

## Lagrangian based methods for coherent structure detection

Michael R. Allshouse and Thomas Peacock

Citation: *Chaos* **25**, 097617 (2015); doi: 10.1063/1.4922968

View online: <http://dx.doi.org/10.1063/1.4922968>

View Table of Contents: <http://scitation.aip.org/content/aip/journal/chaos/25/9?ver=pdfcov>

Published by the [AIP Publishing](#)

---

### Articles you may be interested in

[Reaction enhancement of initially distant scalars by Lagrangian coherent structures](#)

*Phys. Fluids* **27**, 035106 (2015); 10.1063/1.4914467

[Dependence of advection-diffusion-reaction on flow coherent structures](#)

*Phys. Fluids* **25**, 106602 (2013); 10.1063/1.4823991

[Lagrangian coherent structures and internal wave attractors](#)

*Chaos* **20**, 017508 (2010); 10.1063/1.3273054

[Computational analysis of an aortic valve jet with Lagrangian coherent structures](#)

*Chaos* **20**, 017512 (2010); 10.1063/1.3272780

[Invariant-tori-like Lagrangian coherent structures in geophysical flows](#)

*Chaos* **20**, 017514 (2010); 10.1063/1.3271342

---



## Lagrangian based methods for coherent structure detection

Michael R. Allshouse<sup>1,a)</sup> and Thomas Peacock<sup>2,b)</sup>

<sup>1</sup>*Center for Nonlinear Dynamics and Department of Physics, University of Texas at Austin, Austin, Texas 78712, USA*

<sup>2</sup>*Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

(Received 23 April 2015; accepted 4 June 2015; published online 8 July 2015)

There has been a proliferation in the development of Lagrangian analytical methods for detecting coherent structures in fluid flow transport, yielding a variety of qualitatively different approaches. We present a review of four approaches and demonstrate the utility of these methods via their application to the same sample analytic model, the canonical double-gyre flow, highlighting the pros and cons of each approach. Two of the methods, the geometric and probabilistic approaches, are well established and require velocity field data over the time interval of interest to identify particularly important material lines and surfaces, and influential regions, respectively. The other two approaches, implementing tools from cluster and braid theory, seek coherent structures based on limited trajectory data, attempting to partition the flow transport into distinct regions. All four of these approaches share the common trait that they are objective methods, meaning that their results do not depend on the frame of reference used. For each method, we also present a number of example applications ranging from blood flow and chemical reactions to ocean and atmospheric flows.

© 2015 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4922968>]

**The transport of material by advection in fluid systems is a process of vital and ubiquitous importance, underlying scenarios as diverse as pollutant distribution and search-and-rescue operations in the ocean to blood flow in the human body. The rapidly advancing field of Lagrangian based methods for studying flow transport has demonstrated an effective ability to find robust and significant transport features that underly the organization of flow transport in complex, unsteady flow fields. In this review, we present an overview of four of the leading Lagrangian approaches, each with their own strengths and challenges. Details of each method are presented along with an example application to the same model system, the double-gyre flow. Furthermore, we highlight a number of exciting applications and future directions to bring Lagrangian based analysis closer to implementation in real-time, real-world decision making strategies.**

### I. INTRODUCTION

Nature is replete with both inspiring and practically important examples of coherent structures in unsteady flow transport. During the Deepwater Horizon disaster, a 200 km long surface oil filament was dramatically ejected from the main body of the spill over the course of a couple days with the potential to enter the loop current and carry contamination along the east coast of America;<sup>58</sup> despite the turbulent dynamics in Jupiter's atmosphere, the Great Red Spot remains seemingly stable and ever-present;<sup>35</sup> throughout the oceans, vast garbage patches continue to grow in size posing an increasingly significant environmental hazard;<sup>78</sup> and

Agulhas rings transport large quantities of water from the Indian Ocean hundreds-to-thousands of kilometers across the Atlantic Ocean with little significant mixing.<sup>32,41,79</sup> Examples such as these are a driving force behind the development of novel Lagrangian mathematical methods that can identify persistent coherent transport features within even highly unsteady flows, and furthermore can assess the role that such structures play in the overall flow transport. There is the exciting potential that such methods may also yield new predictive capabilities and enable new Lagrangian-based control strategies.

Although there is a rich history of using Eulerian (i.e., instantaneous and field-based) methods for identifying coherent structures in fluid flows, a well known example being the Okubo-Weiss criterion for vortex detection,<sup>56,80</sup> such methods generically come up short when it comes to understanding transport in unsteady flows, for two primary reasons. First, a flow feature that is apparent via the instantaneous velocity field (e.g., in a streamline plot or a vorticity contour plot) may persist for only a short period of time compared to the unsteady timescale of the flow, and as such will have little impact on the actual flow transport. Second, Eulerian-based methods for detecting coherent structures are not objective, i.e., their results depend on the frame of reference used to view the flow field. For example, the velocity field of a flow that appears rotationally dominated in one frame of reference may appear very different from a co-rotating frame of reference.<sup>38</sup> The organization of flow transport, however, is frame independent; if a patch of dyed fluid is released in a flow field, the shape that it assumes as it is advected by the flow is the same irrespective of the frame of reference from which it is viewed (presuming we do not enter relativistic regimes!). As such, methods for detecting coherent transport structures must also be objective.

<sup>a)</sup>Electronic mail: mallshouse@chaos.utexas.edu.

<sup>b)</sup>Electronic mail: tomp@mit.edu.

Given the limitations of, and potential for misdirection by, Eulerian-based methods, there has been a rise in the development of Lagrangian tools to identify coherent structures that organize flow transport.<sup>81</sup> Broadly speaking, there are two categories of Lagrangian approaches that are being pursued. The first category utilizes time-varying velocity field data, which may come from a numerical simulation, laboratory experiment or high-frequency-radar ocean-monitoring system, for example, to calculate the necessary trajectory information. We refer to these as “dense methods” because of the relatively large number of trajectories needed for the calculations. The second category, which we refer to as “sparse methods,” considers the trajectories of only a limited number of advected particles, with typical data sources being particle tracking in laboratory experiments, ocean drifter data sets, or a limited number of numerically calculated trajectories. These approaches each have their own, often related, definitions of what constitutes a coherent structure.

The types of coherent structure sought by the Lagrangian approaches can also be roughly broken into two categories. One type of coherent structure is a region of the fluid that does not significantly mix with the rest of the domain. In this case, one can envisage drawing a boundary around such a region so that there is only modest deformation of the region, as it is advected and all the fluid elements within the region remain in close proximity; examples include the Great Red Spot on Jupiter and the coherent Agulhas rings that travel across the Atlantic Ocean. The second type of coherent structure is a region of the fluid that causes a significant amount of local deformation as it is advected. Such regions attract and/or repel large amounts of nearby fluid; the rapid filament developed during the Deepwater Horizon spill is evidence of the presence of such a structure.

Here, we present an overview of four leading, objective Lagrangian methods based on the aforementioned concepts. First, two of the most prominent and well developed “dense” approaches are summarized; the geometric perspective is presented in Sec. II and the probabilistic approach in Sec. III. Then we present an overview of two recently developed “sparse” approaches in Sec. IV; the cluster and braid theory approaches. For each approach, we demonstrate the corresponding method via application to the double-gyre analytic system<sup>69</sup> and present a number of example applications. In this review, we focus on 2D flows and discuss extensions to 3D flows. Finally, we conclude in Sec. V by considering outstanding challenges and some important research directions.

## II. GEOMETRIC COHERENT STRUCTURES

Geometric approaches seek to identify key material lines in 2D flows (and, correspondingly, material surfaces in 3D flows) that are distinguished by the dominant nature of flow transport in their vicinity. These structures are referred to as Lagrangian coherent structures (LCS).<sup>36,40–42</sup> The properties of the right Cauchy-Green (CG) strain tensor field are the basis for the definition of these structures. Geometric approaches have recently been comprehensively reviewed,<sup>39,61</sup> and here, we give a sufficiently detailed synopsis for 2D flows to

provide the interested reader with an overview of the key concepts, and provide references for details on 3D flows.

Given a velocity field data set, obtaining the CG tensor field for a given spatial domain and time interval  $[t_0, t]$  requires three steps. The first step is to obtain the flow map  $\mathbf{F}_{t_0}^t(\mathbf{x}_0) = \mathbf{x}(t; \mathbf{x}_0, t_0)$ , which maps a fluid element from its initial position  $\mathbf{x}_0$  at time  $t_0$  to its final position  $\mathbf{x}$  at time  $t$ . For real-world flows, the flow map is not generally an analytically derivable function and has to be calculated numerically by solving

$$\frac{d\mathbf{x}}{dt} = \mathbf{u}(\mathbf{x}, t)$$

for a set of passive tracers that follow the flow field  $\mathbf{u}(\mathbf{x}, t)$ , which itself is usually provided as a spatiotemporally discretized data set (e.g., from a numerical model). A standard package for numerical integration is `ode45` in Matlab, which allows for the necessary interpolation from the velocity field data set; care should be taken to ensure that sufficiently stringent tolerances are used for convergence of the numerical integration.<sup>4</sup>

The second step is to calculate the flow map gradient

$$\nabla \mathbf{F}_{t_0}^t(\mathbf{x}_0) = \partial x_i / \partial x_{0,j} \quad \text{for } i, j = 1, 2,$$

which is the spatial derivative of the flow map with respect to the initial tracer particle location (i.e., how much does the final position of a tracer particle change with variation in its initial position). This flow map gradient is a local linearization of the flow transport and has the property that it maps an infinitesimal vector  $\epsilon$  originating at the initial location  $\mathbf{x}_0$  to its final state (i.e., length and orientation) originating at the final location  $\mathbf{F}_{t_0}^t(\mathbf{x}_0)$ . Again, care must be taken when numerically calculating the flow map gradient, and the so-called auxiliary grid method<sup>22</sup> has proven to be an effective way to achieve this.<sup>4</sup>

The third, and final, step is to calculate the CG tensor field from the flow-map gradient field via

$$\mathbf{C}_{t_0}^t(\mathbf{x}_0) = [\nabla \mathbf{F}_{t_0}^t]^\top [\nabla \mathbf{F}_{t_0}^t],$$

where  $^\top$  corresponds to the transpose operator. The eigenvectors,  $\xi_i$ , and eigenvalues,  $\lambda_i$ , of the CG tensor provide the fundamental information regarding stretching due to advection by the flow field. More specifically, if one considers an infinitesimal circle released at location  $\mathbf{x}_0$ , the initial orientation of the principal axes of the resulting ellipse produced by advection and the amount of stretching along these axes are represented by the eigenvectors and eigenvalues, respectively. The eigenvectors and eigenvalues of the CG tensor field form the basis of all the following geometric approaches for identifying key coherent structures. By construction, the eigenvectors of the CG tensor are orthogonal, and if the flow is incompressible then the product of the eigenvalues is unity (i.e.,  $\lambda_1 \lambda_2 = 1$ ,  $\lambda_1 \leq \lambda_2$ ).

