11th International Symposium on Turbulence and Shear Flow Phenomena (TSFP11)
Southampton, UK, July 30 to August 2, 2019

IDENTIFYING THE TURBULENT-BOUNDARY-LAYER INTERFACE IN A TRANSITIONAL FLOW USING A SELF-ORGANIZING MAP

Zhao Wu, Jin Lee, Charles Meneveau* & Tamer Zaki
Department of Mechanical Engineering
Johns Hopkins University
Baltimore, MD 21211, USA
email: meneveau@jhu.edu

ABSTRACT

An unsupervised machine-learning algorithm, the self-organizing map (SOM), has been used to identify the turbulent boundary layer (TBL) and non-TBL regions in bypass transition (Wu *et al.*, 2019). In that study it was found that the SOM can successfully distinguish the near wall turbulence region from the free stream turbulence in the outer flow and the laminar streaky structures in the laminar/transition portion near the wall without choosing detector flow variables or thresholds. The turbulent boundary layer interface (TBLI) separating the TBL and non-TBL regions can be represented as a hyperplane in the high-dimensional space of the input features used by the SOM.

The prior study applied the SOM on the full-resolution DNS data. With an interest to applications in wall modeling for large eddy simulations (LES), in this paper the use of SOM is applied to Gaussian-filtered data in transitional boundary layer. Various approaches are compared and it is found that when using the full 15-dimensional inputs, the hyperplanes representing the TBLI with unfiltered (DNS) and filtered and coarsely-sampled (LES) data are almost the same and yield good results in identifying the interface. The results are not satisfactory, however, when using only surface variables based on equilibrium wall modeling concepts. It is concluded that unlike DNS, where only surface information was sufficient to properly identify the TBLI's wall signature, in coarse-resolution wall modeled LES, data characterizing the flow also away from the wall will be required to properly identify the wall signature of the TBLI.

INTRODUCTION

One of the most striking features of wall turbulence is the turbulent boundary layer interface (TBLI), a surface that separates near-wall turbulent boundary layer (TBL) regions from the non-TBL regions, the latter of which include the laminar boundary layer, laminar portions in the transitional region and the outer flow that can be either nearly irrotational or weakly turbulent. This TBLI has attracted interest in the research community over decades (Corrsin & Kistler, 1955; Anand *et al.*, 2009; Westerweel *et al.*, 2009; de Silva *et al.*, 2013; Chauhan *et al.*, 2014).

Since vorticity is a defining characteristic of turbulence (Bisset *et al.*, 2002; da Silva *et al.*, 2014; Jahanbakhshi & Madnia, 2016), it is often chosen as a detector flow variable. Even so, deciding how to threshold such a variable is challenging. Borrell & Jiménez (2016) normalized the vorticity in a particular way so that the choice of the threshold can be independent to the Reynolds number in the fully turbulent region. However, using vorticity faces difficulties in the transitional boundary layer, especially in the context of bypass transition (Zaki, 2013) for the following reasons: (i) the free-stream turbulence has vorticity which should not be treated as the TBL region; (ii) the streaky structures in the laminar/transition region which contain high wall-normal vorticity also should not be counted as part of the near-wall turbulent region; and (iii) the onset of turbulence is sporadic in space and time in the form of unique spots that result from various secondary instability mechanisms (Hack & Zaki, 2016).

To overcome these difficulties, Nolan & Zaki (2013) proposed a function based on the velocity fluctuations $|v'| + |w'|$ for bypass transition. Nevertheless, a single threshold is not possible in this 3D flow. As a result, these authors had to set different thresholds at different wall-normal heights and rebuilt the 3D turbulent boundary layer structures plane by plane. Lee & Zaki (2018) used a normalized vorticity function to eliminated the free stream turbulence first, but had to use a streamwise vorticity function to eliminated the streaks later.

The choice of the threshold can be still difficult even if a suitable detector flow variable has been selected. Usually, the threshold is chosen within the plateau or minimum of the PDF profile (da Silva *et al.*, 2014; Lee *et al.*, 2017). However, the range of the plateau may be over one order of magnitude (Borrell & Jiménez, 2016). Also in some cases where no plateau or minimum is seen in the PDF profile, selecting a threshold becomes a trial-and-error process and often must rely on the researcher's subjective judgment.

