This is a postprint version of the following published document:

Güemesa, A., Discetti, S. y Ianiro, A. (2019). Sensing the turbulent large-scale motions with their wall signature. *Physics of Fluids*, 31(12), 125112.

DOI: https://doi.org/10.1063/1.5128053

© 2019 AIP Publishing.

# Sensing the turbulent large-scale motions with their wall signature

A. Güemes,[1, a)] S. Discetti,[1] and A. Ianiro[1]
*Aerospace Engineering Research Group, Universidad Carlos III de Madrid, Avenida de la Universidad 30, Leganés, 28911, Spain*

(Dated: 9 November 2019)

This study assesses the capability of extended proper orthogonal decomposition (EPOD) and convolutional neural networks (CNNs) to reconstruct large-scale and very-large-scale motions (LSMs and VLSMs respectively) employing wall-shear-stress measurements in wall-bounded turbulent flows. Both techniques are used to reconstruct the instantaneous LSMs evolution in the flow field as a combination of proper orthogonal decomposition (POD) modes, employing a limited set of instantaneous wall-shear-stress measurements. Due to the dominance of non-linear effects, only CNNs provide satisfying results. Being able to account for nonlinearities in the flow, CNNs are shown to perform significantly better than EPOD both in terms of instantaneous flow field estimation and turbulent statistics reconstruction. CNNs are able to provide a more effective reconstruction performance employing more POD modes, at larger distances from the wall and employing lower wall-measurement resolutions. Furthermore, the capability of tackling non-linear features of CNNs results in estimation capabilities that are weakly dependent on the distance from the wall.

## I. INTRODUCTION

Wall-bounded turbulent flows are present in many industrial and aeronautical applications, including all kind of vehicles immersed in a fluid or turbomachinery flows. Understanding the dynamics of these flows is of utmost importance to enabling sensing and control to improve the performances in terms of skin friction, heat transfer, etc.

Among the spectral pipeline of scales in a turbulent boundary layer, large-scale and very-large-scale motions (LSMs and VLSMs respectively, shown to have a wavelength of up to more than 10 flow characteristics lengths) are one of the main sources of momentum and energy transport. LSMs and VLSMs are regions of the flow that present spatial and temporal correlation, whose evidences are available in several research studies[1–6]. LSMs have been documented as long features of uniform high or low momentum in the logarithmic region of a turbulent boundary layer. In Ref. 6 it was proved using a spanwise rake of hot-wires that these elongated regions extended up to a length of the order of 20 boundary layer thicknesses $\delta_{99}$. These features were also found experimentally in turbulent pipe flows and channel flows, as reported in Refs. 7 and 8.

LSMs have been proved both experimentally[9,10] and numerically[11,12] to have a direct impact on the inner-scales fluctuations close to the wall. It has been shown that the LSMs present in the logarithmic region have an amplitude and frequency modulation effect on the streamwise fluctuations close to the wall[13,14], which has also been extended to the wall-normal and spanwise fluctuations[15]. Ref. 16 showed that it is possible to estimate the wall-shear-stress fluctuations from the one-point measurements of LSMs in the logarithmic region. These works[13–16] proposed a model that assumes that the near-wall small-scale structures are universal[17] and that the effect of Reynolds number can be modelled as an amplitude and frequency modulation effect of the LSMs on the small scales. Furthermore, they showed that the spatial resolution of the measurements was a driving parameter to obtain accurate predictions.

Due to the relevance that LSMs have on the flow, it has become of utmost importance to define techniques that allow to detect and control them. Ref. 18 proposed the use of jets to reduce the skin-friction drag in turbulent boundary layers. Piezo-ceramic actuators to generate a transverse travelling wave along the wall surface[19] or a spanwise-oscillating surface[20] have also been explored for the same purpose. Suction and blowing wall jets that weaken the near-wall streaks[21] have also been proven as an effective solution.

Closed-loop control strategies require proper sensorization strategies for the flow; the clear role played by LSMs thus set the requirements of an effective real-time LSMs-detection solution with minimum intrusiveness. Pointwise probes have been often explored as a good compromise between the needs of low-processing time and minimization of flow disturbance. The flow state is then estimated using only the probe data, thus requiring robust algorithms leveraging statistical evidence of correlations between flow features and probe data. Linear stochastic estimation[22] has been successfully used to estimate turbulent fields [23–28] based on a limited set of probes. One of the most promising approaches that have been proposed is the extended proper orthogonal decomposition (EPOD, see e.g. Ref. 29). Using the projection of the snapshot matrix of a given quantity on the temporal basis corresponding to another one, it has been shown that it is possible to extract the correlation between synchronized measurements through the extended POD modes[29]. The compactness of the POD basis allows data-driven filtering, thus providing more robust estimations than the simple application of linear stochastic estimation. Using this correlation it has been shown that accurate far-field predictions of a turbulent jet can be obtained,[24] as a combination of the flow-field POD modes through time coefficients linearly-estimated based on instantaneous point-wise measurements. The EPOD has been used previously to reconstruct turbulent flow fields using velocity probes in a turbulent channel[30] and in a high-Reynolds-number pipe flow[8].

---

[a)] Electronic mail: guemes.turb@gmail.com

The EPOD approach in Ref. 30, however, does not allow to take into account nonlinearities in the relation between probe data and flow features. More specifically, the amplitude modulation effects that LSMs have on the near-wall scales[9,10] could not be identified using EPOD. One promising path to embed these non-linear effects is the use of artificial neural networks (ANNs), firstly introduced with the *perceptron* concept[31]. ANNs are a set of layers that transform a fixed-size input (for example, a flow velocity field) into a fixed-sized output (for example, a coefficient) by computing a weighted sum of their inputs from the previous layer and applying a nonlinear function to the result[32]. It has been shown that an ANN with a single hidden layer is equivalent to perform a POD on the data[33] while non-linear effects can be introduced making use of several hidden layers.

ANNs have been applied to a large variety of research fields, including fluid mechanics. The main applications include the solution of the Reynolds-Averaged Navier-Stokes (RANS) and of the large-eddy simulation (LES) problems, improving turbulence modelling[34–37]. Ref. 38 used a multilayer perceptron (MLP) to estimate the lift coefficient and the displacement of a freely vibrating cylinder from flow measurements in the surroundings. Additionally, they were able to estimate the flow fields from the concentration of a passive scalar in the flow. Another possible application has been shown in Ref. 39, where streaks in pressure-gradient boundary layers are classified into those that will break into turbulence and those which will not.

As reviewed in Ref. 40, several authors employed machine learning for modelling fluid flows. A promising application of ANNs in turbulence is the direct reconstruction of turbulent flows from external measurements. In particular, Ref. 41 carried out a reconstruction of the near-wall flow using a neural network to approximate the higher-order terms of the second-order model

$$\underline{u}(\underline{x},t) = \underline{\omega} y + \frac{Re}{2}\frac{d\underline{P_w}}{\underline{x}}y^2 + ANN(\underline{P_w},\underline{\tau_w}) \tag{1}$$

where $\underline{\omega}$ is the wall tangential vorticity, $\underline{P_w}$ is the wall pressure field,and $\underline{\tau_w}$ the wall-shear-stress field. The model provided a good reconstruction up to $y^+ \approx 40$ using wall-shear and pressure measurements. Superscript $+$ refers to inner-scaling, based on the friction velocity $u_\tau = \sqrt{\tau_w/\rho}$ and viscous length $l_* = \nu/u_\tau$ (being $\tau_w$, $\rho$, and $\nu$ the mean wall-shear stress, density and viscosity respectively). An ANN model was proposed in Ref. 42 for the deconvolution and regularization of low-pass spatially filtered turbulent fields. In Ref. 43 two different ANN models were used to increase the resolution of turbulent fields; they use DNS fields to train the model to upgrade low-resolution fields into fully-resolved ones. Ref. 44 has shown that ANNs can be used to predict the nonlinear dynamics of the flow around a cylinder. In Ref. 45 both MLP and long short-term memory (LSTM) ANNs have been used to predict the temporal evolution of a low-order model of a turbulent flow. Due to the sequential nature of LSTM, their results outperformed those obtained by the MLP, providing very accurate turbulence statistics. Based on the results of these studies, the ANNs appear to be a perfect candidate to estimate the dynamics of POD modes from a limited set of time-resolved wall measurements.

Among ANNs, convolutional neural networks (CNNs) take advantage of their input hierarchical shape[46–49]. CNNs have been applied to a large variety of research fields[50,51], and are recently being exploited for turbulent-flow estimation. For instance, CNNs were used in Ref. 43 to increase the resolution of the flow field, employing as input the flow field itself but with lower resolution. Ref. 52 reports a remarkable application in which convolutional kernels with different inputs are used to estimate the streamwise velocity fluctuations of a zero-pressure-gradient turbulent boundary layer. In Ref. 53 CNNs have been used to reconstruct instantaneous velocity fields parallel to the wall up to distances of $y^+ = 50$. Their results show that first- and second-order statistics can be computed accurately from the reconstructed flow fields, although the reconstruction error increases with the wall distance.

Aiming at the identification of a suitable sensing device for the detection of large- and very-large-scale motions in a turbulent wall-bounded flow, we explore the performances of CNNs for the identification of the large-scale instantaneous flow fields from a limited set of shear-wall probes. A comparison with the EPOD is carried out to assess whether the non-linearity provided by CNNs can improve the quality of the flow field reconstruction.

This work is based on a computational dataset which is used both for the training of the estimation methodologies and for the assessment of the reconstruction performances. Wall-shear-stress values at a limited number of locations are employed to reconstruct the flow field in a turbulent channel flow at different wall-parallel planes ranging from $y^+ = 50$ to $y^+ = 200$, being the friction-based Reynolds number equal to $Re_\tau = 1000$. The paper is organised as follows: Sec. II describes the different methodologies employed to reconstruct the LSMs in the flow field and the database used to assess their performance. In Sec. III the obtained results are presented and discussed, prior to the conclusions in Sec. IV.