An important feature of this type of analysis is that it can also be applied in reverse time (i.e., from the end of a time interval  $t$  to the start of a time interval  $t_0$ ). This provides a consistent method to identify repelling and attracting

material structures based upon their influence in either forwards or backwards time for the same time interval (e.g., an attracting material line in forwards time is a repelling material line in backwards time, and vice versa).

### A. Finite-time Lyapunov exponents

A basic implementation of the geometric approach is to plot the finite-time Lyapunov Exponent (FTLE) field

$$FTLE_{t_0}^t(\mathbf{x}_0) = \frac{1}{2(t-t_0)} \log \lambda_2(\mathbf{x}_0),$$

which is a rescaled version of the largest eigenvalue field  $\lambda_2(\mathbf{x}_0)$  of  $C_{t_0}^t$ . By definition, this field identifies the regions of greatest relative stretching of material elements. Figure 1 presents the FTLE field for the 2D double-gyre flow advected for the four period time interval [2.5, 42.5] with parameters  $A = 0.1$ ,  $\epsilon = 0.1$ , and  $\omega = \pi/5$ . The standout features of the figure are the FTLE ridges, which correspond to a heuristic diagnostic for identifying hyperbolic LCSs.<sup>36,37</sup> Since the analysis was carried out for forwards time, it reveals the location of repelling (i.e., stretching) ridges at time  $t_0$ ; reverse time calculations would reveal the location at time  $t$  of a similarly complex network of repelling ridges in backwards time, these being the final states of attracting ridges for forwards-time advection over the time window being considered. Material placed in the vicinity of a forwards-time FTLE ridge will undergo substantial stretching and is drawn on to attracting FTLE ridges; conversely, material released in regions of low FTLE values will be advected with little deformation. To locate these structures at any instant throughout the given time interval, it is necessary to advect the ridges with the flow map, as opposed to the historically oft-used approach of repeating analysis on a shifted time interval.<sup>69</sup> If a direct comparison of repelling and attracting structures is desired, for example, the same time interval must be used for calculation and advection of the structures to the same time instant must be implemented since the forwards-time calculations reveal the positions of repelling structures at  $t_0$  and backwards-time calculations reveal the positions of attracting structures at  $t$ .

While the FTLE calculation is an effective and simple method for investigating flow transport, there are several caveats. For example, it has proven challenging to effectively pinpoint the material lines that run along the FTLE ridges as these are, by definition, the most repelling features

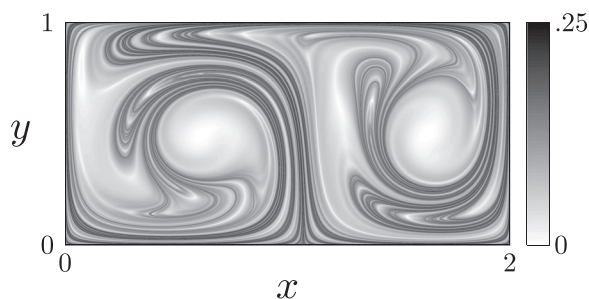


FIG. 1. The FTLE field calculated for the double-gyre system over the time interval [2.5, 42.5] for the parameters  $A = 0.1$ ,  $\epsilon = 0.1$ , and  $\omega = \pi/5$ .

in the entire flow transport field. Effort has been made to identify the FTLE ridges,<sup>4,43,47,50,67,69</sup> but small errors in the initial location of a repelling ridge or the final location of an attracting ridge, which are the basic results of FTLE analysis, are greatly amplified by advection. Given this property, it is challenging to determine the location of the repelling and attracting ridges at intermediate times. It is also the case that FTLE analysis does not reveal what type of stretching takes place, so the impact of the ridges on the surrounding material is unclear from looking at the FTLE field alone. Finally, due to its potential for distortion while being advected, an FTLE ridge does not necessarily, and most likely does not, represent a long lived transport barrier. Indeed, many FTLE ridges simply represent a record of a collection of material elements that flow past a significant, localized Lagrangian flow feature, such as a finite-time hyperbolic core.<sup>58</sup>

### B. Hyperbolic, parabolic, and elliptic geometric structures

In seeking a more rigorous definition of LCS, different types of material lines and surfaces have been identified based on their local<sup>23,40</sup> and global properties.<sup>40,41</sup> By “local,” we mean that these material lines are defined in such a way that for all points in the material line their tangent space maximizes a particular type of local deformation. The “global” perspective uses geodesics (i.e., shortest paths under a particular metric) to identify unique material lines that demonstrate particular types of deformation. While these approaches take different perspectives, some common results are reached, with the global approach providing additional types of structures.

For 2D flows, there are two types of lines that maximize and minimize in-line stretching. Stretch lines are material lines that follow the  $\xi_2$  field (i.e., the eigenvector field corresponding to the largest eigenvalue) for the time interval considered and are the class of material lines that will undergo the most tangential (i.e., in-line) stretching. By virtue of the properties of the CG tensor, the  $\xi_1$  field is always perpendicular to stretch lines, and so the direction normal to a stretch line is a direction of minimal stretching (indeed, most likely, compression, which is guaranteed for incompressible flows). Shrink lines are material lines that follow the  $\xi_1$  field and are material lines that will undergo the minimal in-line stretching. Both shrink lines and stretch lines are hyperbolic material lines, and the kinematically most important such lines are those that have the maximum normal repulsion (for shrink lines) and maximum normal attraction (for stretch lines). The location of shrink and stretch lines can be found in both forwards and backwards time, in which case their locations are determined at the start and end of the time interval, respectively.<sup>23</sup> Their locations at intermediate times are found by advection, which can be unstable in the case of shrink lines, which are strongly normally repelling. Select shrink lines for the double-gyre system are presented in red in Figure 2, closely aligning with the ridges of the FTLE field in this case, although there is no formal requirement that shrink lines and stretch lines need to align with or lie perpendicular to FTLE ridges, respectively.

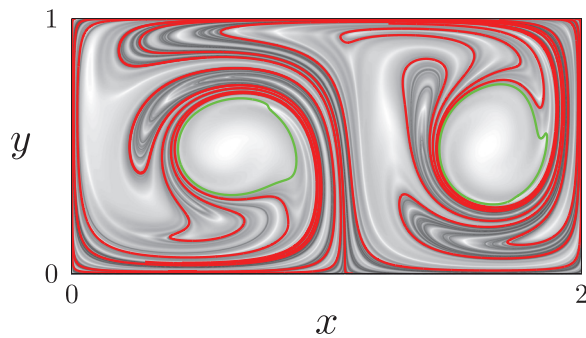


FIG. 2. The stretch lines (red) and objective vortices (green) calculated for the double-gyre system over the time interval  $[2.5, 42.5]$  for the parameters  $A=0.1$ ,  $\epsilon=0.1$ , and  $\omega=\pi/5$ .

In addition to normal repulsion and attraction, another important class of material lines focuses on Lagrangian shear, which is shear that occurs in the reference frame of the trajectories. For the “local” based analysis, it is possible to determine the tangent vector field that will maximize local Lagrangian shear at all points in the domain, and material lines tangent to this field are referred to as shear lines.<sup>40</sup> In some cases, these shear lines will form closed material lines that act as elliptic barriers. While this is possible, the rigid requirement that the line be Lagrangian shear maximizing is relaxed in the “global” analysis.<sup>41</sup> In this case, the closed material lines are allowed to stretch uniformly and are referred to as  $\lambda$ -lines, which are tangent to the vector field

$$\eta_{\pm} = \sqrt{\frac{\lambda_2(x_0) - \lambda^2}{\lambda_2(x_0) - \lambda_1(x_0)}} \xi_1(x_0) \pm \sqrt{\frac{\lambda^2 - \lambda_1(x_0)}{\lambda_2(x_0) - \lambda_2(x_0)}} \xi_2(x_0).$$

It can be shown that this vector field corresponds to the shear line vector field in the case where  $\lambda=1$  and the system is incompressible. Of particular interest are closed  $\lambda$ -lines, in which case the outermost of a series of closed orbits marks the boundary of a Lagrangian vortex that traps the fluid within it as it advects. The presence of closed orbits is linked to the existence of singularities in the CG tensor field, which is used as a means for identifying their location.<sup>44</sup> By virtue of their transport properties, these are also considered elliptical LCS. There are two elliptic barriers in the double-gyre system, and they are presented as green lines in Figure 2.

