PHYSICAL REVIEW FLUIDS **7**, 054703 (2022)

Hierarchical clustering method of volumetric vortical regions with application to the late stage of laminar-turbulent transition

Kazuo Matsuura
Graduate School of Science and Engineering, Ehime University, 790-8577, Japan

Yasuhide Fukumoto
Institute of Mathematics for Industry, Kyushu University, 819-0395, Japan

(Received 29 January 2021; accepted 11 May 2022; published 31 May 2022)

A hierarchical clustering method for volumetric vortical regions is proposed. Point clouds in vortical regions are extracted from an instantaneous flow field, using a series of parallel planes normal to a direction, and are grouped between consecutive planes. Cross-sectional distributions of points are then stacked in the direction. Each group of extracted points is identifiable as an individual entity that is referred to as a *p* cluster, with the letter *p* replaced by the terms proto, sub, super, and hyper, according to the level of clustering. The spatial distribution and temporal evolution of *p* clusters are visualized and automatically tracked. The present method is applied to the late stage of the K-type natural transition of a boundary-layer flow. Our method clearly captures, with admitting a detailed analysis, pairs of vortex structures located in the viscous sublayers near the roots of hairpin vortices. Moreover, it allows the algorithmic identification of dominant mathematical terms in the enstrophy and helicity equations, and a categorization of the extracted points based on local dynamics characterized by sets of the dominant terms. Construction of ensemble-averaged quantities using the present method, an extension of the method to omnidirections, and comparison between the present method and conventional vortex analysis methods are also discussed.

DOI: 10.1103/PhysRevFluids.7.054703

I. INTRODUCTION

Direct numerical simulation (DNS) has prevailed for understanding the mechanism of laminar-turbulent transition in boundary layers [1–4]. The transition has some stages regarding the evolution of disturbances and vortices. For example, in a natural transition that takes place under low free-stream turbulence (FST), the initial stage corresponds to the stage of linear growth of Tollmien-Schlichting waves [5]. Roughly speaking, the middle stage corresponds to the three-dimensionalization of shear layers. The late stage connects the vortex emergence in the middle stage to sustained vortex generation in the turbulent regime. The degree of freedom (DOF) of disturbance evolution in the initial stage is relatively low, and there is much consensus on its dynamics. However, consensus is still lacking on the late stage regarding the mechanism of turbulence production and analytical techniques, for the deep understanding of the mechanism is required.

Conventionally, specific vortices that are picked up nonautomatically have been focused on, and the dynamics of the vortices has been investigated through computational visualization. Although techniques to reduce the DOF of vortices such as isosurface extraction and skeletonization have been proposed, the dynamics of vortices picked up by algorithmic or automatic search has been largely unexplored.

So far, various visualization methods have been proposed for a simplified understanding of vortices [6,7]. Among these methods, one of the most sophisticated methods to compactly represent the feature of a vortex tube is the vortex core representation, or skeletonization. Singer and Banks [8] proposed a predictor-corrector scheme for a vortex skeleton. The technique uses the vorticity vector to move in the direction of the skeleton curve, and the pressure field to correct the location in the plane perpendicular to the skeleton line. Sujudi and Haimes [9] proposed an algorithm for identifying the core of a swirling flow in three-dimensional discretized vector fields on the basis of the critical point theory. The core is found as a stagnation point on a cross section of the velocity fields where the rate-of-deformation tensor has one real and a pair of complex conjugate eigenvalues. Kida and Miura [10] proposed the sectional-swirl-and-pressure-minimum scheme, in which individual vortices are visualized by its skeleton representation. To pick up swirling vortices, a swirl condition to a plane of arbitrary orientation was derived and used with the sectional-pressure-minimum condition. The vortex core was defined as the region where the swirl condition becomes negative around the vortex axis, pointing in the direction of the third eigenvector of the pressure Hessian [11]. Nakayama [12] elucidated by introducting quantities swirlity and sourcity, the topological features of vortices in terms of local flow geometry in an isotropic homogeneous turbulence and relationships between the topology and development of a vortex. In Ref. [13], it is shown that the eigen-vortical-axis line proposed as a vortical axis tends to pass through the core region with strong intensity of swirlity and vorticity and high flow symmetry. Peikert and Roth [14] proposed the parallel vectors operator as a basic tool for vector field visualization, and demonstrated various techniques for computing feature lines, such as vortex cores and extremal curves, from field data using this operator in a unified manner. Gelder and Pang [15] proposed a method called PVsolve, based on a root-finding technique. Schafhitzel *et al.* [16] proposed a vortex core line extraction method based on the λ_2 vortex region criterion. Liu *et al.* [17] and Gao *et al.* [18] proposed, by introducing a tensor Liutex, a series of third-generation vortex identification methods that can give improved vortex structures and vortex core lines that are unique and threshold-free. Zhu and Xi [19], on the basis of the work of Jeong *et al.* [20], proposed a method named vortex axis tracking by iterative propagation (VATIP) for three-dimensional vortices that propagates along the vortex axis lines and iteratively searches for new directions for growth.

The purpose of the methods mentioned above is basically to visualize the centers or the iso-surfaces of vortex structures. Although some studies [10,11] visualize the flow topology of the surrounding shear layer around a vortex core by the integral curves of flow vectors, a vortical region comprising the core and surrounding shear layer that has detailed information on its local dynamics is not extracted as a whole.

If the surrounding shear layers are included in the reduced representation of vortices, the dynamics of the vortical regions, including competition among mathematical terms which evolves the regions in time, can be understood. Numerous similar three-dimensional vortex structures are generally observed in transitional boundary layers. It is also expected that the categorization and division of vortical regions into similar groups enhance the automatic and systematic data mining of common mechanisms between vortices. Because vortical regions are represented by discrete mesh points in flow field computations, the clustering of points belonging to the regions is required for the purpose. Here, clustering is a method of unsupervised learning in the context of machine learning and uncovers a certain kind of natural structure in a given data set without assuming model answers [21,22].

Regarding the point cloud representation of vortices, Haller [23] proposed a Lagrangian coherent structure (LCS), which describes the most repelling, attracting, and shearing material surfaces identified using Lagrangian particle dynamics. By extracting such surfaces from experimental and numerical flow data, a simplified understanding of overall flow geometries, an exact quantification of material transport, and forecasting, or influencing, large-scale flow features and mixing events are conducted. Wilroy *et al.* [24], using spatially measured points, considered the clustering of vortex points and tracked the cluster in time to evaluate the circulation of a leading-edge vortex produced by an impulsively started flat plate. Zhao *et al.* [25] applied a two-step scheme combining vortex feature detection and k-means algorithm to reveal the time-resolved vortex motions on the basis of large data sets of 2D Particle Image Velocimetry (PIV) in-cylinder flow fields of a direct injection engine. Deng *et al.* [26] proposed an approach to locate vortical structures on the basis of canopy [27]+k-means. To identify vortical structures, they selected the IVD (threshold-independent vortex identification) vector [28] for vortex characterization and then normalized the IVD vector using the z-score method followed by a sigmoid function. After that, they identified vortices as clusters of the normalized IVD vector. Ferrari *et al.* [29] proposed a visualization method based on evolution surfaces and clustering to represent quasiperiodic vortex shedding over multiple shedding cycles concurrently.

Matsuura [30] proposed a clustering method of hairpin vortices that reduces the DOF of the vortices on the basis of the results of comparing successive incorporation clustering, k-means, and spectral clustering of point clouds extracted from the surrounding isosurfaces of the second invariance of the velocity gradient tensors (SIVGT) of the vortices. The method stores vortices in a list by putting a cluster identification number to each of them. The method was applied to vortices appearing in the late stage of K-type natural transition. In Ref. [31], Matsuura proposed a method that evaluates dominant local dynamics by skeletonization, mathematical term decomposition, and the recombination of a reduced number of dominant terms around the skeleton points to clarify the dynamics of hairpin vortices generated during the boundary-layer transition under FST. The method enables the automatic finding and categorization of the variations of the sets of dominant terms that govern local dynamics during the evolution of hairpin vortices.

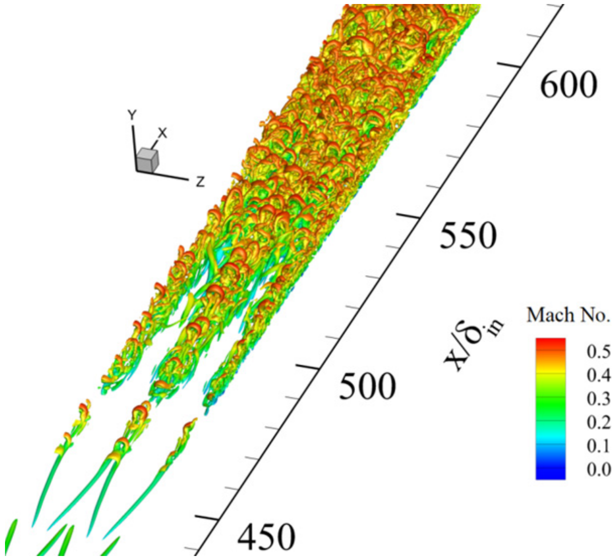
However, clustering methods that can separate three-dimensional vortical regions successfully, recognize them individually on the basis of connectivity among point clouds, and track their temporal motions in boundary layers have not been reported. In this paper, we propose a clustering method to realize such requirements on the basis of the authors' previous works on hierarchical vortex clustering [32,33].

The structure of this paper is as follows: In Secs. II and III, computational cases and computational methods are described. In Sec. IV, the proposed clustering method and p-clusters are explained. In Sec. V, results including the rationality of the present method, the spatial and temporal distribution of p clusters, the functional clustering of the elements of p-clusters, the construction of ensemble-averaged quantities, and the extension of the present clustering method to omnidirection are discussed. In Sec. VI, conclusions are presented. In Appendices A and B, the flow topology analysis of a vector field and the Helmholtz-Hodge decomposition are explained, respectively.

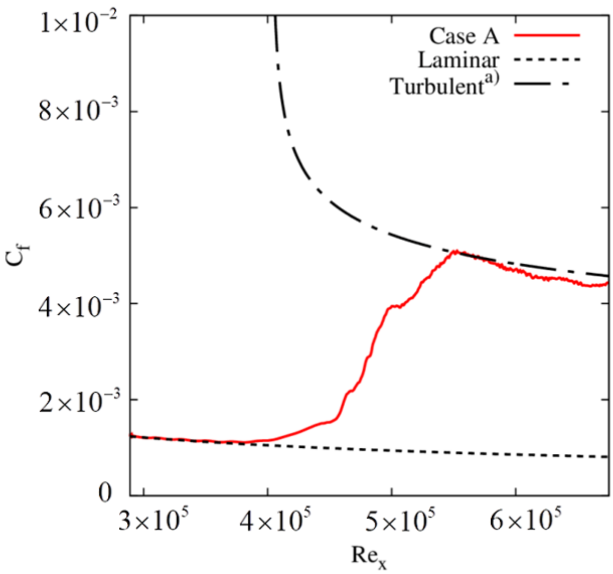
II. COMPUTATIONAL CASE

Before discussing the clustering method proposed in this study, we start with explaining flow field data to which the method is applied. The flow is boundary-layer transition of K-type without FST, which is computed by DNS explained in Sec. III. The K-type is a type of natural transition in which in-phase hairpin vortices, i.e., peak-following-peak, are observed near the beginning of transition [5]. The mean free-stream Mach number is 0.5. Disturbances comprising a two-dimensional Tollmien-Schlichting wave and a pair of oblique waves are superimposed on the Blasius solution [34].

Figure 1 shows vortex structures that appear during the laminar-turbulent transition and associated skin-friction coefficients C_f along the streamwise Reynolds number Re_x for the flow data analyzed in this paper. Compared to the DNS result of Ref. [4], the C_f of this study similarly touches the turbulent correlation around $\mathrm{Re}_x = 5.5 \times 10^5$, and therefore the present DNS is reasonable. The streamwise Reynolds number is defined as $\mathrm{Re}_x \equiv \rho_\infty u_\infty x/\mu_\infty$ using the density ρ_∞, streamwise velocity u_∞, and dynamic viscosity coefficient μ_∞ at the free stream. The vortices are visualized by the isosurface of SIVGT with $Q = 8 \times 10^{-4}$. Here, Q is nondimensionalized by the speed of sound at stagnation c_0 and δ_{in}. Here, δ_{in} is the displacement thickness at a virtual point $x = 300.79\delta_{\mathrm{in}}$. As in Fig. 1(a), numerous hairpin vortices are generated and basically aligned in the streamwise direction. Autogeneration of hairpin vortices takes place downstream of the middle region of transition. As



(a)



(b)

FIG. 1. Vortex structures that appear during laminar-turbulent transition and associated skin-friction coefficients along streamwise Reynolds number. Panels (a) and (b) show vortex structures and skin friction coefficient (C_f), respectively. (a) Vortex structures. Vortices are visualized by the isosurface of SIVGT with $Q = 8 \times 10^{-4}$. The color signifies streamwise Mach number. (b) C_f. Superscript (a) shows a correlation by Cousteix [35].

found from Fig. 1(b), the extracted subdomain for the present data analysis corresponds to the region around the maximum C_f.

Loosely speaking, the transition region can be divided into three stages, i.e., initial, middle, and late stages. The initial stage is the beginning stage of hairpin vortex formation and around

$\mathrm{Re}_x = 4.0 \times 10^5$ in Fig. 1(b). The middle stage is around the beginning of the autogeneration of secondary, tertiary, and successive hairpin vortices, and around $\mathrm{Re}_x = 5.0 \times 10^5$. The late stage is the stage for the amalgamation of hairpin vortex regions and the generation of a turbulent region, which is around $\mathrm{Re}_x = 5.5 \times 10^5$.

To investigate the dynamics of vortical regions, especially in the late stage, the subdomain $\Omega_{\mathrm{ext}} \equiv [543.5, 575.1]\delta_{\mathrm{in}} \times [0., 2.]\delta_{\mathrm{in}} \times [5.37, 16.5]\delta_{\mathrm{in}}$ is extracted. As found from part (b), the extracted subdomain for the present data analysis corresponds to the region around the maximum C_f.

III. COMPUTATIONAL METHODS

Here, the procedure of the present DNS is explained. The governing equations are the unsteady three-dimensional fully compressible Navier-Stokes equations written in general coordinates for body-fitted mesh geometries. The system of equations is closed by the perfect gas law. A constant Prandtl number of $\mathrm{Pr} = 0.72$ is assumed. The equations are solved using a sixth-order finite-difference method. Time-dependent solutions to the governing equations are obtained using the third-order explicit Runge-Kutta scheme. The present numerical method has been extensively validated for the prediction of transitional and turbulent subsonic flows [36,37].

Regarding the boundary condition, inflow profiles are specified at the inlet. Depending on the computational case, disturbances are added to a base inflow profile. The base inflow profile is obtained by solving the boundary layer equation [34]. At the outflow and upper boundaries, nonreflecting boundary conditions with a mean static pressure of $p_\infty = 101, 325$ Pa are imposed. At the wall, the nonslip, isothermal wall condition $T_\infty = 273.15$ K is imposed. Periodicity is imposed in the spanwise direction. The computational domain is a rectangular region with dimensions $416.31\delta_{\mathrm{in}}$, $80.0\delta_{\mathrm{in}}$, and $22.0\delta_{\mathrm{in}}$ in the x (streamwise), y (wall normal), and z (spanwise) directions, respectively. The thickness δ_{in} in the wall coordinate based on the maximum skin-friction coefficient at $x = 575.1\delta_{\mathrm{in}}$, is $\delta_{\mathrm{in}}^+ = 50.5$. The Reynolds number based on the free-stream quantity and the length δ_{in} is 1000. The inlet position of the computational domain is $x = 300.79\delta_{\mathrm{in}}$. The number of mesh points is $1001 \times 241 \times 128$ in the x, y, and z directions, respectively. Here, x, y, and z correspond to the streamwise, wall-normal, and spanwise directions, respectively. Mesh resolutions based on local skin-friction are $\Delta x^+ = 8 - 22$, $\Delta y^+_{\mathrm{min}} = 0.25 - 0.75$ and $\Delta z^+ = 3.5 - 9.5$. The details of the present computation are explained in Ref. [38]. The Courant-Friedrichs-Lewy number of the time integration is around 0.6. After flow reaches equilibrium, flow field data are collected for 1.5–2.0 flow-through times. The time of an instant after flow reaches equilibrium is set to $t = t_1 = 0$.

IV. CLUSTERING METHOD

This section presents the concept of the proposed clustering, its detailed procedures, the freedom of variables for point selection, the extension to three-directional clustering, the construction of ensemble-averaged quantities, and comparison between the present method and conventional vortex analysis methods.

A. Motivation

The motivation of the present paper is to enable, beyond individual human judgment, the automatic and systematic data mining of common mechanisms of the local dynamics of vortices and shear layers around vortices appearing in complex flow fields. It is expected that such mining can be realized by dividing vortical regions, which are represented by discrete mesh points, into separate but connected groups of points and categorizing them into similar groups. More specifically, the authors want to extract, categorize, retrieve, track, and visualize groups of vortical points by tying them hierarchically using an inclusion relationship such as reducing and increasing the number of degrees of freedom. Additionally, the cluster representation of connected vortical points is

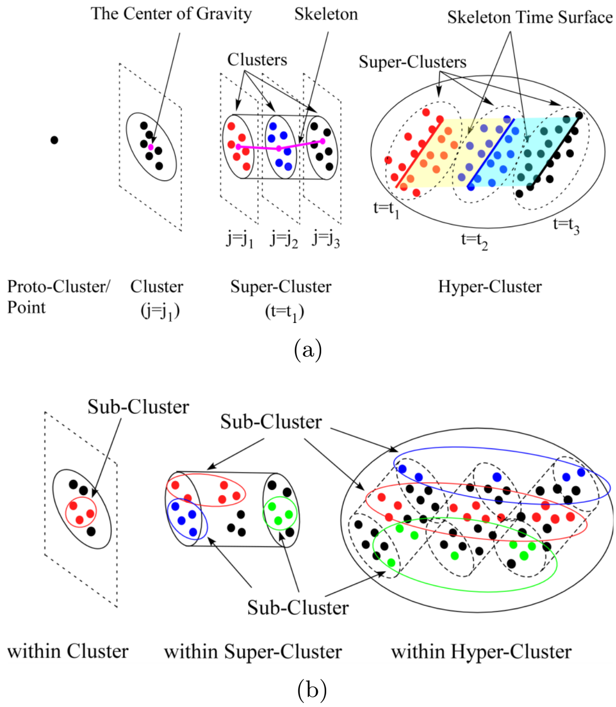


FIG. 2. Hierarchical clusters and its subsets considered in this paper. These clusters are represented in a unified manner as *p* clusters, where *p* is replaced with *proto*, *sub*, *super*, or *hyper* (a simple cluster has no prefix). (a) Fundamental clusters of various levels. Clusters are built up from left to right, and the cluster level increases from lower to upper, accordingly. (b) Subclustering within fundamental clusters.

advantageous when we investigate differences between vortex structures at two different times in the context of the stability theory of vortex shear layers.

Vortical points are divided and categorized by not only the Euclidean distances between points but also dominant mathematical terms (i.e., mathematical terms representing local dynamics) defined at each point.

Generally speaking, vortices have various orientations. However, major vortices sometimes have similar directions in boundary layers such as streamwise and spanwise vortices. Unidirectional clustering is expected to be the most straightforward approach for extracting such vortical regions. Therefore, unidirectional clustering is developed first. The unidirectional method is then extended to the clustering of points for omnidirectional vortices.

B. Concept of clustering

The concept of the present clustering is first described. In this method, mesh points belonging to vortical regions, all of which are originally indivisible, are divided into mutually separated (discrete) and individually recognizable groups (sets) of points with hierarchical properties. Here, hierarchical means that a group belongs to an extended group that includes the group and also some other groups, and the extended group belongs to a further extended group.

To realize this idea, a family of data groups called *p* clusters is introduced, as explained later in Fig. 2. Here, the purpose of *p* clusters is to extract connected strong vortical regions and treat them as discrete and enumerable groups or structures. This is not limited to the core regions of swirling flows. Neighboring shear regions around a core can also be included, depending on the criteria of extracting mesh points for *p* clusters.

The consideration of such data groups has two advantages in terms of visualization and data mining for evaluating mathematical terms:

(1) Individual connected strong vortical regions or partial blocks of the regions can be algorithmically extracted as objects with interior regions from an instantaneous field. The spatial distribution and temporal evolution of the objects can be visualized and automatically tracked, allowing the evolution of the objects to be grasped.

(2) Balances between mathematical terms that govern local dynamics can be evaluated at each point belonging to the objects. Similarity and nonsimilarity of the balances (i.e., equivalence) between the subsets of the objects can be mechanically evaluated by an algorithm, which makes the finding and retrieval of vortical regions with a common mechanism more automatic and systematic than before.

The word cluster in the term *p* cluster is used in a sense similar to that used in the context of computer science [21,22]. The aim of classical cluster analysis is to split a set of m observations (objects) into $K < m$ groups $C = \{C_1, \ldots, C_K\}$, where the ith group C_i is called a cluster. Such a division satisfies three natural requirements [22]:

(1) Each cluster ought to contain at least one object; i.e., $C_i = \varnothing,\ i = 1, \ldots, K$.

(2) Each object ought to belong to a certain cluster.

(3) Each object ought to belong to exactly one cluster; i.e., $C_{i_1} \cap C_{i_2} = \varnothing,\ i_1 \neq i_2$.

