Visual Analysis of Spatio-temporal Relations of Pairwise Attributes in Unsteady Flow

Marzieh Berenjkoub, Lei Zhang, Rodolfo Ostilla Monico, Robert S. Laramee, and Guoning Chen

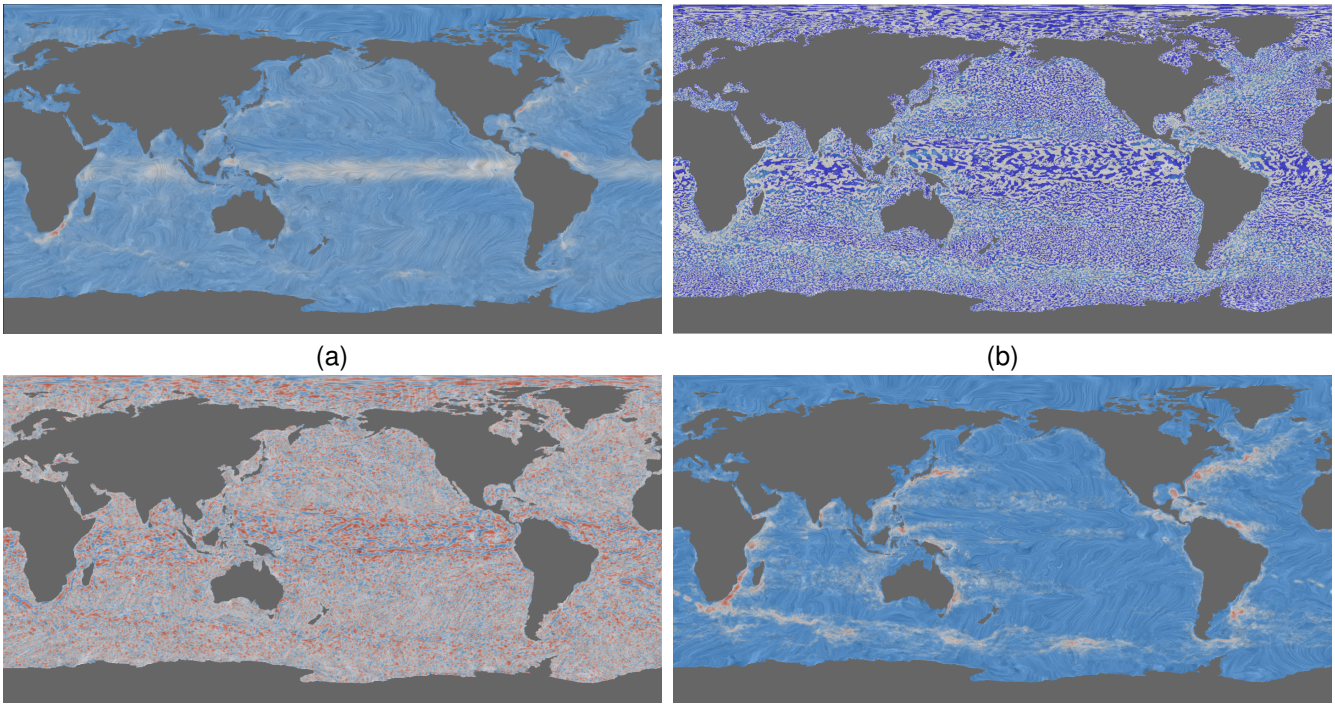


Fig. 1: The ST_LCC linear correlation (c) and ST_MI dependency (d) of the acceleration (a) and Q (b) of the large ocean simulation data with a spatio-temporal kernel size 5×50. The blue-white-red color scheme is used to highlight the different characteristics of attribute correlation and dependency. The ST_LCC result (c) highlights a string of strong vortices (or eddies) around the equator, while the MI result (d) emphasizes the strong offshore currents on the east coast of the individual continents.

Abstract—
Despite significant advances in the analysis and visualization of unsteady flow, the interpretation of it's behavior still remains a challenge. In this work, we focus on the linear correlation and non-linear dependency of different physical attributes of unsteady flows to aid their study from a new perspective. Specifically, we extend the existing spatial correlation quantification, i.e. the Local Correlation Coefficient (LCC), to the spatio-temporal domain to study the correlation of attribute-pairs from both the Eulerian and Lagrangian views. To study the dependency among attributes, which need not be linear, we extend and compute the mutual information (MI) among attributes over time. To help visualize and interpret the derived correlation and dependency among attributes associated with a particle, we encode the correlation and dependency values on individual pathlines. Finally, to utilize the correlation and MI computation results to identify regions with interesting flow behavior, we propose a segmentation strategy of the flow domain based on the ranking of the strength of the attributes relations. We have applied our correlation and dependency metrics to a number of 2D and 3D unsteady flows with varying spatio-temporal kernel sizes to demonstrate and assess their effectiveness.

Index Terms—Unsteady flow, correlation study, mutual information

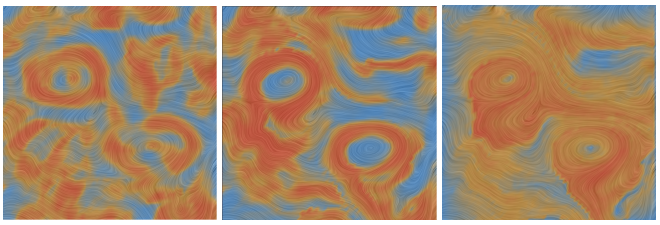
1 Introduction

Despite significant advances in the development of the analysis and visualization techniques for unsteady flow, interpretation of its behavior remains a challenging task. Existing methods either focus on the transport behavior of the unsteady flow by characterizing the behavior of particles and their trajectories (i.e., pathlines) [1–3], or study the features defined under a given coordinate system [4]. Both groups of techniques provide valuable insight into the dynamics of unsteady flow.

In parallel, various local characteristics are widely used in the fluid mechanic community for the study of physical properties of the fluid. In the flow visualization community, local properties either have been utilized for the definition of critical points in a Lagrangian context [5] or have facilitated the understanding of their transport characteristics [6].

- *Marzieh Berenjkoub, Lei Zhang, Rodolfo Ostilla Monico and Guoning Chen are with the University of Houston. E-mails: mberenjkoub@uh.edu, zhanglei.later@gmail.com, rostilla@central.uh.edu, and gchen16@uh.edu.*
- *Robert S. Laramee is with Swansea University. E-mail:r.s.laramee@swansea.ac.uk*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx



(a) (b) (c)

Fig. 2: Different correlation measurements of a region of the ocean simulation that contains eddies. (a) Spatial LCC, (b) Spatio-temporal (Lagrangian) LCC, and (c) mutual information (MI) (Lagrangian) between the *Acceleration* and *Determinant* of Jacobian matrix, respectively. For the first two results, the color shows the correlation characteristics, blue indicates negatively correlated and red positively correlated. For the MI result (c) blue indicates lower dependency and red indicates higher dependency.

Important flow features, such as vortices, can also be identified and characterized by certain local flow attributes [7,8]. However, there lacks an effective way to study the co-varying behaviors and dependencies between various flow attributes over time, which may provide additional insight into the behavior of unsteady flows.

To fill the above gap, this work makes the following contributions. First, we extend the Local Correlation Coefficient (LCC) introduced by Sauber et al. [9] to the spatio-temporal domain and define a spatio-temporal LCC (or ST_LCC) to reveal the time-varying linear correlation of attributes derived from unsteady flow. Figure 2 compares our ST_LCC result (b) with the existing spatial LCC result (a) of two attributes on an ocean current simulation (detailed information is provided in Section 5). From this comparison, we see the correlation field computed with our approach captures more complete structures of the two main vortices in this ocean region. In addition, we introduce a vector-based LCC and its spatio-temporal equivalence (ST_VEC_LCC) for the correlation analysis of the gradient vectors of pairwise attributes over time. This vector-based LCC yields complementary information about the linear correlation among pairs of attributes by focusing on their gradient behavior (see Figure 3c and Figure 3d for an example).

Second, to capture the non-linear dependency between attributes, we extend and compute the mutual information (MI) of pairwise attributes. MI was first applied for image registration [10] then later for uncertainty reduction in visualization [11]. Recently, it has been used for statistical analysis in flow visualization [12]. Compared to the previous work, we provide a systematic discussion on how to apply MI for measuring time-varying dependency of flow attributes, and evaluate its effectiveness in revealing certain characteristics of unsteady flows (e.g., vortices). In contrast to the ST_LCC result shown in Figure 2(b), the MI field (c) in this example shows that the two attributes are highly dependent on each other in a larger region beyond the boundaries of vortices shown by ST_LCC.

Third, we apply the proposed ST_LCC, ST_VEC_LCC, and MI for various time-varying attributes measured in both the Eulerian (i.e., at fixed position) and Lagrangian (i.e., along moving particles) views for the first time. Our results show that both views provide meaningful insight into pairwise attribute relations.

Finally, we also offer a simple yet intuitive way to visualize the correlation and dependency information from the Lagrangian view, which is difficult to show in a static image. In particular, we propose a new pathline visualization to convey the varying correlation between attributes of interest over the course of the particle advection, which provides interesting insight into the detailed particle behavior that has not been revealed. We provide a number of use cases with analytic and simulated unsteady flow to demonstrate the utility of the proposed correlation quantification and visualization. Although we primarily focus on vortices – an important flow structure in fluid mechanics, and effectively reveal detailed vortex structure over time that is non trivial to show using a single attribute at a time (Figures 9 and 12), our method can potentially apply to the study of other flow behaviors.

2 Related Work

A large amount of work has been proposed to address the visualization and analysis of unsteady flow [13, 14]. These techniques either extract the structural information or the local dynamics of the flow.

Structural analysis of unsteady flow aims to study the transport behavior of the flow and identify the boundaries of different regions, such that the particles within each region exhibit similar temporal transport behavior [13]. There are various methods to define and compute the structure of unsteady flow, including topological feature tracking [15–17] based on the bifurcation theory, and pathline-based segmentation [18]. Nonetheless, the most successful and popular method is the computation of *Finite Time Lyapunov Exponent (FTLE)* fields [1, 19–21], whose ridges (i.e., *Lagrangian Coherent structures (LCS))* are extracted as the boundaries of different regions with distinct transport behavior. This method inspires a new direction of defining unsteady vector field topology based on streak lines/surfaces [22–24].

In addition to flow separation behavior in the transport of unsteady flow, the rotational behavior also leads to an important coherent structure, i.e., vortices [25–27]. In contrast to the transport structure, the vortical structure can be suppressed by a strong translational flow [28]. To reveal the vortex structure of interest, a certain reference frame needs to be extracted. Cucitore et al. [29] used a reference frame that moves with a particle. Haller [30] proposed M_z to detect vortices, which is both Galilean invariant and rotation invariant. Bhatia et al. [4] introduced an internal frame that is computed as the harmonic component of a natural Helmholtz-Hodge decomposition [31]. Günther et al. [32] showed how to construct a Galilean invariant rotation invariant technique for vortex detection. Recently, Sauer et al. [33] introduced a novel method to explore spatio-temporal characteristics in both particle and volume data simultaneously. The transport and vortical behaviors in unsteady flow can also be studied via analyzing various attributes associated with pathlines [2, 34]. These attributes can be utilized for pathline and streak line placement [3], and glyph design [35].