Finally, parabolic LCS take the form of so-called jets cores, composed of alternating sequences of shrink and stretch lines. These structures are considered to be geometrically robust, and if located they can serve as important transport barriers. A clear example of the role of a shearless jet in organizing flow transport is seen by studying the classical Bickley-jet problem,<sup>24</sup> but in the case of the double-gyre system there are no parabolic LCS.

All of the aforementioned LCS discussion has focused on structures in 2D flows. Efforts are well underway to define and locate these structures in 3D flows, which is inherently more computationally challenging. For example, the current state-of-the-art calculates the CG tensor for the 3D system, and then calculates the normally hyperbolic LCS material surfaces on a family of 2D slices of the domain; the results are merged to form 2D surfaces in the 3D domain.

Similarly, closed 2D elliptic surfaces within a 3D space are identified by finding closed elliptic barriers in the family of 2D slices.<sup>11</sup> In addition to this transition to 3D calculations, a tool set to calculate many of these material lines has been developed for Matlab.<sup>3,60</sup>

### C. Applications of the geometric approach

There have been a large number of applications of the geometric techniques to ocean surface flows; the following are just a small sample, and more complete reviews of ocean applications exist.<sup>39,65</sup> Using the FTLE ridge approach, it was determined that there was a persistent barrier off the West Florida Coast that periodically insulated the region from mixing. The long residence time of water near the coast occasionally enabled the production of phytoplankton blooms resulting in red tides.<sup>57,59</sup> Another pair of studies in the Gulf of Mexico focused on applying the shrink line approach and identified hyperbolic cores that dominate the local stretching in the field, shedding some insight into the “tiger-tail” filament formed during the Deepwater horizon spill.<sup>20,58</sup> Finally, elliptic barriers were used to locate and track Agulhas rings formed off the southern tip of Africa. These rings were identified as Lagrangian vortices and the structures themselves survived for much longer than the time interval used for calculation.<sup>41,79</sup>

There have also been a number of applications of the geometric methods to atmospheric flows. In a pair of studies analyzing LIDAR data of currents above Hong Kong International Airport, FTLE ridges were identified and their potential impact on aircraft was quantified.<sup>73,74</sup> A preliminary attempt was made at classifying the types of local deformation near these ridges, but no attempt was made at identifying any of the other class of structures. Another atmospheric application of FTLE ridges looked into the impact of Lagrangian transport structures on airborne microbial populations.<sup>70</sup> In conjunction with atmospheric calculations, a remote control aircraft was flown through different regions collecting samples of the microbial populations; the authors were able to identify significant shifts in the population sizes either side of an FTLE surface. Because atmospheric Lagrangian structures can play a significant role in airborne transport, further studies were made investigating the uncertainty of the FTLE calculation and the contributions of unresolved turbulence and other forecast uncertainties.<sup>13,14</sup>

The geometric techniques have been utilized to study several different biological systems. One of the first studies considered flows induced by a swimming jellyfish.<sup>62</sup> By considering escape forces, it was possible to identify regions of the flow where the jellyfish captured particular types of prey based on their ability to escape. Another investigation looked into the flows generated by biologically inspired geometries. Using an experimental set up that pitched a panel at various angles representing fish fins, FTLE fields were calculated to investigate the resulting vortex shedding.<sup>34</sup> Finally, a sequence of studies investigating cardiac blood flow has helped shed light on a number of important problems. An initial study looked at the flow dynamics near an abdominal

aortic aneurysm.<sup>7</sup> Another study utilized FTLE fields to investigate the vortex that develops in the heart’s left ventricle during pumping.<sup>21</sup> There is a recent review of the application of geometric techniques to hemodynamics.<sup>68</sup>

Finally, several studies have applied geometric methods to turbulent experimental data. In a study of a quasi-two-dimensional turbulent system, the overlapping of regions of attracting and repelling structures were designated as hyperbolic cores, or regions of significant Lagrangian hyperbolic deformation.<sup>50</sup> A study of quasi-two-dimensional turbulence investigated the relationship between the direction of scale-to-scale energy transport and strainlines. It was found that shrink lines delineate regions where energy moves up and down the energy cascade.<sup>46</sup>

### III. PROBABILISTIC COHERENT STRUCTURES

A second methodology for identifying coherent structures in flow transport relies on the transfer (or Perron-Frobenius) operator, and we refer to this as the probabilistic approach. For a given time interval, this approach identifies regions of the flow domain for which there is a high probability of starting in one region and ending in another. The structures identified are partitions of the fluid that are advected through the system without significant mixing occurring between material within and without the identified partitions. The transfer operator was initially used to identify stationary structures in autonomous systems,<sup>17</sup> but its application has been extended to time-dependent systems for which the coherent set is not required to be stationary.<sup>25,31</sup> The probabilistic approach has previously been presented in full detail,<sup>29</sup> and the following description outlines the approach as well as presenting several application examples.

As the Cauchy-Green deformation tensor is to the geometric approach, the transfer operator is the foundation for the probabilistic methods. The transfer operator maps a density distribution forward from the initial to final time. In order to numerically calculate this operator, the initial domain is partitioned into the set of connected boxes  $\{B_1, \dots, B_n\}$ , and the final domain into the set  $\{C_1, \dots, C_m\}$ , where  $n$  and  $m$  are the total number of boxes in the initial and final domain, respectively. Using a modification of Ulam’s method<sup>77</sup> and this partition, a finite-dimensional representation of the transfer operator is

$$\left(\mathbf{P}_{t_0}^t\right)_{i,j} = \frac{\mu(B_i \cap F_t^{t_0}(C_j))}{\mu(B_i)},$$

where  $\mu$  is the volume (Lebesgue) measure, and the ratio is the fraction of the area of  $B_i$  covered by the preimage of  $C_j$ . This is numerically achieved by seeding the boxes  $B_i$  with a uniform grid of initial conditions and mapping them to their final time positions. The transfer operator is then

$$\left(\mathbf{P}_{t_0}^t\right)_{i,j} \approx \#\{r : \mathbf{x}_{i,r} \in B_i, F_{t_0}^t(\mathbf{x}_{i,r}) \in C_j\} / N,$$

where the numerator is the number of initial positions,  $\mathbf{x}_{i,r}$ ,  $r = 1, \dots, N$ , in  $B_i$  that are mapped into  $C_j$ . With the converged calculation of this row stochastic matrix, it is possible to create the finite-time coherent sets and estimate the

finite-time entropy (FTE), which is a probabilistic analogy of the FTLE field.

#### A. Finite-time coherent sets

The primary use of the transfer operator for dynamical systems is to find sets of coherent pairs,  $A_i$  and  $\tilde{A}_i$ , where trajectories initially in  $A_i$  are mapped to  $\tilde{A}_i$  with high probability. In addition to this condition, the coherent pairs must be robust to imposed diffusion because the boundary length remains small, and conserved quantities (e.g., mass) must be the same in both sets.<sup>25</sup> This prevents the arbitrary pair  $A_i$  and  $\tilde{A}_i = F_{t_0}^t(A_i)$  from satisfying the conditions with certainty, leaving only structures with boundaries that remain small under advection. Finally, the partition must be a fraction of the domain to prevent the coherent set being empty or consisting of the entire domain. These criteria produce sets that are geometrically regular, meaning that there may be translation and deformation of the set boundary, but not a significant difference of the perimeter between sets  $A_i$  and  $\tilde{A}_i$ .<sup>25</sup> The elliptic structures found in Sec. II form an approximate boundary on an example set  $A_i$  that would satisfy these conditions, where the advected structure would be the approximate boundary of  $\tilde{A}_i$ . For elliptic barriers, a material line is identified that guarantees retention of all trajectories initially inside the structure. Coherent pairs simply identify sets of boxes, however, and yield no bounding material line, hence the potential for trajectories leaving or entering the set.

The coherent pairs form partitions of the initial and final domains. While some systems may have more than one coherent pair, we first focus on the case where there is a single coherent pair dividing the domain in two parts. In this case,  $X$  and  $Y$  represent sets of boxes in the initial and final states of the coherent pair. To identify the partition and optimization problem such that the optimal coherent pair will maximize the following measure

$$\rho(X, Y) = \frac{\mu(X \cap F_t^{t_0} Y)}{\mu(X)} + \frac{\mu(X^C \cap F_t^{t_0} Y^C)}{\mu(X^C)},$$

where  $\rho$  is a measure of the quality of the partition, the raised  $C$  represents the complement operator, and  $\mu$  is generalized to any conserved quantity (e.g., volume for incompressible systems or mass for compressible systems). The first ratio represents the fraction of the coherent pair  $X$  that is covered by the preimage of  $Y$ , and the second is the analogous case for the complement of the coherent pair.<sup>29</sup> The optimal solution would be a partition where all trajectories in  $X$  map to  $Y$ , and all trajectories in  $X^C$  map to  $Y^C$ ; that is to say, there is no transfer from or to the coherent pairs. Retention of all trajectories within the coherent pair will make the first ratio 1, and if no trajectories move from outside the coherent pair into it, then the second ratio is 1 making the maximum value of  $\rho = 2$ . In a general system, particularly when diffusion occurs, trajectories enter and leave the set, so the coherent set is identified through the solution of an optimization problem that minimizes this transfer. While the initial definition requires the coherent set to include approximately half the

domain, it is necessary to relax this condition if multiple coherent pairs are expected in a system.