The number of objects is usually assumed in cluster analysis [22].

Similar to the case of classical cluster analysis, the separation of vortical regions is not unique, which means that the resultant individual cluster objects are different, depending on the criteria of clustering, in the present method. At present, there is no single correct definition of a vortex [39], nor are vortex structures mutually independent. Therefore, there is no single correct clustering. In this paper, a criterion of clustering associated with vortex strength is employed to make the intuitive understanding of the spatial distribution and time evolution of vortical regions simple. One intuitive criterion of successful clustering is that a three-dimensional dense point cloud is recognized as a group different from other groups of points. As an example, when we think of a vortex tube extending from the near-wall region to the boundary layer edge, the clustering is considered successful if a dense group of points belonging to the vortex tube is extracted as a group (i.e., supercluster, which is mentioned later).

On the above basis, the hierarchical clusters and their subsets shown in Fig. 2 are considered in this paper. A mesh point belonging to a strong vortical region is called a *protocluster* or simply a *point*. On a plane, closely located protoclusters with common properties are grouped into a *cluster*. The central point in a cluster is called the *center of gravity* (CG). By connecting clusters in a direction, a three-dimensional volumetric group of protoclusters called a *supercluster* is formed. A median curve connecting the CGs is called a *skeleton*. By connecting superclusters of consecutive times, a larger group of protoclusters called a *hypercluster* is made. The surface generated by the time evolution of a skeleton is called a *skeleton time surface*. The skeleton and skeleton time surface are nearly central structures of the supercluster and hypercluster, respectively. The protocluster, cluster, supercluster, and hypercluster are *fundamental clusters*. In addition to the fundamental clusters, *subclusters* that are derived from them are also considered. The subclusters play an important role in regrouping the elements of similar and non- subsets from different fundamental clusters in terms of local dynamics as discussed in Sec. V C. We represent fundamental clusters and subclusters in a unified manner as *p* clusters, where *p* is replaced with *proto*, *sub*, *super*, or *hyper* (a simple cluster has no prefix).

Figure 3 shows the overview of the present hierarchical vortex clustering. The present clustering methods consist of two methods, i.e., unidirectional clustering and omnidirectional clustering. The latter can be further divided into two methods: the method repeating the x-(uni)directional clustering algorithm in the y and z directions, which will be discussed later in Sec. IV E 1, and the method based on a vortex axis, which will be discussed later in Sec. IV E 2. The unidirectional method stacks clusters on cross sections in a predetermined direction. The omnidirectional method is an extension of the unidirectional method whereby some procedures

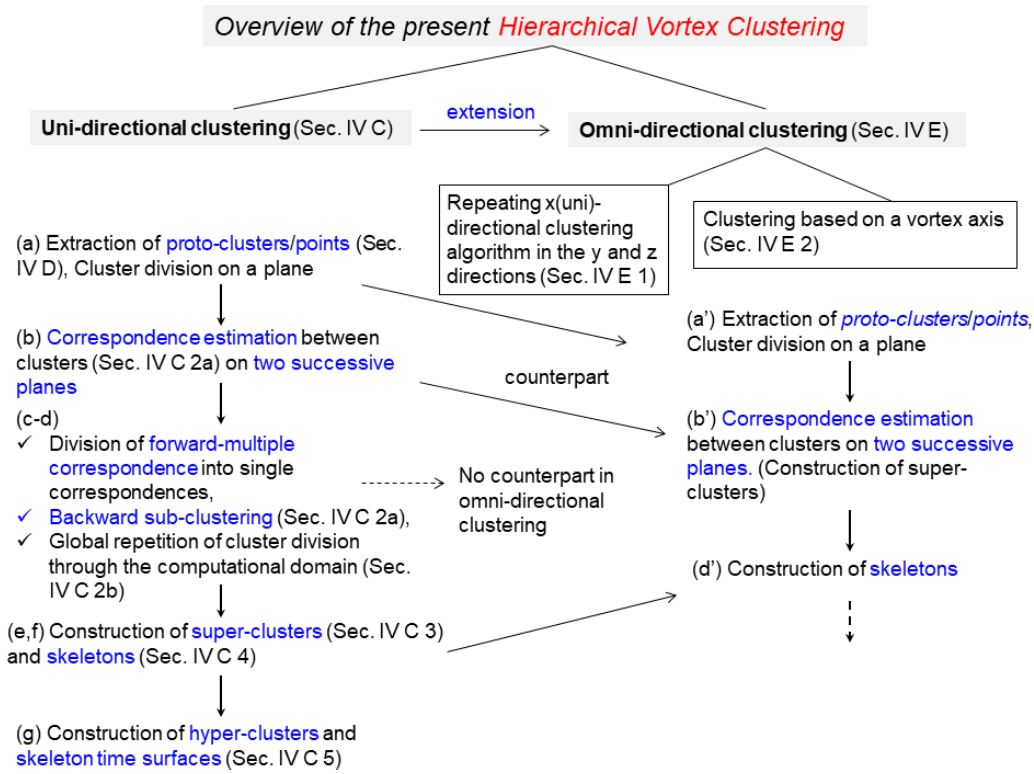


FIG. 3. Overview of the present hierarchical vortex clustering.

are extended and some are eliminated as shown, respectively, by solid and dotted arrows in Fig. 3.

C. Unidirectional clustering method

The procedure of the proposed clustering method for one direction is first outlined and the details of each step are then explained in Secs. IV C 1–IV C 5. Here, the one-directional clustering method is the most fundamental and main clustering method in this paper. The method is extended to omnidirectional methods in Sec. IV E.

The data to be analyzed are originally stored in the three-dimensional structured mesh used for the flow computation, i.e.,

$$\{(x_j, y_k, z_l); j = 1, \ldots, j\text{max}, k = 1, \ldots, k\text{max}, l = 1, \ldots, l\text{max}\}. \tag{1}$$

The proposed method consists of the following steps:

(a) Protoclusters or points. On the jth x plane, where $j = 1, \ldots, j\text{max}$, regions with large-magnitude streamwise vorticity ω_x are divided into several clusters. We call the mesh points constituting the clusters *protoclusters* or *points*. This step will be explained in Sec. IV C 1. Although the choice of a variable in extracting protoclusters is free, in principle, the procedures of the unidirectional method are explained here using ω_x. The extended choice will be explained later in Sec. IV D.

(b) Forward multiple correspondence. Correspondence between each cluster on two successive planes [(i.e., jth and $(j+1)$th] is evaluated.

The correspondence between clusters on successive x planes is not necessarily one-to-one. A single cluster on the jth x plane can have correspondence with multiple clusters on the $(j+1)$th x plane. We refer to this situation as *forward multiple correspondence*.

(c) Backward subclustering. Any forward multiple correspondence is decomposed into *single* correspondences by dividing the single cluster on the jth x plane into multiple *subclusters*, each of which has a one-to-one correspondence with each cluster on the $(j+1)$th x plane. We call this step *backward subclustering*, and explain it in Sec. IV C 2.

Although intercluster correspondences between the jth and $(j+1)$th x planes are decomposed into single correspondences by the above procedure, this may introduce new multiple correspondences between the (j-1)th and jth x planes, even if the original correspondences between these planes are one-to-one.

(d) Global division. The division of the jth x plane is repeated for $j = 1, \dots, j\text{max}$, and the global repetition of the division through the computational domain is iterated until all correspondences between clusters on any two successive x planes become single. This step will be explained in Sec. IV C 2.

(e) Superclusters. By connecting highly correlated clusters in the streamwise direction, a group of highly correlated clusters over multiple successive x planes is generated. These are referred to as *superclusters*. This step will be explained in Sec. IV C 3.

(f) Skeleton. A supercluster is a set of point elements. To reduce the DOFs of each cluster, a median curve, i.e., the curve connecting the centers of gravity of a cluster on each x plane, can be drawn. We call this curve the *skeleton* of a supercluster. This step is explained in Sec. IV C 4.

Steps (a)–(f) extract only regions with ω_x. Because streamwise vortices are dominant in the near-wall regions of boundary layers [40] and are one of the key quantities in the production of coherent structures in near-wall turbulence [41–44], the present method is effective in elucidating the vortex dynamics in the late-stage boundary layer transition that is the focus of this paper. The preponderance of elongated structures with strong ω_x is also supported by the distribution of ω_x, as discussed in Sec. V A.

Thus far, steps (a)–(f) have been conducted for instantaneous flow field data.

(g) Hyperclusters, skeleton time surface. To investigate the temporal behavior of superclusters, time accumulation of superclusters and surfaces swept by the movement of skeletons, which we call *hyperclusters* and *skeleton time surface*, are also constructed. This step is explained in Sec. IV C 5.

The hierarchy of clusters considered in this paper runs as follows: *proto*cluster, *sub*cluster, cluster, *super*cluster, and *hyper*cluster. As mentioned previously, these clusters are represented in a unified manner as p clusters, where p is replaced with *proto*, *sub*, *super*, or *hyper* (a simple cluster has no prefix). Each p cluster has the plus (minus) signature of ω_x, which is *inherited* from that of the *proto*cluster. *p-cluster element* signifies a protocluster (point) belonging to the p cluster throughout the text.

A p cluster is different from a skeleton or core [8–10,14–16], which shows only the central curve of a vortex structure, and also includes shear regions around the skeleton with clear boundaries defined by the threshold value of ω_x used to extract the vortical regions.

1. Cluster

First, mesh points belonging to the regions of large magnitude of ω_x, i.e.,

$$\{|\omega_{x,j,k,l}| > \omega_{x,0};\ j = 1, \dots, j\text{max}, k = 1, \dots, k\text{max}, l = 1, \dots, l\text{max}\}, \tag{2}$$

are extracted on each x plane, i.e., $\{(x_j, y_k, z_l); k = 1, \dots, k\text{max}, l = 1, \dots, l\text{max}\}$ for $j = 1, \dots, j\text{max}$. $\omega_{x,0}$ is a constant threshold value used for extracting vortical regions. ω_x and $\omega_{x,0}$ are dimensionless vorticities in which the velocities and coordinates are scaled by c_0 and δ_{in}, respectively. $\omega_{x,0} = 0.1$ is used in this paper. The process of choosing the value is explained later in this subsection. On each x plane, the protoclusters are divided into clusters by considering the Euclidean distance between them. Regarding the present cluster model, each cluster has a positive or

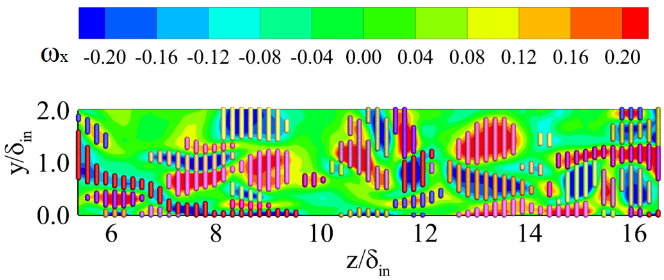


FIG. 4. Clusters generated on the plane $x = 558.9\delta_{\text{in}}$. Regions of $|\omega_x| > 0.1$ and resultant clusters are shown. The surface contour is ω_x, and the points are protoclusters. Different clusters are shown in different colors. The limitation of the number of recognizable colors means that some different clusters are shown in the same color.

negative signature of ω_x, which is inherited by upper p clusters as mentioned previously. Hereafter, a point (x, y, z) is also represented by a vector $\mathbf{x}$ written in boldface. Subscripts are also applied to $\mathbf{x}$ to assign a number to the vector. Although some other quantities can be used to extract mesh points, ω_x is used here because it is the most straightforward quantity for identifying streamwise vortices as mentioned previously.

Initial clustering. Initially, for a point $\mathbf{x} = \mathbf{x}_1$, points $\mathbf{x}_i$, $i \in \{1, \ldots, k\max * l\max\}$ that satisfy $d_p(\mathbf{x}_1, \mathbf{x}_i) < \epsilon$ are categorized into cluster C_1. Here, $d_p(\mathbf{a}, \mathbf{b})$ denotes the Euclidean distance between points $\mathbf{a}$ and $\mathbf{b}$, and we set $\epsilon = 0.2\delta_{\text{in}}$. $I_1 = \{1, \alpha, \beta, \gamma, \ldots\}$ is considered as the set of subscripts of $\mathbf{x}_i$ belonging to cluster C_1. For $m = \min\{\{1, \ldots, k\max * l\max\} \setminus \bigcup_n I_n\}$, the points $\mathbf{x}_i$, $i \in \{1, \ldots, k\max * l\max\}$ that satisfy $d_p(\mathbf{x}_m, \mathbf{x}_i) < \epsilon$ are then grouped into cluster C_m. Each cluster C_q has tags indicating its cluster number q, the total number of elements $n\max$, the set of elements $\mathbf{X}_m = \{\mathbf{x}_{q_1}, \mathbf{x}_{q_2}, \ldots, \mathbf{x}_{q_{n\max-1}}, \mathbf{x}_{q_{n\max}}\}$, and the signature of the vortical region.

Merging. Next, the initial clusters are modified by merging clusters close to each other, i.e., if $d_c(C_q^{\text{old}}, C_r^{\text{old}}) < \epsilon$, $q < r$, then $C_q^{\text{new}} = C_q^{\text{old}} \cup C_r^{\text{old}}$. Here, the superscripts old and new indicate the clusters that exist before and after the merging and updating operations, respectively. $d_c(A, B)$ is the distance between clusters A and B, defined as $d_c(A, B) \equiv \min\{d_p(\mathbf{x}, \mathbf{y}); \mathbf{x} \in A, \mathbf{y} \in B\}$.

Empty cluster removal. After transferring the elements of C_r^{old} to C_q^{new}, C_r^{old} becomes empty. All indices greater than r are shifted down to remove the empty cluster, and the clusters are renumbered, i.e., $C_{r'}^{\text{old}}$, $r' > r$ becomes $C_{r'-1}^{\text{new}}$ with No. $C_{r'-1}^{\text{new}} =$ No. C_r^{old}. Here, No. denotes the total number of elements in a cluster (set). These merging and renumbering operations are repeated for all pairs of clusters until no further modification is necessary.

Figure 4 shows the resultant clusters and the surface contour of ω_x on the plane of $x = 558.9\delta_{\text{in}}$. The points are protoclusters. The vortical regions are successfully divided into different clusters, demonstrating that the proposed algorithm works correctly.

As there is no single right definition of a vortex [39], there are no nonempirical criteria for selecting optimum $\omega_{x,0}$. First, $\omega_{x,0} > 0$ is chosen within $[0, \max(|\omega_{x,\min}|, |\omega_{x,\max}|)]$. Here, $\omega_{x,\min}, \omega_{x,\max}$ are the minimum and maximum streamwise vortices within the domain of the late stage, respectively. The selection of $\omega_{x,0}$ has a trade-off between connectedness and separability of the extracted regions. When $\omega_{x,0}$ is increased, the number of protoclusters extracted becomes smaller, and only connected regions of large ω_x persist. On the other hand, if $\omega_{x,0}$ is decreased, more connected regions are extracted and the separability of the connected region is reduced.

In Fig. 5, the histogram of the number of mesh points at which ω_x belongs to each divided range is shown. The red line shows $\omega_x = \pm 0.1$ which is used for $\omega_{x,0}$ in this paper. The ratio of the number of extracted protoclusters to that of the total mesh points is about 50% when this value of $\omega_{x,0}$ is used, and a reasonable clustering is achieved from the results shown in Sec. V.

$\omega_{x,0}$ is varied as 0.03, 0.1, and 0.19 to show the sensitivity of the present analysis to $\omega_{x,0}$. The ratios of the number of extracted protoclusters to that of all the mesh points are approximately 88%,

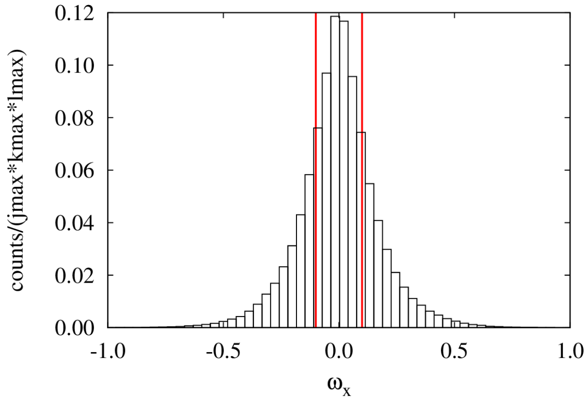


FIG. 5. Histogram of the number of mesh points at which ω_x belongs to each divided range at $t = t_1$. Red lines: $\omega_x = \pm 0.1$.

50%, and 19.4%, respectively. The resultant skeletons are shown in Fig. 6. The numbers of positive and negative superclusters are (74, 72), (62, 73), and (39, 43), respectively. Generally speaking, the number of resultant clusters increases (decreases) because the number of extracted points increases (decreases) as $\omega_{x,0}$ decreases (increases). The present method is robust in generating skeletons in terms of maintaining the lifted distributions of continuous points and directly controlling, through the variation of $\omega_{x,0}$, the number of resultant skeletons (i.e., the resolution of skeletons) while maintaining the inclusion (hierarchical) relationship of the extracted protoclusters.

Regarding the clustering, clustering algorithms such as k-means or spectral clustering can also be considered [21,22,45]. However, because the number of resultant clusters is specified at the outset in these algorithms and is not suitable for the present purpose, the clustering method mentioned above is developed and is used in this paper.

2. *Backward subclustering*

Correspondence between two successive planes. To begin, correspondence between two clusters on two successive planes is judged. The *degree of projected overlap* (DPO) in an x plane is introduced for the purpose as follows: Two clusters C_i and C_{α_l}, respectively, on the jth and $(j+1)$th x planes are considered. Then, the DPO between them is defined by

$$\mathrm{DPO} \equiv \frac{\mathrm{No.}\,\big(\mathrm{Prj}_x(C_i) \cap \mathrm{Prj}_x(C_{\alpha_l})\big)}{\min\big(\mathrm{No.}\ C_i,\, \mathrm{No.}\ C_{\alpha_l}\big)}. \tag{3}$$

Here, $\mathrm{Prj}_x(C_i)$ and $\mathrm{Prj}_x(C_{\alpha_l})$, respectively, denote the orthogonal projection of the elements of C_i and C_{α_l} to the x plane at $x = 0$. If DPO $\geqslant 0.5$, it is considered that C_i and C_{α_l} have correspondence.

When correspondences among clusters between the jth and $(j+1)$th x planes are evaluated, a cluster C_i on the jth x plane can have multiple correspondences with clusters on the $(j+1)$th x plane, e.g., $C_{\alpha_1}, C_{\alpha_2}, \ldots, C_{\alpha_M}$. Here, the division of C_i proceeds by comparing the Euclidean distances $d_m, m = \alpha_1, \alpha_2, \ldots, \alpha_M$, defined by $d_m \equiv \min_r\{d_p(\mathbf{x}_q, \mathbf{x}_r); \mathbf{x}_r \in C_m\}$, for each $\mathbf{x}_q \in C_i$. The element $\mathbf{x}_k$ is considered to have a correspondence with the closest cluster on the $(j+1)$th x-plane. These comparisons are repeated for all elements in C_i.

Figure 7 presents a schematic explanation of the concept of backward subclustering for the case M=2. Because the jth x plane is divided on the basis of the $(j+1)$th x plane, this process is considered to be backward. The cluster C_i^{old} is then divided into $C_{i,1}^{\mathrm{new}}, C_{i,2}^{\mathrm{new}}, \ldots, C_{i,M}^{\mathrm{new}}$, depending on the correspondences with $C_{\alpha_1}, C_{\alpha_2}, \ldots, C_{\alpha_M}$, respectively. Through this backward subclustering procedure, the multiple correspondences are decomposed into a series of single correspondences

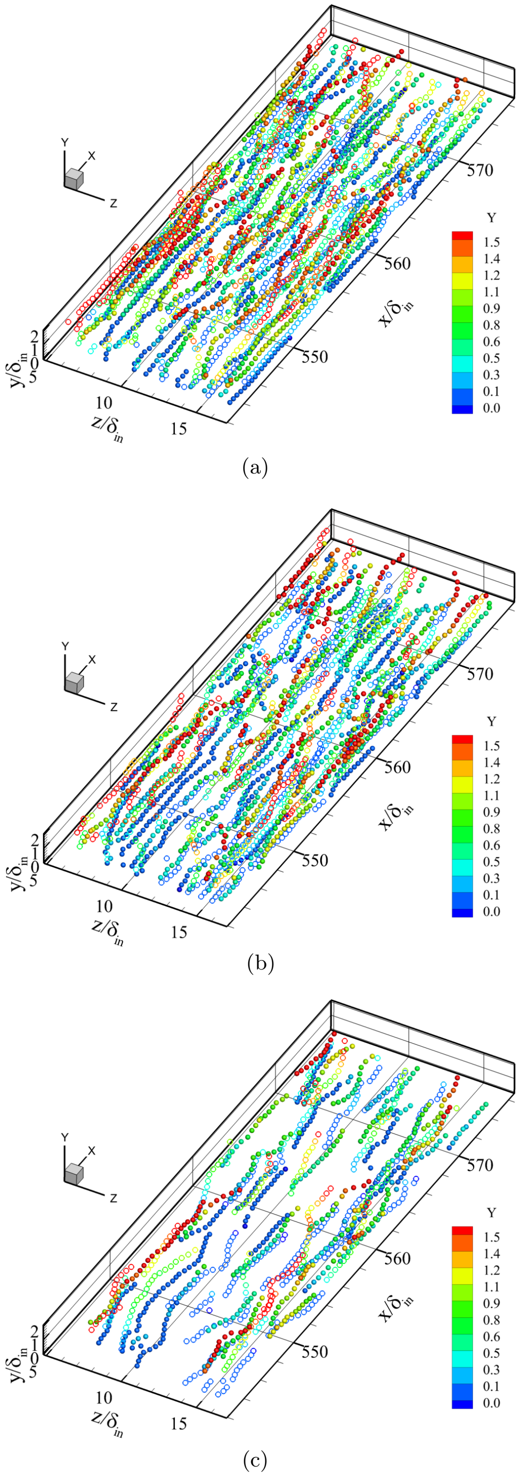


FIG. 6. Effects of varying $\omega_{x,0}$ on the resultant distribution of skeletons at $t = t_1$. Open and filled symbols denote the skeletons of the vortical regions with positive and negative ω_x, respectively. The color indicates the height from the wall. (a) $\omega_{x,0} = 0.03$, (b) $\omega_{x,0} = 0.1$, (c) $\omega_{x,0} = 0.19$.