Local analysis of unsteady flow Local analysis of flow focuses on small-scale behaviors, such as local volume dilation, compression, infinitesimal rotation, stretching, shearing, acceleration and momentum. Due to its locality, it is fast to compute and has been widely applied in the fluid mechanics community, for instance, the well-known λ_2 [8] and Q criterion [7] for the detection of vortices.

However, local analysis may not reveal the global flow structures associated with the transport behavior of the flow, and its visualization typically relies on the selection of a user-specified threshold to highlight the most salient regions, whose boundaries need not be aligned with the flow. Shi et al. [6] developed an advection filter along pathlines to study the transport behavior of various local properties. It is compared with the LCS structures of the corresponding FTLE fields. The set of local characteristics used in that work is adapted for our study. Fuchs et al. [5] introduced a local unsteadiness metric based on the material derivative. A global unsteadiness is then computed using a similar idea to the advection filter along pathlines. Critical points under the Lagrangian view are defined as the pathlines that have minimal unsteadiness.

Different from the above work, we concentrate on the pairwise relations of attributes based on their temporal behavior measured at both fixed locations and along particles.

Correlation analysis of multivariate data Correlation analysis is one of the most important tasks in the study of multi-field data. Sauber et al. [9] analyzed correlations in 3D multi-field scalar data using a vector similarity metric derived from gradient similarity measures. Lee et al. [36] proposed an algorithm to describe the correlation among various attributes in multivariate time-varying data sets based on how the attribute values change over time. Wang at al. [37] studied hierarchical clustering of volumetric samples based on the similarity of their correlation for scientific volume data. Zhang at al. [38] introduced a correlation map, which was expressed as a 2D layout of variables encoding their pairwise correlations. This map is then employed for visual correlation analysis. Chen at al. [39] presented a sampling-based approach to classify correlation patterns, based on which a static volume classification was created to summarize the correlation connection in time-varying multivariate data sets. Recently, Zhang et al. [40] pro-

posed a dissimilarity-preserving clustering algorithm and measured correlation connection in multi-variable and time-varying data sets. Their method can characterize time-varying patterns and spatial positions. The above works typically focus on linear correlation among attributes. In contrast, we study both linear and non-linear relation and/or dependency between attributes in unsteady flow.

Mutual Information (MI) quantifies the dependence between two random variables [41] based on information theory [42]. There are several papers that use information theory for various visualization problems. Xu et al. presented an information-theoretic framework to guide the selection and rendering of integral curves [43]. Biswas et al. proposed a new approach based on information theory through the multivariate data exploration process [44]. Wang et al. proposed a block-wise distribution based representation to preserve important features and alleviate uncertainty in large-scale scientific data visualization [45]. Zhang et al. [46] also analyzed the variable association in ensemble data. Analysis of dependent attributes values from a scalar-based view on multi-field data [12] and a brief overview of relation between the flow attributes [47] have been studied previously. Closely related to our work, Chen and Jaenicke were among the first to introduce mutual information to the visualization community in their seminal work [11]. In particular, they integrated the MI calculation to their proposed framework to reduce the uncertainty in the generated visualization. Their work has a very different goal from our work. Specifically, we apply MI to understand the time-varying dependency of flow attributes. Dutta et al. [12] applied MI to the study of scalar fields. Their analysis is based on pointwise MI (PMI) comparison of time-varying scalar fields to identify salient features. Our work is a logical progression of this in a few ways. Firstly, we extend this idea based on comparison of pairwise integral curves (as opposed to points only). Secondly, we consider both Eulerian and Lagrangian settings and compare the two.

To our best knowledge, our work is the first that explores the spatio-temporal, linear and non-linear relation of unsteady flow attributes using various metrics from both the Eulerian and Lagrangian perspectives. We also compare the ability of various metrics in revealing time-varying relationships between attributes, which can offer guidance for selecting metrics to study and compare amongst the range of flow attributes.

3 Flow Attributes and Their Relation Analysis

In this section, we first review a number of important concepts for unsteady flow and their characterizations (Section 3.1). We then introduce a number of methods to analyze the relation of pairs of unsteady flow attributes in space and time (Section 3.2).

3.1 Unsteady Vector Fields

Consider an unsteady vector field $\mathbf{v}(\mathbf{x},t)$ defined in a space-time domain $\Omega \subset \mathbb{R}^d \times \mathbb{R}$. Its dynamics are determined by the temporal evolution of densely placed massless particles, that is, for a particle $\mathbf{p}$, it satisfies $\frac{d\mathbf{p}(t)}{dt} = \mathbf{v}(\mathbf{x},t)$ ($\mathbf{x} = \mathbf{p}(t)$ is the position of the particle at time t). The trajectory of any massless particle seeded at $(\mathbf{x}_0, t_0)$ is called a *pathline*, which is a solution to the above ordinary different equation.

$$\mathbf{p}_{\mathbf{x}_0,t_0}(t_0+t) = \mathbf{x}_0 + \int_{t_0}^{t_0+t} v(\mathbf{p}_{\mathbf{x}_0,t_0}(\eta), t_0+\eta)d\eta \quad \text{s.t. } \mathbf{p}_{\mathbf{x}_0,t_0}(t_0) = \mathbf{x}_0$$

There are a number of local attributes for an unsteady vector field that are of interest to domain experts. For example, the *acceleration* of $\mathbf{v}$ is defined as $\mathbf{a}(\mathbf{x},t) = \frac{D\mathbf{v}}{Dt} = \frac{\partial \mathbf{v}(\mathbf{x},t)}{\partial t} + (\mathbf{v}(\mathbf{x},t)\cdot\nabla)\mathbf{v}(\mathbf{x},t)$, where $(\mathbf{v}(\mathbf{x},t)\cdot\nabla) = \nabla_{\mathbf{x}}\mathbf{v}$ represents the spatial gradient of $\mathbf{v}$, i.e., *Jacobian*, denoted by $\mathbf{J}$ for simplicity.

Important deformation modes of the flow can be obtained through the decomposition of $\mathbf{J}$. Specifically, the Jacobian matrix can be decomposed as $\mathbf{J} = \mathbf{S} + \mathbf{R}$, where $\mathbf{S} = \frac{1}{2}[\mathbf{J} + (\mathbf{J})^\top]$ and $\mathbf{R} = \frac{1}{2}[\mathbf{J} - (\mathbf{J})^\top]$ are the symmetric and antisymmetric components of $\mathbf{J}$, respectively. From this decomposition, local shear rate is measured as the Frobenius norm of $\mathbf{S}$, Q value at each point can be computed as $Q = \frac{1}{2}(\|\mathbf{R}\|^2 - \|\mathbf{S}\|^2)$ [48], and λ_2 is the second largest eigenvalue of the tensor $\mathbf{S}^2 + \mathbf{R}^2$ [8], respectively. They are typically used to characterize the vortical or rotational behavior of the flow. The attributes we consider in this paper are denoted as follows:

- A_1: vorticity, $\langle \nabla \times \mathbf{v}, \mathbf{z} \rangle$ for 2D, and $||\nabla \times \mathbf{v}||$ for 3D.
- A_2: acceleration magnitude, $||\mathbf{a}(\mathbf{x},t)||$.
- A_3: λ_2 (see above definition).
- A_4: Q (see above definition).
- A_5: local shear rate (i.e., the Frobenius norm of $\mathbf{S}$).
- A_6: determinant of $\mathbf{J}$.
- A_7: norm of $\mathbf{J}$, defined as $\sqrt{\sum_{ij} J_{ij}^2}$.

Note that all the above attributes can be measured (or calculated) at a given position and time in the flow domain in parallel.

3.2 Correlation Between Time-varying Attributes

The aforementioned attributes of unsteady flow are time-dependent. To study their pairwise behavior over time, we first briefly describe how we obtain sequences of attribute values in the Eulerian or Lagrangian sense. Given a specific local attribute $\mathbf{A}$, its value at a spatial position $\mathbf{x} \in \mathbb{M}$ at time $t \in \mathbb{T}$ can be denoted as $\mathbf{A}(\mathbf{x},t)$. Computing $\mathbf{A}(\mathbf{x},t)$ at a location $\mathbf{x}$ over time or along a pathline gives rise to a series of attribute values, which we denote as the attribute value sequence (AVS). The former sequence is obtained in the Eulerian fashion (i.e., measured at fixed location) and the latter is in the Lagrangian fashion (i.e., moving with a particle). With these attribute value sequences, we now describe our metrics to quantify the linear correlation and dependency among attributes over time.

3.2.1 Spatio-temporal Local Correlation Coefficient (LCC)

We extend the LCC [9] for steady scalar field to our unsteady setting, which is the Pearson Product-Moment Correlation Coefficient (PPMCC) [49] extended to the spatio-temporal domain. Specifically, for a local position $\mathbf{p} = (\mathbf{x},t) \in \Omega$ and a window function $G_{(\mathbf{x},t)}$ positioned at $\mathbf{p}$, the correlation value at $\mathbf{p}$ is defined by:

$$\text{ST\_LCC}_{A_i,A_j}(\mathbf{p}) = \rho_{A_i,A_j}(\mathbf{p}) = \frac{\text{cov}_{A_i,A_j}(\mathbf{p})}{\sigma_{A_i}(\mathbf{p})\sigma_{A_j}(\mathbf{p})} \tag{1}$$

$$\text{cov}_{A_i,A_j}(\mathbf{p}) = \int_\Omega G_{(\mathbf{x},t)}(\mathbf{q})(A_i(\mathbf{q}) - \bar{A}_i)(A_j(\mathbf{q}) - \bar{A}_j)d\mathbf{q} \tag{2}$$