To solve the optimization problem, it is necessary to determine the proper partitioning of the initial and final domain. A box  $B_i$  ( $C_j$ ) is therefore categorized as a member of  $X$  ( $Y$ ) if  $i \in I$  ( $j \in J$ ), where  $I$  and  $J$  are the sets of indices for the boxes forming the coherent pairs. Next, the membership vectors  $\mathbf{x} \in \mathbb{R}^n$  and  $\mathbf{y} \in \mathbb{R}^m$  and the corresponding thresholds  $b$  and  $c$  are defined such that for all  $i$  where  $x_i > b$  the corresponding boxes  $B_i$  make up the coherent pair  $X$ , and similarly the set of boxes  $C_j$  where  $y_j > c$  make up the coherent pair  $Y$ . Alternatively, it is possible to create sets by setting a maximum threshold and all boxes with a membership value below the threshold make up the coherent pair. It has been shown that these membership vectors are left and right singular-vectors or the modified transfer operator, presented next.<sup>31</sup>

The transfer operator as calculated even for incompressible systems may not necessarily conserve trajectory densities from the initial to final time across the domain. To maintain the density conservation, a normalization has to be applied. The probability  $p_i$  is the ratio of the volume measure of  $B_i$  to the volume measure of the domain. In the case where all boxes are the same size, then  $p_i = 1/n$ . The mapping forward of this probability measure is  $\mathbf{q} = \mathbf{p}\mathbf{P}'_{t_0}$ , where  $q_j$  is the ratio of volume mapped into  $C_j$  relative to the total volume. The two vectors  $\mathbf{p}$  and  $\mathbf{q}$  are diagonalized to form the square matrices  $\mathbf{\Pi}_p$  and  $\mathbf{\Pi}_q$ , respectively; the modified transfer operator is then  $\mathbf{\Pi}_p^{1/2}\mathbf{P}'_{t_0}\mathbf{\Pi}_q^{-1/2}$ .<sup>31</sup> Because the transfer operator is stochastic under this change of measure transformation, the first singular value is 1, so the second singular-vectors  $\hat{\mathbf{x}}$  and  $\hat{\mathbf{y}}$  form the coherent pair membership vectors  $\mathbf{x} = \hat{\mathbf{x}}\mathbf{\Pi}_p^{-1/2}$  and  $\mathbf{y} = \hat{\mathbf{y}}\mathbf{\Pi}_q^{-1/2}$ , where the rescaling has been undone.

Using the same parameters for the double-gyre system as in Sec. II, we calculate the transfer operator. The domain is partitioned into  $200 \times 400$  equally sized square boxes, which is a tenth of the resolution of the FTLE calculations, and to achieve converged results for the transfer operator it is necessary to advect approximately four times as many trajectories compared to the geometric method. Identifying the coherent pairs simply requires the minor calculation of  $\mathbf{q}$  and

the singular value decomposition of the first four singular values of the modified transfer operator. There are two coherent pairs detected in the second singular-vectors. The optimum thresholds to maximize  $\rho(A_i, \tilde{A}_i)$  result in a positive singular value structure with 99.7% retention and a negative value structure with 99.6% retention. While the optimal coherent pairs result from the second singular-vectors, it is possible to use lower value singular-vectors for the membership vectors. The fourth singular-vector also produces two structures with 99.3% retention. Figure 3 presents the four coherent structures in their initial states (a, c) and their final states (b, d). While internal code was used for this calculation, GAIO,<sup>2,18</sup> an openly available software, can be used to partition a domain and calculate the transfer operator.

For finite-time coherent set detection, it is possible to use a modified but equivalent definition that relies on solving the isoperimetry problem using the newly developed dynamic Laplacian operator and the dynamic Federer-Fleming theorem.<sup>26</sup> This approach defines the boundaries of finite-time coherent sets as curves that maximize the volume to boundary size ratio throughout advection, which eliminates the possibility of boundary filamentation and ensures minimal exchange of passive tracers across the partition even in the presence of numerical diffusion. To identify these curves rapidly and efficiently, radial basis functions have been implemented.<sup>27</sup> It should be noted that radial basis functions can also be used for finding finite-time coherent sets in the advection only scenario, which is important when considering higher dimensional domains. Extending the probabilistic methods from 2D to 3D does not require any new mathematical machinery, but the increased computational demands for the additional number of trajectories per box and the number of boxes in the domain can be significant. Implementation of the radial basis functions results in calculating far fewer trajectories and can alleviate some of the computational demand, making the extension to higher dimensions feasible.

### B. Finite-time entropy

Where the FTLE field directly measures the maximum rate of local stretching, the transfer operator can be used to indirectly calculate stretching by determining the level of

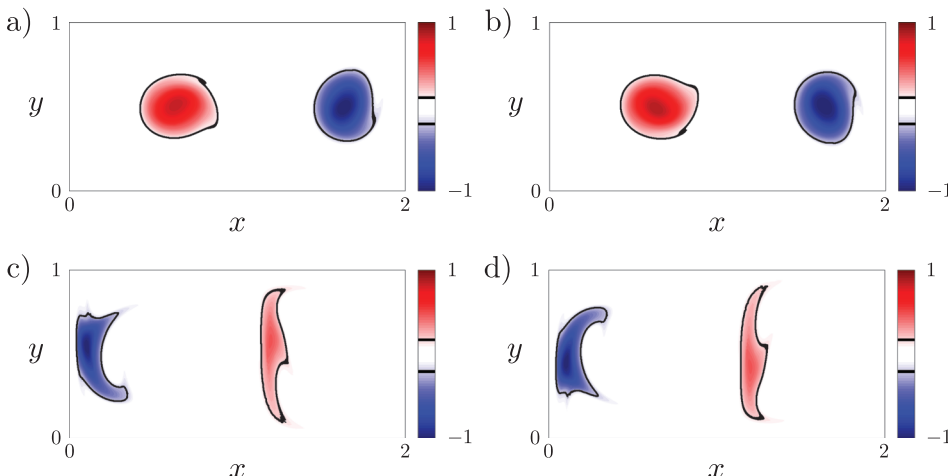


FIG. 3. The second (a) right and (b) left singular-vectors, representing the initial and final membership vectors of two coherent pairs for the double-gyre flow, using the same parameters as for the earlier FTLE studies. The fourth (c) right and (d) left singular-vectors representing the initial and final positions membership vectors of two further coherent pairs. Thresholds in black represent the optimal bounds of the coherent pairs.

uncertainty of the final position of a trajectory starting from a random position within one of the boxes,<sup>28</sup>  $B_i$ . If the transfer matrix has been calculated for a domain partitioned by equal size boxes, the FTE field is

$$FTE_{t_0}^t(B_i) = -\frac{1}{t-t_0} \sum_j \left( \mathbf{P}_{i,j}^t \right) \log \left( \mathbf{P}_{i,j}^t \right).$$

In regions of the field where there is little deformation, trajectories initiated within these boxes will move together and map to a small number of nearby boxes resulting in a low degree of uncertainty in the final position, and thus a small FTE value. In the unique case that all trajectories map to a single box, the FTE value is zero. If there is a large amount of deformation, the particles initially within a given box will be stretched apart and advect to a large number of boxes throughout the domain, resulting in a large final position uncertainty and FTE value. While not generally achievable, there is a theoretical upper bound for the FTE values corresponding to the extreme case of each trajectory being mapped to a different box, resulting in an FTE value of  $\log(n)/(t-t_0)$ .

With the double-gyre transfer operator calculated in Sec. III A, the calculation of the FTE field is effectively a “free” computation, because the transfer operator calculation so dominates the computational demands. Comparing the FTE field presented in Figure 4 to the FTLE field in Figure 1, the fields are qualitatively similar. Despite advecting more trajectories, the FTE field is coarser due to the averaging over the boxes and the values are lower than the corresponding FTLE values, but the FTE field does accurately identify the regions of the domain where large deformations occur. The FTE approach was not intended to be an alternative to the FTLE calculation but meant to provide an alternative method to determine stretching information for systems where the diffusion of passive naturally arises and needs to be accounted for.

### C. Applications of the probabilistic approach

Transfer operator techniques have been applied to a number of real world data sets in order to demonstrate their usefulness for geophysical flows. A study of a two-dimensional atmospheric system was performed to demonstrate that the polar vortex could be identified as a coherent pair.<sup>31</sup> This result showed good agreement with potential

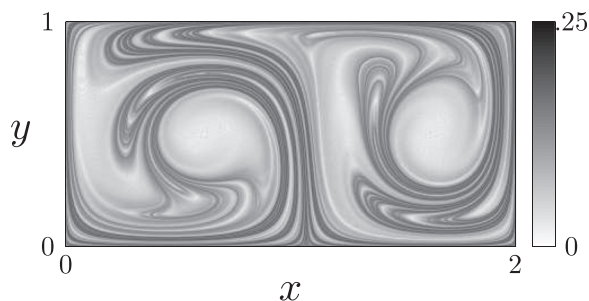


FIG. 4. The finite-time entropy values calculated for the double-gyre system over the interval [2.5, 42.5].

vorticity based measurements, and the corresponding FTE field was also calculated revealing a complex swirl of high FTE values at the edge of the coherent pair, with some moderate FTE filaments extending through the polar vortex.<sup>28</sup> The coherent pair analysis was extended to a three-dimensional partition of the space, indicating that while there is some elevation change of the coherent set there is limited deformation radially outward.<sup>31</sup> In the case of this atmospheric example, the compressibility of the fluid forces consideration of what conserved quantity should be used instead of volume; in this case, the quantity was the air mass.