## II. METHODOLOGY

The reference frame $\underline{x} = [x,y,z]$ adopted in this work is aligned with the streamwise $x$, wall-normal $y$ and spanwise $z$ directions, being their corresponding fluctuating velocity components denoted as $\underline{u} = [u,v,w]$. Velocity fields are decomposed according to the Reynolds decomposition; the fluctuating velocity components are then expressed hereafter as a linear combination of a spatial basis, composed of spatially orthonormal functions $\underline{\phi}_i(\underline{x})$, times a temporal basis, made of temporally orthonormal functions $\psi_i(t)$, by means of the proper orthogonal decomposition (POD)[54]:

$$\underline{u}(\underline{x},t) \approx \sum_{i=1}^{Nm} \psi_i(t)\sigma_i\underline{\phi}_i(\underline{x}) \tag{2}$$

with $N_m$ being the number of POD modes and $\sigma_i$ weighs the contribution of each mode to the snapshots. POD modes are optimal in the least square sense: the first $r$ POD modes form an orthogonal basis which minimizes the Frobenius norm of the flow field reconstruction error, i.e. provides the best rank-$r$ approximation of the snapshots. This implies that POD provides flow field modes ordered by their energy content. Owing to the prevalent role in momentum and energy transport of LSMs, it has indeed been shown that the first POD modes can be thought of as representations of the LSMs present in the flow[55].

The practical implementation with the dataset described in Sec. II C follows the snapshot POD method[56]. The velocity fields are reshaped in a vector form and rearranged in a matrix $\underline{\underline{U}}$ of size $n_t \times n_p$, where $n_t$ is the number of available flow fields and $n_p$ is the number of grid points. The corresponding SVD is:

$$\underline{\underline{U}} = \underline{\underline{\Psi}}\,\underline{\underline{\Sigma}}\,\underline{\underline{\Phi}} \tag{3}$$

.

Assuming statistical convergence, flow fields at a generic time instant (inside and outside of the training dataset) can be quite accurately be described as a linear combination of the modes contained in the $\underline{\underline{\Sigma}}\,\underline{\underline{\Phi}}$ matrix by the corresponding time coefficients, if available. One strategy to obtain these time coefficients for time instants outside of the dataset is the synchronized use of pointwise probes and flow field data, with the probes being trained to *reconstruct* the velocity fields. The velocity field at a given instant can then be expressed as:

$$\underline{u}^* = \underline{\psi}^*\,\underline{\underline{\Sigma}}\,\underline{\underline{\Phi}} \tag{4}$$

Estimated quantities are denoted with the superscript $*$. In the following we describe two methodologies employed for the estimation of the time coefficients $\underline{\psi}^*$.

### A. Extended proper orthogonal decomposition

The snapshot POD can be applied both to the flow field and to any other quantity measured by synchronized probes immersed in the flow or at the wall, e.g. corresponding to a set of probes $\underline{\underline{U}}_{pr}$:

$$\underline{\underline{U}}_{pr} = \underline{\underline{\Psi}}_{pr}\,\underline{\underline{\Sigma}}_{pr}\,\underline{\underline{\Phi}}_{pr} \tag{5}$$

The extended POD modes[29], corresponding to the projection of the probe measurements, on the flow-field temporal basis can be estimated as:

$$\underline{\underline{\Psi}}_{pr}^T\,\underline{\underline{U}} = \underline{\underline{\Psi}}_{pr}^T\,\underline{\underline{\Psi}}\,\underline{\underline{\Sigma}}\,\underline{\underline{\Phi}} = \underline{\underline{\Xi}}\,\underline{\underline{\Sigma}}\,\underline{\underline{\Phi}} = \underline{\underline{\Sigma}}_e\,\underline{\underline{\Phi}}_e \tag{6}$$

where the subscript $e$ refers to extended POD modes and the matrix $\underline{\underline{\Xi}} = \underline{\underline{\Psi}}_{pr}^T\,\underline{\underline{\Psi}}$ is a matrix containing information about the temporal correlation between flow and probe modes. If the dataset is sufficiently large to reach statistical convergence, it is possible to assume that the matrix $\underline{\underline{\Xi}}$ represents a good estimate of the relation between the POD time coefficients at a certain probe snapshot and those of an out-of-sample estimate of the flow field. Taking this and Equation 6 into account, it is possible to estimate the temporal POD coefficients of the flow field using the wall-field POD time coefficient and the temporal correlation matrix:

$$\underline{\underline{\Psi}}^* = \underline{\underline{\Psi}}_{pr}\,\underline{\underline{\Xi}} \tag{7}$$

A flow field can thus be reconstructed by means of a set of probes once the temporal correlation matrix is known. In this work we will explore the use of wall sensors, since it is of more practical application if compared to probes immersed in the flow.

The EPOD reconstruction is equivalent to a linear stochastic estimation and as shown in Ref. 30 it is an energy preserving process. The energy-optimality of POD modes can then be enforced by employing statistical filters[30] and improving the reconstruction with respect to LSE by cleaning the data from uncorrelated features. This approach has been shown to be successful in several application, but it is not able to take into account non-linear effects.

$$\underline{\underline{U}}^* = \underline{\underline{\Psi}}^*\,\underline{\underline{\Sigma}}\,\underline{\underline{\Phi}} \tag{8}$$

### B. Artificial neural network

The alternative approach pursued here to estimate the POD time coefficient is through the use of ANNs. Probe data can be rearranged in a tensor form $\underline{\underline{\underline{W}}}$ of size $n_t \times n_w \times n_h$, where $n_w$ is the number of grid points in the width direction and $n_h$ is the number of grid points in the height direction ($n_p = n_w \times n_h$).

The tensor $\underline{\underline{\underline{W}}}$ is passed through a set of convolutional and max-pooling layers that halve sequentially the field size. The number of convolutional and max-pooling layers depends on the initial wall-shear stress field resolution, thus three different CNN architectures have been set depending on this feature.

When the tensor has been reduced to the minimum possible size, it is reshaped in a vector form $\underline{v}$ by means of a flatten layer, and it is passed through two fully-connected layers with weights $\underline{\underline{w}}$. The last fully-connected layer outputs a single value that corresponds to the temporal POD mode of the flow field associated to the wall-shear-stress field.

$$\underline{\psi}_i^* = \underline{\underline{w}}_2 \times \underline{\underline{w}}_1 \times \underline{v} \tag{9}$$

Finally, the temporal POD-mode estimation obtained through CNNs can be used to reconstruct the flow fields from a set of wall probes applying Eq. 8. Different CNNs have been trained for different probe spacings, wall-distance positions of the flow fields and the first 10 temporal POD modes, resulting

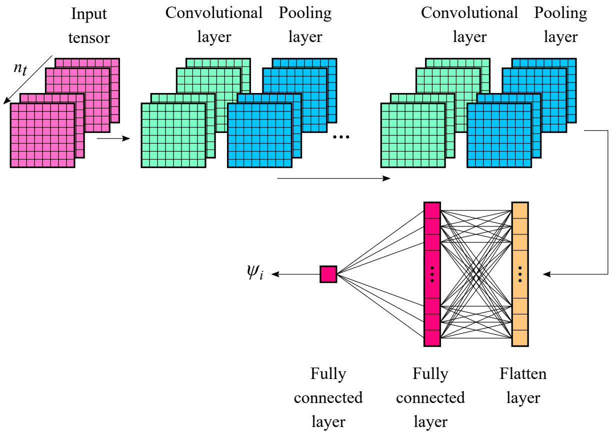


FIG. 1. Schematic view of the CNN architecture.

TABLE I. Dataset description. For each case, 7000 fields have been extracted for training, 500 for validation, and 7500 for testing.

| Case | $y^+$ | Resolution $[pt]$ | EPOD symbol | CNN symbol |
|---|---|---|---|---|
| H050 | 50 | 128 × 128 | (−□−) | (−■−) |
| H100 | 100 | | | |
| H150 | 150 | | | |
| H200 | 200 | | | |
| M050 | 50 | 64 × 64 | (−○−) | (−●−) |
| M100 | 100 | | | |
| M150 | 150 | | | |
| M200 | 200 | | | |
| L050 | 50 | 32 × 32 | (−△−) | (−▲−) |
| L100 | 100 | | | |
| L150 | 150 | | | |
| L200 | 200 | | | |

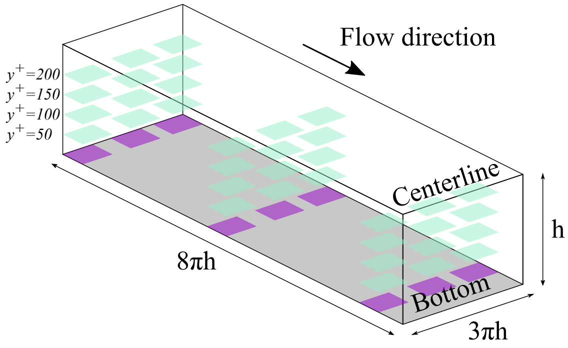


FIG. 2. Schematic representation of the flow and wall fields extracted from the Channel Flow DNS spatial domain. Axes are not equally scaled.

in a total of 120 different CNNs. A simplified sketch of the CNN architectures can be seen in Fig. 1.

A rectified linear unit (*ReLU*) activation function[57] has been used in the fully connected layers of the CNNs architecture to take into account the non-linearities of the flow, while an Adam optimizer[58] with a constant learning rate of 0.001 has been selected to update the weights of the different neuron layers.

$$ReLU(x) = max\{0, x\} \tag{10}$$

### C. Dataset description

The different methods have been tested employing a Direct Numerical Simulation (DNS) dataset that includes full flow-field information from a numerical simulation in a channel flow at friction Reynolds number $Re_\tau \approx 1000$. The channel-flow database is available at the Johns Hopkins Turbulence Database (http://turbulence.pha.jhu.edu/). The channel has a dimension of 2 half-channel-heights $h$ from wall to wall, $3\pi h$ in the span-wise direction and $8\pi h$ in the stream-wise direction. The DNS database covers one channel-flow through-time $8\pi h/U_b$ (where $U_b$ is the channel bulk velocity) with a DNS time step of $\delta_t = 0.0014 h/U_b$. More details on the numerical simulations can be found in Refs. 59 and 60.