Recently, Wu *et al.* (2019) used an unsupervised machine learning algorithm, the self-organizing map (SOM) (Kohonen, 2001), as an automatic classifier to detect the TBLI in a transitional boundary layer with free stream turbulence. The magnitudes of velocity, velocity fluctuations and velocity gradients normalized by their standard deviations are used as the input features to describe the flow. There is no need to choose any particular functions (e.g. normalized vorticity, or the combination of $|v'|$ and $|w'|$) and threshold it, which is an advantage of this method. It has been shown the data points in the entire flow domain are automatically classified into TBL and non-TBL regions by the SOM, and what is more important is that the free-stream turbulence and the near-wall streaky structures are classified as the non-TBL region as desired. Thus the SOM identifies

the turbulent-boundary-layer interface (TBLI) without the usual need for choosing thresholds on e.g. vorticity or velocity fluctuations. The TBLI is found to be a hyperplane in the input space, and the hyperplane functions can be used to very efficiently classify other instances of the same or similar flow; for example, it was successfully applied to data from the simulations by You & Zaki (2019) in order to separate a much higher Reynolds number turbulent boundary layer (up to $Re_\tau = 1000$) from an energetic turbulent free stream.

There are many uses to identify the TBLI: most applications have focused on fundamental questions such the rates of entrainment, the mechanisms and scales responsible for accompanying engulfment processes, and the universality of the turbulence within the spot (Marxen & Zaki, 2019). Another application, which to date has not received attention, is for large eddy simulations (LES). In wall-modeled LES of transitional boundary layer flow, one expects that different wall models will be required depending on whether an LES grid point is in a laminar portion of the flow or inside turbulent spots of the fully turbulent region. Typically the choice in modeling will require binary decisions based on knowledge of in which region a particular LES grid point is located. The question thus arises whether the SOM identified from DNS can also be utilized in the context of filtered (coarse-grained) variables such as filtered velocity fluctuations and gradients.

After describing the data set used in this study and summarizing the results from SOM applied to DNS data, we present various analyses aiming at using SOM for filtered variables to identify a the TBLI as it can be represented in filtered fields and LES resolutions.

TRANSITIONAL BOUNDARY-LAYER DATA

The transitional boundary layer with free-stream turbulence, simulated using DNS by Lee & Zaki (2018) and stored in the FileDB system of the Johns Hopkins Turbulence Database (JHTDB) (`http://turbulence.pha.jhu.edu`) (Perlman *et al.*, 2007; Li *et al.*, 2008), is used in this study, same as that in Wu *et al.* (2019). The flow configuration is shown in Figure 1 and more details about the simulation can be found at `https://doi.org/10.7281/T17S7KX8` and Wu *et al.* (2019). The TBLI identification in this flow is difficult, because the free-stream turbulence and the streaky structures in the laminar/transitional region near the wall both contain non-negligible levels of vorticity, and yet these regions should not be treated as part of the near-wall turbulent region.

The streamwise, wall-normal and spanwise axes are denoted as x, y and z, and the corresponding velocity components are u, v and w. In the rest of paper, x, y and z are normalized by the boundary layer thickness δ_{99} at the inlet of the flow domain and the $(\cdot)^+$ designation represents data normalized by $u_\tau = 0.0444$ at the streamwise location x_0 where the wall intermittency function is equal to 0.5 (i.e. half of the time a point is within the wall turbulent region and outside the other half). The kinematic viscosity is ν. The grid is uniform in the x and z directions, and stretched in the y direction. The grid spacing in the DNS and also stored in the database is $\Delta x^+ = 10.4$, $\Delta y^+_{\text{min}} = 0.26$ and $\Delta z^+ = 4.2$. For reference, the boundary-layer thickness at x_0 is $\delta_{99,x_0} = 5.25$, or $\delta^+_{99,x_0} = 186.5$.