Another study identified the three-dimensional structure of an Agulhas Ring.<sup>32</sup> Using the coherent pair approach, a six month interval for which the Agulhas ring identified as a coherent pair travels approximately 750 km was investigated. In this study, the fourth singular-vectors were the ones that identified the ring boundary, not the second or third, demonstrating that even higher singular-vectors may be of interest and should be investigated. The coherent pair identified had a coherence ratio of 76.3%, demonstrating that this fraction of water remains in a compact space as the ring travels slowly towards South America from the southern tip of Africa.

Finally, a pair of studies investigated global ocean phenomena using the transfer operator. In the first investigation, an average annual transfer operator for ocean surface advection was calculated, and this was used to advect a distribution of debris to show the development of concentrations of debris that correspond to garbage patches observed on the ocean surface.<sup>78</sup> While no coherent pairs were identified, this study demonstrated how quickly the transfer operator can be applied for periodic systems and how the transfer operator can be used to identify asymptotically stable regions of the domain. A related study used the transfer operator to divide the ocean into distinct regions, in an attempt to quantify the connectivity of different parts of the ocean.<sup>33</sup> To better represent the ocean surface dynamics, this study allowed for the loss of trajectories by considering an open domain, which accounted for the beaching of trajectories and advection to the poles. By considering by the Perron-Frobenius operator and its dual, the Koopman operator, the study was extended to both forwards and backwards time, which allowed identification of regions of upwelling and downwelling in the ocean. This modified approach also measured the transfer of material between different regions of the ocean.

### IV. SPARSE TRAJECTORY SET METHODS

For the “dense” geometric and probabilistic approaches, it is necessary to have a well resolved (in time and space) velocity field data set for the time interval of interest, in order to calculate the large number of trajectories necessary for analysis. In real world applications, accurate velocity field information is not always available, but analysis of ocean drifter data or particle tracking experiments, for example, does provide sparse trajectory information. We present two approaches that attempt to glean coherent structure information directly from sparse trajectory information. In both cases, the coherent



structures identified are sets of trajectories that remain in close proximity to each other throughout the time interval.

### A. Cluster based analysis

One sparse method attempts to cluster trajectories based on a distance metric between them in a higher dimensional space.<sup>30</sup> If a set of trajectories is defined on a two-dimensional surface and their position is known at  $T$  distinct times, the entire trajectory can be represented by a single point in  $\mathbb{R}^{2T}$  (i.e., the trajectory  $(x(t), y(t))$  corresponds to  $\mathbf{X}_i = (x(t_1), y(t_1), x(t_2), y(t_2), \dots, x(t_T), y(t_T))$ ). If two trajectories have a small Euclidean distance between the representative points in this higher dimensional space, the distance between the two trajectories must remain small throughout the time interval. The reduction of trajectory information to a set of points also allows for the use of well developed tools from the clustering community.

The particular type of clustering scheme selected is the fuzzy C-means (FCM) algorithm,<sup>9,10</sup> which divides the system into  $K$  clusters based on the distance between a given trajectory point and the cluster’s center. The approach assigns a membership probability

$$u_{k,i} = \left[ \sum_{j=1}^K \left( \frac{\|\mathbf{X}_i - \mathbf{C}_k\|}{\|\mathbf{X}_i - \mathbf{C}_j\|} \right)^{2/(m-1)} \right]^{-1}$$

representing the probability that the  $i$ th trajectory  $\mathbf{X}_i$  is a member of the  $k$ th cluster using the “fuzziness parameter”  $m$  that influences the sharpness of the partition. The membership probability is based on the distance to cluster centers, where the cluster center  $\mathbf{C}_k$  is the weighted average of the trajectory positions

$$\mathbf{C}_k = \frac{\sum_{i=1}^n (u_{k,i})^m \mathbf{X}_i}{\sum_{i=1}^n (u_{k,i})^m}.$$

The algorithm iteratively updates the membership probability,  $\mathbf{u}$ , and the cluster centers,  $\mathbf{C}_k$ , eventually converging on a partition of the trajectories. From a usage perspective, this

approach is ready-made in the sense that FCM algorithms are already programmed and optimized (e.g., the function `fcm` in `Matlab`), and the user only needs to set the number of expected clusters and a “fuzziness” parameter,  $m > 1$ .

For the double-gyre application, using the same parameters and time interval as the previous example applications, we apply this method to two different sets of trajectories. Because the FCM method is so efficient, it can handle large numbers of trajectories well, so we first test a grid of trajectories with an even spacing of  $\delta x = 0.01$ ; these 20 000 trajectories are a small fraction of the tens of millions of trajectories used in the previous methods. Forty-one time slices of the trajectories are taken, creating a trajectory representation in  $\mathbb{R}$ .<sup>80</sup> Expecting four clusters and setting  $m = 1.5$  produces the four membership probabilities presented in Figure 5. Selecting a threshold for membership of 70% produces four closed material lines with similarities to the four coherent sets produced by the transfer operator approach presented in Figure 3, and higher thresholding may yield better results. We note that the FCM approach does generalize to systems with a continuum of initial conditions with continuously sampled trajectories.

The quick FCM analysis produces a reasonable first estimation of the coherent sets for the case where a dense set of trajectories is known, but sparse trajectories are another possible application. We uniformly distribute 300 trajectories using the Poisson-disk sampling of the domain and apply the FCM algorithm to this sparse trajectory data; the results are presented in Figure 6. These results are similar to the results from the dense set of trajectories, but there are some points that fall within the 70% level sets that are not assigned by the scatter result. This is likely due to slight differences in the cluster center location due to the differing levels of information. Because the sparse data approach does not calculate the cluster center as accurately, inclusion and exclusion of points near the cut off level set is prone to discrepancies. As is expected, the more data that is known, the more accurate the coherent set identification becomes.

In addition to a number of analytic examples, the FCM method has been applied to the Global Ocean Drifter Program.<sup>30</sup> One major strength of the approach is that the trajectory data does not need to span the entire time interval.

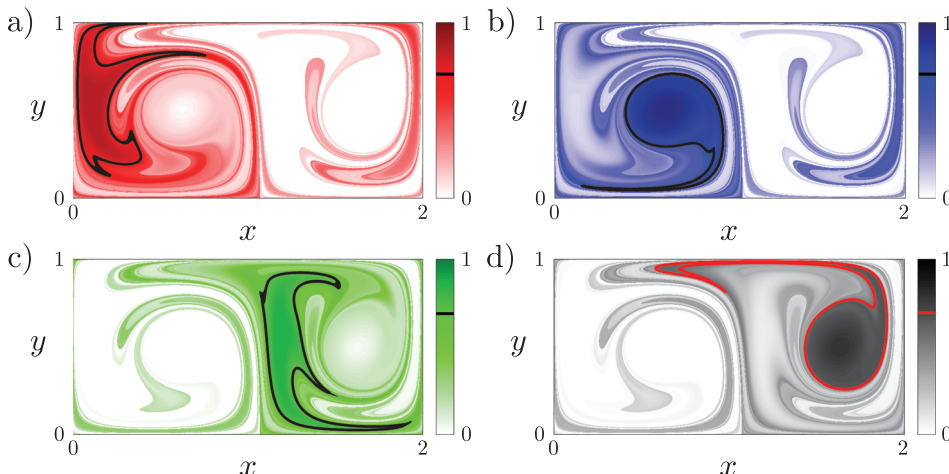


FIG. 5. Membership probabilities for each of the four clusters resulting from an analysis of the double-gyre system. (a)  $u_1$ , (b)  $u_2$ , (c)  $u_3$ , and (d)  $u_4$  all represent the probability of membership in that cluster as a function of position, with the off color line corresponding to the level set of 70% membership.

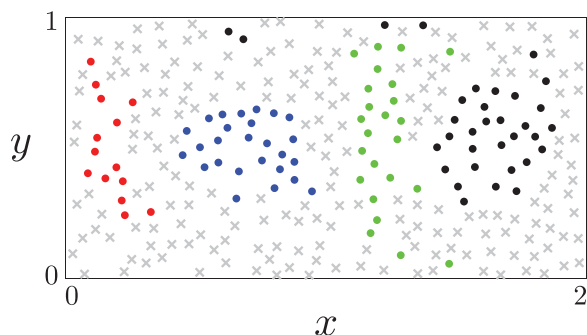


FIG. 6. Initial position of 300 uniformly distributed points advected by the double-gyre system. Membership to the red, blue, green, and black clusters is based on the 70% threshold.

This is vital when analyzing drifter data as drifters have a tendency to run out of battery power, beach, or otherwise break, and relying on only those floats that survived the entire duration of a time interval severely limits analysis. To account for the missing data, a natural modification of the FCM algorithm is necessary and has been derived.<sup>30</sup> Using the global drifter data set, the FCM method was able to identify global ocean partitioning results similar to those found using transfer operator methods.<sup>33</sup>

While the FCM method is quite robust for sparse and intermittent data, there are some limitations. The approach limits the consideration of relative movement between trajectories to just the prescribed distance metric. In an example with periodic islands in a chaotic sea, it is possible for nearby islands to be clustered together because they do stay nearby throughout time despite the fact that there is chaotic mixing occurring in the space between them. Increasing the number of clusters, however, may solve this problem. Another shortcoming is that there is not always an obvious way to bound a cluster with a material line or an approximate material line. As seen in Figure 6, there are a couple of isolated points that are members of the black cluster and it is not obvious how to connect these points with the rest of the domain. One possible remedy is to determine the boundary based on a later time instance for which the points may be more obviously connected. Finally, it is left to the user to select the number of clusters, which is not always trivial. Fortunately, trial and error will reveal robust structures as the cluster number is varied. In spite of these minor drawbacks, the FCM approach is easily applied and gives excellent first order results revealing where further coherent structure study is justified. A final strength of this approach is that extension into higher dimensions simply requires the trajectory representation to account for the extra dimensional information.