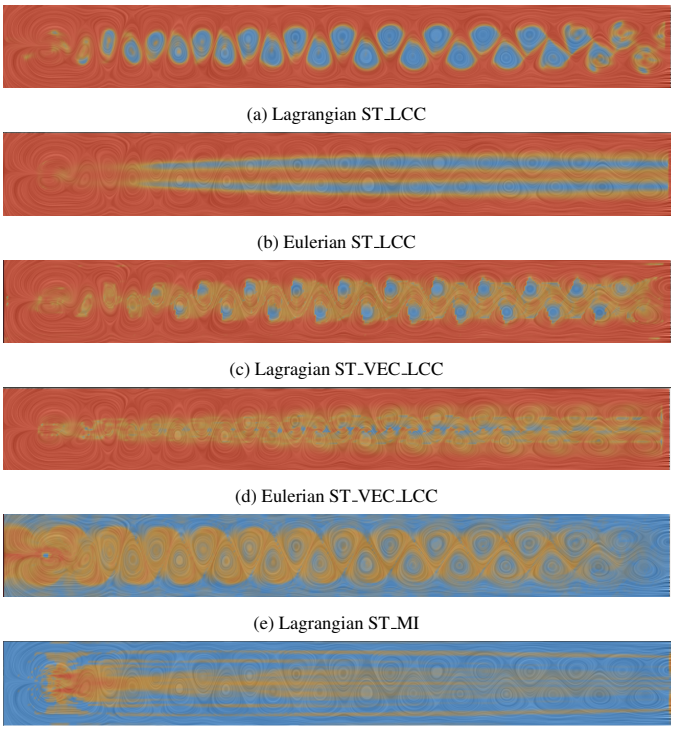
$$\sigma_{A_i}(\mathbf{p}) = \sqrt{\text{cov}_{A_i,A_i}(\mathbf{p})}$$

$$\bar{A}_i = \int_\Omega G_{(\mathbf{x},t)}(\mathbf{q})A_i(\mathbf{q})d\mathbf{q}$$

$\text{ST\_LCC}_{A_i,A_j}(\mathbf{p})$ returns a real number between -1 and 1. The closer to 0, the weaker the correlation between A_i and A_j. A_i and A_j are positively correlated if $\text{ST\_LCC}_{A_i,A_j}(\mathbf{p}) > 0$; otherwise, they are negatively correlated. We use a cylinder filter for 2D unsteady flows (i.e., a disk in space and a deformed cylinder or tube in space-time along the pathline) and a hyper-cylinder for 3D data as the window function $G_{(\mathbf{x},t)}$. We set $G_{(\mathbf{x},t)}(\mathbf{q}) = \frac{1}{n}$ if $\mathbf{q}$ falls in the local region (or kernel) defined around $\mathbf{p}$ and n is the number of points within the region, and $G_{(\mathbf{x},t)}(\mathbf{q}) = 0$ otherwise. Other proper window functions can also be used here. The radius (r) and height (h) of the cylinder determine the kernel size in the spatio-temporal domain Ω. Both r and h can be interactively adjusted by the user. Note that h is defined as the total number of time steps in the kernel, which corresponds to a time range $T = h\tau$ with τ being the time interval between two neighboring frames for a dataset. When $h = 0$, ST_LCC is reduced to the conventional LCC.

One characteristic of the above ST_LCC metric for scalar attributes is that it is scale invariant. That is, $\text{ST\_LCC}_{A_i,A_j} = \text{ST\_LCC}_{kA_i,lA_j}$ where $k,l \in R^+$. This may return very large ST_LCC values for cases where both attributes have small values, which can be altered by multiplying the above result with the absolute value of the normalized attribute values. We provide it as an option for the user to control. By default, this multiplication is disabled.

The above ST_LCC can be applied to both a fixed location (Eulerian) and a particle over time (Lagrangian). Figure 3 compares the ST_LCC fields computed in the Eulerian (b) and the Lagrangian (a) views, respectively, for the 2D flow past a cylinder. The comparison shows



(a) Lagrangian ST_LCC

(b) Eulerian ST_LCC

(c) Lagragian ST_VEC_LCC

(d) Eulerian ST_VEC_LCC

(e) Lagrangian ST_MI

(f) Eulerian ST_MI

Fig. 3: Different correlation measurements between norm (A_7) and shear rate (A_5) of a flow past a square cylinder. Kernel size $r = 3$ and $h = 250$ with $\tau = 0.008s$. For ST_LCC results, blue indicates negatively correlated and red indicates positively correlated. For MI results blue indicates low dependency and red indicates high dependency.

that the time-varying attribute behaviors in the Eulerian view highlight the overall (or average) movement of the vortices in the vortex street region behind the cylinder, while its behaviors in the Lagrangian view emphasize the individual vortices. This observation is similar to what is reported in [50] which compares the behavior of the accumulated attributes in the Eulerian and Lagrangian views, respectively.

3.3 Vector-Based Correlation

In this section, we propose a correlation measurement based on the gradient vector field of the attribute field. Given an attribute field A_i, we define its gradient field as $\mathbf{g}_i = \nabla A_i = (\frac{\partial A_i}{\partial x}, \frac{\partial A_i}{\partial y}, \frac{\partial A_i}{\partial z}, \frac{\partial A_i}{\partial t})$. $\mathbf{g}_i$ points to the direction that the value of A_i increases the fastest, and its magnitude $||\mathbf{g}_i||$ represents how quickly the value changes. Therefore, studying the correlation of the gradient vector field of two attributes may further reveal their similarity in behavior or variation in space and time. First, we define the gradient similarity measure (GC) of two attributes A_i and A_j at position $\mathbf{p} = (\mathbf{x}, t)$ as:

$$\begin{aligned} GC(\mathbf{g}_i, \mathbf{g}_j) &= GC_d(\mathbf{g}_i, \mathbf{g}_j) GC_m(\mathbf{g}_i, \mathbf{g}_j) \\ GC_d(\mathbf{g}_i, \mathbf{g}_j) &= \frac{\langle \mathbf{g}_i, \mathbf{g}_j \rangle}{||\mathbf{g}_i||\ ||\mathbf{g}_j||} \\ GC_m(\mathbf{g}_i, \mathbf{g}_j) &= 4 \frac{||\mathbf{g}_i||\ ||\mathbf{g}_j||}{(||\mathbf{g}_i|| + ||\mathbf{g}_j||)^2} \end{aligned} \tag{3}$$

$GC_d(\mathbf{g}_i, \mathbf{g}_j)$ represents the direction similarity between the two gradient vectors, and $GC_m(\mathbf{g}_i, \mathbf{g}_j)$ measures their magnitude similarity. We use multiplication to combine the two terms, as they may have different value ranges (i.e., g_i is not normalized). $GC(\mathbf{g}_i, \mathbf{g}_j)$ is zero when the two gradients are orthogonal. When the two gradient vectors point in the same direction and have the same magnitude, $GC(\mathbf{g}_i, \mathbf{g}_j)$ returns 1 (i.e., positively correlated). If the two gradient vectors have the same magnitude but point in opposite directions, $GC(\mathbf{g}_i, \mathbf{g}_j)$ returns -1 (i.e., negatively correlated). As opposed to the previous vector similarity metrics that always return positive values [9], our metric can now measure both the positive and negative similarity of two vectors. Compared with the vector similarity metric introduced by Crouzil et al. [51], which relies on the magnitude of the gradient vectors, our metric is able to capture the correlation even when the magnitude of the gradient vectors is small.

Since $GC(\mathbf{g}_i, \mathbf{g}_j)$ is a scalar field in the range of $[-1, 1]$, we can now define a vector-based LCC based on this vector similarity metric. First, the mean gradient field with a given window function $G_{(\mathbf{x},t)}$ is computed as $\bar{\mathbf{g}}_i = \int_{\mathbf{q} \in \Omega} G_{(\mathbf{x},t)}(\mathbf{q}) \mathbf{g}_i(\mathbf{q}) d\mathbf{q}$. Then the co-variance of the two vector-valued attributes (i.e., the gradient vector fields here) can be defined as:

$$\mathrm{cov}_{\mathbf{g}_i, \mathbf{g}_j}(\mathbf{p}) = GC(\bar{\mathbf{g}}_i, \bar{\mathbf{g}}_j) \int_\Omega G_{(\mathbf{x},t)}(\mathbf{q}) GC(\mathbf{g}_i(\mathbf{q}), \bar{\mathbf{g}}_i) GC(\mathbf{g}_j(\mathbf{q}), \bar{\mathbf{g}}_j) d\mathbf{q} \tag{4}$$

Note that multiplying $GC(\bar{\mathbf{g}}_i, \bar{\mathbf{g}}_j)$ is necessary in order to preserve the sign of the correlation. Thus, the vector LCC is defined as:

$$vecLCC_{\mathbf{g}_i, \mathbf{g}_j}(\mathbf{p}) = \frac{\mathrm{cov}_{\mathbf{g}_i, \mathbf{g}_j}(\mathbf{p})}{\sigma_{\mathbf{g}_i}(\mathbf{p}) \sigma_{\mathbf{g}_j}(\mathbf{p})}, \text{ where } \sigma_{\mathbf{g}_i}(\mathbf{p}) = \sqrt{\mathrm{cov}_{\mathbf{g}_i, \mathbf{g}_i}(\mathbf{p})} \tag{5}$$

Compared to the above scalar LCC, our proposed vector based VEC_LCC is scale-sensitive. Figures 3c and 3d show the results for vector-based ST_LCC in Lagrangian and Eulerian domains. The results show that in the core regions of the vortices, a minimal variation of shear rate is exhibited while the norm need not, resulting in negative correlation. In the meantime, the boundary layers of the vortices exhibit stronger variation of shear rate, which is aligned with the behavior of the norm. Therefore, we observe positive correlation values there. This matches previous results [52].

3.4 Measuring Attribute Dependency Using MI

The above correlation computation is limited to the linear co-varying behavior between two attributes. Another metric is needed to capture non-linear relations between attributes. Among many non-linear relations, the dependency between attributes is of interest to fluid experts, which will facilitate understanding the causal relations between attributes. To study the dependency between attributes, we adapt the mutual information metric. In this section, first we review the definition of mutual information, then introduce the spatio-temporal mutual information metrics.