SELF-ORGANIZING MAP (SOM)

Using an SOM, Wu *et al.* (2019) showed successful identification of the TBLI in the transitional boundary layer flow described in the previous section. The SOM (Kohonen, 2001) is an unsupervised machine learning clustering algorithm, which means models using unsupervised learning must learn relationships between elements in a data set and classify the raw data without "help". The data need not be "labeled", meaning that we do not need to know ahead of time how to distinguish the flow regions. The SOM consists of M nodes and in the present study $M = 2$ nodes were used since we only wish to classify each point in the flow as either TBL or non-TBL. Each of the nodes has a position ("weight") in the space of input vectors that need to be classified. The SOM then iteratively finds the locations of these M nodes in the input space that best "clusters" the data around the nodes. And, a linear hyperplane based on Euclidean distance in the data space is constructed to classify all points into belonging to either nodes.

In Wu *et al.* (2019) such an SOM was first tested on a two-dimensional subdomain of the flow, namely the wall. There, only three input variables were used, proportional to the two components of the wall stress and the downstream distance x. In SOM, all input variables were normalized by their root-mean-square (r.m.s.). The application of the SOM to these input data yielded two clusters. One clearly corresponded to the laminar region on the wall including streaks signature, and the other was the fully turbulent region. Then the SOM was applied to a fully 3D domain that included weak outer turbulence, streaky laminar regions near the wall before transition to turbulence, patches of turbulence, and the fully turbulent boundary layer. A 16-dimensional input vector consisted of the following pointwise variables: magnitudes of velocity, their fluctuations, velocity gradients, and x and y position. SOM again classified the 16-dimensional data into two groups. Each point in the flow was compared to the two nodal positions and, depending on its (Euclidean) distance in the space of normalized variables, classified accordingly. Visualizing the resulting two regions and the TBLI between them, it was evident that SOM classification results were consistent with the visual appearance of the flow. The classification was a hyperplane in 16-dimensional state space. It was observed that the respective coefficients were all non-negligible. As a result, it was concluded that none of the input variables used were unimportant and thus could not be discarded.

In the present study, the SOM is applied to filtered data to mimic data available in wall-modeled LES. Results are then compared with the previous ones obtained from DNS. We followed the procedure of SOM described in Wu *et al.* (2019) and the "batch" version of SOM is used. In this work, we only focus on the TBLI identification on a particular wall-parallel slice, i.e. with a constant y. Therefore the y coordinates that were used in Wu *et al.* (2019) as part of the input features is omitted here and only 15 variables remain: they are the magnitudes of filtered velocity, filtered velocity fluctuations and filtered velocity gradients as well as the streamwise coordinate x. In contrast to Wu *et al.* (2019) who used the original data generated from DNS to compute the input features, this work uses filtered velocities to calculate the input features. The DNS data are first interpolated onto a coarse and uniform grid with the grid spacing $\Delta x^+ = \Delta y^+ = \Delta z^+ = 10$. The data are then filtered with a Gaussian filter in the wall-parallel plane with the filter width $\Delta^+ = 60$. The velocity gradients are calculated on

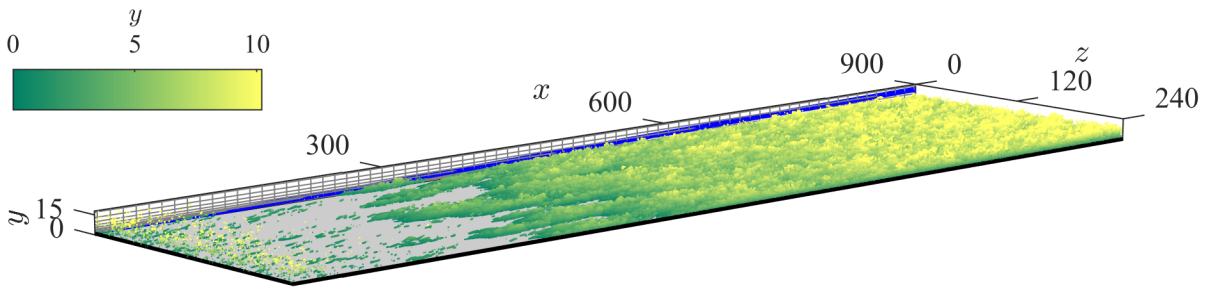

Figure 1. Flow configuration of the transitional boundary layer data set used in the present study. The main flow is in the x direction. The structures are identified by the iso-surface of Q-criterion and colored by the wall-normal heights. The boundary layer thickness δ_{99} is shown as the blue line.

this coarse grid. Finally, these input features are all normalized by their own standard deviation.