## B. Braid based analysis

Another sparse technique that can analyze scattered data utilizes tools from the topological field of braid theory. By reducing the physical space to a topological space, it is possible to deform trajectories and material lines in such a way that lends itself to the rapid analysis of the system dynamics. Braid theory was initially applied to

fluid systems to rapidly study the amount of stirring<sup>12</sup> but was extended to find coherent structures defined as approximate material lines that do not significantly change length during the time interval.<sup>5</sup> Because deformations of the space have been used, this method can only provide approximate positions for material lines that remain geometrically regular. In addition to the coherent structure detection, recent developments allow for the calculation of a finite-time mixing measure referred to as the finite-time braiding exponent (FTBE).<sup>16</sup>

The first step in this analysis is to represent the trajectory information as a braid. Because of the monotonic nature of time, it is possible to extend trajectories from the two-dimensional physical space into three dimensions where the physical space forms a base and time is the vertical coordinate. In this space, the trajectories form strands that weave around each other but do not turn back on themselves or intersect each other. When projected onto one of the physical axes, the trajectories appear to cross in front of or behind other trajectories as the set of strands rearrange themselves over time. By deforming the trajectories, it is possible to isolate individual crossings and define the crossing by the strands involved and the orientation of the cross (clockwise or counterclockwise when viewed from above). These individual crossing events form the building block of the braid, the generator. A trajectory set is converted into a sequence of generators that represents the trajectory information as a list of signed integers.

The other tool of braid theory used is the loop. Where a generator sequence is used to represent the deformed trajectory set, loops are used to represent deformed closed material lines. In a similar manner, there is a coordinate representation of the loops (the Dynnikov coordinates<sup>19</sup>) that lends itself to the simple representation and bijection between coordinates and loops. Because the only rules for deforming a loop are that it cannot be deformed through itself or any of the trajectories, topological loops only change shape when a trajectory forces it to deform during a generator crossing. As such, there are rules for how generators modify the loop coordinates, and this enables the rapid advection of approximate closed material lines as loops under the action of the braid.<sup>75</sup>

A simple measure that can be made with the braid representation and a particular loop, the fundamental-loop, is the FTBE.<sup>16</sup> The FTBE is a measure of the braid complexity during the finite-time interval and is related to the topological entropy of braids. For a given braid  $\mathbf{b}$  representing the dynamics over the time interval  $[t_0, t]$ , the FTBE is

$$FTBE_{t_0}^t(\mathbf{b}) = \frac{1}{t - t_0} \log \frac{|b l_e|}{|l_e|},$$

where  $l_e$  represents the fundamental-loop that will capture all possible mixing of the system, and the length functions in the ratio represent the length of the loop. This measure is a time average rate of exponential growth of the length of the fundamental loop under the action of the braid. By taking into account the dynamics of the entire system as opposed to the local repulsion of individual trajectories, the FTBE

measurement differs from the FTLE or FTE calculation in that it is a global measure of complexity.

The braid based approach is also capable of finding coherent structures and defines a structure as the loop surrounding a set of trajectories that does not grow or shrink significantly under the action of the braid. In order to find these structures, it is necessary to analyze how loops connecting pairs of trajectories grow under the action of the braid. Based on loop entanglement, it is possible to group trajectories together such that there will be a loop bounding them that will not grow during the time interval. This is a direct result of trajectories within the cluster not mixing with trajectories outside of the cluster. The final step of the approach is to construct the corresponding Dynnikov coordinate that represents the bounding loop and confirm that it does not grow under the action of the braid.

Using the same three hundred trajectories used in the FCM approach, a braid is generated consisting of over 140 000 generators. Applying this braid to the fundamental-loop produces an FTBE value of 0.1265 which is slightly smaller than the average FTLE and FTE values, though no claims are made on how exactly these values are related. The application of the coherent structure algorithm produces four large loops that do not grow significantly over time, presented in Figure 7. Compared to the cluster approach, there are more trajectories included in the structures, but their shapes are qualitatively similar to the results presented in Figure 5. While the analysis technique is more complicated than the FCM approach, the benefit of knowing the approximate shape of the closed material lines is significant when trying to understand, locate, and track potentially complex coherent structures.

In addition to the double-gyre flow, the tools of braid theory have been applied to a pair of physical systems. Analyzing a set of trajectory data, estimates of ocean mixing were made using topological entropy, which is similar to the FTBE.<sup>75</sup> The calculation highlighted the challenges of working with ocean drifters from a topological perspective as only the mixing of drifters was captured by the braid. Intermittent trajectories could not be used, and once the trajectories separated no new information was added to the braid. The other application was to a closed viscous system periodically stirred by a rod.<sup>5</sup> A velocity field for this system was analytically known and trajectories were calculated for

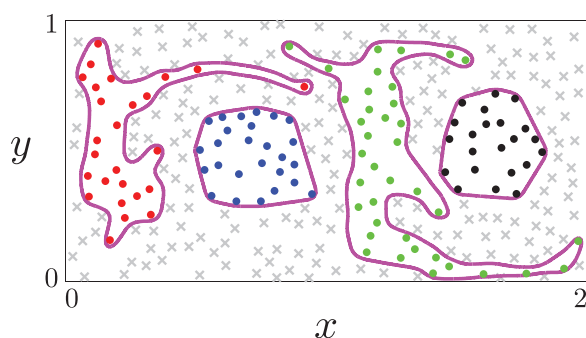


FIG. 7. Initial position of 300 uniformly distributed points advected by the double-gyre system. The red, blue, green, and black points are bound by loops (magenta) that do not grow under the action of the braid.

application to the coherent structure approach. While the method successfully identified a number of structures when compared to the system's Poincare map, the application demonstrated the limitations of working with sparse data; namely, the need to have multiple trajectories within a given structure to enable detection.

There are a number of strengths and weaknesses of the braid based approach. It is able to analyze sparse trajectory data sets but is somewhat limited by the requirements that trajectory information is known for the entire interval and significant mixing must occur. While the method does produce closed material lines that do not grow during advection, these are only approximate representations. Compared to the FCM approach, the braiding method does account for how trajectories interact with each other, but because of the deformations the method neglects the spatial proximity of trajectories when grouping them together. While the analysis is still relatively fast compared to the probabilistic and geometric approaches, the complexity of the braid based coherent structure approach is significant. There is a `Matlab` toolbox available for dealing with braids and loops.<sup>1,76</sup> We note that due to its topological foundations, there is no current scope for extension of the braid theory approach to 3D flows.

## V. CONCLUSION

We have presented an overview of four of the leading objective approaches for identifying the coherent structures that underly the organization of material transport in unsteady fluid flows. Each of the methods has its strengths and weaknesses, and this needs to be appreciated when selecting the approach to be used for a particular application. The geometric and probabilistic techniques have been rigorously developed for situations where full velocity field information and significant computational resources are available. If the goal is to identify key material lines, the geometric approach enables unambiguous determination of those that guide deformations in specific ways; and although simple, basic FTLE analysis remains a practical tool for initial investigations. For systems where the goal is to identify regions that remain unmixed with the rest of the system, the geometric approach is again useful in finding regions where the boundary is uniformly stretched; for more general regions that remain predominantly unmixed during advection the probabilistic approach is practical. In the case where rapid analysis is required and/or there is only sparse trajectory information available, the clustering and braid based approaches are good options. For an efficient first search for regions that remain unmixed, the clustering approach is a useful diagnostic, whereas the braid based approach can provide a bit more insight into the trajectory mixing and approximate shape of boundaries at the cost of added complexity and stricter data demands.

While the aforementioned approaches provide a cross section of Lagrangian based techniques, there are a number of other approaches that are also receiving much attention. Meso-elliptic and meso-hyperbolic measures<sup>54</sup> provide analysis based on time-averaged quantities along trajectories, classifying regions where on average the behavior is more

eddy like, more strain like, or a mixture of the two. While not an instantaneous method, this approach is frame dependent, and, as described in the introduction of this review, this is a concern because the results and accompanying interpretation of coherent features can be altered by changing the frame of reference, and generically there is no predetermined frame of reference from which an unsteady fluid flow should be observed. Another approach utilizes the Koopman operator and observable functions along trajectories to partition the state space into different ergodic regions.<sup>15,52,53</sup> Using trajectory based measures, this approach attempts to cluster trajectories together based on trajectory averaged properties. A recent review covers this approach in more detail.<sup>51</sup> The Complexity Method (CM), which identifies clusters of trajectories with similar Lagrangian properties as a means to determine coherent water masses,<sup>64</sup> has been successfully applied to study near-surface transport and water mass exchange processes around the Philippine Archipelago,<sup>63</sup> and is easy to generalize to 3D flows. Finally, a pair of approaches attempts to quantify the amount of folding that occurs for finite-size passive tracer patches. This measure complements the stretching measures (e.g., FTLE and FTE) and is based on the nonlinearity of the deformed patch<sup>45</sup> or the amount of curvature created in finite length line segments.<sup>48</sup>

There are a number of exciting avenues of research and areas of application beyond those already mentioned. One that is practically important is the detection of coherent structures that organize the transport of inertial particles in unsteady flows, since floating and immersed objects, such as debris and sediment, do not simply act as passive tracers. Examples of work on this topic include a study of coherent structures from the inertial equations of motion as applied to a hurricane<sup>66</sup> and jellyfish predation,<sup>62</sup> and investigations of inertial particle transport near elliptic structures motivated by the behavior of drifters near large scale eddies.<sup>8</sup> Another active research area is the application of coherent structure detection to diffusion-advection-reaction systems. By calculating the FTLE field for an ensemble of runs and comparing the results to the asymptotic states of chemical reactions, it has been shown that there is a strong correlation between regions of high FTLE and dynamically different reaction fates.<sup>71,72</sup> Studies of chemical reaction systems have also been focusing on the identification of burning invariant manifolds, which are one-way barriers in the advection reaction systems that inhibit a chemical reaction.<sup>49,55</sup>

Moving forward, there are several major avenues to be explored. One important question is what, if anything, can identification of these structures contribute in a predictive sense. All of the analysis tools presented consider a given time interval of dynamics and makes statements about the key contributions of the structures for that interval. While in some cases particular structures may continue to have influence after the interval of study,<sup>58</sup> more results on the longevity of structures outside the given time interval would be significant. And while there are now more results for three-dimensional analytic flows,<sup>11</sup> there is a need to exploit GPUs and other numerical techniques<sup>6</sup> to aid in efficient calculation and visualization of this highly parallelizable problem.

