Physics of Fluids

RESEARCH ARTICLE | JUNE 18 2024

Characterizing turbulence structures in convective and neutral atmospheric boundary layers via Koopman mode decomposition and unsupervised clustering

Milad Rezaie ; Mostafa Momen

Physics of Fluids 36, 066605 (2024)
https://doi.org/10.1063/5.0206387

Physics of Fluids ARTICLE pubs.aip.org/aip/pof

Characterizing turbulence structures in convective and neutral atmospheric boundary layers via Koopman mode decomposition and unsupervised clustering

Cite as: Phys. Fluids **36**, 066605 (2024); doi: 10.1063/5.0206387
Submitted: 1 March 2024 · Accepted: 28 May 2024 ·
Published Online: 18 June 2024

Milad Rezaie and Mostafa Momen[a)]

AFFILIATIONS

Department of Civil and Environmental Engineering, University of Houston, Houston, Texas, USA

[a)]**Present address:** Prof. Momen, Room N134, 4726 Calhoun Road, Department of Civil and Environmental Engineering, University of Houston, TX 77204. **Author to whom correspondence should be addressed:** mmomen@uh.edu

ABSTRACT

The atmospheric boundary layer (ABL) is a highly turbulent geophysical flow, which has chaotic and often too complex dynamics to unravel from limited data. Characterizing coherent turbulence structures in complex ABL flows under various atmospheric regimes is not systematically well established yet. This study aims to bridge this gap using large eddy simulations (LESs), Koopman theory, and unsupervised classification techniques. To this end, eight LESs of different convective, neutral, and unsteady ABLs are conducted. As the ratio of buoyancy to shear production increases, the turbulence structures change from roll vortices to convective cells. The quadrant analysis indicated that as this ratio increases, the sweep and ejection events decrease, and inward/outward interactions increase. The Koopman mode decomposition (KMD) is then used to characterize their turbulence structures. Our results showed that KMD can reveal non-trivial modes of highly turbulent ABL flows (e.g., transverse to the mean flow direction) and can reconstruct the primary dynamics of ABLs even under unsteady conditions with only ∼5% of the modes. We attributed the detected modes to the imposed pressure gradient (shear), Coriolis (inertial oscillations), and buoyancy (convection) forces by conducting novel timescale and quadrant analyses. We then applied the convolutional neural network combined with the K-means clustering to group the Koopman modes. This approach is displacement and rotation invariant, which allows efficiently reducing the number of modes that describe the overall ABL dynamics. Our results provide new insights into the dynamics of ABLs and present a systematic data-driven method to characterize their complex spatiotemporal patterns.

I. INTRODUCTION

Real-world atmospheric flows exhibit highly chaotic and complex dynamics due to the interaction of different forces acting on them.[1] The chaotic motions of the atmospheric boundary layer (ABL) flows stem from their highly turbulent nature, which is characterized by a very high Reynolds number of $\sim 10^8$.[2] Accurate representation of turbulence dynamics in these strongly nonlinear systems is becoming increasingly critical for many applications such as the design of wings, flow control, wind energy, and weather and climate models.[3–7] It is thus imperative to characterize the underlying dynamics of coherent turbulent motions in airflows and develop predictive data-driven models that can guide us to forecast turbulence. In the absence of such knowledge, engineering designs can fail to efficiently respond to turbulence fluctuations, and numerical models will be unable to accurately predict the evolution of ABLs.

In addition to the challenges posed by turbulent flows in general, the ABL presents its own set of complexities. For decades, many studies attempted to detect and describe coherence in highly nonlinear ABL flows.[8,9] Many factors can simultaneously influence the dynamics of real-world ABLs such as unsteadiness, baroclinicity, surface heterogeneity, rotation, and stratification.[10–15] Stratification can greatly influence the transport of heat,[16,17] turbulent mixing,[18] the ABL height,[19] and the inclination angle of turbulent structures.[20] Rotation was also shown to have a similar effect to stratification where it can alter the size and orientation of turbulent eddies.[21–24] The unsteadiness of the ABL due to, e.g., variable surface heat flux[25–27] or pressure gradients,[28,29] can significantly impact turbulence dynamics in these flows. Baroclinicity can influence turbulence production in the ABL and lead to different coherent turbulence structures as shown by our prior studies.[30,31]

Roll vortices and convection cells are two well-known coherent turbulence structures in ABLs. Previous large eddy simulation (LES)[8] and observational[32] studies showed that under strong mean shear and near-neutral conditions, horizontal rolls are organized in the ABL. Nevertheless, recent LESs of atmospheric flows have demonstrated that as the surface heat flux increases and mean shear decreases, open cell structures form in convective ABLs.[30,33] Depending on the stability of ABLs, one of these turbulence structures can dominate or both can coexist.[34] Detecting such turbulence structures can be challenging in real-world ABLs where both structures can coexist and other factors in the ABL such as unsteadiness and surface heterogeneities can further complicate the formation of the coherent structures.

The turbulent flow structures are often analyzed through mode decomposition techniques.[35] Given the strongly nonlinear nature of atmospheric flows, there is presently no prevalent data-driven framework for the general characterization and reconstruction of turbulence dynamics and coherent structures in different ABL regimes. Common methods for embedding large-scale dynamics in low-dimensional space include proper orthogonal decomposition (POD) for nonlinear flows[5] and global eigenmodes for linearized dynamics.[36] POD ranks turbulence structures according to their energy content by diagonalizing the covariance matrix. Since this method averages in time, it may not allocate one coherent structure to one mode. In some studies, for the wake flows behind cylinders, higher POD modes were unable to show significant physical mechanisms likely due to contamination with uncorrelated structures.[37]

Koopman mode decomposition (KMD) is another technique that has recently attracted attention in the fluid dynamics community.[38–40] The Koopman operator, $\mathcal{K}$, allows us to shift the focus of the flow analysis from nonlinear finite-dimensional dynamics to linear but infinite-dimensional dynamics. The main advantage of such mapping is that the linear systems can be simply solved and understood by linear operator theory and spectral decomposition.[41] The main idea is that instead of analyzing the original highly nonlinear system, by constructing a Koopman operator, the flow dynamics can be evolved by trivially computing the time dynamics of a linear system using the eigenvalues and eigenfunctions of $\mathcal{K}$.

The capability of Koopman mode approximations in linearizing the representation of many fluid dynamical systems has been demonstrated in recent studies.[36,42,43] KMD has been shown to decouple different frequency components in shear layers more effectively than POD.[36,37,44] Moreover, KMD was able to capture global modes in the flow using local probe signal measurements.[36] To obtain Koopman modes directly from data, the dynamic mode decomposition (DMD) algorithm is often used to approximate the Koopman operator with a best-linear fit model.[36,41] The Koopman operator then generates a linear flow map that approximates the full nonlinear dynamics.

Unraveling and classifying non-trivial dynamics and flow modes in fluids can be a challenging task, especially for strongly nonlinear systems. In recent years, a method that has become popular is the utilization of unsupervised machine learning techniques, such as K-means clustering.[45] K-means clustering has been successfully used to categorize stress anisotropy[46] and POD modes in the wake of wind turbines.[47] It is also shown that despite being sensitive to noise and requiring careful selection of input parameters, this technique is capable of recognizing similarities and discrepancies in various flows, thereby providing an understanding of the flow's underlying physics.[48,49] Despite all challenges, machine learning techniques exhibit significant potential in the analysis and classification of complex turbulent flow structures.

Convolutional neural networks (CNNs) are a powerful machine learning method that can be used to classify turbulence structures. Deep CNNs have become a popular technique for automatically learning hierarchical features from images.[50–52] CNNs have been successfully used in different atmospheric flow detection studies.[53,54] In this paper, we propose to leverage the power of CNNs along with the K-means clustering method in a novel way to classify the Koopman modes of turbulent ABL flows.

Although KMD has been used in previous studies to analyze fluid flows,[37,55,56] its application to detecting coherent turbulence structures in high-Reynolds number ABL flows is not yet well established. Furthermore, a data-driven unsupervised method to categorize the obtained flow modes into reduced distinctive clusters that can adequately describe the underlying dynamics is lacking. In this study, we aim to bridge these knowledge gaps using a suite of LESs, and the KMD approach in conjunction with the K-means clustering method. By conducting different LESs of convective and neutral ABLs, we will characterize the efficiency of these data-driven methods and use them to detect distinctive turbulence structures in ABLs in a controlled numerical environment. In particular, our research questions are as follows:

1. How do coherent turbulence structures vary in convective ABLs as the ratio of buoyancy to shear production increases?
2. How can KMD help us to better understand and detect the spatiotemporal dynamics of coherent turbulence structures in ABLs such as convective rolls and cells under steady and unsteady mean forcing?
3. How can data-driven unsupervised clustering techniques along with KMD reduce the amount of data needed to accurately describe and characterize the overall turbulence structures in the ABL?

The outline of this paper is as follows. In Sec. II, we review the basics of the KMD approach and introduce the suite of conducted LESs and evaluation metrics. In Sec. III, we present the LES results of convective and neutral ABLs and characterize their turbulence features. Next, the KMD is applied to these ABLs, and the obtained dominant coherent structures in Koopman modes are discussed using quadrant analysis, spatial correlation, and timescale assessment. Then, we apply the K-means method to cluster the KMD modes by utilizing CNNs. Finally, a summary of the key findings of the paper is presented in Sec. IV.

II. METHODS

A. Overview of Koopman Mode Decomposition and Clustering

Koopman spectral theory was first introduced for Hamiltonian flows by providing an operator view to nonlinear dynamical systems.[57] Koopman's operator describes how the nonlinear dynamics of a system could be analyzed with an infinite-dimensional linear operator on the Hilbert space of observables. Consider a general dynamical system evolving on a smooth manifold $\mathcal{M}$,

$$\boldsymbol{x}_{k+1} = \boldsymbol{f}(\boldsymbol{x}_k), \tag{1}$$

where $\boldsymbol{x}_k \in \mathcal{M}$ is a state variable. Alternatively, the continuous form of this system $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}(t))$ can be studied; however, since we aim to analyze the discrete-time data, we adopt the discrete-time dynamical system in Eq. (1). The Koopman operator, $\mathcal{K}$, is an infinite-dimensional linear operator that acts on all scalar observable functions $h: \mathcal{M} \rightarrow \mathbb{C}$ such that

$$\mathcal{K}h(\boldsymbol{x}) = h(\boldsymbol{f}(\boldsymbol{x})). \tag{2}$$

The Koopman representation is a transformation of a nonlinear, finite-dimensional state-space representation into a linear, infinite-dimensional representation. This can be seen in Fig. 1 (top panels), in which we depict a schematic of the Koopman operator applied on an autonomous dynamical system. The solutions to the considered dynamical system can be found using the spectral decomposition of the Koopman operator. For this linear operator, we can consider the following eigenvalue problem:

$$\mathcal{K}\varphi_j = \lambda_j \varphi_j, \tag{3}$$

where λ_j denotes the jth eigenvalue and $\varphi_j(\boldsymbol{x})$ represents the jth eigenfunction of the Koopman operator. Consider a vector-valued observable $\boldsymbol{h}(\boldsymbol{x})$ including q components that could represent any quantities of interest such as velocity measurements. This vector may be expressed in terms of Koopman eigenfunctions as follows:

$$\boldsymbol{h}(\boldsymbol{x}) = \begin{bmatrix} h_1(\boldsymbol{x}) \\ h_2(\boldsymbol{x}) \\ \cdot \\ \cdot \\ h_q(\boldsymbol{x}) \end{bmatrix} = \sum\nolimits_{j=1}^{\infty} \varphi_j(\boldsymbol{x})\boldsymbol{v}_j, \tag{4}$$

where $\boldsymbol{v}_j$ denotes the Koopman mode associated with the eigenfunction φ_j and is given by the projection of the observable $\boldsymbol{h}$ onto the eigenfunction φ_j. Hence, observables $\boldsymbol{h}(\boldsymbol{x})$ can be represented as a linear combination of Koopman eigenfunctions $\varphi_j(\boldsymbol{x})$ with weight $\boldsymbol{v}_j$. Now, instead of evolving the nonlinear system in the original space (Fig. 1, top left), one can evolve it by trivially computing the time dynamics using (2)–(4),

$$\mathcal{K}\boldsymbol{h}(\boldsymbol{x}) = \mathcal{K}\sum_{j=1}^{\infty} \varphi_j(\boldsymbol{x})\boldsymbol{v}_j = \sum_{j=1}^{\infty} \mathcal{K}\varphi_j(\boldsymbol{x})\boldsymbol{v}_j = \sum_{j=1}^{\infty} \lambda_j \varphi_j(\boldsymbol{x})\boldsymbol{v}_j. \tag{5}$$

The iterates of $\boldsymbol{x}_0$ can now be calculated as follows:

$$\boldsymbol{y}_k = \boldsymbol{h}(\boldsymbol{x}_k) = \sum_{j=1}^{\infty} \mathcal{K}^k \varphi_j(\boldsymbol{x}_0)\boldsymbol{v}_j = \sum_{j=1}^{\infty} \lambda_j{}^k \varphi_j(\boldsymbol{x}_0)\boldsymbol{v}_j. \tag{6}$$

In this paper, when only a finite number of modes (p) is used for the reconstruction of the main flow field, we refer to it as R_p according to

$$R_p \equiv \sum\nolimits_{j=1}^{p} \lambda_j \varphi_j(\boldsymbol{x})\boldsymbol{v}_j. \tag{7}$$

An important feature of the Koopman operator is that it can capture the entire dynamics of a nonlinear system. However, in many practical cases, it is not possible to find all the modes and the associated eigenvalues of a dynamical system. To evaluate the Koopman modes directly from data, the DMD algorithm can be used to approximate the Koopman operator with a best-linear fit model.[41] It has been shown that the Koopman eigenvalues are the DMD eigenvalues if the set of observables is sufficiently large, and the data are adequately rich.[36]

An algorithm for using the data and DMD to approximate a Koopman operator comes as follows. To apply the DMD in the space of observables, first, the data matrices Y and Y' can be created as

$$\mathbf{Y} = \begin{bmatrix} | & | & & | \\ h(\boldsymbol{x}_1) & h(\boldsymbol{x}_2) & \cdots & h(\boldsymbol{x}_{m-1}) \\ | & | & & | \end{bmatrix}, \tag{8}$$

$$\mathbf{Y}' = \begin{bmatrix} | & | & & | \\ h(\boldsymbol{x}'_1) & h(\boldsymbol{x}'_2) & \cdots & h(\boldsymbol{x}'_{m-1}) \\ | & | & & | \end{bmatrix}, \tag{9}$$

where $\boldsymbol{x}_k$ is an initial condition to Eq. (1) and $\boldsymbol{x}'_k$ denotes the corresponding temporal evolution output after Δt. Next, the DMD algorithm can be performed to calculate $\mathbf{A}_\mathbf{Y} = \mathbf{Y}'\mathbf{Y}^\dagger$ along with the low-rank counterpart $\tilde{\mathrm{A}}_\mathrm{Y}$, where † denotes the Moore–Penrose pseudoinverse. The eigenvalues and eigenvectors of $\mathbf{A}_\mathbf{Y}$ can approximate Koopman eigenvalues and modes. Eventually, the approximation of Koopman modes is given by

$$\boldsymbol{\Phi}_\mathbf{Y} = \mathbf{Y}'\mathbf{V}\boldsymbol{\Sigma}^{-1}\mathbf{W}, \tag{10}$$

where $\mathbf{W}$ can be found from the eigenvalue problem $\tilde{\mathbf{A}}_\mathbf{Y}\mathbf{W}=\mathbf{W}\mathbf{L}$ and $\mathbf{Y}=\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^*$ (see Ref. 41 for more details). This procedure yields approximate Koopman modes, amplitudes, and future time steps in a linear evolution. The outcome of this process is shown as an example in Fig. 1 (middle panels) whereby applying the DMD to actual LES data, the spatiotemporal dynamics of the flow field is represented by the KMD. These decompositions will be further discussed in Sec. III.