Same to the previous study, the number of the clusters is set to $M = 2$, with the expectation that two clusters will represent the TBL and non-TBL regions in the filtered velocity representation, respectively. After training, the SOM would output the position of the two nodes in the input feature space. Those data points closer to one node in the input space are classified as one group (say TBL), while the other points are the other group (say non-TBL). Again, the bisecting plane of the two nodes then corresponds to the TBLI.

RESULTS

Figure 2 shows the results at $y^+ = 40$, which is at the lower-bound of the log-law region and could correspond to the location of the first grid-point away from the wall in a wall-modeled LES. The TBLI identified by the SOM based on the unfiltered data and the gradients calculated on the original DNS grid, which is used as the reference result, is presented in Figure 2 (a) and will be denoted as unfiltered-SOM TBLI hereafter.

We can write the representation of the hyperplane function representing the unfiltered-SOM TBLI according to

$$\sum_i c_i f_i + 1 = 0, \tag{1}$$

where

$$c_i = \{-0.42, 0.41, 0.36, -0.05, 0.41, 0.34, -0.46, \\ -0.07, 0.40, 0.38, 0.40, 0.40, 0.33, 0.37, 0.60\}$$

are the coefficients of the input features (all are normalized by their r.m.s.):

$$f_i = \{|u|, |v|, |w|, |u'|, |v'|, |\partial u/\partial x|, |\partial u/\partial y|, |\partial u/\partial z|, \\ |\partial v/\partial x|, |\partial v/\partial y|, |\partial v/\partial z|, |\partial w/\partial x|, |\partial w/\partial y|, |\partial w/\partial z|\}.$$

The data point would be categorized as TBL if $\sum_i c_i f_i + 1 > 0$, non-TBL otherwise. In Wu *et al.* (2019) the SOM was applied simultaneously in the 3D domain and there all coefficients were found to be relevant. Here, at the single plane at $y^+ = 40$, we find that the coefficients of $|u'|$ and $|\partial u/\partial z|$ are one order of magnitude smaller than others, which indicates that $|u'|$ and $|\partial u/\partial z|$ are less relevant in determining the TBLI. This may be because at this height above the wall $|u'|$ and $|\partial u/\partial z| = |\partial u'/\partial z|$ are appreciable in the streaks and also within the turbulent boundary layer; as a result, they are not good discriminators. The result shows that the SOM has the ability to discover and disregard locally ineffective inputs.

Next, we apply the SOM to determine the hyperplane using filtered input features and gradients calculated on a coarsen grid. Specifically, now the input features are given by

$$\mathbf{f}_\sim = \{|\tilde{u}|, |\tilde{v}|, |\tilde{w}|, |\tilde{u}'|, |\tilde{v}'|, |\partial \tilde{u}/\partial x|, |\partial \tilde{u}/\partial y|, |\partial \tilde{u}/\partial z|, \\ |\partial \tilde{v}/\partial x|, |\partial \tilde{v}/\partial y|, |\partial \tilde{v}/\partial z|, |\partial \tilde{w}/\partial x|, |\partial \tilde{w}/\partial y|, |\partial \tilde{w}/\partial z|\}.$$

Again, each variable is normalized by its own r.m.s. The result from the classification is shown in Figure 2 (b), and will be denoted as filtered-SOM TBLI hereafter. The u and $\tilde{u}$ contours are also shown in (a) and (b) respectively. In both plots, as desired, the streaky structures with streamwise vorticity are not classified as TBL by the SOM. The visualizations of the TBLI confirm that the classification results are consistent with the visual appearance of the flow in both the cases of unfiltered data and filtered data. Figure 2 (c) shows the comparison between the unfiltered-SOM TBLI (in color) and filtered-SOM TBLI (black line) in a zoomed region. The two interfaces are very close to each other. Although the filtered-SOM interface is smoother than the unfiltered one, it still retains significant spatial complexity and details.