Finally, it needs to be recognized that despite being around for almost a decade, to date, studies have been primarily confined to the academic literature and coherent-structure-based tools have yet to significantly impact any real-world decisions making systems. There needs to be a thorough and honest assessment of what aspects of the different approaches are sufficiently robust, insightful, and practically useful to really make a difference in important scenarios such as search-and-rescue operations and oil spill response strategies. When the next major disaster at sea occurs, will this suite of Lagrangian analysis tools for coherent structure detection be ready to play any real-time role?

## ACKNOWLEDGMENTS

The authors would like to thank Gary Froyland for helpful discussions on the application of the transfer operator and fuzzy clustering approaches, and George Haller for clarification on the implementation of the objective vortex algorithm. The authors thank Gary Froyland, George Haller, Jean-Luc Thiffeault, and the anonymous referee for helpful comments that improved the review. This work was supported by ONR grant N000141210665.

<sup>1</sup>See <https://github.com/jeanluct/braidlab> for Braidlab; accessed 01 May 2015.

<sup>2</sup>See [http://web.maths.unsw.edu.au/~froyland/banff\\_tutorial.html](http://web.maths.unsw.edu.au/~froyland/banff_tutorial.html) for GAIO; accessed 01 May 2015.

<sup>3</sup>See <https://github.com/jeixav/LCS-Tool> for LCS tool; accessed 01 May 2015.

<sup>4</sup>M. R. Allshouse and T. Peacock, "Refining and classifying FTLE ridges," Chaos (submitted), e-print [arXiv:1506.07027](https://arxiv.org/abs/1506.07027) [math.DS].

<sup>5</sup>M. R. Allshouse and J.-L. Thiffeault, "Detecting coherent structures using braids," *Physica D* **241**, 95–105 (2012).

<sup>6</sup>S. Ameli, Y. Desai, and S. C. Shadden, "Development of an efficient and flexible pipeline for Lagrangian coherent structure computation," in *Topological Methods in Data Analysis and Visualization III, Mathematics and Visualization*, edited by P.-T. Bremer *et al.* (Springer, 2014).

<sup>7</sup>A. Arzani and S. C. Shadden, "Characterization of the transport topology in patient-specific abdominal aortic aneurysm models," *Phys. Fluids* **24**(8), 081901 (2012).

<sup>8</sup>F. J. Beron-Vera, M. J. Olascoaga, G. Haller, M. Farazmand, J. Trinanes, and Y. Wan, "Dissipative inertial transport patterns near coherent Lagrangian eddies in the ocean," Chaos (submitted), e-print [arXiv:1408.6512](https://arxiv.org/abs/1408.6512) [physics.ao-ph].

<sup>9</sup>J. C. Bezdek, R. Ehrlich, and W. Full, "Fcm: The fuzzy c-means cluster algorithm," *Comput. Geosci.* **10**(2), 191–203 (1984).

<sup>10</sup>J. C. Bezdek, R. J. Hathaway, M. J. Sabin, and W. T. Tucker, "Convergence theory for fuzzy c-means: Counterexamples and repairs," *IEEE Trans. Syst. Man Cybern.* **17**(5), 873–877 (1987).

<sup>11</sup>D. Blazeviski and G. Haller, "Hyperbolic and elliptic transport barriers in three-dimensional unsteady flows," *Physica D* **273–274**, 46–64 (2014).

<sup>12</sup>P. L. Boyland, H. Aref, and M. A. Stremler, "Topological fluid mechanics of stirring," *J. Fluid Mech.* **403**, 277–304 (2000).

<sup>13</sup>A. E. Bozorgmagham and S. D. Ross, "Atmospheric Lagrangian coherent structures considering unresolved turbulence and forecast uncertainty," *Commun. Nonlinear Sci. Numer. Simul.* **22**, 964–979 (2015).

<sup>14</sup>A. E. Bozorgmagham, S. D. Ross, and D. G. Schmale III, "Real-time prediction of atmospheric Lagrangian coherent structures based on forecast data: an application and error analysis," *Physica D* **258**, 47–60 (2013).

<sup>15</sup>M. Budisic and I. Mezic, "Geometry of the ergodic quotient reveals coherent structures in flows," *Physica D* **241**, 1255–1269 (2012).

<sup>16</sup>M. Budisic and J.-L. Thiffeault, "Finite-time braiding exponents," Chaos (submitted), e-print [arXiv:1502.02162](https://arxiv.org/abs/1502.02162) [nlin.CD].

<sup>17</sup>M. Dellnitz and O. Junge, "Almost invariant sets in Chua's circuit," *Int. J. Bifurcation Chaos* **7**(11), 2475–2485 (1997).

<sup>18</sup>M. Dellnitz, G. Froyland, and O. Junge, "The algorithms behind GAIO—Set oriented numerical methods for dynamical systems," in *Ergodic*