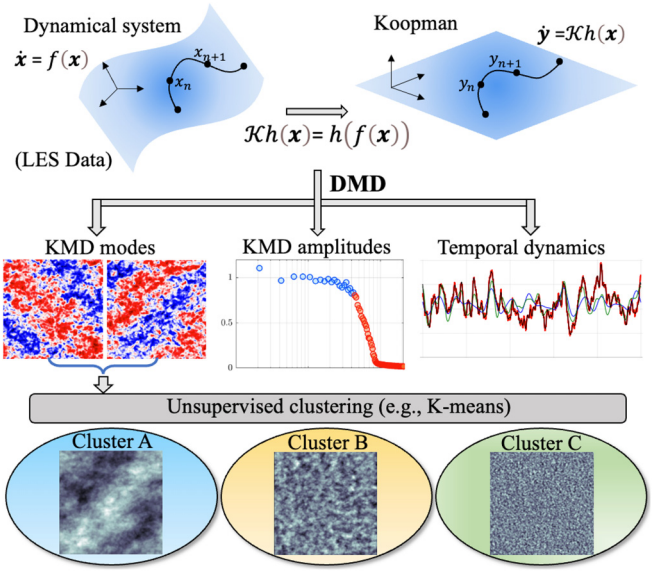

FIG. 1. Schematic of the overall methodology used in this study to decompose the LES data of ABL flows. The Koopman operator analysis through DMD is employed (top) to obtain Koopman modes, amplitudes, and temporal dynamics (middle). K-means clustering is then applied to cluster the resulting Koopman modes (bottom).

When the input data are large, many Koopman modes will be generated. However, the majority of the resulting Koopman modes do not necessarily yield a physical insight and are sometimes interpreted as a sign of chaos in data.[58] The goal of this paper is to come up with a reduced set of modes where they can reveal some non-trivial dynamics about the ABL flows. To this end, after the completion of the KMD, we use unsupervised clustering methods (e.g., K-means) to classify the resulting Koopman modes into reduced sets where we can relate them to the underlying physics. The schematic of applying the unsupervised clustering method to Koopman modes is shown in the bottom row of Fig. 1 and will be extensively discussed in Sec. III. This clustering method enables us to focus on a reduced set of uncorrelated Koopman modes rather than dealing with many outputted modes from the algorithm described above, which can be similar to each other.

For the data analysis, we developed all Koopman mode decomposition, reconstruction, and clustering framework codes for large data (Fig. 1) in Python. A range of Python libraries were leveraged to facilitate our computations and data processing. For applying machine learning algorithms, "tensorflow" and "sklearn.cluster" libraries were used. Furthermore, "shutil" was utilized for file operations. All these post-processing tasks were performed on supercomputers using Intel Xeon G6252 CPU. The total computational cost for the KMD operations was ∼5000 CPU-core hours. The computational cost associated with the K-means clustering and CNN feature extraction was notably lower (∼100 CPU-hours) compared to the KMD computations since we employed pre-trained validated CNN models in this work.

B. Large eddy simulations

In this study, the LES technique is employed to simulate different ABL flow regimes and generate data for the KMD method. LES is the state-of-the-art method for resolving large turbulent eddies in high-Reynolds number ABL flows.[59] In LES, large energetic scales of motion are resolved and the eddies smaller than the grid-filter size are parameterized using a sub-grid scale (SGS) model. The adopted LES code in this paper solves the incompressible continuity equation,

$$\frac{\partial \tilde{u}_i}{\partial x_i} = 0. \tag{11}$$

Navier–Stokes momentum conservation,

$$\frac{\partial \tilde{u}_i}{\partial t} + \tilde{u}_j\left(\frac{\partial \tilde{u}_i}{\partial x_j} - \frac{\partial \tilde{u}_j}{\partial x_i}\right) = -\frac{1}{\rho}\frac{\partial \tilde{p}^*}{\partial x_i} - \frac{\partial \tau_{ij}}{\partial x_j} + g\frac{\tilde{\theta}'}{\theta_r}\delta_{i3} + f\left(U_g - \tilde{u}_1\right)\delta_{i2} - f\left(V_g - \tilde{u}_2\right)\delta_{i1}, \tag{12}$$

and the thermal energy conservation for potential temperature ($\tilde{\theta}$),

$$\frac{\partial \tilde{\theta}}{\partial t} + \tilde{u}_j\frac{\partial \tilde{\theta}}{\partial x_j} = -\frac{\partial \pi_j}{\partial x_j}, \tag{13}$$

where x_i represents the position vector; $\tilde{u}_i = (\tilde{u}, \tilde{v}, \tilde{w})$ is the resolved velocity vector; $\tilde{p}^*$ is a modified pressure defined as $\tilde{p}^* \equiv \tilde{p} + \frac{1}{2}\rho\tilde{u}_j{}^2 + \frac{1}{3}\rho\sigma_{kk}$, where the third term denotes the SGS kinetic energy; τ_{ij} is the deviatoric part of the SGS stress tensor; θ_r is the reference potential temperature; f is the Coriolis parameter ($=9.7 \times 10^{-5}\,\mathrm{s}^{-1}$ here); ρ is a constant fluid density; g is the gravitational constant; δ_{ij} is the Kronecker delta; and π_j is the SGS heat-flux vector. To calculate the SGS eddy-viscosity for τ_{ij}, a dynamic Lagrangian scale-dependent (LASD) Smagorinsky model is used.[60] For calculating the SGS heat-flux, an eddy-diffusivity (ν_T) based model is employed as $\pi_j = (\nu_T/Pr_{SGS})(\partial\theta/\partial x_j)$, where a constant Prandtl number $Pr_{SGS} = 0.4$ is used in the simulations.[61] The employed LASD SGS model has generated reliable velocity spectra and Smagorinsky coefficients that were in good agreement with experiments.[62] To evaluate the spatial derivatives, a Fourier-based pseudospectral approach in the horizontal directions and a second-order centered-difference scheme in the vertical direction are employed. For numerical time integration, the fully explicit second-order Adams–Bashforth method is used. In this study, the mean variable pressure forcing is expressed as a geostrophic wind as follows:

$$(U_g,\ V_g) \equiv \left(-\frac{1}{f\rho}\frac{\partial p}{\partial y}, \frac{1}{f\rho}\frac{\partial p}{\partial x}\right), \tag{14}$$

where U_g and V_g denote the geostrophic wind velocity components in the x and y direction, respectively. For the top boundary, a stress-free and an impermeable boundary condition ($w = 0$) is used, and lateral boundary conditions are periodic. In all the simulations, a sponge layer with a Rayleigh damping method is used at the top of the domain to damp all flow perturbations. For the bottom boundary, a local logarithmic-wall model, with a Monin–Obukhov stability correction for non-neutral conditions, based on the filtered velocities is employed to impose the surface fluxes. For the surface, the aerodynamic roughness length of $z_0 = 0.1$ m is used. More details about the LES model, and the diabatic simulation setups can be found in Refs. 25, 28, 29, 63, and 64. This code has been extensively validated against many observational data under different atmospheric regimes.[60,65–68]

C. Suite of LESs

To evaluate the KMD method for detecting coherent turbulence structures in high-Reynolds number ABL flows, we conducted a suite of different LESs. To this end, ABLs with different stabilities are simulated to generate two distinctive coherent structures. When buoyancy production is dominant in the ABL, thermal plumes—aka cells—become the dominant turbulence structures, while hairpin vortex packets—aka rolls—become the major structures when shear production dominates. Hence, we control the ratio between shear and buoyancy turbulence productions by varying the magnitude of the imposed geostrophic wind in convective ABLs. The summary of all the conducted LESs is presented in Table I. In the first two cases, the geostrophic wind magnitude is set to 4 and 16 m s^{-1} under the same imposed surface heat flux of 0.16 K m s^{-1} (200 w m^{-2}) to vary the ratio of shear to buoyant production. These unstable ABLs are named U4, and U16 cases, respectively, according to the magnitude of their imposed geostrophic wind (Table I). We also tested a completely neutral ABL case denoted as N16 in which the geostrophic wind is set to 16 m s^{-1}, but the surface heat flux is set to zero. Furthermore, another case is conducted under unsteady forcing conditions to assess the KMD performance when the mean flow is transient. To this end, the geostrophic forcing is suddenly changed from 16 to 8 m s^{-1} under neutral conditions, and the case is named N16_uns.

To evaluate the sensitivity of the KMD method to grid resolution, we conducted all these LESs using two different grid resolutions $96 \times 96 \times 96$ and $192 \times 192 \times 192$. Thus, in total, eight different LESs

TABLE I. Details of the performed LESs of convective and neutral ABLs. All the cases were conducted using two resolutions of $192 \times 192 \times 192$ and $96 \times 96 \times 96$ leading to eight total LESs.

Case name	U_g (m s^{-1})	Surface heat flux (K m s^{-1})
U4	4	0.16
U16	16	0.16
N16	16	0
N16_uns	$16 \rightarrow 8$	0

were performed in this study. The domain height in all cases is $L_z = 1250$ m, and the horizontal domain sizes are $L_x = L_y = 2\pi L_z$. For these neutral and unstable ABLs, previous grid convergence tests of our LES code indicated that Δz of 15.6 and 7.8 m resulted in very similar mean profile results.[69] Furthermore, the grid size of $\Delta z \approx 12.5$ m for our LES code was shown to be adequate for convergence of higher-order moments such as variance and skewness in convective ABLs.[33] Hence, our chosen grid resolution of $\Delta z \approx 6.5$ and 13 m should be sufficient for our considered cases. The initial temperature inversion layer height is $z_i = 1000$ m in all cases. All the results of this study are non-dimensionalized using the magnitude of the imposed geostrophic wind $G = \sqrt{U_g^2 + V_g^2}$ and z_i. The inversion layer height will remain constant for neutral cases, but it may vary for unstable cases as the ABL evolves. Nevertheless, we use the initial imposed inversion layer depth z_i as a fixed value for non-dimensionalizing all our results to facilitate their comparison in a dimensional framework especially for experimentalists, consistent with previous ABL studies.[30,69,70]

All our cases are initialized by running the 96^3 resolution cases for a spin-up period of three inertial timescales ($3 \times \tau_{ABL} = 3 \times 2\pi/f \approx 54$ h). Following this phase, the simulation data were interpolated to a finer 192^3 resolution and ran for an additional 2 h as warm-up to remove interpolation effects. Subsequently, the last phase of the simulations is conducted for a final τ_{ABL} period ($\approx$18 h), which are utilized and shown in the paper. In the unsteady scenario, the N16 simulation is used as the initial condition of N16_uns case. The N16 simulation is continued, but the geostrophic wind is suddenly changed to 8 m s^{-1} for a duration of $2 \times \tau_{ABL}$ ($\approx$36 h). This leads to a sudden mean pressure gradient change and generates unsteady flow conditions. We note that this spin-up period is sufficiently large to remove major inertial oscillations in the flow. Our spin-up time for the neutral case (54 h) in terms of the large-eddy turnover timescale ($\tau_* \equiv z_i/u_*$) becomes $\approx 134\tau^*$, which remarkably exceeds the $20\tau^*$ recommended spin-up time,[71] and 60–80 τ_* commonly performed warm-up periods in previous ABL studies.[4,72]

The 192^3 resolution simulations were conducted using 48 CPU cores on an Intel Xeon G6252 CPU. Running one physical hour of each 192^3 case required $\sim$1000 CPU-core hours. Given long simulation runs, the total computational cost of the four cases in the final production phase (excluding warm-up) was $\sim$96 000 CPU-core hours. In total, we used $\sim$10 Terabytes of disk space to store the 3D LES data of all the cases.

D. Evaluation metrics

The performance of the Koopman mode reconstruction is evaluated by calculating two metrics: mean absolute error (*MAE*) and the Pearson correlation coefficient (*Corr*). *MAE* is calculated as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N} \left| R_p(i) - M_{LES}(i) \right|, \tag{15}$$

where N is the number of samples (here the grids in one plane, e.g., $N = 192 \times 192$ or 96×96, depending on the grid resolution), R_p denotes the reconstructed flow field with the first p Koopman modes, and M_{LES} is the actual LES data. *MAE* indicates the sum of absolute prediction errors, while the Pearson coefficient (between -1 and 1) measures the linear correlation between the actual and predicted values as follows:

$$Corr = \frac{\sum_{i=1}^{N} \left(R_p(i) - \overline{R_p}\right)\left(M_{LES}(i) - \overline{M_{LES}}\right)}{\sqrt{\sum_{i=1}^{N} \left(R_p(i) - \overline{R_p}\right)^2} \sqrt{\sum_{i=1}^{N} \left(M_{LES}(i) - \overline{M_{LES}}\right)^2}}, \tag{16}$$

where the overbar denotes the mean of the data. These metrics will be used to evaluate the accuracy of the reconstructed wind velocity against the LES flow field.

To evaluate the performance of the K-means method in clustering the KMD modes, the "*Inertia*" metric is employed. In this paper, we used the within-cluster sum of squared error (WCSS) method to calculate this metric.[73,74] Inertia indicates how closely the data points within a cluster are to one another, and how well separated the clusters are from each other. This metric measures the degree of tightness of the clusters and is defined as follows:

$$Inertia\,(WCSS) = \sum_{k=1}^{K} \sum_{x_i \in C_k} |x_i - \mu_k|^2, \tag{17}$$

where K is the number of clusters, C_k is the set of modes belonging to the kth cluster, x_i is the ith Koopman mode, and μ_k is the centroid of the kth cluster.