In the paper by Wu *et al.* (2019), the SOM method with a hyperplane function trained from one snapshot was successfully applied to another snapshot of the same flow to determine the TBLI. The resulting TBLI was essentially identical to that found by applying SOM to the latter snapshot directly. This was due to the fact that a single snapshot contained sufficient samples to provide a (nearly) complete state-space representation for training an accurate TBLI identification model.

In the present context, one might wonder how a model trained with unfiltered (DNS-level) data performs in the case of filtered data in the LES context. Figure 3 shows the TBLI obtained by applying the unfiltered-SOM TBLI hyperplane to the filtered inputs: that is applying the hyperplane function Eq. 1 to the filtered input features. It is found that the differences of this TBLI and the filtered TBLI are negligible. This suggested the unfiltered-SOM TBLI and the filtered-SOM TBLI should have the same hyperplane functions: we find that the coefficients of the two hyperplanes differ by about 5% (excluding the coefficients of $|u'|$

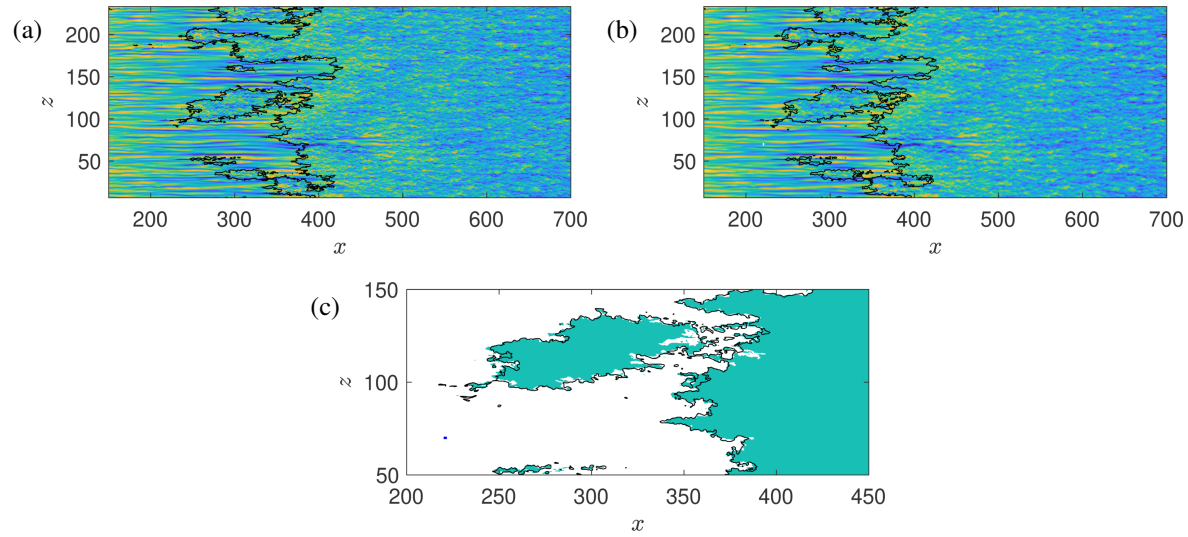

Figure 2. (a) u contour and unfiltered-SOM TBLI (black line) at $y^+ = 40$. (b) $\tilde{u}$ contour and filtered-SOM TBLI (black line) at $y^+ = 40$. (c) The comparison of unfiltered-SOM TBLI (color) and filtered-SOM TBLI (black line) zoomed in the region of $200 < x < 450$ and $50 < z < 150$. The tiny blue dot shows the filter width $\Delta^+ = 60$.

and $|\partial u/\partial z|$, which are one order of magnitude smaller than others).

We further test the performance of SOM on filtered inputs but with larger filter width. Figure 4 shows the results at $y^+ = 40$ but for filter width of $\Delta^+ = 200$. Although one could hardly identify the interface by visually checking the contours directly with such a large filter, the SOM provides acceptable TBLI identification in this case as well. By comparing with the unfiltered-SOM TBLI, it is found that the filtered-SOM TBLI roughly follows the shape of the unfiltered-SOM TBLI, but as expected many details are missing.