3.4.1 Spatio-temporal Mutual Information

In this section, we start with a brief introduction of mutual information based on a book by Chen et al. [41]. Consider X to be a discrete random variable with alphabet $\mathbb{X}$ and probability distribution $p(x)$, where $p(x) = Pr[X = x]$ and $x \in \mathbb{X}$. The *entropy* $H(x)$ of a discrete random variable X is defined by:

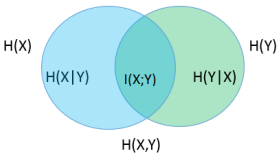
$$H(X) = -\sum_{x \in \mathbb{X}} p(x) \log_2 p(x)$$

for a pair of discrete random variables X and Y with a joint probability distribution $p(X,Y) = \{p(x,y)\}$, the *joint entropy H(X,Y)* is defined by:

$$H(X,Y) = -\sum_{y \in \mathbb{Y}} \sum_{x \in \mathbb{X}} p(x,y) \log_2 p(x,y)$$

where $p(x,y) = Pr[X = x, Y = y]$ is the joint probability of x and y. The *conditional entropy* $H(Y|X)$ of a random variable Y given a random variable X is defined by:

$$\begin{aligned} H(Y|X) &= \sum_{x \in \mathbb{X}} p(x) H(Y|X = x) = \sum_{x \in \mathbb{X}} p(x) (-\sum_{y \in \mathbb{Y}} p(y|x)) \log_2 p(y|x) \\ &= -\sum_{x \in \mathbb{X}} \sum_{y \in \mathbb{Y}} p(x,y) \log_2 p(y|x) \end{aligned}$$



where $p(y|x) = Pr[Y = y|X = x]$ is the conditional probability of y given x. The entropy measures the average amount of information or uncertainty in a random variable X. We want to quantify how much uncertainty the realization of a random variable X has if the outcome of another random variable Y is known. The inset illustrates the additive and subtracting relationships for different measurements associated

with two correlated variables X and Y. The area contained by both circles is the joint entropy (X,Y) and the circle on the left (blue and cyan) is the individual entropy (X), with the blue being the conditional entropy $(X|Y)$. The circle on the right (green and cyan) is (Y), with the green being $(Y|X)$. The cyan overlap represents the mutual information $I(X;Y)$ which measures the common uncertainty between two random variables X and Y and is defined as:

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) = H(X) + H(Y) - H(X,Y) = H(X,Y) - H(X|Y) - H(Y|X) \quad (6)$$

where $H(X)$ and $H(Y)$ are the marginal entropies, $H(X|Y)$ and $H(Y|X)$ are the conditional entropies, and $H(X,Y)$ is the joint entropy of X and Y. In our case, X and Y are two attributes and $I_{\mathbf{p}}(X;Y)$ (i.e., ST_MI) is computed within a spatio-temporal kernel located at $\mathbf{p}$.

Using the MI metric to study attribute dependency Figures 3e and 3f show the ST_MI fields computed based on the Lagrangian and the Eulerian attribute sequences of the Frobenius norm of the Jacobian and local shearing, respectively. The Lagrangian ST_MI field (Figures 3e) exhibits a similar vortex street pattern to its ST_LCC counterpart (Figure 3a), while the ST_MI in the Eulerian view reveals the average movement of the vortices. However, compared to the result of ST_LCC, ST_MI has large values at the vortex regions behind the cylinder. This is because the MI values are always non-negative, therefore, MI essentially reveals the strength of the dependency of the attributes, ignoring whether the attributes are positively or negatively correlated. Also, there are relations, e.g., the non-linear relations, between attributes that cannot be captured by ST_LCC, but can be captured by ST_MI. See Figure 7 for an example, in which the non-linear dependency of the two attributes within the vortex regions is revealed by ST_MI (Figure 7d, h), but ST_LCC only highlights regions where the two attributes have strong co-varying behavior (Figure 7a, e). Note that the MI and entropy computation is dependent on the number of bins for distribution estimation. As most other papers, we used equal-width bins for all our results to automatically choose the optimal bin size [53].

4 Correlation Visualization

To visualize the scalar fields computed using either the ST_LCC (Section 3.2.1) or the ST_MI (Section 3.4.1), we utilize a blue-white-red (BWR) color coding scheme. For the ST_LCC correlation fields, since the range of the correlation value is [-1, 1], the BWR color scheme is particularly useful to highlight the different characteristics of correlation, i.e., blue for negative values, white for zero and red for positive values. For ST_MI fields, while the values are always non-negative, we still opt for the BWR scheme to highlight places with large correlation values (i.e., in red). Both the 2D color plots and 3D volume renderings shown in this paper are the results of the S_LCC, ST_LCC, ST_VEC_LCC or ST_MI in spatio-temporal domain starting from a given time.

Pathline rendering with attribute relation information To reveal more detailed attribute relation during particle advection, we visualize the seeded pathlines using color coding based on the quantified relation of the pairwise attributes associated with the particles over time. Figures 4a and 4b provide a few examples of pathlines that are color-mapped based on the ST_LCC and ST_VEC_LCC computed along them, respectively. The two attributes used for these experiments are acceleration and Q.

Specifically, we seed these pathlines at the outer layers of two vortices and near a vortex core, as indicated by the LIC patterns, respectively. Interestingly, the colors along the pathline seeded near the vortex core exhibit little variation, while those pathlines seeded at the boundaries of vortices or near the saddle-like regions exhibit large fluctuation, especially when gradient similarity (i.e, vector-based ST_LCC) (Eq.(3)) is employed for the color coding. In contrast, at the boundaries of vortices and near the saddle-like regions where the flow exhibits large shearing and stretching behavior, the co-varying behavior of the pairwise attributes may not be similar over time. These correlations can be better understood by inspecting the accompanying plots of attributes. The same scenario occurs with the mutual information based result (Figure 4c). As shown in the accompanying histograms in Figure 4c,

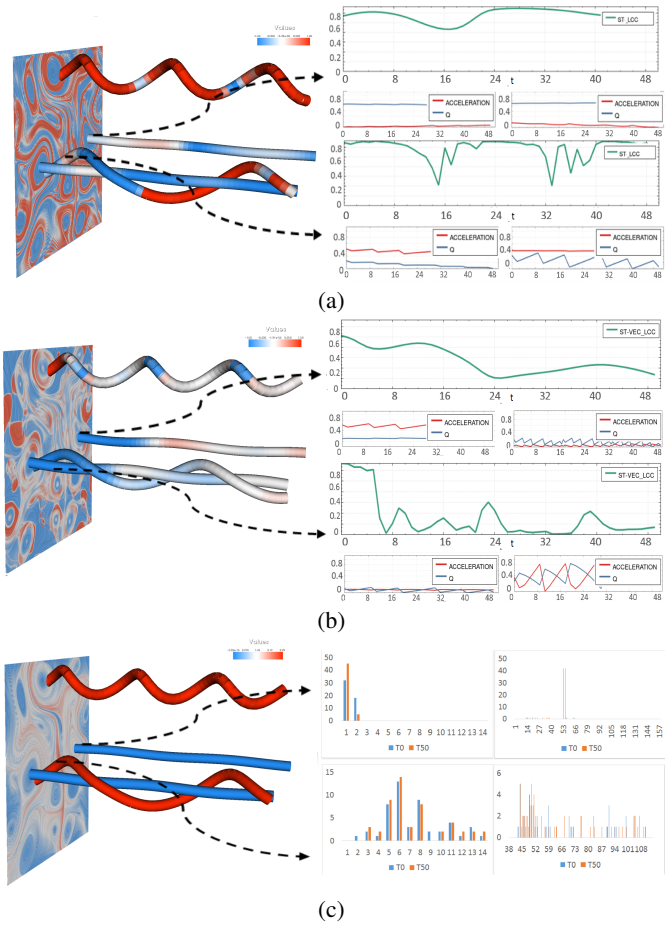
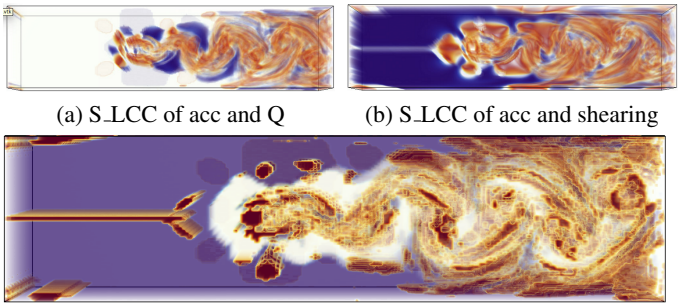


Fig. 4: Sampled pathlines colored based on the ST_LCC (a), ST_VEC_LCC (b) and ST_MI (c) of acceleration and Q, respectively. Each top plot in (a) and (b) shows the ST_LCC/ST_VEC_LCC values along the respective pathlines indicated by the arrows. The two subplots beneath each large plot show the trends of the two given attributes within a kernel. For the MI result (c), the left two histograms show the distributions of the attribute values of Q within kernels on two pathlines (indicated by the arrows), while the right two show the distributions of the values of acceleration within the same kernels. Each histogram shows two distributions of the values of the respective attribute within two different kernels sampled at t=0 (blue) and t=50 (orange) on the corresponding pathline, respectively. Since we used equal-width bins, the number of bins are different for different attributes.

we see a wide range and non-uniform (or fluctuating) distribution of the values for attributes associated with pathlines seeded at the outer layers of the vortices, indicating the high entropy (i.e., high uncertainty) and joint probability of the two attributes, which result in high MI values. In contrast, the distributions of values for attributes associated with pathlines seeded at the vortex center are very concentrated, leading to low entropy (i.e., low uncertainty) and low MI value. Such an interpretation is not easy to obtain without the proposed visualization.

Ranking strategy One interesting question that experts may ask in practice is which pairs of attributes are more related in certain areas (e.g., having large absolute correlation value or having high dependency), and how this information is related to certain flow characteristics (e.g., vortex cores and outer layers of vortices). Knowing this will in turn help develop an effective way to characterize regions with distinct flow behavior.

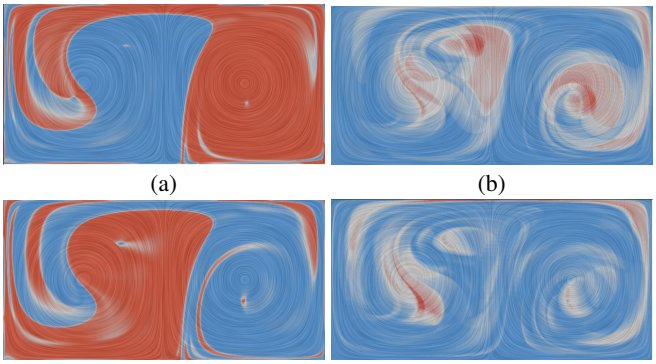
To achieve so, for each spatial position, we compute the ST_LCC (or MI) of various attribute-pairs, and sort their ST_LCC values in descending order. We then color this position based on the top-ranked pair of attributes. This gives rise to a domain segmentation, which we hope enables us to identify the most dominant behaviors (characterized by the most correlated attributes) in the individual flow regions. Here, let



(a) S_LCC of acc and Q (b) S_LCC of acc and shearing

(c) segmentation

Fig. 5: Spatial LCC of acceleration and Q (a) vs acceleration and shearing (b) and the segmentation result for three pairs of these attributes (c). In (c), red regions correspond to places where the correlation between acceleration and Q is the strongest. White regions visualize the places where acceleration and shearing has the highest value of correlation, and purple segments correspond to the regions where correlation between shearing and Q is the strongest.