To train the proposed methodologies and to assess their capability of reconstructing the LSMs, we collected both wall-shear-stress fields and $[u, v]$ fields in streamwise-spanwise planes located at four distances from the wall. These fields were representative of square $h \times h$ domains parallel to the wall located at $y^+ = [50, 100, 150, 200]$. The fields resolution is equal to $128 \times 128$ points. The wall-friction fields were extracted at the same temporal instants in the same $[x, z]$ region with three different spatial resolutions, $[128 \times 128, 64 \times 64, 32 \times 32]$, covering in all cases the same $h \times h$ area.

Only 24 flow and wall-shear fields, separated by $t^+ = 1000$ and therefore uncorrelated, can be extracted in the same position due to the limited timespan of the simulation. To acquire a well-converged dataset, the flow and wall-shear fields have been extracted at different positions in the streamwise and spanwise directions of the bottom part of the channel domain, taking advantage of the simulation homogeneity in those directions. Specifically, flow fields have been obtained in 26 streamwise locations and in 10 spanwise locations. Fields have been extracted each 0.25 convective times, leading to 100 different temporal fields at each location. The total amount of available flow and wall fields is 26000. Taking into account that this quantity is more than enough to reach statistical convergence and that the dataset is aimed to be used for training, validation and testing of different LSMs estimation methods, three field subsets have been extracted: 7000 fields for training, 500 fields for validation (only used in the CNN) and 7500 fields for testing. It is important to mention that the field subsets have been separated in the time domain, i.e. belong to snapshots sets extracted respectively at the beginning and at the end of the sequence. A sketch of the fields extracted from the DNS domain is given in Fig. 2, while a description of each dataset is provided in Table I.

The POD modes have been computed using the fluctuating velocity fields. Fig. 3 reports both the modes energy distribution (i.e the eigenvalue distribution $\lambda_i = \sigma_i^2$) and the cumulative energy content for the flow-field snapshots. It is well known that the first POD modes correspond to the large-scale

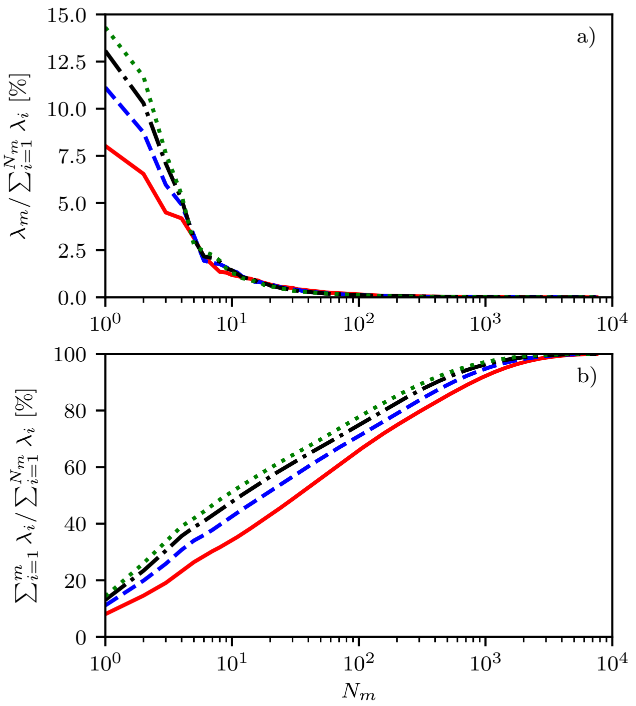


FIG. 3. a) Energy spectrum and b) cumulative energy contribution of the flow field modes of the channel simulation. Colours refer to the wall distance position of the flow fields: $y^+ = 50$ (——), $y^+ = 100$ (– – –), $y^+ = 150$ (–·–·), $y^+ = 200$ (······).

motions present in the flow[55,61] while following modes are associated with smaller and less-energetic structures. According to previous literature, LSMs and VLSMs account for approximately 40% of the turbulent kinetic energy[8,62,63] which approximately correspond here to the first 10 POD modes. In the following, we thus focus on the estimation of the time coefficients for these modes only, resulting in a low-order reconstruction of the LSMs and VLSMs. Turbulent flows are characterized by having a rich range of spatial scales, which in this case makes necessary a large number of modes (between 400 and 900) to reach 90% of the cumulative energy spectrum. When analyzing the energy contribution of each mode at different wall-normal positions, it can be seen that, the closer to the wall, the smaller is the contribution of the LSMs, which is not surprising since close to the wall the smallest scales energy contribution is more significant[15].

## III. RESULTS

This section presents the LSM reconstruction with 10 POD modes using EPOD and CNN techniques. The main implementation difference is that EPOD is able to estimate simultaneously the time coefficients of all the POD modes, while a CNN needs a separate training for each mode. From the practical viewpoint, EPOD and CNN work with the same input. As explained in Sec. II the input data are the streamwise wall-shear-stress fields, while the predicted output is the first 10 POD time coefficients. As argued in the previous section, this is normally sufficient to detect the large-scale motions. Having available the target fields to be reconstructed, the actual temporal POD modes of each field can be computed and it is possible to employ the low-order reconstruction of the fields obtained with the first 10 modes as reference field.

### A. A-priori assessment

Both methods (EPOD and CNN) allow an *a priori* assessment of the reconstruction ability, without having to rely on a direct comparison with a ground truth, which is of course not available in experiments. The estimation performances of the EPOD can be deduced for instance from the temporal correlation coefficient matrix $\underline{\underline{\Xi}}$. Fig. 4 shows the $\Xi_{i,i}$ values for the first 20 temporal POD modes for all the tested wall-distance positions and wall-shear-stress-field resolutions. It can be seen that the matrices exhibit a diagonal dominance in the north-west corner, which corresponds to a direct relation between the most energetic modes of flow fields and wall-shear stress. While this dominance is not a necessary condition for high-quality reconstruction, it is an indicator of a "healthy" correspondence of the most energetic features in flow fields and probe data. This is needed to achieve a compact and efficient reconstruction, since the energy of each mode is not spread over a range of modes. However, this dominance disappears starting from the 5th-6th mode. Moreover, this correlation between low-frequency features in the flow and low-frequency features at the wall impedes the EPOD to capture the amplitude modulation effects that the LSMs produce over high-frequency near-wall scales[10]. The wall-distance is the most relevant parameter; as expected, the peak of the correlation coefficient of the first modes is larger for the flow fields located closer to the wall. On the other hand, the resolution level of the wall measurements seems to have a minor effect, which could indicate that coarser meshes could be employed without significantly affecting the estimation quality. This is clearly an important indication to simplify the practical implementation of this method in experiments.

For CNNs, the validation error is a directly interpretable indicator of the quality of the reconstruction. Fig. 5 shows the evolution of both training and validation datasets for the first mode of the H050 case (see Table I). During training, a callback function was used initially to select the number of iterations, also called epochs, based on the mean-square error evolution of the validation dataset. However, after the first training it was noticed that the validation mean-square error had a significant scatter and a low rate of decrease, while the training error at the stopping epoch was still clearly descending for increasing number of epochs. Because of this, it was decided to fix the number of epochs equal to 100.

An important metric for comparison is the training time of EPOD and CNN. Both data-driven techniques have been trained in a standard workstation (Intel® Core™ i9-8950K CPU @ 2.9 GHz with NVIDIA GeForce® GTX 1050 Ti GPU with Max-Q design). EPOD has required 7 minutes to train

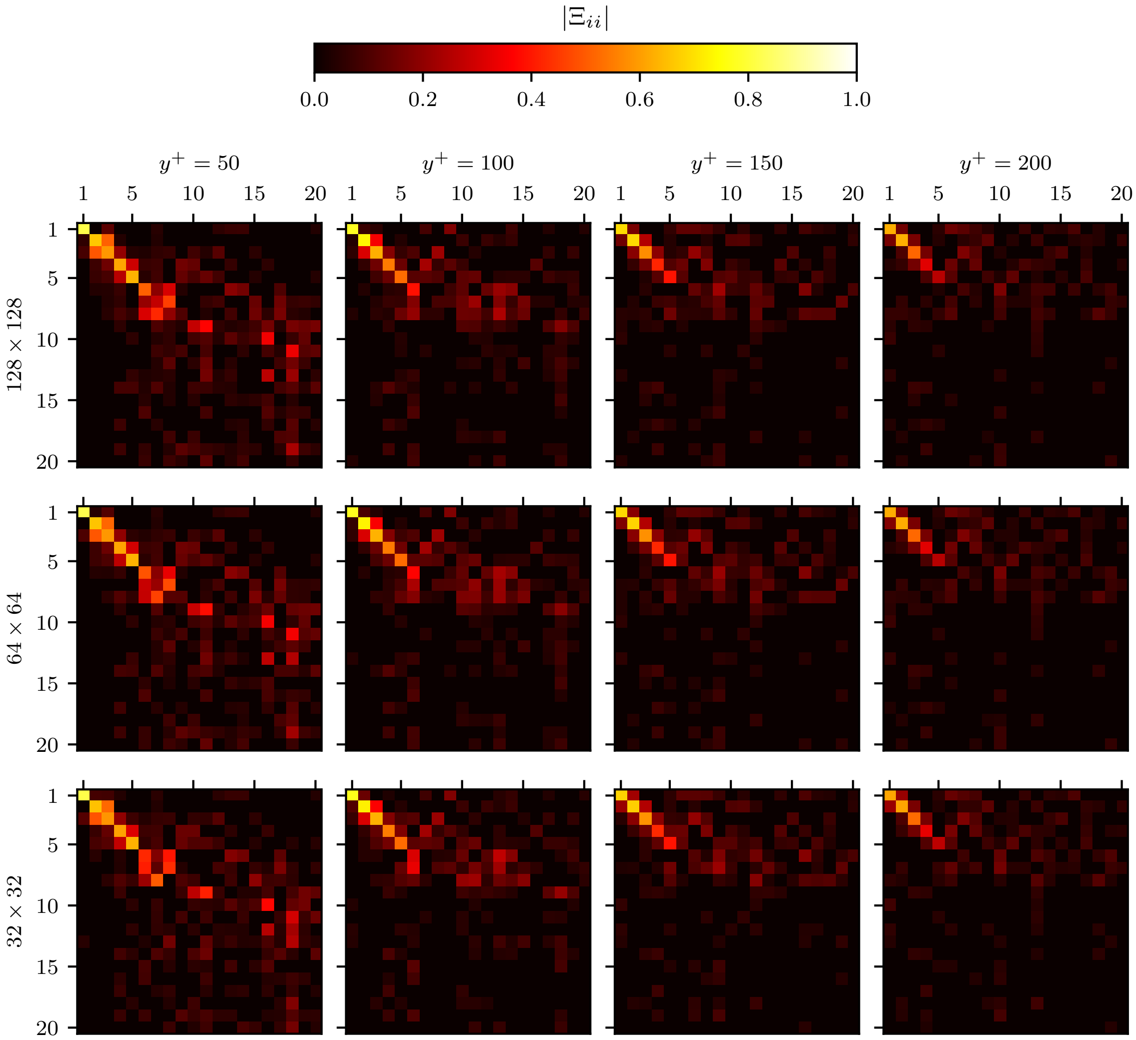


FIG. 4. Time correlation coefficient matrix $\underline{\underline{\Xi}}$ for EPOD estimation. Columns report data at fixed wall distance $y^+$. Rows report data at fixed wall-shear-stress resolution.