III. RESULTS

A. The LES results of the convective and neutral ABLs

First, we describe the distinctions between the mean and turbulence characteristics in the considered ABL cases. The normalized mean wind speed profiles of the conducted LESs are shown in Fig. 2(a). The vertical profiles in this figure are averaged in horizontal directions as well as in time over the last inertial timescale $\tau_{ABL} = 2\pi/f \approx 18$ h of the simulations. By increasing the buoyancy to shear production ratio, the wind profiles show a stronger mean wind shear near the surface. This is consistent with the Monin–Obukhov similarity theory (MOST) prediction that unstable ABL wind profiles are concave upward in a semi-log plot compared to neutral ABLs.[2] Furthermore, turbulent mixing in convective ABLs increases, and thus, the profiles become more uniform compared to neutral ABLs.[30] This increased turbulent mixing and eddy viscosity in convective ABLs causes a strong coupling of the ABL with the surface. Hence, due to this strong vertical coupling with the surface and higher eddy diffusivity, the mean wind profiles can decrease above the surface ($z/z_i \gtrsim 0.3$) in unstable ABLs compared to the neutral case [compare blue and black line in Fig. 2(a)]. In the neutral case N16, the atmosphere becomes stably stratified in the top imposed capping inversion layer. Hence, some inertial oscillations may occur above the ABL ($z/z_i \gtrsim 1$) and can lead to supergeostrophic jets near the top of our domain. However, such inertial oscillations are very weak compared to the mean flow within the ABL ($z/z_i \lesssim 1$) for the N16 case since our spin-up period was long enough ($134\tau^*$) to remove such mean oscillations. These obtained

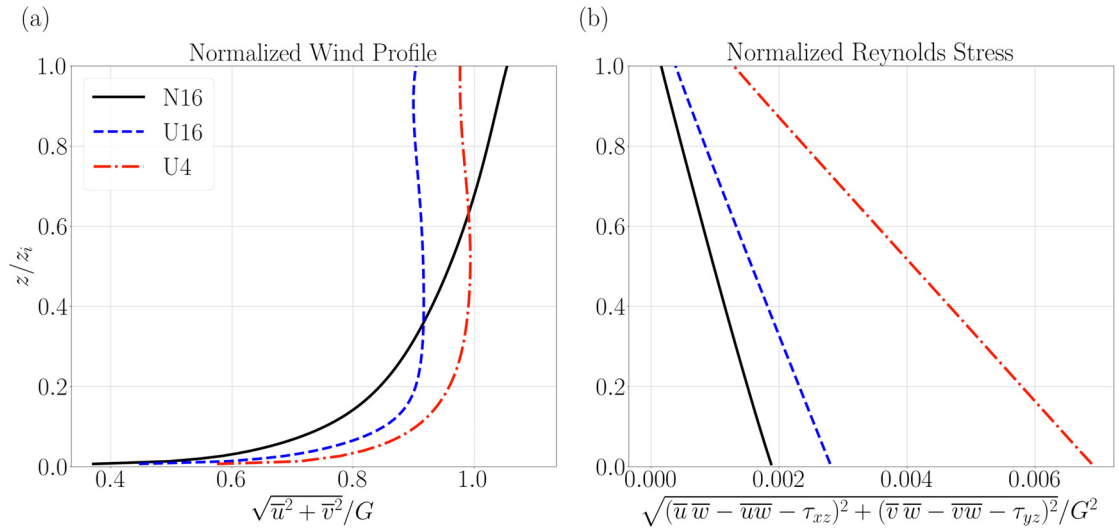

FIG. 2. (a) Non-dimensionalized mean horizontal wind speed and (b) Reynolds stress profiles for the considered LES cases. The height is normalized with z_i, which is 0.8 L_z in our cases, and the wind velocities are normalized with G.

wind profile results are also consistent and similar to previous LES studies of convective and neutral ABLs.[33] The profiles of wind velocity components in the x and y directions can be seen in supplementary material, Fig. S1.

The turbulence dynamics of the considered ABL cases are also remarkably different. Figure 2(b) displays the averaged Reynolds stress profiles in the conducted LES cases. By increasing the buoyancy to shear production ratio, the non-dimensionalized Reynolds stress profiles increase. A similar trend can be seen in the shear production (SP) term of the turbulent kinetic energy (TKE) budget (see Momen and Bou-Zeid[25] for more details about the TKE calculations). It is found that while the absolute magnitude of the SP increases with increasing shear, the normalized SP by the imposed geostrophic wind increases as the ABL becomes more unstable (see supplementary material, Fig. S2).

B. Turbulence characteristics and quadrant analysis of the convective ABL cases

Turbulence production mechanisms have considerable differences between a neutral and a convective ABL. In a neutral ABL, the turbulence is primarily produced by the shear production, while in a strongly convective ABL, it is mainly produced by buoyancy. In our considered cases, the ratio of shear to buoyancy is varied and this influences the turbulence properties in the ABL such as the wall shear stress. This can be seen from the friction velocity, u_*, that has been calculated and shown for the steady cases in Table II. The table indicates that by increasing the imposed geostrophic wind speed (from U4 to U16), u_* also increases. Under the same geostrophic wind, convective ABL case U16 has a higher u_* than the corresponding neutral ABL case N16. This is because both shear and buoyancy produce TKE in the U16 case, while in N16 only shear produces TKE. Furthermore, the higher vertical mixing in convective ABL leads to a higher surface wind shear [compare blue and black lines in Fig. 2(a)] that consequently leads to a larger u_* in the U16 case than N16. The table also shows the convective velocity scale $w_* \equiv \left[g z_i \overline{w'\theta'}_s/\overline{\theta}\right]^{1/3}$, calculated for all cases. Since the imposed surface heat flux is the same for both convective cases, w_* is the same for both cases, while for the neutral case, it is zero since there is no imposed surface heat flux in the ABL.

The ratio of shear to buoyancy forces is typically characterized by the Obukhov length (L), which is a measure of the atmospheric stability in the ABL community. The Obukhov length is defined as follows:

$$L \equiv -{u_*^3\overline{\theta}}\Big/{\kappa g \overline{w'\theta'}}, \tag{18}$$

where κ is the von Karman constant (here $\kappa = 0.4$) and $\overline{w'\theta'}$ denotes the heat flux. The values of L are calculated for our considered cases and shown in Table II. As the table indicates, increasing the shear-to-buoyancy ratio increases $-L$ from 17.4 m (U4 case) to ∞ (N16 case). The Richardson flux number (Ri_f) is a common dimensionless number to determine the stability of the ABL. It represents the negative ratio of the buoyancy to shear production of TKE and is defined as

$$Ri_f \equiv \frac{(g/\overline{\theta})\overline{w'\theta'}}{\overline{u'w'}\left(\frac{\partial \overline{u}}{\partial z}\right) + \overline{v'w'}\left(\frac{\partial \overline{v}}{\partial z}\right)}. \tag{19}$$

TABLE II. Friction velocity, convective velocity, Obukhov length, flux Richardson number, and velocity fluctuation correlation for the considered steady LES cases.

Case name	u_* (m s^{-1})	w_* (m s^{-1})	L (m)	Ri_f	u'–w' correlation
U4	0.332	1.81	−17.4	−0.481	−0.11
U16	0.847	1.81	−288.4	−0.016	−0.42
N16	0.693	0	$-\infty$	0.000	−0.48

A similar trend for this non-dimensional number can be seen in Table II in which $Ri_f = 0$ in the neutral ABL, and as we increase the ratio of buoyancy to shear production, the absolute magnitude of Ri_f increases to ≈0.48. In this table, Ri_f is calculated in horizontal directions and then averaged over the entire ABL depth. Note our simulated neutral case is the conventionally neutral ABL[75] since we also solve for the potential temperature equation and impose a capping inversion layer at the top of our domain. Hence, the ABL can be stable in this capping inversion layer unlike other studies in which such layer is not imposed, and the impacts of θ are fully eliminated from LESs.[28] Our simulated conventionally neutral condition is more prevalent in real-world ABLs.

To further characterize the turbulence structures in the considered cases, we conducted a quadrant analysis for the ABL flows. Quadrant analysis is a simple yet powerful method[76] that yields insights into the Reynolds stress production mechanisms and contributions of organized eddy motions to the TKE production and transport. This technique categorizes the products of velocity fluctuations u' and w' into four quadrants as follows:[77]

- Quadrant 1 (Q1) where $u' > 0$ and $w' > 0$, indicating outward interaction,
- Quadrant 2 (Q2) where $u' < 0$ and $w' > 0$, indicating ejection,
- Quadrant 3 (Q3) where $u' < 0$ and $w' < 0$, indicating inward interaction, and
- Quadrant 4 (Q4) where $u' > 0$ and $w' < 0$, indicating sweep.

These are typically called the quadrants of the Reynolds shear stress plane. The scatterplot of u' and w' has been shown near the surface at height $z = 85$ m ($z/z_i = 0.085$) in Fig. 3. The figure shows that for the neutral ABL, sweeps (Q4) and ejections (Q2) have the largest contributions to the Reynolds shear stress. This is consistent with typical channel flow studies.[78,79] However, as the magnitude of the Ri_f increases, the contributions of sweeps and ejections structures decrease [compare Figs. 3(a) and 3(c)].

To quantify the contributions of each quadrant, we calculated the frequency of each of these events and showed them in Fig. 4. The percentage of the points that lie in Q2 + Q4 for the neutral case N16 is 65.2%, while it is 55.5% for the convective case U4 at $z = 85$ m. These contributions can be further understood by analyzing the correlation of u' and w' in Fig. 3. These correlations are calculated and shown in Table II. As the table indicates, by increasing the buoyancy to shear production ratio, the magnitude of the correlation between u' and w' decreases. In the strongly unstable ABL U4 case, the correlation is very low. The reason for this is mainly because, in near-neutral ABLs, hairpin vortices exist that contribute to gradient-type motions, while in convective ABLs, thermal plumes increase ABL mixing and counter-gradient motions, consistent with previous studies.[80] The enhanced vertical mixing and counter-gradient motions (Q1 and Q3) of such thermal plumes contribute to the observed scatter of the velocity fluctuations in Fig. 3(c). The increased turbulent motion variability and decreased correlation of u' and w' in U4 case can also be linked to the decreased efficiency of momentum transport as the convection increases, which agrees with prior studies.[33] These findings also agree with prior research,[81] which presented the vertical profiles of mean flux contributions from the four quadrants.

C. Applying the KMD method to characterize turbulent eddies in convective ABLs

In this study, the KMD method is applied to the horizontal snapshots (x–y plane) of the normalized wind speed and vertical velocity separately, at different heights. The input data to the KMD method (Fig. 1) are a two-dimensional matrix in which the first dimension is related to the observables (here the velocity field) representing the spatial grid points (e.g., 192×192), and the second dimension denotes the number of snapshots representing the temporal time steps. To capture the whole mean dynamics of the ABL cases, a full inertial cycle of the LES data ($\tau_{ABL} \approx 18$ h) is extracted and ingested into the KMD method. To save computational and storage resources, the 3D LES data are averaged for 5 s and then outputted during the simulation period. This leads to 12 960 3D snapshots (192^3) for one inertial cycle of each run, which generated ∼10 Terabytes of data in total for all the considered cases. Another reason for choosing $dt = 5$ s is that according to the Nyquist sampling theorem, the maximum frequency that KMD can resolve is $\frac{1}{2dt}$, which in our case would be 0.1 Hz. This frequency is significantly higher than turbulent eddy turnover frequency (u_*/z_i) in our LES cases. We also conducted a sensitivity to outputting frequency and found that the choice of $dt = 5$ or $dt = 10$ s does not remarkably affect the decompositions, especially the predominant modes (see supplementary material, Fig. S3). On the other hand, the

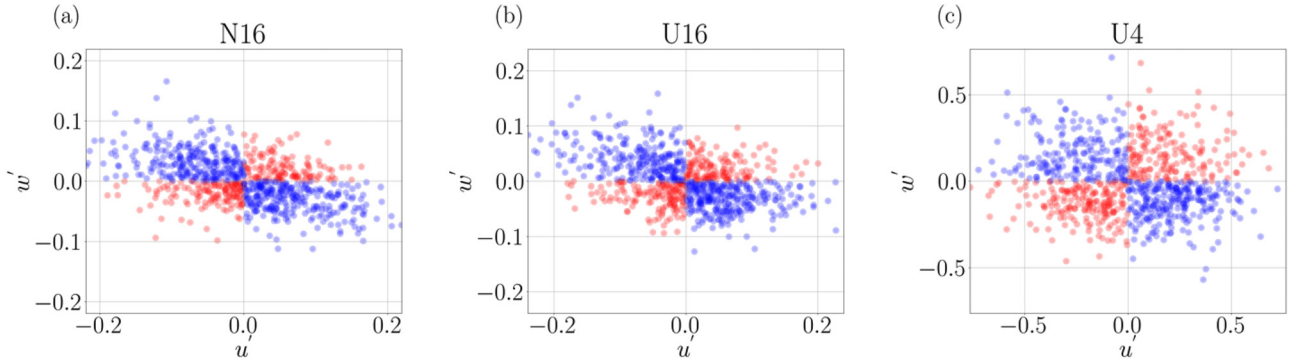

FIG. 3. The scatterplot of u' and w' for (a) N16, (b) U16, and (c) U4 case at $z = 85$ m ($z/z_i = 0.085$). The red points lie in the first and third quadrants (red color represents counter-gradient-type motions), and the blue points lie in the second and fourth quadrants (blue color represents gradient-type motions). The points are colored with transparency so more points can be seen.

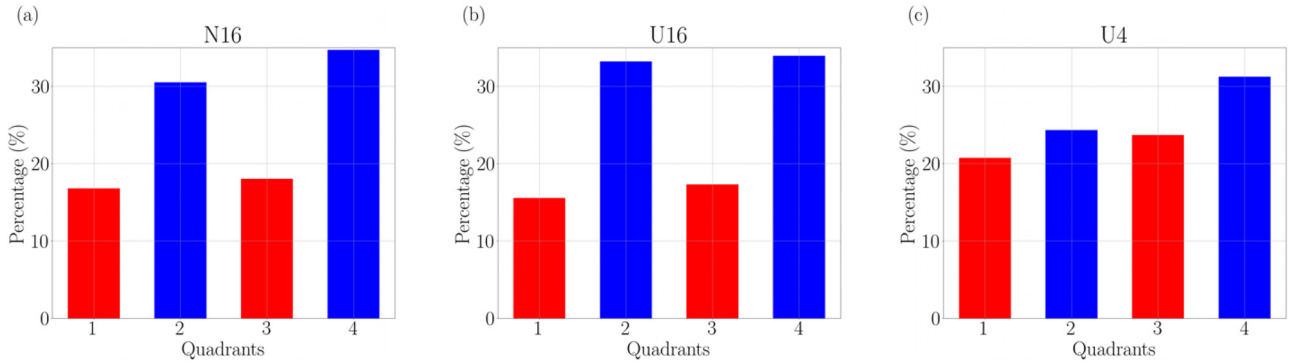

FIG. 4. Bar plot of the contribution of each quadrant to Reynolds shear stress for (a) N16, (b) U16, and (c) U4 case at $z = 85$ m ($z/z_i = 0.085$).

smallest frequency of Koopman modes is proportional to the length of the input dataset, which in our case would be $1/\tau_{ABL}$. In this section, we will present the results for the mode amplitudes, frequencies, and temporal dynamics in capturing the LES results. In Sec. III D, a comprehensive analysis of the decomposed modes and the spatial coherent structures will be presented.