(a) (b)

(c) (d)

Fig. 6: ST_LCC results of curl and Q (a) and curl and shearing (c); ST_MI results of curl and Q (b) and curl and shearing (d). All results are computed based on the Lagrangian attributes using a spatial kernel r = 3 and temporal kernel h = 100 with $\tau = 0.01$.

us take the 3D flow past a cylinder as an example. As we already know the dominant characteristic of this flow is the von Kármán vortex street, we concentrate on the three relevant attributes, i.e., acceleration, Q and local shear rate for the ranking and perform the segmentation based on the ranking. Figure 5a and 5b show the ST_LCC results between acceleration and Q vs. acceleration and shearing, respectively, while the ranking based segmentation is visualized in Figure 5c. Different colors correspond to different top-ranked pairs of attributes. From this result, we see that at the boundary layers of the vortex street, the attribute pair of acceleration and shearing is ranked top (colored by white), while in the vortex core areas the attribute pair of Q and accelerating dominates (colored in red). The rest of the domain (in purple) is dominated by the the pair of shearing and Q. This result matches the knowledge in fluid mechanics [54]. This example demonstrates that our simple ranking-based segmentation can provide an overview of the dominant flow dynamics (e.g., rotational versus stretching or shearing) in different flow regions for an effective unsteady flow exploration.

5 Results

We have applied our correlation analysis and visualization to a number of 2D and 3D analytic and real-world simulated flow. In the following, we provide some detailed discussion on some of the results.

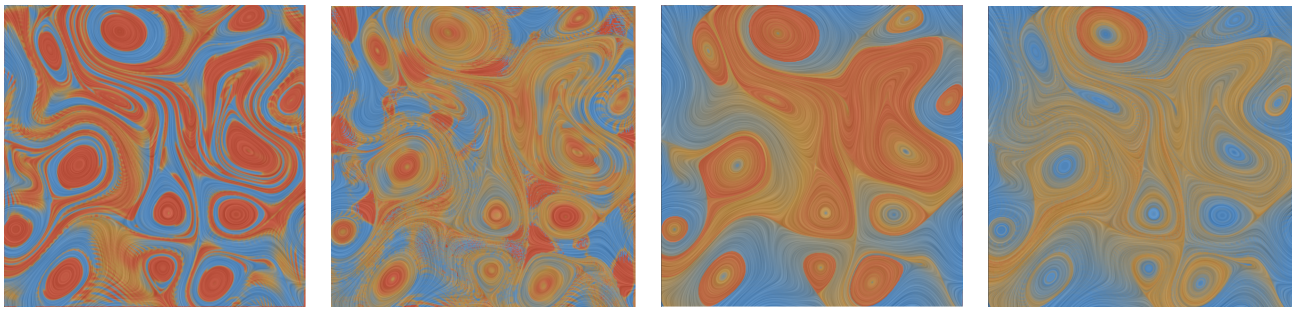
Double gyre Figure 6 shows the ST_LCC correlation (6a and 6c) and ST_MI (6b and 6d) of the vorticity (A_1) vs. Q ((A_4)) (top row) and the vorticity vs. shearing (A_5) (bottom row), respectively, of the double gyre flow [55] starting from $t = 0$ and with time window $T = 100 \times 0.01$. Lagrangian attribute value sequences are used. The spatial sample resolution is 256×128. Figures 6a and 6c convey similar structure, but the colors are reversed. This is because in regions where flow exhibits strong vortical (or rotational) behavior (i.e., with large Q values), its shear characteristic is low (i.e., with low shear values). Similarly, when the shearing motion dominates the flow behavior, its rotational behavior is weak. This can also be derived from the definition of $Q = \frac{1}{2}(\|\mathbf{R}\|^2 - \|\mathbf{S}\|^2)$ [48], where $\|\mathbf{S}\|^2$ denotes the square of the local shear rate. In this definition, Q and shearing have relatively negative (but non-linear) correlation. From the ST_MI results shown in Figures 6b and 6d, we find that both of them exhibit a similar pattern; however, the result between vorticity and Q highlights the LCS structure (not shown here) that is stronger due to the high vorticity in that region. Also, the interleaving layout of the negative (blue) and positive (red) regions are aligned with the LCS structure, indicating the partitioning of regions with different dominant rotational behaviors.

2D flow past a cylinder This simulation covers a subset of the spatio-temporal domain, i.e., $[-0.5, 7.5] \times [-0.5, 0.5] \times [15, 23]$. The time range is chosen such that the von Kármán vortex street is fully developed behind the cylinder. The resolution of the data set is $400 \times 50 \times 1001$. Total time is 8 seconds for all 1001 frames. We measured the time from when a vortex is created close to the cylinder until it leaves the boundary. The lifespan of a vortex is 925 frames. We use the same spatial resolution, i.e., 400×50 to compute the attribute fields and the subsequent correlation fields. Figure 3 provides the correlation results for the local shear rate (A_5) and norm of Jacobian (A_7) . Based on their definitions, in 2D vector fields, $A_7 = \sqrt{J_{11}^2 + J_{12}^2 + J_{21}^2 + J_{22}^2}$, and $A_5 = \sqrt{J_{11}^2 + J_{22}^2 + \frac{(J_{12}+J_{21})^2}{2}}$. The difference of the squares of these two norms is $\sigma^2 = \frac{(J_{12}-J_{21})^2}{2}$. That said, in regions that possess strong rotational characteristics (i.e., $|(J_{12} - J_{21})|$ is large), the two norms have larger difference, while in regions that are close to curl-free (i.e., $|(J_{12} - J_{21})| \approx 0$), the two norms are almost identical. Using this observation, we find that the correlation results shown in Figure 3 match our expectation. That is, the correlation of these two attributes in the vortex regions should be small, while in other places it should be relatively large. One interesting point is that due to the dissipation of energy, attributes are going to be depreciated at the end (far right) of the flow. This results in lower correlation values close to the outlet on the right of the cylinder flow.

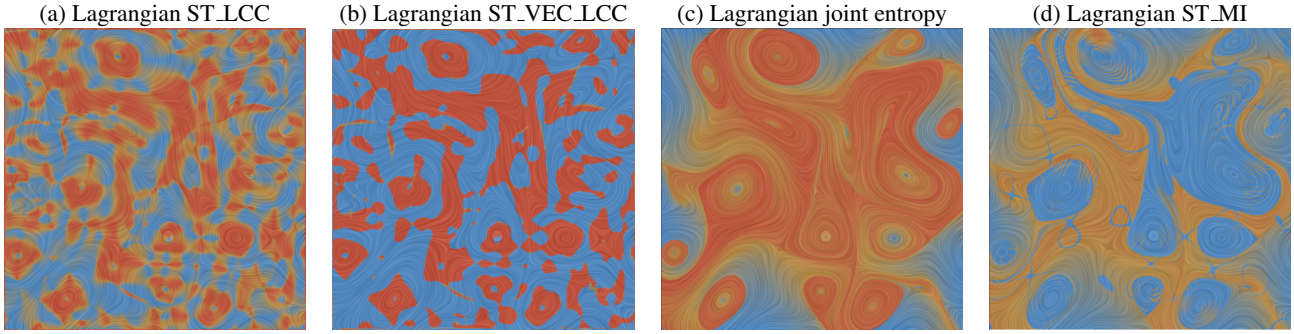
Figures 3a, 3c and 3e show the ST_LCC, ST_VEC_LCC, and ST_MI results in the Lagrangian view, which highlight the structure of vortex street. In contrast, the results shown in Figures 3b, 3d and 3f, using a spatio-temporal kernel in the Eulerian view, reveals that the individual vortex structure is merged into a corridor, which corresponds to the area that the vortex street sweeps through. Comparing the results of ST_MI and ST_LCC, we see that ST_MI reveals the boundaries of the vortex structure clearer than ST_LCC. This is due to the fluctuating pattern of the attribute values in the boundary of the vortex structure which indicates a higher dependency between these attributes in the boundary. Looking closer at the ST_LCC field computed in the Lagrangian view, we see that the two attributes involved, i.e., acceleration and Q, exhibit a positive correlation along the boundaries of the individual vortices, while having negative correlation near the center of the vortices. This also matches the observation in previous literature [56]. That is, the acceleration magnitude at the centers of vortices are minimal, while the Q values there are large.

Ocean data The next data set is taken from the top layer of a 3D simulation of global oceanic eddies for 350 days of the year 2002 [57]. Each time step corresponds to one day. The 2D vector field has a spatial resolution of 3600×2400. We extract tiles from the central Atlantic Ocean (600×600). Figure 2 shows the results of the acceleration and the determinant of Jacobian matrix (i.e., Q in 2D [48]). As we see, these two attributes are positively correlated in the boundary of the vortices, while negatively correlated at the vortex center. Figure 2c shows the Lagrangian MI computed for these two attributes. Compared to ST_LCC the entire vortex regions are pronounced except at the vortex cores due to the high dependency of these two attributes at vortex regions whether they are positively or negatively correlated.

HCCI data This data simulates a homogeneous charge compression ignition (HCCI) engine combustion [58], which has a spatial resolution of 640×640 and 299 time steps. Figure 7 shows a number of correla-

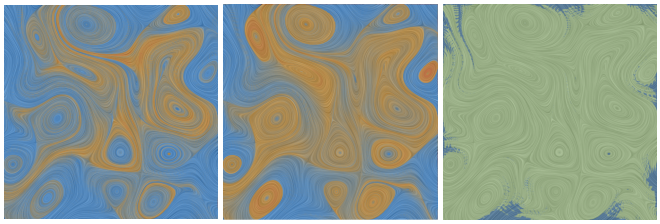


(a) Lagrangian ST_LCC (b) Lagrangian ST_VEC_LCC (c) Lagrangian joint entropy (d) Lagrangian ST_MI



(e) Eulerian ST_LCC (f) Eulerian ST_VEC_LCC (g) Eulerian joint entropy (h) Eulerian ST_MI

Fig. 7: Different correlation measurements between λ_2 and acceleration of HCCI data set. Kernel size $r = 3$ and $h = 50$ (with $\tau = 0.05$)).



(a) $H(A_7|A_3)$ (b) $H(A_3|A_7)$ (c) Comparison.

Fig. 8: The comparison of conditional entropy for norm of the Jacobian (A_7) vs. λ_2 (A_3). Blue-white-red color coding is applied to (a) and (b), while for (c), light green indicates $H(A_7|A_3) < H(A_3|A_7)$ while blue means $H(A_3|A_7) < H(A_7|A_3)$.

tion fields for acceleration (A_2) and λ_2 (A_3). Similar to the results of the cylinder data, the ST_LCC computed using the Lagrangian attribute sequences highlights a more complete structure of vortices, while the Eulerian ST_LCC emphasizes the areas the vortex cores sweep through over time. Also, we find that acceleration and λ_2 are positively correlated at vortex regions which is expected.

Figures 7b and 7f show the ST_VEC_LCC results using the Lagrangian and Eulerian attribute sequences for acceleration and λ_2, respectively. The Lagrangian ST_VEC_LCC highlights the vortex centers and places with strong stretching flow (i.e., places with saddle-like patterns), indicating the gradients of the two attributes have similar time-varying behavior. Similarly to its ST_LCC counter part, the Eulerian ST_VEC_LCC does not show much meaningful structure. Figures 7c and 7g show the Lagrangian and Eulerian joint entropy of these two attributes, respectively. The joint entropy highlights regions where both attributes have strong relation, either dependency or independency. In this comparison, Lagrangian joint entropy is more focused on vortex regions than the Eulerian result, as the Eulerian result tends to highlight regions where those vortices sweep through, similar to the Eulerian ST_LCC we have seen earlier (Figure 3b).