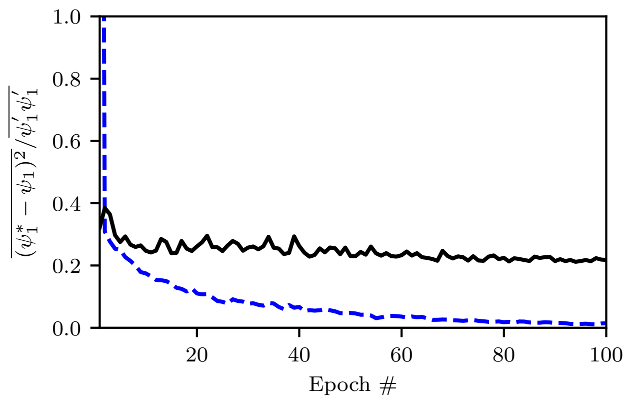


FIG. 5. Mean-squared-error evolution for the training (– – –) and validation (——) datasets for the first temporal POD-mode coefficient in the case H050 (see Table I).

case M050, while CNNs have required 44 minutes for the training of the same case. It has to be reminded that 10 different CNNs have been trained, one for each of the first 10 first temporal POD modes. Considering the required time to train the CNNs for a single temporal POD mode, it is reduced to 4 minutes. The larger computational cost required by CNNs is counterbalanced by the possibility of tackling non-linear effects related with the LSMs present in the flow. The estimation time for each field is of the order of $0.01s$ per snapshot for CNN, while it is practically real-time for EPOD. This is important to allow the possibility of active closed-loop flow control based on flow field estimation. However, it has to be remarked that the selected architecture is not conceived for direct implementation in active flow control, and significant improvements could be obtained employing dedicated high-performance hardware.

### B. LSMs-reconstruction performance

For the *a posteriori* assessments of the LSM reconstruction techniques, several metrics have been used. The first one is the squared correlation coefficient $R^2$ of the temporal mode coefficients, which quantifies the quality of the reconstruction of $\psi_i^*$ before reconstructing the flow fields. It is defined as:

$$R^2_{\psi_i} = \frac{\overline{\psi_i^* \psi_i}^2}{\overline{\psi_i^{*2}}\,\overline{\psi_i^2}} \tag{11}$$

The squared correlation coefficient is reported in Fig. 6 for both EPOD and CNN techniques. Here and in the following, ensemble-averaged values are indicated as $\bar{\cdot}$. For colour code, the reader is referred to Table I. CNN abundantly surpasses EPOD in reconstruction quality for all the analyzed cases, independently of wall-distance and wall-shear-stress resolution. The squared correlation coefficient achieved by EPOD for the case H50 (the one closer to the wall and with higher resolution, thus the most favourable case) is lower than the value reached by CNN for case L200 (the most challenging case, further from the wall and with lower resolution) for all the first 10 modes. Additionally, EPOD shows a steep drop in performance beyond the third POD mode in most cases, while CNNs are significantly more robust. While the analysis here is limited to the first 10 POD modes due to the available computational resources, it can be expected that CNNs can pave the way to successful reconstruction of also higher-order modes.

Overall, the correlation coefficient spotlights the limits of EPOD for LSM detection. This is most likely to be addressed to non-linear effects of LSMs on the wall signature, which are not modelled by EPOD and become more intense when moving away from the wall. CNNs exhibit a negligible decrease of the squared correlation coefficient when going away from the wall, thus extending the capability of flow-field estimation from the wall signature. However, this effect could be an artifact of the number of selected modes. The energy content extracted with the first 10 POD modes is different at each distance from the wall, as shown in Fig. 3. In the regions closer to the wall ($y^+ = 50$) the first modes do not represent the LSMs but the near-wall scales affected by them. Thus, the reconstruction of the full instantaneous fields would feel a significant larger effect of the wall distance on the reconstruction performance. In terms of wall-shear-stress resolution, CNNs show a weak sensitivity, though stronger than EPOD. This can be probably ascribed to the fact that the non-linear effects require larger datasets in order to be adequately characterized. This trend can be explained considering that reducing the resolution would actually tend to the asymptotic result corresponding to a $2 \times 2$ resolution which would correspond to a simple MLP. It was shown in Ref. 33 that this neural-network architecture produces the same output as POD. Consequently low-resolution configurations were not explored also because, with such low-resolution inputs, it would be much more interesting to employ temporal signals as input in both EPOD (as shown in Ref. 8 and Ref. 30) and ANNs (using long short-term memory for example[64]).

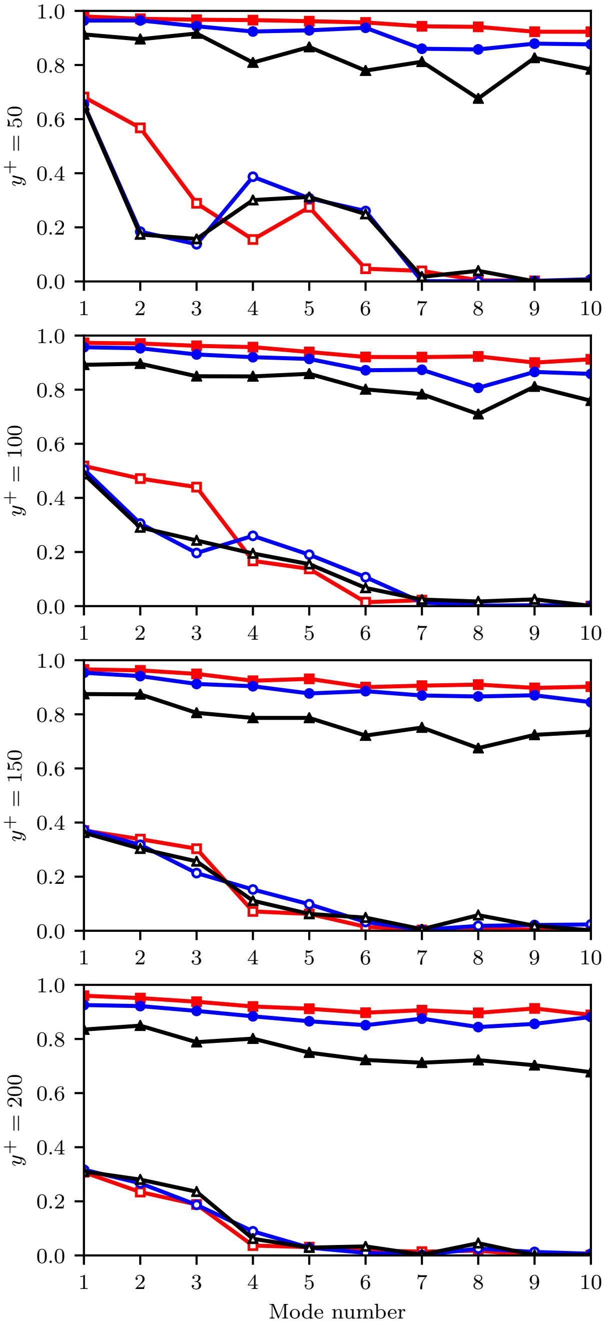


FIG. 6. Squared correlation coefficient $R^2_{\psi_i}$ of the reconstructed time coefficients. From top to bottom, wall-distance positions correspond to $y^+ = [50, 100, 150, 200]$. For colour code, see Table I.

Fig. 7 reports the reconstruction of the streamwise velocity fluctuation fields at $y^+ = 150$ using the largest wall-shear-stress resolution for the estimation. This wall distance was chosen since it has been proven that the geometrical center of the logarithmic region (defined as $y^+ = 3.9\sqrt{Re_\tau}$) is the point where high-momentum LSMs start providing a dominating effect over the near-wall small-scale fluctuations[10] (while for smaller wall distances low-momentum events are dominant).

The streamwise velocity fluctuations at this point have been used in previous studies to predict near-wall[13,14] and wall-shear-stress[16] turbulent statistics. In Fig. 7 an instantaneous field is compared with the low-order reconstruction performed with EPOD and with that performed with CNNs. The exact reconstruction using the first 10 modes is included as ground truth for reference. Visual inspection shows that the CNNs are able to capture the high- and low-momentum regions of the turbulent field with high accuracy. On the other hand, as expectable from the low $R^2$ values, EPOD reconstruction does not capture any significant flow feature.

A quantitative analysis is carried out by defining the error in the reconstruction of the instantaneous velocity fields as the root mean squared error of the fields scaled with the streamwise Reynolds stress:

$$error = \frac{\sqrt{\overline{(u'^{*} - u')^2}}}{\sigma_{u'}} \tag{12}$$

The errors reported in Fig. 8 show that the reconstruction quality of the instantaneous fields with EPOD is rather poor, reaching levels around 100% of streamwise Reynolds stress deviation. The error does not exhibit any significant dependence with the wall-shear resolution. Regarding CNN, it can be seen that wall-shear-stress resolution has a significant larger impact than the wall-distance position, reaching a 40% error in the L050 case. This result is in agreement with the behaviour of squared correlation coefficient shown in Fig. 6.