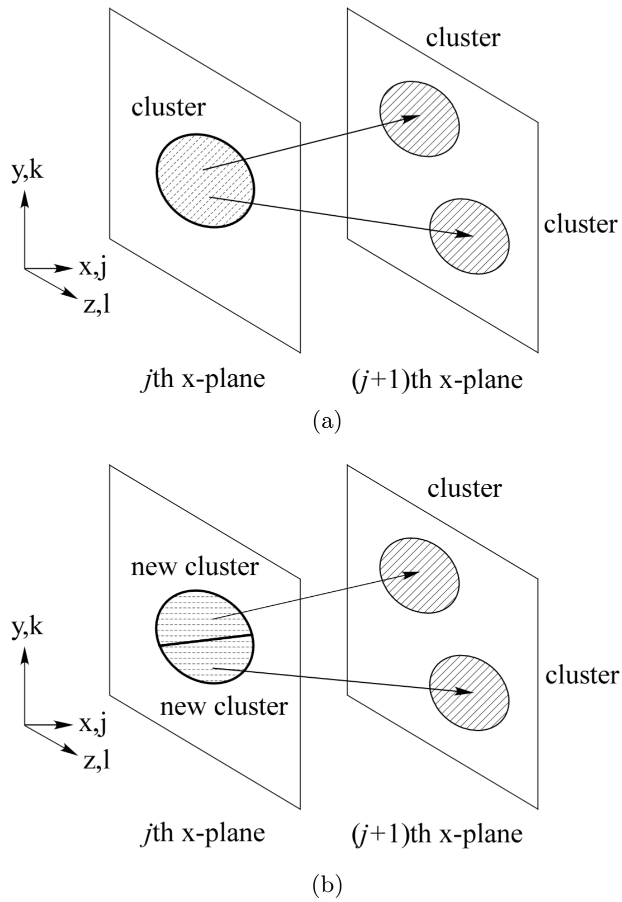


FIG. 7. Backward subclustering. (a) Before backward division. (b) After backward division.

such as $C_{i,1}^{\text{new}} \sim C_{\alpha_1}, C_{i,2}^{\text{new}} \sim C_{\alpha_2}, \ldots, C_{i,M}^{\text{new}} \sim C_{\alpha_M}$ between the jth x plane and the $(j+1)$th x plane. Here, the symbol $\sim$ indicates a single correspondence.

The reason for conducting this backward subclustering is to refine the correspondences over two successive x planes in the case of multiple correspondences. This process clarifies the upstream (starting) ends of the superclusters.

Iteration through the computational domain. In Sec. IV C 2, forward multiple correspondences were decomposed into a series of single correspondences between the jth and $(j+1)$th x planes. However, if we then think of the correspondences between the (j-1)th and jth x planes, the previously generated clusters $C_{i,1}, C_{i,2}, \ldots, C_{i,M}$ introduce new forward multiple correspondences between these x planes, even if there were originally no forward multiple correspondences between them. Figure 8 shows the mechanism of the backward propagation of multiple correspondences and further backward subclustering.

To resolve the problem of these new forward multiple correspondences and achieve single correspondences over the whole domain, the backward subclustering procedure is iterated until single correspondence is attained through the whole domain, i.e., for $j = 1, \ldots, j\text{max} - 1$. Because the number of iterations required depends on the flow structures and is unknown in advance, the process terminates after a sufficiently large number of iterations or according to some algorithmic criterion.

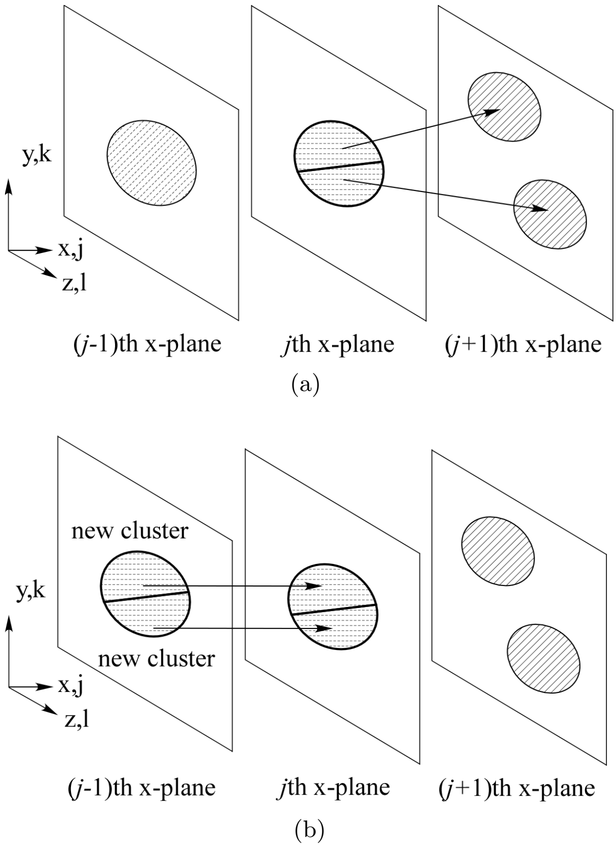


FIG. 8. Backward propagation of multiple correspondence. (a) Backward subclustering on the jth x plane according to the forward multiple correspondences with the $(j+1)$th x plane. (b) Emergence of new forward multiple correspondences between the $(j$-$1)$th and jth x planes, and backward subclustering on the $(j$-$1)$th x plane.

3. Supercluster

By connecting highly correlated clusters extending over successive x planes in the streamwise direction, extended clusters that have a streamwise dimension are constructed. We call these superclusters. By identifying such connectivity using the depth-first principle [46], all superclusters can be found.

4. Skeleton

Because a supercluster consists of numerous elements, a reduced-order model representation is sometimes convenient for understanding its dynamics.

By weight averaging the positions of elements within a cluster on an x plane, the weight-averaged gravity center of the cluster $\overline{\mathbf{x}}^W$ can be computed as follows:

$$\overline{\mathbf{x}} \equiv \sum_{r=1}^{\text{No. } C_p} \mathbf{x}_r / \text{No. } C_p, \mathbf{x}_r \in C_p, \tag{4}$$

$$\overline{\mathbf{x}}^W \equiv \sum_{r=1}^{\text{No. } C_p} \mathbf{x}_r \theta(|\mathbf{x}_r - \overline{\mathbf{x}}|) / \sum_{r=1}^{\text{No. } C_p} \theta(|\mathbf{x}_r - \overline{\mathbf{x}}|), \tag{5}$$

$$\theta(r) \equiv \exp(-r^2/(h/2)^2). \tag{6}$$

In this paper, $h = 0.5\delta_{\mathrm{in}}$. By connecting the weight-averaged gravity centers through successive x planes in the streamwise direction within a supercluster, the median curve, i.e., the skeleton of a supercluster, is obtained.

5. Hypercluster and skeleton time surface

To understand the temporal behavior of superclusters, time accumulation of superclusters and the surface swept by each skeleton are considered. We call them a hypercluster and a skeleton time surface, respectively. Especially, the skeleton time surface is convenient for simplified visualization because they show the translation, angular rotation, and rate of shear deformation of a vortical region. It is not necessarily simply connected, as found from the construction procedure described in this section.

The correspondence between skeletons is evaluated at successive points in time, and if the skeleton of a supercluster S_B at $t = t_2$ is the result of movement from the skeleton of a supercluster S_A at $t = t_1$, S_A and S_B are connected by surface elements. Here, the difference between successive points in time, i.e., $t_2 - t_1$, is assumed to be sufficiently small. All skeleton time surfaces are obtained at successive times in the domain of interest by the same procedure, which is explained below. In Fig. 9, some intermediate steps in the construction of a skeleton time surface are illustrated schematically.

Algorithm for Constructing Skeleton Time Surfaces

Step 1: Smooth interpolation of the centers of gravity. First, the centers of gravity on successive cross sections in the streamwise direction (which generate the skeleton of a supercluster) are interpolated by a parametric cubic spline to make the resultant hypercluster as smooth as possible. For each smooth skeleton-generating set (SSGS) of spline-interpolated points $\tilde{C}^{\sigma}_{sk,q,t=t_1} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_1}\}$ at $t = t_1$, the corresponding SSGS of spline-interpolated points $\tilde{C}^{\sigma}_{sk,r,t=t_2} = \{\mathbf{x}'_1, \mathbf{x}'_2, \ldots, \mathbf{x}'_{N_2}\}$ is sought at $t = t_2$. Here, the superscript σ explicitly indicates the positive (negative) sign of ω_x. The subscripts q and r denote the identification numbers of the SSGS at $t = t_1$ and t_2, respectively.

Step 2: Finding corresponding skeletons. The corresponding skeleton is sought from candidate skeletons constructed from ω_x that has the same sign. The corresponding SSGS is selected by the following algorithm: For the point $\mathbf{x}_1 \in \tilde{C}^{\sigma}_{sk,q,t=t_1}$, the closest point $\mathbf{x}'_\alpha$ in a candidate SSGS $\tilde{C}^{\sigma}_{sk,r,t=t_2}$ and the distance between $\mathbf{x}_1$ and $\mathbf{x}'_\alpha$, which is denoted as d_1, are found. The same operation is then repeated for the point $\mathbf{x}_2 \in \tilde{C}^{\sigma}_{sk,q,t=t_1}$, and $\mathbf{x}'_\beta \in \tilde{C}^{\sigma}_{sk,r,t=t_2}$ and d_2 are found. The closest points and their distances are then found for all other $\mathbf{x}_i \in \tilde{C}^{\sigma}_{sk,q,t=t_1}$, $i = 3, \ldots, N_1$. This process is shown in Fig. 9(a). Using d_i, $i = 1, \ldots, N_1$, the average distance $\overline{d}$ from the points in $\tilde{C}^{\sigma}_{sk,q,t=t_1}$ to those in $\tilde{C}^{\sigma}_{sk,r,t=t_2}$ is defined as follows:

$$\overline{d} \equiv \sum_{p=1}^{N_1} d_p/\mathrm{No.}\ \tilde{C}^{\sigma}_{sk,q,t=t_1}. \tag{7}$$

The average distance $\overline{d}$ is not always small because superclusters may change discontinuously with time. To exclude nonphysical distances ascribable to the superclustering algorithm, only distances of less than $\overline{d}$ are selectively averaged,

$$\bar{\bar{d}}_{q\to r} \equiv \sum_{p\in M} d_p \mathrm{No.}\ M, M \equiv \{\alpha; 0 < d_\alpha < \bar{d}\}, \tag{8}$$

where the subscript $q \to r$ indicates the distance from SSGS q at $t = t_1$ to candidate SSGS r at $t = t_2$. Suppose that, at $t = t_2$, the candidate SSGSs are $O = \{r_1, r_2, \ldots, r_q\}$. The closest SSGS

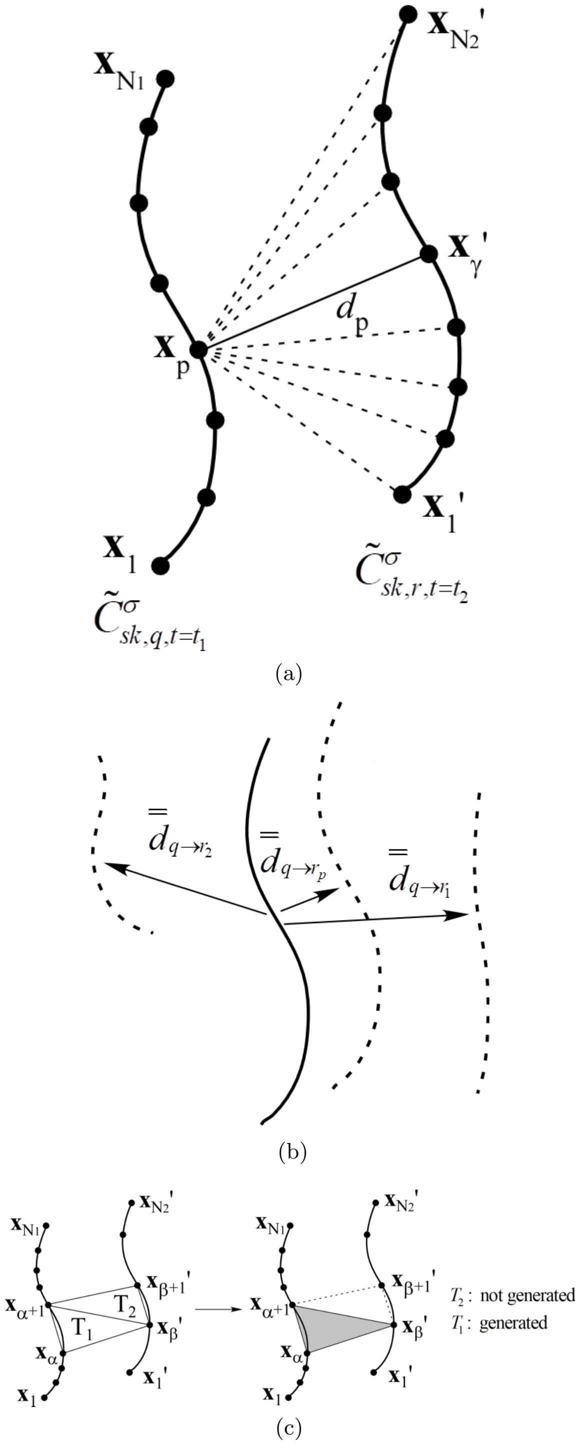


FIG. 9. Schematic explanation of the construction of a skeleton time surface. (a) Finding the closest point to $\mathbf{x}_p$, (b) Selection of the closest SSGS from candidate SSGSs. Solid and dotted curves represent SSGSs at $t = t_1$ and t_2, respectively, (c) Generation of a surface element satisfying the admissible condition from the candidate triangles among $\{\mathbf{x}_\alpha, \mathbf{x}_{\alpha+1}, \mathbf{x}'_\beta, \mathbf{x}'_{\beta+1}\}$. The gray element is the generated surface element T_1.

is selected as $\tilde{C}^{\sigma}_{sk,r_p,t=t_2}$ such that $\bar{\bar{d}}_{q\to r_p} = \min\{\bar{\bar{d}}_{q\to r_1}, \bar{\bar{d}}_{q\to r_2}, \ldots, \bar{\bar{d}}_{q\to r_q}\}$. This process is shown in Fig. 9(b). The same procedures are repeated for all other SSGSs at $t = t_1$, and the closest SSGSs at $t = t_2$ are identified.

Step 3: Surface generation. After finding the closest SSGSs, surface elements are generated between SSGS pairs. Here, the generation of surface elements among $\{\mathbf{x}_\alpha, \mathbf{x}_{\alpha+1}, \mathbf{x}'_\beta, \mathbf{x}'_{\beta+1}\}$ is performed using two triangular elements $T_1 \equiv \text{triag}\{\mathbf{x}_\alpha, \mathbf{x}_{\alpha+1}, \mathbf{x}'_\beta\}$ and $T_2 \equiv \text{triag}\{\mathbf{x}_{\alpha+1}, \mathbf{x}'_\beta, \mathbf{x}'_{\beta+1}\}$. An admissible condition that the inter-SSGS edge is shorter than the local convection distance is imposed to generate a triangular element. In mathematical notation, element T_1 is generated if

$$d_p(\mathbf{x}_\alpha, \mathbf{x}'_\beta) < |\mathbf{u}_\alpha|\Delta t, \quad d_p(\mathbf{x}_{\alpha+1}, \mathbf{x}'_\beta) < |\mathbf{u}_{\alpha+1}|\Delta t, \tag{9}$$

and element T_2 is generated if

$$d_p(\mathbf{x}_{\alpha+1}, \mathbf{x}'_\beta) < |\mathbf{u}_{\alpha+1}|\Delta t, \quad d_p(\mathbf{x}_{\alpha+1}, \mathbf{x}'_{\beta+1}) < |\mathbf{u}_{\alpha+1}|\Delta t, \tag{10}$$

where $\Delta t = t_2 - t_1$ and $\mathbf{u}_\alpha$ is the convection velocity at $\mathbf{x}_\alpha$ when $t = t_1$. This process is shown in Fig. 9(c).

In the field of flow visualization, the surfaces of streamlines and streaklines, which are similar to the present surface-tiling problem, can be accurately represented [47–51] because those lines always have one-to-one corresponding counterparts after a small time step. However, the present skeletons do not necessarily have such counterparts because superclusters may change discontinuously depending on instantaneous flow fields. Consequently, the surface elements do not satisfy the admissible condition at some locations, and can be disconnected.

A hypercluster is formed when the supercluster of a skeleton has a counterpart.

D. Freedom of variable choice for extracting protoclusters

Various quantities can be considered for extracting protoclusters associated with vortical regions. Because the present paper aims at also extracting neighboring shear layers of vortices as mentioned previously in the Introduction, variables for extracting protoclusters should not be limited to those for extracting only the central regions (core) of vortices, resulting in the present method being different from and broader than conventional vortex identification.

In previous vortex identification, efforts are made to achieve objectivity (frame) invariance, remove vorticity due to shear flows, and avoid the use of an arbitrary threshold [39]. These are not necessarily pursued strictly in the present paper.

Candidate criteria for extracting protoclusters are (i) the Δ criterion (i.e., the velocity gradient tensor $\nabla\mathbf{u} \equiv (\mathbf{u})_{ij} = \partial u_i/\partial x_j$ has one real eigenvalue and two complex conjugate eigenvalues), (ii) the isosurfaces of the second invariance of $\nabla\mathbf{u}$, and (iii) the vorticity magnitude $||\omega||$ or its component is greater than a certain threshold.

Criterion (i) means that flow swirls around the axis pointing in the direction of the eigenvector corresponding to the real eigenvalue. This criterion is used in Sec. IV E 2. In criterion (ii), the second invariant Q of $\nabla\mathbf{u}$ is $Q = \frac{1}{2}(||\mathbf{\Omega}||^2 - ||\mathbf{S}||^2)$ for incompressible flows. Here, $\mathbf{\Omega}$ and $\mathbf{S}$ are, respectively, the vorticity and strain tensors and $\nabla\mathbf{u} = \mathbf{\Omega} + \mathbf{S}$. Q is a measure of the local rotation rate in excess of the strain rate. The region inside the isosurface $Q = Q_0$ corresponds to the locations where $Q \geqslant Q_0$; i.e., the local rotation rate is much greater than that at the isosurface. The extraction of points using this criterion was adopted in Refs. [30,31]. Criterion (iii) includes vorticity due to shear flows in addition to a pure vortical region with a swirl that is considered as a vortex in a narrow sense.

The inclusion is an advantage of criterion (iii) because capturing neighboring shear layers of a vortex in addition to the vortex core is part of the present objective. As mentioned previously, the threshold of ω_x can be used as the lower limit for extracting protoclusters in Sec. IV C 1, and the criterion has a desirable property that the size of the vortical regions to be analyzed can be controlled by retaining strong vortical regions. This control introduces hierarchical reduction and extension in the present clustering methods different from the variation in the p clusters as mentioned previously in Sec. IV C 1. In addition, because vortical regions always have nonzero vorticity, the criterion can extract wide regions including vortex candidates.

E. Extension to omnidirectional clustering

So far, unidirectional clustering has been discussed. Because of the preponderance of vortical regions with ω_x, the method is effective in extracting strong vortical regions connected by ω_x. Meanwhile, there are vortical structures that have dominant vorticity components other than ω_x and a more generalized clustering method is demanded. In this subsection, we demonstrate that p clusters can be constructed from more general three-dimensional vortical regions as shown in Fig. 3. It is shown that the unidirectional clustering algorithm explained in Sec. IV C can be extended to omnidirectional clustering in two ways. One way is to repeat the application of the x-directional algorithm in the y and z directions. The other way is to evaluate the direction of a vortical axis at each protocluster and project surrounding protoclusters on a plane normal to the axis. Both ways are explained below. In both ways, the unidirectional algorithm plays a fundamental role. The effectiveness of the above two ways is discussed in Sec. V A.

1. Repeating x-directional clustering algorithm in the y and z directions

The extension of the x-directional clustering algorithm explained in Sec. IV C to the y and z directions is straightforward; i.e., the application of the x-directional algorithm is simply repeated for the y and z directions. An extended p-cluster considering multiple directions can be constructed by merging superclusters overlapping with each other. From the authors' experience, reasonable extended superclusters can be constructed by merging superclusters obtained from x- and z-directional clustering. However, merging superclusters from all x-, y-, and z-directional clustering results in the connection of almost all mesh points, which does not allow the construction of meaningful extended superclusters.