The mode amplitude corresponds to the magnitude or energy content of the spatial mode structures. In Figs. 5(a)–5(c), Koopman mode amplitude is shown vs the mode frequency for the first 1000 decomposed modes of three cases N16, U16, and U4 (the rest of the modes have very small amplitudes and represent very low-energy eddies). In this figure, the mode frequencies are shown in dimensionless form using the Strouhal number, $St \equiv Mode\ frequency \times z_i/G$. As expected, lower frequencies have larger amplitudes implying that the large eddies in the flow contain more energy compared to the smaller scales. Moreover, the smaller amplitude modes have a shorter timescale (higher frequency) and sustain shorter in the flow field. This result is consistent with Kolmogorov's spectrum of velocity fluctuations representing the source, inertial, and viscous dissipation ranges. This similarity is shown in [Figs. 5(d)–5(f)] by computing the power spectrum vs the wave number for the LES data of the same cases. Furthermore, Figs. 5(a)–5(c) indicate that the Koopman modes in convective case U4 have higher amplitudes compared to the near-neutral U16 and neutral N16 cases. This suggests that the coherent spatial structures decomposed in the convective ABL are more energetic than in the near-neutral case. These more energetic structures in U4 are also associated with lower frequencies, which indicate that the energy-containing eddies in U4 are larger than U16 and N16 cases. A similar trend can be seen when comparing the energy peak of the largest eddies (lowest wave numbers) in the power spectrum [Figs. 5(d)–5(f)]. These findings agree with other studies that showed that coherent turbulence structures in convective ABLs are larger than neutral ABLs.[69] We will further characterize these spatial mode structures in Sec. III D.

D. The strength of KMD in describing ABL dynamics and interpretation of main Koopman modes via timescale analysis

In this section, we will use the timescale analysis to identify and interpret the characteristics of important Koopman modes. Different forces, such as pressure gradient, Coriolis, friction, and buoyancy, interact nonlinearly with each other to determine the dynamics of ABLs. Our goal is to associate some of the Koopman modes with these forces to better understand how they impact the ABL dynamics. To this end, we first show the Koopman modes of the neutral cases, which do not include surface buoyancy forces. Figure 6 depicts the Koopman modes of the N16 case in which they are numbered according to their amplitude magnitude.

We first considered the inertial timescale of the ABL, which is associated with the Coriolis frequency. The closest mode to this frequency is found to be mode 10 [blue dot in Fig. 6(a)]. Figure 6(b) displays the timeseries of the demeaned and normalized fluctuations of horizontal wind speeds in this case. The figure shows the actual LES results (black line) vs various selected Koopman modes. In this figure, mode 10 is shown as a blue line where it captures the inertial oscillations of the ABL flow with a timescale equal to τ_{ABL} ($\approx$18 h here). The figure shows that this mode captures mean flow fluctuations resulting from Coriolis effects. Figures 6(c)–6(e) depict the spatial structure of three selected modes. To better compare the structure of each mode in a non-dimensional space, they are normalized with their maximum values in each case, e.g., max (mode 1, mode 3, mode 10) = (1.57, 1.49, 1.06). The spatial structure in mode 10 [Fig. 6(e)] is slightly tilted horizontal structures, which are aligned with the mean flow velocity direction (compare the black arrow with the spatial structure of this mode). This tilt in the mean flow and subsequently in this mode is due to the Coriolis force effect that induces Ekman-like spirals in the ABL.[2]

The next important timescale that we explored is related to the turbulent eddy timescale, which can be found as z/u_*. The closest decomposed Koopman mode to this timescale is interestingly mode 1, which has the highest amplitude. This has been shown as a yellow dot in Fig. 6(a) and as a yellow line in Fig. 6(b), which has a higher frequency and larger magnitude oscillations compared to mode 10 (compare blue and yellow lines in Fig. 6b). The other significant timescale is related to the largest turbulent eddy in the ABL, which can be found as z_i/u_* ($\approx$1443 s for this case). The closest Koopman mode to this timescale is found to be mode 3, which is shown as a green dot in Fig. 6(a) and a green line in Fig. 6(b). The spatial structure of these two modes is also shown in Figs. 6(c) and 6(d). Both structures indicate x-direction-oriented modes, which are likely induced by the imposed strong pressure gradient via the geostrophic wind ($U_g = 16$ m s^{-1}). The size of eddy structures in mode 1 seems to be smaller than mode 3, which

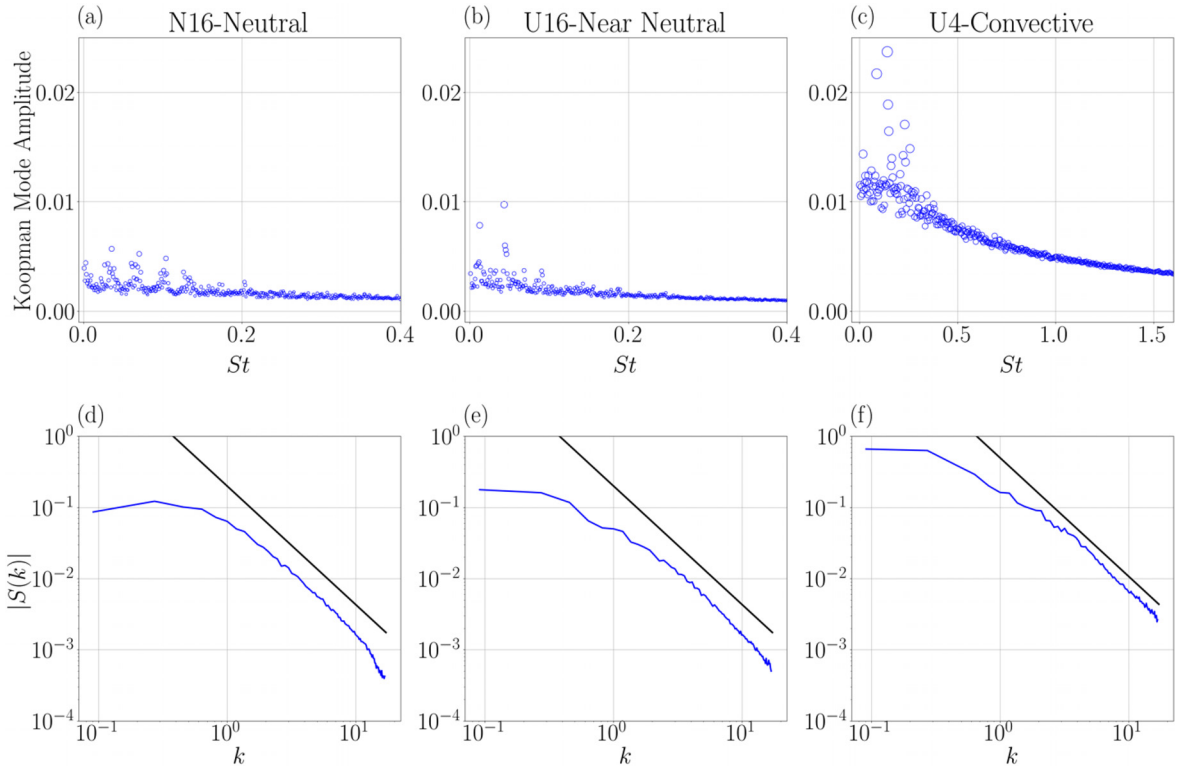

FIG. 5. (Top) The amplitude vs the frequency for the first 1000 Koopman modes for (a) N16, (b) U16, and (c) U4 case at height $z = 345$ m. Each circle corresponds to one mode, and the size of the circle corresponds to the mode amplitude. (Bottom) Power spectrum vs the wave number of the LES data at the same elevation for (d) N16, € U16, and (f) U4 case for one time step in the middle of the simulations. The black lines in (d)–(f) indicate the −5/3rd Kolmogorov's power law. The *St* here represents a non-dimensional frequency, and *k* denotes a dimensionless wave number (wave number × z_i).

could be associated with their difference in timescale (mode 1 has a smaller timescale than mode 3).

Note to corroborate our discussion about these Koopman modes, we also performed some sensitivity tests in which we altered the domain size. We noticed that the frequency of the major discussed modes did not change when altering the horizontal domain size by doubling the L_x and L_y (not shown). In another sensitivity test, we doubled the vertical domain size and z_i. This time, the frequency of z_i/u_* mode was changed indicating that the timescale of this mode directly depends on z_i. These results demonstrate that these modes represent the underlying dynamics of ABLs and are not numerical artifacts.

A strength of the KMD is that once we determine the Koopman modes, we can reconstruct the nonlinear flow field by a linear combination of the obtained modes. To capture the prevailing flow dynamics, we do not need to use all the obtained Koopman modes as higher frequency modes represent smaller lower energetic eddies that do not influence the overall flow field that much. To this end following Eq. (7), we define R_τ as the reconstruction of the flow field using all the modes with frequencies smaller than $10\,u_*/z_i$. This definition accounts for all major turbulent eddy time scales as well as the nonlinear interactions of the mean (inertial timescale here) and turbulence time scales. In this case, R_τ includes only ∼7% of all the decomposed Koopman modes. The red dashed line in Fig. 6(b) indicates that the reconstruction of the ABL flow using these R_τ modes agrees well with the LES results and captures its nonlinear fluctuations.

To further evaluate the accuracy of the Koopman mode reconstruction, we also calculated the average error and spatial correlation between the actual LES snapshots and reconstructed flow fields. The time series of the correlation between the reconstructed flow and LES data for three different sets of employed modes is presented in Fig. 7(a) for the U4 case. Reconstructing the ABL flow using R_τ Koopman modes for the U4 case also shows more than 80% correlation. For consistency, we further examined if by increasing the number of Koopman modes, we could obtain the exact LES flow field. To this end, we only reconstructed one snapshot at $t = \tau_{ABL}/2$ by increasing the number of modes from 1 to 12 957 (all modes). As Fig. 7(b) indicates, by increasing the number of modes, the correlation between the reconstructed snapshot and the actual LES snapshot also increases rapidly and approaches 1 when ∼40% of the modes are used. Furthermore, the average error of the reconstruction rapidly decreases by increasing the number of used modes and reaches ≈0 when all the modes are used for reconstruction [Fig. 7(c)].

E. The performance of the KMD in unsteady ABL flows

In this section, we will characterize the performance of the KMD in detecting modes of an unsteady ABL under transient pressure

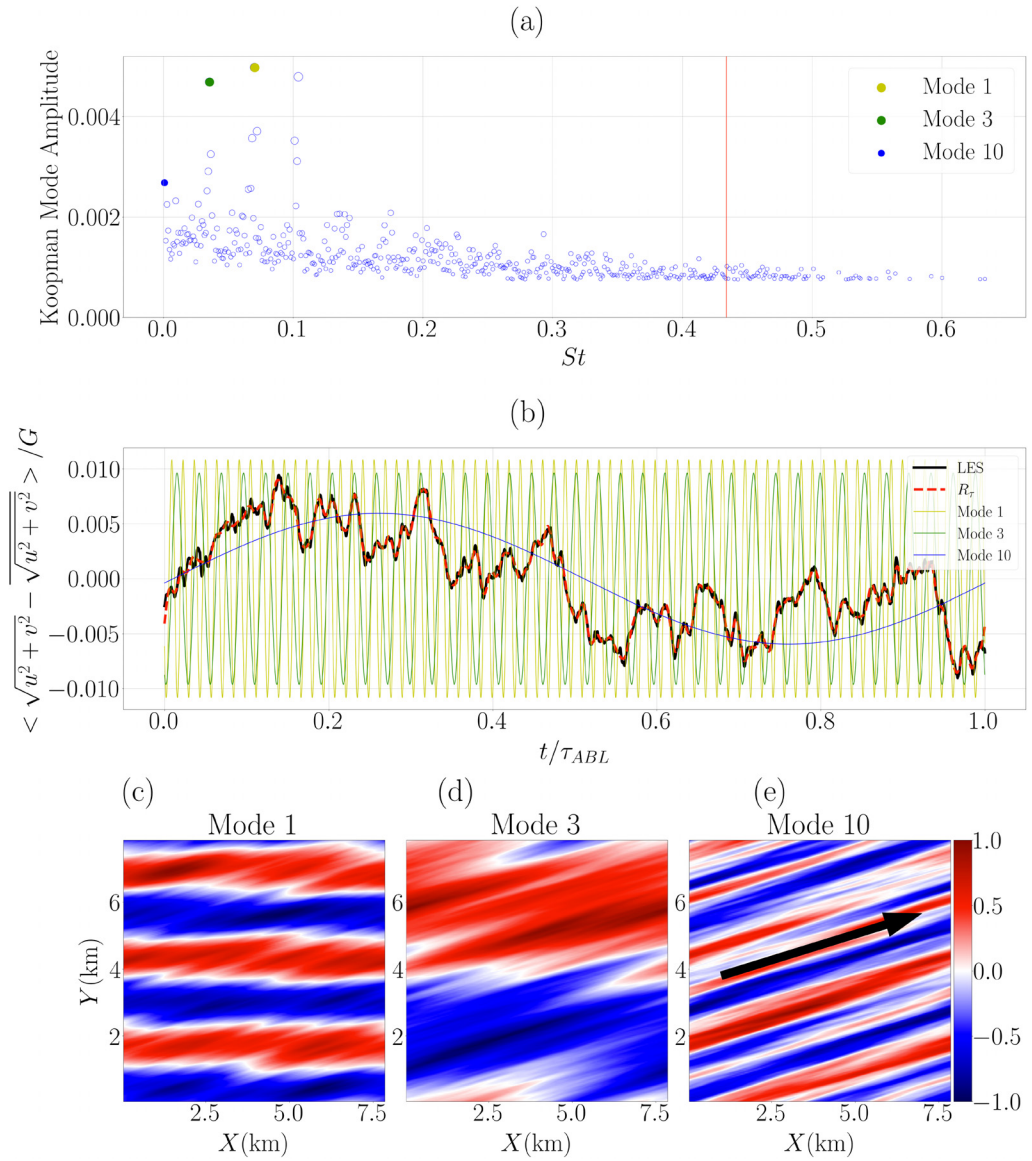

FIG. 6. (a) Koopman mode amplitude of case N16 vs frequency at elevation $z = 710$ m. The vertical red line shows the frequency of $10\times$ z_i/u_* (the maximum frequency used to reconstruct the flow field in R_τ). (b) Time series of the fluctuations in wind speed for LES, three important Koopman modes, and the reconstructed Koopman flow (R_τ). (c)–(e) The spatial structure of Koopman mode 1, mode 3, and mode 10 is shown. The black line in (e) depicts the mean flow velocity direction from LES data.

gradient forcing. To this end, we applied KMD to the N16_uns case where the pressure gradient is suddenly decreased to half. Figure 8(a) shows the Koopman mode amplitude vs frequency for this case. Upon comparing the magnitude of the Koopman modes of this case [Fig. 8(a)] vs the steady state case [Fig. 6(a)], we find that the Koopman modes of the equivalent unsteady case have a larger amplitude, especially at lower frequencies. This could be due to the fact that the mean flow is varying in this case and Koopman modes represent the mean variations, which are more energetic here than turbulence fluctuations. The most energetic mode of this case (mode 1) has a frequency close to the Coriolis frequency indicating that the prevailing dynamics of this case is governed by inertial oscillations. This mode is shown as a blue circle in Fig. 8(a) indicating a major peak. This is unlike the steady state case where the Coriolis-related mode was mode number 10 showing that the turbulence eddies have more energy than inertial oscillations in such a steady mean ABL flow.