The inner-scaled streamwise Reynolds stresses estimated in each of the four planes, accounting for the reconstructed LSM velocity fields, are reported in Fig. 8. It can be observed that CNNs provided a reasonably good estimation independently of the wall-shear-stress resolution. Similarly, the turbulent statistics computed with EPOD are not affected by the wall-shear-stress resolution, which is in line with the results of Fig. 6. EPOD is significantly underestimating the streamwise Reynolds stress values, with this effect being stronger at larger distances from the wall. This result is not surprising, since EPOD is a linear operator which is expected to be severely challenged when the relation between probe and field fluctuations become strongly non-linear.

## IV. CONCLUSIONS

The ability of two different data-driven techniques to reconstruct the LSMs present in a turbulent channel flow using wall-shear stress measurements has been assessed. Both techniques have been used to perform a low-order estimation within the POD basis described by the flow field measurements. The effect of the wall-shear-stress field resolution and of the wall-distance of the plane to be reconstructed have been addressed.

The extended proper orthogonal decomposition (EPOD) has shown a limited capacity to reconstruct the streamwise Reynolds stress of the LSMs. The reconstruction performance is severely affected by the wall distance, a limit that can be addressed to impossibility to model non-linearity with EPOD. The reconstruction error reaches levels of 100% streamwise Reynolds stress deviation, with a corresponding significant underestimation of the streamwise velocity fluctuations of LSMs.

Convolutional neural networks (CNNs) have shown an outstanding capability to reconstruct LSMs in terms of turbulent statistics and instantaneous flow fields. The reconstruction performance is only weakly affected by the wall distance and the wall-shear-stress resolution. Nevertheless, this effect could be addressed to the number of selected POD modes, as long as their energy content varies with wall distance. In all tested cases the reconstruction error is significantly lower than the error achieved by EPOD. It is impressive that the reconstruction error of CNN of the most challenging case (minimum wall-shear resolution and maximum wall distance) is about one half of the error of the simplest case for EPOD (i.e. maximum wall-shear resolution and minimum wall distance). This opens the path to more robust estimation of flow fields from wall-data.

It would be interesting, in future investigations, to analyze the capabilities of CNNs to reconstruct the LSMs in a flow with a different Reynolds number to that of the training dataset. This would enable flow-field reconstructions based on a single computational or experimental training. One key enabler is that POD spatial modes have been shown to be independent of the Reynolds number[65]. In the same way, it is worth analyzing whether the reconstruction algorithms trained in a turbulent channel flow would be able to obtain similar results in other canonical flows, i.e. turbulent-boundary-layer or pipe flows, suggesting a universal mechanism of non-linear interactions in wall-bounded flows. Both purposes could be tackled using the transfer-learning technique, that allows to transfer knowledge acquired from a solved task, like channel-flow reconstruction, to a new one, such as the reconstruction of a pipe flow. For example, Ref. 53 has shown that transfer learning can be used to reduce the computational cost of training a CNN for the reconstruction at a new wall distance by using a previously-trained CNN at a different wall distance.

With respect to a future experimental set-up to test the capabilities of CNNs to reconstruct LSMs, recent investigations have shown that it is possible to obtain time-resolved heat-transfer maps with resolutions of the order of 100 × 100[66–68]. Whereas the relationship between heat-transfer maps and wall-shear-stress measurements have been shown in Ref. 69, it might be expected that this input resolution is enough for CNNs to provide accurate LSM reconstructions.

## ACKNOWLEDGMENTS

This work has been partially supported by the Grant DPI2016-79401-R funded by the Spanish State Research Agency (SRA) and European Regional Development Fund (ERDF). AG acknowledges Dr. A. Sánchez for insightful discussions about CNN architecture. The authors acknowledge Dr. R. Vinuesa for insightful comments and discussions.

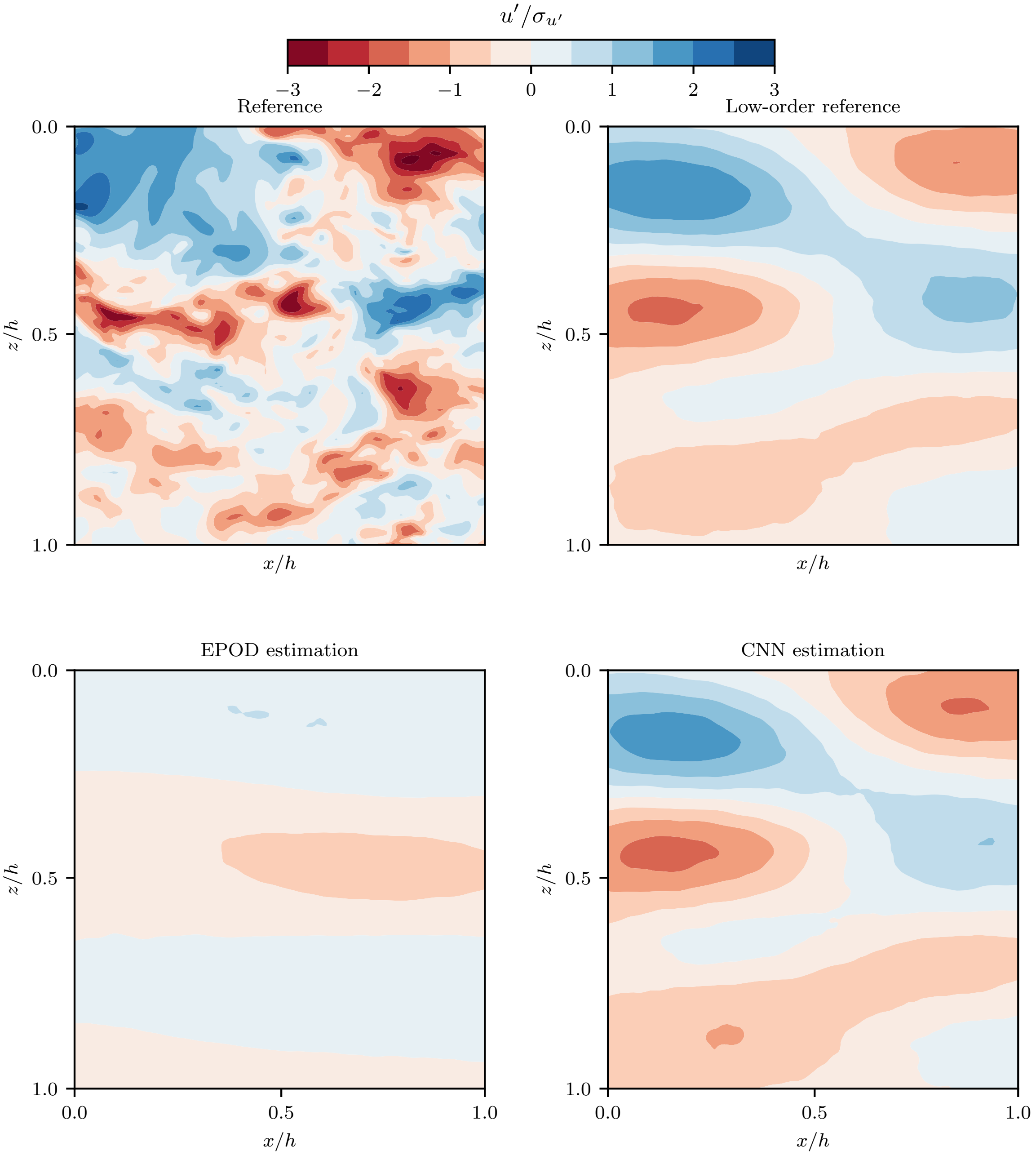


FIG. 7. Instantaneous streamwise velocity fluctuations fields reconstructed with 10 POD modes for case H150. The first row shows the instantaneous field and the reference low-order reconstruction with 10 modes. The second row shows, from left to right, EPOD reconstruction, and CNN reconstruction. All fields have been scaled with the streamwise velocity fluctuation standard deviation $\sigma_{u'}$ of the nominal case.

[1] S. J. Kline, W. C. Reynolds, F. Schraub, and P. Runstadler, "The structure of turbulent boundary layers," J. Fluid Mech. **30**, 741–773 (1967).

[2] A. Favre, J. Gaviglio, and R. Dumas, "Structure of velocity space-time correlations in a boundary layer," Phys. Fluids **10**, S138–S145 (1967).

[3] R. F. Blackwelder and L. S. Kovasznay, "Time scales and correlations in a turbulent boundary layer," The Physics of Fluids **15**, 1545–1554 (1972).

[4] C. E. Wark and H. Nagib, "Experimental investigation of coherent structures in turbulent boundary layers," J. Fluid Mech. **230**, 183–208 (1991).

[5] K. Kim and R. Adrian, "Very large-scale motion in the outer layer," Phys. Fluids **11**, 417–422 (1999).

[6] N. Hutchins and I. Marusic, "Evidence of very long meandering features in the logarithmic region of turbulent boundary layers," J. Fluid Mech. **579**, 1–28 (2007).

[7] J. Monty, J. Stewart, R. Williams, and M. Chong, "Large-scale features in turbulent pipe and channel flows," J. Fluid Mech. **589**, 147–156 (2007).

[8] S. Discetti, G. Bellani, R. Örlü, J. Serpieri, C. S. Vila, M. Raiola, X. Zheng, L. Mascotelli, A. Talamelli, and A. Ianiro, "Characterization of very-large-scale motions in high-re pipe flows," Exp. Therm. Fluid Sci. **104**, 1–8 (2019).

[9] N. Hutchins and I. Marusic, "Large-scale influences in near-wall turbulence," Philos. T. R. Soc. A **365**, 647–664 (2007).

[10] R. Mathis, N. Hutchins, and I. Marusic, "Large-scale amplitude modulation

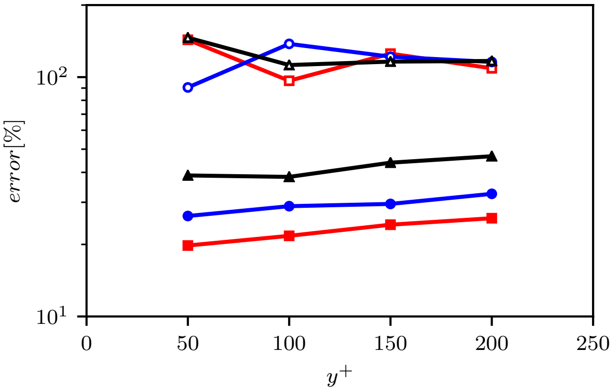


FIG. 8. Mean error (according to Eq. 12) of the instantaneous streamwise velocity fluctuation fields reconstructed with the first 10 POD modes versus wall-normal positions. For colour code, see Table I.