2. Clustering based on a vortex axis

The x-directional clustering algorithm explained in Sec. IV C has the following four steps: (a) protoclusters are extracted on the basis of ω_x. (b) Correspondence between any two clusters on two successive planes in the x direction is evaluated. (c) Any forward multiple correspondence is decomposed into single correspondences by the backward subclustering. (d) A series of the CGs within clusters on the x planes is connected to form a skeleton.

Four alternative steps are adopted to develop an omnidirectional method. (a′) protoclusters are extracted using a method that does not assume a specific direction. (b′) The original step (b) is used. (c′) The backward subclustering is rejected and multiple correspondence is allowed because a specific direction is not assumed in the present omnidirectional algorithm. (d′) A series of CGs within clusters on planes normal to vortex axes are used for constructing a skeleton. Details of the procedures are as follows: (a′) protoclusters are extracted using criterion (i), (ii), or (iii) mentioned in Sec. IV D. A unit vector pointing in the direction of a vortex is defined at each point. As an example, if criterion (i) is adopted, a unit vector pointing in the direction of the eigenvector corresponding to the real eigenvalue, which has a positive inner product with the vorticity vector, is defined at each point. (b′) Correspondence between each cluster on two successive x planes is evaluated. Here, using the x planes does not lose the generality of the method for constructing superclusters because the accumulation of the x planes restores original vortical regions, which is also shown in Sec. V A. Mutually separated superclusters are constructed from links between clusters. (d′) A protocluster in a supercluster is selected. Using the vector averaged within a spherical neighborhood with the radius d_r around the protocluster, a plane element that is perpendicular to the vector is generated. The size of the element is arbitrary as long as the cross section of the original vortical region with the plane is fully included in the plane element. Because the protoclusters extracted in step (a) do not necessarily lie on the plane, surrounding protoclusters located within a distance d_n from the plane are

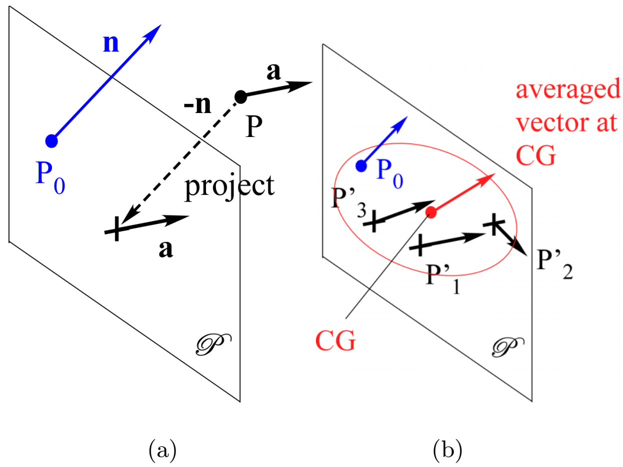


FIG. 10. Clustering on the plane going through point $\mathbf{P_0}$. (a) Projection of point P to plane $\mathscr{P}$ going through the selected point $\mathbf{P_0}$ and having the normal vector $\mathbf{n}$. Here, $\mathbf{n}$ is the vector averaged around point $\mathbf{P_0}$. (b) Clustering of points P_1', P_2', and P_3' on plane $\mathscr{P}$, the CG, and the averaged vector at the CG.

projected onto the plane in the direction of the normal vector. The projected protoclusters are divided into clusters adopting the method mentioned in Sec. IV C 1. The CG of the cluster possessing the originally selected protocluster is kept for constructing a skeleton for the supercluster. The above procedure is schematically explained in Fig. 10. Figure 10(a) shows the projection and Fig. 10(b) shows the clustering on the plane, the CG, and the averaged vector at the CG. Specific combinations (d_r, d_n) are used in Sec. V A.

The above procedure is repeated for all protoclusters within a supercluster. As a result, the CGs are distributed in narrow regions. The procedure is repeated to obtain more converged CGs. The CGs evaluated from a supercluster constitute a skeleton of the supercluster. The above procedure of constructing a skeleton is repeated for all superclusters.

F. Construction of ensemble-averaged quantities

This paper introduces the hierarchy of p clustering. Each constituent of a p cluster can be individually identified by a unique number. The realization of direct access to the hierarchical p clusters enables the construction of unique ensemble-averaged quantities.

For example, in Sec. V B, superclusters with positive (negative) ω_x that have upstream origins in the near-wall region are selectively extracted. On the basis of the results, near-wall vortical regions associated with off-wall vortical regions under a specific flow topology are selected and statistically considered. In Sec. V C, supercluster elements with specific combinations of mathematical terms predominantly representing the RHSs (Right Hand Sides) of the enstrophy and helicity equations are extracted. Collecting these p clusters or elements with specific properties enables a unique ensemble E associated with a p-= cluster and the ensemble-average $\langle f \rangle_E$ of a quantity f to be constructed.

A function defined on a p cluster can be used to divide the p cluster itself. We refer to this as *functional clustering*. For example, the backward subclustering introduced in Sec. IV C 2 is a functional clustering for clusters based on distances between clusters on the $(j$-1)th and jth x planes. Suppose that a p cluster E can be divided into subclusters E_k such that $E = \bigcup_{m=1} E_m$(which is different from a fundamental cluster) using some function f or the parts of some function f_i such that $f = \sum_{m=1} f_m$. Ensemble averages defined by the subcluster or function part f_i can also be constructed. These ensembles enable us to elucidate the dynamical similarity between different p clusters and examine the statistically averaged dynamics of p clusters. In Sec. V C, the functional clustering of supercluster elements is discussed.

G. Comparison between the present method and conventional vortex analysis method

First, the present method is compared with conventional vortex analysis methods qualitatively to clarify the extra insight gained using the present method. Quantitative comparison will be conducted later in Sec. V A. Among vortex analysis methods, variable-interval time-average (VITA) [52], variable-interval space-average (VISA) [53], four-quadrant analysis [54], invariants of the velocity gradient tensor [55], skeletonization by the sectional-swirl-and-pressure-minimum scheme [10,11], proper orthogonal decomposition (POD) [56], snapshot POD [57], dynamic mode decomposition (DMD) [58], linear stochastic estimation (LSE) [59], and Lagrangian coherent analysis (LCA) [28] are well-used.

The VITA method [52] extracts instants at which the highest velocity fluctuations occur and identifies the instants at which the variance of the velocity data in a significant time interval is greater than the variance in the entire series. The VISA [53] method is the spatial counterpart of the VITA method.

In four-quadrant analysis [54], the u-v sample space of fluctuating velocities is divided by axes into four quadrants. In quadrants 2 and 4, the product uv is negative and events consequently correspond to positive turbulence production. Quadrants 2 and 4, respectively, correspond to ejection and sweep.

The invariants of the velocity gradient tensor characterize the eigenvalues of the tensor. Here, the characteristic equation is written as

$$\lambda^3 + P\lambda^2 + Q\lambda + R = 0. \tag{11}$$

Positivity of Q corresponds to the positive Laplacian of pressure. The condition of having complex conjugate roots is

$$(Q/3)^3 + (R/2)^2 > 0. \tag{12}$$

In POD [56], a fluctuating velocity field is decomposed using empirical eigenfunctions of the integral equation whose kernel is the averaged autocorrelation function. The empirical eigenvalues are non-negative, owing to the symmetry of the kernel. The ranking of the magnitude of the eigenvalues corresponds to the dominance of a mode, which is a product of the eigenfunction with the expansion coefficient of the fluctuating velocity field into the series of the empirical eigenfunction. In snapshot POD [57], time-series data of velocity fluctuations in a domain are used.

In DMD [58], data are collected in the form of a snapshot sequence that is given by a matrix $\mathbf{V}_1^N$,

$$\mathbf{V}_1^N = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_N\}, \tag{13}$$

where $\mathbf{v}_i$ is the ith flow field. By assuming a linear mapping A that connects the flow field $\mathbf{v}_i$ to the subsequent flow field $\mathbf{v}_{i+1}$, i.e.,

$$\mathbf{v}_{i+1} = A\mathbf{v}_i \tag{14}$$

$$\mathbf{V}_1^N = \{\mathbf{v}_1, A\mathbf{v}_1, A^2\mathbf{v}_1, \ldots, A^{N-1}\mathbf{v}_1\}. \tag{15}$$

The dynamic characteristics of the dynamical process described by A are extracted in the form of global modes based on the sequence $\mathbf{V}_1^N$.

The LSE technique [59] aims at obtaining the best linear approximation to a conditionally averaged flow field $\langle u_i'(x_i')|u_i(x_i)\rangle$, where $u_i(x_i)$ is a velocity event specified at point x_i, upon which the flow is conditioned. The best linear estimation of the fluctuating flow field $\hat{u}_i'(x_i')$ is expressed in terms of an event vector $u_i(x_i)$:

$$\hat{u}_i'(x_i') = \sum_{j=1}^{3} L_{ij}(x_i', x_i)u_j. \tag{16}$$

Here, L_{ij} denotes the linear estimation coefficients that are determined so as to minimize the mean square error between the linear estimate $\hat{u}_i'(x_i')$ and the conditional average $\langle u_i'(x_i')|u_i(x_i)\rangle$. The estimated flow field depends on the event vector.

LCS [28] is a coherent structure identified using Lagrangian methods and is a material surface that divides a flow into dynamically distinct regions. Here, a coherent structure in a turbulent flow is, according to Robinson's definition [40], a region of the flow over which at least one fundamental flow variable (e.g., the velocity component, density, or temperature) exhibits significant correlation with itself or with another variable over a range of space and/or time that is significantly larger than the smallest local scales of the flow. Simply stated, a coherent structure is a region of the flow that moves together in a statistical sense. A vortex is then an example of a coherent structure in which fluid particles orbit a common axis.

The present method differs from all the above methods. Specifically, the present method and p clusters have the following differences and features when compared with Robinson's definition and the vortex analysis methods mentioned above. Robinson's definition implies that a coherent structure is not judged in terms of an instantaneous flow field because the property of moving together can only be judged in a space-time sense. In addition, a coherent structure is a material volume of fluid. Meanwhile, the present p clusters at levels lower than a hypercluster are constructed in an instantaneous flow field. Time comes into play in hyperclusters. The supercluster extracts a volume region by accumulating clusters. Concerning visualization, the present method allows the algorithmic extraction of individual connected vortical regions or their partial regions as the hierarchical data objects with the interior regions. This point is a unique and innovative aspect of the present method that does not apply to other well-used methods, including LCA, which automatically tracks only the motion of a material surface. The spatial distribution and temporal evolution of the objects can be visualized and automatically tracked. Additionally, some objects are selectively extracted by filtering out other objects for distinct visualization. Concerning data mining, the present method allows the evaluation of balances between mathematical terms that characterize local dynamics at each point belonging to the objects. Similarity and nonsimilarity of the balances (i.e., equivalence) between the subsets of the objects can be mechanically evaluated and retrieved as shown in Sec. V C, which makes it more automatic and systematic to find vortical regions with a common mechanism.

V. RESULTS AND DISCUSSION

Vortical regions that appear in the late stage of the boundary-layer transition are mainly analyzed adopting the clustering method proposed in Sec. IV. After demonstrating the rationality of the present method, the spatial and temporal distributions of p clusters and the functional clustering of supercluster elements are discussed.

A. Rationality of the present method

First, to verify the rationality of the extended clustering methods proposed in Secs. IV E 1 and IV E 2, a vortex ring is compared with resultant superclusters and skeletons. The vortex ring is generated as follows: A torus expressed by the following equations is generated:

$$(\sqrt{x^2+y^2}-R)^2+z^2=r^2(R>r>0), \quad r=0.3,\ R=3. \tag{17}$$

The torus is then rotated through angles of 45° around both y and z axes. The torus is represented on a Cartesian mesh that has a mesh width of 0.0269 in all directions. A vector is defined at each mesh point inside the torus. This vector is the unit normal vector of the plane on which the major axis lies and passes through the mesh point. Figure 11 shows the vortex ring and its point cloud with vectors. Figures 11(a) and 11(b), respectively, show the isosurface of the vortex ring and the points with vectors.

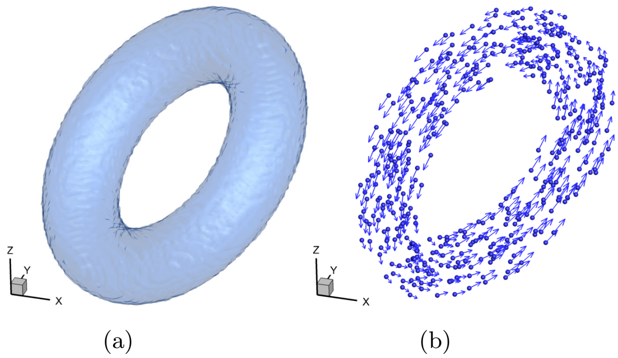


FIG. 11. Vortex ring and its point cloud with vectors. (a) Isosurface. (b) Vortex ring. Points are plotted every 200 points for visualization purposes.

Figure 12 shows superclusters and skeletons constructed by clustering in x, y, and z directions and the accumulated plots of the skeletons. In the results for all directions, points are clearly split into two superclusters depending on the signs of the vorticity components. The skeletons reasonably represent the approximate shapes of the superclusters. The accumulated plots of the skeletons shown in Fig. 12(d) reasonably restore the approximate shape of the torus.

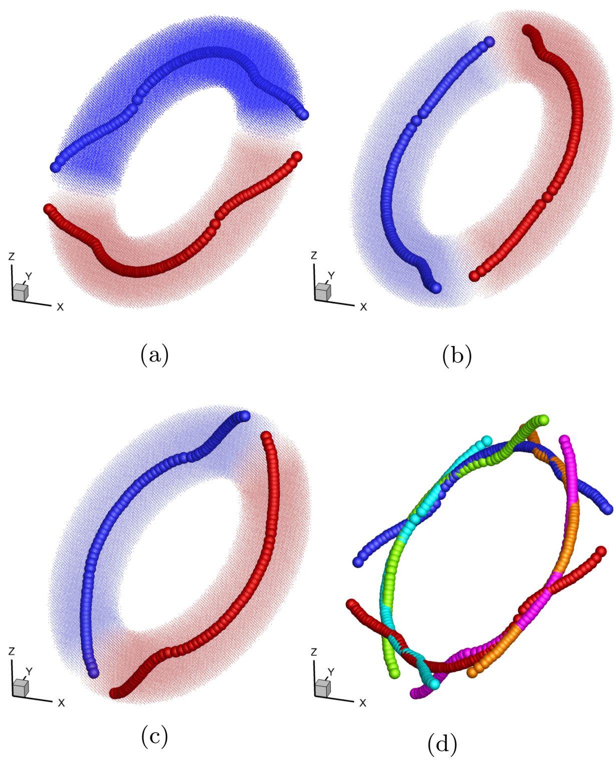


FIG. 12. Superclusters and skeletons constructed by the x-, y-, and z-directional clusterings. (a) Superclusters and skeletons constructed by the x-directional clustering. Red: $\omega_x > 0.1$, Blue: $\omega_x < -0.1$. (b) Superclusters and skeletons constructed by the y-directional clustering. Red: $\omega_y > 0.1$; blue: $\omega_y < -0.1$. (c) Superclusters and skeletons constructed by the z-directional clustering. Red: $\omega_z > 0.1$; blue: $\omega_z < -0.1$. (d) Accumulated plots of the skeletons constructed by the x-, y-, and z-directional clusterings.

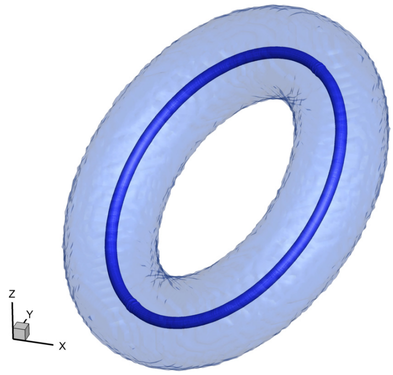


FIG. 13. Skeleton extracted by the clustering method based on a vortex axis.

Meanwhile, small deviations of a skeleton from the major circle near supercluster boundaries, which is due to the finite thickness of the vortex ring, are observed in the unidirectional clustering. The locations of the skeletons can be improved so as to lie on the major circle when the clustering based on a vortex axis is used. Figure 13 shows the skeleton extracted by the clustering method. This skeleton lies precisely on the major circle.

Although protoclusters are split in two superclusters in the unidirectional clustering, this division is not due to the property of the clustering but is due to the variable used for extracting the protoclusters. By using components ω_x, ω_y, and ω_z, the appearance of multiple clusters on a cross section is prevented. By merging superclusters of the unidirectional clustering in different directions, the reconstruction of the original protoclusters can be achieved. Although the small deviations of skeletons near the ends of the superclusters are observed, a skeleton is just a subsidiary member of the clustering and it affects neither the construction of the p cluster nor the reconstruction of the original protocluster distribution. While the two methods give characteristically different clustering results, they are both successful in constructing superclusters and skeletons.

Second, the clustering based on a vortex axis is then applied to a flow field of the late stage of boundary-layer transition to show the rationality and applicability of the clustering method to the flow. Figure 14 shows the original protoclusters and clustering results. Figure 14(a) shows protoclusters extracted using the Δ criterion with the imaginary parts of eigenvalues greater than 0.1. Here, the velocity gradient tensor is nondimensionalized by c_0 and δ_{in}. The extracted protoclusters account for about 24% of protoclusters that have imaginary parts of eigenvalues greater than 10^{-10}. Figures 14(b) and 14(c), respectively, show superclusters and skeletons. In Fig. 14(b), discrete superclusters are successfully obtained. In Fig. 14(c), the locations of CGs which constitute skeletons are well converged in narrow regions. Here, step (d′) explained in Sec. IV E 2 is applied twice. In the first and second applications, the combinations (d_r, d_n) are (2, 0.4) and (2, 1.0), respectively. Each skeleton is constructed from protoclusters belonging to each supercluster. The color of a skeleton corresponds to that of the corresponding supercluster. Third, the resultant skeletons are compared with supercluster elements and the distribution of ω_x to verify the rationality of the unidirectional clustering method regarding the late stage of transition. Figure 15 shows supercluster elements, the boundary surfaces of the superclusters, and their resultant skeletons. Numerous superclusters are extracted by the present clustering method; for the clarity of visualization, only a few superclusters are shown. The skeletons pass through the superclusters' elements, indicating that the skeletons have been successfully constructed. Figure 16 shows hyperclusters constructed from superclusters for $t/(\delta_{\text{in}}/c_0) = 0$, 1.6, and 3.2. The shapes of the superclusters remain largely similar, and small variations in the supercluster locations are captured as time elapses. The computational observation indicates that grouping superclusters as a hypercluster makes sense. Figure 17 shows the distribution of ω_x on the streamwise series of x planes, and also the skeletons of superclusters with $\omega_x > \omega_{x,0}$ at $t = 0$. The long streamwise continuation of the regions of ω_x supports the use of ω_x for extracting

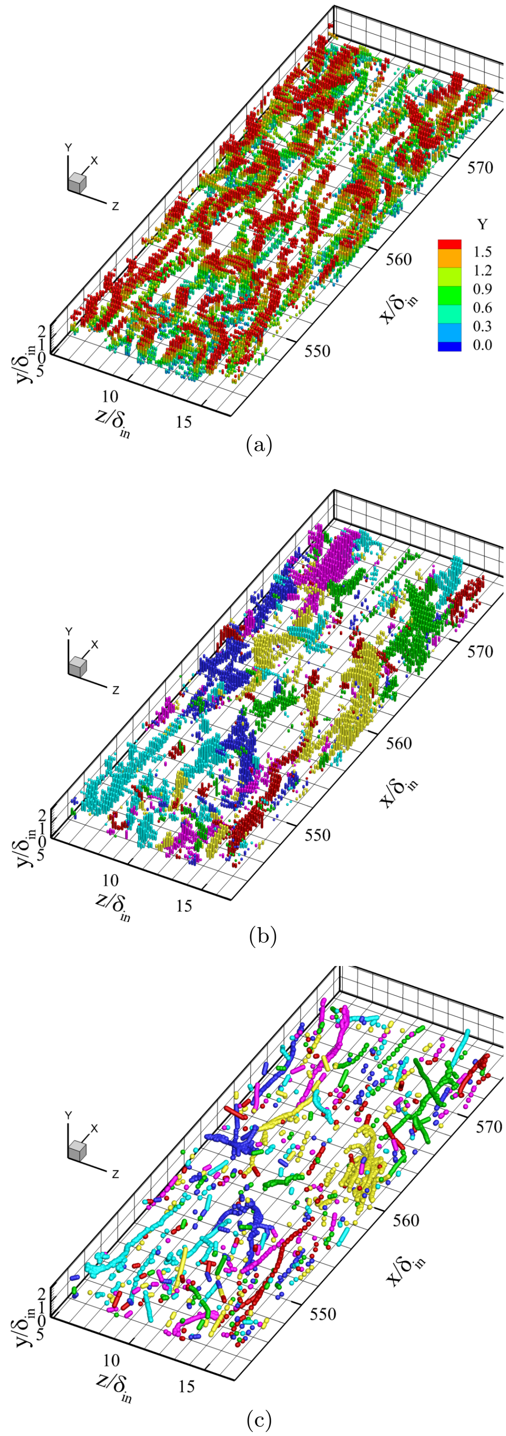


FIG. 14. Protoclusters in vortical regions, and superclusters and skeletons constructed by the clustering based on a vortex axis. Different colors in (b) and (c) show different superclusters and skeletons. The limitation of the number of recognizable colors means that some different superclusters (skeletons) are shown in the same color. (a) Protoclusters extracted by the Δ criterion with the imaginary parts of eigenvalues greater than 0.1. (b) Superclusters. (c) Skeletons.