For the turbulence timescale, unlike the steady-state cases, u_* and relevant time scales are also varying in this case to reach an equilibrium with the new imposed pressure gradient. Hence, a range of time scales can represent turbulent eddies as they evolve in time. Nevertheless, we considered u_* of the average of an inertial cycle, and also found the closest mode to turbulence timescale $z_i/u_* \approx 2564$ s. This mode is

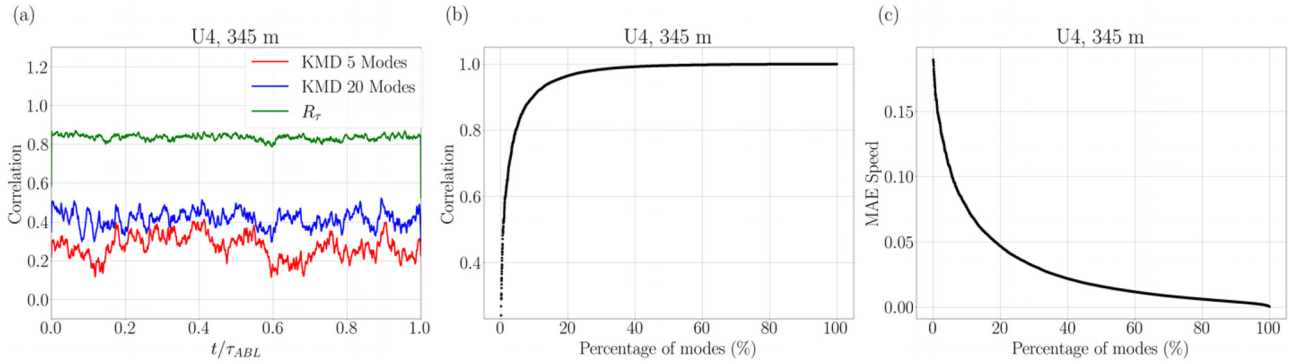

FIG. 7. Correlation and mean average error of wind speed between the reconstructed flow and LES data are calculated using different numbers of modes for case U4 at $z = 345$ m. (a) The correlation is calculated using 5, 20, and R_τ modes over one inertial timescale, (b) the correlation is calculated vs the number of modes used for reconstruction, and, (c) the MAE of wind speed is shown vs the number of modes used for reconstruction.

shown as a cyan circle in Fig. 8(a), which has a local peak in the Koopman magnitude.

The horizontally averaged timeseries of the wind fluctuations for this case is shown in Fig. 8(b). The temporal behavior of the two identified modes is also shown in this figure. As the figure indicates, mode 1 that is related to the Coriolis frequency has significant oscillations with the largest magnitude among the modes over the inertial timescale of τ_{ABL}. On the other hand, mode 12 has smaller magnitude oscillations but with a higher frequency (or a smaller timescale) related to turbulence fluctuations. When compared with the LES data, the figure indicates that the primary variations of the wind are related to the inertial flow oscillations of mode 1.

To ensure that KMD is capable of reconstructing the unsteady ABL flow, we also combined these modes by generating the reconstruction as R_τ explained above. The result of this reconstruction is shown as the red dashed line in Fig. 8(b). As the figure indicates, this reconstruction using ∼4% of the total Koopman modes agrees well with the LES data [compare red dashed and black lines in Fig. 8(b)]. The spatial structure of this reconstruction is also shown in Fig. 8(c) vs the actual LES data in Fig. 8(e) in the middle of the simulation. The figure shows that reconstructing the flow field using only ∼4% of the total Koopman modes could capture the primary flow structures such as large streaks while filtering some of the very small and high-frequency flow fluctuations. As shown before, we expect the accuracy of the reconstruction to increase with increasing the number of modes. To demonstrate this, we used 50% of the Koopman modes to reconstruct the flow field and displayed in Fig. 8(d). The figure exhibits a very good agreement between the reconstruction [Fig. 8(d)] and actual LES data [Fig. 8(e)] by capturing even tiny and high-frequency fluctuations. These results underscore the capability of KMD to reconstruct and capture complex ABL flow dynamics even under unsteady mean flow conditions.

F. The strength of KMD in characterizing roll and cell structures in ABLs

In this section, we will describe the performance of the KMD in characterizing the convective cell and roll structures in ABL, and determine some prevailing Koopman modes that are responsible for these coherent turbulence structures. To this end, we calculate the velocity statistics from a conditionally averaged flow field. We first calculate the 2D spatial correlation for velocity fluctuations at a reference height, z_{ref} (=260 m here) following Momen *et al.*[68] Then, a 3D spatial correlation function is calculated with respect to that reference height by changing z similar to previous studies.[82,83] The results of these correlations for two LES cases N16 and U4 are shown in Figs. 9(b) and 9(d) where the regions of positive correlations are marked as red and negative correlations are colored blue.

The results indicate that N16 case [Fig. 9(b)] includes roll vortices and low- and high-momentum streamwise-elongated streaks, which are the dominant flow mechanism of shear-induced boundary layers.[84,85] We also conducted a similar analysis using the Koopman reconstruction R_τ to evaluate the KMD performance in detecting these structures as shown in Fig. 9(a). The figure clearly shows these large streaks that flank each other in the cross-stream direction. The structures in R_τ are smoother than in the actual LES case since we did not use all the modes to create R_τ and we did not consider higher-frequency modes that are responsible for very small turbulence fluctuations. This will also lead to larger cleaner correlations in the KMD R_τ case since the noisy data of very small-scale eddies are filtered. Note again if we use all the Koopman modes, we will obtain almost exact LES results as shown in supplementary material, Fig. S4.

Next, the 3D autocorrelation is applied to the convective case U4. The results of LES case in Fig. 9(d) indicate a cell like structure unlike the roll structure in N16. This cell structure is a characteristic of the convective ABLs when the buoyancy over shear production of turbulence increases. We also conducted the 3D autocorrelation using the Koopman reconstruction R_τ as shown in Fig. 9(c). The results show this cell structure in the Koopman modes as well, but with a slightly bigger structure and higher correlations similar to the previous case. Hence, KMD is capable of detecting these coherent turbulence structures and will be very useful for detecting such large structures especially in very noisy data, e.g., real-world measurements.

G. Detecting Koopman modes related to roll and cell structures in ABLs via quadrant analysis

In this section, we will use KMD in a novel way to distinguish which prevailing modes correspond to roll and which to cell structures. To this end, we will use quadrant analysis with timescale

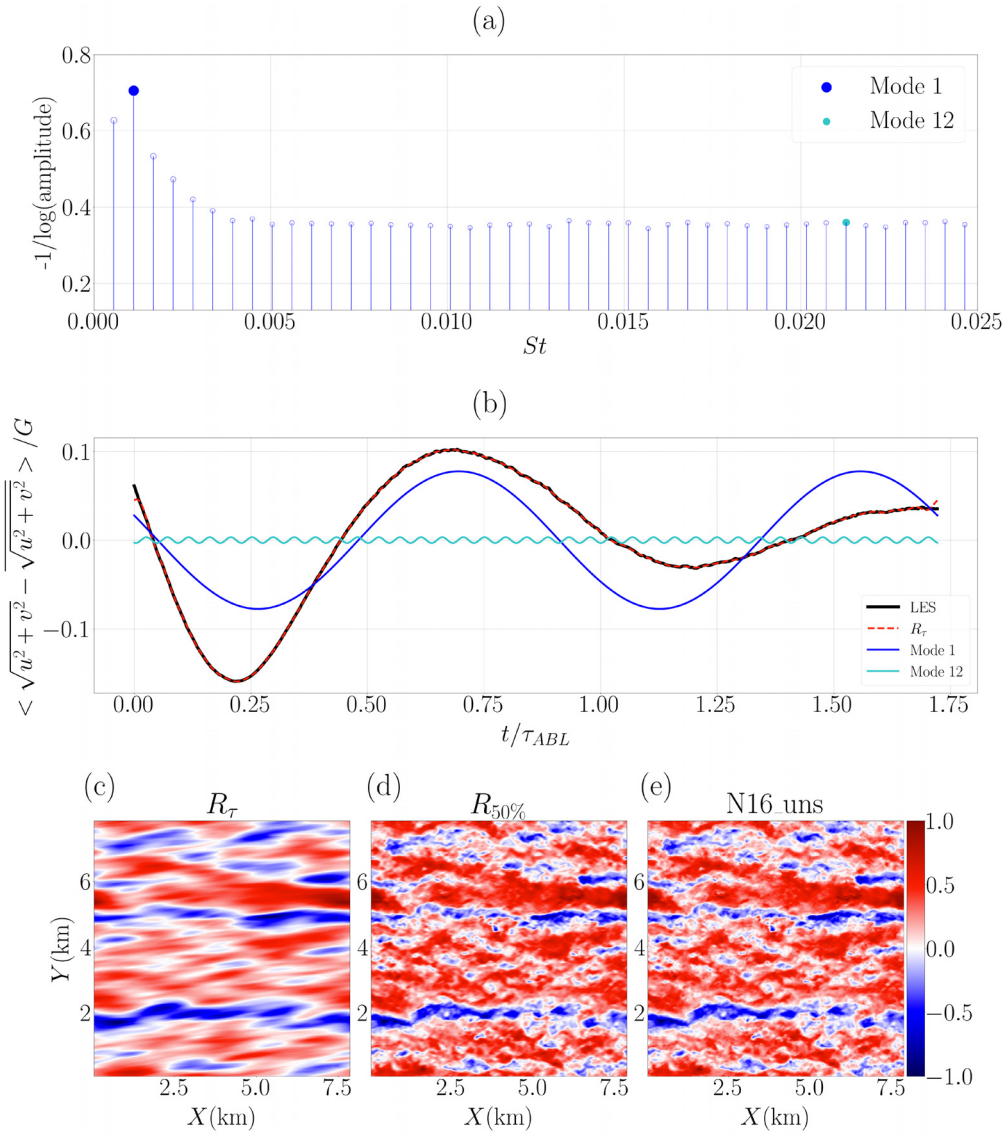

FIG. 8. (a) Koopman mode amplitude of unsteady ABL case vs frequency; note that the amplitude is shown in logarithmic scale. (b) Time series of the fluctuations in horizontal wind speed for LES, two important Koopman modes, and the reconstructed flow (R_τ) are shown. (c)–(e) The spatial structure of R_τ, $R_{50\%}$, and actual LES snapshot are depicted in the middle of the simulation.

interpretations of the Koopman modes. First, we show whether we would obtain similar quadrant analysis findings as in Figs. 3 and 4 when we use Koopman reconstructed flow. Hence, we conducted another KMD where we used the vertical velocity component in addition to the horizontal velocity as our observable datasets. After performing the decomposition on the u–w dataset, and reconstructing u and w separately, we performed quadrant analysis for the reconstructed flow. The Koopman reconstruction of the flow field yielded a similar result to the LES findings in Figs. 3 and 4, where the contribution of the Q2 + Q4 to the total Reynolds stress tensor decreases by increasing the ratio of buoyancy to shear production in the ABL (see the supplementary material, Figs. S5 and S6).

We then use this feature of convective ABLs to detect the Koopman modes that are responsible for roll- and cell-like structures. For the near neutral case U16, the first Koopman mode is shown in Fig. 10(a). The timescale of this mode is very close to the shear turbulence timescale z_i/u_* (≈1181 s for U16). Furthermore, we displayed the quadrant analysis of this mode in Fig. 10(c). As the figure shows, this mode mostly accounts for sweeps and ejection events in which Q2 + Q4 include more than 70% of the Reynolds shear stress tensor. These quadrants indicate gradient-type motions, which are a characteristic of typical shear-induced boundary layers or roll structures. The shape of this structure in Fig. 10(a) is also parallel to the mean flow direction [black arrow in Fig. 10(a)], which again shows this Koopman

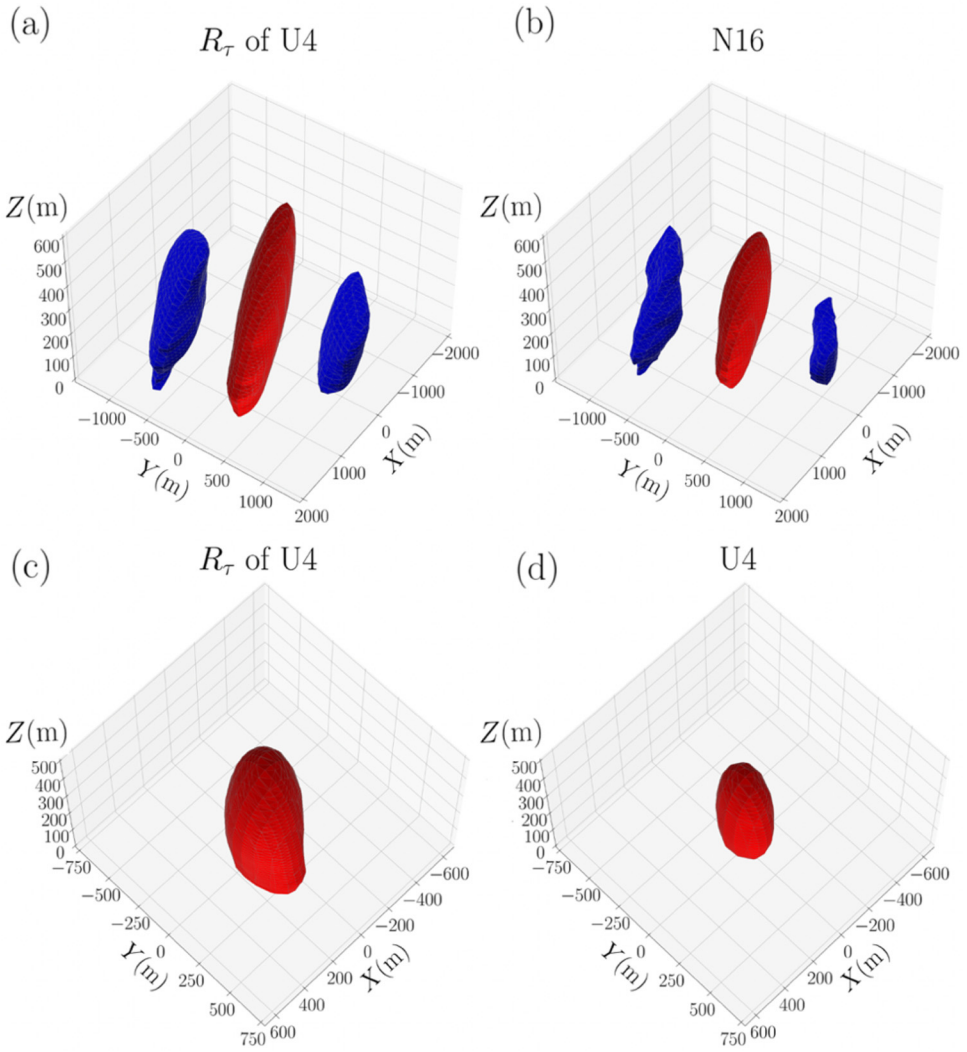

FIG. 9. 3D spatial correlation of (b) and (d) LES cases vs (a) and (b) Koopman mode reconstructions. Here, red isosurfaces show positive correlations [>0.35 for (a) and (b) and >0.5 for (c) and (d)], while blue isosurfaces show negative correlations [<-0.25 for (a) and (b)].

mode 1 is responsible for generating roll structures in the near-neutral ABL.