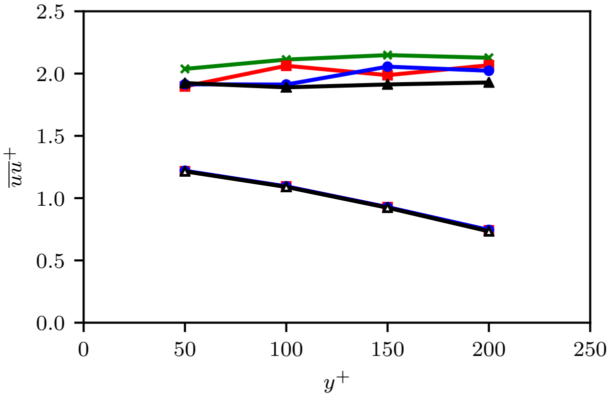


FIG. 9. Reconstruction with 10 POD modes of inner-scaled streamwise velocity fluctuations versus the wall-normal positions. For colours, see Table I. Nominal POD reconstruction is shown as (—×—).

of the small-scale structures in turbulent boundary layers,” J. Fluid Mech. **628**, 311–337 (2009).

[11] S. Hoyas and J. Jiménez, “Scaling of the velocity fluctuations in turbulent channels up to $Re_\tau$= 2003,” Phys. Fluids **18**, 011702 (2006).

[12] S. Hoyas and J. Jiménez, “Reynolds number effects on the reynolds-stress budgets in turbulent channels,” Phys. Fluids **20**, 101511 (2008).

[13] I. Marusic, R. Mathis, and N. Hutchins, “Predictive model for wall-bounded turbulent flow,” Science **329**, 193–196 (2010).

[14] R. Mathis, N. Hutchins, and I. Marusic, “A predictive inner–outer model for streamwise turbulence statistics in wall-bounded flows,” J. Fluid Mech. **681**, 537–566 (2011).

[15] B. Ganapathisubramani, N. Hutchins, J. Monty, D. Chung, and I. Marusic, “Amplitude and frequency modulation in wall turbulence,” J. Fluid Mech. **712**, 61–91 (2012).

[16] R. Mathis, I. Marusic, S. I. Chernyshenko, and N. Hutchins, “Estimating wall-shear-stress fluctuations given an outer region input,” J. Fluid Mech. **715**, 163–180 (2013).

[17] J. Jiménez and A. Pinelli, “The autonomous cycle of near-wall turbulence,” J. Fluid Mech. **389**, 335–359 (1999).

[18] M. Abbassi, W. Baars, N. Hutchins, and I. Marusic, “Skin-friction drag reduction in a high-reynolds-number turbulent boundary layer via real-time control of large-scale structures,” Int. J. Heat Fluid Fl. **67**, 30–41 (2017).

[19] H. Bai, Y. Zhou, W. Zhang, S. Xu, Y. Wang, and R. Antonia, “Active control of a turbulent boundary layer based on local surface perturbation,” J. Fluid Mech. **750**, 316–354 (2014).

[20] K. Gouder, M. Potter, and J. F. Morrison, “Turbulent friction drag reduction using electroactive polymer and electromagnetically driven surfaces,” Exp. Fluids **54**, 1441 (2013).

[21] C. Lee, J. Kim, D. Babcock, and R. Goodman, “Application of neural networks to turbulence control for drag reduction,” Phys. Fluids **9**, 1740–1747 (1997).

[22] R. J. Adrian, “Stochastic estimation of the structure of turbulent fields,” in *Eddy structure identification* (Springer, 1996) pp. 145–195.

[23] C. Tinney, F. Coiffet, J. Delville, A. Hall, P. Jordan, and M. N. Glauser, “On spectral linear stochastic estimation,” Exp. Fluids **41**, 763–775 (2006).

[24] C. E. Tinney, L. Ukeiley, and M. N. Glauser, “Low-dimensional characteristics of a transonic jet. Part 2. Estimate and far-field prediction,” J. Fluid Mech. **615**, 53–92 (2008).

[25] W. J. Baars, N. Hutchins, and I. Marusic, “Spectral stochastic estimation of high-reynolds-number wall-bounded turbulence for a refined inner-outer interaction model,” Phys. Rev. Fluids **1**, 054406 (2016).

[26] T. Suzuki and Y. Hasegawa, “Estimation of turbulent channel flow at $Re_\tau =$ 100 based on the wall measurement using a simple sequential approach,” J. Fluid Mech. **830**, 760–796 (2017).

[27] S. J. Illingworth, J. P. Monty, and I. Marusic, “Estimating large-scale structures in wall turbulence using linear models,” J. Fluid Mech. **842**, 146–162 (2018).

[28] M. P. Encinar, A. Lozano-Durán, and J. Jiménez, “Reconstructing channel turbulence from wall observations,” Procs. CTR Summer School, p. in press. Stanford Univ (2018).

[29] J. Borée, “Extended proper orthogonal decomposition: a tool to analyse correlated events in turbulent flows,” Exp. Fluids **35**, 188–192 (2003).

[30] S. Discetti, M. Raiola, and A. Ianiro, “Estimation of time-resolved turbulent fields through correlation of non-time-resolved field measurements and time-resolved point measurements,” Exp. Therm. Fluid Sci. **93**, 119–130 (2018).

[31] F. Rosenblatt, “The perceptron: a probabilistic model for information storage and organization in the brain.” Psychol. Rev. **65**, 386 (1958).

[32] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature **521**, 436 (2015).

[33] P. Baldi and K. Hornik, “Neural networks and principal component analysis: Learning from examples without local minima,” Neural Networks **2**, 53–58 (1989).

[34] J. Ling, A. Kurzawski, and J. Templeton, “Reynolds averaged turbulence modelling using deep neural networks with embedded invariance,” J. Fluid Mech. **807**, 155–166 (2016).

[35] J.-L. Wu, J.-X. Wang, H. Xiao, and J. Ling, “A priori assessment of prediction confidence for data-driven turbulence modeling,” Flow Turbul. Combust. **99**, 25–46 (2017).

[36] R. Maulik, O. San, A. Rasheed, and P. Vedula, “Subgrid modelling for two-dimensional turbulence using neural networks,” J. Fluid Mech. **858**, 122–144 (2019).

[37] Z. Nikolaou, C. Chrysostomou, L. Vervisch, and S. Cant, “Progress variable variance and filtered rate modelling using convolutional neural networks and flamelet methods,” Flow Turbul. Combust. **103**, 1–17 (2019).

[38] M. Raissi, Z. Wang, M. S. Triantafyllou, and G. E. Karniadakis, “Deep learning of vortex-induced vibrations,” J. Fluid Mech. **861**, 119–137 (2019).

[39] M. P. Hack and T. A. Zaki, “Data-enabled prediction of streak breakdown in pressure-gradient boundary layers,” J. Fluid Mech. **801**, 43–64 (2016).

[40] S. Brunton, B. Noack, and P. Koumoutsakos, “Machine learning for fluid mechanics,” Annu. Rev. Fluid Mech. (2020).

[41] M. Milano and P. Koumoutsakos, “Neural network modeling for near wall turbulent flow,” J. Comput. Phys. **182**, 1–26 (2002).

[42] R. Maulik and O. San, “A neural network approach for the blind deconvolution of turbulent flows,” J. Fluid Mech. **831**, 151–181 (2017).

[43] K. Fukami, K. Fukagata, and K. Taira, “Super-resolution reconstruction of turbulent flows with machine learning,” J. Fluid Mech. **870**, 106–120 (2019).

[44] B. Lusch, J. N. Kutz, and S. L. Brunton, “Deep learning for universal linear embeddings of nonlinear dynamics,” Nat. Commun. **9**, 4950 (2018).

[45] P. Srinivasan, L. Guastoni, H. Azizpour, P. Schlatter, and R. Vinuesa, “Predictions of turbulent shear flows using deep neural networks,” Phys. Rev. Fluids **4**, 054603 (2019).

[46] K. Fukushima, “Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position,” Biol. Cybern. **36**, 193–202 (1980).

[47] Y. LeCun, “A learning scheme for asymmetric threshold networks,” Proceedings of Cognitiva **85**, 599–604 (1985).

[48] K. Fukushima, “Neocognitron: A hierarchical neural network capable of visual pattern recognition,” Neural Networks **1**, 119–130 (1988).

[49] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, “Backpropagation applied to handwritten zip code recognition,” Neural Comput. **1**, 541–551 (1989).

[50] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, “A survey of deep neural network architectures and their applications,” Neurocomputing **234**, 11–26 (2017).

[51] J. Ou and Y. Li, “Vector-kernel convolutional neural networks,” Neurocomputing **330**, 253–258 (2019).

[52] K. Sasaki, R. Vinuesa, A. V. Cavalieri, P. Schlatter, and D. S. Henningson, “Transfer functions for flow predictions in wall-bounded turbulence,” J. Fluid Mech. **864**, 708–745 (2019).

[53] L. Guastoni, M. P. Encinar, P. Schlatter, H. Azizpour, and R. Vinuesa, “Prediction of wall-bounded turbulence from wall quantities using convolutional neural networks,” (submitted).

[54] J. L. Lumley, “The structure of inhomogeneous turbulent flows,” Atmospheric turbulence and radio wave propagation (1967).

[55] Y. Wu, “A study of energetic large-scale structures in turbulent boundary layer,” Phys. Fluids **26**, 045113 (2014).

[56] L. Sirovich, “Turbulence and the dynamics of coherent structures. I. Coherent structures,” Q. Appl. Math. **45**, 561–571 (1987).

[57] V. Nair and G. E. Hinton, “Rectified linear units improve restricted boltzmann machines,” in *Proceedings of the 27th international conference on machine learning (ICML-10)* (2010) pp. 807–814.