Figures 8a and 8b show the conditional entropy of two attributes λ_2 (A_3) and the norm of the Jacobian (A_7), for the HCCI data. The conditional entropy $H(A_3|A_7)$ measures the amount of dependency of A_3 on A_7, and usually, $H(A_7|A_3) \neq H(A_3|A_7)$. We compare the two conditional entropy for the two attributes and visualize it in Figure 8c. We see that in most regions, *Norm* is dependent on λ_2, while λ_2 is less dependent on *Norm*, since $H(A_7|A_3) < H(A_3|A_7)$. This is expected, as regions with strong vortical flow result in larger norm of the Jacobian, while the inverse is not always true. More discussion on the conditional entropy can be found in the supplemental document.

Ocean circulation This dataset is a model of oceanic circulation, provided by *the Estimating the Circulation and Climate of the Ocean, Phase II (ECCO2): High-Resolution Global-Ocean and Sea-Ice Data Synthesis project* [59]. This data provides us the opportunity to examine the evolution of ocean eddies and circulation through time. The data has a spatial resolution of 1440×720 with 50 time steps (equals 150 days). Figure 1 shows the Lagrangian ST_LCC and ST_MI for acceleration and Q. The spatial kernel window size is 5 and the temporal kernel size is 50 ($\tau = 0.004$). The outline of the African Continent is visible at the far left of that slice, with South America visible at the far right. The result shows that around the equator, there exists a series of strong vortices (i.e., eddies). More results of this data are in the supplemental document.

Axisymmetric vortex ring impact This flow is an axisymmetric simulation of a vortex ring hitting a no-slip wall. During the interaction process, the vortex ring approaches the wall and causes a boundary layer to appear. As the vortex slides against the wall, the boundary layer is lifted up as a secondary vortex, which in turns lifts up the primary vortex. This dataset helps us analyze the role of coherent structures and their interaction with walls, as well as the generation of turbulence in wall-bounded flows [60]. Figure 9 shows the segmentation results for pairs of attributes involving pressure (i.e., the pressure versus the other attributes) over time. From the result, we see that at the vortex center region, the correlation between pressure and vorticity is the highest (i.e., the dark green areas), while there is a layer around the vortex center (i.e., the pink region) in which the correlation between pressure and kinetic energy is the strongest. Outside the vortex area (i.e., the blue regions), the correlation between pressure and shearing dominates.

3D flow behind a square cylinder This flow was simulated by placing a square cylinder in a fluid flow. By subtracting the average velocity from the flow, we see interesting swirling structures [61]. The dimensions are $192 \times 64 \times 48 \times 101$ in the volume of $[-12, 20] \times [-4, 4] \times [0, 6]$. Figures 10a and 10b show the vorticity and acceleration of the flow, respectively, and Figure 10c shows the S_LCC of these two attributes. From this result, we see that acceleration and vorticity are positively correlated in the outer layer of vortices, which matches our previous observation in the 2D flow past a cylinder. Figure 11 compares the results of Eulerian and Lagrangian ST_MI between acceleration and Q with different kernel sizes.

In Section 3.2 we show that for 2D datasets, larger temporal kernel

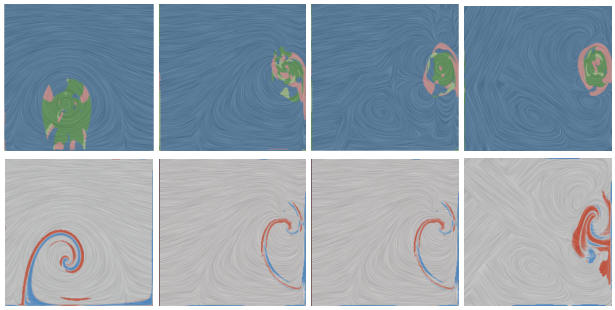
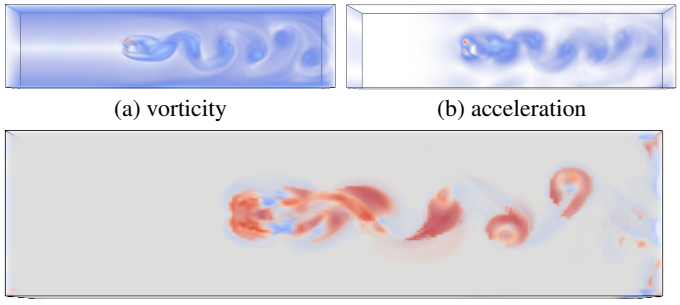


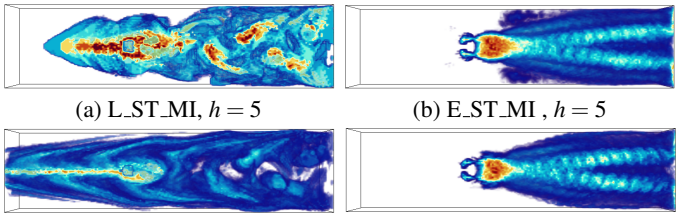
Fig. 9: The top row shows the ranking based segmentation for the spatial correlation between pressure and other attributes over time (from left to right). The green color indicates the regions where the correlation of pressure and vorticity is the highest, the pink color corresponds to the pair of pressure and kinetic energy, and blue for the pair of pressure and shearing. The second row shows the result of ST_LCC for acceleration and kinetic energy. The results highlight regions similar to the FTLE ridges. The spatial kernel size is $r = 3$ and the temporal kernel size is $h = 40$.



(a) vorticity (b) acceleration

(c) spatial LCC

Fig. 10: Spatial LCC of *acceleration* and vorticity for the 3D flow behind cylinder at time step 40. Note that the LCC result highlights the outer layer of the 3D vortices of the flow as expected.



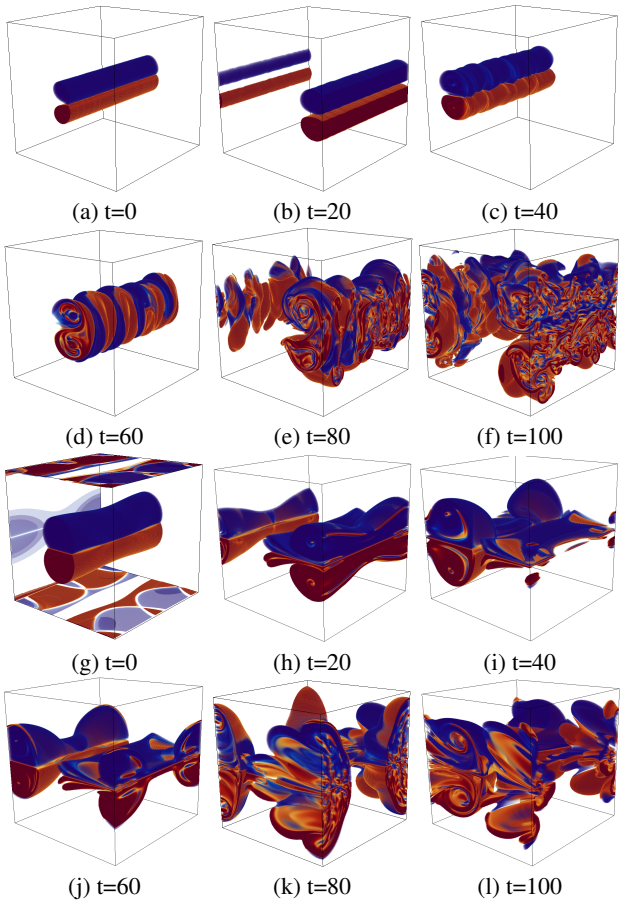
(a) L_ST_MI, $h = 5$ (b) E_ST_MI , $h = 5$

(c) L_ST_MI, $h = 20$ (d) E_ST_MI , $h = 20$

Fig. 11: Comparison of Lagrangian (L_) ST_MI and Eulerian (E_) ST_MI for the 3D cylinder data with different time windows h. Left column shows L_ST_MI for acceleration and Q, and the right shows E_ST_MI. We see that larger h makes both L_ST_MI and E_ST_MI smoother, but it does not change the structure of E_ST_MI much.

sizes reveal the flow structure clearer. However, this need not be the case for some 3D data. This is because for most 2D data we have experimented with (e.g., Double Gyre, 2D cylinder flow, HCCI etc.), the vortex structures are relatively stable over time, while in some 3D flows the vortices undergo large changes. For instance, the 3D cylinder flow contains the state before the vortex street is formed, thus, the pathlines started in early time may exhibit large change in shape during the transition from the laminar flow to turbulence flow. In that case, our proposed metrics work well for a small temporal kernel size in order to be able to capture the behavior of pathlines in a specific region. Larger temporal kernels will average (or smooth) the values of correlation or MI in different regions. Hence, we see Figure 11c does not reveal the vortices better than Figure 11a using ST_MI. In contrast, the ST_MI result in the Eulerian view is not affected much by the kernel size, since particle movement is not considered there.

3D vortex tube data This dataset simulates two separate cases of parallel, counter-rotating vortex tubes at a circulation-based flow with



(a) t=0 (b) t=20 (c) t=40

(d) t=60 (e) t=80 (f) t=100

(g) t=0 (h) t=20 (i) t=40

(j) t=60 (k) t=80 (l) t=100

Fig. 12: The spatial correlation of dye and vorticity for the elliptical instability simulation (a-f) and the Crow / reconnection simulation (g-l). In the top, we can see how the dye first tracks vorticity, until the non-linear elliptical instability kicks in and destroys the vortex, and the correlation is lost. In the bottom, we can see how the dye follows the vorticity much better even after the change in topology, and there is less mixture between red and blue regions. The spatial kernel size is 3.

Re=3500 and a distance of 2.5 radii apart, which undergo either an elliptical instability [62] that ends with a vortex disintegration, or a Crow instability which ends with a vortex reconnection [63] depending on the initial conditions. The two vortices interact mainly through strain. The datasets have dimensions of $360 \times 360 \times 360 \times 120$. A comparison between the two instabilities is made by adding a tracer, or dye, simulated as a passive scalar with a Schmidt number of unity. This scalar will track the fluid where it was originally injected. For the reconnection simulation, there are two counter-rotating vortex tubes, which after the reconnection changes topology, half of each tube reconnects with the half of the other tube, to form two tubes. For this case, the dye follows the vorticity from each tube, while for the elliptical case, the instability generates perpendicular filaments of vorticity that have a lower dye intensity, so the correlation between dye and vorticity will be weaker. Figure 12 compares the spatial LCC of vorticity and dye over time for the two simulations. From this comparison, we see that in both flows, the two attributes have high correlation at vortex regions, indicating similar co-variance behavior of the two attributes over time despite the flow in the elliptical instability decorrelating dye and vorticity more than the other one. This discrepancy from the expectation is interesting and asks for an in-depth investigation. We provide additional results on these two simulations in the supplemental document.