The spatial structure of Koopman mode 15 in convective case U4 is also shown in Fig. 10(b). We selected this mode because of the following two reasons. First, the timescale of this mode is close to z_i/w_* ($\approx$551 s for U4 case), where w_* is the convective velocity scale. Second, unlike the previous mode, most of the Reynolds shear stress of this mode is from counter-gradient motions by having more than 75% of the velocity fluctuations in Q1 + Q3 quadrants as shown in Fig. 10(d). This figure demonstrates that most of the velocity fluctuations of this mode are outward and inward interactions, which we showed are higher in convective mixed ABLs. Hence, these two reasons can attribute this mode to the creation of the convective cell structures. In fact, the spatial structure of this mode in Fig. 10(b) indicates cell-like structures in the ABL. This structure has a non-trivial shape, which is transverse to the mean flow direction [black arrow in Fig. 10(b)]. Highly convective thermal plumes alter roll structures and lead to the formation of cell structures in convective ABLs. Such transverse turbulence structures in unstable cases can be associated with these convective cell structures. Although this selected mode is not necessarily the most energetic mode of this case, since its timescale is close to the convective velocity timescale, and it displays counter-gradient motions, it could be a good representative of cell structures. Furthermore, we reconstructed the flow using all the modes with roll and cell-like structure features among the first 20 largest amplitude Koopman modes for both cases, and found similar quadrant analysis and spatial structure results (see supplementary material, Fig. S7). Therefore, these results indicate that the KMD along with the timescale and quadrant analyses can provide significant new insights into the underlying dynamics of coherent turbulence structures in ABLs.

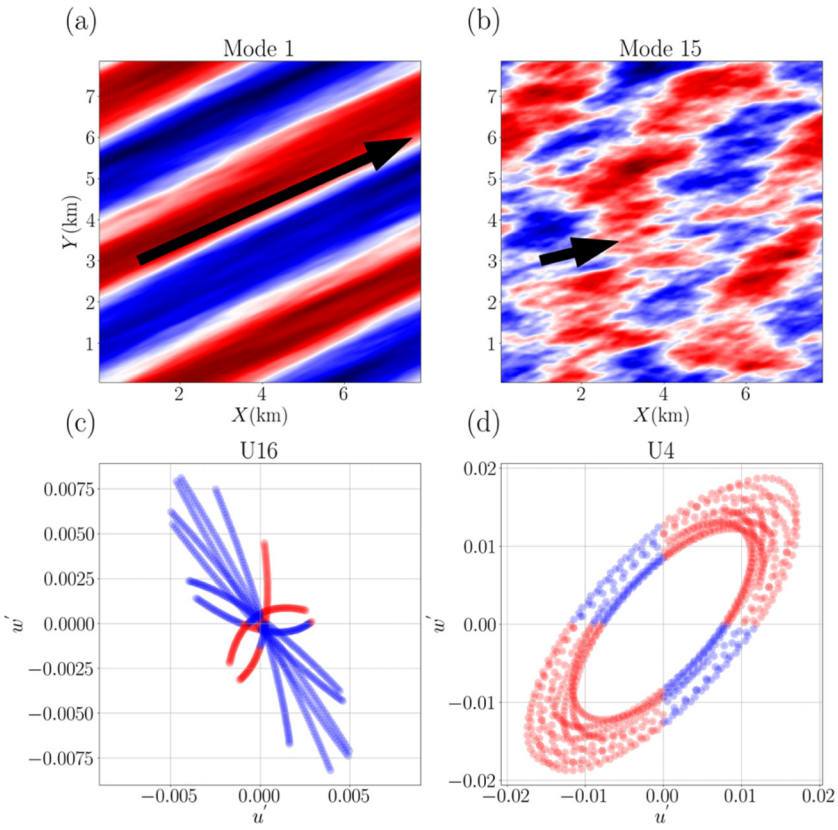

FIG. 10. (a) and (b) Two important Koopman modes for cases U16 and U4 and (c) and (d) the quadrant analysis of these modes are presented at elevation $z = 345$ m. The percentage of the velocity fluctuation points, which are in Q2 + Q4 for mode 1 of the near neutral case U16, is 77.8%, and for mode 15 of the convective case U4 is 23.5%.

H. Characterizing convective cells and rolls in vertical velocity via KMD

The structure of convective cells and rolls in ABLs can be better noticed from the vertical velocity components where they depict the organization of updraft and downdraft motions.[86] Figure 11 exhibits the instantaneous snapshots of fluctuating normalized vertical velocity contours for the U16 and U4 cases. The figure clearly shows roll vortices in near neutral U16 case (Fig. 11, top right panel), and as the ratio of buoyancy to shear production increases, the convective cells form in the ABL (Fig. 11, bottom right panel). These coherent turbulence structures in the ABL can also be seen at other elevations as shown in supplementary material, Figs. S8 and S9. The formation of these convective rolls and cells in ABLs is also consistent with previous studies that associated them with z_i/L values.[30,33]

The resulting KMD modes can be examined for these two distinctive coherent turbulence structures. Figure 11 shows the first and second Koopman modes, as well as the reconstructed flow using R_τ. The first two Koopman modes in the near-neutral case U16 are aligned with the mean wind direction (compare their direction with the black arrow in the right panel of this figure), and thus represent the roll structures. However, the first mode of the convective case U4 is slightly tilted from the mean wind velocity direction (compare with the arrow shown in the bottom right panel). Furthermore, for the convective case U4, some Koopman modes are even transverse to the mean wind direction. These transverse structures (e.g., mode 2 of U4) are associated with the existence of cell structures, and the presence of convection in the ABL. In fact, all such modes together add up to give rise to the cell structures in the U4 case. These coherent transverse structures cannot be trivially noticed from the actual LES snapshots in the right panels of Fig. 11. Hence, KMD can detect such non-trivial turbulence structures in the ABL by decomposing the spatiotemporal dynamics of the LES data. Similar roll and cell structures exist at other elevations as shown in supplementary material, Figs. S8 and S9.

We also applied the KMD to the horizontal velocity components of three N16, U16, and U4 cases and obtained a similar finding. In the neutral and near-neutral cases, the primary Koopman modes were related to the roll structures. However, in the convective ABL, some non-trivial Koopman modes transverse to the mean flow direction were observed. These results are shown for four elevations in supplementary material, Figs. S10–S13.

To corroborate the generality of the findings for different grid resolutions, another set of LES cases was performed using a reduced grid resolution of $96 \times 96 \times 96$. First, to verify if the LES mean wind and Reynolds stress profiles are well converged, we plotted Fig. 2 for both resolutions on top of each other. We found that our chosen resolutions are sufficiently converged (see the supplementary material, Fig. S14), consistent with previous studies, which used a similar LES model to our code.[33,69,87] Subsequently, the KMD was applied to these coarse resolution cases. Furthermore, the overall range of the obtained Koopman mode amplitudes and frequencies of these two resolutions agree with each other (see supplementary material, Fig. S15). The results demonstrate that while the low-energy finer decomposed

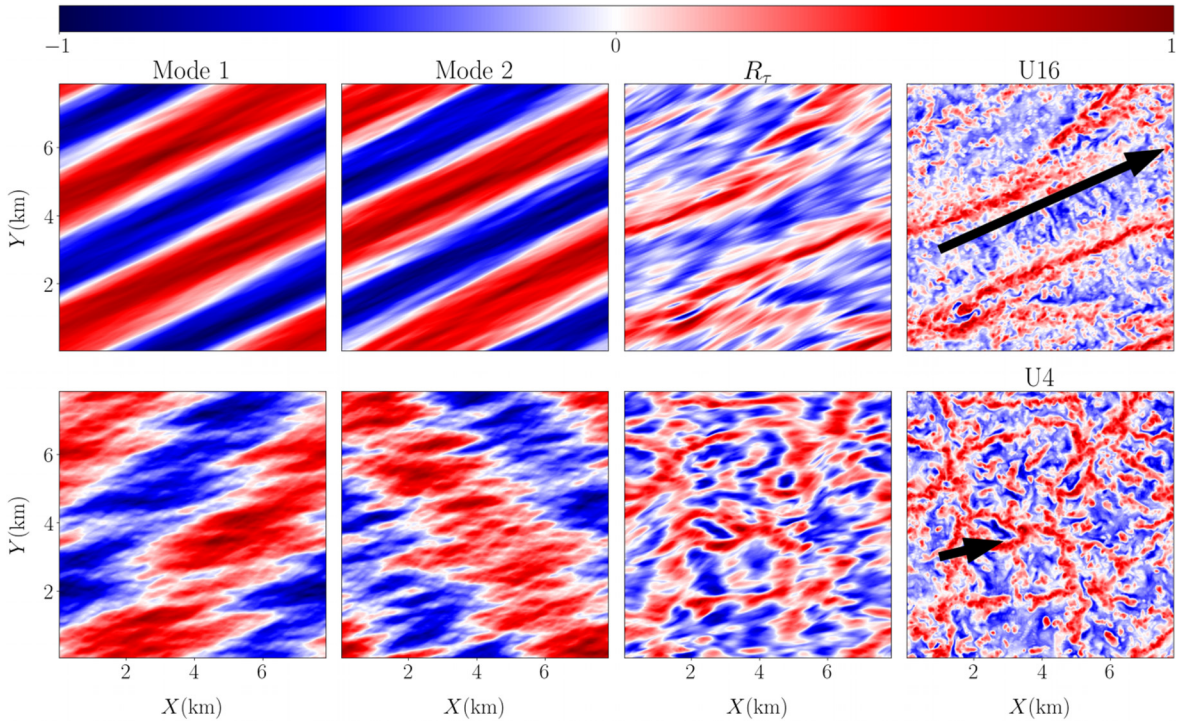

FIG. 11. Contours of the first and second Koopman modes (first two columns), reconstructed flow using the first R_τ Koopman modes (third column), and instantaneous snapshots of normalized fluctuating vertical velocity component (fourth column) for (top) U16 and (bottom) U4 cases at $z = 189$ m. The black arrows in the right panels depict the direction of the mean wind velocity vector.

structures may vary between the two resolutions, the large-scale energy-containing structures of the KMD modes remain consistent with those obtained from the higher grid resolution of $192 \times 192 \times 192$ (see supplementary material, Fig. S16). Therefore, the obtained findings are not grid dependent based on the performed grid dependence test. Hence, the KMD provides a general consistent method for decomposing non-trivial turbulence structures in the ABL.

I. A novel data-driven approach to categorize Koopman modes

One drawback of the KMD approach is that it produces many modes (e.g., 12 957 modes in the considered cases) for which some of them are correlated with each other. Hence, distinguishing the Koopman modes and the underlying dynamics associated with them becomes challenging. Moreover, sometimes a physically meaningful mode may have a low amplitude and only appear in a high mode number. Therefore, by investigating only the first few modes, we might lose some important information about the flow dynamics. These issues led us to perform a systematic data-driven method to classify all the KMD modes into smaller uncorrelated groups, so that we can attribute them to primary flow dynamics. To this end, we use unsupervised K-means clustering, which is a well-known method for clustering large data without *a priori* training.[88–90]

First, we applied the K-means clustering to the spatial Koopman mode data, resulting in the classification of modes into distinct, uncorrelated groups. The Koopman mode data are the spatial mode structure (e.g., Fig. 11), which is represented as a matrix derived from KMD. These matrices are then treated as high-dimensional vectors by flattening them. Hence, each vector's dimension is equal to the product of the matrix's dimensions (e.g., 192×192 here). This process treats each mode as a distinct data point, with the algorithm iterating to minimize the clustering error through centroid recalibration, by calculating the distance of the high-dimensional vectors. The clusters were ordered based on the largest mode amplitude within each group, yielding physically meaningful results (please see supplementary material, Text S8 and Fig. S17 for more information). Despite the effectiveness of this method in revealing modes with similar spatial structures, such as streamwise streaks and transverse structures, it exhibited several issues. For example, this method does not account for the displacement and rotation in structures. Hence, some structures with similar physical dynamics were classified into separate groups. To overcome these limitations, we used K-means clustering with deep convolutional neural networks in a novel way. CNNs account for the displacement and rotation of the spatial structures and hence will be very useful for classifying the obtained Koopman modes.

We utilize the Inception V3 CNN architecture to extract features from the set of Koopman modes. The Inception V3 network was introduced in 2015 and has been widely used for various image recognition tasks.[53,91,92] This CNN is trained on the large ImageNet dataset, which makes it a good candidate for transfer learning.[93] This implies that it can be fine-tuned on a smaller dataset for a specific task, such as object detection or image segmentation, and still achieve a good performance. This CNN reached a 93.7% top-5 accuracy for image classification of

this large dataset.[91] Furthermore, the idea behind developing Inception V3 was to reduce the computational cost without compromising the ability of the network to generalize at deeper levels. On the other hand, its architecture is designed to be scalable, making it adaptable to different datasets.[91] This model had also the best performance among four different tested CNNs for accurately detecting different weather patterns.[53]

The input of the Inception V3 CNN is the images derived from the spatial Koopman mode matrices. Each Koopman mode is converted into an image format and is resized to 299×299 pixels to align with the standard input size of the Inception V3 architecture. Then, the CNN determines the features of each image, and in the final average pooling layer of the network outputs 2048-dimensional feature vectors for each mode. These feature vectors serve as a compact yet comprehensive representation of the spatial patterns of the Koopman modes. Finally, rather than the Koopman mode images, we provide their corresponding feature vector space from CNN as inputs of the K-means clustering method. This approach clusters the images into an arbitrary number of distinct groups based on the similarity of their feature vectors. By calculating the inertia and using the elbow method, four clusters appeared to be sufficient to categorize the resulting Koopman modes in the considered cases (supplementary material, Text S9 and Fig. S19). Figure 12 presents two representative modes of each of the four clusters for U4 and U16 cases.