[58] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” arXiv preprint arXiv:1412.6980 (2014).

[59] Y. Li, E. Perlman, M. Wan, Y. Yang, C. Meneveau, R. Burns, S. Chen, A. Szalay, and G. Eyink, “A public turbulence database cluster and applications to study lagrangian evolution of velocity increments in turbulence,” J. Turbul. **9**, N31 (2008).

[60] H. Yu, K. Kanov, E. Perlman, J. Graham, E. Frederix, R. Burns, A. Szalay, G. Eyink, and C. Meneveau, “Studying lagrangian dynamics of turbulence using on-demand fluid particle tracking in a public turbulence database,” J. Turbul. **13**, N12 (2012).

[61] Z. Liu, R. Adrian, and T. Hanratty, “Large-scale modes of turbulent channel flow: transport and structure,” J. Fluid Mech. **448**, 53–80 (2001).

[62] M. Guala, S. Hommema, and R. Adrian, “Large-scale and very-large-scale motions in turbulent pipe flow,” J. Fluid Mech. **554**, 521–542 (2006).

[63] C. Sanmiguel Vila, R. Örlü, R. Vinuesa, P. Schlatter, A. Ianiro, and S. Discetti, “Adverse-pressure-gradient effects on turbulent boundary layers: statistics and flow-field organization,” Flow Turbul. Combust. **99**, 589–612 (2017).

[64] Z. Deng, Y. Chen, Y. Liu, and K. C. Kim, “Time-resolved turbulent velocity field reconstruction using a long short-term memory (lstm)-based artificial intelligence framework,” Phys. Fluids **31**, 075108 (2019).

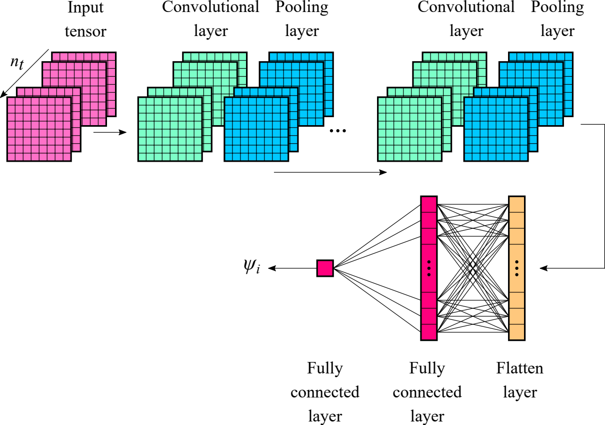
[65] Z.-C. Liu, R. J. Adrian, and T. J. Hanratty, “Reynolds number similarity of orthogonal decomposition of the outer layer of turbulent wall flow,” Phys. Fluids **6**, 2815–2819 (1994).

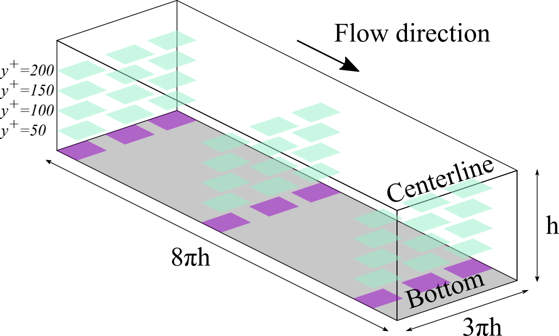
[66] H. Nakamura and S. Yamada, “Quantitative evaluation of spatio-temporal heat transfer to a turbulent air flow using a heated thin-foil,” Int. J. Heat Mass Tran. **64**, 892–902 (2013).

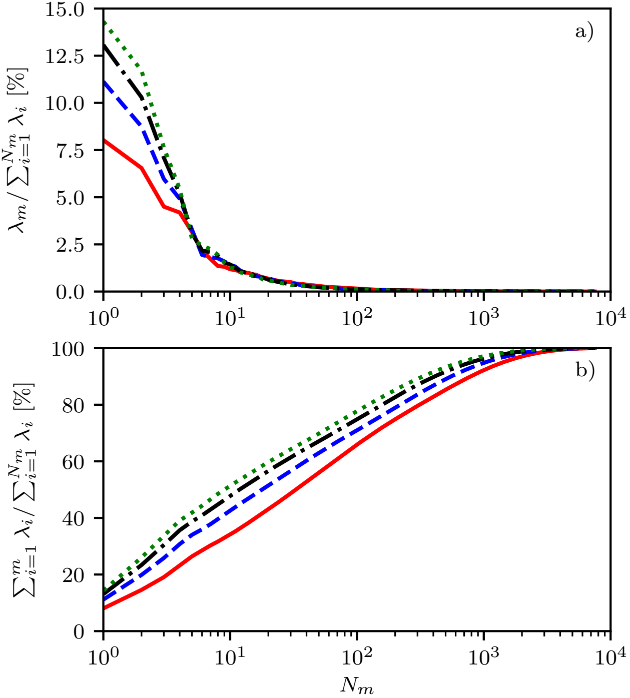
[67] M. Raiola, C. S. Greco, M. Contino, S. Discetti, and A. Ianiro, “Towards enabling time-resolved measurements of turbulent convective heat transfer maps with ir thermography and a heated thin foil,” Int. J. Heat Mass Tran. **108**, 199–209 (2017).

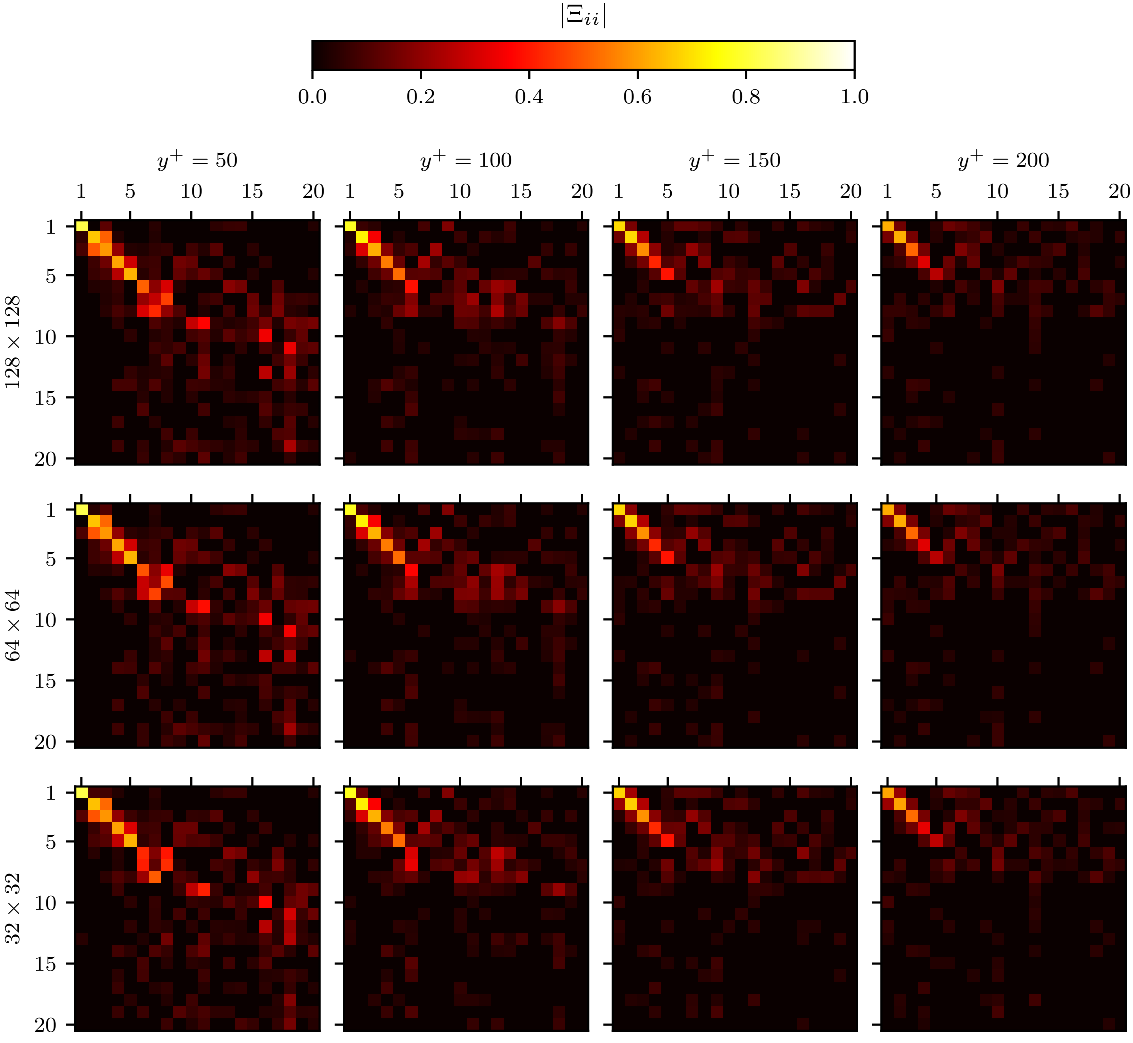
[68] F. Mallor, M. Raiola, C. S. Vila, R. Örlü, S. Discetti, and A. Ianiro, “Modal decomposition of flow fields and convective heat transfer maps: An application to wall-proximity square ribs,” Exp. Therm. Fluid Sci. **102**, 517–527 (2019).

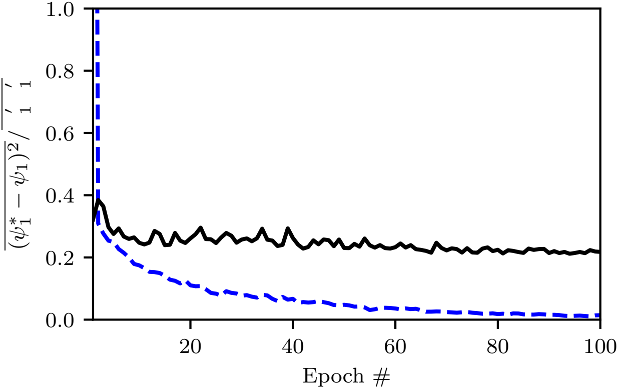
[69] R. Mayer, R. A. W. M. Henkes, and J. L. Van Ingen, “Quantitative infrared-thermography for wall-shear stress measurement in laminar flow,” Int. J. Heat Mass Tran. **41**, 2347–2356 (1998).

