves reproduction of intermittency at smaller scales and geometry of the flow at large scales.

References

[1] R. King, P. Graf, and M. Chertkov. Creating turbulent flow realizations with generative adversarial networks. *70th Annual Meeting of the APS Division of Fluid Dynamics*, 62(14), 2017.

[2] O. Hennigh. Lat-Net : Compressing Lattice Boltzmann Flow Simulations using Deep Neural Networks arXiv :. *arXiv:1705.09036*, 2017.

[3] Johns hopkins turbulence database. `http://turbulence.pha.jhu.edu/`, 2008.

[4] Spectral DNS. `https://github.com/spectralDNS/spectralDNS`, 2016.

[5] I. Goodfellow, J. Pouget-Abadie, and M. Mirza. Generative Adversarial Networks. *arXiv preprint arXiv: . . .* , pages 1–9, 2014.

[6] A. Radford, L. Metz, and S. Chintala. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks, 2016.

[7] S. Reed, Z. Akata, X. Yan, L. Logeswaran, B. Schiele, and H. Lee. Generative Adversarial Text to Image Synthesis. *arXiv:1605.05396*, May 2016.

[8] M. Chertkov, A. Pumir, and B. I. Shraiman. Lagrangian tetrad dynamics and the phenomenology of turbulence. *Physics of Fluids*, 11(8):2394–2410, 1999.

[9] A. Barati Farimani, J. Gomes, and V. S. Pande. Deep Learning the Physics of Transport Phenomena. *arXiv:1709.02432*, 94305, 2017.

[10] T. P. Miyanawala and R. K. Jaiman. An Efficient Deep Learning Technique for the Navier-Stokes Equations: Application to Unsteady Wake Flow Dynamics. *arXiv:1710.09099*, 2017.

[11] R. Maulik and O. San. A neural network approach for the blind deconvolution of turbulent flows. *Journal of Fluid Mechanics*, 831:151–181, 2017.

[12] Y. Xie, E. Franz, M. Chu, and N. Thuerey. Tempogan: A temporally coherent, volumetric gan for super-resolution fluid flow. *ACM Transactions on Graphics (TOG)*, 37(4):95, 2018.

[13] S. Wiewel, M. Becher, and N. Thuerey. Latent-space Physics: Towards Learning the Temporal Evolution of Fluid Flow. *arXiv:1802.10123*, February 2018.

[14] A. T. Mohan and D. V. Gaitonde. A Deep Learning based Approach to Reduced Order Modeling for Turbulent Flow Control using LSTM Neural Networks. *arXiv:1804.09269*, April 2018.

[15] E. J. Parish and K. Duraisamy. A paradigm for data-driven predictive modeling using field inversion and machine learning. *Journal of Computational Physics*, 305:758–774, 2016.

[16] K. Duraisamy. A Framework for Turbulence Modeling using Big Data. 2015.

[17] B. Tracey, K. Duraisamy, and J. Alonso. Application of Supervised Learning to Quantify Uncertainties in Turbulence and Combustion Modeling. In *51st AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, 2013.

[18] J.-X. Wang, J.-L. Wu, and H. Xiao. A Physics Informed Machine Learning Approach for Reconstructing Reynolds Stress Modeling Discrepancies Based on DNS Data. *Physical Review Fluids*, pages 1–36, 2016.

[19] J.-X. Wang, J.-L. Wu, and H. Xiao. Physics-informed machine learning approach for reconstructing reynolds stress modeling discrepancies based on dns data. *Phys. Rev. Fluids*, 2:034603, Mar 2017.

[20] A. N. Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large reynolds numbers. *Dokl. Akad. Nauk SSSR*, 30:299–303, 1941.

[21] A. N. Kolmogorov. On the degeneration of isotropic turbulence in an incompressible viscous fluids. *Dokl. Akad. Nauk SSSR*, 31:538–541, 1941.

[22] A. N. Kolmogorov. Dissipation of energy in isotropic turbulence. *Dokl. Akad. Nauk SSSR*, 32:19–21, 1941.

[23] G. Falkovich. *A Voyage Through Turbulence*, chapter Kolmogorov and the Russian school of turbulence. Cambridge University Press, 2011.

[24] A. N. Kolmogorov. A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incom-pressible fluid at high reynolds number. *Journal of Fluid Mechanics*, 13:82–85, 1962.

[25] U. Frisch. *Turbulence: The Legacy of A.N. Kolmogorov*. Cambridge University Press, 1995.