The resulting KMD clusters seem to be significantly different and represent a distinctive flow dynamic. As the figure indicates, this technique resolves the issue of the previous method by accounting for displacement and rotation in Koopman modes. For example, cluster A of U16 (Fig. 12 top left) displays two displaced modes classified in one cluster, while cluster A of U4 (Fig. 12 bottom left) shows two rotated modes grouped into one cluster. Note that the clustering of displaced modes into one group is consistent with the fundamental KMD basis. Koopman modes will be displaced in space when evolving in time because their spatial structures, encoded by the eigenfunctions, are multiplied by exponential terms representing their temporal dynamics. This causes the modes to shift or propagate spatially as time progresses, while they carry the same structure. Moreover, this approach resolves the slow decrease in inertia in the previous method by ensuing a quick decrease in inertia, which shows that four clusters are sufficient to describe the overall structure of these KMD modes (supplementary material, Figs. S18 and S19). The numbers in each box of Fig. 12 show the number of Koopman modes that belong to that cluster. As the figure indicates unlike the previous method, there are considerably more modes in each four clusters with a smoother distribution when using CNN. Essentially, based on this approach, the modes are divided into large (e.g., cluster A of U4), medium (e.g., cluster B of U4), and small-scale (e.g., cluster D of U4) structures. The larger modes represent primary mean flow large-scale dynamics, such as convective rolls and cells, while the smaller modes display small-scale turbulent fluctuations. To further corroborate this finding, we analyzed the Koopman amplitude and frequency of each of these four clusters (see supplementary material, Text S10 and Fig. S20). Our results show that the average Koopman amplitude of cluster A is 0.012, cluster B 0.008, cluster C 0.005, and cluster D 0.003 for case U4, indicating a cascade of energy content. Furthermore, the average frequency increases from clusters A to D demonstrating that the timescale of the eddies decreases as we go from clusters A to D. Similar results were found for case U16 corroborating our finding that cluster A represents large-energetic eddies, B medium, and D small-scale eddies. Thus, by combining KMD and CNN, the introduced approach is capable of disentangling and classifying non-trivial coherent turbulence structures in a fully data-driven approach, which can be used to improve our understanding of the underlying dynamics in ABL flows.

IV. CONCLUSIONS

In this paper, coherent turbulence structures of neutral and convective ABL flows were characterized using a combination of LES cases and data-driven methods. In total, eight LESs were conducted by varying the ratio of the buoyancy to shear production of TKE and mean flow forcing under two coarse and fine resolutions. Next, the KMD method was applied to the LES results to determine the nonlinear spatiotemporal dynamics of turbulent eddies in ABL flows. Then, using unsupervised machine learning methods (K-means) in conjunction with deep CNNs, the resulting Koopman modes were categorized into a reduced number of clusters, which are associated with distinctive flow dynamics. In summary, the key findings of this study are as follows:

1. The quadrant analysis of ABL flows indicates that unlike the neutral cases in which sweeps and ejections (gradient type) are dominant eddy motions (∼65.2%), their contribution decreases in convective ABLs (∼55.5%), and the contribution of outward and inward interactions (counter-gradient type) increases. This is due to the thermal plumes in convective ABLs that impact hairpin vortices and roll structures in neutral ABLs and lead to the formation of convective cells.
2. The reconstructed flows using Koopman modes demonstrated the KMD's effectiveness in capturing the primary dynamics of turbulent eddies in ABLs. The largest amplitude Koopman modes were able to decompose unique dynamics in ABL flows such as the imposed pressure gradient, buoyancy convection, and Coriolis force. Using the time scale analysis, we determined Koopman modes that were related to the inertial oscillations (Coriolis frequency), shear turbulence fluctuations (z_i/u_*), and convection (z_i/w_*). Moreover, the Koopman modes with frequencies lower than $10 \times z_i/u_*$ were able to accurately reconstruct the overall ABL flow field and 3D spatial correlation of turbulent roll and cell structures in neutral and convective cases.
3. The efficacy of KMD in reconstructing unsteady ABL conditions under transient pressure gradient forcing was also evaluated. The reconstructed Koopman flow field was able to capture the highly nonlinear mean-turbulence interactions under unsteady forcing. The largest Koopman mode was associated with the inertial oscillations (Coriolis frequency) in unsteady ABL conditions as expected from the underlying dynamics. Using only ∼5% of the Koopman modes, the reconstructed Koopman flow field agreed well with the actual LES data.
4. KMD was shown to detect non-trivial turbulence structures in ABLs. In strongly convective cases, some non-trivial modes were obtained that are transverse to the mean wind direction. This was associated with the formation of cell structures in convective ABLs. To corroborate this, a quadrant analysis of the Koopman modes was conducted in a novel way and modes that were associated with roll and cell structures were characterized. The time scale and spatial structure of the detected modes using this approach further verified the findings.

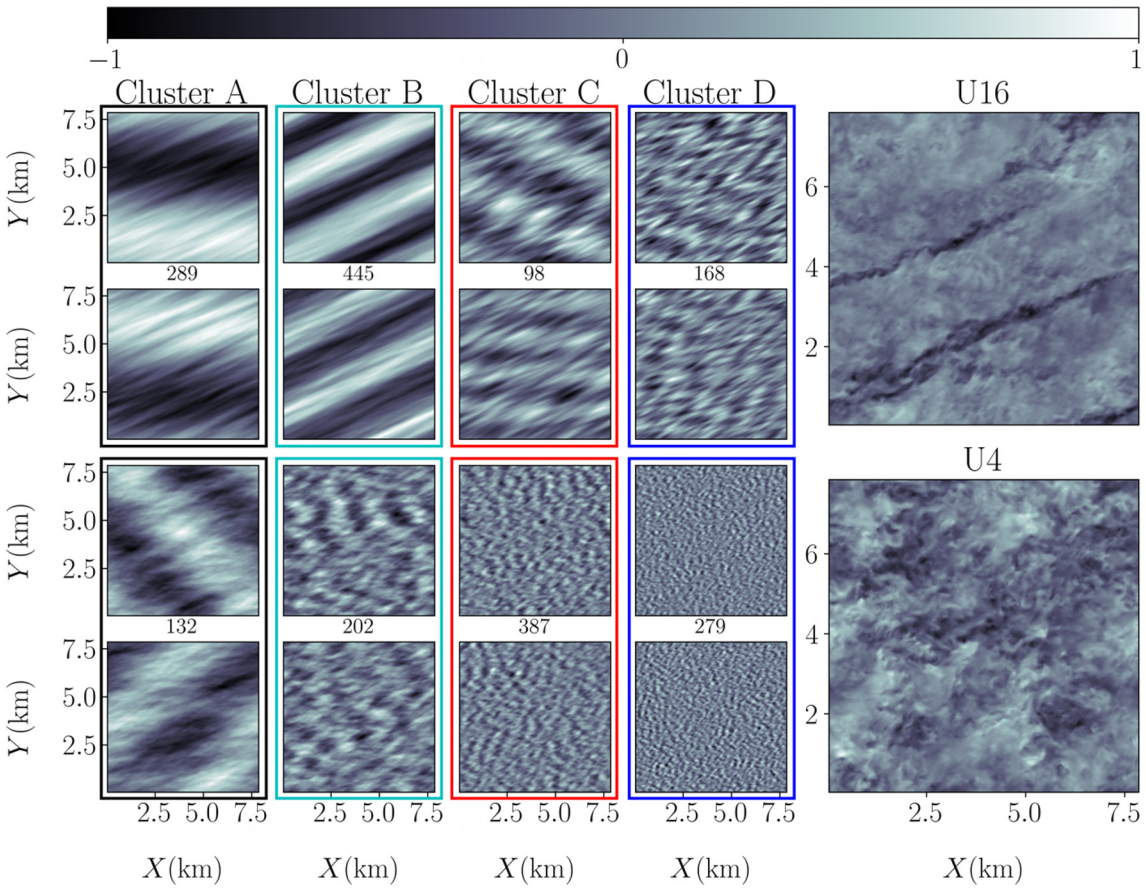

FIG. 12. The contours of two representative Koopman modes belonging to each cluster, and instantaneous snapshots of the normalized fluctuating horizontal velocity component for (top) U16 and (bottom) U4 cases at $z = 345$ m. The colored rectangular boxes show that the modes in each box belong to one cluster. The numbers in each rectangle indicate the number of modes in that cluster (e.g., cluster A of U16 case includes 289 modes in which two of them are shown in the top left panel).

5. Using less than ∼10% of the Koopman modes yielded a reconstruction with over 80% correlation with the LES data in the considered cases. This can help us to better understand the underlying dynamics of complex ABL flows using a few reduced modes. Furthermore, it can be used as a tool for reducing the storage of large LES data by capturing the primary spatiotemporal dynamics of turbulent flows.
6. A novel data-driven approach was also introduced to categorize many obtained Koopman modes by combining K-means clustering with deep CNNs. The introduced method is rotation and displacement free and can classify the spatial structure of the modes with similar physics that are tilted or displaced into one cluster. The resulting clusters of this method led to unique turbulence structures in each cluster, representing different dynamics in the ABL. The inertia of this method indicated that a few clusters (here ∼4) could represent the primary dynamics of all the Koopman modes in the considered ABL cases.

Overall, our findings underscore the potential strength of KMD and data-driven techniques to better understand and determine the complex dynamics of ABL flows. Using these methods, new insights were provided into the coherent turbulence structures of neutral and convective ABLs, and these nonlinear turbulent flows were reconstructed using only a small number of Koopman modes. Our results can be used to increase the understanding of ABL dynamics, characterize coherent turbulence structures in meteorological and wind energy applications, and improve the parametrization of ABLs in weather/climate models.

SUPPLEMENTARY MATERIAL

See the supplementary material for figures that provide further supporting details about the discussions, which include mean wind velocity and shear production profiles, KMD *dt* sensitivity, Koopman reconstructions of the spatial correlations and quadrant analysis, and K-means clustering details of the Koopman modes.

ACKNOWLEDGMENTS

The simulations were performed on the University of Houston's computing clusters (Carya and Sabine). The authors also acknowledge computing cluster support from the National Center for Atmospheric Research (NCAR) under project number

UHOU0002 as well as Bridges-2 at Pittsburgh Supercomputing Center through allocation EES230054 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which is supported by NSF. The authors acknowledge financial support from the Department of Civil and Environmental Engineering at the University of Houston via startup funds. M.M. was also supported by the Physical and Dynamic Meteorology Program of the National Science Foundation under grant AGS-2228299.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Milad Rezaie: Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Validation (equal); Visualization (equal); Writing—original draft (equal); Writing—review & editing (equal). **Mostafa Momen:** Conceptualization (lead); Formal analysis (equal); Funding acquisition (lead); Investigation (equal); Methodology (equal); Project administration (lead); Resources (lead); Supervision (lead); Writing—original draft (equal); Writing—review & editing (lead).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1]J. R. Garratt, "The atmospheric boundary layer," Earth-Sci. Rev. **37**(1–2), 89–134 (1994).
[2]R. B. Stull, *An Introduction to Boundary Layer Meteorology* (Springer Science & Business Media, 1988).
[3]M. J. Churchfield, S. Lee, J. Michalakes, and P. J. Moriarty, "A numerical study of the effects of atmospheric and wake turbulence on wind turbine dynamics," J. Turbul. **13**, N14 (2012).
[4]M. Calaf, C. Meneveau, and J. Meyers, "Large eddy simulation study of fully developed wind-turbine array boundary layers," Phys. Fluids **22**(1), 15110 (2010).
[5]P. Holmes, J. L. Lumley, G. Berkooz, and C. W. Rowley, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry* (Cambridge university Press, 2012).
[6]J. W. Kurelek, A. R. Lambert, and S. Yarusevych, "Coherent structures in the transition process of a laminar separation bubble," AIAA J. **54**(8), 2295–2309 (2016).
[7]L. Mahrt and W. Gibson, "Flux decomposition into coherent structures," Boundary-Layer Meteorol. **60**(1–2), 143–168 (1992).
[8]C.-H. Moeng and P. P. Sullivan, "A comparison of shear-and buoyancy-driven planetary boundary layer flows," J. Atmos. Sci. **51**(7), 999–1022 (1994).
[9]S. Khanna and J. G. Brasseur, "Three-dimensional buoyancy- and shear-induced local structure of the atmospheric boundary layer," J. Atmos. Sci. **55**(5), 710–743 (1998).
[10]R. Floors, A. Peña, and S. E. Gryning, "The effect of baroclinicity on the wind in the planetary boundary layer," Q. J. R. Meteorol. Soc. **141**(687), 619–630 (2015).
[11]L. Mahrt, "Stratified atmospheric boundary layers," Boundary-Layer Meteorol. **90**(3), 375–396 (1999).
[12]F. T. M. Nieuwstadt, "The atmospheric boundary layer," in *Environmental Stratified Flows*, CISM International Center for Mechanical Sciences Courses and Lectures Vol. 479 (Springer, 2005), pp. 179–232.
[13]P. Ramamurthy, E. Bou-Zeid, J. A. Smith, Z. Wang, M. L. Baeck, N. Z. Saliendra, J. L. Hom, and C. Welty, "Influence of subfacet heterogeneity and material properties on the urban surface energy budget," J. Appl. Meteorol. Climatol. **53**(9), 2114–2129 (2014).
[14]I. Stiperski, M. Calaf, and M. W. Rotach, "Scaling, anisotropy, and complexity in near-surface atmospheric turbulence," J. Geophys. Res. Atmos. **124**(3), 1428–1448, https://doi.org/10.1029/2018JD029383 (2019).
[15]R. Stoll and F. Porté-Agel, "Surface heterogeneity effects on regional-scale fluxes in stable boundary layers: Surface temperature transitions," J. Atmos. Sci. **66**(2), 412–431 (2009).
[16]M. Sharan and S. G. Gopalakrishnan, "Mathematical modeling of diffusion and transport of pollutants in the atmospheric boundary layer," Pure Appl. Geophys. **160**(1–2), 357–394 (2003).
[17]T. Mauritsen, G. Svensson, S. S. Zilitinkevich, I. Esau, L. Enger, and B. Grisogono, "A total turbulent energy closure model for neutrally and stably stratified atmospheric boundary layers," J. Atmos. Sci. **64**(11), 4113–4126 (2007).
[18]G. S. Poulos, W. Blumen, D. C. Fritts, J. K. Lundquist, J. Sun, S. P. Burns, C. Nappo, R. Banta, R. Newsom, and J. Cuxart, "CASES-99: A comprehensive investigation of the stable nocturnal boundary layer," Bull. Am. Meteorol. Soc. **83**(4), 555–582 (2002).
[19]J. D. Mirocha and B. Kosović, "A large-eddy simulation study of the influence of subsidence on the stably stratified atmospheric boundary layer," Boundary-Layer Meteorol. **134**, 1–21 (2010).
[20]J. A. Gibbs, R. Stoll, and S. T. Salesky, "Inclination angles of turbulent structures in stably stratified boundary layers," Boundary-Layer Meteorol. **186**(1), 27–41 (2023).
[21]D. J. Tritton, "Stabilization and destabilization of turbulent shear flow in a rotating fluid," J. Fluid Mech. **241**, 503–523 (1992).
[22]O. Romdhani, J. A. Zhang, and M. Momen, "Characterizing the impacts of turbulence closures on real hurricane forecasts: A comprehensive joint assessment of grid resolution, horizontal turbulence models, and horizontal mixing length," J. Adv. Model. Earth Syst. **14**(9), e2021MS002796 (2022).
[23]F. Sabet, Y. R. Yi, L. Thomas, and M. Momen, "Characterizing mean and turbulent structures of hurricane winds via large-eddy simulations," Center for Turbulence Research Proceedings of the Summer Program 2022:311–321, 2022; see https://web.stanford.edu/group/ctr/ctrsp22/v01_Sabet.pdf.
[24]L. Matak and M. Momen, "The role of vertical diffusion parameterizations in the dynamics and accuracy of simulated intensifying hurricanes," Boundary-Layer Meteorol. **188**(3), 389–418 (2023).
[25]M. Momen and E. Bou-Zeid, "Analytical reduced models for the non-stationary diabatic atmospheric boundary layer," Boundary-Layer Meteorol. **164**, 383–399 (2017).
[26]B. J. H. Van de Wiel, A. F. Moene, and H. J. J. Jonker, "The cessation of continuous turbulence as precursor of the very stable nocturnal boundary layer," J. Atmos. Sci. **69**(11), 3097–3115 (2012).
[27]A. A. M. Holtslag, G. Svensson, P. Baas, S. Basu, B. Beare, A. C. M. Beljaars, F. C. Bosveld, J. Cuxart, J. Lindvall, and G. J. Steeneveld, "Stable atmospheric boundary layers and diurnal cycles: Challenges for weather and climate models," Bull. Am. Meteorol. Soc. **94**(11), 1691–1706 (2013).
[28]M. Momen and E. Bou-Zeid, "Large-eddy simulations and damped-oscillator models of the unsteady Ekman boundary layer," J. Atmos. Sci. **73**(1), 25–40 (2016).
[29]M. Momen and E. Bou-Zeid, "Mean and turbulence dynamics in unsteady Ekman boundary layers," J. Fluid Mech. **816**, 209–242 (2017).
[30]M. Momen, E. Bou-Zeid, M. B. Parlange, and M. Giometto, "Modulation of mean wind and turbulence in the atmospheric boundary layer by baroclinicity," J. Atmos. Sci. **75**(11), 3797–3821 (2018).
[31]M. Momen, "Baroclinicity in stable atmospheric boundary layers: Characterizing turbulence structures and collapsing wind profiles via reduced models and large-eddy simulations," Q. J. R. Meteorol. Soc. **148**(742), 76–96 (2022).
[32]T. M. Weckwerth, T. W. Horst, and J. W. Wilson, "An observational study of the evolution of horizontal convective rolls," Mon. Weather Rev. **127**(9), 2160–2179 (1999).
[33]S. T. Salesky, M. Chamecki, and E. Bou-Zeid, "On the nature of the transition between roll and cellular organization in the convective boundary layer," Boundary-Layer Meteorol. **163**, 41–68 (2017).