At the wall, in the DNS level SOM, the number of SOM input features reduces to three ($|\partial u/\partial y|$, $|\partial w/\partial y|$ and x), because all other variables are zero. This approach also gave good results and the TBLI's intersection with the wall appeared to be well predicted by the SOM applied to the three-variable case.

In LES using a wall model, the wall shear stresses (or vertical gradients $|\partial u/\partial y|_{y=0}$, $|\partial w/\partial y|_{y=0}$) are in principle unknown and need to be modeled based on resolved information, typically the LES velocity at the first (or second) grid point away from the wall. The equilibrium wall model uses the assumption of a locally valid logarithmic law and solves for the wall stress based on the velocity at the grid point away from the wall. Then the stress is essentially proportional to the (square) of filtered velocity there. Thus the unknown vertical gradients $|\partial u/\partial y|_{y=0}$, $|\partial w/\partial y|_{y=0}$ are related to the filtered velocity components at the first grid point away from the wall in LES. Thus, to mimic this behaviour and evaluate the performance of SOM in this scenario, we use $|\tilde{u}|$ and $|\tilde{w}|$ and x as input variables, as wall model surrogates of $|\partial u/\partial y|_{y=0}$, $|\partial w/\partial y|_{y=0}$ and x. The result is shown at Figure 5. As can be seen, this identification method is entirely unsatisfactory. Specifically, it is found that the streaky structures are now erroneously included in the TBL region. This behavior is reminiscent of earlier findings that motivated detector flow variables. Since the Klebanoff streaks are predominately streamwise velocity perturbations Nolan & Zaki (2013) constructed their detector flow variables for the transitional boundary layer based on the sum of the absolute values of the wall-normal and spanwise fluctuation field, so as to exclude the streamwise component. Clearly, using the equilibrium wall modeled inspired variables to characterize the TBLI at the wall does not work. Instead, using full 15-dimensional 3D information inside the flow near the wall does provide good results.

CONCLUSIONS

In the present study, we investigated the performance of SOM-based TBLI identification in the context of large eddy simulations. A DNS data set of transitional boundary layer is used. The data are filtered with 2D Gaussian filter in the wall-parallel plane and represented on a coarsened grid which is uniform in all three directions. The velocity gradients are evaluated on this coarse and uniform grid.

A horizontal plane in the log-law region is selected to apply the SOM using the filtered data as input. Results show that the streaky structures are not selected into the TBL region. The identification of TBLI is satisfactory when visually compared with the $\tilde{u}$ contour. In addition, the filtered- and unfiltered-SOM TBLI are almost the same, only that the former is smoother, as can be expected in LES. Only a few small structures are missing in the filtered-SOM TBLI since the Gaussian filter removes these small features.

The hyperplane function obtained from training the SOM on unfiltered DNS data is also applied on the filtered data, and a satisfactory TBLI is obtained. In fact, by comparing this TBLI and the filtered-SOM TBLI, one could hardly see any differences between the two. While using the 15-dimensional input variables in the flow field near the wall, a method inspired by wall modeling using variables only 'at' the wall did not yield good results. We showed that it is insufficient to use only $\tilde{u}$, $\tilde{w}$ and x as the input features for the TBLI identification on a plane in the log-law region (which using wall-modeling concepts could be interpreted as surrogates for the modeled wall stress): streaky structures would be classified as TBL in this case. Hence, results suggest that more information about the flow is required to properly distinguish between TBL and non-TBL regions in such a flow.