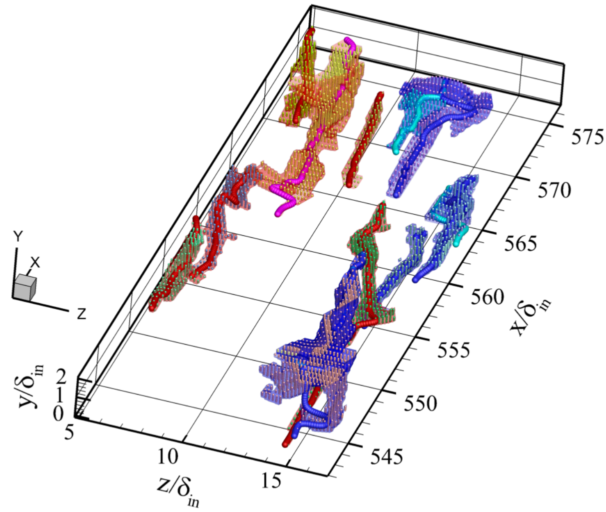


FIG. 15. Supercluster and their skeletons at $t = t_1$. Superclusters are, in particular, shown by both isosurfaces and points. Only a few superclusters are shown for clarity. The skeletons of superclusters with $\omega_x > \omega_{x,0}$ are shown in red and pink, and those with $\omega_x < -\omega_{x,0}$ are shown in blue and cyan.

vortex structures. The skeletons pass through the middle of the series of the regions of large ω_x. It is clear from Figs. 15–17 that the present algorithm successfully constructs superclusters that connect regions of large ω_x.

Fourth, to show the relationship between the unidirectional and omnidirectional clustering regarding the detailed structures of superclusters in wall-bounded flows, the protoclusters extracted using the Δ criterion in the late stage are clustered below. Here, the unidirection is the x direction.

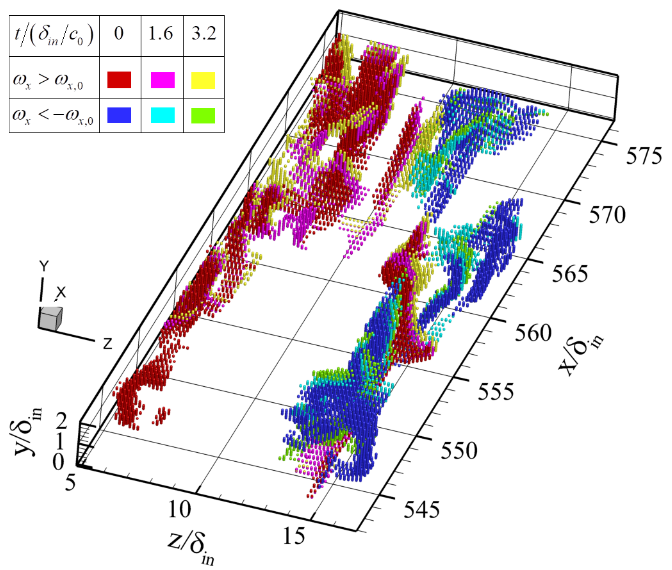


FIG. 16. Hyperclusters constructed from superclusters at $t/(\delta_{\text{in}}/c_0) = 0$, 1.6 and 3.2. Regions of $\omega_x > \omega_{x,0}$ are shown in red, pink, and yellow for $t/(\delta_{\text{in}}/c_0) = 0$, 1.6, and 3.2, respectively. Regions of $\omega_x < -\omega_{x,0}$ are shown in blue, cyan, and yellow green for $t/(\delta_{\text{in}}/c_0) = 0$, 1.6, and 3.2, respectively. The limitation of the number of recognizable colors means that different superclusters are shown in the same color.

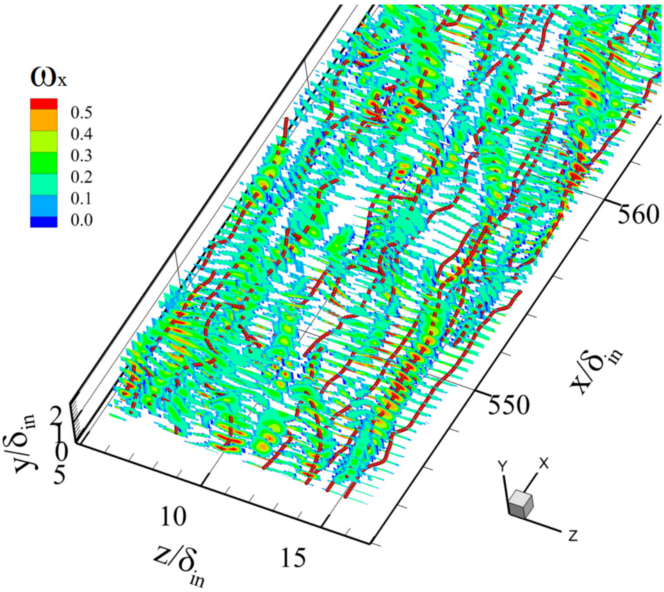


FIG. 17. Distribution of ω_x on the streamwise series of x planes and the skeletons of superclusters with $\omega_x > \omega_{x,0}$ at $t = t_1$. Only the regions with $\omega_x > \omega_{x,0}$ are shown for the series of x planes.

Figure 18 presents the resultant superclusters obtained in the omniclustering. In particular, we focus on the superclusters and skeletons A and B shown in green and light blue, respectively. Figure 19 shows resultant superclusters and skeletons of the superclusters obtained in omnidirectional and unidirectional clustering. Connected groups of protoclusters are extracted. Superclusters extracted

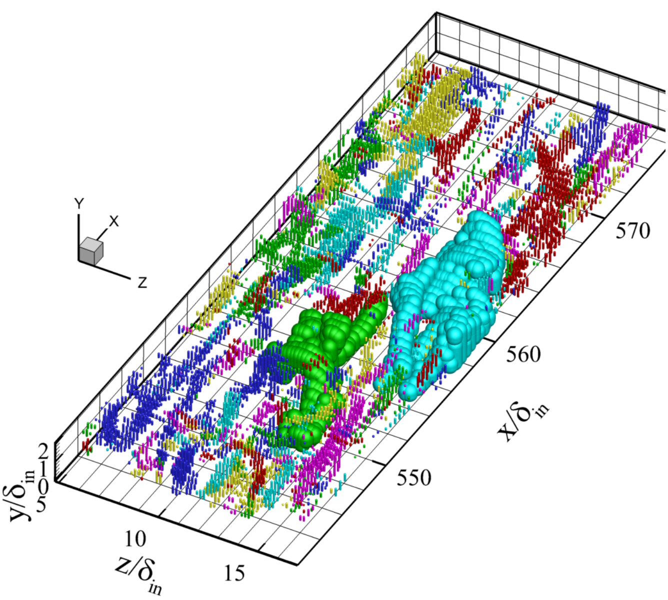
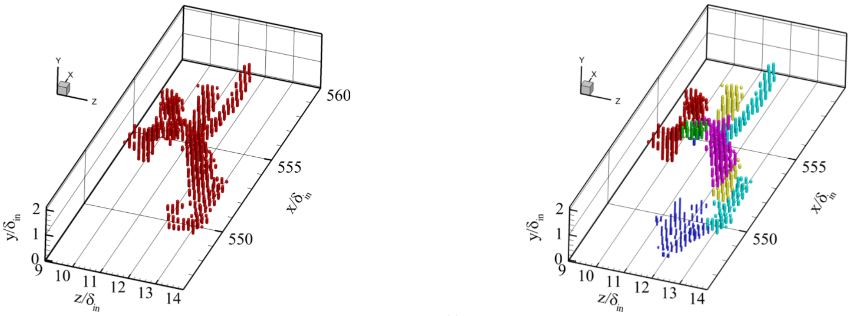
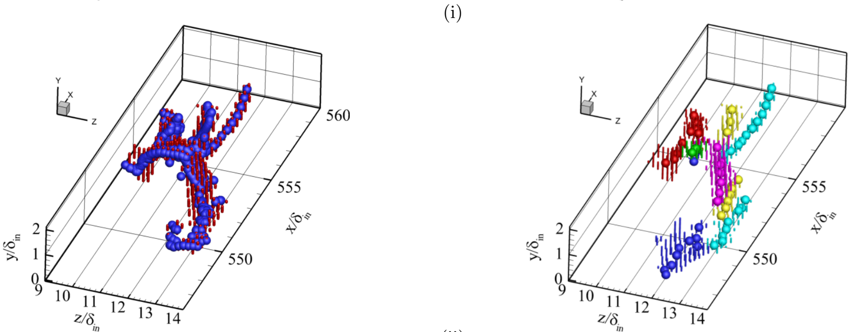


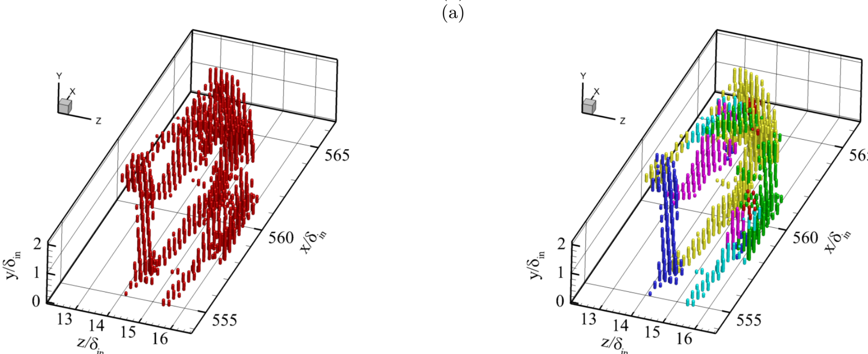
FIG. 18. Superclusters extracted by the omnidirectional clustering at $t = t_1$ (green: supercluster A; light blue: supercluster B). Different colors show different superclusters. The limitation of the number of recognizable colors means that some different superclusters are shown in the same color.



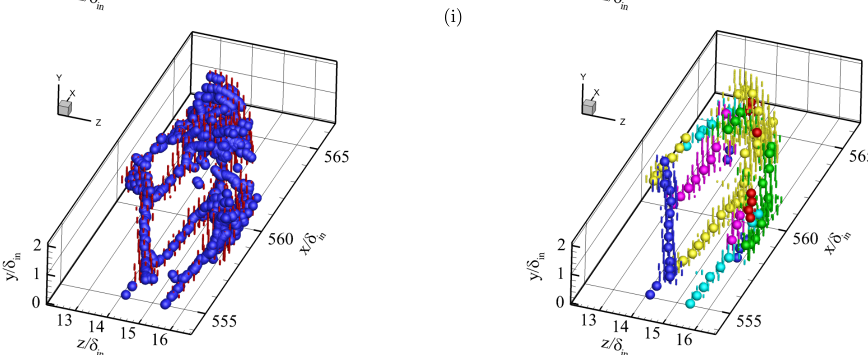
(i)



(ii)
(a)



(i)



(ii)
(b)

FIG. 19. Resultant superclusters (small spheres) and skeletons (large spheres) of superclusters A and B obtained by both omnidirectional and unidirectional clustering. Different colors show different superclusters. The limitation of the number of recognizable colors means that some different superclusters are shown in the same color. (i) Superclusters (left: omnidirectional clustering; right: unidirectional clustering). (ii) Skeletons (left: omnidirectional clustering; right: unidirectional clustering). (a) Supercluster A. (b) Supercluster B.

using the omnidirectional method are successfully represented by the union of multiple superclusters extracted using the unidirectional method. In particular, each branch of the structures of the superclusters obtained using the omnidirectional method is represented by superclusters obtained using the unidirectional method. The skeletons obtained using the unidirectional method represent the nearly mean locations of planar clusters stacked in the x direction. Additionally, the superclusters extracted using the omnidirectional method are simplified through the component-wise representation by skeletons.

The clustering based on a vortex axis can construct p clusters without assuming a specific component of vorticity nor a specific direction of a vortex axis. However, the structure of a p cluster becomes much more complex than that of a p cluster constructed by unidirectional clustering. Positive and negative signs of a vorticity component are mixed and superclusters are intricately bifurcated. The backward subclustering cannot be defined clearly. Meanwhile, the unidirectional clustering algorithm categorizes protoclusters depending on the sign of ω_x without losing points. Additionally, the missing connection between separated superclusters can be inferred by other unidirectional clustering in different directions if there are vortical regions connecting the separated superclusters. For these reasons, we adopt the unidirectional clustering hereafter in this paper to make the recognition of vortical regions simple. The extension of the present analysis to general p clusters constructed by the clustering based on a vortex axis will be investigated in future work.

Fifth, the results of the present method are quantitatively compared with those of conventional vortex analysis methods. Among the methods mentioned above, VITA, VISA, four-quadrant analysis, POD, DMD, and LSE are statistical methods and not suitable for comparison with the present method, which extracts vortical regions from instantaneous flow fields. Although the other methods are also not directly comparable with the present clustering methods, which have a hierarchical property, some methodological differences are discussed by applying them to the same flow field of the initial stage. In Fig. 20, results of the present method are compared with those of the Kida-Miura method, LCS, and community model [60].

Figure 20(a) shows resultant superclusters generated in the omnidirectional clustering based on a vortex axis for protoclusters extracted using the Δ criterion with the imaginary parts of eigenvalues greater than 3×10^{-2}. There are 34 534 protoclusters. When there is no arch over the two hairpin legs of the protoclusters, the two legs are independently recognized as different superclusters as found around $x/\delta_{\text{in}} = 420 - 430$. When legs are connected by arches, the connected system is recognized as a single supercluster as found at $x/\delta_{\text{in}} = 440$.

Figure 20(b) shows resultant superclusters generated in the x-directional clustering for the same protoclusters in Fig. 20(a). When there is no arch over the two hairpin legs, the two legs are recognized as two different superclusters similar to what is seen in Fig. 20(a). When the legs are connected by arches, the connected system is decomposed into different superclusters near the center in the spanwise direction, owing to the prevention of the multiple correspondence as explained in Sec. IV C 2.

Figure 20(c) shows vortical points extracted using the Kida-Miura method [10,11]. The purpose of this method is to extract and track swirling low-pressure vortices which are observed as universal and elementary vortices in turbulent flows. It extracts a chain of vortex cores according to the criterion that the flow has swirl on the plane normal to the third/smallest eigenvector of the pressure Hessian with two positive eigenvalues. The swirl criterion D [10] is

$$D = \frac{1}{4}(W_{11} - W_{22})^2 + W_{12}W_{21} < 0. \tag{18}$$

Here, W_{ij} $(i, j = 1, 2)$ is the velocity gradient tensor defined as $W_{ij} = \partial u_i/\partial x_j$ using a velocity vector (u_1, u_2) on a $x_1 - x_2$ plane. W_{ij} is nondimensionalized by c_0 and δ_{in}. The number of extracted points depends on the value of D. Vortical points with $D < -100$ are shown in the figure. Except for small vortex fragments that depend on the value of D, as indicated by the arrows, this method can extract protoclusters very sharply, similar to the Δ method. Compared to the Kida-Miura method, our method has the functions of the hierarchy-based clustering and tracking, which the Kida-Miura method does not possess.

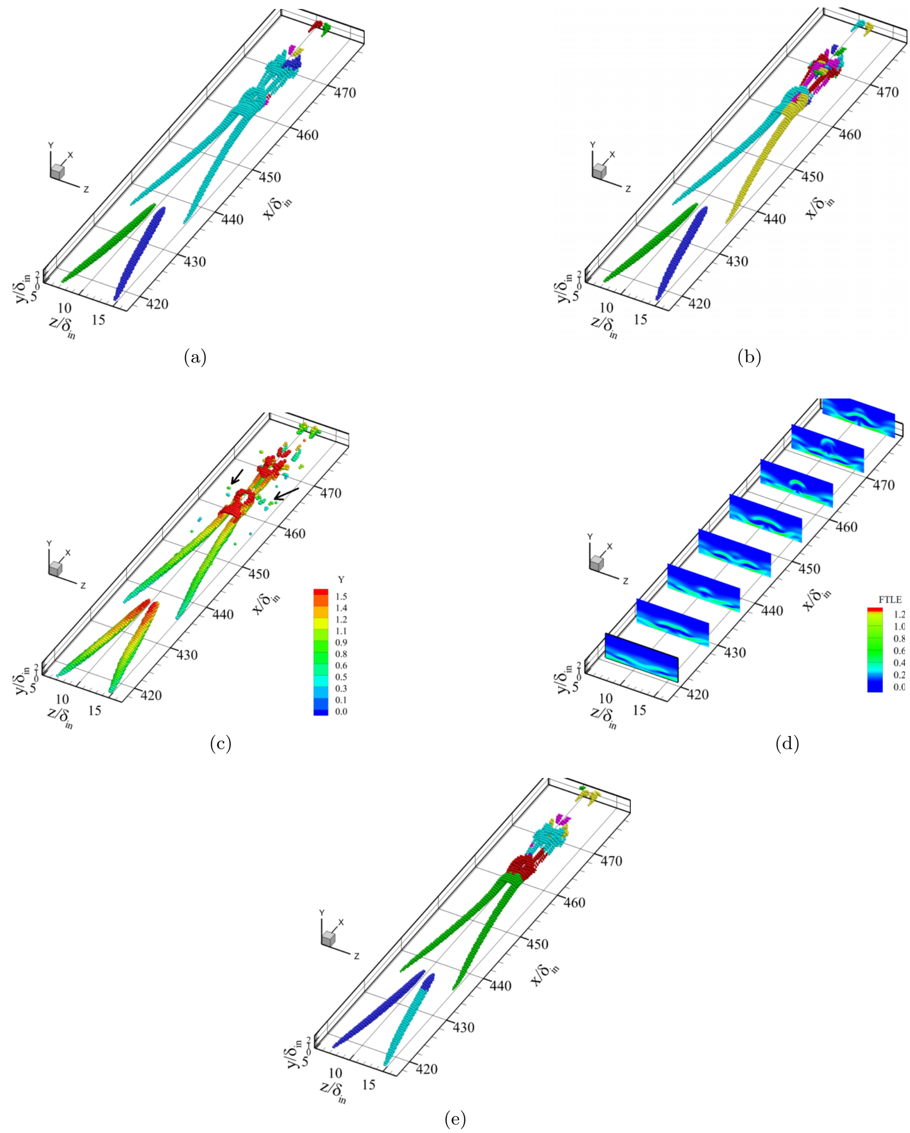


FIG. 20. Comparison of the present method with the Kida-Miura method, LCS, and the community model for an instantaneous flow field of the initial stage. (a) Resultant superclusters generated by the omnidirectional clustering based on a vortex axis for protoclusters extracted by the Δ criterion. (b) Resultant superclusters generated by the x-directional clustering for protoclusters extracted by the Δ criterion. (c) Vortical points extracted by the Kida-Miura method ($D < -100$). (d) Scalar contours of FTLE at each cross section. (e) Vortical points divided by the community model.

Figure 20(d) shows the finite-time Lyapunov exponent (FTLE), which is the time-averaged exponent of the largest possible stretching at point x_0 within the time period $[t_0, t_1]$ and defined as

$$\Lambda_{t_0}^{t_1}(x_0) = \frac{1}{t_1 - t_0} \ln\sqrt{\lambda_n(x_0)}. \tag{19}$$

Here, λ_n is the largest eigenvalue of the right Cauchy-Green strain tensor defined as the product of the transpose of the gradient tensor of the flow map and the gradient tensor. The lifting up of the near-wall structure, the arches of hairpin vortices, and multiple hairpin (lambda) vortices are neatly captured in Fig. 20(d). On the other hand, the LCS is defined as a ridge of the scalar field $\Lambda_{t_0}^{t_1}(x_0)$ [61] in contrast to the present method extracting volumetric vortical regions rationally separated from each other. The above formulation is especially oriented for constructing and tracking accurate and robust LCSs not only from computation but also from measured data such as time-resolved velocity fields extracted from video footage by space probes [23]. In fact, the structural stability of the FTLE ridge under perturbations (noise) is guaranteed by Karrasch and Haller [61]. Thus, the remarkable property of the LCS makes the method intrinsically different from our method.

Figure 20(e) shows the results of applying the community model [60] to the same points (protoclusters) considered in Figs. 20(a) and 20(b). In this method, the strength of interaction between two nodes is given by the induced velocity from the vorticity distribution in the local neighborhood of one node to the other. The points are then divided into communities so as to maximize the modularity Q [62–64]. In applying this method to the present problem, the modularity resolution parameter [60] was set at 1. As seen in Fig. 20(e), an efficient division of the original points into several communities is realized. However, the right hairpin leg around $x/\delta_{\mathrm{in}} = 435$ is divided into two communities, with the downstream community belonging to the same community as that of the left hairpin leg. Therefore, two symmetric groups of points separated from each other are not necessarily divided into two independent and equivalent communities contrary to the visual separation, which was also pointed out as a characteristic in the spectral clustering of vortical points [30]. The division that does not follow intuition loses advantages of connectedness and symmetry within the p cluster in pursuing categorization, retrieval, tracking, and visualization for the present objective. Additionally, simplicity in evaluating the distribution of a certain physical quantity and analyzing mathematical terms over the p cluster is lost. These problems are successfully overcome using the method proposed in this paper.