Conclusive remarks: Based on the above experiments, we can summarize that for finding the center region of vortices, computing ST_LCC and/or ST_VEC_LCC of acceleration with another relevant attribute (e.g., Q or λ_2) in a Lagrangian view, works well (e.g., Figures 7a-b and 5c). Computing ST_LCC of the norm of the Jacobian and local shearing or the ST_MI of various attribute pairs better highlights the outer layer

Table 1: Performance result for different datasets.

Dataset/Time(s)	cylinder		HCCI		Isabel2D		VortexRing		Large Ocean		Cylinder3D		Tube Reconnection	
	GPU	CPU	GPU	CPU	GPU	CPU	GPU	CPU	GPU	CPU	GPU	CPU	GPU	CPU
ST_LCC	0.285	117	1.784	363	0.341	176	0.679	91	8.327	2670	2.239	671	7.13	1902
ST_VEC_LCC	1.219	687	4.019	1093	0.732	451	2.76	310	20.341	3100	2.981	1589	12.496	3045
ST_MI	0.341	132	2.01	540	0.428	204	1.102	129	12.41	2540	3.812	1890	8.925	2013

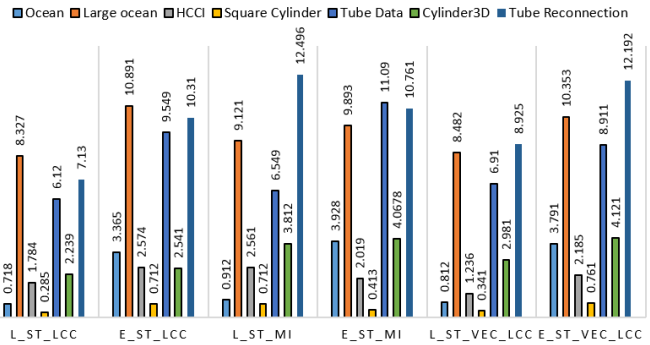


Fig. 13: Performance result for different metrics and different datasets.

of vortices (Figures 3e and 7d). For the general understanding of attribute relation and the overview of the flow structure, ST_LCC (and S_LCC for 3D data) should be used. For highlighting the full vortex regions as well as more complicated types of dependency, ST_MI should be applied. Also, our ranking based segmentation can help identify different layers of vortex structure (Figures 5 and 9). Furthermore, joint / conditional entropy of attributes may provide explanation on the dependency of two attributes (Figure 8).

5.1 Performance Analysis

We conducted a performance analysis on a number of datasets for the three metrics (i.e., ST_LCC, ST_VEC_LCC, and ST_MI) to demonstrate the performance of our method. We used a workstation with an Intel Xeon(R) CPU E5-2640 v3 @ 2.60GHz, 128GB RAM and an Nvidia Quadro M4000 graphics card. Using the GPU accelerated implementation of our measurements, we are able to generate the flow movement animation from large sets of particles even for Isabel and large scale ocean in real time. Figure 13 and Table 1 provide the timing results for the correlation computation with respect to the given resolutions, kernel sizes and datasets. The performance gain obtained using a GPU accelerated version of our method is between 20-50 times, compared to a CPU implementation. This does not include data loading time. The kernel sizes for different datasets were selected empirically to ensure a clear representation of flow features (e.g., vortex structures).

5.2 Parameter Study

The spatial size, r, of the kernel and the temporal length, h, are two parameters that the user may need to explore in order to obtain optimal results. In particular, a small spatial kernel size may lose some large features such as large vortices or LCS structures. Figure 1c shows that we only capture many small size vortices although some of them may be artifacts due to a relatively small r. Increasing r may not address this limitation, as some small but important vortices could be smoothed out. Figure 14 shows the results of different spatial kernel sizes on the hurricane simulation. As we can see, using a larger kernel size removes some shearing effect in the upper right of the images. The temporal length may affect the completeness of the captured features. The larger the h is the more pronounced the structure will be. Figure 15 shows the result of double gyre using various temporal kernel lengths. Comparing these results we can see that after increasing the temporal window size over a certain threshold, the results are somewhat converged.

6 Conclusion and Future Work

In this paper, we study the pairwise relations among attributes in unsteady flow. Our approach provides a general and powerful way to explore linear relation, i.e., correlations and non-linear relation, i.e., dependency among attribute-pairs derived from unsteady flow. We also compared the relations of attribute value sequences in a fixed location over time (Eulerian) or along pathlines (Lagrangian). We find that in

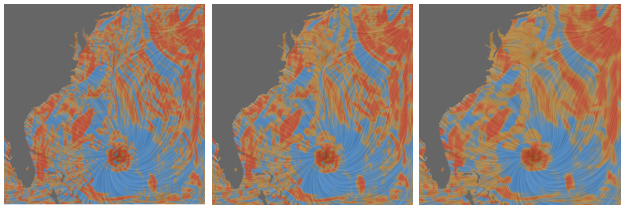

(a) (b) (c)

Fig. 14: The effect of spatial kernel size, r, on Lagrangian ST_LCC. From the left image to right, the r values for the results are 3, 5, and 10, respectively.

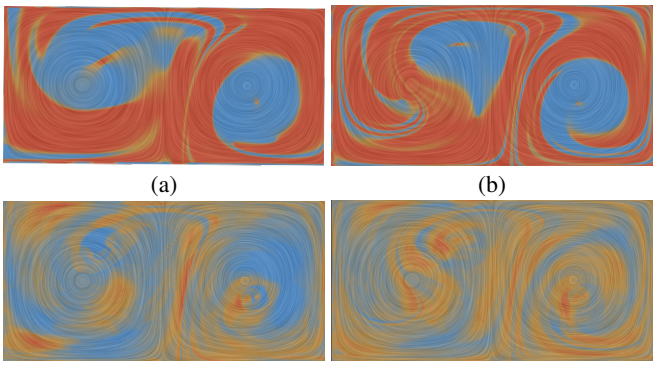

(a) (b) (c) (d)

Fig. 15: The effect of different temporal kernel lengths, h, for ST_LCC (a-b) and ST_MI (c-d). The h values are 50 (left column) and 100 (right column), respectively, considering $\tau = 0.05$.

general, the correlation and dependency fields reveal more coherent and meaningful flow structures in the Lagrangian view than in the Eulerian view. Finally, we applied various metrics to extract the statistical linear and non-linear relation and dependency for a set of attributes in a number of unsteady flow. Our results indicate that our ST_LCC is good at providing overview of the linear correlation among attributes, while mutual information (MI) can reveal more complex and non-linear relations. In addition, MI is scale-invariant and can measure relation between three or more attributes by expanding the conditional and joint entropy, which we plan to explore in the future. Furthermore, we may include tensor-valued attributes, such as the Jacobian into our framework as long as a proper tensor similarity metric can be developed.

Although our proposed method works across many datasets, there are some limitations. First, we use a fixed kernel size throughout the computation for one dataset, which may cause the loss of some important features that often have varying scales in space and time. Second, our method is sensitive to the temporal sampling rate provided by the simulation. If the temporal sampling rate is coarse, we see a lot of fluctuation in the pathline computation which interpolation cannot alleviate. We plan to address these limitations in the future and extend our framework to study the attribute relations in ensemble flow data. Finally, we plan to migrate our current implementation onto clusters so that extremely large-scale data can be processed more effectively.

Acknowledgment

We thank all the anonymous reviewers for their constructive comments. This research was supported by NSF IIS 1553329.

References

[1] G. Haller. Lagrangian coherent structures and the rate of strain in two-dimensional turbulence. *Phys. Fluids A*, 13:3365–3385, 2001.

[2] Kuangyu Shi, Holger Theisel, Helwig Hauser, Tino Weinkauf, Kresimir Matkovic, Hans-Christian Hege, and Hans-Peter Seidel. Path line attributes

- an information visualization approach to analyzing the dynamic behavior of 3D time-dependent flow fields. In Hans-Christian Hege, Konrad Polthier, and Gerik Scheuermann, editors, *Topology-Based Methods in Visualization II*, Mathematics and Visualization, pages 75–88, Grimma, Germany, 2009. Springer.

[3] T. Weinkauf, H. Theisel, and O. Sorkine. Cusps of characteristic curves and intersection-aware visualization of path and streak lines. In R. Peikert, H. Hauser, H. Carr, and R. Fuchs, editors, *Topological Methods in Data Analysis and Visualization II*, Mathematics and Visualization, pages 161–176. Springer, 2012.

[4] Harsh Bhatia, Valerio Pascucci, Robert M Kirby, and P-T Bremer. Extracting features from time-dependent vector fields using internal reference frames. In *Computer Graphics Forum*, volume 33, pages 21–30. Wiley Online Library, 2014.

[5] Raphael Fuchs, Jan Kemmler, Benjamin Schindler, Jürgen Waser, Filip Sadlo, Helwig Hauser, and Ronald Peikert. Toward a lagrangian vector field topology. In *Computer Graphics Forum*, volume 29, pages 1163–1172. Wiley Online Library, 2010.

[6] Kuangyu Shi, Holger Theisel, Tino Weinkauf, Hans-Christian Hege, and Hans-Peter Seidel. Visualizing transport structures of time-dependent flow fields. *IEEE Computer Graphics and Applications*, 28(5):24–36, 2008.

[7] Julian CR Hunt, AA Wray, and Parviz Moin. Eddies, streams, and convergence zones in turbulent flows. In *Studying Turbulence Using Numerical Simulation Databases, 2*, volume 1, pages 193–208, 1988.

[8] Jinhee Jeong and Fazle Hussain. On the identification of a vortex. *Journal of fluid mechanics*, 285:69–94, 1995.

[9] Natascha Sauber, Holger Theisel, and H-P Seidel. Multifield-graphs: An approach to visualizing correlations in multifield scalar data. *Visualization and Computer Graphics, IEEE Transactions on*, 12(5):917–924, 2006.

[10] Josien Petra Willemijn Pluim. *Mutual information based registration of medical images*. uitgever niet vastgesteld, 2001.

[11] Min Chen and Heike Jaenicke. An information-theoretic framework for visualization. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1206–1215, 2010.

[12] Soumya Dutta, Xiaotong Liu, Ayan Biswas, Han-Wei Shen, and Jen-Ping Chen. Pointwise information guided visual analysis of time-varying multifields. In *SIGGRAPH Asia 2017 Symposium on Visualization*, page 17. ACM, 2017.