[26] C. Meneveau. Lagrangian dynamics and models of the velocity gradient tensor in turbulent flows. *Annual Review of Fluid Mechanics*, 43(1):219–245, 2011.

[27] Xingjian, S. H. I.. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems.*, 2015.

Supplementary Materials (SM) for "From Deep to Physics-Informed Learning of Turbulence: Diagnostics"

In the two supplemental Sections we give some additional details on the general Physics Informed Machine Learning approach in Section A, this manuscript is the first step of, and then in the Section B we present the primer on physics rationale behind the turbulence metric used to diagnose the three ML schemes in the main part of the manuscript.

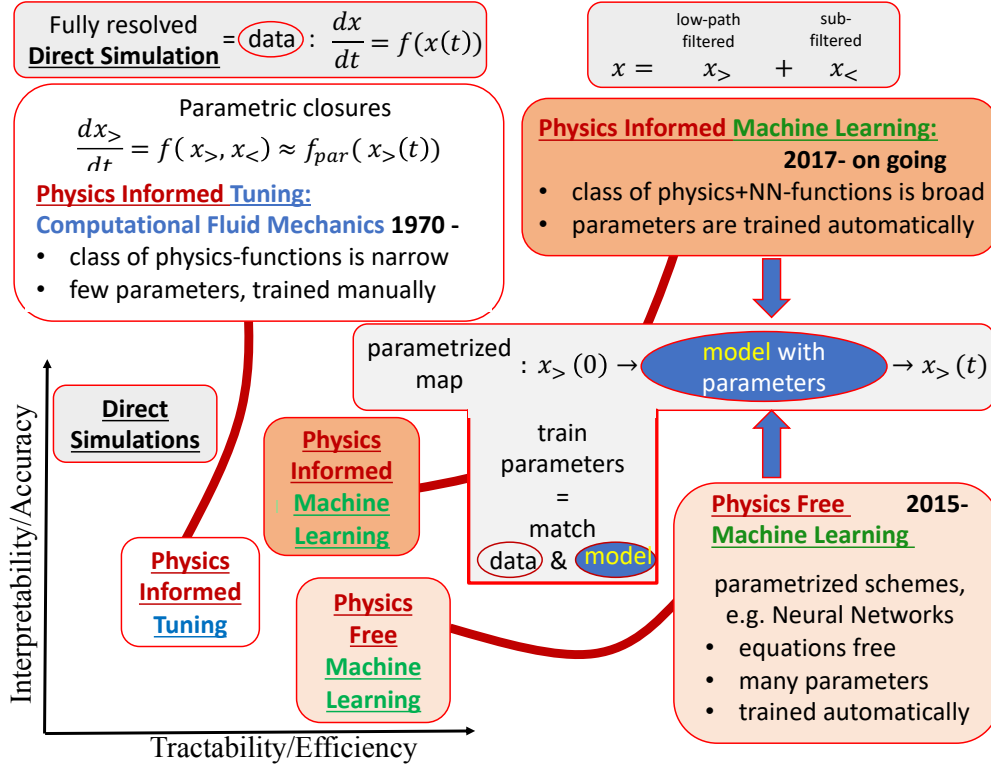## A Physics Informed Machine Learning: The Methodology

Figure 4: Essence of the PIML framework.

The essence of the Physics Informed Learning (PIML), illustrated in Fig. (4), is in capturing physics with the right amount of domain specificity and interpretability of the Physics Informed Tuning (PIT) non-automatic approach, expressed e.g. in symmetries and constraints, while also providing prediction power and computational tractability which is on pier with state of the art Machine Learning, such as Deep Learning (DL), which is typically Physics Free, thus PFML. Development of the PIML approach, which we expect to be transferable to many areas of natural and engineering sciences, is anchored in this manuscript to fluid mechanics. Majority of fluid flows of interest are turbulent, i.e. containing rich spatio-temporal correlations. Goverments, Industry and Science and Engineering communities continue to make significant investments in computational software and hardware to solve turbulence in multiple situations of interest. We conjecture that, when developed, the PIML approach will allow to accelerate turbulent computations by orders of magnitude through separation of scales into macro-scales that are simulated and sub-scales that are efficiently emulated based on PIML learned dynamics. Moreover, the acceleration will be automatic allowing program experts to focus on fewer ad-hoc adjustments of reduced models than is custom today. Also, since a low-dimensional sub-scale state is known the development of closure models (coupling of scales) is greatly simplified.