B. Spatial and temporal distribution of *p* clusters

To understand the variation in vortex distribution leading to the late stage, Fig. 21 shows all extracted skeletons for $\mathrm{Re}_x = 300.79 - 575.1$.

Figure 21(a) shows the initial stage ($\mathrm{Re}_x = 300.79 - 479.8$), Figure 21(b) shows the middle stage ($\mathrm{Re}_x = 480.2 - 540.6$) and Fig. 21(c) shows the late stage ($\mathrm{Re}_x = 543.5 - 575.1$). Figure 21(a) is further divided into subpanels (i) before hairpin vortex generation and (ii) after hairpin vortex generation. In Fig. 21(a), the construction of superclusters and skeletons is conducted alternately in the x and z directions to extend the present method to multiple directions, and also to illustrate the three-dimensional shear structures of the vortical regions. As found from $x/\delta_{\mathrm{in}} \geqslant 320$, near-wall shear layers with negative ω_z are perturbed by the existence of ω_x, which is shown by the skeletons of ω_x, and the deformation grows downstream. These structures are strongly similar to hairpin vortices observed in experiments [65]. In Figs. 21(a) and 21(b), there is a degree of symmetry in the skeleton distribution around $z = 11\delta_{\mathrm{in}}$. In particular, Fig. 21(a) shows that, as the skeletons of a hairpin vortex lifts up, new skeletons labeled A and B′ emerge near the wall. The skeletons with positive and negative signatures then approach each other and lift up downstream. Skeletons of ω_z appear around the archlike heads and legs of the symmetric hairpin vortices. In Fig. 21(b), the skeleton distribution becomes more complex than that in Fig. 21(a). The area of the skeleton distribution grows in both the spanwise and downstream directions. Symmetric skeleton pairs such as C and D appear near the center of symmetry around $z = 11\delta_{\mathrm{in}}$, although the peripheral vortical regions appear to be asymmetric if viewed locally. In Fig. 21(c), the skeleton distribution becomes more disorganized and uniform than in Figs. 21(a) and 21(b) in both the x and z directions.

In Sec. V C, the dynamics of the vortical regions is investigated by evaluating balances between mathematical terms that categorize local dynamics at points belonging to the regions. Prior to the

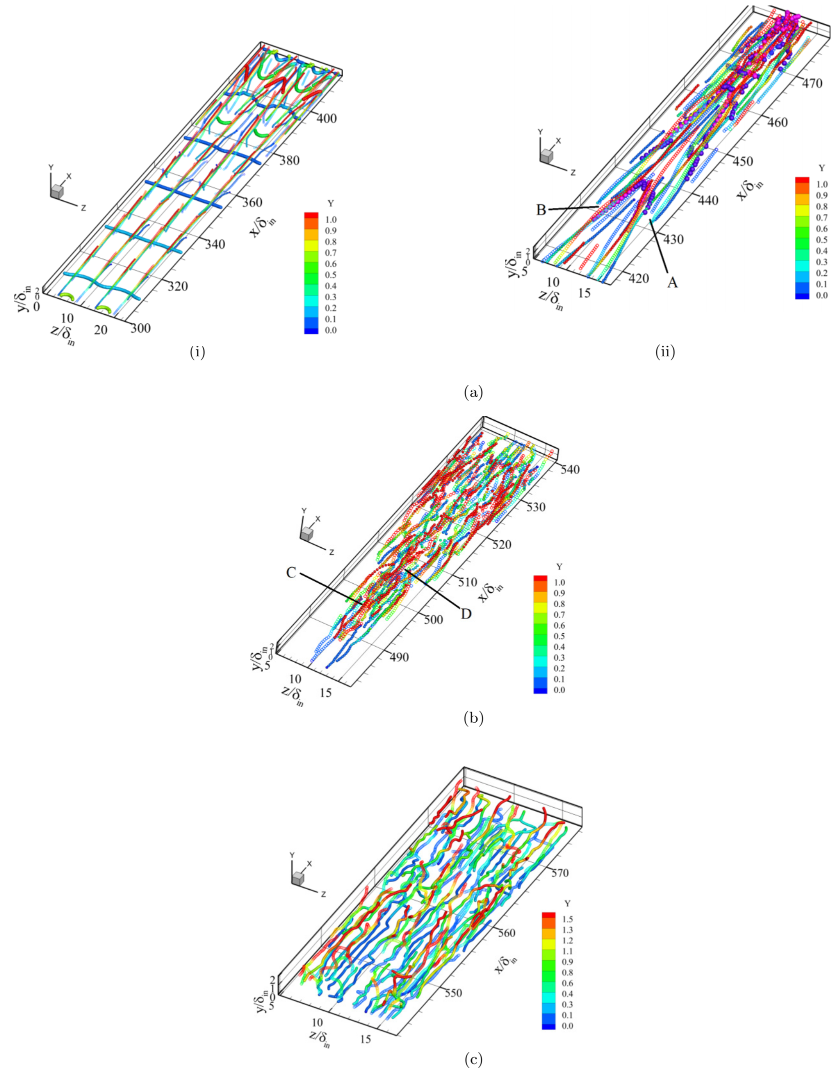


FIG. 21. All extracted skeletons for the initial, middle, and late stages ranging from $\mathrm{Re}_x = 300.79 - 575.1$ at $t = t_1$. (i) Before hairpin vortex generation. Skeletons are obtained by alternate clustering in the x and z directions on the basis of ω_x and ω_z, respectively. Small open circles and small spheres are skeletons of positive and negative ω_x, respectively. Big spheres are skeletons of negative ω_z. The threshold values for ω_x and ω_z are 0.001 and -0.3, respectively. (ii) After hairpin vortex generation. Skeletons are obtained by alternate clustering in the x and z directions on the basis of ω_x and ω_z, respectively. Open circles and small spheres are skeletons of positive and negative ω_x, respectively. Pink and purple spheres are skeletons of positive and negative ω_z, respectively. The threshold value for ω_x is 0.01. The threshold value for ω_z depends on whether ω_z is positive or negative because of bias in the statistical mean value; if positive, the threshold is 0.01, if negative the threshold is -0.5. (a) Initial stage. (b) Middle stage. Open circles and small spheres denote the skeletons of positive and negative ω_x, respectively. The threshold value for clustering is 0.1. (c) Late stage. Open circles and small spheres represent the skeletons of positive and negative ω_x, respectively. The threshold value for clustering is $\omega_{x,0}$.

subsection, a matrix for producing the balances is investigated in terms of the spatial and temporal distribution of p clusters.

To investigate the origins of near-wall vortical regions and observe how regions with ω_x develop downstream, a series of near-wall vortical regions were extracted. Some vortices lift up and grow while others remain attached to the wall (this behavior is confirmed by Fig. 25, which is discussed later).

To extract lifted-up or horizontal vortical regions separately, the local slopes of the skeletons are evaluated by the least-squares method using successive skeleton points:

$$L_i = \mathrm{LSQ}\{\mathbf{x}_i, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_{p_{\max}}\}, \quad \text{for } i = 1, \ldots, p_{\max} - 1. \tag{20}$$

Here, LSQ evaluates the slope L_i of a regression line in the x-y plane using points $\mathbf{x}_i, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_{p_{\max}}$, and $p_{\max}$ is the total number of point elements in a skeleton. If all regression lines have a positive inclination, the supercluster is judged to have lifted up.

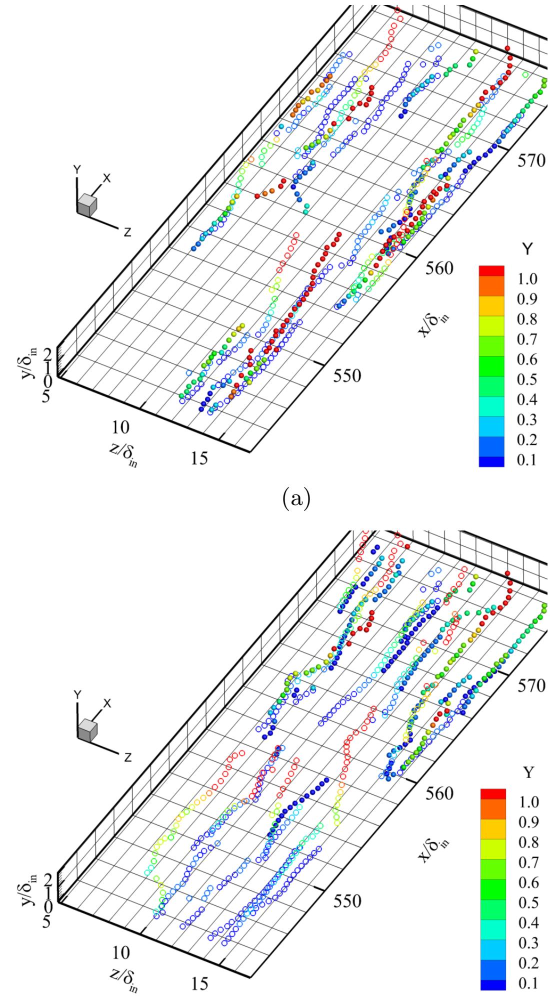
Figure 22 shows all near-wall skeletons and off-wall skeletons closest to and/or associated with each near-wall skeleton. Here, a near-wall skeleton is defined as a skeleton in which some parts are less than $0.1\delta_{\text{in}}$ from the wall. Off-wall skeletons have heights that are uniformly greater than $0.1\delta_{\text{in}}$. Figure 22(a) shows near-wall skeletons with positive ω_x and their associated off-wall skeletons, and Fig. 22(b) shows near-wall skeletons with negative ω_x and their associated off-wall skeletons.

The associated skeleton for any near-wall skeleton is found by the following procedure: First, within a skeleton with positive ω_x, an element of the lowest height is selected from the initial three points near the upstream origin. A skeleton with positive (negative) ω_x containing the closest element to the near-wall skeleton is then selected as the associated skeleton. Although the height of an associated skeleton is not considered in the above selection procedure, the associated skeleton is likely to contain heights greater than $0.1\delta_{\text{in}}$.

A remarkable feature of the figure is that, in the majority of cases, a near-wall skeleton is associated with an off-wall skeleton with the opposite sign of ω_x, which is considered to be the counterpart of a skeleton pair, and the generation of the upstream ends of near-wall vortical regions are associated with the lift-up from the wall of the off-wall vortical regions. These characteristics are detailed below from the viewpoint of the present clustering method.

Figure 23 shows the heights and ω_x values of skeletons along the x direction. Only the skeletons of near-wall vortical regions with $\omega_x > \omega_{x,0}$ and those of the associated off-wall vortical regions are shown. Here, ω_x values at the skeleton points of vortical regions with $\omega_x > \omega_{x,0}$ and $\omega_x < -\omega_{x,0}$ are not necessarily positive and negative, respectively, because skeleton points are not necessarily protoclusters for those regions. From Fig. 23(a), the angles (the tangents of the angles) of the skeletons of near-wall vortical regions are around $0.1^\circ(0.00175) - 17^\circ(0.306)$, and those of off-wall vortical regions are around $0.3^\circ(0.00524) - 22.4^\circ(0.412)$. In addition, the lift up of near-wall vortical regions coincides with the lift up of the associated off-wall vortical regions, which will be investigated further below. From Fig. 23(b), it appears that ω_x takes particularly high values near the lift up of skeletons from the wall.

Figure 24 shows the local slopes and products between the slopes of the skeletons of near-wall vortical regions and the associated off-wall vortical regions. Figure 24(a) shows the local slopes of the skeletons of near-wall vortical regions and the associated off-wall vortical regions. The number of skeleton elements with a positive slope is 2.8 times greater than that of skeleton elements with a negative slope, which means that the number of lifted-up vortical regions is much greater than that of descending vortical regions. To show the correlations between the slopes of near-wall vortical regions and the associated off-wall vortical regions, the products of the slopes of skeleton pairs are shown in Fig. 24(b). The products are shown at the x coordinates of the skeletons of the near-wall vortical regions. The distributions in the x direction are different between the paired skeletons. When there are no skeleton elements of the associated off-wall vortical regions that correspond to the streamwise locations of the skeleton elements of the near-wall vortical regions, the products are set to zero. In mathematical notation, this operation can be described as follows: Suppose that there is a skeleton pair, i.e., skeleton n and skeleton o, where n is the skeleton of a near-wall vortical region and



(a)

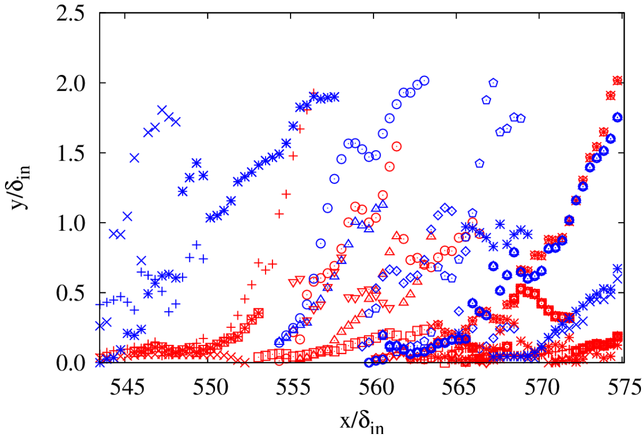
(b)

FIG. 22. All near-wall skeletons with the magnitude of ω_x greater than $\omega_{x,0}$ and their associated off-wall skeletons at $t = t_1$. (a) and (b) show the cases of near-wall vortical regions with positive and negative vorticities, respectively. Open and filled symbols denote the skeletons of the vortical regions with positive and negative ω_x, respectively. The color indicates the height from the wall. (a) Near-wall skeletons with $\omega_x > \omega_{x,0}$. (b) Near-wall skeletons with $\omega_x < -\omega_{x,0}$.

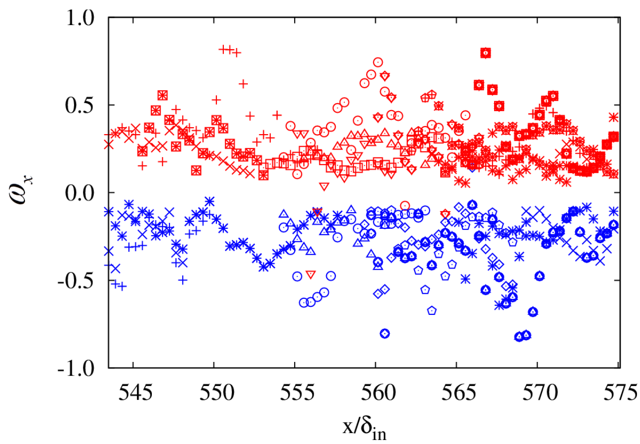
o is the skeleton of its associated off-wall vortical region. The positions of the skeleton elements and slopes at each position are denoted as $\mathbf{X}_n \equiv \{\mathbf{x}_{1,n}, \mathbf{x}_{2,n}, \ldots, \mathbf{x}_{n\max,n}\}$ and $\{b_{1,n}, b_{2,n}, \ldots, b_{n\max,n}\}$ for skeleton n and $\mathbf{X}_o \equiv \{\mathbf{x}_{1,o}, \mathbf{x}_{2,o}, \ldots, \mathbf{x}_{o\max,o}\}$ and $\{b_{1,o}, b_{2,o}, \ldots, b_{o\max,o}\}$ for skeleton o. The products of the local slopes of the skeletons $C_n \equiv \{c_{1,n}, c_{2,n}, \ldots, c_{n\max,n}\}$ are computed by

$$\begin{matrix} b_{q,n}b_{r,o} & \text{if } \mathbf{x}_{q,n} = \mathbf{x}_{r,o} \in \mathbf{X}_o \\ 0 & \text{if } \mathbf{x}_{q,n} \notin \mathbf{X}_o. \end{matrix} \tag{21}$$

From the figure, the distribution of positive products is dominant, which means that the slopes of skeleton pairs are highly correlated. Taking the preponderance of lifted-up vortical regions into



(a)



(b)

FIG. 23. Heights and ω_x values of the skeletons of near-wall vortical regions along x at $t = t_1$. Only the skeletons of near-wall vortical regions with $\omega_x > \omega_{x,0}$ and those of the associated off-wall vortical regions are shown. Different skeletons are denoted by different symbols. The same symbols in different colors indicate skeleton pairs. The same symbols in (a) and (b) denote the same skeletons. Red and blue symbols indicate skeletons with $\omega_x > \omega_{x,0}$ and $\omega_x < -\omega_{x,0}$, respectively. (a) Skeleton heights along x. (b) ω_x values of skeletons along x.

consideration, the generation of near-wall vortical regions is attributed to the lift up of off-wall vortical regions.

Figure 25 shows skeleton time surfaces constructed for $t = 0 - 3.2\delta_{\text{in}}/c_0$. Excluding surface elements that do not satisfy the admissible condition mentioned in Sec. IV C 5, the surface elements have been successfully generated. The figure illustrates skeleton time surfaces that remain close to the wall, those that lift up from the near-wall region to the free stream and exhibit some surface twisting, those that lift up from the off-wall region to the free stream, and also spatial entanglement between the surfaces, including the pairing of skeletons.

To understand the major flow topology that generates the near-wall vortical regions, the flow topology of a near-wall supercluster is categorized around the equilibrium point closest to the intersection of the upstream origin of the skeleton with a spanwise plane. The spatial range of analysis is $\pm 1.0\delta_{\text{in}}$ in the y and z directions around each intersection point. The flow topology is

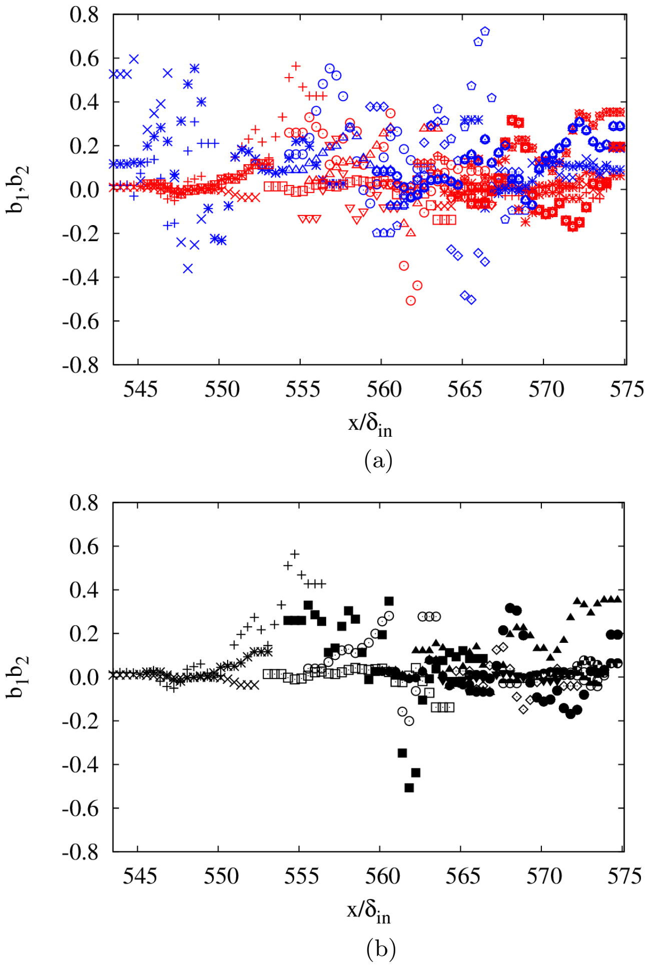


FIG. 24. Local slopes and products between the local slopes of the skeletons of near-wall vortical regions and the associated off-wall vortical regions at $t = t_1$. Only the skeletons of near-wall vortical regions with $\omega_x > \omega_{x,0}$ and those of the associated off-wall vortical regions are shown. By b_1 and b_2, we denote the local slope of a near-wall skeleton and the associated off-wall skeleton in the x-y plane, respectively. (a) Local slopes of the skeletons of near-wall vortical regions and the associated off-wall vortical regions. Red and blue symbols indicate skeletons with $\omega_x > \omega_{x,0}$ and $\omega_x < -\omega_{x,0}$, respectively. (b) Products of the local slopes of the skeletons of near-wall vortical regions and the associated off-wall vortical regions.

categorized in terms of the eigenvalues of the dynamical system of a local velocity field, as explained in Appendix A.

Similar dynamical system analyses can be conducted for other near-wall vortical regions. Figure 26 shows the histogram of the resultant categories of flow topology. The dominant category on the cross-sections near the upstream ends of the near-wall vortical regions (27 out of 41 cases) is spiral sinks. The other category in Fig. 26 corresponds to cases where no equilibrium points are found. Excluding these negligible cases, spiral sinks constitute 65% of all vortex structures.