[34]B. Brümmer, "Roll and cell convection in wintertime Arctic cold-air outbreaks," J. Atmos. Sci. **56**(15), 2613–2636 (1999).
[35]P. Benner, S. Gugercin, and K. Willcox, "A survey of projection-based model reduction methods for parametric dynamical systems," SIAM Rev. **57**(4), 483–531 (2015).
[36]C. W. Rowley, I. Mezi, S. Bagheri, P. Schlatter, and D. S. Henningson, "Spectral analysis of nonlinear flows," J. Fluid Mech. **641**, 115–127 (2009).
[37]Q. Zhang, Y. Liu, and S. Wang, "The identification of coherent structures using proper orthogonal decomposition and dynamic mode decomposition," J. Fluids Struct. **49**, 53–72 (2014).
[38]S. Bagheri, "Koopman-mode decomposition of the cylinder wake," J. Fluid Mech. **726**, 596–623 (2013).
[39]D. Giannakis, A. Kolchinskaya, D. Krasnov, and J. Schumacher, "Koopman analysis of the long-term evolution in a turbulent convection cell," J. Fluid Mech. **847**, 735–767 (2018).
[40]I. Mezić, "Spectrum of the Koopman operator, spectral expansions in functional spaces, and state-space geometry," J. Nonlinear Sci. **30**(5), 2091–2145 (2020).
[41]J. N. Kutz, J. L. Proctor, and S. L. Brunton, "Koopman theory for partial differential equations," arXiv:1607.07076 (2016).
[42]M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, "A data–driven approximation of the koopman operator: Extending dynamic mode decomposition," J. Nonlinear Sci. **25**, 1307–1346 (2015).
[43]M. O. Williams, C. W. Rowley, I. Mezić, and I. G. Kevrekidis, "Data fusion via intrinsic dynamic variables: An application of data-driven Koopman spectral analysis," Europhys. Lett. **109**(4), 40007 (2015).
[44]A. Towne, O. T. Schmidt, and T. Colonius, "Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis," J. Fluid Mech. **847**, 821–867 (2018).
[45]A. Likas, N. Vlassis, and J. J. Verbeek, "The global k-means clustering algorithm," Pattern Recognit. **36**(2), 451–461 (2003).
[46]C. Jiang, R. Vinuesa, R. Chen, J. Mi, S. Laima, and H. Li, "An interpretable framework of data-driven turbulence modeling using deep neural networks," Phys. Fluids **33**(5), 55133 (2021).
[47]G. V. Iungo, R. Maulik, S. A. Renganathan, and S. Letizia, "Machine-learning identification of the variability of mean velocity and turbulence intensity for wakes generated by onshore wind turbines: Cluster analysis of wind LiDAR measurements," J. Renewable Sustainable Energy **14**(2), 23307 (2022).
[48]E. S. Rosa, R. M. Salgado, T. Ohishi, and N. Mastelari, "Performance comparison of artificial neural networks and expert systems applied to flow pattern identification in vertical ascendant gas–liquid flows," Int. J. Multiphase Flow **36**(9), 738–754 (2010).
[49]E. Kaiser, B. R. Noack, L. Cordier, A. Spohn, M. Segond, M. Abel, G. Daviller, J. Östh, S. Krajnović, and R. K. Niven, "Cluster-based reduced-order modelling of a mixing layer," J. Fluid Mech. **754**, 365–414 (2014).
[50]R. Chauhan, K. K. Ghanshala, and R. C. Joshi, "Convolutional neural network (CNN) for image detection and recognition," in *ICSCCC 2018—First International Conference on Secure Cyber Computing and Communication* (IEEE, 2018), pp. 278–282.
[51]Y. H. Liu, "Feature extraction and image recognition with convolutional neural networks," J. Phys. **1087**(6), 062032 (2018).
[52]W. Wang, Y. Sheng, J. Wang, X. Zeng, X. Ye, Y. Huang, and M. Zhu, "HAST-IDS: Learning hierarchical spatial-temporal features using deep neural networks to improve intrusion detection," IEEE Access **6**, 1792–1806 (2018).
[53]Y. Li and M. Momen, "Detection of weather events in optical satellite data using deep convolutional neural networks," Remote Sens. Lett. **12**(12), 1227–1237 (2021).
[54]J. Miller, U. Nair, R. Ramachandran, and M. Maskey, "Detection of transverse cirrus bands in satellite imagery using deep learning," Comput. Geosci. **118**, 79–85 (2018).
[55]I. Mezić, "Analysis of fluid flows via spectral properties of the Koopman operator," Annu. Rev. Fluid Mech. **45**, 357–378 (2013).
[56]T. Krake, D. Weiskopf, and B. Eberhardt, "Dynamic mode decomposition: Theory and data reconstruction," arXiv:1909.10466 (2019).
[57]B. O. Koopman, "Hamiltonian systems and transformation in Hilbert space," Proc. Natl. Acad. Sci. **17**(5), 315–318 (1931).
[58]H. Arbabi and I. Mezić, "Study of dynamics in post-transient flows using Koopman mode decomposition," Phys. Rev. Fluids **2**(12), 124402 (2017).
[59]R. Stoll, J. A. Gibbs, S. T. Salesky, W. Anderson, and M. Calaf, "Large-eddy simulation of the atmospheric boundary layer," Boundary-Layer Meteorol. **177**, 541–581 (2020).
[60]E. Bou-Zeid, C. Meneveau, and M. Parlange, "A scale-dependent Lagrangian dynamic model for large eddy simulation of complex turbulent flows," Phys. Fluids **17**(2), 025105 (2005).
[61]V. Kumar, G. Svensson, A. A. M. Holtslag, C. Meneveau, and M. B. Parlange, "Impact of surface flux formulations and geostrophic forcing on large-eddy simulations of diurnal atmospheric boundary layer flow," J. Appl. Meteorol. Climatol. **49**(7), 1496–1516 (2010).
[62]J. Kleissl, V. Kumar, C. Meneveau, and M. B. Parlange, "Numerical study of dynamic Smagorinsky models in large-eddy simulation of the atmospheric boundary layer: Validation in stable and unstable conditions," Water Resour. Res. **42**(6), W06D10, https://doi.org/10.1029/2005WR004685 (2006).
[63]J. D. Albertson and M. B. Parlange, "Surface length scales and shear stress: Implications for land-atmospheric interaction over complex terrain," Water Resour. Res. **35**(7), 2121–2132, https://doi.org/10.1029/1999WR900094 (1999).
[64]J. D. Albertson and M. B. Parlange, "Natural integration of scalar fluxes from complex terrain," Adv. Water Resour. **23**, 239–252 (1999).
[65]V. Kumar, J. Kleissl, C. Meneveau, and M. B. Parlange, "Large-eddy simulation of a diurnal cycle of the atmospheric boundary layer: Atmospheric stability and scaling issues," Water Resour. Res. **42**(6), W06D09, https://doi.org/10.1029/2005WR004651 (2006).
[66]V. Sharma, M. B. Parlange, and M. Calaf, "Perturbations to the spatial and temporal characteristics of the diurnally-varying atmospheric boundary layer due to an extensive wind farm," Boundary-Layer Meteorol. **162**(2), 255–282 (2017).
[67]M. Chamecki, C. Meneveau, and M. B. Parlange, "Large eddy simulation of pollen transport in the atmospheric boundary layer," J. Aerosol Sci. **40**(3), 241–255 (2009).
[68]M. Momen, M. B. Parlange, and M. G. Giometto, "Scrambling and reorientation of classical boundary layer turbulence in hurricane winds," Geophys. Res. Lett. **48**(7), e2020GL091695, https://doi.org/10.1029/2020GL091695 (2021).
[69]S. Shah and E. Bou-Zeid, "Very-large-scale motions in the atmospheric boundary layer educed by snapshot proper orthogonal decomposition," Boundary-Layer Meteorol. **153**(3), 355–387 (2014).
[70]C. Q. Dias-Junior, E. P. Marques Filho, and L. D. A. Sa, "A large eddy simulation model applied to analyze the turbulent flow above Amazon forest," J. Wind Eng. Ind. Aerodyn. **147**, 143–153 (2015).
[71]H. Wurps, G. Steinfeld, and S. Heinz, "Grid-resolution requirements for large-eddy simulations of the atmospheric boundary layer," Boundary-Layer Meteorol. **175**(2), 179–201 (2020).
[72]B. Jayaraman and J. G. Brasseur, "Transition in atmospheric boundary layer turbulence structure from neutral to convective, and large-scale rolls," J. Fluid Mech. **913**, A42 (2021).
[73]A. Serani, M. Diez, F. van Walree, and F. Stern, "URANS analysis of a free-running destroyer sailing in irregular stern-quartering waves at sea state 7," Ocean Eng. **237**, 109600 (2021).
[74]D. D'Agostino, M. Andre, P. Bardet, A. Serani, M. Felli, and M. Diez, "Observing piv measurements through the lens of data clustering," in *Proceedings of the 33rd Symposium on Naval Hydrodynamics*, Osaka, Japan, 18–23 October 2020.
[75]S. Zilitinkevich, I. Esau, and A. Baklanov, "Further comments on the equilibrium height of neutral and stable planetary boundary layers," Q. J. R. Meteorol. Soc. **133**(622), 265–271 (2007).
[76]J. M. Wallace, "Quadrant analysis in turbulence research: History and evolution," Annu. Rev. Fluid Mech. **48**, 131–158 (2016).
[77]S. Chowdhuri and P. K. Deb Burman, "Representation of the Reynolds stress tensor through quadrant analysis for a near-neutral atmospheric surface layer flow," Environ. Fluid Mech. **20**(1), 51–75 (2020).
[78]S. Chowdhuri and T. V. Prabha, "An evaluation of the dissimilarity in heat and momentum transport through quadrant analysis for an unstable atmospheric surface layer flow," Environ. Fluid Mech. **19**(2), 513–542 (2019).
[79]F. Alcántara-Ávila and S. Hoyas, "Direct numerical simulation of thermal channel flow for medium–high Prandtl numbers up to $Re_\tau = 2000$," Int. J. Heat Mass Transfer **176**, 121412–122021 (2021).

[80]D. Li and E. Bou-Zeid, "Coherent structures and the dissimilarity of turbulent transport of momentum and scalars in the unstable atmospheric surface layer," Boundary-Layer Meteorol. **140**(2), 243–262 (2011).
[81]E. O. Nilsson, A. Rutgersson, A. S. Smedman, and P. P. Sullivan, "Convective boundary-layer structure in the presence of wind-following swell," Q. J. R. Meteorol. Soc. **138**(667), 1476–1489 (2012).
[82]B. Giacomini and M. G. Giometto, "On the suitability of second-order accurate finite-volume solvers for the simulation of atmospheric boundary layer flow," Geosci. Model Dev. **14**(3), 1409–1426 (2021).
[83]J. Fang and F. Porté-Agel, "Large-eddy simulation of very-large-scale motions in the neutrally stratified atmospheric boundary layer," Boundary-Layer Meteorol. **155**, 397–416 (2015).
[84]B. Ganapathisubramani, E. K. Longmire, and I. Marusic, "Characteristics of vortex packets in turbulent boundary layers," J. Fluid Mech. **478**, 35–46 (2003).
[85]A. Lozano-Durán, O. Flores, and J. Jiménez, "The three-dimensional structure of momentum transfer in turbulent channels," J. Fluid Mech. **694**, 100–130 (2012).
[86]Q. Li, P. Gentine, J. P. Mellado, and K. A. McColl, "Implications of nonlocal transport and conditionally averaged statistics on Monin–Obukhov similarity theory and Townsend's attached eddy hypothesis," J. Atmos. Sci. **75**(10), 3403–3431 (2018).
[87]F. Porté-Agel, C. Meneveau, and M. B. Parlange, "A scale-dependent dynamic model for large-eddy simulation: Application to a neutral atmospheric boundary layer," J. Fluid Mech. **415**, 261–284 (2000).
[88]R. M. Alguliyev, R. M. Aliguliyev, and L. V. Sukhostat, "Efficient algorithm for big data clustering on single machine," CAAI Trans. Intell. Technol. **5**(1), 9–14 (2020).
[89]A. G. Nair, C.-A. Yeh, E. Kaiser, B. R. Noack, S. L. Brunton, and K. Taira, "Cluster-based feedback control of turbulent post-stall separated flows," J. Fluid Mech. **875**, 345–375 (2019).
[90]L. Weber, S. Gabriel, and A. G. Class, "On the statistical evaluation of bubbly flows using Voronoi cells grouped in clusters with fixed cell count," Phys. Fluids **35**(5), 053311 (2023).
[91]C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2016), pp. 2818–2826.
[92]X. Xia, C. Xu, and B. Nan, "Inception-v3 for flower classification," in *Second International Conference on Image, Vision and Computing—ICIVC 2017* (IEEE, 2017), pp. 783–787.
[93]L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M. A. Fadhel, M. Al-Amidie, and L. Farhan, "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions," J. Big Data **8**, 53 (2021).