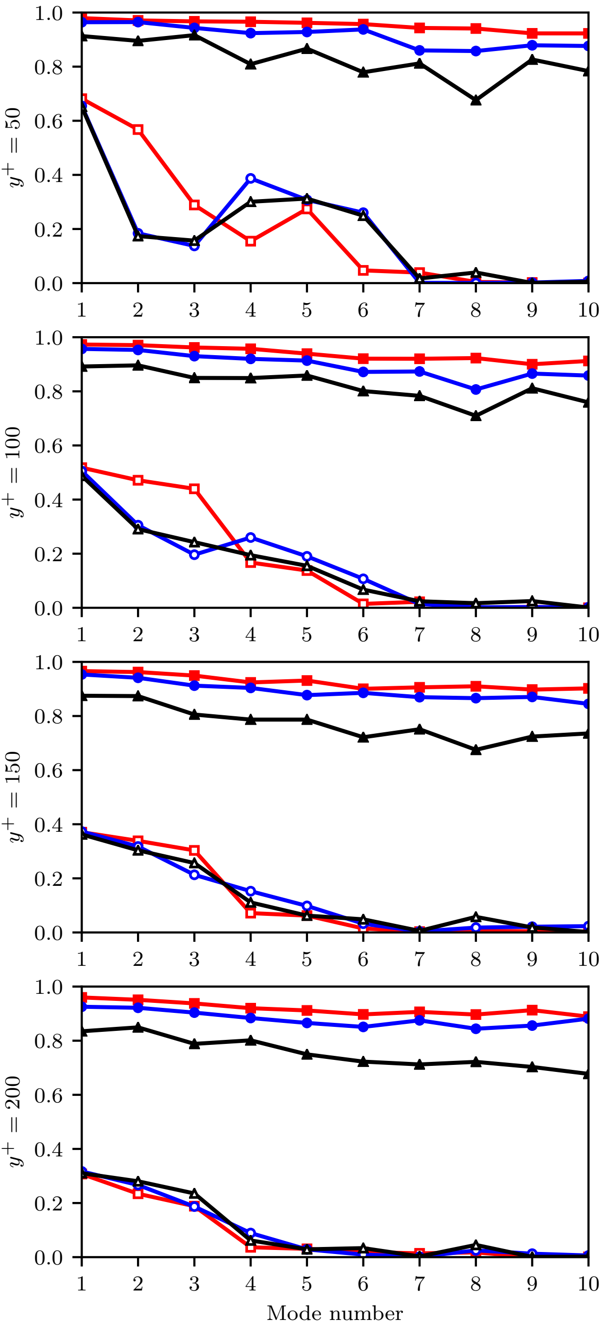

Input tensor
Convolutional layer
Pooling layer
Convolutional layer
Pooling layer
$n_t$
...
$\psi_i$
Fully connected layer
Fully connected layer
Flatten layer

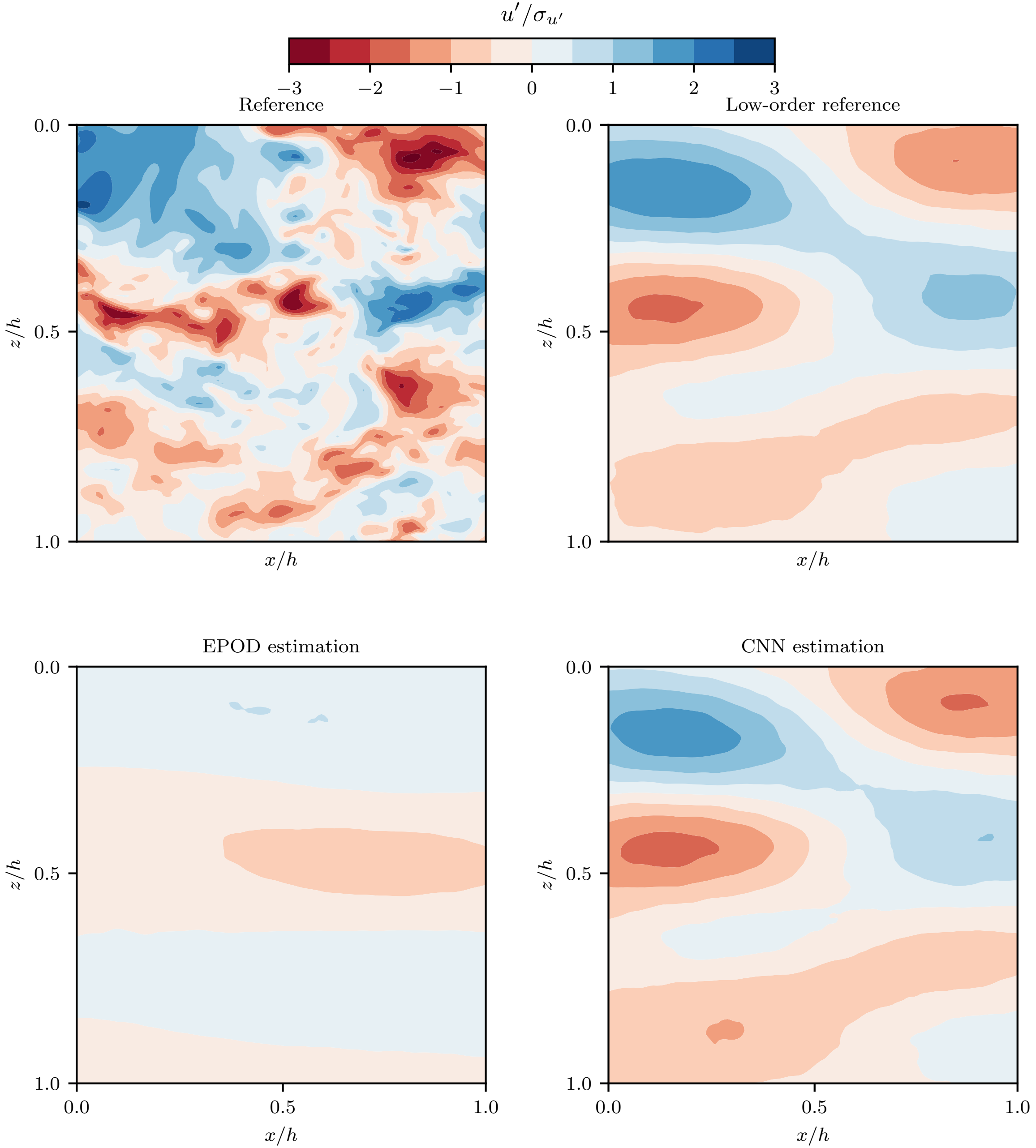

Flow direction
$y^+$=200
$y^+$=150
$y^+$=100
$y^+$=50
Centerline
h
8πh
Bottom
3πh

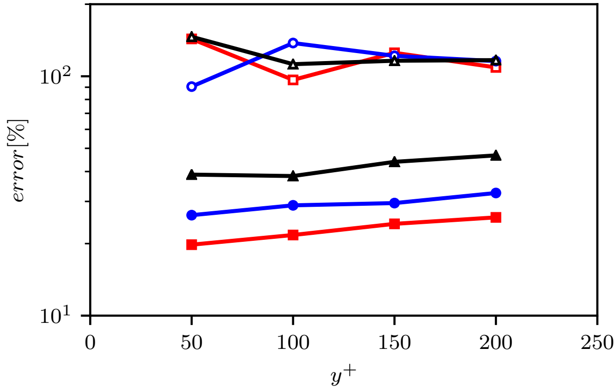

a)
15.0
12.5
10.0
7.5
5.0
2.5
0.0
$\lambda_m / \sum_{i=1}^{N_m} \lambda_i$ [%]
$10^0$
$10^1$
$10^2$
$10^3$
$10^4$
b)
100
80
60
40
20
0
$\sum_{i=1}^{m} \lambda_i / \sum_{i=1}^{N_m} \lambda_i$ [%]
$10^0$
$10^1$
$10^2$
$10^3$
$10^4$
$N_m$

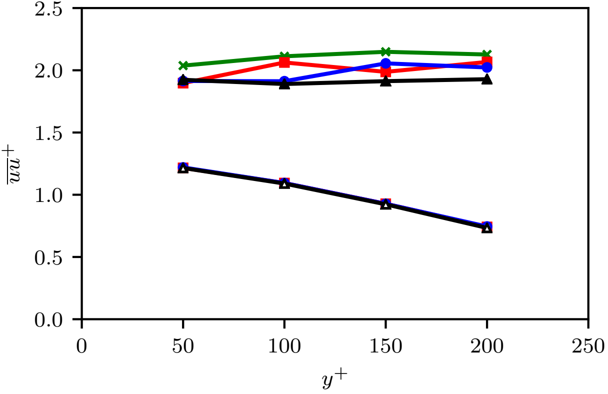

$|\Xi_{ii}|$
0.0
0.2
0.4
0.6
0.8
1.0
$y^+ = 50$
$y^+ = 100$
$y^+ = 150$
$y^+ = 200$
1
5
10
15
20
$128 \times 128$
$64 \times 64$
$32 \times 32$


$\overline{(\psi_1^* - \psi_1)^2}/\overline{{}'_1{}'_1}$
0.0
0.2
0.4
0.6
0.8
1.0
20
40
60
80
100
Epoch #


$y^+ = 50$
$y^+ = 100$
$y^+ = 150$
$y^+ = 200$
0.0
0.2
0.4
0.6
0.8
1.0
1
2
3
4
5
6
7
8
9
10
Mode number


$u'/\sigma_{u'}$
−3
−2
−1
0
1
2
3
Reference
Low-order reference
EPOD estimation
CNN estimation
$z/h$
$x/h$
0.0
0.5
1.0


$10^2$
$10^1$
$error[\%]$
0
50
100
150
200
250
$y^+$


2.5
2.0
1.5
1.0
0.5
0.0
$\overline{uu}^+$
0
50
100
150
200
250
$y^+$