In the simplest setting of interest we train parameters of a ML model on the data from Direct Numerical Simulations (DNS) to predict important features of the flow faster than the DNS. Our main hypothesis is that tremendous acceleration is achievable with PIML. The hypothesis is based on a number of recent reports (1; 9; 10; 2; 11; 12; 13; 14) that DL schemes developed originally for popular IT industry applications, when applied to DNS data as is, do provide an acceleration. Developing scientific diagnostics for testing and juxtaposing the bare DL approaches and then augmenting DLs based on the results of the diagnostics becomes critical. Complementarily we expect, based on another set of recent publications (15; 16; 17; 18; 19) that a significant acceleration is also achievable with properly relaxed and parameterized fluid-mechanics phenomenological models, e.g. of Large Eddy Simulations (LES) and Reynolds Averaged Navier Stocks (RANS) types, with much fewer parameters than in the DL models. PIML connects the two approaches to extract robust winning schemes from a number of analyzed synergistic solutions.

Future PIML solutions should be customized to data from DNS of a sequence of turbulence models of increasing complexity, starting from steady homogeneous, isotropic turbulence and extending to non-stationary, non-isotropic turbulent flows involving active (chemical and nuclear) mixing. This strategy will allow us, in particular, to probe and validate transferability of the PIML models trained on simpler cases to more complex situations.

B   Turbulence Metric

In this Section we review basic statistical concepts commonly used in the modern literature to analyze results of theoretical, computational and experimental studies of homogeneous isotropic incompressible turbulence in three dimensions. Combination of these concepts are used in the main part of the manuscript as a metric to juxtapose results of the two (static and dynamic) DL methods.

We assume that a $3d$ snapshot, or its $2d$ slice, or a temporal sequence of snapshots of the velocity field, $\boldsymbol{v} = (v^{(i)}(\boldsymbol{r})|i = 1, \cdots, 3)$, is investigated. We will focus here on analysis of static correlations within the snapshots.

We consider various objects of interest, e.g. correlation functions of second, third and fourth orders

$$C_2^{(i,j)}(\mathbf{r}_1, \mathbf{r}_2) = \langle v^{(i)}(\mathbf{r}_1)v^{(j)}(\mathbf{r}_2)\rangle, \tag{1}$$

$$C_3^{(i,j,k)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \langle v^{(i)}(\mathbf{r}_1)v^{(j)}(\mathbf{r}_2)v^{(k)}(\mathbf{r}_3)\rangle, \tag{2}$$

$$C_4^{(i,j,k,l)}(\mathbf{r}_1, \cdots, \mathbf{r}_4) = \langle v^{(i)}(\mathbf{r}_1)v^{(j)}(\mathbf{r}_2)v^{(k)}(\mathbf{r}_3)v^{(l)}(\mathbf{r}_4)\rangle. \tag{3}$$

Rich tensorial structure of the correlation functions carry a lot of information. It also suggests a number of useful derived objects, each focusing on a particular feature of the turbulent flows. In particular, we may be to discuss structure functions, defined as tensorial moments of the increments between two points separated by the radius-vector $\mathbf{r}$:

$$S_2^{(i,j)}(\boldsymbol{r}) = \langle (v^{(i)}(\mathbf{r}) - v^{(i)}(\mathbf{0}))(v^{(j)}(\mathbf{r}) - v^{(j)}(\mathbf{0}))\rangle, \tag{4}$$

$$S_3^{(i,j,k)}(\boldsymbol{r}) = \langle (v^{(i)}(\mathbf{r}) - v^{(i)}(\mathbf{0}))(v^{(j)}(\mathbf{r}) - v^{(j)}(\mathbf{0}))$$
$$\times (v^{(k)}(\mathbf{r}) - v^{(k)}(\mathbf{0}))\rangle, \tag{5}$$

$$S_4^{(i,j,k,l)}(\boldsymbol{r}) = \langle (v^{(i)}(\mathbf{r}) - v^{(i)}(\mathbf{0}))(v^{(j)}(\mathbf{r}) - v^{(j)}(\mathbf{0}))$$
$$\times (v^{(k)}(\mathbf{r}) - v^{(k)}(\mathbf{0}))(v^{(l)}(\mathbf{r}) - v^{(l)}(\mathbf{0}))\rangle. \tag{6}$$

Other objects of interest, derived from the correlation function by spatial differentiation and then merging the points are moments of the velocity gradient tensor, $m^{(i,j)} = \nabla^{(i)} v^{(j)}$

$$D_2^{(i_1,i_2;j_1,j_2)} = \langle m^{(i_1,j_1)}m^{(i_2,j_2)}\rangle, \tag{7}$$

$$D_3^{(i_1,i_2,i_3;j_1,j_2,j_3)} = \langle m^{(i_1,j_1)}m^{(i_2,j_2)}m^{(i_3,j_3)}\rangle, \tag{8}$$

$$D_4^{(i_1,i_2,i_3,i_4;j_1,j_2,j_3,j_4)} = \langle m^{(i_1,j_1)}m^{(i_2,j_2)}m^{(i_3,j_3)}m^{(i_3,j_3)}\rangle. \tag{9}$$