[13] Armin Pobitzer, Ronald Peikert, Raphael Fuchs, Benjamin Schindler, Alexander Kuhn, Holger Theisel, Kresimir Matkovic, and Helwig Hauser. The state of the art in topology-based visualization of unsteady flow. *Computer Graphics Forum*, 30(6):1789–1811, September 2011.

[14] F. H. Post, B. Vrolijk, H. Hauser, R. S. Laramee, and H. Doleisch. The state of the art in flow visualization: feature extraction and tracking. *Computer Graphics Forum*, 22(4):775–792, Dec. 2003.

[15] X. Tricoche, G. Scheuermann, and H. Hagen. Topology-based visualization of time-dependent 2D vector fields. In *Proceedings of the Joint Eurographics - IEEE TCVG Symposium on Visualization (VisSym '01)*, pages 117–126, May 28–30 2001.

[16] Xavier Tricoche, Thomas Wischgoll, Gerik Scheuermann, and Hans Hagen. Topology tracking for the visualization of time-dependent two-dimensional flows. *Computers & Graphics*, 26(2):249–257, 2002.

[17] Holger Theisel and H-P Seidel. Feature flow fields. In *Proceedings of the symposium on Data visualisation 2003*, pages 141–148. Eurographics Association, 2003.

[18] Holger Theisel, Tino Weinkauf, Hans-Christian Hege, and Hans-Peter Seidel. Stream line and path line oriented topology for 2d time-dependent vector fields. In *Proceedings of the conference on Visualization'04*, pages 321–328. IEEE Computer Society, 2004.

[19] F. Lekien, S.C. Shadden, and J.E. Marsden. Lagrangian coherent structures in n-dimensional systems. *Journal of Mathematical Physics*, 48(6):Art. No. 065404, 2007.

[20] S.C. Shadden, F. Lekien, and J.E. Marsden. Definition and properties of lagrangian coherent structures from finite-time lyapunov exponents in two-dimensional aperiodic flows. *Physica D*, 212(3–4):271–304, 2005.

[21] C. Garth, A. Wiebel, X. Tricoche, K. I. Joy, and G. Scheuermann. Lagrangian visualization of flow-embedded surface structures. *Computer Graphics Forum*, 27(3):1007–1014, 2008.

[22] Filip Sadlo and Daniel Weiskopf. Time-dependent 2-d vector field topology: An approach inspired by lagrangian coherent structures. *CoRR*, abs/1105.5678, 2011.

[23] Thomas Ertl, Filip Sadlo, and Markus Uffinger. A time-dependent vector field topology based on streak surfaces. *IEEE Transactions on Visualization and Computer Graphics*, 19(3):379–392, 2013.

[24] Raphael Fuchs, Jan Kemmler, Benjamin Schindler, Filip Sadlo, Helwig Hauser, and Ronald Peikert. Toward a Lagrangian Vector Field Topology. *Computer Graphics Forum*, 29(3):1163–1172, 2010.

[25] Ming Jiang, Raghu Machiraju, and David Thompson. Detection and visualization of vortices. In *The Visualization Handbook*, pages 295–309. Academic Press, 2005.

[26] Raphael Fuchs, Ronald Peikert, Helwig Hauser, Filip Sadlo, and Philipp Muigg. Parallel vectors criteria for unsteady flow vortices. *IEEE Transactions on Visualization and Computer Graphics*, 14(3):615–626, 2008.

[27] J. Sahner, T. Weinkauf, and H.-C. Hege. Galilean invariant extraction and iconic representation of vortex core lines. In K. Joy K. Brodlie, D. Duke, editor, *Proc. Eurographics / IEEE VGTC Symposium on Visualization (EuroVis '05)*, pages 151–160, Leeds, UK, June 2005.

[28] T. Weinkauf and H. Theisel. Streak lines as tangent curves of a derived vector field. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1225–1234, November - December 2010.

[29] R Cucitore, M Quadrio, and A Baron. On the effectiveness and limitations of local criteria for the identification of a vortex. *European Journal of Mechanics-B/Fluids*, 18(2):261–282, 1999.

[30] George Haller. An objective definition of a vortex. *Journal of Fluid Mechanics*, 525:1–26, 2005.

[31] Harsh Bhatia, Valerio Pascucci, and Peer-Timo Bremer. The natural helmholtz-hodge decomposition for open-boundary flow analysis. *IEEE Transactions on Visualization and Computer Graphics*, 20(11):1566–1578, 2014.

[32] Tobias Günther, Maik Schulze, and Holger Theisel. Rotation invariant vortices for flow visualization. *IEEE transactions on visualization and computer graphics*, 22(1):817–826, 2016.

[33] Franz Sauer and Kwan-Liu Ma. Spatio-temporal feature exploration in combined particle/volume reference frames. *IEEE Transactions on Visualization and Computer Graphics*, 23(6):1624–1635, 2017.

[34] Armin Pobitzer, Alan Lez, Kresimir Matkovic, and Helwig Hauser. A statistics-based dimension reduction of the space of path line attributes for interactive visual flow analysis. In *PacificVis*, pages 113–120, 2012.

[35] Marcel Hlawatsch, Filip Sadlo, Hajun Jang, and Daniel Weiskopf. Pathline glyphs. In *Computer Graphics Forum*, volume 33, pages 497–506. Wiley Online Library, 2014.

[36] Teng-Yok Lee and Han-Wei Shen. Visualization and exploration of temporal trend relationships in multivariate time-varying data. *IEEE Transactions on Visualization and Computer Graphics*, 15(6), 2009.

[37] Yi Gu and Chaoli Wang. A study of hierarchical correlation clustering for scientific volume data. *Advances in Visual Computing*, pages 437–446, 2010.

[38] Zhiyuan Zhang, Kevin T McDonnell, Erez Zadok, and Klaus Mueller. Visual correlation analysis of numerical and categorical data on the correlation map. *IEEE transactions on visualization and computer graphics*, 21(2):289–303, 2015.

[39] Cheng-Kai Chen, Chaoli Wang, Kwan-Liu Ma, and Andrew T Wittenberg. Static correlation visualization for large time-varying volume data. In *Pacific Visualization Symposium (PacificVis), 2011 IEEE*, pages 27–34. IEEE, 2011.

[40] Huijie Zhang, Yafang Hou, Dezhan Qu, and Quanle Liu. Correlation visualization of time-varying patterns for multi-variable data. *IEEE Access*, 4:4669–4677, 2016.

[41] Min Chen, Miquel Feixas, Ivan Viola, Anton Bardera, Han-Wei Shen, and Mateu Sbert. *Information Theory Tools for Visualization*. CRC Press, 2016.

[42] Claude Elwood Shannon. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1):3–55, 2001.

[43] Lijie Xu, Teng-Yok Lee, and Han-Wei Shen. An information-theoretic framework for flow visualization. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1216–1224, 2010.

[44] Ayan Biswas, Soumya Dutta, Han-Wei Shen, and Jonathan Woodring. An information-aware framework for exploring multivariate data sets. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2683–2692, 2013.

[45] Ko-Chih Wang, Kewei Lu, Tzu-Hsuan Wei, Naeem Shareef, and Han-Wei Shen. Statistical visualization and analysis of large data using a value-based spatial distribution. In *Pacific Visualization Symposium (PacificVis), 2017 IEEE*, pages 161–170. IEEE, 2017.

[46] Huijie Zhang, Dezhan Qu, Quanle Liu, Qi Shang, Yafang Hou, and Han-

Wei Shen. Uncertainty visualization for variable associations analysis. *The Visual Computer*, pages 1–19, 2017.

[47] Marzieh Berenjkoub Lei Zhang, Guoning Chen. Correlation study on attributes of unsteady flows. In *2017 IEEE Visualization Conference (VIS), 1-6 October, Phoenix, AZ, USA*, 2017.

[48] Tobias Günther and Holger Theisel. The state of the art in vortex extraction. In *Computer Graphics Forum*. Wiley Online Library.

[49] Stephen M. Stigler. Francis galton's account of the invention of correlation. *Statistical Science*, 4(2):73–79, 05 1989.

[50] Lei Zhang, Duong Nguyen, David Thompson, Robert Laramee, and Guoning Chen. Enhanced vector field visualization via lagrangian accumulation. *Computers & Graphics*, 2017.

[51] Alain Crouzil, Ouis Massip-pailhes, and Serge Castan. A new correlation criterion based on gradient fields similarity. In *Pattern Recognition, 1996., Proceedings of the 13th International Conference on*, volume 1, pages 632–636. IEEE, 1996.

[52] Cambridge Aerospace Series. Principles of helicopter aerodynamics; 2000; 5 pages; new york, ny us.

[53] Bernard W Silverman. *Density estimation for statistics and data analysis*. Routledge, 2018.

[54] Charles HK Williamson. Vortex dynamics in the cylinder wake. *Annual review of fluid mechanics*, 28(1):477–539, 1996.

[55] Shawn C Shadden, Francois Lekien, and Jerrold E Marsden. Definition and properties of lagrangian coherent structures from finite-time lyapunov exponents in two-dimensional aperiodic flows. *Physica D: Nonlinear Phenomena*, 212(3):271–304, 2005.

[56] Jens Kasten, Jan Reininghaus, Ingrid Hotz, and Hans-Christian Hege. Two-dimensional time-dependent vortex regions based on the acceleration magnitude. *Transactions on Visualization and Computer Graphics (Vis'11)*, 17(12):2080–2087, 2011.

[57] Primoz Skraba, Bei Wang, Guoning Chen, and Paul Rosen. 2d vector field simplification based on robustness. In *Pacific Visualization Symposium (PacificVis), 2014 IEEE*, pages 49–56. IEEE, 2014.

[58] Jacqueline H Chen, Evatt R Hawkes, Ramanan Sankaran, Scott D Mason, and Hong G Im. Direct numerical simulation of ignition front propagation in a constant volume with temperature inhomogeneities: I. fundamental analysis and diagnostics. *Combustion and flame*, 145(1-2):128–144, 2006.

[59] NASA. Estimating the Circulation and Climate of the Ocean, Phase II. `http://ecco2.org`. [].

[60] Paolo Orlandi and Roberto Verzicco. Vortex rings impinging on walls: axisymmetric and three-dimensional simulations. *Journal of Fluid Mechanics*, 256:615–646, 1993.

[61] Simone Camarri, M Buffoni, A Iollo, and Maria Vittoria Salvetti. Simulation of the three-dimensional flow around a square cylinder between parallel walls at moderate reynolds numbers. In *XVII Congresso AIMeTA di Meccanica Teorica e Applicata*, volume 1, pages 23–34. -Firenze University Press, Florence, Italy, 2005.

[62] Nathanaël Schaeffer and Stéphane Le Dizès. Nonlinear dynamics of the elliptic instability. *Journal of Fluid Mechanics*, 646:471–480, 2010.

[63] Fazle Hussain and Karthik Duraisamy. Mechanics of viscous vortex reconnection. *Physics of Fluids*, 23(2):021701, 2011.