As an example of a spiral sink, the y and z components of velocity vectors within the analysis domain, the intersection points, and the skeleton of the near-wall vortical region are shown in Fig. 27. As mentioned previously, the near-wall vortical regions are associated, in most cases,

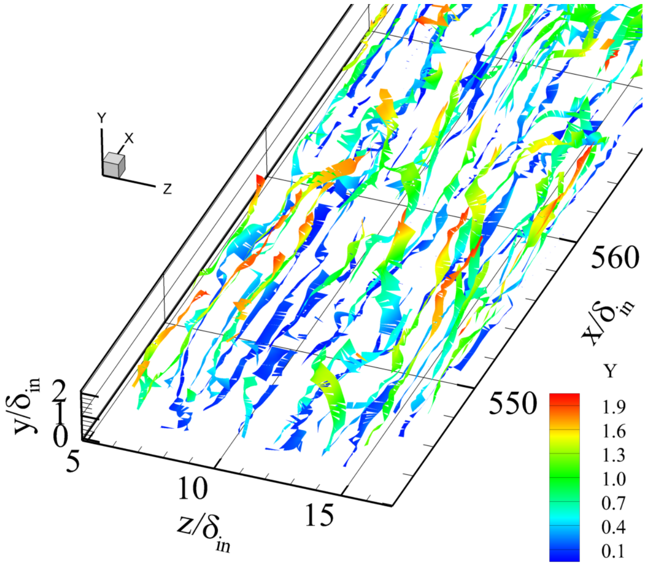


FIG. 25. Skeleton time surfaces constructed for $t = 0 - 3.2\delta_{\text{in}}/c_0$.

with off-wall vortical regions with opposite signs of ω_x. To see this, the skeleton of the associated vortical region that has detached from the wall is plotted. The near-wall vortical region is generated under the influence of the swirls of the off-wall vortex because of the no-slip wall condition.

To understand the downstream evolution of the system of paired vortical regions with ω_x, the streamwise array of cross-sectional streamlines obtained from the y and z components of velocity vectors, i.e., (v, w), within the x planes is shown in Fig. 28.

From Fig. 28, the center of the off-wall vortical region corresponding to the blue points moves in a circular manner within the array of the cross sections. The red skeleton is distributed in the region where the vortical region contacts the wall. With the movement of the off-wall vortical region, the near-wall red skeleton moves accordingly.

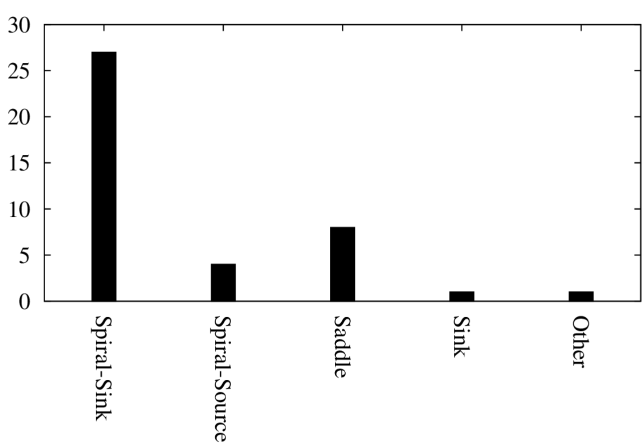


FIG. 26. Histogram of categories at equilibrium points on x planes closest to the intersections between the x planes and near-wall skeletons at $t = t_1$.

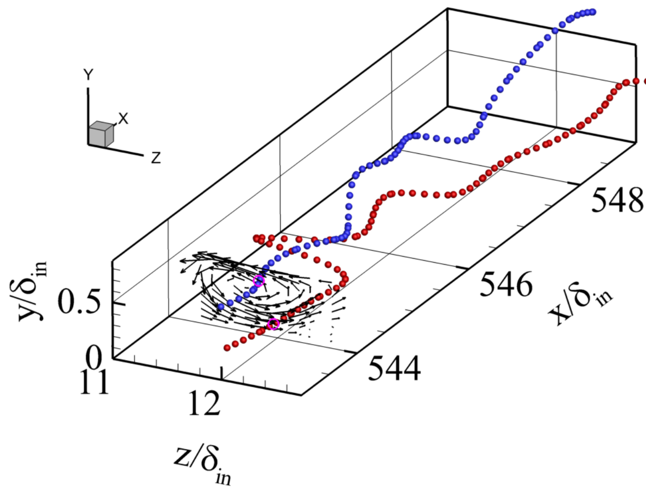


FIG. 27. Typical case of a spiral sink, showing two-dimensional velocity vectors within the analysis domain, intersection points (pink circles), and the skeleton of the near-wall vortical region (red) and its associated off-wall vortical region (blue) at $t = t_1$.

C. Functional clustering of supercluster elements

Balances between mathematical terms that characterize local dynamics can be evaluated at protoclusters belonging to superclusters. Similarity and nonsimilarity of the balances, i.e., equivalence, can be mechanically evaluated between the protoclusters. As a result, subclusters and unique ensembles of subclusters from different superclusters which have same balance inside them can be constructed to gain insight into the local dynamics of vortices and shear layers around vortices.

First, dominant mathematical terms associated with vortex dynamics are visualized in superclusters. For this purpose, enstrophy and helicity equations [66], which are scalar expressions associated with vortex strength, are investigated. The equation for the enstrophy density can be written as follows:

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\omega \cdot \omega\right) = \sum_{n=1}^{5} T_n, \quad T_1 = -\{(\mathbf{u} \cdot \nabla)\omega\} \cdot \omega, \quad T_2 = (\omega \cdot \nabla)\mathbf{u} \cdot \omega,$$

$$T_3 = -\{(\nabla \cdot \mathbf{u})\omega\} \cdot \omega, \quad T_4 = \left\{\nabla p \times \nabla\left(\frac{1}{\rho}\right)\right\} \cdot \omega, \quad T_5 = \left\{\nabla \times \left(\frac{1}{\rho}\nabla \cdot \tau\right)\right\} \cdot \omega, \tag{22}$$

where ω is the vorticity vector, $\mathbf{u}$ is the velocity vector, p is the pressure, and ρ is the density. In T_5, τ is the viscous shear stress tensor, the elements of which are defined as follows:

$$\tau_{ij} = \mu\left[\left(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j}\right) - \frac{2}{3}(\nabla \cdot \mathbf{u})\delta_{ij}\right], \quad i, j = 1, 2, 3, \tag{23}$$

where μ is the dynamic viscosity and δ_{ij} is the Kronecker delta.

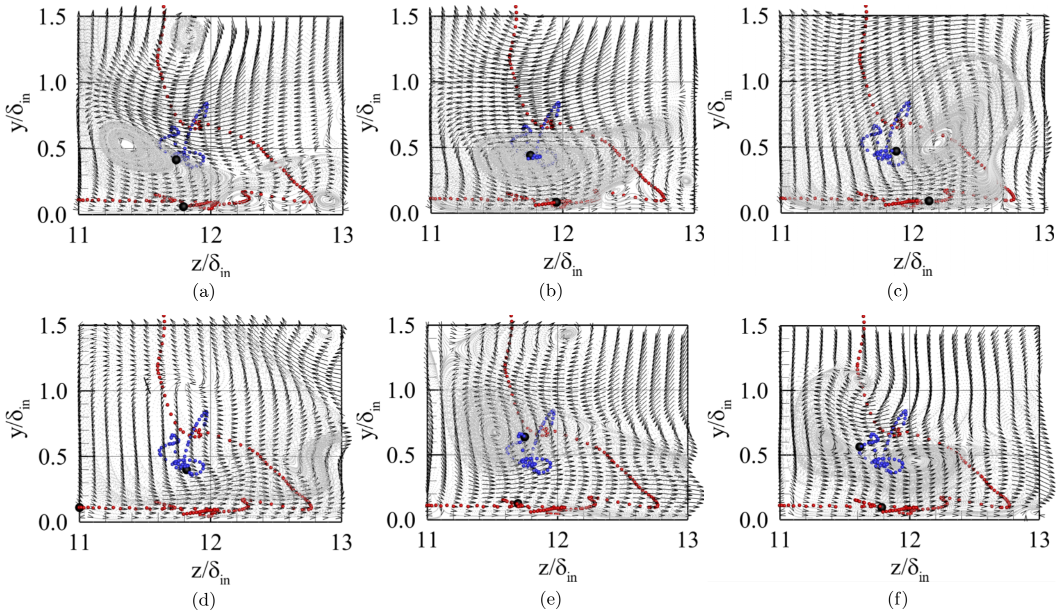


FIG. 28. Streamwise array of cross-sectional streamlines and velocity vectors obtained from the y and z components, i.e., (v, w), within x planes at $t = t_1$ for the typical case of a spiral sink. Panels (a)–(f) show each streamwise location in the range $x = 543.49\delta_{\text{in}} - 547.05\delta_{\text{in}}$. Red and blue points are the trajectories of the skeleton elements on each cross section. Black points are the intersections of the skeletons with each cross section. (a) $x = 543.49\delta_{\text{in}}$. (b) $x = 544.20\delta_{\text{in}}$. (c) $x = 544.92\delta_{\text{in}}$. (d) $x = 545.63\delta_{\text{in}}$. (e) $x = 546.34\delta_{\text{in}}$. (f) $x = 547.05\delta_{\text{in}}$.

The equation for the helicity density is given as follows:

$$\frac{\partial}{\partial t}(\mathbf{u}\cdot\omega) = \sum_{n=1}^{9} T_n, \quad T_1 = -\{(\mathbf{u}\cdot\nabla)\omega\}\cdot\mathbf{u}, \quad T_2 = (\omega\cdot\nabla)\mathbf{u}\cdot\mathbf{u},$$
$$T_3 = -\{(\nabla\cdot\mathbf{u})\omega\}\cdot\mathbf{u}, \quad T_4 = \left\{\nabla p \times \nabla\left(\frac{1}{\rho}\right)\right\}\cdot\mathbf{u}, \quad T_5 = \left\{\nabla\times\left(\frac{1}{\rho}\nabla\cdot\tau\right)\right\}\cdot\mathbf{u},$$
$$T_6 = -\frac{1}{\rho}(\nabla p\cdot\omega), \quad T_7 = -\nabla\left(\frac{u^2}{2}\right)\cdot\omega, \quad T_8 = \frac{1}{\rho}(\nabla\cdot\tau)\cdot\omega. \tag{24}$$

The dominant mathematical terms are evaluated at each element of the superclusters. For this purpose, a local neighborhood $W_q (q = 1, \ldots, N_s)$ is generated for each element. To identify the dominant terms in each W_q, the RHSs of the equations are decomposed into terms 1–5 for the enstrophy equation and terms 1–8 for the helicity equation. The terms of T_q that have large absolute values (LAVs) in W_q are then selected, and the minimum combination (MC) of the LAV terms necessary to adequately approximate the RHS is determined.

A rectangular-shaped neighborhood, specified in terms of mesh indices $W_q = [j_q - \Delta, j_q + \Delta] \times [k_q - \Delta, k_q + \Delta] \times [l_q - \Delta, l_q + \Delta]$, is generated for the supercluster element closest to mesh point (j_q, k_q, l_q). Here, Δ is the half width of W_q in the index space. In this paper, $\Delta = 3$.

Not all terms on the RHS are important. First, LAV terms are selected. The norm of an RHS term T_n, $n = 1, 2, \ldots$, in W_q (denoted as $||T_n||_{W_q}$) is defined as the absolute value of the T_n averaged over W_q, i.e., $||T_n||_{W_q} \equiv |\overline{T}_{n,q}|$, $\overline{T}_{n,q} \equiv \sum_{(j,k,l)\in W_q} T_n(j,k,l)/\text{No. } W_q$. The LAV terms are then defined as RHS terms T_n such that $||T_n||_{W_q} \geqslant \epsilon_T \max_n\{||T_n||_{W_q}\}$. In this paper, $\epsilon_T = 0.3$.

The selected LAV terms do not necessarily determine the variation of the RHS in W_q because some terms may cancel out, that is, there could be a possibility that only a small subset of the LAV terms truly control the variation of the RHS. Thus, the MC of LAV terms that has a high correlation with the RHS in W_q is finally selected. The reasons for selecting the LAV terms first are to exclude negligible terms and to focus on a combination of the dominant terms. Flow variation in W_q is driven by competition within the MC. Because the MC is selected only within the neighborhood W_q belonging to a supercluster element, the MC will be different for different supercluster elements. The following sums of the RHS terms are defined:

$$\mathrm{RHS}^{\mathrm{EXACT}} = \sum_{q=1}^{ex\max} T_q, \quad \mathrm{RHS}^{\mathrm{LAV}} = \sum_{r=1}^{\mathrm{LAVmax}} T_{\sigma(r)}, \quad \mathrm{RHS}^{\mathrm{MC}} = \sum_{s=1}^{\mathrm{MCmax}} T_{\tau(s)}. \tag{25}$$

Here, $\mathrm{RHS}^{\mathrm{EXACT}}$ is the sum of all T_q terms, $ex\max$ is a constant (taking a value of 5 or 8 for the enstrophy or helicity equation, respectively), $\mathrm{RHS}^{\mathrm{LAV}}$ is the approximation of $\mathrm{RHS}^{\mathrm{EXACT}}$ given by all LAV terms, and $\mathrm{RHS}^{\mathrm{MC}}$ is the approximation of $\mathrm{RHS}^{\mathrm{LAV}}$ given by the MC of the LAV terms. Here, the MC is defined as the smallest subset of LAV terms $T_{\mathrm{MC}} \equiv \{\tau(p), p = 1, \ldots, \mathrm{MCmax}\}$ satisfying the following inequality:

$$\mathrm{corr}(\mathrm{RHS}^{\mathrm{MC}}, \mathrm{RHS}^{\mathrm{EXACT}}) \geqslant \epsilon_C \mathrm{corr}(\mathrm{RHS}^{\mathrm{LAV}}, \mathrm{RHS}^{\mathrm{EXACT}}), \tag{26}$$

where $\mathrm{corr}(\mathbf{a}, \mathbf{b})$ is the cross-correlation between $\mathbf{a}, \mathbf{b}$, which are one-dimensionalized arrays of quantities within W_q. In this paper, $\epsilon_C = 0.8$.

The combination of $\mathrm{RHS}^{\mathrm{MC}}$ and the time-derivative term constitutes a reduced-order or data-compressed model for the system dynamics within W_q. Denoting the set of T_q terms, all LAV terms, and the MC of the LAV terms in W_q as $\sum_{T,W_q}$, $\sum_{\mathrm{LAV},W_q}$, $\sum_{\mathrm{MC},W_q}$, respectively, the hierarchical relationship $\sum_{T,W_q} \supset \sum_{\mathrm{LAV},W_q} \supset \sum_{\mathrm{MC},W_q}$ holds.

Figure 29 shows the histograms of the number of protoclusters at which each combination of the mathematical terms becomes the MC for the enstrophy and helicity equations. Figures 30 and 31 show the paired skeletons for near-wall vortical regions with $\omega_x > \omega_{x,0}$ and the locations in the near-wall superclusters where each combination of the terms is the MC for the enstrophy and helicity equations, respectively, at $t = t_1$.

For the enstrophy equations, combinations $\{1\}, \{2\}, \{5\}, \{2, 5\}, \{1, 2, 5\}$ become the MC at $85\%, 8.3\%, 2.9\%, 2.3\%$, and 0.99% of the locations of protoclusters, respectively (see Fig. 29). Terms 3 and 4 are negligible in the superclusters. They correspond to the dilatation effect, i.e., compressibility, and the baroclinic production of vorticity when temperature and entropy gradients are not parallel [67]. From Fig. 30, term 1 is the MC in almost all regions of the near-wall superclusters. Term 2 becomes the MC around locations where superclusters begin to lift up from the wall. Term 5, i.e., the viscous term, only becomes the MC near the upstream origins of the skeletons. In the helicity equation, the MCs $\{1\}, \{2\}, \{6\}, \{7\}, \{2, 7\}, \{6, 8\}$ account for $92.1\%, 0.78\%, 0.96\%, 2.1\%, 0.63\%$, and 0.64% of locations, respectively. Compared to the enstrophy equation, the contrast between MCs other than $\{1\}$ is weak because $\{1\}$ is predominant. This result is compatible with the fact that the helicity is usually treated in a Lagrangian manner [68]. Terms 3 and 4 are negligible similarly to the case of the enstrophy equation. From Fig. 31, term 1 is the MC in almost all regions of the near-wall superclusters. Term 6 becomes the MC around locations where the near-wall superclusters bend in the spanwise direction. The combination of terms 6 and 8 becomes the MC around the region where y is less than $0.1\delta_{in}$.

In these analyses, the divergence of velocity, i.e., $\nabla \cdot \mathbf{u}$, which generally reflects the compressibility effect, does not appear in the MCs. However, this term is expected to be large at some point within a flow because sound propagation is prevalent at the Mach number as large as 0.5 [69]. One possible reason for the disparity is that only those regions with strong vorticity have been analyzed in this paper, and the region where $\nabla \cdot \mathbf{u}$ is large lies outside the vortical regions. This inference is supported by the Helmholtz-Hodge decomposition, which is explained in Appendix B.

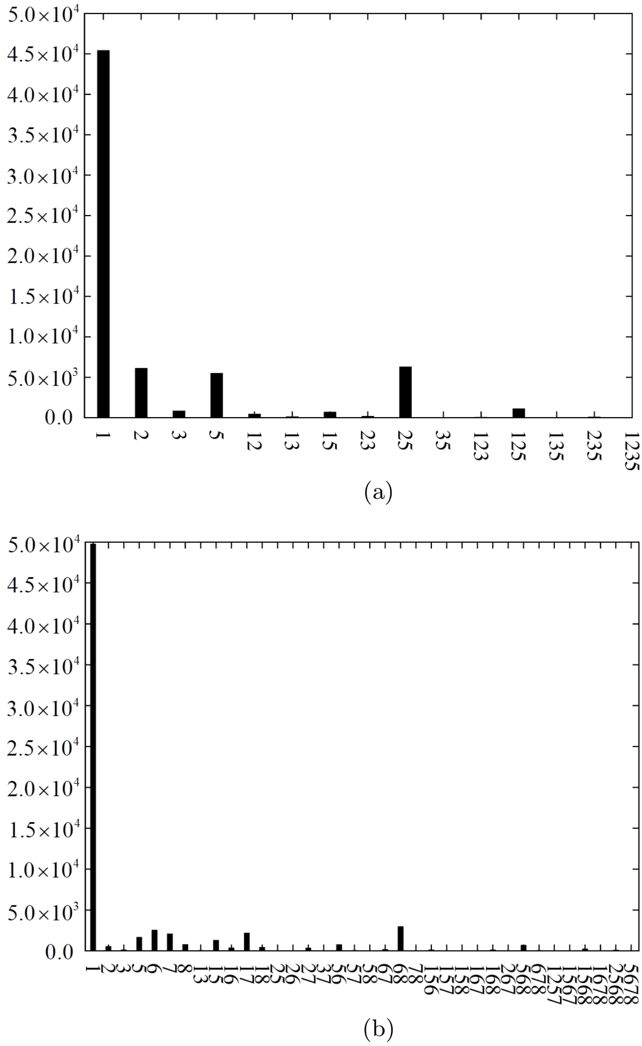


FIG. 29. Histograms of the number of protoclusters at which each combination of mathematical terms becomes the MC at $t = t_1$. (a) and (b) show the cases for the enstrophy and helicity equations, respectively. (a) Enstrophy equation. (b) Helicity equation.

The present method is also considered effective in extracting the region of large $\nabla \cdot \mathbf{u}$, and a detailed investigation into the region will be reported in the future.

In terms of the MC, supercluster elements are divided into subclusters by considering the subclusters as equivalent classes under the relation of possessing a same MC. This means that subclusters with a common MC belonging to different superclusters can be extracted. In mathematical notation, such subclusters can be represented as $E_\alpha|_{S_A}$ and $F_\beta|_{S_B}$ if we think of two superclusters S_A and S_B and their respective subclusters $\bigcup_k E_k = S_A$, $\bigcup_l F_l = S_B$. Here, $\cdot|_{S_A}$ explicitly shows that a subcluster is restricted (belongs) to S_A. The present method enables us to automatically extract $E_\alpha|_{S_A}$ and $F_\beta|_{S_B}$ that have a common MC, which are difficult to extract directly from the original set of protoclusters shown in Eq. (1). A set consisting of these subclusters, i.e., $\{E_\alpha|_{S_A}, F_\beta|_{S_B}\}$ becomes a unique ensemble, as mentioned in Sec. IV F.

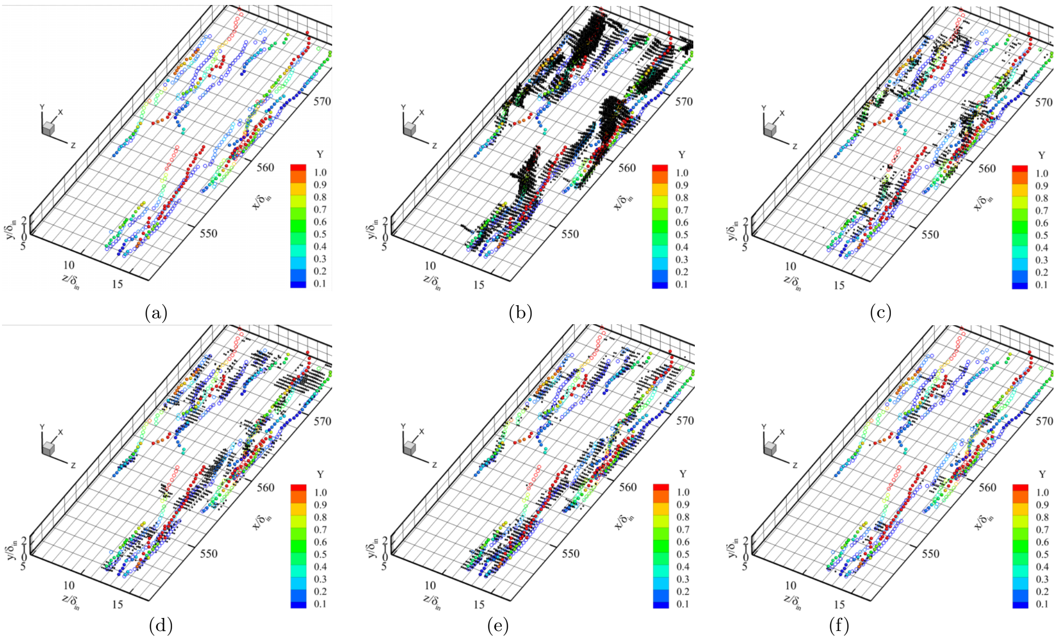


FIG. 30. Paired skeletons for near-wall vortical regions with $\omega_x > \omega_{x,0}$ and the locations in the near-wall superclusters where each combination of the terms is the MC for the enstrophy equation at $t = t_1$. Panels (a)–(f) show different combinations. (a) Baseline. (b) Term 1. (c) Term 2. (d) Term 5. (e) Terms 2 and 5. (f) Terms 1, 2, and 5.