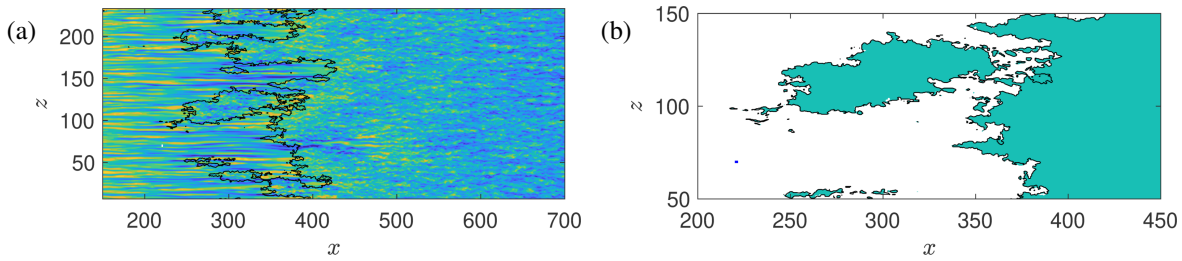

Figure 3. (a) $\tilde{u}$ contour at $y^+ = 40$ and the TBLI (black line) obtained by applying unfiltered-SOM TBLI model to the filtered inputs. (b) The comparison of filtered-SOM TBLI (color) and the TBLI in (a) (black line) zoomed in the region of $200 < x < 450$ and $50 < z < 150$. The small blue dot shows the filter width $\Delta^+ = 60$.

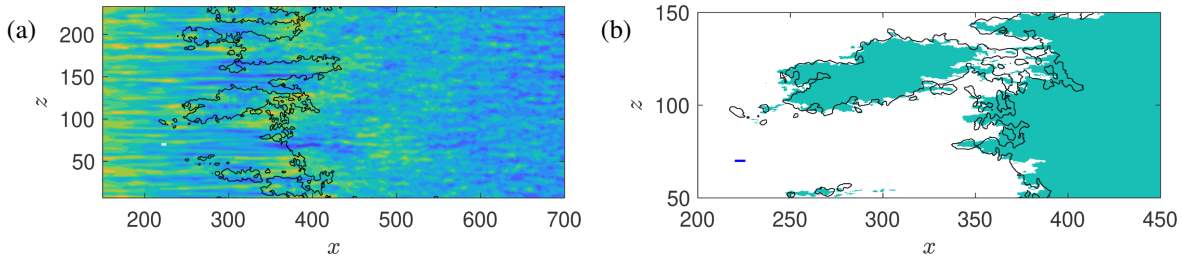

Figure 4. (a) $\tilde{u}$ contour and filtered-SOM TBLI (black line) at $y^+ = 40$. (b) The comparison of unfiltered-SOM TBLI (color) and filtered-SOM TBLI (black line) zoomed in the region of $200 < x < 450$ and $50 < z < 150$. The blue bar shows the filter width $\Delta^+ = 200$.

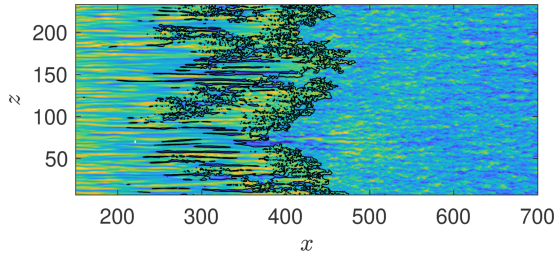

Figure 5. Filtered streamwise velocity ($\tilde{u}$) contours and filtered-SOM TBLI (black line) at $y^+ = 40$. Inspired byu wall-modeled LES using the equilibrium wall model, here only three variables are used as the input features: $\tilde{u}$, $\tilde{w}$ and x coordinates. The filter width is $\Delta^+ = 60$.

ACKNOWLEDGMENTS

The authors acknowledge funding from the Office of Naval Research (grant N00014-17-1-2937) and the National Science Foundation (grant OCE-1633124). Computations were made possible by the Maryland Advanced Research Computing Center (MARCC). Development of FileDB was also supported by NSF grant OAC-1261715.

REFERENCES

Anand, Ravi Kumar, Boersma, B. J. & Agrawal, Amit 2009 Detection of turbulent/non-turbulent interface for an axisymmetric turbulent jet: Evaluation of known criteria and proposal of a new criterion. *Experiments in Fluids* **47** (6), 995–1007.

Bisset, David K., Hunt, Julian C. R. & Rogers, Michael M. 2002 The turbulent/non-turbulent interface bounding a far wake. *Journal of Fluid Mechanics* **451**, 383–410.