We may also be interested to study mixed objects, e.g. the so-called energy flux

$$F(\mathbf{r}) = \langle v^{(j)}(\mathbf{0})v^{(i)}(\mathbf{r})m^{(i,j)}(\mathbf{r})\rangle. \tag{10}$$

The remainder of this Section is organized as follows. We describe important turbulence concepts mentioned in the main part of the paper one by one, starting from simpler ones and advancing towards more complex concepts.

4/5 **Kolmogorov law and the Energy Spectra** Main statement of the Kolmogorov theory of turbulence (in fact it is the only formally proved statement of the theory) is that asymptotically in the inertial range, i.e. at $L \gg r \gg \eta$, where L is the largest, so-called energy-containing scale of turbulence and η is the smallest scale of turbulence, so-called Kolmogorov (viscous) scale, $F(r)$ does not depend on r. Moreover, the so-called 4/5-law states for the third-order moment of the longitudinal velocity increment

$$L \gg r \gg \eta: \quad S_3^{(i,j,k)} \frac{r^i r^j r^k}{r^3} = -\frac{4}{5} \varepsilon r, \tag{11}$$

where $\varepsilon = \nu D_2^{(i,j;i,j)}/2$ is the kinetic energy dissipation also equal to the energy flux.

Self-similarity hypothesis extended from the third moment to the second moment results in the expectation that within the inertial range, $L \gg r\eta$, the second moment of velocity increment scales as, $S_2(r) \sim v_L(r/L)^{2/3}$. This feature is typically tested by plotting the energy spectra of turbulence (expressed via $S_2(r)$) in the wave vector domain, e.g. as shown in Figs. (1a,3a) of the main text.

Normal and Anomalous scaling of velocity increments One expects moments of the velocity increment to show the following scaling behavior inside the inertial range of turbulence

$$L \gg r \gg \eta: \quad S_n(r) \sim (v_L)^n \left(\frac{r}{L}\right)^{n/3-\Delta_n} \tag{12}$$

where L is the energy containing (largest) scale of turbulence, $\eta \sim (\nu/v_L)^{3/4}L^{1/4}$ is the Kolmogorov (viscous) scale, v_L is the typical velocity fluctuations at the energy containing scale, ν is the viscosity coefficient and Δ_n is the so-called anomalous scaling. $\Delta_3 = 0$, and Δ_n is a growing function of n. In 1941 A. Kolmogorov (20; 21; 22) hypothesized self-similarity of turbulence, i.e. that the anomalous scaling is absent, $\forall n = 3, \cdots: \quad \Delta_n = 0$. Criticized by L. Landau (see e.g. (23) for related history notes), Kolmogorov reconsidered self-similarity in 1962 (24), suggesting instead the so-called refined (log-normal) similarity hypothesis, $\Delta_n = \mu n(n-3)/16$, which was then confirmed to be rather accurate in experiments and simulations, see e.g. (25).

The intermittency test was implemented for both static and dynamic NN schemes discussed in the main part of the manuscript. Both static and dynamic NN schemes passed the test sucessfully. We do not show results in the main part of the manuscript because of the space limitations.

Intermittency of Velocity Gradient Consistently with Eq. (12), estimation of the moments of the velocity gradient results in

$$D_n \sim \frac{S_n(\eta)}{\eta^n}, \tag{13}$$

where dependence of the r/L- and v_L- independent pre-factors in both Eq. (12) and Eq. (13) on n is ignored. Intermittency (extreme non-Gaussianity) of turbulence is expressed the strongest in Eq. (13).

Statistics of coarse-grained velocity gradients: $Q-R$ plane. Isolines of probability in the $Q-R$ plane, expressing intimate feature of the turbulent flow geometry, has a nontrivial shape documented in the literature. See (8; 26) and references therein. Different parts of the $Q-R$ plane are associated with different structures of the flow. Thus lower right corner (negative Q and R), which has higher probability than other quadrants, corresponds to a pancake type of structure (two expending directions, one contracting) with the direction of rotation (vorticity) aligned with the second eigenvector of the stress. This tea-drop shape of the probability isoline becomes more prominent with decrease of the coarse-graining scale.