- Theory, Analysis and Efficient Simulation of Dynamical Systems*, edited by B. Fiedler (Springer, 2001), pp. 145–174.
- <sup>19</sup>I. A. Dynnikov, “On a Yang-Baxter map and the Dehornoy ordering,” *Russian mathematical Surveys*, 2002.
  - <sup>20</sup>M. J. Olascoaga *et al.*, “Drifter motion in the Gulf of Mexico constrained by altimetric Lagrangian coherent structures,” *Geophys. Res. Lett.* **40**, 6171–6175, doi:10.1002/2013GL058624 (2013).
  - <sup>21</sup>S. Hendabadi *et al.*, “Topology of blood transport in the human left ventricle by novel processing of doppler echocardiography,” *Ann. Biomed. Eng.* **41**(12), 2603–2616 (2013).
  - <sup>22</sup>M. Farazmand and G. Haller, “Computing Lagrangian coherent structures from variational LCS theory,” *Chaos* **22**, 013128 (2012).
  - <sup>23</sup>M. Farazmand and G. Haller, “Attracting and repelling Lagrangian coherent structures from a single computation,” *Chaos* **23**, 023101 (2013).
  - <sup>24</sup>M. Farazmand and G. Haller, “How coherent are the vortices of two-dimensional turbulence?,” e-print [arXiv:1402.4835](https://arxiv.org/abs/1402.4835) [math.DS].
  - <sup>25</sup>G. Froyland, “An analytic framework for identifying finite-time coherent sets in time-dependent dynamical systems,” *Physica D* **250**, 1–19 (2013).
  - <sup>26</sup>G. Froyland, “Dynamic isoperimetry and the geometry of Lagrangian coherent structures,” e-print [arXiv:1411.7186](https://arxiv.org/abs/1411.7186) [math.DS].
  - <sup>27</sup>G. Froyland and O. Junge, “On fast computation of finite-time coherent sets using radial basis functions,” *Chaos* (submitted), e-print [arXiv:1505.05056](https://arxiv.org/abs/1505.05056) [math.NA].
  - <sup>28</sup>G. Froyland and K. Padberg-Gehle, “Finite-time entropy: A probabilistic approach for measuring nonlinear stretching,” *Physica D* **241**, 1612–1628 (2012).
  - <sup>29</sup>G. Froyland and K. Padberg-Gehle, “Almost-invariant and finite-time coherent sets: directionality, duration, and diffusion,” in *Ergodic Theory, Open Dynamics, and Coherent Structures*, edited by W. Bahsoun, C. Bose, and G. Froyland (Springer, 2014), Vol. 70.
  - <sup>30</sup>G. Froyland and K. Padberg-Gehle, “A rough-and-ready cluster-based approach for extracting finite-time coherent sets from sparse and incomplete trajectory data,” *Chaos* (submitted), e-print [arXiv:1505.04583](https://arxiv.org/abs/1505.04583) [math.DS].
  - <sup>31</sup>G. Froyland, N. Santitissadeekorn, and A. Monahan, “Transport in time-dependent dynamical systems: Finite-time coherent sets,” *Chaos* **20**, 043116 (2010).
  - <sup>32</sup>G. Froyland, C. Horenkamp, V. Rossi, N. Santitissadeekorn, and A. S. Gupta, “Three-dimensional characterization and tracking of an Agulhas ring,” *Ocean Modell.* **52–53**, 69–75 (2012).
  - <sup>33</sup>G. Froyland, R. M. Stuart, and E. van Sebille, “How well-connected is the surface of the global ocean?,” *Chaos* **24**, 033126 (2014).
  - <sup>34</sup>M. A. Green, C. W. Rowley, and A. J. Smits, “Using hyperbolic Lagrangian coherent structures to investigate vortices in bioinspired fluid flows,” *Chaos* **20**, 017510 (2010).
  - <sup>35</sup>A. Hadjighasem and G. Haller, “Geodesic transport barriers in Jupiter’s atmosphere: A video-based analysis,” *SIAM Rev.* (submitted), e-print [arXiv:1408.5594](https://arxiv.org/abs/1408.5594) [nlin.CD].
  - <sup>36</sup>G. Haller, “Distinguished material surfaces and coherent structures in 3D fluid flows,” *Physica D* **149**, 248–277 (2001).
  - <sup>37</sup>G. Haller, “Lagrangian coherent structures from approximate velocity data,” *Phys. Fluids* **14**(6), 1851–1861 (2002).
  - <sup>38</sup>G. Haller, “An objective definition of a vortex,” *J. Fluid Mech.* **525**, 1–26 (2005).
  - <sup>39</sup>G. Haller, “Lagrangian coherent structures,” *Annu. Rev. Fluid. Mech.* **47**, 137–162 (2015).
  - <sup>40</sup>G. Haller and F. J. Beron-Vera, “Geodesic theory of transport barriers in two-dimensional flows,” *Physica D* **241**, 1680–1702 (2012).
  - <sup>41</sup>G. Haller and F. J. Beron-Vera, “Coherent Lagrangian vortices: The black holes of turbulence,” *J. Fluid Mech.* **731**, R4 (2013).
  - <sup>42</sup>G. Haller and G. Yuan, “Lagrangian coherent structures and mixing in two-dimensional turbulence,” *Physica D* **147**, 352–370 (2000).
  - <sup>43</sup>D. Karrasch and G. Haller, “Do finite-size Lyapunov exponents detect coherent structures?,” *Chaos* **23**, 043126 (2013).
  - <sup>44</sup>D. Karrasch, F. Huhn, and G. Haller, “Automated detection of coherent Lagrangian vortices in two-dimensional unsteady flows,” *Proc. R. Soc. A* **471**, 20140639 (2015).
  - <sup>45</sup>D. H. Kelley and N. T. Ouellette, “Separating stretching from folding in fluid mixing,” *Nat. Phys.* **7**, 477–480 (2011).
  - <sup>46</sup>D. H. Kelley, M. R. Allshouse, and N. T. Ouellette, “Lagrangian coherent structures separate dynamically distinct regions in fluid flows,” *Phys. Rev. E* **88**, 013017 (2013).
  - <sup>47</sup>D. Lipinski and K. Mohseni, “A ridge tracking algorithm and error estimate for efficient computation of Lagrangian coherent structures,” *Chaos* **20**, 017504 (2010).
  - <sup>48</sup>T. Ma, N. Ouellette, and E. M. Bollt, “Stretching and folding in finite time,” *Chaos* (submitted).
  - <sup>49</sup>J. Mahoney, D. Bargeil, M. Kingsbury, K. Mitchell, and T. Solomon, “Invariant barriers to reactive front propagation in fluid flows,” *EPL* **98**, 44005 (2012).
  - <sup>50</sup>M. Mathur, G. Haller, T. Peacock, J. E. Ruppert-Felsot, and H. L. Swinney, “Uncovering the Lagrangian skeleton of turbulence,” *Phys. Rev. Lett.* **98**(14), 144502 (2007).
  - <sup>51</sup>I. Mezic, “Analysis of fluid flows via spectral properties of the Koopman operator,” *Annu. Rev. Fluid. Mech.* **45**, 357–378 (2013).
  - <sup>52</sup>I. Mezic and A. Banaszuk, “Comparison of systems with complex behavior,” *Physica D* **197**, 101–133 (2004).
  - <sup>53</sup>I. Mezic and S. Wiggins, “A method for visualization of invariant sets of dynamical systems based on the ergodic partition,” *Chaos* **9**, 213 (1999).
  - <sup>54</sup>I. Mezic, S. Loire, V. A. Fonoberov, and P. Hogan, “A new mixing diagnostic and Gulf oil spill movement,” *Science* **330**(6003), 486–489 (2010).
  - <sup>55</sup>K. A. Mitchell and J. R. Mahoney, “Invariant manifolds and the geometry of front propagation in fluid flows,” *Chaos* **22**, 037104 (2012).
  - <sup>56</sup>A. Okubo, “Horizontal dispersion of floatable particles in the vicinity of velocity singularities such as convergences,” *Deep-Sea Res.* **17**, 445–454 (1970).
  - <sup>57</sup>M. J. Olascoaga, “A favorable test of isolation on the west florida shelf with implication for red tides,” *Nonlinear Processes Geophys.* **17**, 685–696 (2010).
  - <sup>58</sup>M. J. Olascoaga and G. Haller, “Forecasting sudden changes in environmental contamination patterns,” *Proc. Natl. Acad. Sci.* **109**, 4738–4743 (2012).
  - <sup>59</sup>M. J. Olascoaga, F. J. Beron-Vera, L. E. Brand, and H. Kocak, “Tracing the early development of harmful algal blooms with the aid of Lagrangian coherent structures,” *J. Geophys. Res.* **113**, C12014, doi:10.1029/2007JC004533 (2008).
  - <sup>60</sup>K. Onu, F. Huhn, and G. Haller, “LCS tool: A computational platform for Lagrangian coherent structures,” *J. Comput. Sci.* (submitted), e-print [arXiv:1406.3527](https://arxiv.org/abs/1406.3527) [nlin.CD].
  - <sup>61</sup>T. Peacock and G. Haller, “Lagrangian coherent structures: The hidden skeleton of fluid flows,” *Phys. Today* **66**(2), 41–47 (2013).
  - <sup>62</sup>J. Peng and J. O. Dabiri, “Transport of inertial particles by Lagrangian coherent structures: Application to predator-prey interaction in jellyfish feeding,” *J. Fluid Mech.* **623**, 75–84 (2009).
  - <sup>63</sup>I. I. Rypina, L. J. Pratt, J. Pullen, J. Levin, and A. Gordon, “Chaotic advection in an archipelago,” *J. Phys. Oceanogr.* **40**(9), 1988–2006 (2010).
  - <sup>64</sup>I. I. Rypina, S. Scott, L. J. Pratt, and M. G. Brown, “Investigating the connection between complexity of isolated trajectories and Lagrangian coherent structures,” *Nonlinear Processes Geophys.* **18**, 977–987 (2011).
  - <sup>65</sup>R. M. Samelson, “Lagrangian motion coherent structures, and lines of persistent material strain,” *Annu. Rev. Mater. Sci.* **5**, 137–163 (2013).
  - <sup>66</sup>T. Sapsis and G. Haller, “Inertial particle dynamics in a hurricane,” *J. Atmos. Sci.* **66**, 2481–2492 (2009).
  - <sup>67</sup>C. Senatore and S. D. Ross, “Detection and characterization of transport barriers in complex flows via ridge extraction of the finite time Lyapunov exponent field,” *Int. J. Numer. Methods Eng.* **86**, 1163–1174 (2011).
  - <sup>68</sup>S. C. Shadden and A. Arzani, “Lagrangian postprocessing of computational hemodynamics,” *Ann. Biomed. Eng.* **43**(1), 41–58 (2014).
  - <sup>69</sup>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**, 271–304 (2005).
  - <sup>70</sup>P. Tallapragada, S. D. Ross, and D. G. Schmale III, “Lagrangian coherent structures are associated with fluctuations in airborne microbial populations,” *Chaos* **21**, 033122 (2011).
  - <sup>71</sup>W. Tang and C. Luna, “Dependence of advection-diffusion-reaction on flow coherent structures,” *Phys. fluids* **25**, 106602 (2013).
  - <sup>72</sup>W. Tang and A. Dhumuntarao, “Bistability in inhomogeneity - effects of flow coherent structures on the fate of a bistable reaction,” *Phys. Fluids* (accepted).

- <sup>73</sup>W. Tang, P. W. Chan, and G. Haller, "Lagrangian coherent structure analysis of terminal winds detected by lidar. Part I: Turbulence structures," *J. Appl. Meteorol Climatol.* **50**, 325–333 (2011).
- <sup>74</sup>W. Tang, P. W. Chan, and G. Haller, "Lagrangian coherent structure analysis of terminal winds detected by lidar. Part II: Structure evolution and comparison with flight data," *J. Appl. Meteorol Climatol.* **50**, 2167–2183 (2011).
- <sup>75</sup>J.-L. Thiffeault, "Braids of entangled particle trajectories," *Chaos* **20**, 017516 (2010).
- <sup>76</sup>J.-L. Thiffeault and M. Budisic, "Braidlab: A software package for braids and loops," e-print [arXiv:1410.0849](https://arxiv.org/abs/1410.0849) [math.GT].
- <sup>77</sup>S. Ulam, *Problems in Modern Mathematics* (Interscience, 1964).
- <sup>78</sup>E. van Sebille, M. H. England, and G. Froyland, "Origin, dynamics and evolution of ocean garbage patches from observed surface drifters," *Environ. Res. Lett.* **7**, 044040 (2012).
- <sup>79</sup>F. J. Beron-Vera, Y. Wang, M. J. Olascoaga, J. G. Goni, and G. Haller, "Objective detection of oceanic eddies and the Agulhas leakage," *J. Phys. Oceanogr.* **43**, 1426–1438 (2013).
- <sup>80</sup>J. Weiss, "The dynamics of enstrophy transfer in two-dimensional hydrodynamics," *Physica D* **48**, 273–294 (1991).
- <sup>81</sup>S. Wiggins, *Chaotic Transport in Dynamical systems, Interdisciplinary Applied Mathematics* (Springer, 1992).