Borrell, Guillem & Jiménez, Javier 2016 Properties of the turbulent/non-turbulent interface in boundary layers. *Journal of Fluid Mechanics* **801**, 554–596.

Chauhan, Kapil, Philip, Jimmy, De Silva, Charitha M., Hutchins, Nicholas & Marusic, Ivan 2014 The turbulent/non-turbulent interface and entrainment in a boundary layer. *Journal of Fluid Mechanics* **742**, 119–151.

Corrsin, Stanley & Kistler, Alan L 1955 Free-stream boundaries of turbulent flows.

Hack, M. J.Philipp & Zaki, Tamer A. 2016 Data-enabled prediction of streak breakdown in pressure-gradient boundary layers. *Journal of Fluid Mechanics* **801**, 43–64.

Jahanbakhshi, Reza & Madnia, Cyrus K. 2016 Entrainment in a compressible turbulent shear layer. *Journal of Fluid Mechanics* **797**, 564–603.

Kohonen, Teuvo 2001 *Self-Organizing Maps*, *Springer Series in Information Sciences*, vol. 30. Berlin, Heidelberg: Springer Berlin Heidelberg.

Lee, Jin, Sung, Hyung Jin & Zaki, Tamer A. 2017 Signature of large-scale motions on turbulent/non-turbulent interface in boundary layers. *Journal of Fluid Mechanics* **819**, 165–187.

Lee, Jin & Zaki, Tamer A 2018 Detection algorithm for turbulent interfaces and large-scale structures in intermittent flows. *Computers and Fluids* **175**, 142–158.

Li, Yi, Perlman, Eric, Wan, Minping, Yang, Yunke, Meneveau, Charles, Burns, Randal, Chen, Shiyi, Szalay, Alexander & Eyink, Gregory 2008 A public turbulence

database cluster and applications to study Lagrangian evolution of velocity increments in turbulence. *Journal of Turbulence* **9**, N31.

Marxen, O. & Zaki, T.A. 2019 Turbulence in intermittent transitional boundary layers and in turbulence spots. *Journal of Fluid Mechanics* **860**, 350–383.

Nolan, Kevin P. & Zaki, Tamer A. 2013 Conditional sampling of transitional boundary layers in pressure gradients. *Journal of Fluid Mechanics* **728**, 306–339.

Perlman, Eric, Burns, Randal, Li, Yi & Meneveau, Charles 2007 Data exploration of turbulence simulations using a database cluster. *Proceedings of the 2007 ACM/IEEE Conference on Supercomputing (SC '07)* p. 1.

da Silva, Carlos B., Hunt, Julian C.R., Eames, Ian & Westerweel, Jerry 2014 Interfacial Layers Between Regions of Different Turbulence Intensity. *Annual Review of Fluid Mechanics* **46** (1), 567–590.

de Silva, Charitha M., Philip, Jimmy, Chauhan, Kapil, Meneveau, Charles & Marusic, Ivan 2013 Multiscale geometry and scaling of the turbulent-nonturbulent interface in high reynolds number boundary layers. *Physical Review Letters* **111** (4), 044501.

Westerweel, J., Fukushima, C., Pedersen, J. M. & Hunt, J. C.R. 2009 Momentum and scalar transport at the turbulent/non-turbulent interface of a jet. *Journal of Fluid Mechanics* **631**, 199–230.

Wu, Zhao, Lee, Jin, Meneveau, Charles & Zaki, Tamer A. 2019 Application of a self-organizing map to identify the turbulent-boundary-layer interface in a transitional flow. *Physical Review Fluids* **4** (2), 023902.

You, Jiho & Zaki, Tamer A 2019 Conditional statistics and flow structures in turbulent boundary layers buffeted by free-stream disturbances. *Journal of fluid Mechanics* **866**, 526–566.

Zaki, Tamer A. 2013 From streaks to spots and on to turbulence: Exploring the dynamics of boundary layer transition. In *Flow, Turbulence and Combustion*, , vol. 91, pp. 451–473. Springer Netherlands.