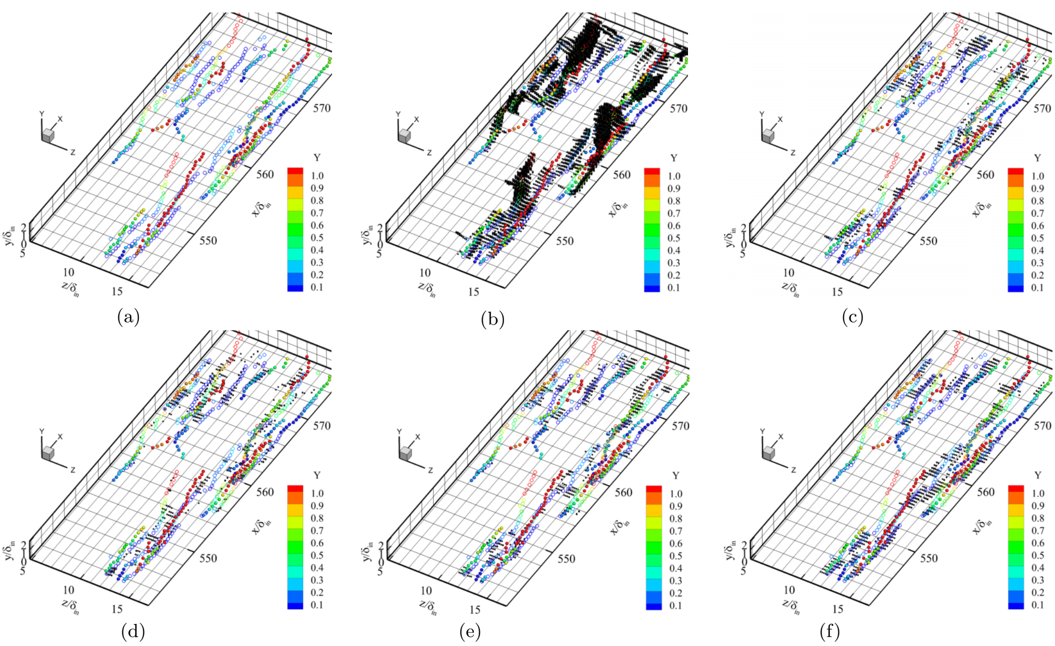


FIG. 31. Paired skeletons for near-wall vortical regions with $\omega_x > \omega_{x,0}$ and the locations in the near-wall superclusters where each combination of the terms is the MC for the helicity equation at $t = t_1$. (a)–(f) show different combinations. (a) Baseline. (b) Term 1. (c) Term 6. (d) Term 7. (e) Terms 1 and 7. (f) Terms 6 and 8.

The above procedure can be generalized to the elements of nonlowest p clusters. The present clustering method successfully restricts the evaluation of dominant subsystems to p clusters and thereby clarifies spatially nonuniform local dynamics that drive the deformation of the p clusters. A subcluster of a p cluster with a specific MC can be used to extract p clusters that have common local dynamics. Additionally, p-cluster elements with a specific subsystem can be used to extract specific p clusters upper or lower than the p cluster through the inheritance property. The lower-order subclusters can also be constructed. Collecting the subclusters from different p clusters, unique ensembles and statistical quantities can be constructed.

VI. CONCLUSIONS

A hierarchical clustering method of vortical regions is proposed. The clusters of the regions of large streamwise vorticity are extracted at each streamwise location and connected in the streamwise direction forming extended clusters named p clusters. The multiple correspondence is decomposed into a single correspondence. Global single correspondence is achieved by the backward subclustering and iteration through the domain. By connecting superclusters over successive times, hyperclusters are generated. At each supercluster element, the mathematical terms in the enstrophy and helicity equations are evaluated, and the MCs of dominant mathematical terms approximating the equations are obtained. As a result, the elements are clustered within each supercluster, which enables to extract subclusters having a common MC of the dominant mathematical terms from different superclusters. Compared with the conventional vortex analysis methods, the present method has the advantages of the following: (i) Individual connected strong vortical regions or partial blocks of the regions can be extracted as data objects with hierarchical properties from an instantaneous field. The spatial distribution and temporal evolution of them can be visualized and tracked. (ii) Balances between mathematical terms that characterize local dynamics can be evaluated at points belonging to the objects. Then, equivalence between the subsets of the objects can be mechanically evaluated by an algorithm. The proposed method is applied to the late stage of K-type transitions. The spatial and temporal distribution of p clusters, the typical origin of near-wall vortical regions, functional clustering of supercluster elements, the construction of ensemble-averaged quantities and the extension of the present clustering method to omnidirection are discussed.

Linkage between vortical regions, its spatial distribution, and the time variation of the vortical regions in the computational domain are successfully extracted by the proposed method.

Due to the realization of the individual identification of p clusters with unique identification numbers, the algorithmic analysis of its dynamics and statistical properties is realized. By calculus on p clusters and the extraction of supercluster pairs or flow topology around them, similar categories of p clusters can be algorithmically extracted.

In the majority of cases, the near-wall origins of superclusters are attributable to the circulation of off-wall superclusters and the nonslip condition at the wall. The generation of the near-wall vortical regions is correlated with the lift up of its associated off-wall vortical regions. Some vortical regions lift up to the edge of the boundary layer, and other vortical regions remain near the wall or come back to the near-wall region after lift up.

The present method enables the evaluation of dominant subsystems on p clusters. Terms related to the dilatation effect, the production of vorticity when temperature and entropy gradients are not parallel, and the Lamb vector are negligible in the present case. The performance of the functional clustering depends on selected equations because of its separatability between RHS terms.

The p-cluster elements with a specific subsystem can be used to extract specific p clusters upper or lower than the p cluster. The subclustering of p-cluster elements, which is different from the fundamental clusters, can also be conducted. The present clustering method is capable of establishing statistical relations using uniquely generated ensembles.

The present unidirectional method can be extended to the clustering of omnidirectional vortical regions in two ways. One is to repeat the x-directional clustering algorithm to the y and z directions. The other is the clustering based on a vortex axis. Both methods are successful in constructing superclusters and skeletons pertaining to the vortex ring and the vortical regions of the late stage of transition.

ACKNOWLEDGMENTS

This study is a result of an IMI Joint Research Programs on development of data mining methods for vortex dynamics in laminar-turbulent transition. It is also supported by the Institute of Statistical Mathematics (ISM) Cooperative Research Program 2019 No. ISM-CRP2019 and JSPS KAKENHI Grant No. JP19K03672. Computations were partly conducted using a supercomputer system at the Japan Aerospace Exploration Agency (JAXA-JSS2) and ISM.

APPENDIX A: FLOW TOPOLOGY OF A VECTOR FIELD

To categorize vector fields in terms of flow topology, a velocity vector field (u, v) is formulated as follows:

$$X' = AX, \quad X = (x, y) \in \mathbf{R}^2. \tag{A1}$$

Here, the prime($'$) on the left-hand side denotes differentiation with time t. A is a 2×2 real matrix. Suppose the system has an equilibrium point of $X_0 = (0, 0)$. The eigenvalues of A are real or complex. Flow topology is categorized as follows [70]:

(i) When λ is real and $\lambda_1 \neq \lambda_2$,

$$\text{saddle} : \lambda_1 < 0 < \lambda_2, \quad \text{sink} : \lambda_1 \leqslant \lambda_2 < 0, \quad \text{source} : 0 < \lambda_1 \leqslant \lambda_2. \tag{A2}$$

(ii) When λ is complex, $\lambda = \alpha \pm i\beta$,

$$\text{center} : \alpha = 0, \quad \text{spiral sink} : \alpha < 0, \quad \text{spiral source} : \alpha > 0. \tag{A3}$$

APPENDIX B: HELMHOLTZ-HODGE DECOMPOSITION

By the Helmholtz-Hodge decomposion [71], the fluid velocity $\mathbf{u}$ is uniquely decomposed into the irrotational and the solenoidal parts along with the gauge condition as

$$\mathbf{u} = \nabla\phi + \nabla \times \mathbf{A}, \quad \nabla \cdot \mathbf{A} = 0. \tag{B1}$$

Here, ϕ and $\mathbf{A}$ are a scalar potential and a vector potential, respectively.

---

[1] M. M. Rai and P. Moin, Direct numerical simulation of transition and turbulence in a spatially evolving boundary layer, J. Comput. Phys. **109**, 169 (1993).
[2] F. Ducros, P. Comte, and M. Lesieur, Large-eddy simulation of transition to turbulence in a boundary layer developing spatially over a flat plate, J. Fluid Mech. **326**, 1 (1996).
[3] X. Wu and P. Moin, Direct numerical simulation of turbulence in a nominally zero-pressure-gradient flat-plate boundary layer, J. Fluid Mech. **630**, 5 (2009).
[4] C. Liu and L. Chen, Parallel DNS for vortex structure of late stages of flow transition, Comput. Fluids **45**, 129 (2011).
[5] F. M. White, *Viscous Fluid Flow*, 3rd ed. (McGraw-Hill, New York, 2006).
[6] M. Jiang, R. Machiraju, and D. Thompson, Detection and visualization of vortices, in *The Visualization Handbook*, 1st ed., edited by C. D. Hansen and C. R. Johnson (Academic Press, Massachusetts, 2004), pp. 295–309.
[7] T. Günther and H. Theisel, The state of the art in vortex extraction, Comput. Graphics Forum **37**, 149 (2018).
[8] B. A. Singer and D. C. Banks, A predictor-corrector scheme for vortex identification, NASA-CR. No. 194882 (NASA, Langley Research Center, VA, 1994).

[9] D. Sujudi and P. Haimes, Identification of swirling flow in 3-D vector fields, *Proceedings of the 12th Computational Fluid Dynamics Conference, San Diego, CA, USA* (AIAA, 1995), paper AIAA-95-1715.
[10] S. Kida and H. Miura, Identification and analysis of vortical structures, Eur. J. Mech. B/Fluids **17**, 471 (1998).
[11] H. Miura and S. Kida, Identification of tubular vortices in turbulence, J. Phys. Soc. Jpn. **66**, 1331 (1997).
[12] K. Nakayama, Topological features and properties associated with development/decay of vortices in isotropic homogeneous turbulence, Phys. Rev. Fluids **2**, 014701 (2017).
[13] K. Nakayama and H. Hasegawa, Relationships between eigen-vortical-axis line and vorticity line, Fluid Dyn. Res. **50**, 011417 (2018).
[14] R. Peikert and M. Roth, The "parallel vectors" operator—a vector field visualization primitive, in *Proceedings of the Visualization '99 (Cat. No.99CB37067)* (IEEE, Piscataway, NJ, 1999), pp. 263–270.
[15] A. V. Gelder and A. Pang, Using PVsolve to analyze and locate positions of parallel vectors, IEEE Trans. Vis. Comput. Graph **15**, 682 (2009).
[16] T. Schafhitzel, J. E. Vollrath, J. P. Gois, D. Weiskopf, A. Castelo, and T. Ertl, Topology-preserving λ2-based vortex core line detection for flow visualization, Comput. Graphics Forum (Eurovis 2008) **27**, 1023 (2008).
[17] C. Liu, Y. Gao, X. Dong, Y. Wang, J. Liu, Y. Zhang, X. Cai, and N. Gui, Third generation of vortex identification methods: Omega and Liutex/Rortex based systems, J. Hydrodynamics **31**, 205 (2019).
[18] Y. Gao, J. Liu, Y. Yu, and C. Liu, A Liutex based definition of vortex rotation axis line, J. Hydrodynamics **31**, 445 (2019).
[19] L. Zhu and L. Xi, Vortex axis tracking by iterative propagation (VATIP): A method for analysing three-dimensional turbulent structures, J. Fluid Mech. **866**, 169 (2019).
[20] J. Jeong, F. Hussain, W. Schoppa, and J. Kim, Coherent structures near the wall in a turbulent channel flow, J. Fluid Mech. **332**, 185 (1997).
[21] B. S. Everitt and S. Landau, *Cluster Analysis*, 5th ed. (Wiley, New Jersey, 2011).
[22] S. Wierzchoń and M. Kłopotek, *Modern Algorithms of Cluster Analysis* (Springer, Cham, Switzerland, 2018).
[23] G. Haller, Lagrangian coherent structures, Annu. Rev. Fluid Mech. **47**, 137 (2015).
[24] J. Wilroy, R. A. Wahidi, and A. Lang, Effect of butterfly-scale-inspired surface patterning on the leading edge vortex growth, Fluid Dyn. Res. **50**, 045505 (2018).
[25] F. Zhao, D. L. S. Hung, and S. Wu, K-means clustering-driven detection of time-resolved vortex patterns and cyclic variations inside a direct injection engine, Appl. Therm. Eng. **180**, 115810 (2020).
[26] L. Deng, Y. Wang, C. Chen, Y. Liu, F. Wang, and J. Liu, A clustering-based approach to vortex extraction, J. Vis. **23**, 459 (2020).
[27] A. McCallum, K. Nigam, and L. H. Ungar, Efficient clustering of high-dimensional data sets with application to reference matching, in *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '00* (ACM, New York, 2000), pp. 169–178.
[28] G. Haller, A. Hadjighasem, M. Farazmand, and F. Huhn, Defining coherent vortices objectively from the vorticity, J. Fluid Mech. **795**, 136 (2016).
[29] S. Ferrari, Y. Hu, C. Morton, and R. J. Martinuzzi, Visualizing three-dimensional vortex shedding through evolution surface clusters, J. Vis. **23**, 17 (2020).
[30] K. Matsuura, Clustering analysis of vortices appearing in the process of laminar-turbulent transition (Turbulence and transition: structures, multi-scale, models), RIMS Kôkyûroku **2117**, 21 (2019).
[31] K. Matsuura, Algorithmic exploration of dominant terms around hairpin vortices generated during boundary-layer transition under free-stream turbulence, Int. J. Environ. Impacts **3**, 69 (2020).
[32] K. Matsuura and Y. Fukumoto, Development of a vortex clustering algorithm in a transitional boundary-layer, in *Proceedings of the 25th Conference of Japan Soc. Fluid Mech., Chushikoku & Kyushu Branch* (The Japan Society of Fluid Mechanics, Tokyo, Japan, 2020), p. 1.
[33] K. Matsuura and Y. Fukumoto, A study on vortex clustering in the late-stage of boundary-layer transition, in *Proc. Ann. Conf. Jpn. Soc. Fluid Mech.* (The Japan Society of Fluid Mechanics, Tokyo, Japan, 2020), p. 1.

[34] T. Cebeci and A. M. O. Smith, *Analysis of Turbulent Boundary Layer* (Academic Press, Massachusetts, 1974).
[35] J. Cousteix, *Turbulence et Couche Limite* (Cepadues, Toulouse, France, 1989), p. 625.
[36] K. Matsuura and C. Kato, Large-eddy simulation of compressible transitional flows in a low-pressure turbine cascade, AIAA J. **45**, 442 (2007).
[37] K. Matsuura and M. Nakano, A throttling mechanism sustaining a hole tone feedback system at very low Mach numbers, J. Fluid Mech. **710**, 569 (2012).
[38] K. Matsuura, M. Matsui, and N. Tani, Effects of free-stream turbulence on the global pressure fluctuation of compressible transitional flows in a low-pressure turbine cascade, Int. J. Num. Meth. Heat Fluid Flow **28**, 1187 (2018).
[39] B. P. Epps, Review of vortex identification methods, in *55th AIAA Aerospace Sciences Meeting 9–13 January 2017, Grapevine, Texas* (AIAA, Virginia, 2017), pp. 1–22.
[40] S. K. Robinson, Coherent motions in the turbulent boundary layer, Annu. Rev. Fluid Mech. **23**, 601 (1991).
[41] J. M. Hamilton, J. Kim, and F. Waleffe, Regeneration mechanisms of near-wall turbulence structures, J. Fluid Mech. **287**, 317 (1995).
[42] G. Kawahara, Theoretical interpretation of coherent structures in near-wall turbulence, Fluid Dyn. Res. **41**, 064001 (2009).
[43] K. Matsuura, Direct numerical simulation of a straight vortex tube in a laminar boundary-layer flow, Int. J. Comp. Meth. and Exp. Meas. **4**, 474 (2016).
[44] J. Jimenez, Coherent structures in wall-bounded turbulence, J. Fluid Mech. **842**, P1 (2018).
[45] U. V. Luxburg, A tutorial on spectral clustering, Stat. Comput. **17**, 395 (2007).
[46] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 3rd ed. (The MIT Press, Massachusetts, 2009).
[47] T. McLoughlin, R. S. Laramee, and E. Zhang, Constructing streak surfaces for 3D unsteady vector fields, in *Proceedings of the 26th Spring Conference on Computer Graphics, SCCG '10* (ACM, New York, 2010), pp. 17–26.
[48] H. Krishnan, C. Garth, and K. I. Joy, Time and streak surfaces for flow visualization in large time-varying data sets, IEEE Trans. Visualization Comput. Graphics **15**, 1267 (2009).
[49] T. Schafhitzel, E. Tejada, D. Weiskopf, and T. Ertl, Point-based stream surfaces and path surfaces, in *Proceedings of Graphics Interface, GI '07* (ACM, New York, 2007), pp. 289–296.
[50] C. Garth, X. Tricoche, T. Salzbrunn, T. Bobach, and G. Scheuermann, Surface techniques for vortex visualization, in *Proceedings of the Sixth Joint Eurographics-IEEE TCVG Conference on Visualization* (Eurographics Association, Goslar, Germany, 2004), pp. 155–164.
[51] S. K. Ueng, K. Sikorski, and K. L. Ma, Fast algorithms for visualizing fluid motion in steady flow on unstructured grids, in *Proceedings of the 6th IEEE Visualization Conference* (IEEE Computer Society, Washington, DC, 1995), pp. 313–320.
[52] R. F. Blackwelder and R. E. Kaplan, On the wall structure of the turbulent boundary layer, J. Fluid Mech. **76**, 89 (1976).
[53] J. Kim, Turbulence structures associated with the bursting event, Phys. Fluids **28**, 52 (1985).
[54] W. W. Willmarth and S. S. Lu, Structure of the Reynolds stress near the wall, J. Fluid Mech. **55**, 65 (1972).
[55] M. S. Chong, A. E. Perry, and B. J. Cantwell, A general classification of three-dimensional flow fields, Phys. Fluids **2**, 765 (1990).
[56] P. Holmes, J. L. Lumley, and G. Berkooz, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry* (Cambrige University Press, Cambridge, UK, 1998).
[57] L. Sirovich and J. D. Rodriguez, Coherent structures and chaos: A model problem, Phys. Lett. A **120**, 211 (1987).
[58] P. J. Schmid, Dynamic mode decomposition of numerical and experimental data, J. Fluid Mech. **656**, 5 (2010).
[59] R. J. Adrian, Conditional eddies in isotropic turbulence, Phys. Fluids **22**, 2065 (1979).
[60] M. G. Meena and K. Taira, Identifying vortical network connectors for turbulent flow modification, J. Fluid Mech. **915**, A12 (2021).

[61] D. Karrasch and G. Haller, Do finite-size lyapunov exponents detect coherent structures? Chaos **23**, 043126 (2013).

[62] M. Newman, *Networks*, 2nd ed. (Oxford University Press, Oxford, UK, 2018).

[63] M. E. J. Newman and M. Girvan, Finding and evaluating community structure in networks, Phys. Rev. E **69**, 026113 (2004).

[64] M. E. J. Newman, Analysis of weighted networks, Phys. Rev. E **70**, 056131 (2004).

[65] C. B. Lee and J. Z. Wu, Transition in wall-bounded flows, Appl. Mech. Rev. **61**, 030802 (2008).

[66] C. Hirsch, *Numerical Computation of Internal & External Flows*, 2nd ed. (Butterworth-Heinemann, Oxford, UK, 2007).

[67] M. S. Howe, *Acoustics of Fluid-Structure Interactions*, 1st ed. (Cambridge University Press, Cambridge, UK, 1998).

[68] H. K. Moffatt, Structure and stability of solutions of the Euler equations: A lagrangian approach, Phil. Trans. R. Soc. Lond. A **333**, 321 (1990).

[69] A. J. Smits and J.-P. Dussauge, *Turbulent Shear Layers in Supersonic Flow*, 2nd ed. (Springer, Berlin, Germany, 2006).

[70] M. W. Hirsch, S. Smale, and R. L. Devaney, *Differential Equations, Dynamical Systems, and an Introduction to Chaos*, 3rd ed. (Academic Press, Massachusetts, 2012).

[71] A. Chorin and J. Marsden, *A Mathematical Introduction to Fluid Mechanics*, 3rd ed. (Springer, Berlin, Germany, 1992).